UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECHNOLOGIE

DIPARTIMENTO DI CHIMICA

DOCTORATE SCHOOL IN

CHEMICAL SCIENCES AND TECHNOLOGIES

*Curriculum* Chemical Sciences (XXVII cycle)

# COMBINED NMR/LC-MS APPROACH FOR STRUCTURAL CHARACTERIZATION OF HEPARINS AND NON-ANTICOAGULANT GLYCOL-SPLIT HEPARINS

Ph.D thesis of

Anna Alekseeva

R09725

Tutor: Prof. Luigi Lay

Co-tutor: Dr. Annamaria Naggi

Coordinator: Prof. Emanuela Licandro

Academic year 2013/2014

*To my family*

## ABBREVIATIONS

2D – two-dimensional

[η] – intrinsic viscosity

A* ($A_{NS3S(6S)}$) – glucosamine-*N*,3-*O*-disulfate, mainly 6-*O*-sulfated

1,6a$A_{NS}$ – 1,6-anhydro-D-glucosamine-*N*-sulfate

1,6a$M_{NS}$ – 1,6-anhydro-D-mannosamine-*N*-sulfate

$A_{NS(6S)}$ – D-glucosamine-*N*-sulfate, mainly 6-*O*-sulfated in heparins

$A_{NAc(6S)}$ – *N*-acetyl-D-glucosamine, optionally 6-*O*-sulfated

aM.ol – 2,5-anhydro-D-mannitol

AT – antithrombin III

ATBR – antithrombin-binding region

BMH – bovine mucosal heparin

BLH – bovine lung heparin

D – polydispersity

DBA – dibutylamine

dp – degree of polymerization

ECM – extracellular matrix

ESI-Q-TOF-MS – electrospray ionization – quadrupole – time-of-flight mass spectrometry

G – D-glucuronic acid

GAG – glycosaminoglycan

Gal – galactose

GalA – D-galacturonic acid

$G_{LR}$ – D-glucuronic acid within the linkage region tetrasaccharide sequence

gsG/gsI/gsU – glycol-split glucuronic/iduronic/uronic acid

gs-heparin – glycol-split heparin

$^{1}J_{C\text{-}H}$ – one-bond proton-carbon coupling constant

HA fraction – high affinity fraction (toward AT)

HILIC – hydrophobic interaction liquid chromatography

HMBC NMR – heteronuclear multiple bond correlation NMR

HS – heparan sulfate

HSQC – (two-dimentional) heteronuclear single quantum coherence NMR

$I_{(2S)}$ – L-iduronic acid, mainly 2-*O*-sulfated in heparins

IPRP-HPLC – ion-pair reversed phase high performance liquid chromatography

LMWH – low molecular weight heparin

LOD – limit of detection

LOQ – limit of quantification

LOL – limit of linearity

LR – linkage region

$M_{NS}$ – D-mannosamine-*N*-sulfate

$M_w$ – weight average molecular weight

$M_n$ – number average molecular weight

NA fraction – no affinity fraction (toward AT)

NMR – nuclear magnetic resonance

NRE – non-reducing end

OMH – ovine mucosal heparin

R – remnant generated by the hydrolysis of glycol-split uronic acids

R(Unr) – remnant generated by oxidation of non-sulfated uronic acid at the non-reducing end

RALLS – right-angle laser light scattering

RE – reducing end

RI – refractive index

PMH – porcine mucosal heparin

RSD – relative standard deviation

TDA – triple detector array

TOCSY – total correlation spectroscopy

TSD – trisulfated disaccharide

S/N – signal-to-noise

Ser – serine of the LR

$Ser_{ox}$ – oxidized serine reside of the LR

SEC – size-exclusion chromatography

UFH – unfractionated heparin

U – uronic acid

$U_{2S}$ – 2-*O*-sulfated uronic acid

$\Delta U_{2S}$ – 4,5-unsaturated 2-*O*-sulfated uronic acid

Xyl – xylose

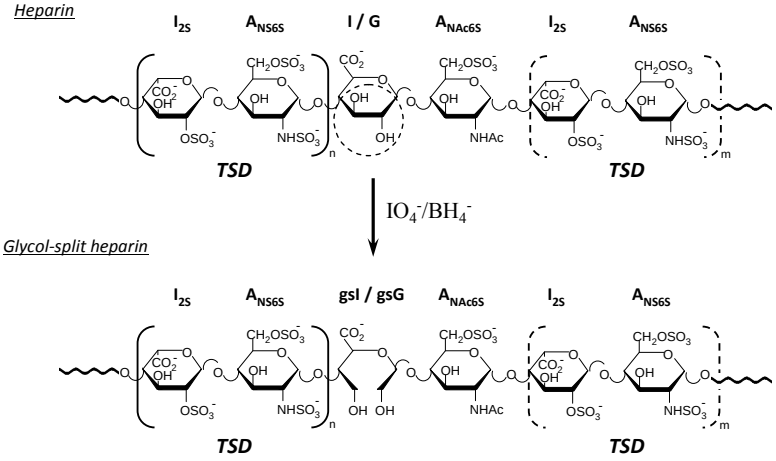
**CONTENT** **Page**

**Page**

**Page**

# INTRODUCTION

Heparin is a sulfated polysaccharide extracted from animal organs, which is successfully used since more than fifty years in clinics as anticoagulant and antithrombotic drug (*Casu* and *Lindahl*, 2001) There is also a growing interest in exploiting new activities of heparin (*Fuster* et al., 2005; *Lever* et al., 2002). Among the heparin derivatives possessing additional, pharmaceutically exploitable activities, glycol-split (gs) heparins, obtained by periodate oxidation of non-sulfated uronic acids usually followed by borohydride reduction, are arousing interest in anticancer and anti-inflammatory therapies (*Casu* et al., 2010). While heparins are mainly composed of repeating disaccharide sequences of 2-*O*-sulfated L-iduronic acid ($I_{2S}$) linked to D-glucosamine-*N,6-O*-disulfate ($A_{NS6S}$) interrupted by some units containing non-sulfated uronic acids (D-glucuronic acid (G) and L-iduronic acid (I)) and *N*-acetyl-D-glucosamine ($A_{NAc}$) residues (*Casu* and *Lindahl*, 2001), in glycol-split heparins the C(2)-C(3) bonds of G and I residues are "opened" (*Naggi*, 2005).




Simplified structure of heparin and its glycol-split derivative obtained by periodate oxidation/borohydride reduction

I and G – iduronic and glucuronic acids, $I_{2S}$ – 2-*O*-sulfated iduronic acid, A – glucosamine (*N*-sulfated or *N*-acetylated), TSD – trisulfated disaccharide

Given the increased local flexibility of gs residues, the interactions with proteins of the heparin sequences that remain substantially unaffected by the splitting reaction are expected to be facilitated (*Casu* et al., 2002). One important heparin-binding protein is heparanase. Overexpressed by tumors, this enzyme facilitates the release and activation of angiogenic growth factors stored by heparan sulfate proteoglycans (HSPGs) through partial cleavage of heparan sulfate (HS) chains. Nowadays, this enzyme represents an attractive target for

anticancer therapy and its inhibition is thought to be a very promising strategy in the development of new anti-cancer drugs (*Pisano* et al, 2014).

The search of effective inhibitors of heparanase started from studying the inhibitory activity of heparin, the closest mimic of HS that is commercially available. Early studies demonstrated that heparin does inhibit heparanase (*Lapierre* et al, 1996) in experimental animal models. However, the high doses needed for an effective anticancer treatment may cause undesirable effects, such as bleeding complications, associated with heparin anticoagulant activity. Moreover, some of the G-$A_{NS/NAc}$ linkages of heparin are susceptible to heparanase cleavage, resulting in a partial depolymerisation and loss of inhibitory activity. In contrast, gs-heparins, structural analogs of heparin and HS, maintain affinity toward heparanase, but do not represent substrates for the enzyme due to the induced modification of glucuronic acid. The new medical uses of gs-heparins are favored by the loss of the anticoagulant activity, associated with the splitting of the essential glucuronic acid within the active site responsible for the anthrombin III (AT) binding (*Islam* et al., 2002). A number of non-anticoagulant gs-heparins, differing for the extent of glycol-splitting, and molecular weight, are being developed as potential antiangiogenic and antimetastatic drugs (*Zhou* et al., 2011; *Ritchie* et al., 2011).

In spite of the numerous studies on their biological activity, detailed structural analysis of gs heparins as well as *structure–activity* relationship have not been reported yet. Their structural characterization remains a very challenging task, firstly, due to the high microheterogeneity of the starting heparin (significantly depending on tissue and animal sources), and to the various side reactions that can occur under sometimes uncontrolled reaction conditions. In the case of more bioavailable and, consequently, more pharmacologically attractive, low molecular weight heparins (LMWHs) used as starting material, the end groups introduced by the depolymerisation processes induce additional

heterogeneity of the corresponding gs-derivatives. Evidently, an appropriate analytical approach is required for a pharmaceutical development of these potential drugs.

The present study is focused on the development and application of an efficient analytical approach combining advanced techniques, such as NMR and LC-MS methods, for structural characterization of various gs-heparins and gs-LMWHs. The structural information obtained for gs-samples was expected to provide structural details about the sequences of the parent heparin/LMWH material. The long-term aim of this project includes the exploration of structural basis of heparin-related gs-oligosaccharide sequences for their optimal binding motif with heparanase and other target proteins.

In **Chapter I** the state-of-art of the present topic is overviewed. Structural features, biosynthesis and interaction with proteins of heparin and HS as well as structure and bioactivity of LMWHs are discussed. The potential anticancer and anti-inflammation activity of gs-heparins/gs-LMWHs and the main problems in their preparation and structural analysis are also highlighted. In the last part of the Chapter I the main analytical techniques used for analysis of heparin-related samples (NMR, LC-MS, SEC-TDA) are reviewed.

The overall goal of the study and the short-term objectives are described in **Chapter II**.

**Chapter III** is focused on the development of the combined NMR/LC-MS approach using a gs-derivative of a porcine mucosal heparin (PMH) sample as a model. The potentiality of this method for both characterization and differentiation of gs-products originated from heparins of various tissue and animal sources are discussed in terms of average monosaccharide composition determined by NMR as well as LC-MS qualitative and quantitative analysis of the enzymatically digested samples. The advantages of the combination of glycol-splitting and enzymatic depolymerisation for the characterization of gs-products and starting heparin samples are also discussed. A particular attention has been dedicated to the relevance of the diversity of antithrombin binding region (ATBR) in different heparins and to the length of undersulfated domains within their chains. Molecular weight

measurements as well as the dependence of the hydrodynamic properties on the introduced glycol-split units (determined by SEC-TDA) are also described. The possible control and monitoring of glycol-splitting reaction (including side-reactions) using the applied techniques is discussed.

The applications of the developed approach for the characterization of the more biologically attractive but often more complex gs-LMWHs (gs-tinzaparin, gs-enoxaparin, gs-dalteparin) are reported in **Chapter IV**. The developed LC-MS method for direct qualitative profiling of gs-LMWHs (possible because of the molecular weight lower than for unfractionated heparins) is discussed as a possible additional analytical tool for their compositional and structural analysis. The studied susceptibility to periodate oxidation/borohydride reduction of particular end-groups (at both reducing and non-reducing ends) of all the three LMWHs is highlighted. Among those, one of the most unexpected result was the oxidation of the enoxaparin 1,6-anhydro-D-mannosamine-*N*-sulfate, which was studied in detail. The potential application of the LC-MS method for the heparin sequencing is shown for sequences adjacent to the linkage region tetrasaccharide.

General conclusions and future perspectives are discussed in **Chapter V.**

The experimental details are reported in **Chapter VI.**

The present study was carried out at Institute for Chemical and Biochemical Research "Ronzoni". The project was financially supported by the National Institute of Health (grant number R01-CA138535) and the Italian Association for Cancer Research (AIRC, grant number IG10569).

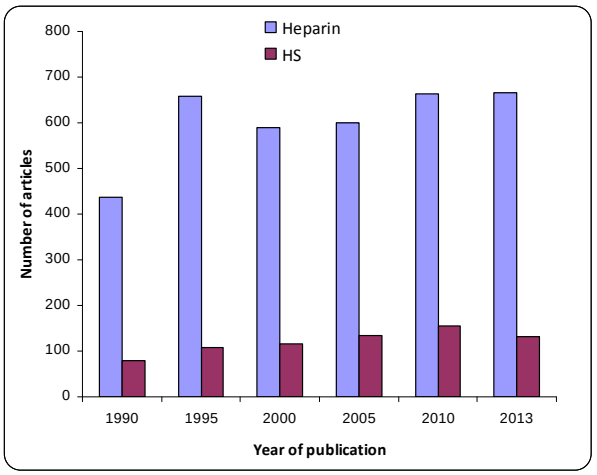
The results have been published in three scientific journals:

- Alekseeva A., Casu B., Torri G., Pierro S., Naggi A. Profiling glycol-split heparins by high-performance liquid chromatography/mass spectrometry analysis of their heparinase-generated oligosaccharides. *Analytical Biochemistry.* 434 (2013) 112-122.
- Alekseeva A., Casu B., Cassinelli G., Guerrini M., Torri G., Naggi A. Structural features of glycol-split low-molecular-weight heparins and their heparin lyase generated fragments. *Analytical and Bioanalytical Chemistry.* 406 (2014) 249-265.
- Alekseeva A., Elli S., Cosentino C., Torri G., Naggi A. Susceptibility of enoxaparin reducing end amino sugars to periodate oxidation. *Carbohydrate Research. 400* (2014) 33-43.

# CHAPTER I: SCIENTIFIC BACKGROUND

## I.1. Heparin and heparan sulfate

Heparin, clinically used as an anticoagulant in surgery and in both the treatment and prophylaxis of thrombosis, is one of the most widely used drugs. Given the large number of proteins of physiological and pathophysiological importance interacting with heparin and the structurally related HS, the range of their potential therapeutic applications (treatment of cancer, neurological disorders, inflammatory diseases, malaria treatment (*Fuster* et al., 2005; *Lever* et al., 2002)) is raising. Such an increasing interest in the exploiting of new activities of heparin requires development of analytical methods for their better structural characterization, even if, nowadays, heparin represents one of the most studied glycosaminoglycans (GAGs). This trend is reflected in the diagram (on the right), showing an almost constant high number of publications on heparin and a constant, but much lower publications number on HS (publication search with words "heparin" or "HS" in title/abstract/keywords found by sciencedirect searching system are shown).

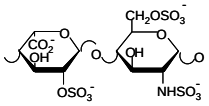
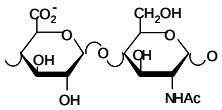
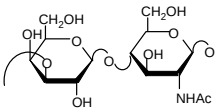
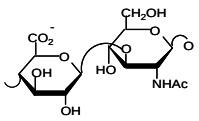
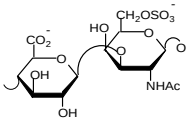
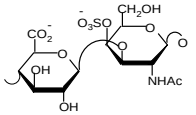
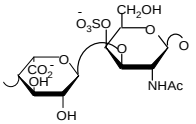



### I.1.1. Glycosaminoglycan family

GAGs are natural linear, polydisperse and heterogeneous polysaccharides that are present in every mammalian tissue and affect cell properties and functions, acting directly on cell receptors or via interactions with extracellular proteins. All GAGs are composed of repeating uronic acid and glycosamine disaccharide subunits that can be either *O*-/*N*-sulfated or not. Differences in the type of monosaccharide in the repeating unit as well as their sulfation pattern distinguish the major categories of GAGs: *glucosaminoglycans* (heparin, HS,

hyaluronic acid and keratan sulfate) characterized by D-glucosamine and *galactosaminoglycans* (chondroitin sulfate and dermatan sulfate), characterized by D-galactosamine. Each GAG type, described in terms of its prevalent disaccharide unit, is shown in Table 1.1.

**Table 1.1**. The most prevalent repeating disaccharide unit characteristic for each type of GAG

| GAG | Uronic acid | Galactose | Glycosamine | Main repeating disaccharide unit |
|---|---|---|---|---|
| GLUCOSAMINOGLYCANS | | | | |
| HEPARIN | L-iduronic acid 2-*O*-sulfate ($I_{2S}$) | – | D-glucosamine *N*-sulfate ($A_{NS}$) |  $I_{2S}$ α(1→4) $A_{NS}$ α(1→4) |
| HEPARAN SULFATE | D-glucuronic acid (G) | – | *N*-acetyl-D-glucosamine ($A_{NAc}$) |  G β(1→4) $A_{NAc}$ α(1→4) |
| KERATAN SULFATE | – | D-galactose (Gal) | *N*-acetyl-D-glucosamine ($A_{NAc}$) |  Gal β(1→4) $A_{NAc}$ β(1→3) |
| HYALURONIC ACID | D-glucuronic acid (G) | – | *N*-acetyl-D-glucosamine ($A_{NAc}$) |  G β(1→3) $A_{NAc}$ β (1→4) |
| GALACTOSAMINOGLYCANS | | | | |
| CHONDROITIN SULFATE | D-glucuronic acid (G) | – | *N*-acetyl-D-galactosamine-6-*O*-sulfate ($Gal_{NAc6S}$) |  G β(1→3) $Gal_{NAc6S}$ β (1→4) |
| | | | *N*-acetyl-D-galactosamine-4-*O*-sulfate ($Gal_{NAc4S}$) |  G β(1→3) $Gal_{NAc4S}$ β (1→4) |
| DERMATAN SULFATE | L-iduronic acid (I) | – | *N*-acetyl-D-galactosamine-4-*O*-sulfate ($Gal_{NAc}$) |  I α(1→3) $Gal_{NAc4S}$ β (1→4) |

In each GAG type molecular size, structural sequence, composition and sulfation pattern may vary significantly, depending on the tissue origin, age and pathological conditions. Due to their biological functions in both physiological and pathological conditions, GAGs are being considered important biomarkers as well as potential pharmacological targets. Their involvement in the cell signalling and cancer progression and some *structure–biological activity* relationship has been recently reviewed (*Afratis* et al., 2012).

### I.1.2. Heparin/heparan sulfate: biosynthesis and structural microheterogeneity

Heparin and HS are built up of repeating disaccharide units consisting of glucosamine (A) and uronic acid (U). Such a repeating structure is the result of a similar multistep biosynthesis pathway, occurring in the Golgi apparatus, which includes initiation, elongation and modification steps, described in detail below (Fig. 1.1).

*Initiation* – the formation of a tetrasaccharide sequence $G_{LR}$-Gal-Gal-Xyl (where, $G_{LR}$ – glucuronic acid of the linkage region (LR), Gal – galactose, Xyl – xylose) linking the heparin/HS chain to the core protein at the level of the serine (Ser) residues (*Iacomini* et al., 1999). The LR can undergo the enzymatic modifications, as phosphorylation at C-2 of Xyl or sulfation at C-4 or C-6 of Gal residues. Remarkably, the 2-*O*-phosphorylation of the xylose residue, even if rare, was observed in both heparin/HS and chondroitin sulfate (CS), while sulfated galactose residues were found in the LR of CS and dermatan sulfate (DS), but not in heparin or HS (*Tone* et al., 2008). This finding suggests that the sulfate groups on the Gal residues may be involved in the selective chain assembly of CS/DS, while its absence leads to the synthesis of heparin and HS (*Tone* et al., 2008).

*Elongation* – the first step of the generation of polysaccharide chain consisted in the subsequent addition of D-glucuronic acid (G) β1 → 4 linked to a *N*-acetyl-D-glucosamine ($A_{NAc}$) building blocks catalyzed by glycosyltransferases.

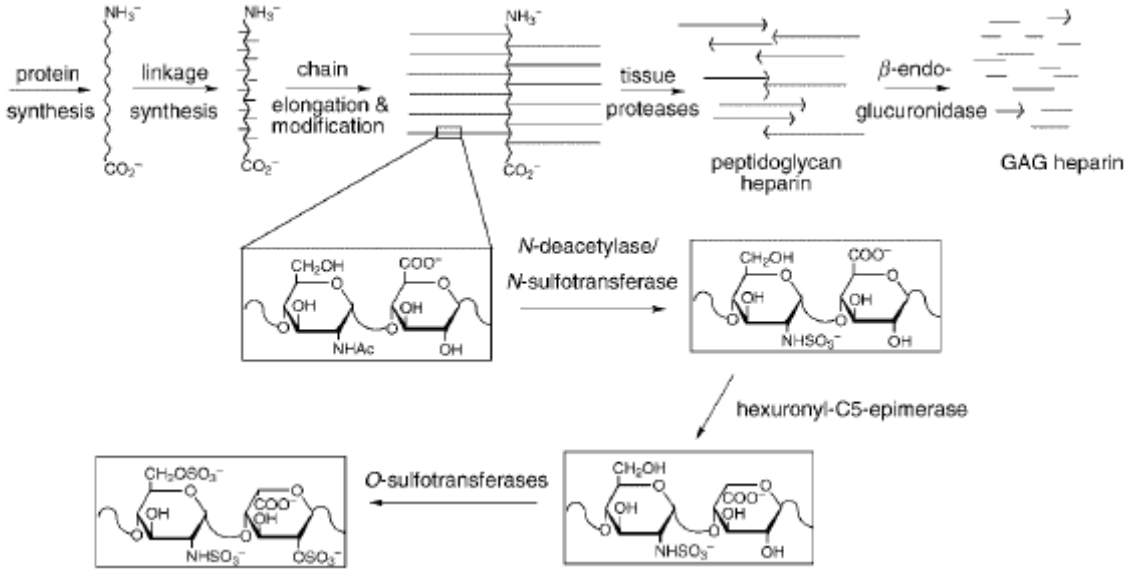



**Fig.1.1.** Biosynthesis scheme of heparin/HS (taken from *Capila et al., 2002*)

*Modification* – the highly specific enzymatic modification of these chains, involving *N*-deacetylation and *N*-sulfation of glucosamine residues, $C_5$-epimerization of G to iduronic acid (I), and *O*-sulfation at either the C-2 of the U or the C-6 of the A residues (Fig.1.1). Since both $C_5$-epimerization and *O*-sulfation depend on a prior *N*-sulfation of glucosamine units, the modified sequences tend to be clustered in the NS or NS/NA-domains (*Capila et al., 2002, Esko et al., 2001*), as shown for heparin in Fig.1.2. As a final step, the *O*-sulfation at the C-3 of some A units can occur. This latter process plays a role in the synthesis of a specific pentasaccharide $A_{NAc6S}$-G-$A_{NS3S6S}$-$I_{2S}$-$A_{NS6S}$ (AGA*IA) mainly within the *N*-sulfated domain (Fig.1.2) responsible for the interaction of heparin with antithrombin III (AT) (see I.1.4). The percentage of 3-*O*-sulfated glucosamine ranges from about 4 up to 8 % in commercial heparins (*Guerrini* et al., 2005, *Zhang* et al., 2011). As a minor modification, also G within heparin/HS chains may be 2-*O*-sulfated (*Razi* et al., 1995; *Guerrini* et al., 2007).

Heparin is modified at the last "*modification*" step to a higher extent than HS. Their structural differences are discussed in the section I.1.3

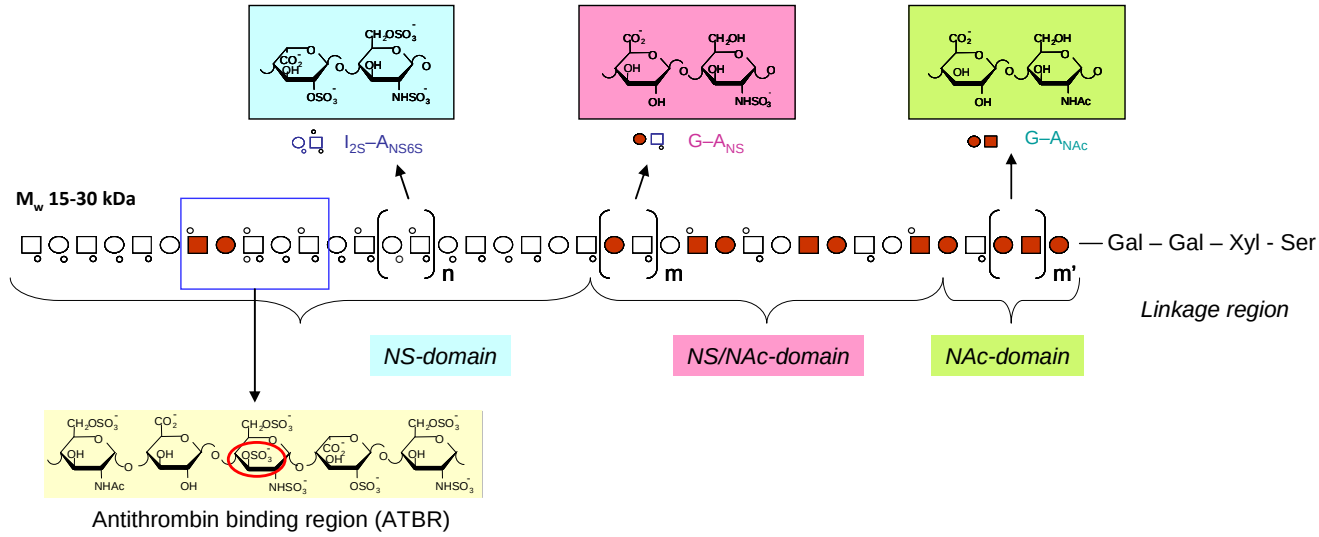



**Fig.1.2**. Schematic representation of heparin domain structure

G – glucuronic acid, $I_{2S}$ –iduronic acid 2-*O*-sulfate, $A_{NS6S}$ – glucosamine *N*,6-*O*-disulfate, $A_{NS}$ – glucosamine-*N*-sulfate, $A_{NAc}$ – *N*-acetyl-glucosamine, Gal – galactose, Xyl – xylose, Ser – serine, ATBR – antithrombin binding region, $M_w$ – weight average molecular weight

Such a sequential enzymatic modification determines the formation of the specific saccharide sequences that mediate the interactions with plasma and tissue proteins and, consequently, the biological properties of heparin and HS (*Casu* et al, 1989; 2001). Each of these biosynthetic reactions usually does not proceed to completion, resulting in chains containing sequences modified to different extent. The structural microheterogeneity may further increase due to the cleavage of a macromolecular precursor of heparin/HS by heparanase, an endogeneous glucuronidase. The enzyme acts at the G-A glycosidic bond (see also *I.1.4.3*), leading to the formation of a glucosamine at the non-reducing end (NRE) (*Gong* et al, 2003; *Sandbäck-Pikas* et al., 1998). Because the extent of such a specific cleavage can vary under different pathological conditions, characterization of these disease-specific NRE sequences may amplify medical diagnosis possibilities (*Lawrence* et al., 2011). HS chains are generally longer than those of heparin, and have an average molecular weight of ~30 kDa as compared to 15-20 kDa for heparin (*Shriver* et al, 2012; *Kjellén et al*, 2000; *Mulloy* et al., 2000). It is worth noting that heparin/HS structure (specific sulfation/*N*-acetylation pattern, polydispersity, molecular weight, etc.) varies for different tissues. Pharmaceutical heparin is

most often prepared from porcine intestinal mucosa (porcine mucosal heparin, **PMH**), even if historically it has also been obtained from bovine lung (bovine lung heparin, **BLH**) and can be also prepared from ovine intestine (ovine mucosal heparin, **OMH**) (*Fu* et al, 2013). Muslim countries use a substitute of PMH obtained from bovine organs (BLH and bovine mucosal heparin, **BMH**). The processes used during the heparin production can contribute to its structural complexity as well. The LR, present at the reducing end (RE) of some heparin chains, is the most sensitive sequence to such treatments, and its content and structure depends on the strength of oxidative "bleaching" treatment (*Mourier* et al, 2005, Fig.1.3). Usually a few percents (1.0-1.3%) of intact or partially modified LR remain in heparins (*Zhang* et al, 2011) and in some LMWHs. .

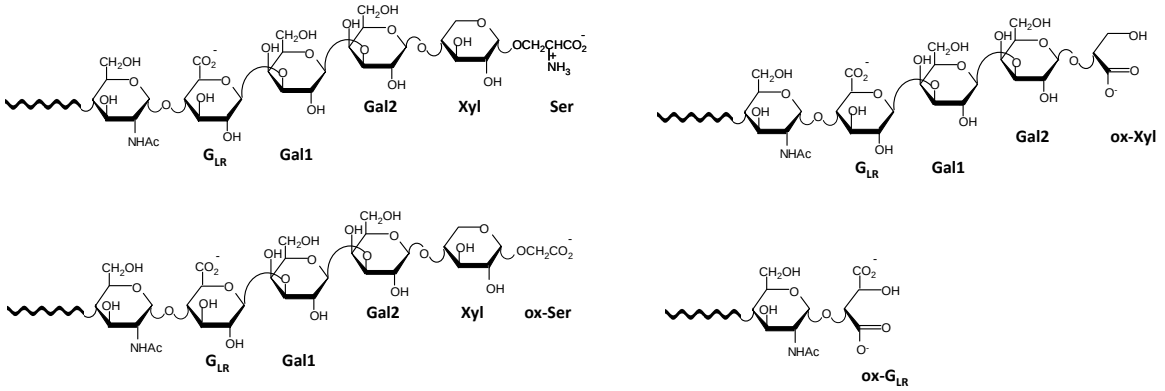



**Fig. 1.3**. Linkage region structures found in heparin after bleaching by permanganate treatment
$G_{LR}$ – glucuronic acid of the LR sequence, Gal1/Gal2 – galactoses, Xyl – xylose, Ser – serine, ox-Xyl , ox-Ser, ox-$G_{LR}$ – corresponding oxidized forms
Permanganate may oxidize vicinal hydroxyl-groups of $G_{LR}$ and Xyl as well as the serine amino-group, leading to shortening of LR sequence

Because of the high microheterogeneity, the to-date partially undefined heparin structure needs further studies. Recently, several new details, better elucidating the heparin structure, have been discovered. For example, the results obtained by *Xiao* et al (2011) suggest that pharmaceutical heparin contains two undersulfated regions, an internal one, separating two highly sulfated domains, and the second one in the closest proximity to the LR, both susceptible to cleavage by the heparin lyase III (Fig. 1.4).

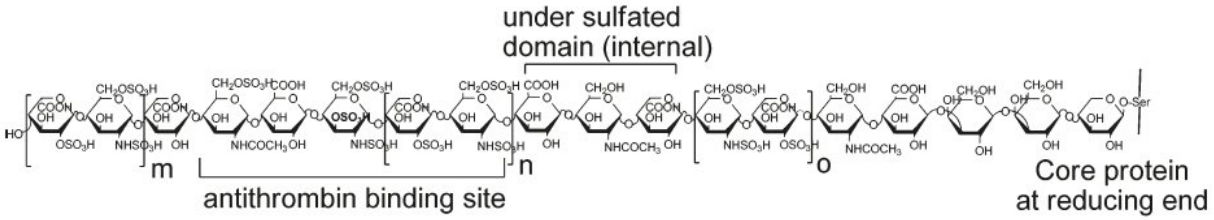



**Fig. 1.4.** Schematic representation of heparin chains containing two undersulfated domains (from *Xiao* et al, 2011)

Interestingly, the oligosaccharides containing two and three adjacent AGA*IA sequences (Fig. 1.5) were isolated from a LMWH (*Viskov* et al., 2013; *Mourier* et al., 2014).

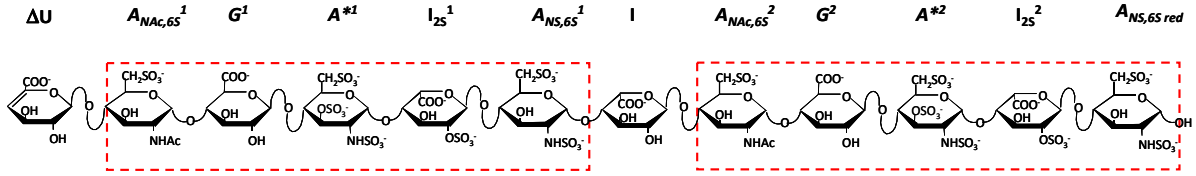



**Fig. 1.5.** Structure of a particular enoxaparin sequences containing two adjacent ATBR pentasaccharide sequences (*Viskov* et al., 2013)

I.1.3. Structural differences between heparin and heparan sulfate

Heparin attached to the serglycin core protein is synthesized exclusively in mast cells and cleaved from its core protein to be released into the extracellular matrix (ECM). On the contrary, HS, produced by virtually all the cells of human organism, can be covalently attached to a variety of core proteins, forming heparan sulfate proteoglycans (HSPGs), that are either secreted to the extracellular space, or associated with cell surface (Fig. 1.6).

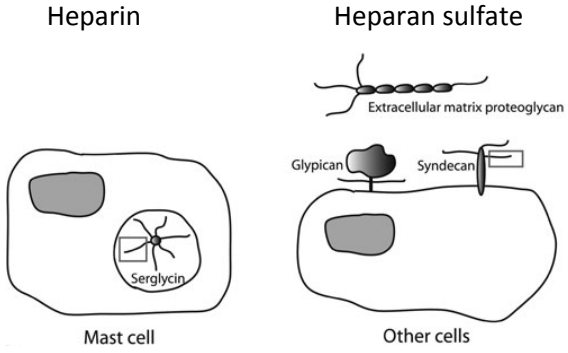



**Fig. 1.6**. Heparin and HS formation in the human organism

Mast cells produce highly sulfated heparin attached to the serglycin core protein, stored in mast cell granules that can be released into the extracellular environment. Other cells synthesize less negatively charged HS, which is found at the cell surface as glypican and syndecan proteoglycans, and in the extracellular matrix.

Despite this difference, heparin and HS share the same precursor, *N*-acetyl-heparosan, consisting of $A_{NAc}$ and G residues. Heparin is often referred to as the more completely modified version of HS. Indeed, it is known to be more *N*- and *O*-sulfated than HS (Fig.1.7), and, consequently, to possess a higher negative charge density, the highest of any discovered biological macromolecules (*Capila* et al, 2002). Other than a higher *N*-sulfation degree, heparin is characterized by a higher content of iduronic acid due to a higher $C_5$-epimerization extent. In fact, the prevalent heparin sequences are represented by a trisulfated disaccharide (TSD) L-iduronic acid 2-*O*-sulfate α1 → 4 D-glucosamine-*N*,6-*O*-disulfate ($I_{2S}$-$A_{NS6S}$) in heparin (Fig.1.7a) and D-glucuronic acid β1→ 4 *N*-acetyl-D-glucosamine (G-$A_{NAc}$) in HS (Fig.1.7b), so that the heparin disaccharides contain an average of ~2.7 sulfate groups, whereas those of HS only ~1 sulfate group (*Capila* et al., 2002; *Shriver* et al., 2012). The structural heterogeneity of HS is much greater due to the lower sulfation (Fig.1.7c), thought to be a highly regulated process, producing very frequently alternated *N*-sulfated and *N*-acetylated domains (Fig.1.7c), and, thus, determining the HS role in various biological processes.

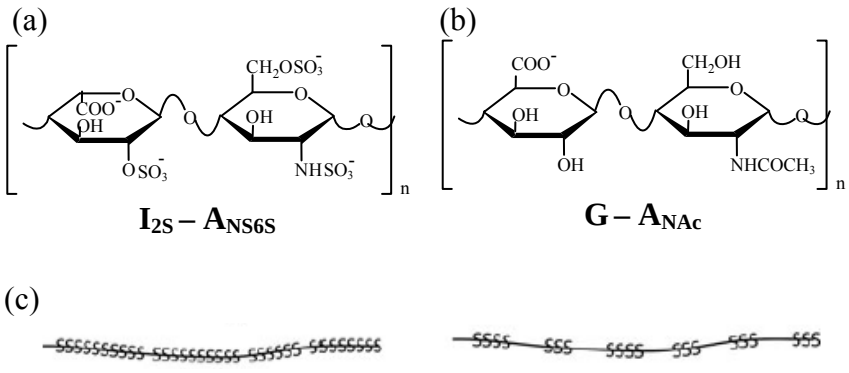



**Fig.1.7.** The major repeating disaccharide unit in heparin (a) and HS (b) Panel (c) illustrates the difference in sulfation pattern and heterogeneity between heparin and HS (*Carlsson* et al, 2012)

Heparin and HS contain 80-90% and 30-60% of $A_{NS(6S)}$, respectively (*Carlsson* et al, 2012). $C_5$-epimerization degree of heparin is much higher than for HS. In fact, I is prevalent within heparin chains, while its epimer G represents the majority of the uronic acids in HS (*Lindahl* et al., 1998).

### I.1.4. Interactions with proteins

The structural diversity of heparin/HS relies on the core of the wide range of biological processes that these molecules modulate by their interaction with numerous proteins, such as AT, fibroblast growth factors (FGFs), vascular endothelial growth factors (VEGFs), receptors, chemokines, etc. (*Bernfield* et al., 1999). The understanding of these interactions, which play important role in both normal physiological (blood coagulation, cell proliferation, etc.) as well as pathological (inflammation, tumor metastasis, angiogenesis etc.) processes, at the molecular level is of fundamental importance in the design of new highly specific therapeutic agents.

Heparin/HS are known to exist as a helical structure that cannot fold into a tertiary structure due to their extremely high negative charge density (*Capila et al., 2002)*. However, their interactions with some proteins are not aspecific and require defined sequences within their chains. The interaction specificity is thought to be associated with a particular orientation of sulfate and carboxyl groups within defined oligosaccharide sequences.

An additional structural property, through which heparin/HS chains can alter the spatial orientation of their functional groups, is the conformational flexibility of $I_{2S}$ within linear sequences (*Ferro* et al., 1986 and 1990), which prevalently exist in chair $^{1}C_4$ and skew-boat $^{2}S_0$ forms (Fig. 1.8). For example, $I_{2S}$ surrounded by two $A_{NS6S}$ residues, as frequently found in heparin-like sequences, showed the two forms $^{1}C_4$ and $^{2}S_0$ populated approximately at 60 and 40%, respectively (*Ferro* et al., 1990). Importantly, the conformation of individual residues may be influenced by the adjacent sequences (*Ferro* et al., 1986, 1990; *Yates* et al., 2000). For $I_{2S}$ residue the higher the number of sulfate groups in the adjacent glucosamines, the higher the percentage of the $^{2}S_0$ conformation. In contrast, both $A_{NS/NAc}$ and G are stable in $^{4}C_1$ chair irrespective of substitution (*Powell* et al., 2004). From this point of view, the HS backbone with a lower content of $I_{2S}$ residues, possesses lower local flexibility. However, its lower sulfation degree and, consequently, the lower charge density, does not “block” its chain providing flexibility sufficient to interact with proteins (*Khan* et al., 2012).

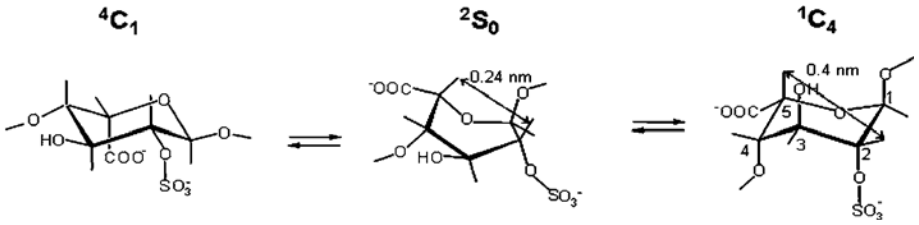



**Fig. 1.8**. Equilibrium of $^4C_1$, $^2S_0$ and $^1C_4$ forms of $I_{2S}$ residues (*Guerrini* et al, 2006)

The specificity of interactions between heparin/HS and proteins can be either a restrictive process, as for example, the highly specific heparin/HS–AT interaction (see *I.1.4.1*), or a more relaxed one, as in the case of their binding to FGFs (*Kreuger* et al., 1999).

Among the proteins interacting with heparin/HS, there are also some enzymes that specifically cleave their chains leading to their depolymerization. Two types of such enzymes are known, the prokaryotic polysaccharide lyases (acting through an eliminative mechanism), also called heparin lyases (see *I.1.4.2*), and the eukaryotic glucuronyl hydrolases (acting through a hydrolytic mechanism), first of all, heparanase (see *I.1.4.3*) mentioned above as the enzyme degrading macroheparin.

Hereafter, the interactions of heparin and HS with AT, heparin degrading heparinases and heparanase are discussed in detail.

*I.1.4.1. Heparin – antithrombin (III) interaction*

The first reported and the best studied example of the specific heparin–protein interaction of physiological significance is the "lock and key" interaction between heparin/HS and AT, leading to the inactivation of thrombin, factor Xa and other serine protease of the coagulation cascade. A specific pentasaccharide containing the disaccharide unit *glucuronic acid – trisulfated glucosamine* (G-A*) with the unusual but essential 3-*O*-sulfation (Fig. 1.2) is responsible for the interaction between heparin and AT, causing a conformational change in the serine protease inhibitor, thus allowing the loop with the reactive centre to interact with key coagulation enzymes, such as thrombin and factor Xa. Recent studies demonstrate that also the I residue preceding the antithrombin binding region (ATBR) pentasaccharide (Fig.1.9) may influence the AT-binding properties (*Guerrini* et al., 2006). It is worth noting that the

sulfation pattern of the ATBR sequence differs for heparins of different animal and tissue origin. For instance, its sulfated analogue ($I_{2S}$)-$A_{NS6S}$-G-$A_{NS3S6S}$-$I_{2S}$-$A_{NS6S}$ was reported to represent the prevalent ATBR sequence in bovine lung heparin (*Loganathan* et al., 1990). In Fig.1.9 the functional groups essential for AT-binding are shown.

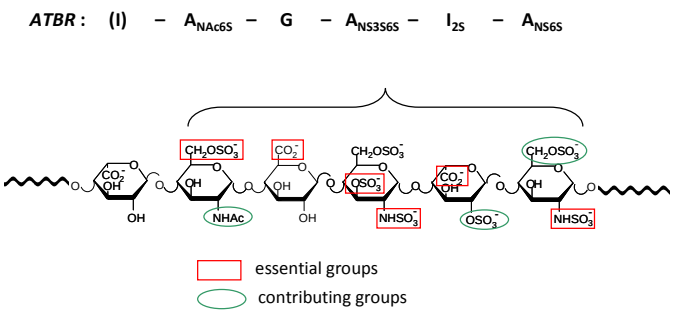



**Fig. 1.9**. Structure of the major ATBR sequence found in PMH and the role of the individual charged groups in the AT activation
(adapted from *Petitou and van Boeckel*, 2004)
The first glucosamine can be also *N*-sulfated instead of *N*-acetylated. According to the heparin biosynthesis, in such a case, the iduronic acid preceeding the pentasaccharide sequence is almost always 2-*O*-sulfated (*Loganathan* et al., 1990)

The pentasaccharide stimulates exclusively the AT-mediated inactivation of factor Xa (anti-Xa activity), whereas longer heparin fragments comprising both the ATBR sequence and a thrombin-binding domain are required for stimulating AT activity (Fig.1.10).

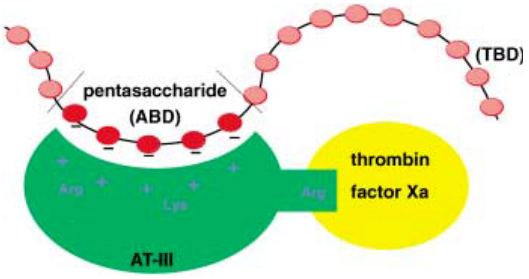



**Fig.1.10**. Tertiary complex heparin – antithrombin (III) – coagulation factor
(from *Petitou and van Boeckel*, 2004)

ABD – antithrombin (III) binding domain responsible for exclusively the AT-III-mediated inactivation of factor Xa (anti-Xa activity),
TBD – thrombin-binding domain

*I.1.4.2. Heparin-degrading lyases (or heparinases) I, II and III*

*Flavobacterium heparinum* produces three heparinases specific for different sequences of heparin (*Gallagher* et al., 1981). Heparinases I, II and III cleave A – U glycosidic bond leading to the formation of two oligosaccharides, one bearing a glucosamine at the RE and the second one bearing an unsaturated uronic acid at the NRE (Fig.1.11).

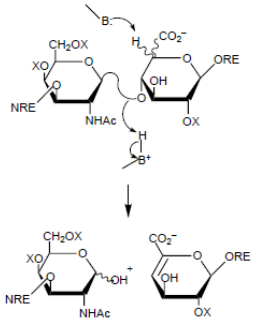



**Fig.1.11.** Mechanism of the eliminative depolymerization of heparin/HS chains by heparinases (*Capila et al., 2002)*
RE – reducing end, NRE – non-reducing end, B – enzyme basic centre

The specificity of heparinases largely depends on the sulfation/acetylation pattern of the heparin/HS sequence (Annex 1). Heparinase I cleaves highly sulfated regions of heparin/HS at $I_{2S}$, whereas heparinase II has a broader substrate specificity and cleaves glycosidic linkages containing both 2-*O*-sulfated and non-sulfated U residues. As a result of its broader specificity, heparinase II provides a most complete digestion of PMH (up to 95%), in comparison with ~75% obtained by heparinase I (*Xiao* et al., 2011; *Wang* et al., 2012). Heparinase III acts prevalently at the glycosidic linkages $A_{NAc}$-U within the undersulfated regions of heparin/HS, therefore, heparinase III results in the least digestion (~5%) of heparin samples. The present data indicate that 100% digestion cannot be performed, because some particular sequences are resistant to heparinases. It is generally accepted that the linkage at the non-reducing side of G-$A_{NS3S(6S)}$ disaccharide unit is resistant toward heparinases I, II and III (*Shriver* et al., 2000; *Yu* et al., 2000). On the contrary, the linkage at the RE of this unit is extremely sensitive to heparinases I and II (*Xiao, Zhao* et al, 2011) as confirmed using the synthetic pentasaccharide $A_{NS6S}$-G-$A_{NS3S6S}$-$I_{2S}$-$AOMe_{NS6S}$ as a model (*Yu* et al., 2000). As a result of this specificity, the exhaustive digestion of heparin/HS "truncates" the ATBR sequence generating oligosaccharides containing the important G-A* unit at the RE (Fig.1.12). Unlike heparinases I and II, heparinase III does not recognized the linkage of the highly sulfated -$A_{NS3S6S}$-$I_{2S}$- unit and, consequently, cannot cut the ATBR. *Xiao* et al (2011) suggested that the heparinase III specificity depends on the oligosaccharide sequence at the NRE of the linkage $A_{NAc}$-U. It is likely that fully sulfated residues flanking the cleavage site from the NRE may prevent the action of heparinase III.

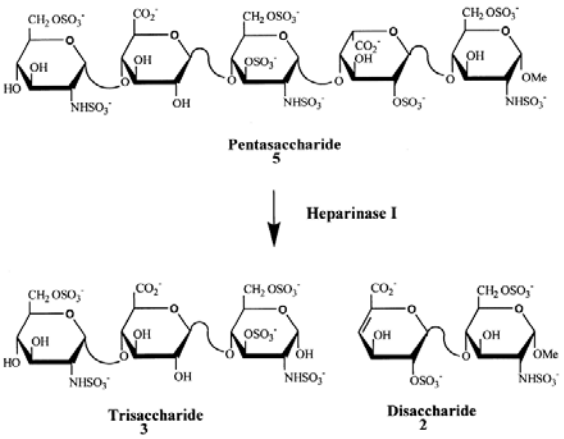



**Fig.1.12**. Synthetic heparin pentasaccharide (5) treated with by heparinase I affords a disaccharide (2) and a trisaccharide (3) (reported from *Yu* et al., 2000)

Me – methyl group at the RE

Results of many investigations on the heparinases substrate specificity have contributed to the structural elucidation of heparin/HS and understanding of the *structure-function* relationships.

*I.1.4.3. Heparanase*

Heparanase is a β-D-endoglucuronidase, which cleaves HS side chains of HSPGs, releasing fragments of 10-20 sugar units (3-6 kDa). The heparanase cDNA encodes for a polypeptide of 543 amino acids that appears as a heparanase inactive form of ~ 65 kDa. The protein undergoes a proteolytic cleavage that is likely to occur at two potential cleavage sites, $Glu^{109}$ – $Ser^{110}$ and $Gln^{157}$ – $Lys^{158}$, yielding an 8 kDa polypeptide at the N-terminus and a 50 kDa polypeptide at the C-terminus that are non-covalently associated (Fig.1.13), forming the active heparanase form (*Fairbanks et al., 1999; Levy*-Adam et al., 2003).

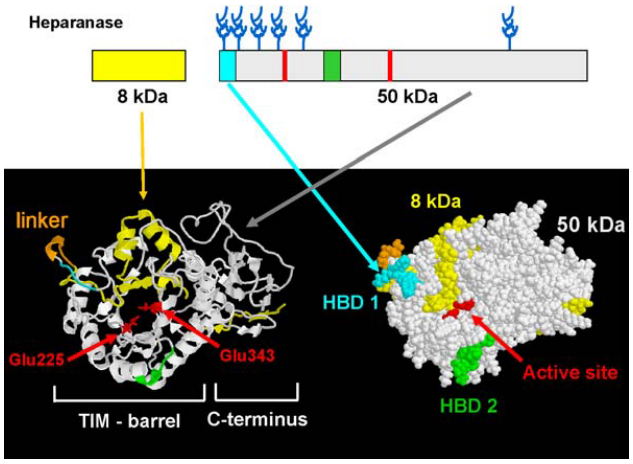



**Fig.1.13**. Primary structure, critical amino acids and predicted three-dimensional structure of the heparanase heterodimer

8 kDa (Gln36–Glu109) and 50 kDa (Lys158–Ile543) (gray) protein subunits (yellow and grey respectively) heterodimerize to yield an active enzyme. The heparin binding domains (HBD 1 and 2, blue and green) are in close proximity to the enzyme active site (red)

(reported from *Vlodavsky* et al., 2007)

The substrate specificity is not well-characterized, however, it is known that the linkage between G and $A_{NS}$ is susceptible to heparanase, while disaccharide units containing $I/I_{2S}$ are resistant. The heparanase cleavage is size dependent: the degradation rate is higher for longer sequences (dp 9 > dp 7 > dp 5), and the minimum size is thought to be trisaccharide sequence (*Peterson and Liu*, 2012). Moreover, the substrate specificity significantly depends on the sulfation/acetylation pattern of the HS sequences in proximity of the cleavage site (Annex 2, *Okada* et al., 2002; *Peterson and Liu*, 2013). This enzyme was shown to degrade substrates in two modes: both consecutive and gapped cleavage. In the first mode, heparanase cleaves each susceptible linkage starting from the NRE, while in the gapped mode the enzyme may skip few cleavage sites (*Peterson and Liu*, 2012). This latter mechanism may contribute to preserve the integrity of the sulfated domains of HS, necessary for its interaction with various biomolecules of the extracellular matrix (ECM).

Both cell-surface and ECM HSPGs store various growth factors, cytokines, chemokines and other proteins (Fig.1.14). Heparanase-degradation of HSPGs releases these bioactive molecules, physiologically and pathologically important (Fig.1.14). While heparanase is confined inside specific cells under the normal conditions, it is overexpressed and found in the circulation of most human oncological and inflammation diseases (*Pisano* et al., 2014). In experimental tumors heparanase promotes tumor metastasis and angiogenesis. Moreover, also chronic inflammatory conditions, induced by heparanase and present in the most tumor

microenvironments, may contribute to cancer progression. Considering all its activities, heparanase has generated interest as target for anti-cancer and anti-inflammatory therapy. Current approaches for heparanase inhibition include development of chemically-modified heparins (such as glycol-split heparins, *see I.3*) and small molecule inhibitors, including small oligosaccharides such as PI-88 and PG545, *Ferro* et al, 2007 and 2012). Some of these compounds are currently in various stages of preclinical/clinical development.

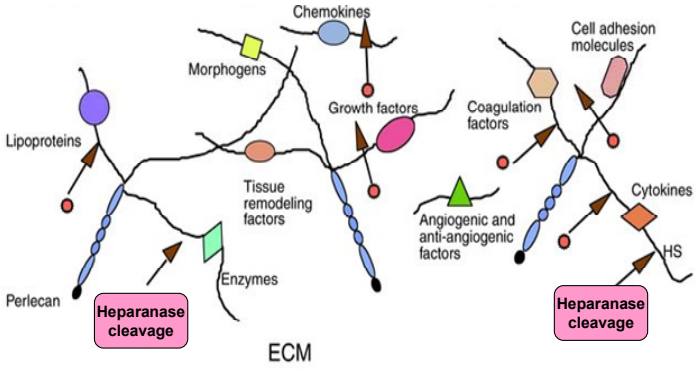



**Fig.1.14.** Heparanase-mediated release of bioactive molecules from the extracellular matrix (ECM)

From *(Vlodavsky and Friedman,* 2001)

***HIGHLIGHTS***

*The biosynthesis pathway determines structural microheterogeneity and polydispersity of heparin and heparan sulfate, which can be further increased during the manufacture processes. Detailed investigations are still required for the structural elucidation of heparin and the more complex HS, and their multifunctionality in terms of numerous biological activities they modulate.*

### I.2. Low molecular weight heparins

Despite of the increasing interest in the heparin therapeutic applications, its use is often limited by some undesirable side effects, such as bleeding complications or heparin-induced thrombocytopenia (HIT). This drawback stimulated detailed studies of the coagulation cascade and led to the development of low molecular weight heparins (LMWHs) at the end of the 20$^{th}$ century. The pharmaceutical differences between LMWH and unfractionated heparin (UFH), such as better bioavailability, more predictable anticoagulant activity and longer half-life, can

be explained by the decreased propensity of the former to bind to plasma proteins, entothelial cells, and macrophages due to their shorter chains (*Weitz*, 1997). For example, a shorter-chain heparin, containing an ATBR, but unable to accommodate thrombin binding, would retain AT-mediated inactivation of factor Xa but prevent the formation of a ternary heparin-AT-thrombin complex (Fig.1.10). More selective anticoagulant/antithrombotic agents were synthesized by enzymatic and chemical depolymerisation. Various LMWHs (5–8 kDa) and ultra-LMWH (3 kDa) were also generated (*Shriver* et al, 2000), and "at the beginning of the new millennium, LMW heparins have largely displaced heparin as the major clinical anticoagulant" (*Capila* et al., 2002).

Four LMWHs, tinzaparin ("Innohep", LEO Pharma), enoxaparin ("Clexane", Sanofi-Aventis), dalteparin ("Fragmin", Pfizer) and nadroparin ("Fraxiparin", GlaxoSmithKline) (Fig.1.15), are currently commercially available.

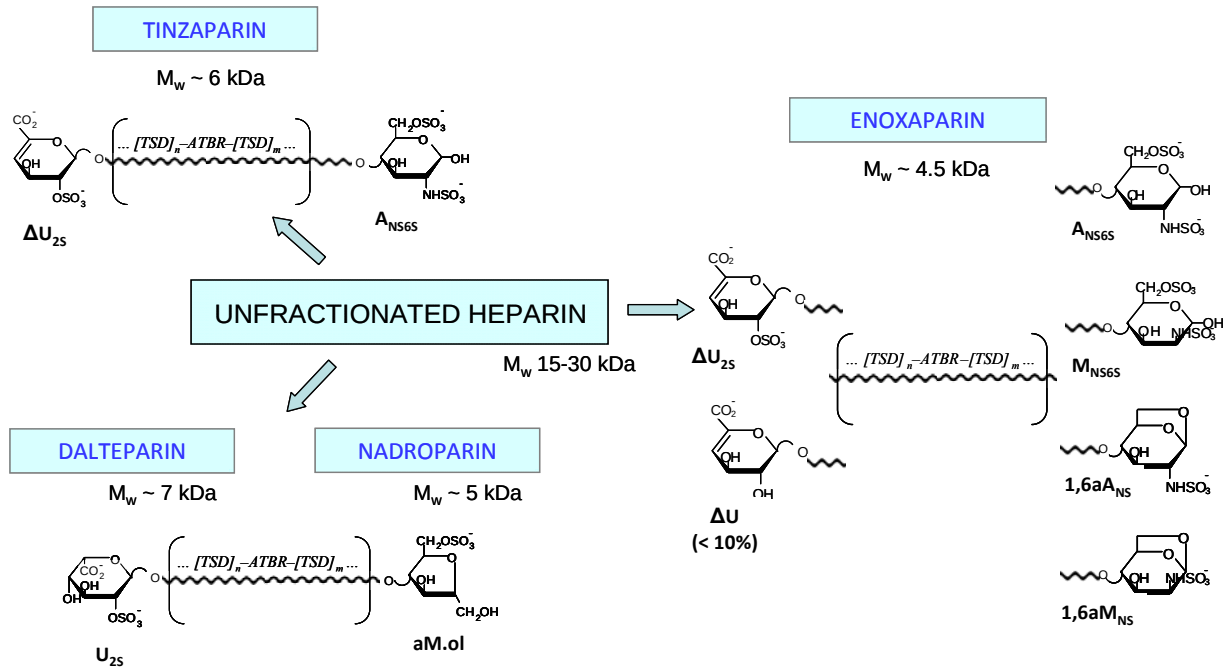



Fig.1.15. Simplified structures of tinzaparin, enoxaparin, dalteparin and nadroparin

Internal sequences are substantially the same as in non-depolymerized heparin

TSD – trisulfated disaccharide, ATBR – antithrombin binding region, ΔU – 4,5-unsaturated uronic acid, $\Delta U_{2S}$ – 4,5-unsaturated 2-*O*-sulfated uronic acid, U – saturated uronic acid, $A_{NS}$ – glucosamine-*N*-sulfate, $M_{NS}$ – mannosamine-*N*-sulfate, $1,6aA_{NS}$ – 1,6-anhydro-glucosamine-*N*-sulfate, $1,6aM_{NS}$ – 1,6-anhydro-mannosamine-*N*-sulfate, aM.ol – 2,5-anhydro-mannitol.

*Tinzaparin,* obtained by the enzymatic beta-eliminative depolymerisation by heparinase I (Fig.1.11) of heparin at the level of the $I_{2S}$, is characterized by unsaturated sulfated uronic acid ($\Delta U_{2S}$) at the NRE and *N*-sulfated glucosamine at the RE of its chains (Fig.1.15). Due to the highly selective and "soft" enzymatic depolymerization, tinzaparin is considered the LMWH most similar to UFH. LR present in the starting heparins are also preserved.

*Enoxaparin* is obtained by chemical beta-elimination, a less selective process, that includes two critical steps: esterification of heparin to the corresponding heparin benzyl ester followed by alkaline β-eliminative cleavage of the intermediate benzyl ester (*Lee* et al., 2013). As a result of this treatment, the major enoxaparin oligosaccharides bear unsaturated uronic acids ($\Delta U_{2S}$ or $\Delta U$) at the NRE, while a partial epimerisation of the A residues leads to formation of its isomer D-mannosamine (M) residues at the RE. The mechanism of epimerization was studied by *Toida* and co-workers (1996) using heparan sulfate tetrasaccharides containing terminal *N*-acetylated glucosamine (Fig.1.16). It is worth noting that mannosamine residues found in enoxaparin are mainly *N*-sulfated (*Guerrini* et al., 2007), probably because of the high sulfation degree of heparin starting material, so that, the basic β-eliminative depolymerisation, a statistically occurring process, cleaves with higher probability the glycosidic bond after *N*-sulfated glucosamine.

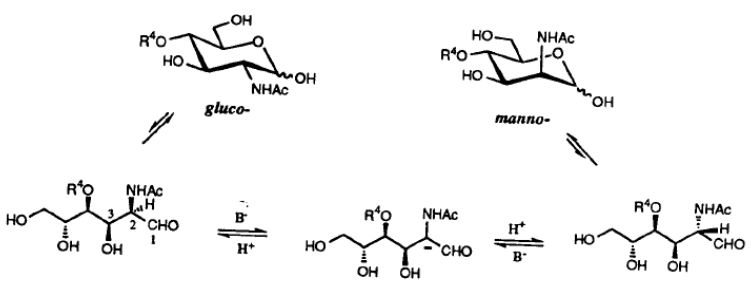



**Fig.1.16.** Proposed mechanism of base-catalyzed epimerization of terminal glucosamine into mannosamine
(reported from *Toida* et al., 1996)

Basic treatment may also cause desulfation and dehydration of the RE amino sugars, so that several enoxaparin oligosaccharides bear unnatural epimeric 1,6-anhydro-glucosamine (1,6aA) and mannosamine (1,6aM) (Fig.1.15), which are prevalently *N*-sulfated (*Guerrini* et al., 2007). Other minor residues that can be generated during basic depolymerisation are an

C(2)-C(3) epoxy uronic acid, L-galacturonic acid GalA (as a result of the opening of the epoxide ring) and $G_{2S}$ residues: ~2.5% of total uronic acid content for GalA and $G_{2S}$, and < 0.5% for the epoxy uronic acid (*Guerrini* et al, 2007).

*Dalteparin* and *nadroparin* are obtained by controlled deaminative treatment with nitrous acid followed by borohydride reduction leading to the formation of 2,5-anhydro-mannitol (aM.ol) ring at the RE and, unlike the other two LMWHs, saturated I or G at the NRE (Fig.1.17). As a minor modification, the deamination may lead to the formation of the internal ring contracted (Rc) 2-deoxy-2-C-hydroxymethyl pentofuranosidic residues (*Baer* et al., 1985; *Horton* et al., 1973; *Shively* et al, 1970).

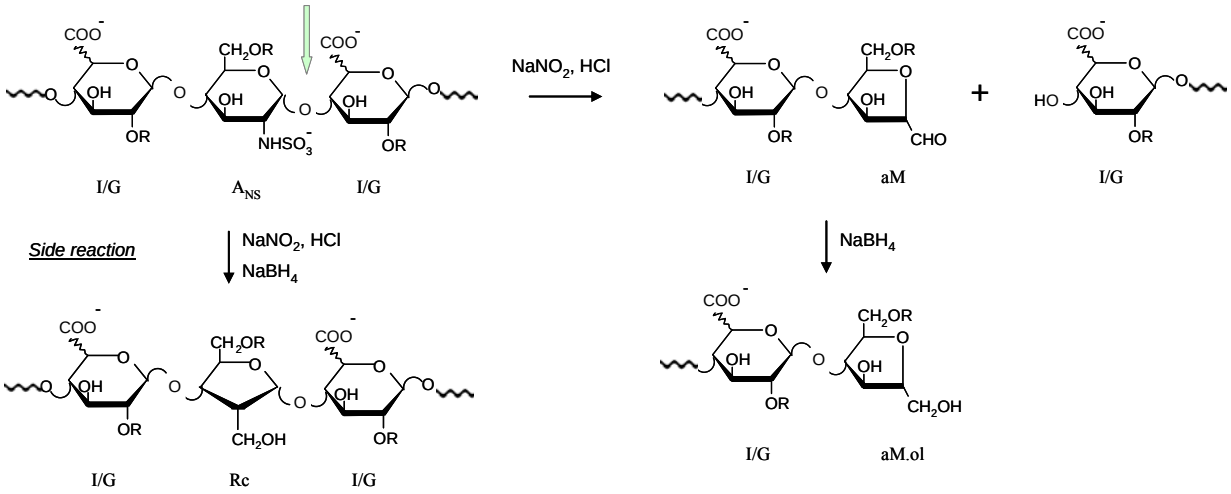



**Fig.1.17.** Deaminative heparin depolymerization used for dalteparin preparation

Unlike *N*-sulfated glucosamines, *N*-acetylated glucosamines are resistant to nitrous acid (*Guo* and *Conrad* et al., 1989; *Turnbull* et al., 1999).
I/G – iduronic/glucuronic acid, $A_{NS}$ – glucosamine-*N*-sulfate, aM – 2,5-anhydro-mannose, aM.ol – 2,5-anhydro-mannitol, Rc – ring contracted residue

***HIGHLIGHTS***

*In general, even if it may slightly modify the internal sequences, depolymerisation processes preserves in the LMWHs the major biological functions of heparin. However, both chemical and enzymatic depolymerisation, modifying the heparin sequences at the site of chain*

*cleavage, make each LMWH structurally (Fig.1.15), and, to some extent biologically, different from each other (Linhardt et al., 1990). Consequently, the structural analysis of LMWHs is more complicated due to their higher heterogeneity induced by depolymerisation processes.*

### I.3. Glycol-split heparins. Structure and biological activity

In the last decade, several new therapeutic uses of heparins and LMWHs have been envisaged in various fields (*Lever* et al, 2002), especially in inflammation and cancer (*Fuster* et al., 2005). To reduce risks of bleeding in non-anticoagulant/non-antithrombotic uses of heparins and LMWHs, removal or inactivation of their ATBR is required. Although removal of the ATBR-containing chains could be achieved exploiting their interaction with AT (*Höök* et al., 1976) or by enzymatic digestion (*Xiao* et al, 2011), a more practical way of reducing anticoagulant properties, without impairing other biological activities, is periodate oxidation of heparins (*Naggi*, 2005; *Mousa* et al., 2006).

Periodate oxidation is a classical reaction in carbohydrate chemistry that has been widely applied to determine vicinal carbons bearing unsubstituted hydroxyl and/or amino groups (*Perlin* et al., 2006). The mechanism includes formation of a monoester between periodate ion and hydroxyl group, the formation of a cyclic ester, and, finally, the ester decomposition into the oxidized products (Fig.1.18).

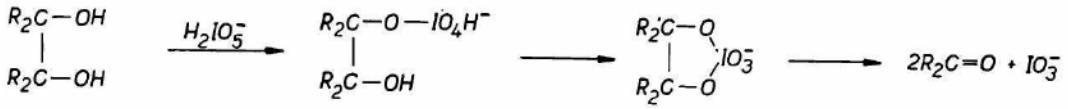



**Fig.1.18**. Scheme of the mechanism of 1,2-diol oxidation by periodate (*Babor* et al., 1973)

When applied to heparins, periodate oxidation mainly leads to the splitting of the C(2)-C(3) bond of non-sulfated U, including the essential G within the ATBR (*Islam* et al., 2002) (Fig.1.19). Also $G_{LR}$ and Xyl residues of the LR sequence, containing vicinal hydroxyls, should be considered possible sites of oxidation (Fig.1.19c,d). The primary products of the glycol-

splitting reaction (the polyaldehydic oxy-heparins) are usually stabilized by borohydride reduction, leading to reduced oxy-heparins (RO-heparins) (*Naggi*, 2005). Saturated NRE-groups may be also affected by glycol-splitting due to the presence of 3,4-diol group (Fig.1.19). Under uncontrolled reaction conditions, first of all, pH, side hydrolytic reactions may take place, contributing in increased microheterogeneity and polydispersity of the gs-heparins. Both acidic and basic conditions can cause the cleavage of gs-heparin chains. Even if not studied in detail, acetal fragments of the gs-units are known to be labile and susceptible to hydrolysis (*Conrad* et al., 1992). Their hydrolytic cleavage should generate two shorter oligosaccharides, the first one, bearing a remnant of gsU (R) and the second one, bearing a glucosamine at its NRE (Fig.1.19e). These reactions rely on the principle of the so-called Smith degradation of carbohydrates that, historically, has been widely used for heparin depolymerisation. Extremely strong acid conditions may cleave also the glycosidic bonds (*Jandik* et al, 1996). It is worth noting that partial hydrolysis is used to reduce the molecular weight and, consequently, to improve the bioavailability of some gs-heparin under development (*Leitgeb* et al., 2011; *Zhou* et al., 2011).

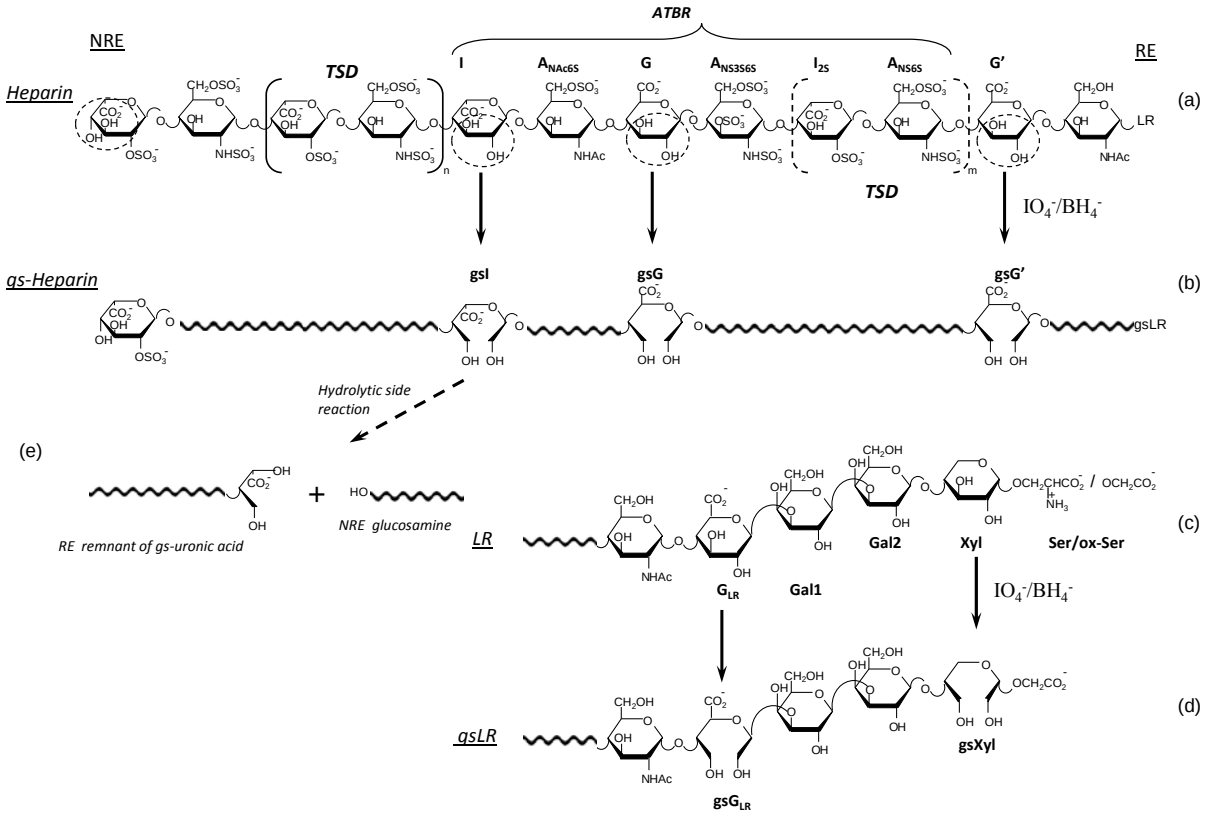



**Fig.1.19.** Simplified formula of a representative ATBR-containing chain of PMH (a) and the corresponding gs derivative (b) obtained by periodate oxidation/borohydride reduction

Structure (a) is adapted from (*Xiao* et al., 2011). The structure of the "full" linkage region LR is reported in (c) and (d) for heparin and gs-heparin, respectively, where $G_{LR}$ is the G residue in the LR, Gal1 and Gal2 are two galactose residues, Xyl is xylose, Ser is serine (*Iacomini* et al., 1999). The two glycol-split residues ($gsG_{LR}$ and gsXyl) in the LR of gs-heparin were not previously described (see text). Circled residues in the unmodified heparin structure (a) and (c) highlight unsubstituted diol groups susceptible of glycol-splitting

Other properties that are not dependent on the intact structure of the AT-binding sequence are usually preserved in gs-heparins. Moreover, most of the literature data indicate that the glycol-splitting gives rise to a higher protein-binding activity of gs-heparins than their parent GAGs (*Casu* et al., 2002; *Casu* et al., 2010; *Naggi*, 2005). One of the factors is the increased local mobility of gs residues, which act as flexible joints along the polysaccharide chains, thus, facilitating interactions with the heparin/HS binding sites of proteins (Fig.1.20).

(a)

(b)

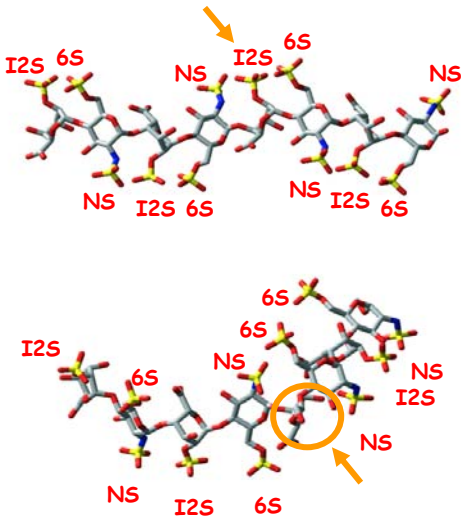



**Fig.1.20**. Molecular models of the most favored conformations of heparin octasaccharide $I_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-**$I_{2S}$**-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ (a) and a "mono-glycol-split" octasaccharide $I_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-**gsI**-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ (b)

*(Casu and Naggi, 2003)*

The ability of gs-heparins to mimic heparin biological functions and their decreased anticoagulant activity led to the generation of a number of biologically active non-anticoagulant gs-heparins, differing for molecular weight and extent of glycol-splitting (Table 1.2). Some of these potential drugs under development (such as M402, *Zhou* et al, 2011) were obtained from LMWHs. However, their structural features have not been reported yet.

**Table 1.2.** Glycol-split heparins and their biological activity

| Glycol-split derivative | Biological activity | Reference |
|---|---|---|
| gs-heparin | Retention of antilipemic activity | *Casu* et al., 1986 |
| gs-heparin | Antiangiogenic, antimetastatic activity (FGF2 and heparanase inhibitor) | *Lapierre* et al, 1996 |
| gs-LMWH ("Vasoflux") | Anticoagulant activity | *Weitz*, 1999 |
| undersulfated gs-LMWH | Antiangiogenic activity (FGF2 antagonists) | *Casu* et al, 2004 |
| undersulfated gs-LMWH | Antiangiogenic activity (VEGF antagonists) | *Pisano* et al, 2005 |
| gs-LMWH | Antimalarial activity (desegregating and inhibiting parasite adhesion) | *Vogt* et al, 2006 |
| gs-LMWH | Antimetastatic activity | *Mousa* et al, 2006 |
| gs-heparins | Antimetastatic activity (P-selectin inhibitors) | *Hostetter* et al, 2007 |
| gs-LMWH ("DF01", obtained by hydrolysis of a gs-heparin) | Labor pain reducing effect | *Ekman-Ordeberg* et al, 2009 |
| *N*-acetyl-gs-heparin ("SST0001") | Antiangiogenic, antimetastatic (heparanase inhibitor) | *Naggi* et al, 2005<br>*Ritchie* et al, 2011<br>*Shafat* et al, 2011 |
| gs-LMWH(s) | Anticancer activity<br>Increased tumor chemo responsiveness | *Phillips* et al, 2011 |
| gs-LMWH(s) ("M402") | Anticancer activity through (multiple inhibition pathway) | *Zhou* et al, 2011 |
| hydrolyzed gs-heparin ("DFX232" and "Sevapurin") | Antimalarial activity | *Leitgeb* et al., 2011 |

***HIGHLIGHTS***

*In spite of the current growing interest on gs-heparins as potential anticancer and anti-inflammatory drugs, their structural characterization is still a challenging task, firstly, due to the high microheterogeneity of the starting material (determined predominantly by the biosynthesis pathway and structural complexity introduced during the manufacture processes). Moreover, uncontrolled reaction conditions may cause side reactions, such as hydrolysis at the level of gs residues* (Fig.1.19e; *Conrad* et al., 1992) *as well as hydrolytic cleavage of*

*glycosidic bonds (Jandik et al, 1996). In the case of LMWHs, the end-groups generated during the depolymerisation processes induce additional heterogeneity, making their analysis even more complex. Development of efficient and complementary analytical techniques has become increasingly important to structurally characterize these molecules and to identify motifs responsible for their interactions with proteins, in order to rationalize the seqrch of new potential drugs..*

### I.4. Analytical methods for characterization of heparins and their derivatives

Despite their medical and biological importance, heparin and HS are still incompletely characterized in terms of their molecular sequences (see section I.1). Most work concerning the structural characterization of heparin and HS has been performed on the former, largely as result of its abundance, commercial availability and lower heterogeneity than HS. While numerous studies have been published about the structural characterization of heparin in the last decades, detailed structural characterization of gs-heparins and gs-LMWHs have not been reported yet (see section I.3).

Multistep and hybrid approaches are required for structurally characterizing heparin/HS and their derivatives. Due to the heparin/HS microheterogeneity (epimerization degree, sulfation and acetylation pattern) and polydispersity, their sequences are usually characterized in statistical terms, by nuclear magnetic resonance (NMR) spectroscopy and high-performance liquid chromatography coupled with mass spectrometry (HPLC-MS). Monodimensional ($^{1}$H and $^{13}$C) and multidimensional NMR have been employed for the characterization of the intact chains (the so-called *top-down* approach) of both unfractionated and LMWHs. In the case of LC-MS analysis, the *bottom-up* approach, based on the identification/quantification of the fragments obtained by either partial or complete digestion, is usually applied. Together with NMR and LC-MS analysis, molecular weight distribution analysis and hydrodynamic properties determination may provide complementary structural information.

### I.4.1. Nuclear magnetic resonance spectroscopy for heparin characterization

NMR spectroscopy is recognized to be one of the most effective tools for the structural elucidation of oligosaccharides as well as heparins and LMWHs (*Guerrini* et al., 2005 and 2007; *Fu* et al., 2013). For unmodified unfractionated heparins and LMWHs it was demonstrated that the monodimensional $^{1}H$ NMR spectroscopy permits their structural characterization in terms of the monosaccharide composition, sulfation/acetylation pattern, position linkages, and relative $I_{(2S)}$ and G amounts (*Guerrini* et al., 2005*, Guerrini* et al., 2007) (Fig.1.21). The anomeric region (4.6 – 5.6 ppm) is the most informative and can be used for distinguishing the same residues with different environment (so-called sequence effect). $^{1}H$ NMR was successfully used for the differentiation of PMH and BLH (*Guerrini* et al., 2001; *Fu* et al., 2013), for the control of heparin contaminations by over-sulfated CS (*Keire* et al., 2010). $^{13}C$ NMR analysis, utilized to determine the animal origin of heparin (*Casu* et al., 1996), was also successfully applied for the gs-heparin analysis (*Casu* et al., 2002). Determination of total content of gs-uronic acids (gsU) was possible due to the difference in chemical shifts between anomeric carbons of unmodified non-sulfated uronic acids (~105 ppm) and those of the generated gs-units (~107 ppm). This approach can turn out useful for both quantifying total gsU content and monitoring the reaction completeness (Fig.1.22). However, the use of monodimensional NMR analysis of complex GAGs is often limited by strong signal overlapping, preventing quantitative analysis. The spectrum heterogeneity increases for LMWHs due to the presence of unnatural residues (such as end groups in LMWHs, Fig.1.15) causing additional signals (*Guerrini* et al., 2007).

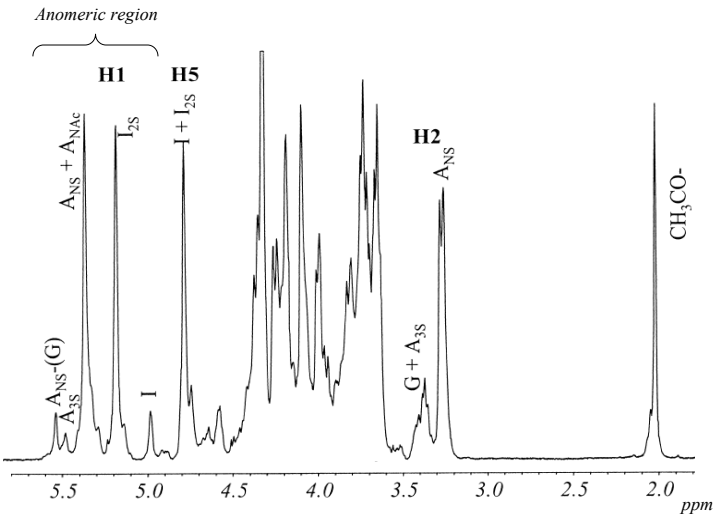



**Fig.1.21.** $^{1}H$ monodimensional NMR spectrum of unmodified pig mucosal heparin (from *Guerrini et al., 2001*)

$I/I_{2S}$ – non-sulfated/2-*O*-sulfated iduronic acid, G – glucuronic acid, $A_{NS}/A_{NAc}$ – *N*-sulfated/*N*-acetylated glucosamine, $A_{3S}$ – glucosamine-*N*,3-*O*-disulfate (mostly, 6-*O*-sulfated)

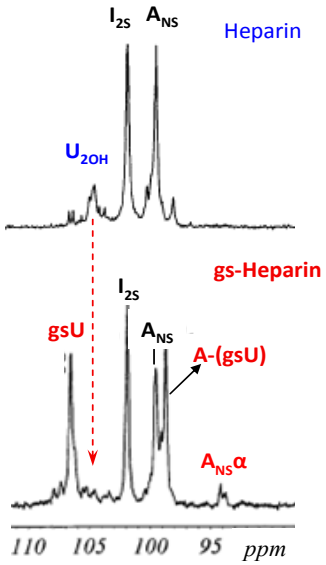



**Fig.1.22.** $^{13}C$ monodimensional NMR spectrum of heparin and 50%-2-*O*-desulfated gs-heparin (adapted form *Casu* et al., 2002)

$I_{2S}$ – 2-*O*-sulfated iduronic acid, $A_{NS}$ – glucosamine-*N*-sulfate, $U_{2OH}$ – non-sulfated uronic acid, $A_{NS}\alpha$ – reducing glucosamine-*N*-sulfate

Two-dimentional (2D) NMR has wider applications. For example, heteronuclear single quantum coherence (HSQC) spectroscopy, based on the $^{1}H$ – $^{13}C$ correlation, permits to resolve signals that overlap in the monodimensional NMR spectra of heparins and LMWHs (*Guerrini* et al., 2005; *Guerrini* et al., 2007). Higher signal resolution, in turn, permits both identification of the constituent monosaccharide residues (Fig.1.23) and their quantitative evaluation.

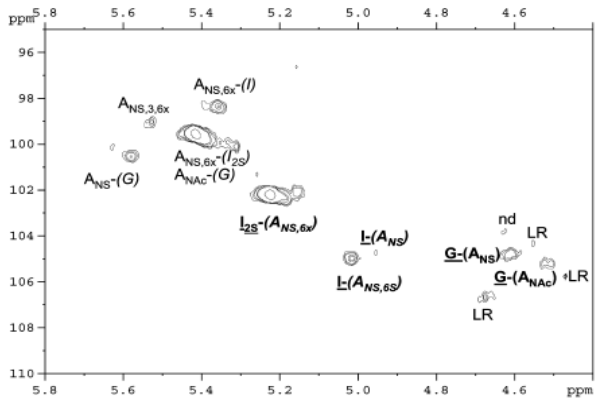



**Fig.1.23.** Partial HSQC spectrum of porcine mucosal heparin (from *Guerrini* et al., 2005)

By 2D NMR analysis, sequence effect can be observed in the anomeric region: the same residues 1→4 linked to different units (shown in the parenthesis) have different and well-defined cross-peaks

*A – glucosamine, I – iduronic acid, G – glucuronic acid, LR – linkage region, nd – not determined*

Quantitative analysis is based on the selection of analytical signals where the effect of $^1J_{C-H}$ coupling is minimal and differences in relaxation effects are sufficiently small to be neglected. The $^1J_{C-H}$ coupling constants of H1/C1 and H2/C2 were shown not to differ significantly for various internal and RE residues of heparin and LMWHs (Table 1.3). Hence, integration of well-resolved anomeric signals can be performed for quantifying the monosaccharide residues in heparins and LMWHs. Unlike the others moieties, for quantifying 2-*O*-sulfated or non-sulfated 4,5-unsaturated uronic acid ($\Delta U_{(2S)}$) typical for tinzaparin and enoxaparin, the integration of the H4/C4 cross peaks that are well-separated from the others is recommended (*Guerrini* et al., 2007), even if its $^1J_{C-H}$ value slightly differs from the other signals of uronic acids (Table 1.3). In contrast, the corresponding anomeric signals (H1/C1 $\Delta U_{(2S)}$) are in a complex region, so that their integration leads to higher errors of quantitative analysis (*Guerrini* et al., 2007). However, when high resolution NMR spectrometers are used, the integration of the anomeric signals may provide more accurate quantitative results. For quantifying aM.ol residues typical for dalteparin, the integration of H5/C5 signals was performed, because its $^1J_{C-H}$ value was found to be much closer to that of the other glucosamine residues (Table 1.3), permitting to compare sample differences with good precision.

**Table 1.3.** $^1J_{C-H}$ coupling constants (Hz) of some internal residues and end groups of LMWHs (*Guerrini* et al, 2007)

| | Tinzaparin, enoxaparin, dalteparin | | | Enoxaparin | | Dalteparin | Tinzaparin, enoxaparin, dalteparin | | Tinzaparin, enoxaparin | Enoxaparin |
|---|---|---|---|---|---|---|---|---|---|---|
| | Glucosamines | | | | | | Uronic acids | | | |
| | $A_{NS}$-($I_{2S}$) | $A_{NS}$-(-I) | A* | 1,6a$A_{NS}$ | 1,6a$M_{NS}$ | aM.ol | $I_{2S}$ | G | $\Delta U_{2S}$ | $G_{2S}$ |
| H1/C1 | 173 | 173 | 174 | 177 | 178 | | 173 | 165 | 176 | 166 |
| H4/C4 | | | | | | 148 | | | 168 | |
| H2/C2 | 139 | | 140 | 138 | 142 | 147 | | | | 151 |
| H5/C5 | | | | | | *151* | | | | |

$A_{NS}$ – glucosamine-*N*-sulfate, A* – glucosamine-*N*,3-*O*,6-*O*-trisulfate, 1,6a$A_{NS}$ – 1,6-anhydro-glucosamine-*N*-sulfate, 1,6a$M_{NS}$ – 1,6-anhydro-mannosamine-*N*-sulfate, aM.ol – 2,5-anhydro-mannitol, $I_{2S}$ – iduronic acid 2-*O*-sulfate, G – glucuronic acid, $\Delta U_{2S}$ – 4,5-unsaturated uronic acid 2-*O*-sulfate, $G_{2S}$ – glucuronic acid 2-*O*-sulfate

Among the other 2D NMR methods that are being used for heparin analysis, there are correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY), heteronuclear

multiple-bond correlation spectroscopy (HMBC), and nuclear Overhauser effect spectroscopy (NOESY) (*Guglieri* et al., 2008). Modern methods for the mathematical analysis of multiple complex datasets can provide objective, systematic interpretations of the spectra, and have the advantage that more than two types of spectra can be combined. Principal component analysis (PCA) has been applied to NMR and CD of heparin and other glycosaminoglycans (*Rudd* et al. 2012).

Apart from its use for qualitative and quantitative analysis, NMR can be a powerful tool for studying interactions between heparin and heparin/HS-binding proteins. The main drawback is that isolation or synthesis of structurally homogeneous oligosaccharides is required. NMR studies were performed on the interaction of heparin oligosaccharides with AT (*Guerrini* et al, 2006; *Viskov* et al, 2013) and growth factors (*Guglieri* et al., 2008), etc.

***HIGHLIGHTS***

*The 2D NMR method can disclose structural differences and reveal particular sequences of heparin/HS and their derivatives, in terms of the relative distribution of the major monosaccharide residues and certain functional groups (N-, 6-O- and 3-O-substitutions). However, it provides information only about the average composition. Moreover, due to its relatively low sensitivity, this method could have challenges for analyzing intact heparin samples, containing unusual sequences. In favourable cases, when standard oligosaccharides are available, NMR can be a powerful technique for heparin oligosaccharide–protein interaction studies.*

I.4.2. Hybrid LC/ESI-MS analysis for heparin structural characterization

Direct ESI-MS analysis can be performed for simple mixtures of oligosaccharides (*Behr* et al., 2005; *Camara* et al, 2007). However, *hybrid methods*, combining LC separation and MS identification, are most often used to overcome the problem of the high heterogeneity and

polydispersity. The main problem in the MS of heparins is that desulfation may occur during the spectrum acquisition, because of high acidity of the sulfate-groups and their lability. However, with the development of soft *electrospray ionization* (ESI) methods, it was shown that desulfation can be minimized by optimizing the ionization parameters and solvent composition (*Huber and Krajete*, 2000). Careful control of the sample solution and mobile phase pH is still necessary to avoid the desulfation during the chromatographic separation and MS detection. Nowadays, negative-mode ESI/MS is being increasingly applied to heparin oligosaccharides structural characterization (*Thanawiroon* et al, 2004; *Doneau* et al., 2009; *Langeslay* et al, 2013; *Wang* et al., 2012). Due to the development of high resolution mass analyzers, such as quadrupole – time-of-flight (Q-TOF) and Fourier transform ion cyclotron resonance (FT-ICR) MS, accurate molecular weight measurements can be made. In favourable cases, also *matrix assisted laser desorption ionization time-of-flight* (MALDI-TOF) mass spectrometric analysis, characterized by very soft ionization mechanism, can be performed without previous chromatographic separation (*Sturiale* et al., 2001). However, it works only for simple oligosaccharide fractions and cannot resolve complex mixtures due to signal overlapping.

In comparison with other analytical techniques, the MS method has several advantages: high sensitivity, detection specificity, low sample consumption, etc. Detection specificity is very important because oligosaccharide standards are generally not available, except for disaccharides. The tandem mass spectrometry (MS/MS) has provided new possibilities to determine oligosaccharide sequences. Several successful studies were performed (*Zaia* et al., 2003; *Wolff* et al., 2007), however, labile sulfate groups still limit MS/MS analysis.

Despite of the progress in heparin structural characterization, analysis of intact heparin is still complicated due to the sample complexity and high molecular weight as well as absence of databases for heparin oligosaccharides identification. Most often, in order to simplify the analytical assay, chemical- or/and enzymatic depolymerisation is employed for its sequencing.

In fact, most of the work was done on the disaccharides generated by heparin depolymerisation (*Wang* et al., 2012; *Zhang* et al., 2011; *Kuberan* et al., 2002).

Among chemical depolymerisation processes, nitrous acid treatment (*Turnbull* et al., 1999; *Stringer* et al., 2003) is the most widely used. The degradation process relies on the reaction of deamination of *N*-sulfated glucosamine, often followed by borohydride reduction, leading to the formation of the 2,5-anhydromannitol (aM.ol) RE residues typical for dalteparin (*see I.2*). The 2,5-anhydromannose (aM) residues are easily distinguishable from natural glucosamine residues by MS analysis due to the mass difference of 17 Da (or 15 Da for the reduced aM.ol form). However, the absence of a chromophore may limit the adoption of this technique for UV detection. It is important to note that *N*-acetylated glucosamines are resistant to oxidation by nitrous acid, providing a certain selectivity of such a treatment (*Guo* and *Conrad* et al., 1989; *Turnbull* et al., 1999). In order to obtain more structural information, chemical degradation is often combined with enzymatic cleavage (*Turnbull* et al., 1999; *Stringer* et al., 2003).

For the enzymatic depolymerisation heparin lyases I, II and III are usually used (*see I.1.4*). Enzymatic reactions afford more specific cleavage than the chemical ones, and as a result of this specificity, exhaustive depolymerisation of heparins is also possible. Disaccharide building blocks analysis of completely depolymerised heparin is widely used for characterizing individual heparin samples and comparing heparins of different animal and tissue origins (*Zhang* et al., 2011; *Fu* et al., 2013) as well as for controlling the LMWHs preparation (*Lee* et al., 2013). Identification is favored and quantitative analysis is possible because standards of disaccharides with different sulfation/*N*-acetylation patterns are commercially available. This methodology provides average disaccharide composition, but does not give information about heparin sequences. Because the functional sequences of heparin/HS are usually longer, partial depolymerisation (by each of the three heparinases) and analysis of larger oligosaccharides may be also useful for heparin/LMWH mapping (*Mourier* et al., 2004; *Zhang* et al., 2011; *Fu*

et al, 2013; *Lee* et al., 2013) mapping. A preliminary reduction of generated oligosaccharides is often used to mark the disaccharide unit located at the RE (*Mourier* et al., 2004).

However, it is still complicated to determine, for example, the sulfate group distribution.

*Strong anion exchange (SAX) chromatography* has been traditionally used for the highly charged heparin oligosaccharide separation as sample preparation step (*Pervin* et al., 1995, *Xiao* et al., 2011). It was also applied for the quantification of oversulfated CS contaminating heparin samples (*Keire* et al, 2010). The main drawback of this method is that it cannot be directly coupled with MS due to the use of strong electrolytes as eluents. However, it can be used with a UV detector for analytes bearing chromophores when their standards are available, for example, for analysis of unsaturated disaccharides (*Zhang* et al., 2011).

Among the other chromatographic approaches amenable to online ESI-MS analysis there are *size exclusion chromatography* (SEC), *hydrophilic interaction liquid chromatography* (HILIC), and *ion-pair reversed phase liquid chromatography* (IPRP-LC).

An online ***SEC/ESI-MS*** method was developed for determining the molecular weights of tinzaparin oligosaccharides (*Henriksen* et al., 2004). One of the drawbacks of SEC is that the very high salt concentration in mobile phases leads to more complex mass spectra due to the formation of numerous adducts with cations. For solving this problem and minimizing salts influence on mass spectra an online cation exchange column has to be coupled with the SEC one (*Henriksen* et al., 2004). Another limitation of SEC is its low resolution that does not provide complete separation of oligosaccharides with similar chain length. In spite of low resolution power and technical problems in coupling SEC with MS, nowadays, SEC is widely used in either high or low pressure mode, first of all, for sample preparation (fractionation and desalting) as well as for molecular weight determination (see I.4.3). It can be used also for studying heparin–protein interactions when structurally homogeneous oligosaccharides are available. For example, SEC-MS, along with NMR, was successfully applied for studying the interactions between ATBR-containing octasaccharides and AT (*Viskov* et al., 2013).

Recently, the ***HILIC/ESI-MS*** method was developed for screening and quantitative analysis of a library of AT-binding heparin/HS oligosaccharides (*Naimy* et al, 2008) and for sequencing and profiling heparins of different origin (*Fu* et al, 2013). Generally, typical HILIC stationary phases consist of classical silica gels modified with polar functional groups (*Buszewski* et al., 2012). The separation is based on different overall polarity determined by size, sulfation pattern and N-acetylation content. However, this technique is characterized by a rather low resolution, insufficient for isomer separation.

Nowadays, one of the most widely used chromatographic approaches for the heparin oligosaccharides analysis is an ***ion pair reversed phase high performance liquid chromatography (IPRP-HPLC)*** that can be easily coupled with ESI-MS. For separating strong anions, like sulfated oligosaccharides, octadecyl (C18) columns, dynamically modified by lipophilic alkyl ammonium salt present in eluent, are often used. Briefly, the separation is based on the different analyte distribution between eluent and stationary reversed (e.g., C18) phase dynamically modified by an aliphatic amine (Fig.1.24). Given the dynamical mechanism of the column modification, the equilibria occurring in the eluent should be considered. So, the separation mechanism may be also explained as the distribution of oligosaccharide-amine ion pair (having total neutral charge) between reversed phase and eluent (Fig.1.24). Given the separation mechanism, oligosaccharides are eluted from the column in dependence, first of all, on their charge. It is also sensitive to the sulfate distribution and steric hindrance, providing positional isomers separation (*Kuberan* et al, 2002; *Thanawiroon* et al., 2004; *Lawrence* et al, 2004; *Doneanu* et al., 2009; *Jones* et al., 2010).The method allows the direct analysis of aqueous solutions of oligosaccharides without derivatization.

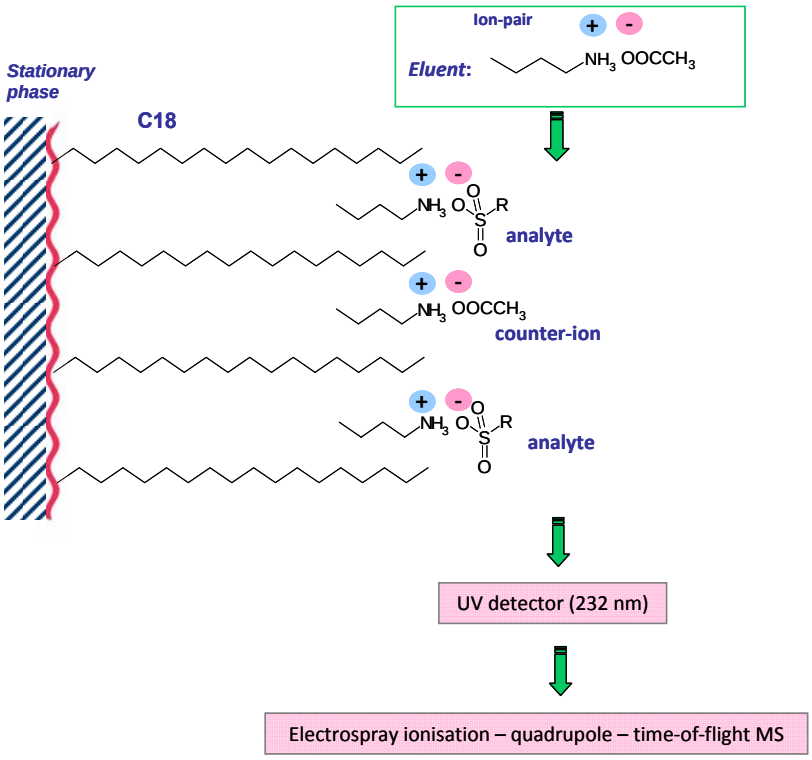



**Fig.1.24** Separation mechanism in ion-pair reversed phase mode

RPIP separation is believed to occur through a combination of two mechanisms.
In the classical model, the hydrophobic ion pair agent (an amine) and the analyte (sulfated oligosaccharide) form in the eluent a neutral species which then interacts with the hydrophobic stationary phase. The dynamic ion-exchange model suggests that the ion pair amine is first adsorbed onto the stationary phase, dynamically modifying its surface, which then acts as ion exchange sites for the oppositely charged oligosaccharides.

This method has the advantage of simplifying the mass spectra by preventing the formation of adducts of oligosaccharides with alkali and earth alkali cations, almost always present in the heparin samples. Minimizing cation adduction provides higher sensitivity of MS detection. Like the other methods, this technique limits the detection of larger oligosaccharides due to the very low ionization. For instance, several studies reported that it was possible to obtain mass spectra only up to tetradecasaccharides (*Thanawiroon* et al, 2004). The use of *n*-hexylamine with 1,1,1,3,3,3-hexafluoro-2-propanol as a counter-ion (*Doneau et al.,* 2009) and dibutylamine (*Langeslay* et al, 2013) allowed to detect oligosaccharides up to dp 22.

The high specificity of both separation mechanism and negative-mode ESI-MS permits to avoid matrix effects and facilitates the identification of oligosaccharides. Moreover, sulfated oligosaccharides are characterized by specific isotope pattern and monoisotopic exact mass due to the presence of sulfur atoms allowing their indentification.

***HIGHLIGHTS***

*Nowadays, ESI-MS coupled with HPLC is a powerful technique that, in part, permits to overcome the problems of heparin structural analysis associated with its compositional microheterogeneity. Generally, it is performed on the fragments generated by specific enzymatic or/and chemical treatment. However, often pre-digestion is not sufficient for heparin analysis, and coupling of different chromatographic techniques, such as SEC-fractionation and further LC/ESI-MS analysis, is required to improve the resolving power. It is worth noting that LC-MS analysis cannot solve several tasks that can be tackled by NMR spectroscopy, and vice-versa. For example, NMR permits to distinguish iduronic and glucuronic acids as well as enoxaparin typical terminal groups 1,6-anhydro-glucosamine and 1,6-anhydro-mannosamine, whereas LC-MS analysis, even if can separate such isomers, cannot distinguish them because of the same mass. On the contrary, LC-MS allows to determine individual heparin oligosaccharides and to fingerprint oligosaccharide mixtures, whereas NMR provides information on average monosaccharide composition. Evidently, these two powerful analytical methods may provide more structural information, when used in combination.*

I.4.3. Molecular weight distribution and hydrodynamic properties of heparin in solution by SEC-TDA

Molecular weight distribution is one of the parameters used for describing the polydispersity of heparin, contributing to its structural heterogeneity. A convenient way to describe molecular weight distribution is a combination of number-average mean molecular weight ($M_n$) and weight average mean molecular weight ($M_w$), while ratio $M_w/M_n$ expresses the polydispersity of polymers.

$$M_n = \frac{\sum_i N_i M_i}{\sum_i N_i} \quad \text{and} \quad M_w = \frac{\sum_i N_i M_i^2}{\sum_i N_i M_i}$$ where $N_i$ – the number of molecules of molecular weight $M_i$

SEC represents the most widely used method for measuring molecular weight and size distribution (*Mulloy* et al., 1997; *Bertini* et al., 2005; *Gray* et al., 2008; *Sommers* et al., 2011). However, a set of heparin reference standards with well-defined molecular weight, which are not commercially available, is required for column calibration. In order to overcome this problem, SEC can be coupled with triple detector array (TDA), including three online detectors, right-angle laser light scattering (RALLS), refractometer (measuring refractive index (RI)) and viscosimeter. Combination of these detectors permits to characterize the molecular weight distribution without calibrating the chromatographic column (Fig. 1.25, *Bertini* et al., 2005). The elaboration of LS and concentration RI detector responses gives molecular weight values:

$$RI = K \cdot dn/dc \cdot C \quad \text{and} \quad LS = K' \cdot Mw \cdot (dn/dc)^2 \cdot C ,$$

where *RI and LS* – refractometer and light-scattering response, *K/K'* – curve constants, *dn/dc* – differential index of refraction equal to 0.12 ml/g (*Bertini* et al., 2005), *C* – concentration, $M_w$ - weight average mean molecular weight. The dn/dc parameter is used to convert *RI* response to the concentration value, that, then, is used to calculate $M_w$ using LS data.

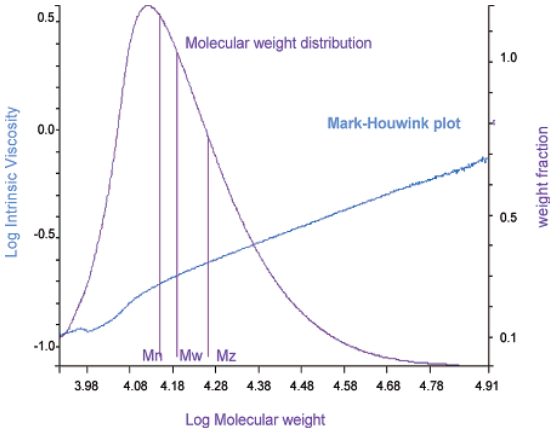



**Fig.1.25.** Molecular weight distribution of an unfractionated heparin and Mark-Houwink plot (from *Bertini* et al., 2005)

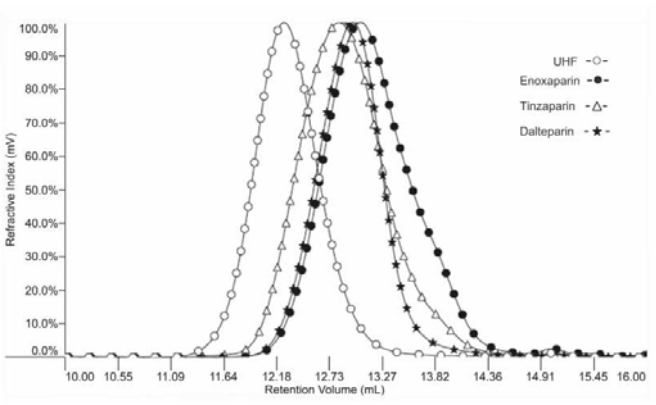



**Fig.1.26.** SEC-TDA profile of three LMWHs (tinzaparin. enoxaparin, dalteparin) compared with an unfractionated heparin (UFH) (from *Bisio* et al., 2009)

The method was successfully applied for the molecular weight determination of unfractionated heparins and dermatan sulfates (*Bertini* et al., 2005) as well as commercially available LMWHs and their fractions with and without affinity towards AT (*Bisio* et al, 2009).

The SEC-TDA approach can also provide parameters describing the hydrodynamic properties of GAGs. $M_w$ value together with the viscosity data permits to calculate the value of the intrinsic viscosity ([η]): $\eta = K \cdot [\eta] \cdot C$, where $\eta$, $K$, and $C$ refer to the measured viscosity, curve constant, analyte concentration, respectively. The obtained results, in turn, may be used to calculate the parameters $a$ and $K$ for the Mark-Houwink relationship: $[\eta] = K \cdot M^a$ through the double logarithmic plot of intrinsic viscosity versus molecular weight (Fig.1.25, *Bertini* et al., 2005).

SEC coupled with multiple detector was applied to monitor the Mark-Houwink parameters of neutral polysaccharides upon glycol-splitting (*Vold* et al., 2006). The reported results show that the persistence length, generally used for describing the stiffness of the macromolecule, decreases with increasing degree of glycol-splitting. This is in agreement with the molecular modelling data obtained for gs-oligosaccharides (Fig.1.20, *Casu* et., 2004).

***HIGHLIGHTS***

*Molecular weight of polydisperse heparin-related samples has to be described using number average $M_n$ and weight average $M_w$ molecular weight as well as their ratio. The use of TDA as detector after SEC separation permits to overcome the problem of the absence of standards for microheterogeneous and polydisperse polymers such as heparin. Combination of LS detector and viscosimeter may provide information also about hydrodynamic properties that may be extremely important for studying biological functions of heparins and their related compounds (for example, gs-heparins).*

## CHAPTER II: OBJECTIVES OF THE WORK

In spite of the increasing interest on the biological activity of gs heparins, their structural characterization has not been reported yet. As mentioned in the Introduction, difficulties in their analysis are due to, first of all, high chemical heterogeneity and polydispersity of starting material, increasing for LMWH during depolymerisation processes. Glycol-splitting reactions, along with possible side reactions, are expected to introduce further local modifications and further contribute to the microheteregeneity and polydispersity of the final gs-product. Evidently, pharmaceutical development of gs-heparins requires setting up appropriate analytical methods to control and guarantee preparation reproducibility and to establish correlations between structure and biological activity. The main analytical techniques used for the heparin characterization (NMR and LC-MS) should be considered the best candidates for solving this challenging task.

The **overall goal** of the Ph.D project is to develop an efficient analytical approach, combining NMR and LC-MS, to structurally characterize gs-derivatives generated from various heparin/LMWH sources. The future prospective of this study includes the exploration of structural basis of gs-oligosaccharides for their optimal binding with heparanase.

To achieve the main goal of the study the following objectives have been defined:

*1. To assign all the signals in the 2D HSQC NMR spectra characteristic for gs-heparins and to compare gs-samples obtained from heparins of different sources*

For this purpose, gs-derivatives of the UFHs of different origin (differing by sulfation pattern, N-acetylation degree, relative content of G and I residues) have been prepared by periodate oxidation followed by borohydride reduction.

2. *To develop a selective and reproducible LC-MS method for gs-heparin analysis*

Among the other separation techniques used for heparin sequencing, such as strong anion exchange chromatography and capillary electrophoresis, ion-pair mode of the reversed phase

liquid chromatography has an important advantage because it can be easily coupled with mass spectrometry, that is extremely important considering the high heterogeneity of heparin samples and the absence of the oligosaccharide standards. The HPLC system was coupled with an ESI-Q-TOF mass spectrometer providing high resolution and mass accuracy.

The combination of enzymatic digestion and glycol-splitting before LC-MS analysis is expected to provide more structural information also for differentiating unmodified heparins of different sources. Possibility of the quantitative analysis should be also evaluated.

*3. To apply the developed LC-MS method and NMR analysis to the structural characterization of gs-LMWHs, considered more pharmaceutically attractive than gs-heparins*

Three commercially available LMWHs (tinzaparin, enoxaparin and dalteparin), differing mainly in molecular weight and chemical nature of the end-groups, were chosen for characterization by the developed method. The susceptibility of the particular terminal groups should be also verified for each LMWH.

*4. To determine molecular weight distribution and hydrodynamic parameters by SEC-TDA of gs-derivatives and compare them with the values obtained for unmodified samples*

Biological functions of heparin-related samples may depend on the length and conformation in solution. Given the higher local flexibility of glycol-split uronic acids, hydrodynamic properties may be affected by glycol-splitting reaction. In this context, SEC-TDA represents a promising tool for characterization of gs-heparins.

*5. To apply the developed approach for preparing and screening potential heparanase inhibitors.*

# CHAPTER III: DEVELOPMENT OF THE ANALYTICAL APPROACH FOR CHARACTERIZING GLYCOL-SPLIT HEPARINS AND THEIR STARTING MATERIALS

*Part of the present results has been published in: Alekseeva A., Casu B., Torri G., Pierro S., Naggi A. Profiling glycol-split heparins by high-performance liquid chromatography/mass spectrometry analysis of their heparinase-generated oligosaccharides. Analytical Biochemistry. 434 (2013) 112-122.*

A gs-derivative of PMH (gs-PMH), prepared by periodate oxidation followed by borohydride reduction of porcine mucosal heparin as described in the Experimental section, was chosen as a model for the IPRP-HPLC/ESI-MS method optimization and for signal assignment in the NMR spectra. Additionally, gs-derivatives of heparins of different animal and tissue origin (OMH, BMH, BLH) as well as those of the PMH fraction with high affinity towards AT, *N*-acetyl-heparosan (K5), *N*-acetylated bovine lung heparin (NAc-BLH), and partially 2-*O*-desulfated PMH were studied to verify the specificity of both the LC-MS and NMR methods. The latter gs-derivative, called also "heparanase-inhibitor-III" (HI-3), was chosen for its high inhibitory activity towards heparanase (*Casu* et al., 2004).

The wide range of the model compounds, obtained from natural and semi-synthetic heparin samples was expected to be sufficient to evaluate the potential of NMR, SEC-TDA and LC-MS techniques for both the structural characterization of individual gs-heparins and the differentiation of samples of different animal and tissue origin. Combination of glycol-splitting and specific enzymatic digestion before LC-MS analysis was expected to provide further important information about sequences in the original heparins.

## III.1. Molecular weight determination by SEC-TDA

Molecular weight determinations were performed by HP-SEC-TDA using two TSKgel G2500 and G3000 columns in order to achieve a sufficient chromatographic resolution, and 0.1 M $NaNO_3$ as eluent. All molecular weight parameters (number-average mean molecular mass

($M_n$), weight-average mean molecular mass ($M_w$), and polydispersity ($M_w/M_n$, D)) were determined for each sample (Table 3.1). In Table 3.1 also the cumulative fraction weights for chains with $M_w$ below 8 kDa and above 20 kDa are reported.

**Table 3.1.** Average molecular weight ($M_w$), polydispersity (D), hydrodynamic radius ($R_h$) and fractions weights for fractions with molecular weight < 8 kDa and >20 kDa, obtained for parent heparins and their glycol-split derivatives by SEC-TDA

| Sample | | $M_w$, kDa | D | $R_h$, nm | Fraction weight, % mass | |
|---|---|---|---|---|---|---|
| | | | | | < 8 kDa | > 20 kDa |
| Parent heparins | PMH | 19.1 | 1.3 | 4.2 | 3 | 27 |
| | BLH | 15.8 | 1.4 | 3.6 | 16 | 18 |
| | OMH | 14.8 | 1.6 | 3.5 | 24 | 18 |
| | BMH | 17.8 | 1.5 | 4.0 | 4 | 26 |
| | | | | | | |
| Glycol-split heparins | gs-PMH | 10.8 | 1.5 | 2.7 | 54 | 7 |
| | gs-BLH | 11.4 | 1.4 | 2.8 | 50 | 8 |
| | gs-OMH | 9.9 | 1.6 | 2.1 | 78 | 3 |
| | gs-BMH | 10.5 | 1.6 | 2.5 | 62 | 5 |

- $R_h$ and fraction weight (% mass) were determined using 0.01M $NaNO_3$ as eluent
- $M_w$ and D were measured using 0.1M $NaNO_3$ as eluent in order to avoid aggregation

Upon glycol-splitting the molecular weight of heparins decreased of about 30-40%, while polydispersity (D) slightly increased for the analyzed samples (Table 3.1), along with increase of the content of shorter-chain oligosaccharides (Table 3.1), indicating that hydrolytic side reactions occurred during gs-derivative preparation. Based on the heparin structure proposed by Xiao and co-workers (Fig. 1.4), the internal undersulfated domain should be enriched in gs-units separating two long sulfated regions. In their study 7 kDa and 12 kDa fragments were observed after enzymatic hydrolysis of a heparin with heparinase III. Consequently, even slight hydrolysis of gs-units, which may statistically occur within that region, should lead to the significant decrease of $M_w$.

### III.2. Nuclear magnetic resonance for structural characterization of gs-heparins

As reported in the Introduction, structural characterization of heparins has relied largely on $^{1}H$ and $^{13}C$ NMR analysis. Monodimensional NMR spectroscopy was shown to be a useful tool for determining the extent of glycol splitting (*Casu* et al, 2002*; Casu* et al, 2004). However, due to the significant signal overlapping, monodimensional spectra cannot be used for distinguishing gsG from gsI residues, nor for quantitative analysis. In contrast, 2D NMR spectra, especially in the $^{1}H/^{13}C$ HSQC approach previously applied to quantify typical residues and substituent groups of heparins (*Guerrini* et al, 2005) and LMWHs (*Guerrini* et al., 2007), was expected to provide a more complete characterization of gs-heparins. Its possibilities were studied in both qualitative and quantitative aspects.

#### III.2.1. Qualitative HSQC NMR analysis. Signal assignment

Initially, the gs-derivative of a PMH sample was chosen as model for the signal assignment in the HSQC spectrum. In the 2D NMR spectra cross peaks of typical residues of both unmodified PMH and its gs-derivative are located in characteristic subregions (Fig. 3.1). The cross-peaks of the unmodified PMH and those not affected by glycol-splitting in gs-PMH are at the positions previously assigned by *Guerrini* et al. (2005 and 2007). For both gs-PMH and PMH, the strongest cross peaks in the anomeric region (4.4-5.5/98-107 ppm (*Guerrini* et al., 2005)) are those of the internal "regular" sequences $I_{2S}$-$A_{NS(6S)}$, not affected by glycol-splitting (Fig.3.1). The same can be observed for the "ring" region signals of $I_{2S}$ and $A_{NS/NAc}$ (Fig.3.1).

Anomeric cross-peaks of non-sulfated G (4.5-4.7/104-107 ppm) and I (4.9-5.0/105 ppm) and those associated with these residues ($A_{NS}$-(I), $A_{NS}$-(G)) were shown to disappear upon glycol-splitting (Fig.3.1a,c). Also the H2/C2 cross-peak of G (3.4/76 ppm) may be used as a marker of the reaction completeness (Fig.3.1b,d). New anomeric signals at 4.98/106.9 ppm and at 4.87/106.6 ppm, appeared in the spectrum region typical for heparin uronic acids, were

supposed to be those of gs-residues. The first one is compatible with previous assignments for gsI residues of a gs heparin (*Casu* et al, 2002). Both cross-peaks were shown to correlate with a primary alcohol group at ~ 3.7/65 ppm in the HSQC-TOCSY-DEPT spectrum (Fig. 3.2), confirming their nature.

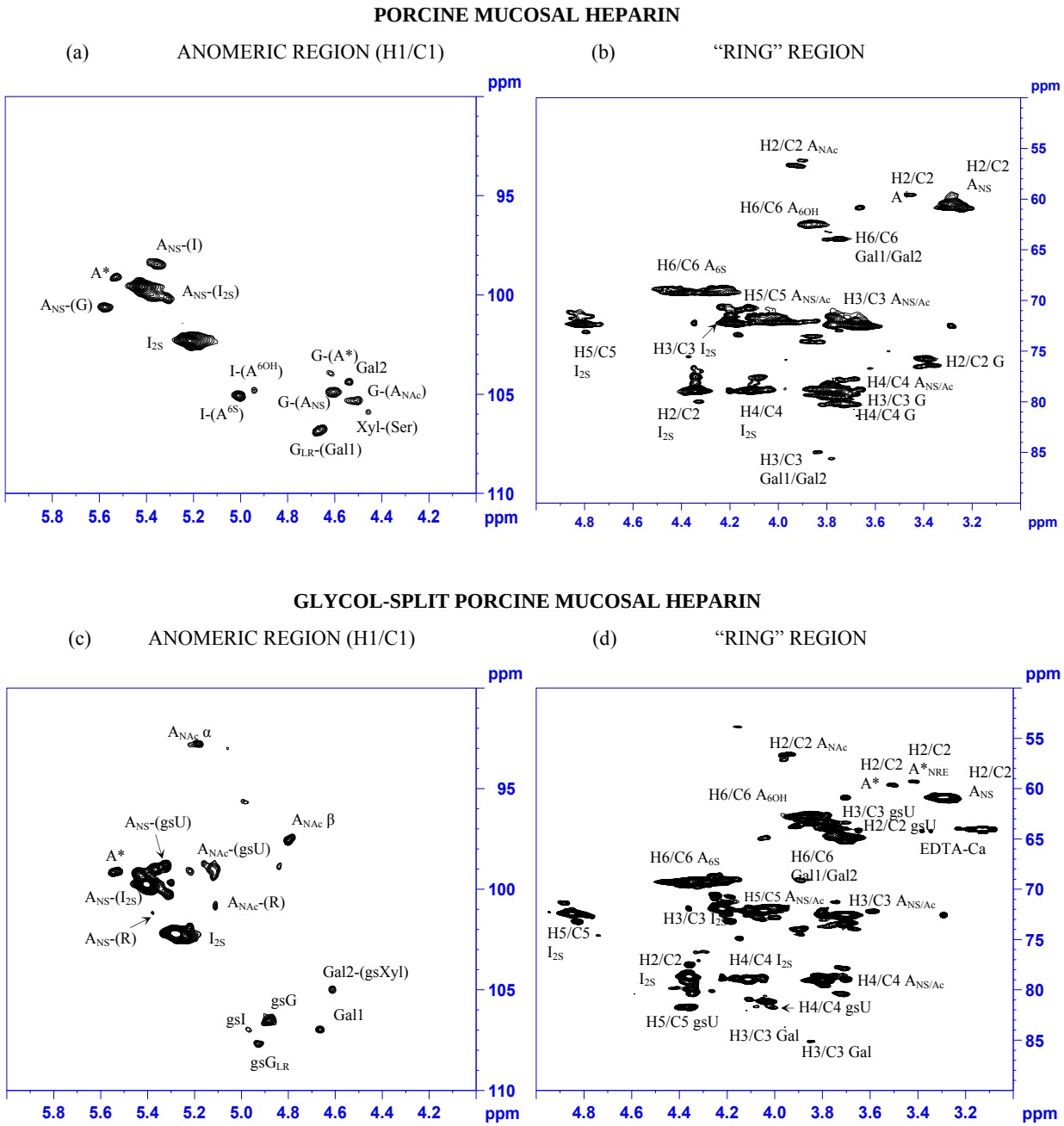



**Fig.3.1.** Two-dimensional HSQC NMR spectra of porcine mucosal heparin (a, c) and the corresponding gs-derivative (b, d)

Together with the monosaccharide residue, the unit following it is indicated in parenthesis, in order to show the sequence effect. A – glucosamine, I – iduronic acid, G – glucuronic acid, gsU – glycol-split uronic acid, Gal1,Gal2, Xyl and Ser – linkage region residues (see Fig.1.19), R – remnant of the hydrolyzed gsU (Fig.1.19)

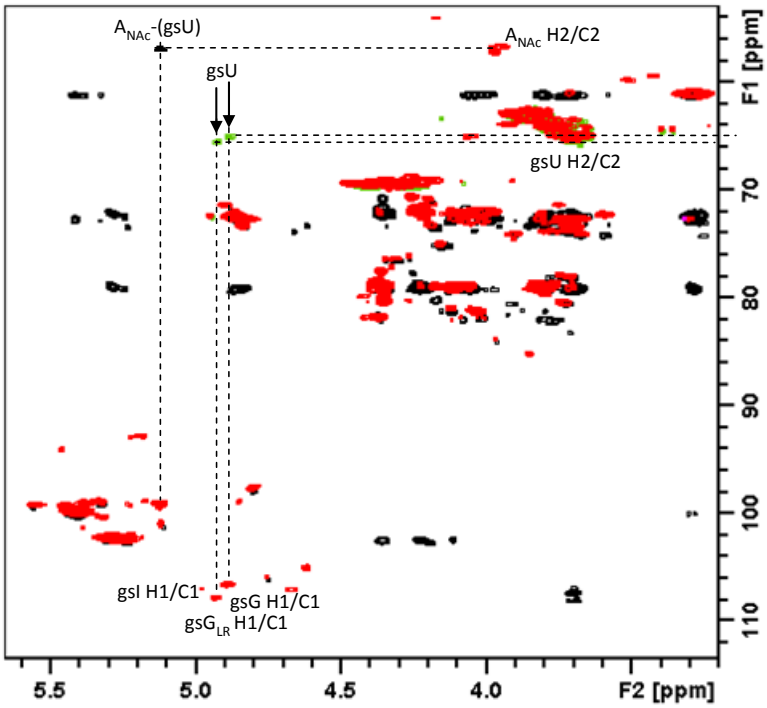



**Fig.3.2**. HSQC-TOCSY-DEPT and HSQC spectra of glycol-split porcine mucosal heparin
The HSQC-TOCSY-DEPT spectrum is shown in black (positive signals) and green (negative signals of $CH_2$ moiety), while the HSQC spectrum is shown in red

Correlation between anomeric signal of gsI and a primary alcohol is not shown in the figure because of its low amount (low signal-to-noise ratio)

HI-3 and gs-derivatives of *N*-acetyl-heparosan (K5) and *N*-Ac-BLH (this latter previously obtained in our laboratory) were used to distinguish between gsI and gsG. In the case of HI-3, a gs heparin derived from partially 2-*O*-desulfated PMH, the expected intensity increase of the signal at 4.98/106.9 ppm (Fig. 3.3a) with respect to the gs-derivative of unmodified PMH (Fig. 3.1c) was observed. Consequently, it has been attributed to gsI residues linked to $A_{NS}$ (this latter being mainly 6-*O*-sulfated).

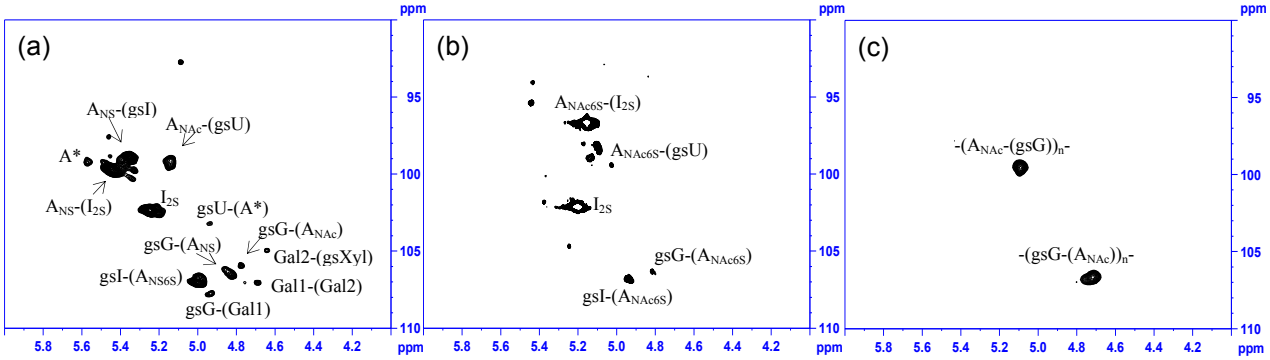



**Fig.3.3.** Two-dimensional HSQC NMR spectra of HI-3 (a), gs-NAc-BLH (b) and gs-K5 (c)

Together with the monosaccharide residue, the residue following it is indicated in parenthesis, in order to show the sequence effect. A – glucosamine, I – iduronic acid, G – glucuronic acid, gsU – glycol-split uronic acid, Gal1, Gal2, Xyl and Ser – linkage region residues (see Fig.1.19).
6-*O*-Sulfation of glucosamine residues is indicated only for gs-NAc-BLH (b) where its presence is known to be very high (>90 %, *Guerrini* et al., 2001)

The spectrum of the gs-derivative of the totally *N*-acetylated BLH (Fig. 3.3b), known to be highly 6-*O*-sulfated and to contain an extremely low quantity of glucuronic acid (*Guerrini* et al., 2001), was also obtained. Accordingly, the cross peak with chemical shift 4.92/106.9 ppm was attributed to gsI residues followed by *N*-acetylated and 6-*O*-sulfated glucosamine ($A_{NAc6S}$). The second minor signal at 4.81/106.2 ppm was assigned to gsG linked to $A_{NAc6S}$. The chemical shifts gsI-($A_{NS6S}$) and gsI-($A_{NAc6S}$) are slightly different from each other due to the so-called sequence effect. It is worth noting that the sequences (I-$A_{NAc}$) are rare in natural heparins (*Yates* et al, 1996) due to the specific biosynthesis pathway, during which epimerization of G to I mainly occurs within the NS-domain (see I.1.2). However, their characterization may be important for the development of the semi-synthetic *N*-acetylated gs-heparins, such as the potential drug SST0001 possessing high anticancer activity (*Naggi* et al, 2005; *Ritchie* et al, 2011; *Shafat* et al, 2011).

As for the gsI residue, also $A_{NS}$ followed by gsI (~5.34/99.0 ppm, Fig.3.3a) is almost at the same position published by Casu (2002), and its intensity increased in the HI-3 spectrum due to the high content of gsI units within NS-domains.

The HSQC spectrum of the gs-derivative of polysaccharide K5, constituted only by the repeating disaccharide units $A_{NAc}$-G, was also informative (Fig. 3.3c). The cross-peak observed at 4.71/106.7 ppm should correspond to the gsG of the repeated gsG-$A_{NAc}$ units. Interestingly, it differs from the corresponding signal found in the spectrum of the gs-NAc-BLH (4.81/106.2 ppm). This may be caused by a sequence effect – due to the presence of 6-*O*-sulfated $A_{NAc}$ residues within gs-NAc-BLH chains, which are not present in the gs-K5.

The second signal observed in the gs-K5 spectrum (5.09/99.5 ppm, Fig.3.3c) corresponds to $A_{NAc}$ of the repeating gsG-$A_{NAc}$ units. Its chemical shift is slightly different in the gs-PMH and HI-3 spectra (5.12/99.0 ppm) due to the sequence effect. In the case of gs-NAc-BLH this cross peak is at 5.12/98.2 ppm probably because of the 6-*O*-sulfation of $A_{NAc}$ (>90% for BLH, *Guerrini* et al., 2001).

Hexa- and octasaccharide fractions of gs-tinzaparin (i.e., the gs derivative of a LMWH rich in *N*-sulfated regions) were used to determine the chemical shift of gsG followed by $A_{NS}$. Among commercially available LMWHs, tinzaparin is considered one of the most similar to the unfractionated heparin (see I.2). Additionally, it is known that $A_{NAc}$ residues are rarely present within its shorter oligosaccharides (*Henriksen* et al., 2004). The absence of the H2/C2 signal at ~3.9/56.0 ppm in the HSQC spectrum (*Guerrini* et al., 2005) confirmed that dp 6 and dp 8 gs-tinzaparin fractions contained only trace amounts of $A_{NAc}$. The chemical shift of the gsG-($A_{NS}$) cross peak varies from 4.90/106.4 ppm (dp 8) to 4.88/106.6 ppm (dp 6), corresponding to the value obtained for the gs-PMH (4.87/106.6 ppm). From the spectrum the value of H1/C1 chemical shift for $A_{NS}$-(gsG) is 5.37/98.9 ppm, similar to the value 5.32/98.9 ppm previously found for $A_{NS}$-(gsI) (*Casu* et al., 2002).

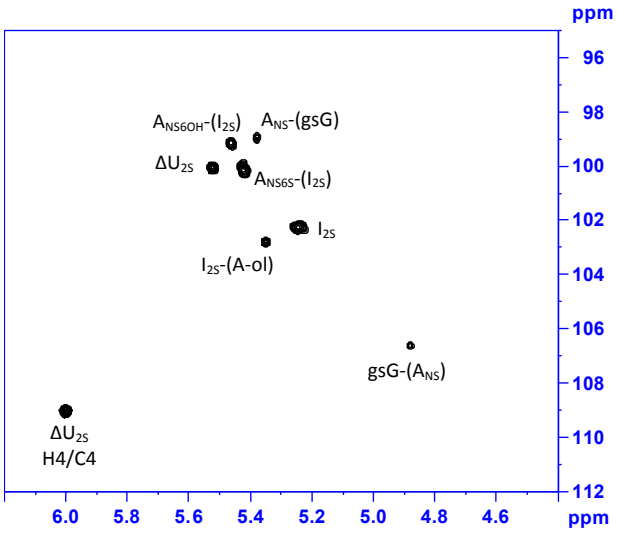



**Fig.3.4.** Two-dimensional HSQC NMR spectrum of the hexasaccharide fraction of gs-tinzaparin

Together with the monosaccharide residue, the residue following it is indicated in parenthesis, to highlight the sequence effect, when observed.
A – glucosamine, I – iduronic acid, G – glucuronic acid, gsU – glycol-split uronic acid, $\Delta U_{2S}$ – 2-*O*-sulfated 4,5-unsaturated uronic acid (see Fig.1.15). A-ol indicates a reduced glucosamine (alditol form) at the RE of oligosaccharide chains (see chapter IV)

Also the chemical shift for the anomeric signal of gsG-(A*), 4.93/103.5 ppm, was previously determined using a gs-derivative of the synthetic pentasaccharide AGA*IA$O_{Me}$ (*Master degree thesis of Monica Ferro*). A signal with similar chemical shift (4.94/103.1 ppm) is detectable in the HI-3 spectrum (Fig. 3.3a). Based on the chemical nature of HI-3, with a high content of gsI residues, the observed cross peak may correspond to gsI-(A*) that is likely to be generated by 2-*O*-desulfation, and further splitting, of sequences initially containing $I_{2S}$ that precedes trisulfated glucosamine A*.

The chemical shifts of the anomeric H1/C1 signals of both gs-uronic acids and glucosamines linked to the gs-units are summarized in Table 3.2.

**Table 3.2.** $^{1}H$ and $^{13}C$ chemical shifts in the anomeric region (H1/C1) characteristic for gs-heparins

| RESIDUE | $^{1}H$, ppm | $^{13}C$, ppm |
|---|---|---|
| **Glycol-split uronic acids** | | |
| -(gsG-($A_{NAc}$))$_n$- [a] | 4.71 | 106.7 |
| gsG-($A_{NAc6S}$) [b] | 4.81 | 106.2 |
| gsG-($A_{NS}$) [c] | 4.88 (dp 6)<br>4.90 (dp 8) | 106.6 (dp 6)<br>106.4 (dp 8) |
| gsG [d] | 4.87 | 106.6 |
| gsI-($A_{NAc6S}$) [b] | 4.92 | 106.9 |
| gsI-($A_{NS}$) [d,e] | 4.98 | 106.9 |
| gsG-(A*) [f] | 4.93 | 103.5 |
| $gsG_{LR}$ [g] | 4.92 | 107.7 |
| gsXyl [g] | 4.75 | 105.9 |
| **Amino sugars** | | |
| -($A_{NAc}$-(gsG))$_n$- [a] | 5.09 | 99.5 |
| $A_{NAc6S}$-(gsI) [b] | 5.12 | 98.2 |
| $A_{NAc}$-(gsU) [d] | 5.12 | 99.0 |
| $A_{NS}$-(gsG) [c] | 5.37 (dp 6)<br>5.38 (dp 8) | 98.9 (dp 6)<br>98.8 (dp 8) |
| $A_{NS}$-(gsI) [e] | 5.32 | 98.9 |

[a] assignment consistent with the 2D NMR spectrum of gs-K5 (gs-derivative of *N*-acetyl-heparosan)
[b] assignment consistent with the 2D NMR spectrum of gs-derivative of *N*-acetylated BLH
[c] determined using hexa- and octasaccharide fractions of gs-tinzaparin
[d] value observed for gs-heparins in the present study
[e] published by Casu and co-workers (2002)
[f] obtained for using gs-derivative of AGA*$IA_{OMe}$; A* = $A_{NS3S(6S)}$
[g] observed in gs-heparins and some gs-LMWHs (*see also Chapter IV, Fig.4.6*)

Together with internal non-sulfated uronic acid residues, also $G_{LR}$ and Xyl of the LR can be split because of the vicinal OH-groups in their structures. A cross-peak at 4.92/107.7 ppm was attributed to the $gsG_{LR}$ because: i) it correlates with the primary alcohol group in the HSCQ-TOCSY spectrum (Fig. 3.2); ii) the ratio of peak volumes $gsG_{LR}$ : Gal1 : Gal2 is about 1:1:1 in each sample where these signals are present; iii) it is absent in gs-heparins with no or

low LR content (K5, BLH); iv) its quantity increases substantially in gs-tinzaparin, known to have a high content of LR sequence preserved during soft depolymerisation process (*Bisio* et al, 2009), and a fraction of the gs-enoxaparin digest enriched in gsLR-bearing oligosaccharides (*see Chapter IV*). This latter sample was used for determining the chemical shift of gs-Xyl (Table 3.2, *see also Chapter IV, Fig.4.6*). The other two cross-peaks clearly seen in the anomeric region of gs-PMH were assigned to the two Gal residues. The long-range correlation of H1/C1 cross-peaks with H3/C3 signals typical for galactose of the heparin LR (*Xiao* et al., 2011) was shown by HMBC spectroscopy (Fig.3.5). Gal residue with the chemical shift (4.67/106.9 ppm) unaffected by glycol-splitting was attributed to Gal1 followed by Gal2 (Fig. 1.19). In contrast, the Gal2 residue shifted after glycol-splitting (4.62/104.9 ppm), probably because it is affected by the adjacent gs-Xyl.

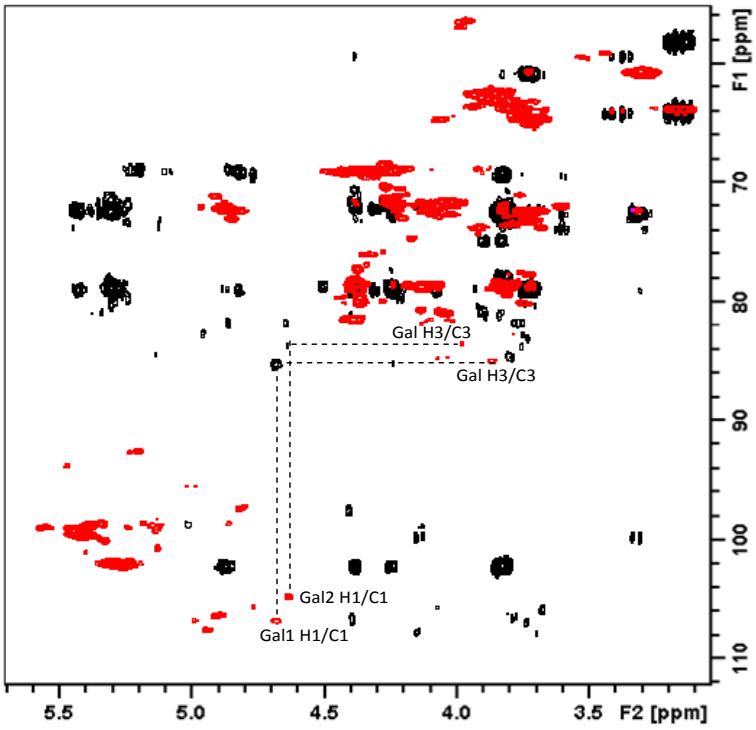



**Fig.3.5.** HSQC (red) and HMBC (black) NMR spectra of glycol-split porcine mucosal heparin (gs-PMH)

Correlation between anomeric cross peaks of two galactose residues of the LR sequence with the corresponding H3/C3 signals is shown

Notably, a cross-peak not related to the TSD regular sequences (H1/C1 5.53/99.1 ppm), attributable to the 3-*O*-sulfated aminosugar $A_{NS3S(6S)}$ (A*) is common to both gs and unmodified PMH. However, in the spectrum of gs-PMH, two H2/C2 cross-peaks, one of which corresponding to internal A* and the other to NRE-terminal A* ($A^*_{NRE}$) residues (*Rudd*

et al, 2013), are observed. The notable increase of $A^*_{NRE}$ may be explained by partial hydrolysis of gs-PMH at the level of gsG/gsI followed by A* (Fig. 1.19e). Indeed, the decrease of the molecular weight after glycol-splitting as determined by SEC-TDA (see III.1) suggested that a hydrolytic side-reaction occured. Interestingly, the *Linhard*'s group reported that the ATBR is the first sequence to be hydrolyzed in a gs-heparin under basic conditions (*Islam* et al., 2002).

Unexpectedly, signals in the "reducing end region" (4.7-5.2/91-98 ppm) are present in the gs-PMH (Fig.3.1). By comparison with the spectra of heparin oligosaccharides (*Guerrini* et al, 2008), they are attributable to anomers α (5.16/92.7 ppm) and β (4.77/97.5) of $A_{NAc}$ residues. Also less intense spots are observable at fields typical for RE residues of $A_{NS}$, i.e., at 5.44/93.9 (α) and 4.72/98.7 (β) ppm. The appearance of RE cross-peaks, which are usually absent in the spectra of UFHs, may indicate that also hydrolysis of some glycosidic bonds occurred (especially, at the level of *N*-acetylated regions).

During gs-heparins preparation the pH was adjusted by adding 0,1 M HCl solution to 4 to destroy $NaBH_4$ in excess, and then increased up to 7 with NaOH to neutralize the acid excess in order to avoid desulfation. The pH adjustment, during which an acidic environment may be locally created, should be one of the most critical steps. To identify marker signals related to these side reactions, we completely hydrolyzed a gs-PMH sample by Smith degradation. SEC-TDA analysis of its hydrolysis products showed a $M_w$ decrease from 10.8 to 5.4 kDa, indicating that the sample was partially depolymerized. As expected, the gsG and gsI cross-peaks and the associated signals ($A_{NS}$-(gsI), $A_{NAc}$-(gsU)) totally disappeared upon hydrolysis (Fig.3.6). The cross-peak at 5.11/100.8 ppm resistant to the hydrolysis was shown to correspond to $A_{NAc}$ followed by R. It correlates with H2/C2 of the $A_{NAc}$ unit in the TOCSY spectrum (Fig. 3.6c). Also long-range correlation between this signal and C(3) of the remnant was shown in the HMBC spectrum (Fig. 3.6b). The same experiments were performed for determining the position of $A_{NS}$-(R) in the HSQC spectrum (5.37/101 ppm, Fig. 3.6).

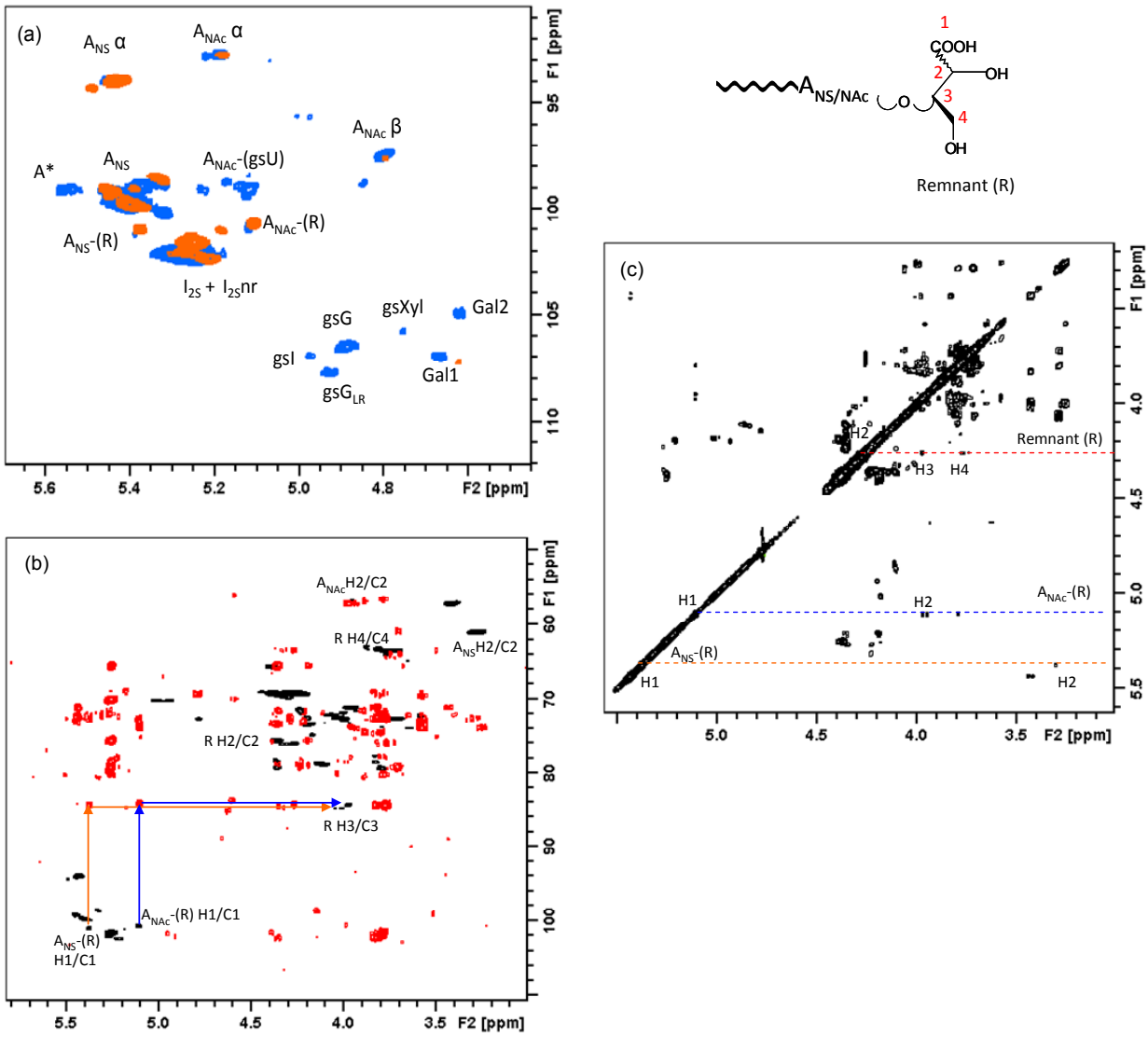



**Fig. 3.6**. NMR study of the hydrolyzed glycol-split porcine mucosal heparin

(a) Superimposed 2D HSQC NMR spectra of gs-PMH before (blue spots) and after Smith degradation (orange spots), (b) HSQC (black spots) and HMBC (red spots) of the hydrolyzed gs-PMH, (c) TOCSY NMR spectrum of the hydrolyzed gs-PMH

A – glucosamine, I – iduronic acid, G – glucuronic acid, gsI/gsG/gsU – glycol-split iduronic/glucuronic/uronic acid, Gal1/Gal2– galactose residues of the linkage region, gsXyl – glycol-split xylose, R – remnant of gsU. Together with the monosaccharide residue, the residue following it is indicated in parenthesis, in order to show the sequence effect.

Together with the glucosamines bearing remnant, also strong RE-signals ($A_{NS}\alpha$, $A_{NAc}\alpha$, $A_{NAc}\beta$) appeared in the anomeric region of the HSQC spectrum (Fig.3.6a), indicating chain shortening. As mentioned in the Introduction (I.3), the hydrolysis of the glycosidic bonds may occur under aggressive acidic conditions (*Jandik* et al., 1996).

III.2.2. Quantitative HSQC NMR analysis. Comparison of different gs-heparins

Though with different relative intensities, the HSQC spectra of gs-BLH, gs-OMH, gs-BMH (Fig.3.7) display cross peaks similar to those of gs-PMH. Except for the highly sulfated and very low acetylated BLH, the gs-heparins are difficult to be distinguished at first glance. The major differences include intensities of cross peaks typical for gsG and gsI as well as their ratio. To better characterize these samples, quantitative $^{1}H/^{13}C$ HSQC NMR was applied as described for unmodified heparins and other GAGs (*Guerrini* et al, 2005 and 2007).

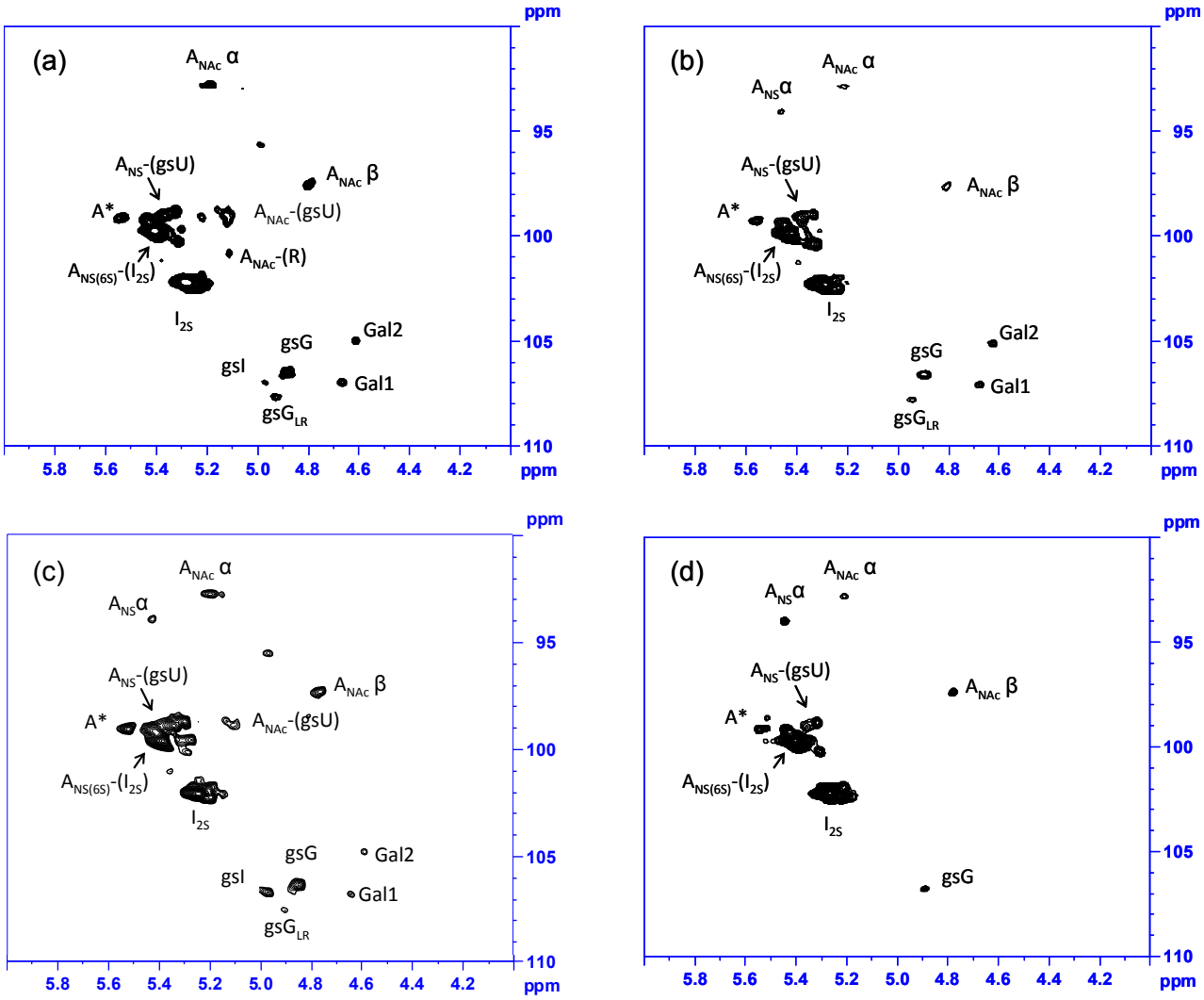



**Fig. 3.7.** Anomeric region of 2D HSQC NMR spectra of gs-heparins of different origin

(a) – gs-PMH, (b) – gs-BMH, (c) – gs-OMH, (d) – gs-BLH

A – glucosamine, I – iduronic acid, G – glucuronic acid, gsI/gsG/gsU – glycol-split iduronic/glucuronic/uronic acid, Gal1/Gal2– galactose residues of the linkage region, gsXyl – glycol-split xylose, R – remnant of gsU. Together with the monosaccharide residue, the residue following it is indicated in parenthesis, in order to show the sequence effect.

Even if the HSQC spectra are currently exploited for quantitative evaluation of individual monosaccharide and disaccharide units in heparins (*Guerrini* et al., 2005) and LMWHs (*Guerrini* et al., 2007), it was necessary to consider that the local flexibility of gs residues may affect their relaxation times and scalar coupling values. Relaxation time $T_2$ values, previously measured for a heparin and a gs-heparin samples (*Casu* et al., 2002), were found to be higher for both unmodified and gs residues in the gs-heparin with respect to the unmodified sample. However, no significant differences were observed for various residues (for example, $I_{2S}$ and gsU) within the same sample. The one bond coupling constants ($^1J_{C\text{-}H}$) for gs-residues gsI and gsG were measured in the present study using a gs-PMH sample and were found equal to 162 Hz, value very close to the values obtained for anomeric signals of unmodified uronic acids (Table 1.3, *Guerrini* et al, 2005). From this point of view, the gs-residue response in HSQC spectra should be the same that for unmodified residues present within heparin chains, indicating that quantitative compositional analysis can be performed by peak volume integration, as currently being applied for heparin and LMWH quantitative analysis (*Guerrini* et al, 2005 and 2007).

Selected values, expressed as percent moles of typical residues, are reported in Table 3.3.

The content of $I_{2S}$ residues, components of the regular TSD sequences (*Casu* B., 2005), is ~90% of total uronic acids for gs-PMH and even higher for gs-BMH (~94%) and gs-BLH (~99%). The $I_{2S}$ content is lower (about 87%) for gs-OMH, and especially for the gs-heparin variant HI-3 (62.6%) derived from a partially 2-*O*-desulfated PMH. The contents of total gs uronic acids (gsI + gsG) were correspondingly in the order: gs-BLH (1.0) < gs-BMH (5.6) < gs-PMH (10.3) < gs-OMH (13.3) < HI-3 (37.4). The obtained values are in accord with the sulfation degree of the starting heparin (for example, the total content of non-sulfated uronic acids followed the order BLH (4.8) < BMH (23.8) ~ PMH (26.7) < OMH (35.5)). However, it is necessary to consider that partial hydrolysis may affect the results. Except for HI-3, the

decrease of gsI was more significant than that of gsG, suggesting that the hydrolysis of the first residue is likely to be facilitated with the respect to its isomer.

**Table 3.3.** Compositional quantitative NMR analysis of parent and glycol-split heparins

| Parent heparins | Uronic acids | | | | Hexosamines | | |
|---|---|---|---|---|---|---|---|
| | $I_{2S}$ | I | G | GalA | $A_{NS}$ | $A_{NAc}$ | $A^{6S}$ |
| | H1/C1 | | | | H2/C2 | | H6/C6 |
| PMH | 73.5 | 9.8 | 16.9 | < LOD [b] | 87.0 | 12.9 | 80.8 |
| BMH | 79.9 | 9.1 | 14.7 | < LOD [b] | 92.4 | 7.7 | 68.5 |
| OMH | 64.5 | 6.7 | 20.2 | 8.6 | 94.7 | 5.3 | 90.3 |
| BLH | 95.2 | 2.7 | 2.1 | < LOD [b] | 99.9 | < LOQ [b] | 92.4 |
| **Glycol-split heparins** | **Uronic acids** | | | | **Hexosamines** | | |
| | $I_{2S}$ | gsI | gsG | | $A_{NS}$ | $A_{NAc}$ | $A^{6S}$ |
| | H1/C1 | | | | H2/C2 | | H6/C6 |
| gs-PMH | 89.5 | 2.0 | 8.3 | | 90.4 | 9.6 | 77.0 |
| gs-BMH | 94.4 | < LOQ [b] | 5.6 | | 94.8 | 5.2 | 69.5 |
| gs-OMH | 86.7 | 4.4 | 8.9 | | 95.0 | 5.0 | 80.0 |
| gs-BLH | 98.9 | < LOQ [b] | 1.0 | | 99.9 | < LOQ [b] | 87.9 |
| Extensively gs-heparin HI-3[a] | 62.6 | 26.7 | 10.7 | | 89.7 | 10.3 | 65.7 |

Data are expressed in moles percent (*Guerrini* et al, 2005); $I_{2S}$ –iduronic acid 2-*O*-sulfate; gsI/gsG – glycol-split iduronic/glucuronic acids (not including gsG of LR, if present); GalA – galacturonic acid; $A_{NS}$ – glucosamine-*N*-sulfate (including the cross-peak of trisulfated glucosamine $A_{NS3S6S}$)); $A_{NAc}$ – *N*-acetylated glucosamine
Uronic acids compositional analysis was performed by integrating peak volumes of the anomeric region; total *N*-sulfation/*N*-acetylation of glucosamine was performed by integrating the signals H2/C2 of the ring, while 6-*O*-sulfation was determined by integration of the corresponding H6/C6 cross-peaks. Galacturonic acid was present in the ovine mucosal heparin contributing to the final content of gsG residues.
a) Obtained from ~40% 2-*O*-desulfated PMH
b) LOD – limit of detection, LOQ – limit of quantification

Though at somewhat lower overall levels, contents found for the $A_{NS}$ residues practically follow the same order observed for the $I_{2S}$ residues. Table 3.3 reports the values for all the *N*-sulfated glucosamines, including 3-*O*-sulfated $A_{NS3S(6S)}$ residues (4.5-6.8 %). Complement to 100% of total glucosamines is accounted for by $A_{NAc}$ residues (from practically zero for gs-BLH up to 9.6% for gs-PMH and ~10% for HI-3). Notably, $A_{NAc}$ contents found in gs-samples are lower than those found for unmodified heparins (Table 3.3). This result is probably

associated with particular hydrolytic side reactions affecting the *N*-acetylated regions in higher degree than the *N*-sulfated sequences. Such an observation is in agreement with the appearance of $A_{NAc}\alpha/\beta$ anomeric cross-peaks upon glycol-splitting. Islam and co-workers observed the disappearance of *N*-acetylated residues as a result of the work-up (acidic treatment) after periodate oxidation, used in order to obtain LMW products (2002).

### III.3. Determination of hydrodynamic properties of glycol-split heparins by SEC-TDA

Mark-Houwink parameters *a* and *logK* were also obtained for all the samples. Hereafter, the trends observed for the parameter *a* are discussed, while complete data, containing *logK* values, are shown in Annex 3. The analyses were repeated using 0.01 M $NaNO_3$ solution as eluent, because under the experimental conditions used for molecular weight measurements (eluent 0.1 M $NaNO_3$) the effect of the gs residues within the heparin chains is suppressed due to the high ion strength. Several trends can be observed from the obtained data. First of all, the Mark-Houwink plot is not linear and 3-4 distinct linear ranges may be observed for each heparin/gs-heparin sample, suggesting the conformational diversity between oligosaccharides of different length. An example for PMH is shown in Fig. 3.8. It is worth noting that internal slices (13-16 and 16-21 kDa in Fig.3.8) of the molecular weight distribution may be influenced by two well-defined adjacent linear regions (9-13 and 21-36 kDa in Fig. 3.8). We characterized these regions separately because of their close proximity to $M_n$ and $M_w$ values (Fig. 3.8).

The higher the molecular weight, the lower the *a* value. Accordingly, they are in the range of typical flexible polymers in random coil form for higher masses, while appeared very high for shorter oligosaccharides, indicating more rigid structures. Different chain length and sulfation degree as well as different flexibility of iduronic and glucuronic acids may contribute to the conformational equilibrium. It is likely that sulfation degree may provide a significant impact. The longer chain oligosaccharides statistically contain longer undersulfated domains

and random coil formation is favored, while shorter chains likely contain highly sulfated negatively-charged sequences that may impair their flexibility.

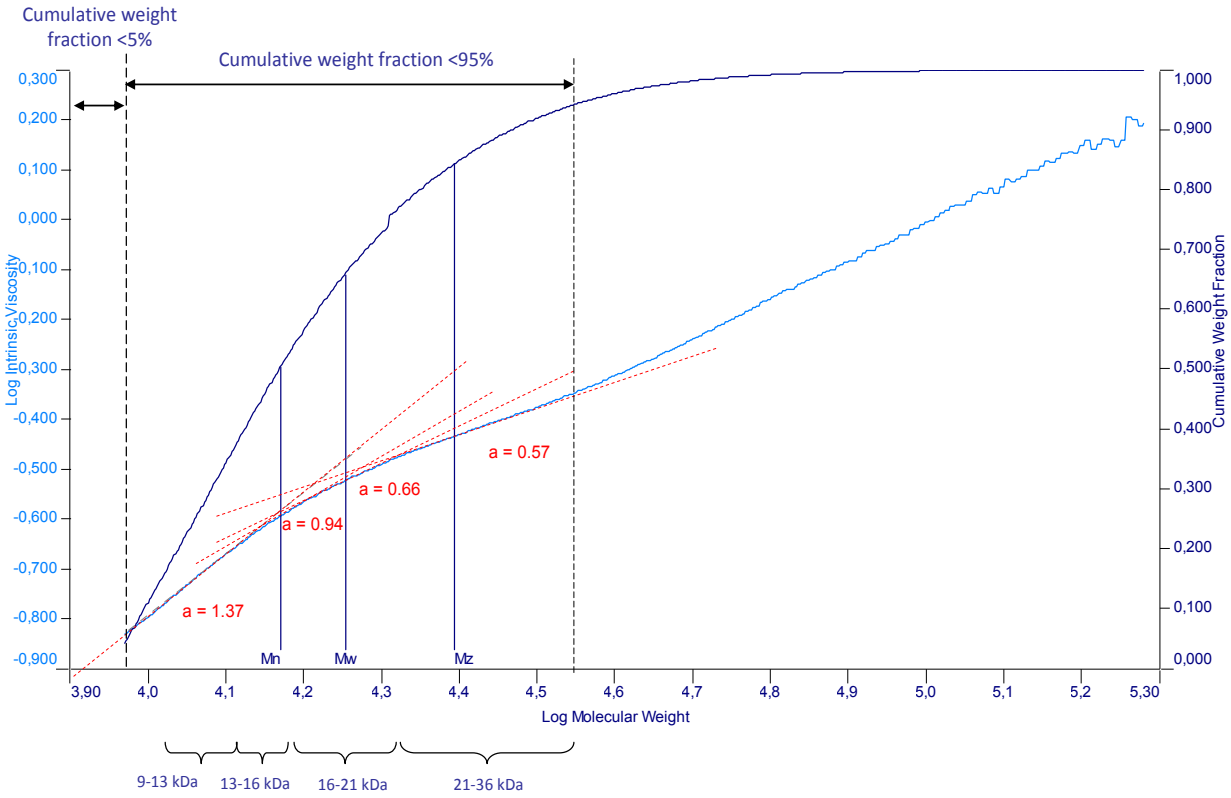



**Fig.3.8.** Mark-Houwink plot for PMH obtained by SEC-TDA using 0,01 M $NaNO_3$ as eluent
Parameter *a* was determined for each linear region (kDa), except for the "critical" zones where sample concentration is low (under 5%)

Interestingly, the number of regions characterized by distinct *a* values was found different for different parent heparins (Fig. 3.9). In Table 3.4 the fraction weight corresponding to each defined molecular weight range and the determined *a* value are shown. Another evident trend is that the *a* value decreases for gs-samples, suggesting the glycol-splitting-induced higher flexibility (Fig. 3.10, Table 3.4). The flexibility of BLH chains containing very low amount of non-sulfated I and G susceptible to periodate, was less affected by glycol-splitting (Fig. 3.10, Table 3.4).

**Table 3.4.** Mark-Houwink parameter *a* obtained for heparins of different animal sources and their glycol-split derivatives

| | Sample | NMR (mol.%) | | | Range, kDa | % | a | Range, kDa | % | a | Range, kDa | % | a | Range, kDa | % | a |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $I_{2S}$ | $A_{NS}$ | $A_{6S}$ | | | | | | | | | | | | |
| Heparins | PMH | 73.5 | 87.0 | 80.8 | 9 – 13 | 28 | 1.37 | 13 – 16 | 21 | 0.94 | 16 – 21 | 17 | 0.66 | 21 – 36 | 17 | 0.57 |
| | BLH | 95.2 | 99.9 | 92.4 | 8 – 10 | 24 | 1.32 | 10 – 12 | 15 | 0.99 | 12 – 23 | 28 | 0.74 | 23 – 37 | 10 | 0.63 |
| | OMH | 64.5 | 94.7 | 90.3 | 6 – 7 | 12 | 1.41 | 7 – 14 | 50 | 0.99 | 14 – 32 | 31 | 0.66 | > 32 | nd | nd |
| | BMH | 79.9 | 92.4 | 68.5 | < 8 | nd | nd | 8 – 17 | 60 | 1.08 | 17 – 22 | 13 | 0.74 | 22 – 35 | 14 | 0.63 |
| | | | | | | | | | | | | | | | | |
| Gs-heparins | gs-PMH | 89.5 | 90.4 | 77.0 | 3.5 – 4.5 | 15 | 1.44 | 4.5 – 8.5 | 38 | 0.90 | 8.5 – 22 | 36 | 0.53 | > 22 | nd | nd |
| | gs-BLH | 98.9 | 99.9 | 87.9 | < 4.5 | nd | nd | 4.5 – 8 | 43 | 1.20 | 8.0 – 23 | 44 | 0.55 | > 23 | nd | nd |
| | gs-OMH | 86.7 | 95.0 | 80.0 | 2.5 – 3 | 20 | 1.33 | 3 – 6.5 | 42 | 0.71 | 6.5 – 17 | 29 | 0.55 | > 17 | nd | nd |
| | gs-BMH | 94.4 | 94.8 | 69.5 | < 3 | nd | nd | 3 – 7.5 | 51 | 0.97 | 7.5 – 21 | 38 | 0.56 | > 21 | nd | nd |

% - the fraction weight in mass.%, for each selected molecular weight range. The columns, better showing the trends discussed in III.3, are in yellow.

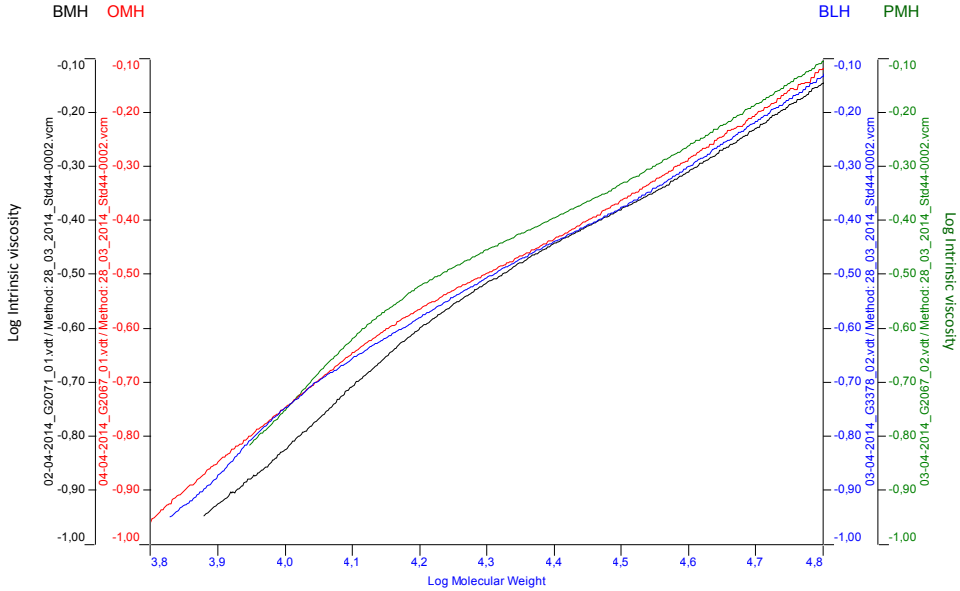



**Fig. 3.9** Mark-Houwink plots obtained for heparins of different sources
BMH – black, OMH – red, BLH – blue, PMH – green

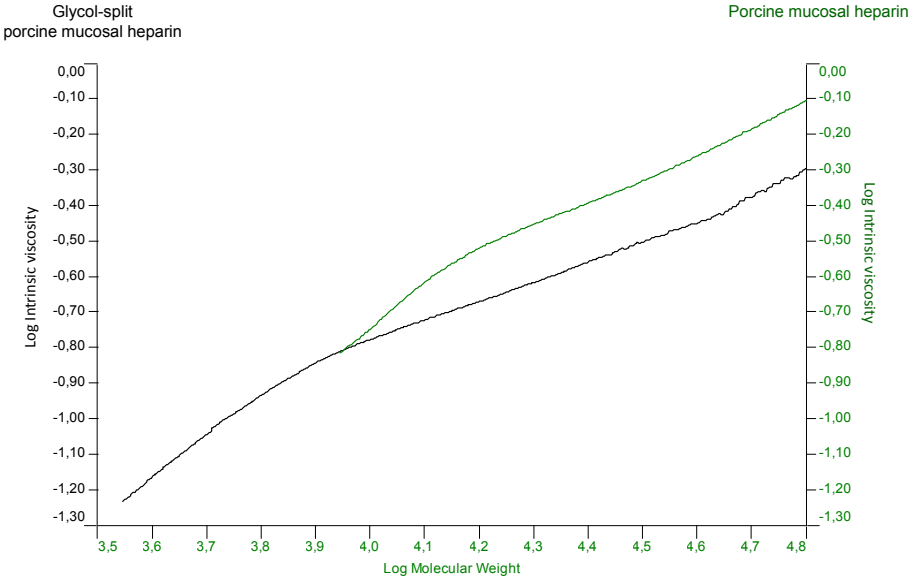



**Fig.3.10.** Mark-Houwink plots obtained for PMH (green) and its gs-derivative (black)

### III.4. LC-MS analysis for structural characterization of gs-heparins and their starting materials

As mentioned in the Introduction, to better study heparin sequences and unravel minor modifications of such a microheterogeneous polysaccharide, enzymatic depolymerisation followed by LC-MS analysis of the obtained mixture is often necessary. Among the other chromatographic techniques IPRP-HPLC has important advantages, first of all, it is easy to couple with MS, permitting the separation and identification of oligosaccharides in complex heparin-derived mixtures. This application demands the use of a mass spectrometer of high resolution and accuracy in mass determination. In the present work ESI-Q-TOF MS was chosen. The potential of the developed LC-MS method for qualitative and quantitative analysis of different gs-heparins has been studied. The obtained results were compared with NMR data.

A previous exhaustive digestion with heparin lyases (heparinases) was performed in order to simplify the analyzed mixture. Given the high specificity of heparinases that preferentially cleave glycosidic bonds of unmodified heparin sequences (*see I.1.4.2*), these enzymes were expected to be unable to cleave bonds between glucosamine and gs-uronic acid residues. Exhaustive digests of gs-heparins with heparin lyases I, II, and III are accordingly expected to consist prevalently of disaccharides generated from unmodified heparin chains, as well as of several small oligosaccharides containing gsG/gsI residues (Fig.3.11). The combined approach of analyzing both non-glycol-split and glycol-split heparins should provide more detailed structural information on sequences of unmodified heparin (for example, on the number of subsequent non-sulfated uronic acids) with respect to the traditional method based on analysis of enzymatic digests of unmodified heparins.

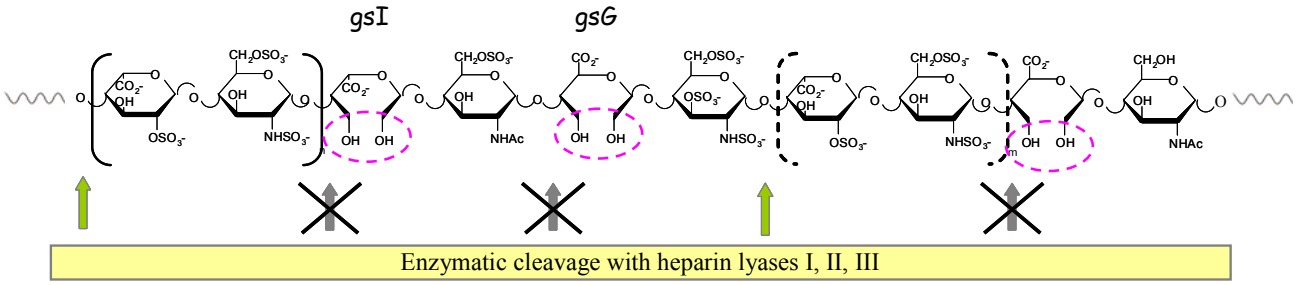



**Fig. 3.11.** Generation of gs-oligosaccharides by exhaustive heparinases-digestion of gs-heparins

Green arrows indicate glycoside bonds that can be cleaved by heparinases; grey arrows show glycosidic bonds supposed to be resistant to enzymes because of the glycol-splitting induced modifications

A specific nomenclature reflecting the first-level structure obtained by LC-MS analysis of heparin oligosaccharides was used. The **abbreviation system** is similar to that of *Henriksen* and co-workers (2004) and indicates the number of monosaccharide residues, sulfate groups, and *N*-acetyl groups. In our case, a number and the symbol "gs" were also added to indicate the number of the gs-residues. "Remnants" of the hydrolyzed internal gs residues were indicated by the symbol R (see Fig. 1.19e). Symbol ΔU and A indicate a 4,5-unsaturated uronic acid and a glucosamine-residue, respectively, at the NRE.

III.4.1. Optimization of IPRP-HPLC/ESI-Q-TOF method

For optimizing separation and detection parameters, gs-PMH was chosen as a model sample. Firstly, ESI-Q-TOF parameters (capillary voltage and transfer line voltages) were optimized to avoid desulfation and significant adduct formation (first of all, with an ion-pair amine present in high concentration in the eluent). Both positive and negative ion mode were tested. The positive one was more sensitive, however, led to the formation of numerous adducts with dibutylamine (DBA) used as ion-pair agent. For this reason, negative ion mode, providing less complex spectra, was chosen. Values slightly lower than the typical ESI voltages were chosen as a compromise (*see Experimental*). A more crucial decrease of these voltages leads to the sensitivity "fall", whereas higher values may cause significant desulfation.

As expected, enzymatic digests of gs-heparins are mainly constituted by disaccharides and by a number of tetra- and higher saccharides containing undigested gs-residues (Fig.3.12).

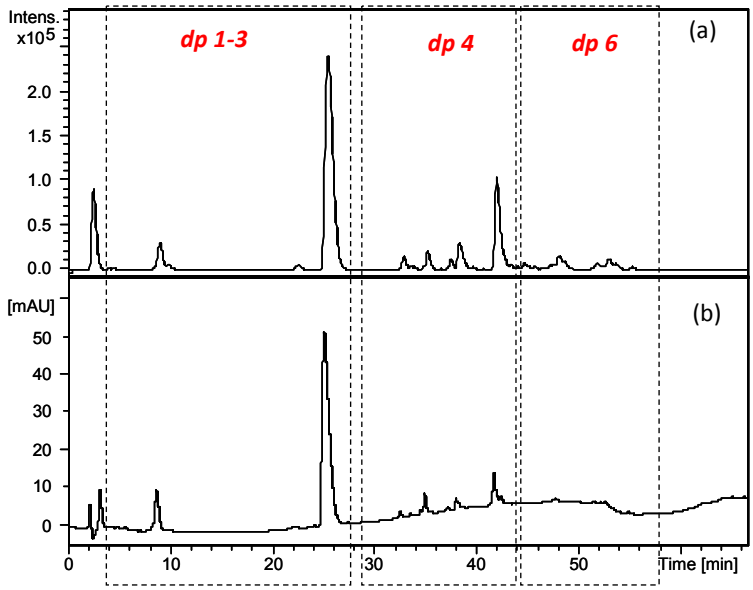



**Fig. 3.12.** LC-MS (a) and LC-UV (b) chromatograms of heparinase-digest of glycol-split porcine mucosal heparin under the optimized conditions

HPLC-UV Conditions: Column C18 100x2.1 mm, 2.6 µm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in water-methanol = 9:1 (v/v) and methanol, respectively; multi-step gradient: 0 min – 17%B, 10 min – 17%B, 30 min – 42%, 50 min – 50%B, 65 min – 90%B, 75 min – 90%B, 76 min – 10%B, 90 min – 10%B; flow-rate – 100 µl/min; UV detection – 232 nm
ESI-Q-TOF-MS conditions: negative ion mode, capillary voltage – 3.2 kV, nebulizer gas – 0.9 bar, drying gas – 7.0 l/min, 180°C, *m/z* range 200-2000

Optimization of chromatographic conditions was focused on the simultaneous separation of disaccharides and other peculiar oligosaccharides. Different aliphatic amines, among which dibutyl-, *n*-pentyl and *n*-hexylamine, have been tested in their acetate salt form, compatible with ESI-MS, as ion pair agents. The other amines that were historically used in IPRP mode are tri- and tetra-butylamines. However, these molecules are strongly adsorbed on the C18 stationary phase, increasing the total analysis time. To compensate this effect, higher concentration of the organic solvent should be used, that, unfortunately, leads to worse resolution and signal suppression (*Langeslay* et al, 2013). DBA was chosen because it has been successfully used for the analysis of HS oligosaccharides (*Wu and Lech*, 2005; *Kuberan* et al., 2002) and for pentasaccharidic and octasaccharidic heparanase substrates and their heparanase-cleaved products (*Bisio* et al., 2007). *n*-Pentyl- and *n*-hexylamines have been chosen to be compared with DBA, since their effectiveness in the IPRP separations of some oligosaccharides has been recently demonstrated (*Doneau* et al., 2009, *Langeslay* et al, 2013).

A linear gradient of methanol was applied to compare the resolution provided by different amines (their concentration was kept constant, 10 mM). However, only the DBA acetate salt provided both the simultaneous separation of di- and oligosaccharides and a good resolution of positional isomers including gs species. Probably, it is more "selective" for the interaction with sulfate groups of different positional isomers than its mono-alkyl-analogs due to higher stereochemical hindrance. It should be noted that our results do not contradict the work of *Doneau* and co-workers (2009), where mono-alkyl-amines provided better results. The authors were focused on the sensitivity improvement of the MS detection for oligosaccharides with high polymerization degree (up to dp 22), while in our study we put special emphasis on the separation of short oligosaccharides, including isomeric components. For this reason, DBA was chosen for a further optimization of separation conditions. The dependence of the resolution of gs tetra- and hexasaccharides generated from gs-PMH by enzymatic cleavage, the DBA concentration in eluents has been investigated using a linear gradient changing from 10% B to 90% B in 90 min. It has been shown that the change of the ion pair agent concentration from 5 to 10 mM was crucial for the chromatographic resolution of gs oligosaccharides, while further increase of its concentration up to 15 mM did not influence significantly the chromatographic separation (Fig.3.13) but led to an essential decrease (~ 2 times) of MS detection sensitivity and to an increase of the total analysis time. In Fig.3.13 the resolution coefficients *Rs* as function of DBA concentration are shown for pairs of oligosaccharides with similar structure and, consequently, similar chromatographic behavior (in other words, difficult to be separated). The mass spectra interpretation and compositional analysis of gs-heparin digests with all the identified oligosaccharides are discussed below. For the species shown in Fig.3.13 see the abbreviation system in the footnote.

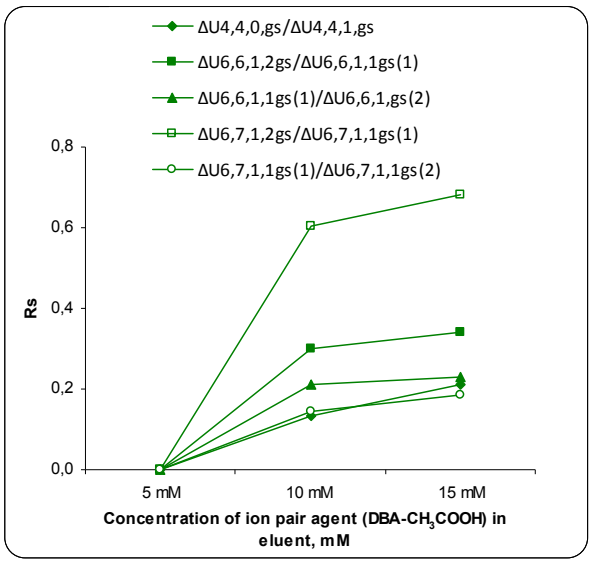



**Fig.3.13.** Dependence of the resolution coefficient (Rs) between target pairs of analytes on DBA concentration in eluent

Conditions: Column C18 100x2.1 mm, 2.6 μm; eluents A and B – x mM DBA and x mM $CH_3COOH$ (x = 5, 10, 15 mM) in water-methanol = 9:1 (v/v) and methanol, respectively; linear gradient: 0 min – 10%B, 90 min – 90%B at 100 μl/min; UV detection – 232 nm

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE.

Having found the optimal concentration of DBA (10 mM), the gradient has been adapted to achieve better resolution for the simultaneous separation of di-, tetra- and hexasaccharides (Fig.3.12). The resolution coefficients (*R*s) for six oligosaccharide pairs, components of heparinase digests of gs-PMH, were found to be ***1.67*** for ΔU4,4,0,1gs/ΔU4,4,1,1gs, ***1.30*** for ΔU4,6,0,1gs/ΔU6,6,1,1gs, ***1.33*** for ΔU6,6,1,2gs/ΔU6,6,1,1gs(1), ***1.21*** for ΔU6,6,1,1gs(1)/ΔU6,6,1,1gs(2), ***1.55*** for ΔU6,7,1,2gs/ΔU6,7,1,1gs(1) and ***1.33*** for ΔU6,7,1,1gs(1)/ΔU6,7,1,1gs(2). This resolution is sufficient for fingerprinting the enzymatic digests obtained from gs derivatives of different heparin samples through their LC-MS profiles.

Inter-day variations have been also evaluated for the twelve target compounds (di-, tetra-, and hexasaccharides) eluted from the chromatographic column with different retention times in the range from 8 up to about 60 min. As shown in the Table 3.5, the value of day-to-day variability (RSD) does not exceed 1.1 % for each target compound.

**Table 3.5.** Retention times (RT) and inter-day variability (RSD) for the target analytes

| Analyte | RT, min (inter-day measurements, n=3, P=0.95) | |
|---|---|---|
| | x ± Δx | |
| ΔU2,2,0 | 9.0 ± 0.2 | 1.1 |
| ΔU2,3,0 | 26.0 ± 0.6 | 1.0 |
| ΔU4,3,1,1gs (1) | 33.2 ± 0.7 | 0.8 |
| ΔU2,3,0-R | 35.6 ± 0.7 | 0.8 |
| ΔU4,4,0,1gs | 37.8 ± 0.7 | 0.7 |
| ΔU4,4,1,1gs | 38.8 ± 0.8 | 0.8 |
| ΔU4,5,0,1gs | 42.5 ± 0.9 | 0.8 |
| ΔU6,6,1,2gs | 47.4 ± 1.0 | 0.9 |
| ΔU4,6,0,1gs | 47.8 ± 0.9 | 0.8 |
| ΔU6,7,1,2gs | 52.7 ± 1.0 | 0.8 |
| ΔU6,7,1,1gs (1) | 53.7 ± 1.0 | 0.8 |
| ΔU6,7,1,1gs (2) | 54.8 ± 0.9 | 0.7 |

The abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues, when present. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol R is used to indicate the presence of a remnant of gsG/gsI unit at the RE

III.4.2. Mass spectra interpretation and identification of oligosaccharides

By exact mass, each analytical signal has been identified as "first-level" structure, including the number of monosaccharide residues, sulfates, acetyl groups, and gs-units when present (Table 3.6). Three examples of negative ESI mass spectra, characteristic for sulfated oligosaccharides by loss of one or more sulfate groups and formation of dibutylamine adducts, are shown in Fig. 3.14 for ΔU2,3,0 (a), ΔU4,5,0,1gs (b), ΔU6,7,1,2gs and ΔU6,7,0,2gs (c).

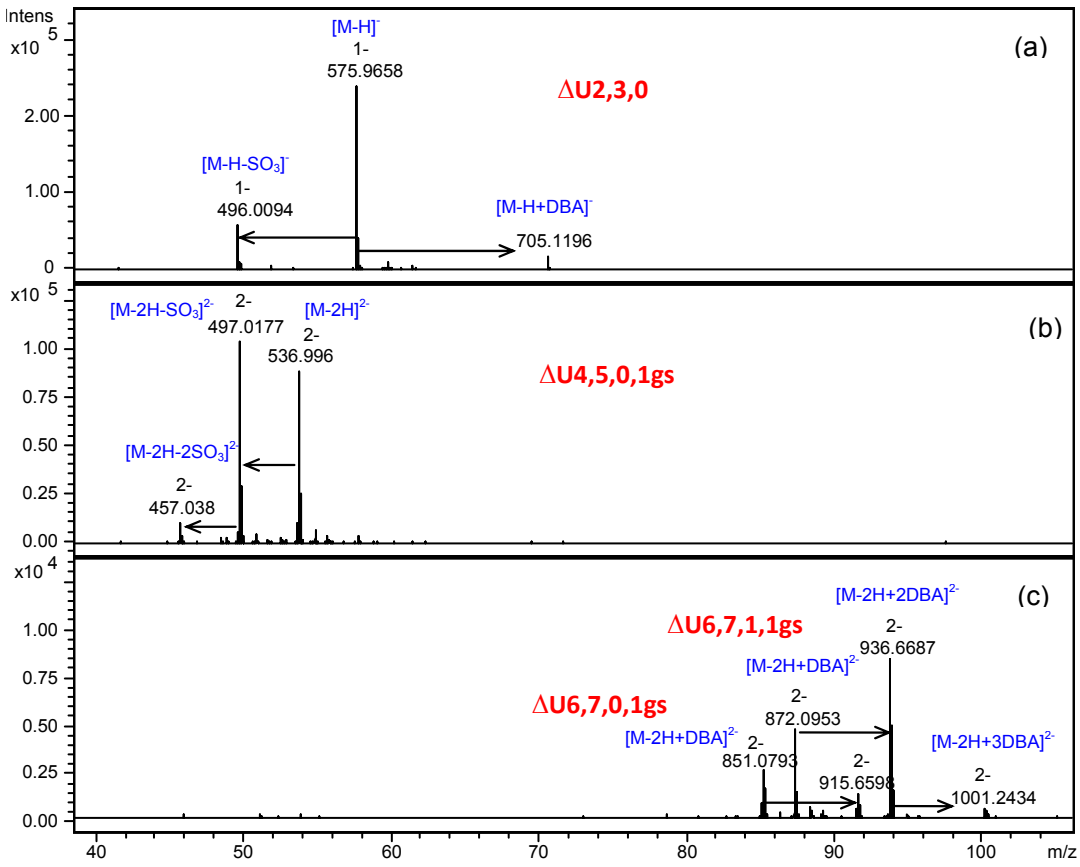



**Fig. 3.14.** Mass spectra of the major disaccharide ΔU2,3,0 (a), gs-tetrasaccharide ΔU4,5,0,1gs (b), and gs-hexasaccharides ΔU6,7,1,2gs and ΔU6,7,0,2gs (c) observed by IPRP-HPLC/ESI-Q-TOF

HPLC-UV/ESI-Q-TOF-MS conditions: see Fig.3.12
The abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE.

Retention times, molecular formulae, monoisotopic exact masses and corresponding first-level structures of the major components of heparinase digests of gs-PMH are reported in Table 3.6. Some components with the same "monoisotopic exact mass" values, corresponding to different isomers, are present in the heparinase-digests of gs-heparins.

Their formula was only accepted if the *m/z* of the candidate analyte matched the theoretical value within 5 ppm. The presence of gs-residues can be detected by the increase of the mass values by two daltons. 4,5-unsaturation of the uronic acids (typical for the heparinase-generated oligosaccharides, see Fig.1.11), detectable by the loss of 18 Da (corresponding to one water molecule), causes the absorption of the UV-light at 232 nm. This property was used as a control during the structure assignment.

**Table 3.6.** Major fragments in heparinase-digests of gs-PMH found by LC-MS

| | **Monoglucosamines** | | **Disaccharides** | | | **Trisaccharides** | | |
|---|---|---|---|---|---|---|---|---|
| **First-level structure** | **A1,2,0** | **A1,3,0 (A*)** | **ΔU2,2,0** | **ΔU2,3,0** | **ΔU2,3,0-R** | **A3,5,0,1gs** | **ΔU3,4,0-ol** | **ΔU3,5,0-$R_{am}$** |
| **Formula** | $C_6H_{13}N_1O_{11}S_2$ | $C_6H_{13}N_1O_{14}S_3$ | $C_{12}H_{19}N_1O_{16}S_2$ | $C_{12}H_{19}N_1O_{19}S_3$ | $C_{16}H_{25}N_1O_{23}S_3$ | $C_{18}H_{34}N_2O_{30}S_5$ | $C_{18}H_{29}N_1O_{28}S_4$ | $C_{22}H_{35}N_1O_{34}S_5$ |
| **Exact mass** | 338.9930 | 418.9498 | 497.0145 | 576.9713 | 694.9979 | 917.9800 | 834.9759 | 1016.9644 |
| **Retention time, min** | 4.4 | 9.9 | 8.9 | 25.9 | 35.5 | 38.8 | 40.8 | 43.8 |
| | **Glycol-split tetrasaccharides** | | | | | | | |
| **First-level structure** | **ΔU4,3,1,1gs** | **ΔU4,4,1,1gs** | **ΔU4,4,0,1gs** | **ΔU4,5,0,1gs** | **ΔU4,5,0,1gs-R** | **ΔU4,6,0,1gs** | | |
| **Formula** | $C_{26}H_{42}N_2O_{30}S_3$ | $C_{26}H_{42}N_2O_{33}S_4$ | $C_{24}H_{40}N_2O_{32}S_4$ | $C_{24}H_{40}N_2O_{35}S_5$ | $C_{28}H_{46}N_2O_{39}S_5$ | $C_{24}H_{40}N_2O_{38}S_6$ | | |
| **Exact mass** | 958.0985 | 1038.0553 | 996.0447 | 1076.0015 | 1194.0281 | 1155.9583 | | |
| **Retention time, min** | (1) 33.0<br>(2) 34.0 | 38.6 | 37.7 | 42.3 | 46.0 | 47.8 | | |
| | **Glycol-split hexasaccharides** | | | | | | | |
| **First-level structure** | **ΔU6,5,1,1gs** | **ΔU6,6,1,1gs** | **ΔU6,6,1,2gs** | **ΔU6,7,0,2gs** | **ΔU6,7,1,1gs** | **ΔU6,7,1,2gs** | **ΔU6,8,0,1gs** | **ΔU6,8,0,2gs** |
| **Formula** | $C_{38}H_{61}N_3O_{46}S_5$ | $C_{38}H_{61}N_3O_{49}S_6$ | $C_{38}H_{63}N_3O_{49}S_6$ | $C_{36}H_{61}N_3O_{51}S_7$ | $C_{38}H_{61}N_3O_{52}S_7$ | $C_{38}H_{63}N_3O_{52}S_7$ | $C_{36}H_{59}N_3O_{54}S_8$ | $C_{36}H_{61}N_3O_{54}S_8$ |
| **Exact mass** | 1455.1130 | 1535.0698 | 1537.0854 | 1575.0317 | 1615.0266 | 1617.0423 | 1652.9729 | 1654.9885 |
| **Retention time, min** | (1) 43.8<br>(2) 44.5<br>(3) 45.3 | (1) 48.4<br>(2) 49.1<br>(3) 49.9 | 47.4 | (1) 47.9<br>(2) 51.8 | (1) 53.6<br>(2) 54.8 | 52.4 | (1) 59.3<br>(2) 60.3 | 57.9 |

* Abbreviation system for the "first-level structures" includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol R is used to indicate the presence of a remnant of gsG/gsI residue at the RE. $R_{am}$ indicates a remnant formed by splitting and further hydrolysis of RE glucosamine residues (described in Chapter IV.4)

The comparison of the experimentally obtained **isotope pattern** with the theoretical one was also used to confirm the assigned structure. This is possible due to the specific sulfur isotope abundance in nature. It exists mainly in two forms, $^{32}S$ (95.02%) and $^{34}S$ (4.21%). An example is shown in Fig.3.15. Oligosaccharides ΔU4,6,0,1gs and ΔU4,4,1,1gs-R, having the same nominal molecular mass (1156 Da) can be only distinguished by the accurate determination of monoisotopic *m/z* value and comparison of the experimental isotope pattern with the theoretical one.

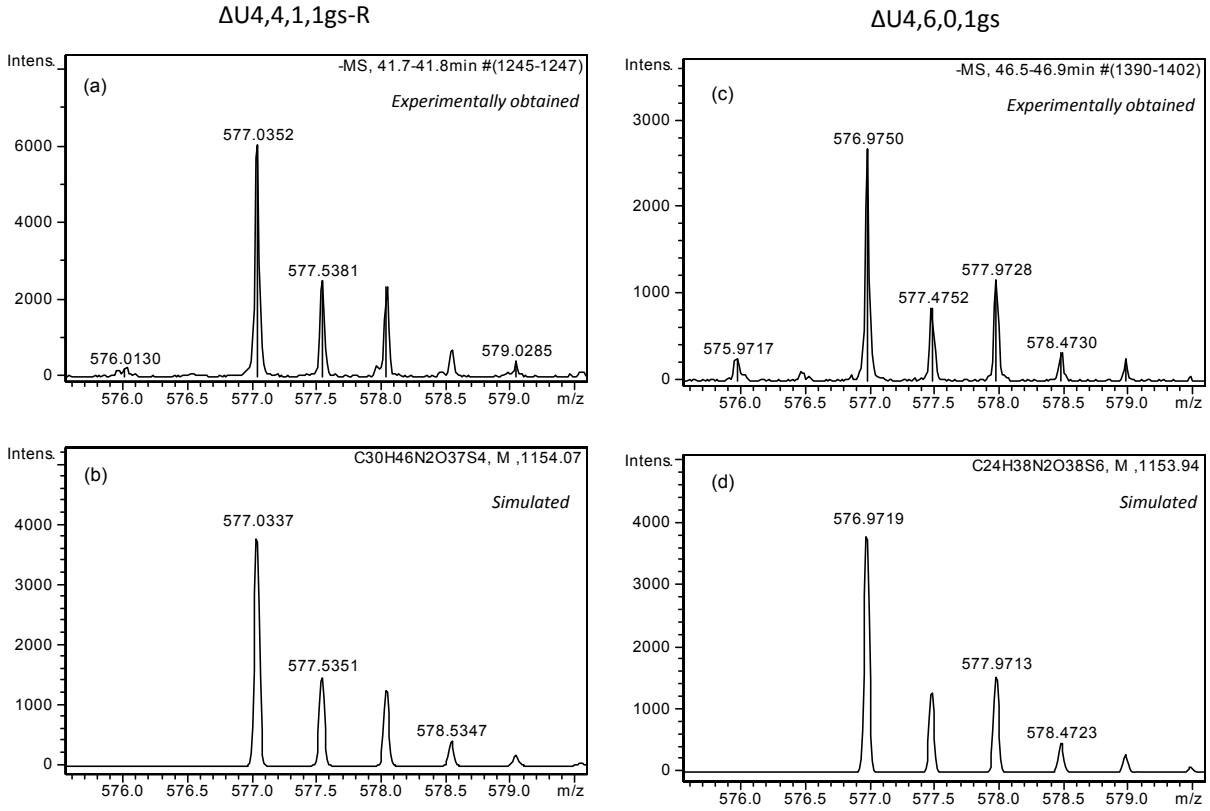



**Fig.3.15.** Experimental and simulated isotope patterns of ΔU4,4,1,1gs-R (a,b) and ΔU4,6,0,1gs (c,d)

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol R is used to indicate the presence of a remnant of gsG/gsI residue at the RE

II.4.3. Comparison of heparinase-digests of porcine mucosal heparin and its glycol-split derivative

After having optimized the separation conditions and identified each analytical signal, the LC-MS profile of the gs-PMH heparinase-digest was compared with the corresponding digest of the unmodified PMH (Fig.3.16) in order to better characterize the gs-sample.

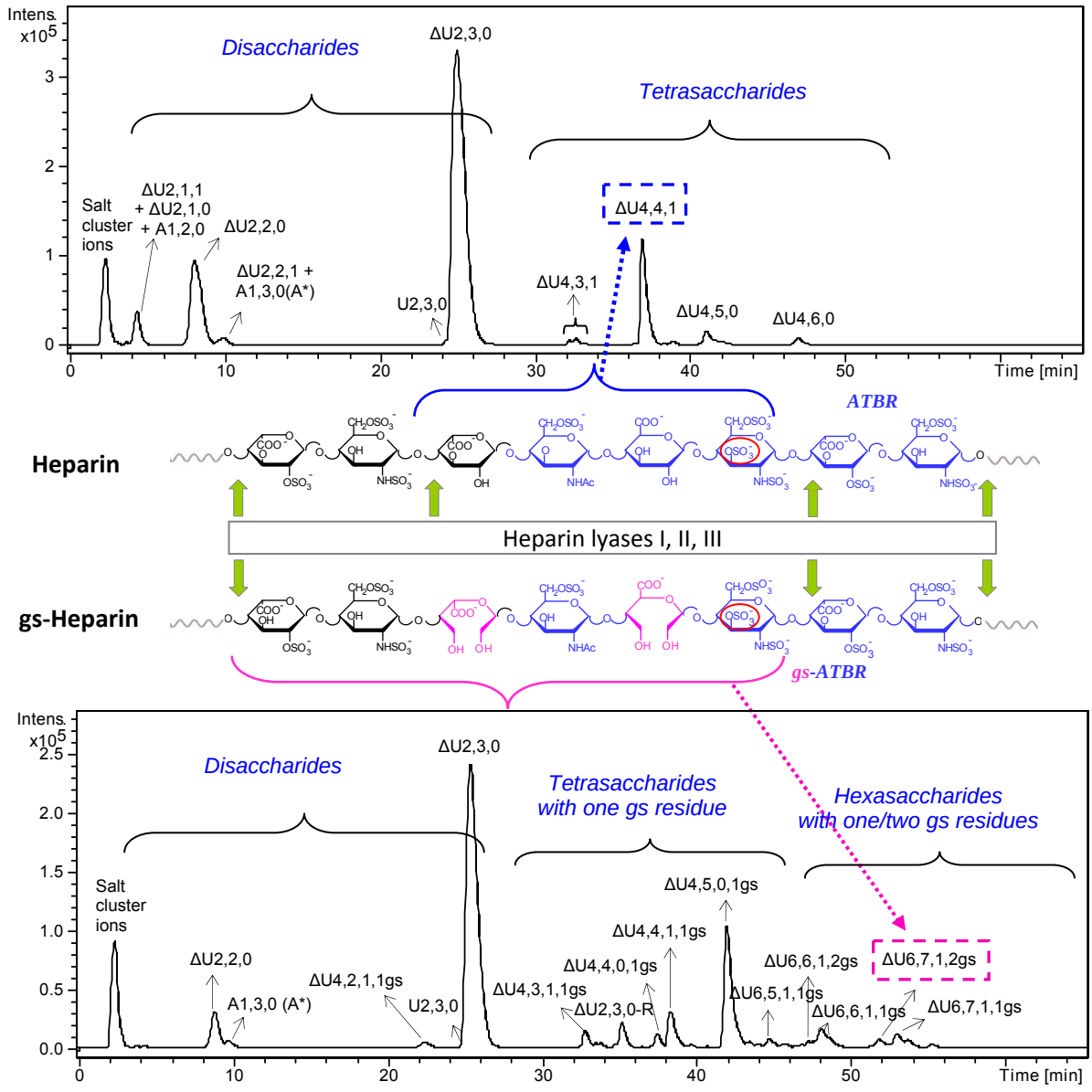



**Fig.3.16.** Comparison of LC-MS profiles of PMH (upper chromatogram) and its glycol-split derivative (lower chromatogram) after exhaustive enzymatic digestion

In the figure only major compounds are shown.

HPLC-UV/ESI-Q-TOF-MS conditions: see Fig. 3.12.

Abbreviation system for the "first-level structures" includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol R is used to indicate the presence of a remnant of gsG/gsI residue at the R

Most of the peaks generated upon enzyme treatment by both PMH and gs-PMH are attributable to oligomers initiating with a $\Delta U_{2S}$ residue (Table 3.6). Even if with different relative intensities, only the major unsaturated trisulfated disaccharide $\Delta U_{2S}$-$A_{NS6S}$ (ΔU2,3,0), characterizing the regular highly sulfated TSD sequences (Fig.1.19), and the minor disulfated unsaturated disaccharide $\Delta U_{2S}$-$A_{NS}$ (ΔU2,2,0) are common components of the digestion products of both unmodified and gs-species.

Two isomers ΔU2,2,0, characterized by slightly different retention time values (8.2 and 8.6 min), were distinguished using the corresponding standard solutions independently by monitoring LC-UV chromatogram (see below, Fig.3.20). The resolution in LC-MS was lower (Fig. 3.16) because of a possible mixing of isomers in the tube connecting UV and MS detectors. While $\Delta U_{2S}$-$A_{NS}$ is preserved, its positional non-2-*O*-sulfated isomer ΔU-$A_{NS6S}$ (ΔU2,2,0) as well as ΔU-$A_{NS}$ (ΔU2,1,0) and ΔU-$A_{NAc6S}$ (ΔU2,1,1) disappear completely in the gs-PMH digest due to the splitting of non-sulfated uronic acids residues. The disaccharides that disappeared upon glycol-splitting were expected to be incorporated in the longer oligosaccharides containing the generated gs-residues. Indeed, heparin digests do not contain oligosaccharides longer than tetramers, while gs-heparin digests contain disaccharides preserved after glycol-splitting, **gs-tetrasaccharides** and **gs-hexasaccharide** fragments (Fig.3.16). Also traces of several **gs-octasaccharides** with three gs-residues were observed in the gs-PMH digest (*not shown, see Discussion*).

The most prominent oligosaccharide peak of gs-PMH corresponds to the pentasulfated ΔU4,5,0,1gs containing a TSD unit followed by a gs-residue and $A_{NS6S}$ (Fig.3.17a). It is a marker of the TSD units that were adjacent to non-sulfated uronic acids in the original heparin chains. This component was isolated and its structure was confirmed by NMR (Fig. 3.17b). The corresponding unmodified tetrasaccharide $\Delta U_{2S}$-$A_{NS6S}$-G-$A_{NS6S}$ was previously found by Xiao et al. (2011).

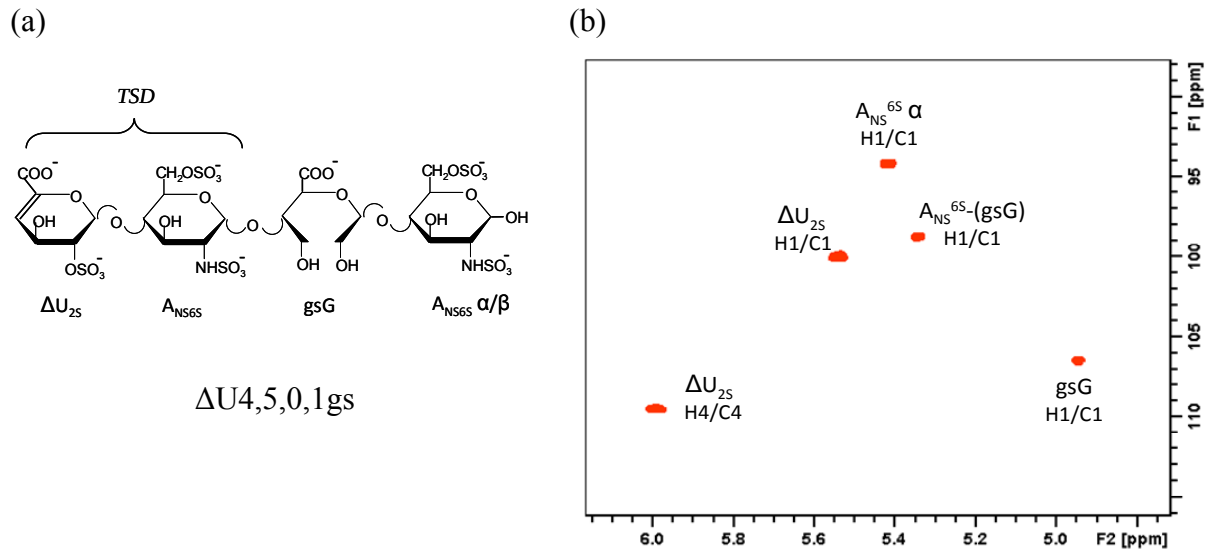
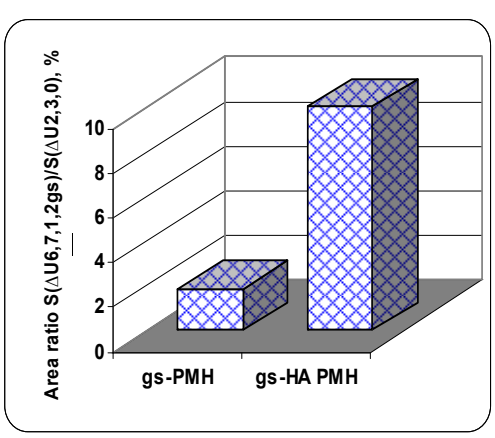



**Fig. 3.17**. Structure of the major gs-tetrasaccharide ΔU4,5,0,1gs of the heparinase-digest of gs-heparins and its partial HSQC NMR spectrum

It is worth noting that in several cases it was possible to determine the exact structure of components of the enzymatic digest, as for example oligosaccharides derived from the ATBR. The resistance to enzymatic cleavage of the glycosidic bond between glucosamine and glucuronic acid when this latter is followed by a 3-*O*-sulfated glucosamine is well known (*see* I.1.4.2). Accordingly, the most prominent tetrasaccharide peak ΔU4,4,1 (ΔU-$A_{NAc6S}$-G-$A_{NS3S6S}$) in the heparin digest is derived from the ATBR (Fig.3.16a, *Viskov* and *Mourier*, 2004). Gs residues induced an additional resistance to heparinase leading to the formation of a monoacetylated hexasaccharide with two internal gs-residues ΔU6,7,1,2gs (Fig.3.16b). The intensity of this signal increased in the gs-derivative of the heparin fraction with high affinity towards AT (fraction enriched in the sequences containing the active site), confirming the proposed structure (Fig.3.18).

The obtained results show also that ΔU4,4,1,1gs is not associated with the ATBR sequence (it is incorporated in the ΔU6,7,1,2gs), but should correspond to the sequence $\Delta U_{2S}$-$A_{NS6S}$-gsG/gsI-$A_{NAc6S}$.



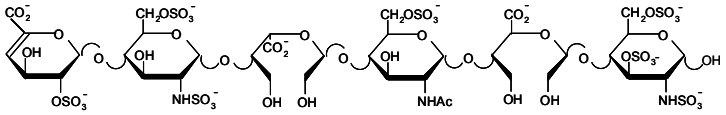




**Fig. 3.18.** Area ratio between peaks of ΔU6,7,1,2gs and ΔU2,3,0 for glycol-split derivatives of porcine mucosal heparin (gs-PMH) and its high AT(III)-affinity fraction (gs-HA-PMH)

Normalization to ΔU2,3,0 peak area was done in order to both evaluate the ATBR content with respect to regular highly sulfated TSD-sequences and to avoid errors associated with the stability of MS detector

As detailed in Table 3.6 for gs-PMH, heparinase-digestion of gs-heparins generate also minor, small-size fragments such as the monosaccharides A1,2,0 and A1,3,0(A*) (where A is $A_{NS6S}$ and A* is a 3-*O*-sulfated $A_{NS3S6S}$) and disaccharide ΔU2,3,0-R (where R is the remnant at the RE, Fig.3.19). All these compounds are markers of a hydrolytic process at the level of gs-units. For example, monosaccharides A1,2,0 and A1,3,0(A*) are usually present in the digest of PMH (indicating that several heparin chains have A at the NRE), but in low amount (*unpublished observation from our laboratory*). This is in agreement with NMR data, where the formation of $A^*_{NRE}$ upon glycol-splitting was observed (discussed in III.2). Trisaccharide ΔU3,4,0-ol (assigned to a compound terminating with a reduced uronic acid, Fig.3.19) is likely to be generated from the heparin chains bearing uronic acid at the RE. The presence of A and U at the NRE and RE, respectively, is in agreement with the possible depolymerization of the macroheparin precursor by heparanase.

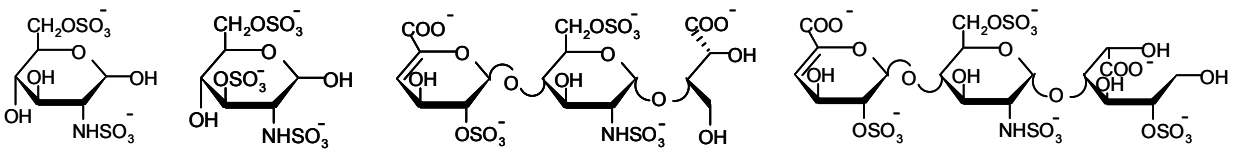



**Fig. 3.19.** Structure of minor small-sized fragments found in the heparinase-digests of gs-heparins

### III.4.4. LC-MS differentiation of gs-heparins obtained from heparins of different sources

After having identified all components of the heparinase-digests of PMH and gs-PMH, the developed approach has been applied for profiling gs-heparins of other sources (gs-OMH, gs-BMH, gs-BLH), including gs-derivatives of a ~40% glycol-split derivative of PMH (HI-3) to verify the sensitivity of the approach for detecting structural differences within gs-heparins chains. For comparing these samples, we focused our attention on three aspects: *(1)* comparison of the disaccharides composition, *(2)* comparison of tetra- and hexasaccharide fractions and, *(3)* finally, the absence or presence of octasaccharidic components.

*Comparison of the disaccharides composition by quantitative LC-MS analysis*

As opposed to tetra- and hexasaccharides, the whole series of disaccharide standards are commercially available, thus permitting their quantitative analysis. Because the enzymatically generated disaccharides have a chromophore group absorbing at a characteristic wavelength (232 nm), the quantitative analysis using both MS (Fig.3.16) and UV (Fig.3.20) detector was performed. The regression equations, correlation coefficients, limit of detection (LOD), limit of quantification (LOQ) and area variability (RSD) for the major disaccharides present in gs-heparin digests ($\Delta U_{2S}$-$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$) are given in Table 3.7.

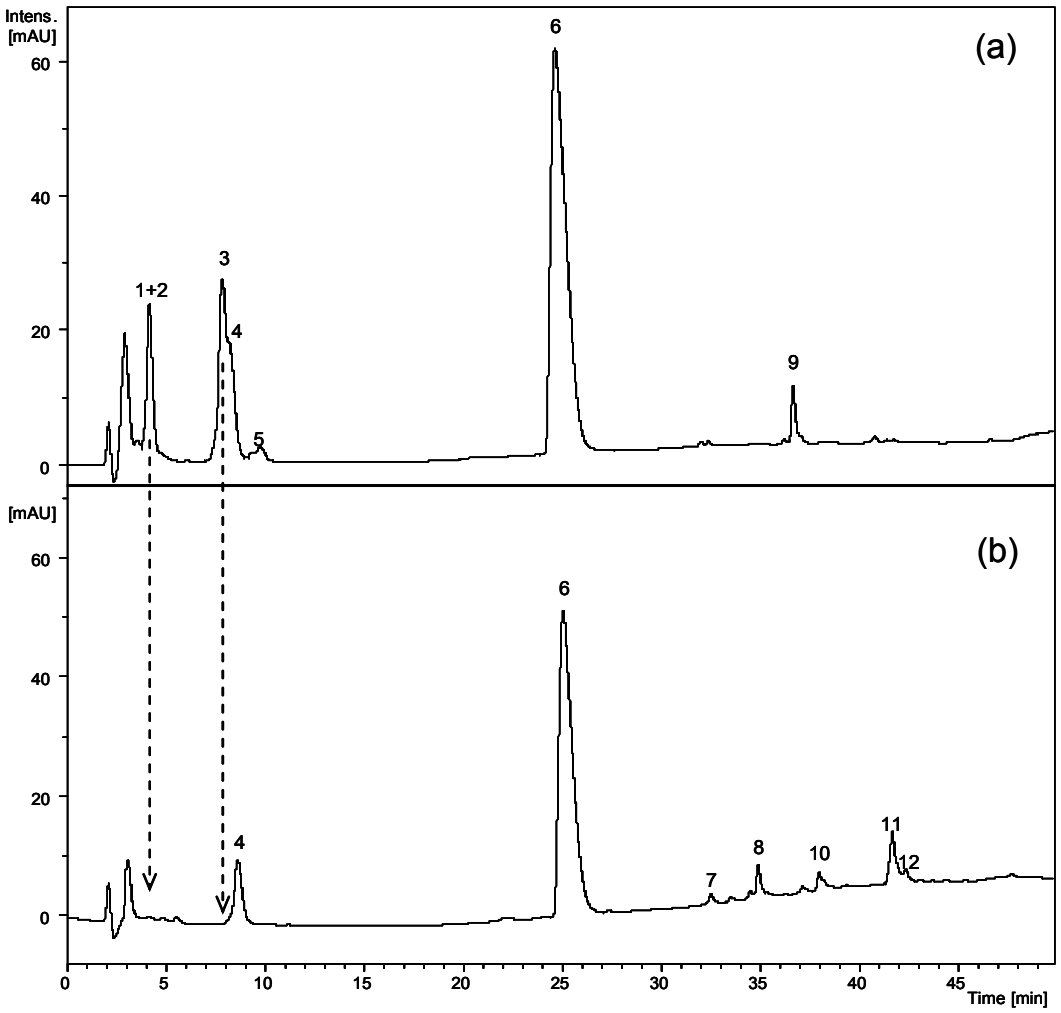



**Fig. 3.20.** LC-UV chromatograms (232 nm) of heparinase-digests of PMH (a) and gs-PMH (b)
Arrows evidentiate that glycol-splitting induced disappearance of disaccharides 1, 2 and 3 originally containing non-sulfated G/I. Upon glycol-splitting, these heparin units become part of tetra- and hexasaccharides containing gs residues. 1 – ΔU2,1,0 (ΔU–$A_{NS}$), 2 – ΔU2,1,1(ΔU–$A_{NAc6S}$), 3 – ΔU2,2,0 (ΔU–$A_{NS6S}$), 4 – ΔU2,2,0 ($\Delta U_{2S}$–$A_{NS}$), 5 – ΔU2,2,1 ($\Delta U_{2S}$–$A_{NAc,6S}$), 6 – ΔU2,3,0 ($\Delta U_{2S}$–$A_{NS6S}$), 7 – ΔU4,3,1,1gs, 8 – ΔU2,3,0-R, 9 – ΔU4,4,1, 10 – ΔU4,4,1,1gs, 11 – ΔU4,5,0,1gs, 12 – ΔU4,5,0.

**Table 3.7.** Regression equations, correlation coefficients ($R^2$), limit of detection (LOD), limit of quantification (LOQ) and area variability (RSD) for $\Delta U_{2S}$-$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$ detected by UV and MS

| **UV** (232 nm) | | | | | |
|---|---|---|---|---|---|
| | Regression equation ($y - S \cdot 10^{-1}$, mAU; x – C, μg/ml) | $R^2$ | LOD, μg/ml | LOQ, μg/ml | Area RSD*,% |
| $\Delta U_{2S}$-$A_{NS}$ | y = 2.137x – 0.120 | 0.9994 | 4.3 | 21.0 | 0.3 |
| $\Delta U_{2S}$-$A_{NS6S}$ | y = 1.621x – 0.122 | 0.9991 | 5.2 | 40.2 | 0.3 |
| **MS** (neb. 0.9 bar, dry gas 7 ml/min, 180ºC; +3.2 kV; m/z 200 – 2000 Da; flow-rate 100 μl) | | | | | |
| | Regression ($y - S \cdot 10^{-4}$, mAU; x – C, μg/ml) | $R^2$ | LOD, μg/ml | LOQ, μg/ml | Area RSD*,% |
| $\Delta U_{2S}$-$A_{NS}$ | y = 7.682x – 1.005 | 0.9989 | 0.13 | 0.24 | 3.5 |
| $\Delta U_{2S}$-$A_{NS6S}$ | y = 7.839x - 1.447 | 0.9989 | 0.11 | 0.20 | 3.4 |

* inter-day measurements, n=3, P=0.95
S – peak area

For the major trisulfated disaccharide $\Delta U_{2S}$-$A_{NS6S}$ the dynamic range of the calibration curve (range between LOQ and limit of linearity (LOL)), was from 40 to 1500 μg/ml when using UV detection at 232 nm and from 0.20 up to 550 μg/ml for MS detection. Because the $\Delta U_{2S}$-$A_{NS6S}$ concentration in digested gs-heparins varied from 300 to 600 μg/ml, each sample was diluted two times before LC/UV/MS analysis in order to avoid saturation of the MS detector and the chromatographic column. Injection of higher amounts of sample occasionally led to lower accuracy of quantitative analysis as well as peak broadening and decreasing efficiency. Both UV and MS calibration curves obtained for the less abundant $\Delta U_{2S}$-$A_{NS}$ were linear in the whole concentration range typical for non-modified and gs-heparins, namely 3 – 100 μg/ml. The lower limit of the dynamic range (LOQ) was found equal to 21.0 μg/ml for the UV detector and 0.24 μg/ml for the MS detector. As expected, the area variability (RSD) for UV detection is lower than for the MS (Table 3.7). So, for quantitative analyses, the calibration curves obtained with UV detector have been chosen. The MS detection cannot be used, especially, when saturated disaccharides, such as the trisulfated disaccharide 2,3,0 ($I_{2S}$-$A_{NS6S}$, RT 24.8 min) which partially overlaps with its unsaturated analog ΔU2,3,0 (RT 26.0 min, Table 3.5), are present. While they did not absorb at 232 nm, the MS detector is sensitive to these analytes that may lead to incorrectly higher results.

For PMH the ΔU-$A_{NS6S}$, ΔU-$A_{NAc6S}$, ΔU-$A_{NS}$ (absent in gs-PMH due to their incorporation into longer chain oligosaccharides) were also evaluated. Since the extinction coefficient at 232 nm for all mentioned disaccharides were the same under the applied analytical conditions, it was possible to use the calibration curve obtained for $\Delta U_{2S}$-$A_{NS}$ for calculating the total amount of two positional isomers ($\Delta U_{2S}$-$A_{NS}$ and ΔU-$A_{NS6S}$), that are eluted at the same retention time, and of ΔU-$A_{NAc6S}$, ΔU-$A_{NS}$.

The contents of disaccharides in digests of gs-heparins obtained from starting materials of different origin are reported in Table 3.8. The values are expressed as a percentage of the digested heparin or gs-heparin material (100 μg). Such an expression permits not only to

compare the different samples, but also to evaluate the content of the regular highly sulfated sequences, not adjacent to non-sulfated uronic acids, within heparin chains.

**Table 3.8**. Disaccharide compositional analysis determined by HPLC/UV, of PMH and different gs-heparins (*normalized to quantity of digested material*)*

| | x ± Δx, % mass. | | | | |
|---|---|---|---|---|---|
| | **ΔU-$A_{NS}$** | **ΔU-$A_{NAc6S}$** | **ΔU-$A_{NS6S}$** | **Δ$U_{2S}$-$A_{NS}$** | **Δ$U_{2S}$-$A_{NS6S}$** |
| | RT 4.3 min | RT 4.4 min | RT 8.2 min | RT 8.6 min | RT 25.6 min |
| PMH | 7.7 ± 0.3 | | 17.5 ± 0.3 # | | 66.7 ± 0.9 |
| gs-PMH | < LOD | < LOD | < LOD | 4.8 ± 0,2 | 49.2 ± 0.9 |
| gs-OMH | < LOD | < LOD | < LOD | 2.1 | 42.3 |
| gs-BMH | < LOD | < LOD | < LOD | 9.5 | 45.9 |
| gs-BLH | < LOD | < LOD | < LOD | 3.9 | 80.4 |
| HI-3 | < LOD | < LOD | < LOD | 2.7 | 32.0 |

*Analyses of digests of PMH and gs-PMH were performed in duplicate (n=2, P=0.95), for other samples single analyses were performed. The RT values shown in this table were obtained using UV detector.
LOD = limit of detection.
# The total amount of these two isomers was determined using the calibration curve obtained for $\Delta U_{2S}$-$A_{NS}$, the isomer that is preserved in gs-heparins. Even if ΔU-$A_{NS6S}$ is characterized by an extinction coefficient somewhat higher that of its isomer leading to an overestimated value, the percentage found for the sum of the two isomers is in agreement with the results obtained by other authors (*Wang* et al., 2012)

As expected, since the glycol-splitting reaction converts the non-sulfated uronic acids to glycol-split residues, only $\Delta U_{2S}$-$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$ are preserved during gs-reaction and, consequently, gs-heparin digests showed an overall content of disaccharides lower than unmodified heparin. Also, with the exception of gs-BLH, the total disaccharide content in digested gs-heparins of different type (43-55%) is consistently lower than for PMH (~92%). The latter sample contains, as opposed to the digests of the gs-heparins, non-2-*O*-sulfated disaccharides as well as a higher quantity of the 2-*O*-sulfated ones. The lower content of $\Delta U_{2S}$-$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$ in the gs-heparin digests with respect to the PMH digest is explained by their incorporation into tetra- (for example, ΔU4,5,0,1gs in Fig. 3.17) and higher oligosaccharides containing gs-residues. The high disaccharide content of gs-BLH (~85%) correlates with its high sulfation degree reported in the literature (*Guerrini* et al., 2001). The extensively gs-PMH variant (HI-3) shows the lowest content of disaccharides (~35%) due to the highest non-sulfated uronic acid content in the starting material.

*Comparison of tetra- and hexasaccharide fractions*

The second step consisted in the comparison of the tetra-and hexasaccharide contents. As illustrated in Fig.3.21, the LC-MS profiles of tetra- and hexasaccharide regions of the heparinase digests of gs-heparins, obtained from heparin samples of different tissue/species, significantly differ from each other.

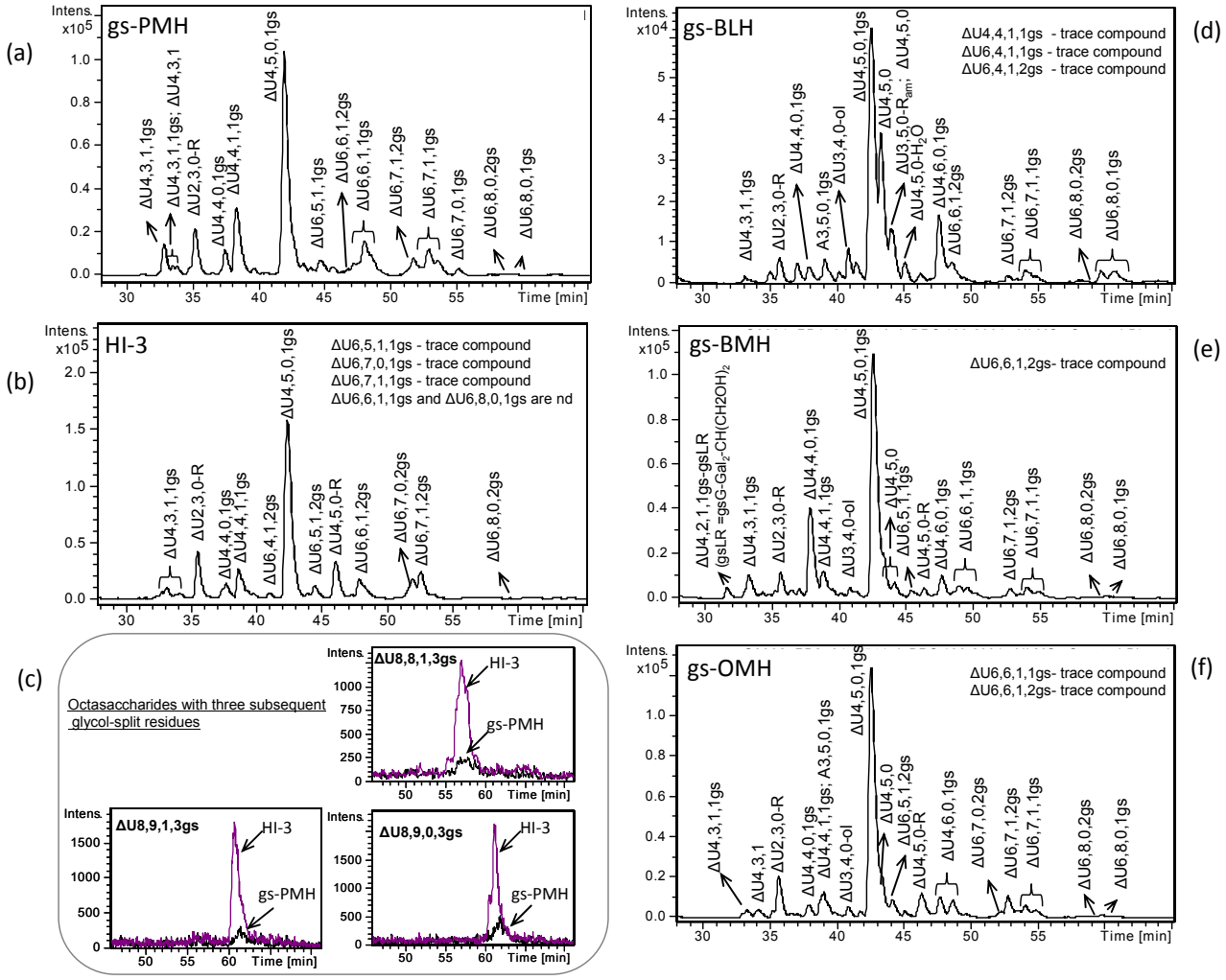
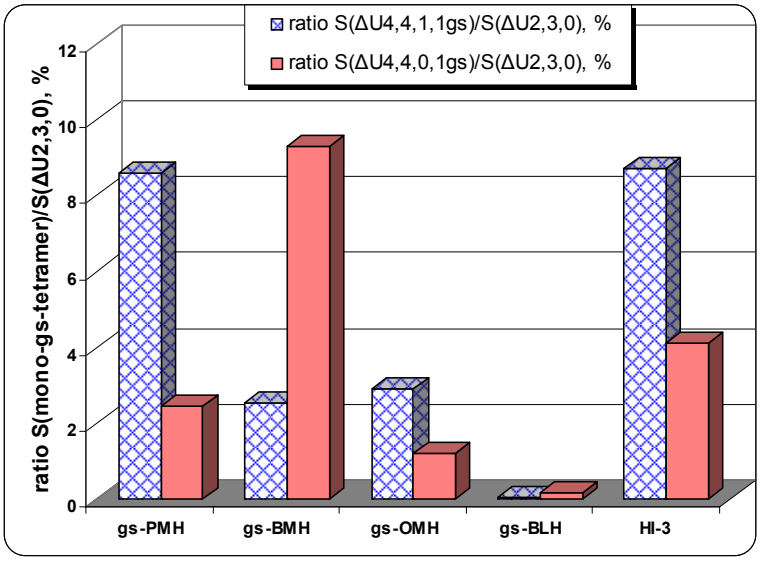



**Fig.3.21.** LC-MS profiles of heparinase digested tetra- and hexasaccharides of glycol-split-heparins of different animal and organ sources including HI-3 derived from partially 2-*O*-desulfated PMH. In panel (c) extraction ion chromatograms (EICs) of octasaccharides with three gs residues found in gs-PMH and HI-3 digests are shown.
HPLC-UV/ESI-Q-TOF-MS conditions: see Fig.3.12.

Even if most of the fragments contain gs residues, the relative proportion of peaks varies for the different heparins and their overall profiles constitute characteristic fingerprints. The

chromatogram of gs-BLH digest displays the largest differences with respect to those of reference gs-PMH. Generally, the signal intensities of the oligosaccharidic region of the chromatogram, where most of the peaks of gs-compounds are observed, are the lowest for gs-BLH (Fig. 3.21). This is in agreement with the high sulfation degree of gs-BLH showed by NMR and the high content of ΔU2,3,0 determined by LC-MS quantitative analysis (Table 3.8). The low *N*-acetylation degree of BLH is also reflected in the relative content of tetrasulfated tetrasaccharides ΔU4,4,0,1gs and ΔU4,4,1,1gs (Fig. 3.22). In the gs-PMH the *N*-acetylated form appeared much more intense, while their intensities are inverted in gs-BMH, and, finally, in the highly sulfated gs-BLH the acetylated tetrasaccharide appeared only in trace. This was found to be in total agreement with the quantitative NMR data about the content of the $A_{NAc}$. In the HI-3 digest the relative content of ΔU4,4,0,1gs increases with respect to ΔU4,4,1,1gs (when comparing with their contents in the gs-PMH digest) because of the higher content of non-sulfated uronic acids within its chains.




$A_{NAc}$ (%): **9.6** **5.2** **5.0** **< LOQ** **10.3**

**Fig. 3.22**. Peak area ratio of mono-gs-tetrasaccharides ΔU4,4,0,1gs and ΔU4,4,1,1gs found in the enzymatic digests of gs-heparins of different sources

Normalization to ΔU2,3,0 peak area was done in order to both evaluate the gs-tetrasaccharide content with respect to regular highly sulfated TSD-sequences and to avoid errors associated with the stability of MS detector.

$A_{NAc}$ content (molar %) determined by quantitative 2D NMR is shown for each sample (Table 3.2).

One of the most interesting difference observed for gs-BLH is a low content of ATBR derived hexasaccharide ΔU6,7,1,2gs (Fig.3.21), even if NMR data showed a content of the trisulfated glucosamine $A_{NS3S6S}$ (6.6 molar %) comparable with that of gs-PMH (5.8 molar %).

This apparent disagreement, can be explained by different sulfation/*N*-acetylation pattern of the ATBR of the highly sulfated bovine lung heparin, which is known to contain *N*-sulfated glucosamine ($A_{NS6S}$-G-$A_{NS3S6S}$-$I_{2S}$-$A_{NS6S}$) instead of the *N*-acetylated one, that, unlike in PMH, is preceded by 2-*O*-sulfated iduronic acid (*Loganathan* et al., 1990). Combination of glycol-splitting and enzymatic digestion generates this tetrasaccharide, instead of the hexasaccharide found in gs PMH digest (Fig.3.23). Accordingly, a significant increase of ΔU4,6,0,1gs (Fig.3.21), corresponding to the highly sulfated ATBR of BLH, is observed.

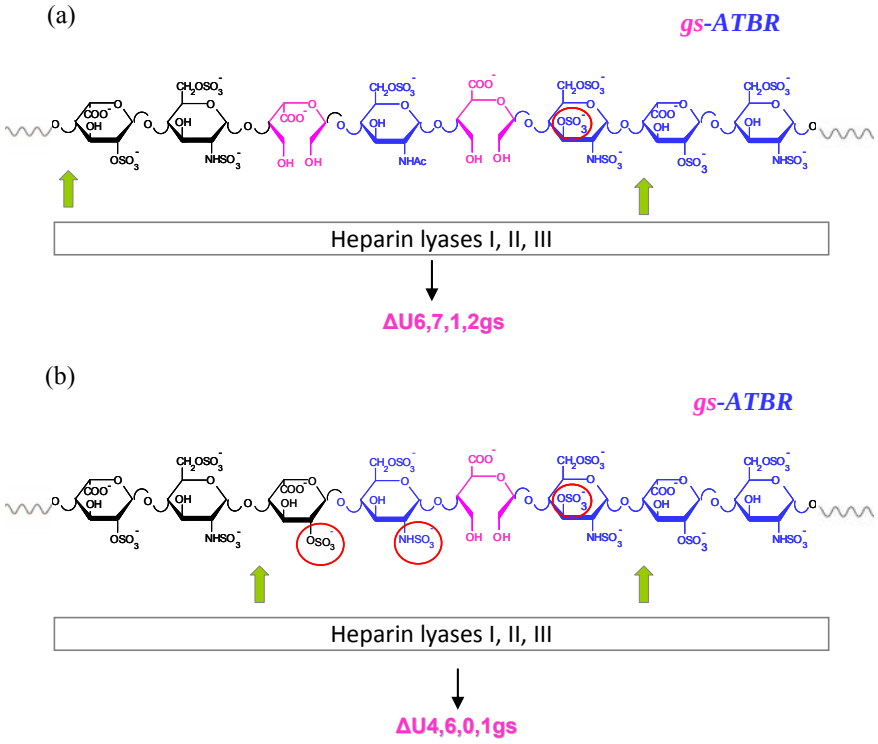



**Fig. 3.23**. Scheme of generation of ΔU6,7,1,2gs (a) and ΔU4,6,0,1gs (b) from different types of ATBR-sequences by exhaustive enzymatic digestion of gs-heparins

The first type (a) is characteristic for PMH, while the second type (b) is typical for highly sulfated BLH

In contrast to gs-BLH, the chromatographic profile of HI-3 (extensive gs-PMH with ~40% of gs-residues) is characterized by the higher intensities of tetra- and hexasaccharide fractions with respect to PMH (Fig. 3.21). This result correlates with the quantitative data on the disaccharide content (Table 3.8). The detectable hexasaccharides of HI-3 are mainly those that contain two gs residues.

Some octasaccharides containing three gs residues were detectable in the enzymatic digests of gs-heparins only in trace, however, their content increases in HI-3 (Fig. 3.21c). Also higher masses (dp > 8) were not detectable in the gs-PMH digest, but were observed in trace

amounts in HI-3. Their identification became feasible only after previous isolation of the corresponding fractions by SEC from the exhaustive enzymatic digest of HI-3 (Fig.3.24, *see* Discussion).

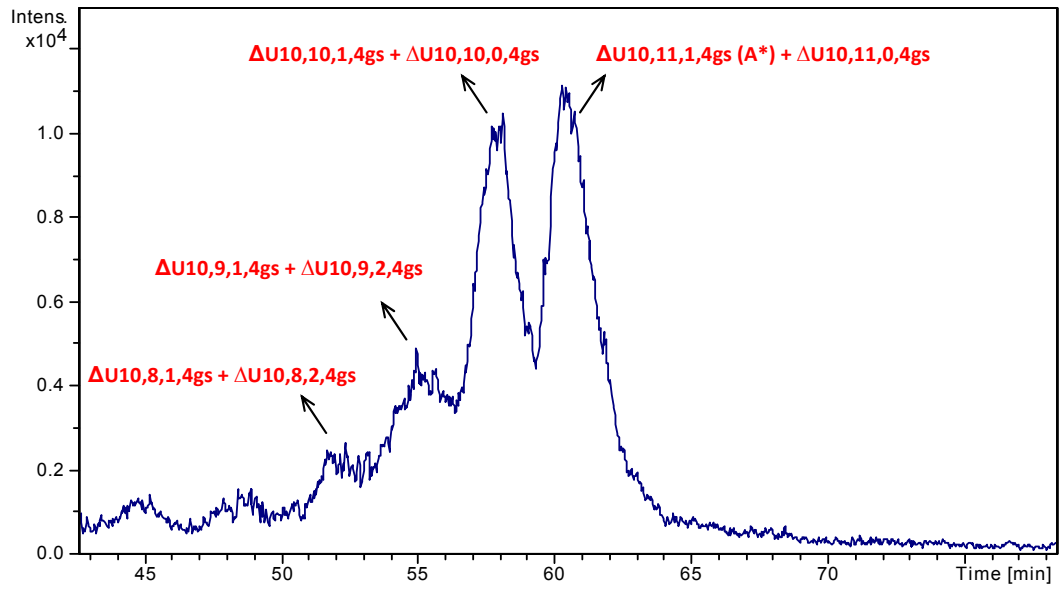



**Fig.3.24.** Extracted ion chromatogram of the decasaccharide fraction obtained by SEC-fractionation of HI-3 heparinase-digest

The work is in progress (see Chapter V). Only major components are shown in these preliminary results. Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE.

### III.5. Sensitivity of the developed approach to the glycol-splitting reaction conditions

Having analyzed all the gs-heparin samples, it appeared evident that the reaction conditions led to partial hydrolysis of heparin chains. The $M_w$ decreases substantially for each heparin sample after glycol-splitting reaction (*see* III.1). Moreover, the quantity of generated gs residues obtained by 2D NMR was lower than the original amount of non-sulfated uronic acids in the starting materials (*see* III.2). LC-MS analysis of the digested gs-heparins showed the presence of species bearing remnant of gsU residues, indicating a partial hydrolysis at the level of gs-residues. We supposed that the addition of 0,1 M HCl solution for pH neutralization after borohydride treatment might create locally very strong acidic conditions causing Smith degradation.

In order to improve the reaction procedure and minimize partial hydrolysis, HCl could be substituted with a mild organic acid. As previously, PMH was chosen as a model and its gs-derivative was prepared using 5% acetic acid for pH neutralization, and then characterized by the developed method. As expected, the obtained sample had higher $M_w$ (16.6 kDa) and content of gs-units (22.5 %). SEC-TDA and NMR data obtained for the two gs-PMH prepared by different methods are shown in Table 3.9. The NMR spectrum of the less hydrolyzed gs-PMH shows higher amounts of gs residues, and even gsG followed by trisulfated glucosamine became detectable (Fig.3.25b). The marker of hydrolysis ΔU2,3,0-R appeared in LC-MS chromatogram of the digest gs-PMH prepared under mild conditions only in trace amounts (Fig.3.26).

**Table 3.9.** Main parameters ($M_w$, gs uronic acid content and the relative content of the hydrolysis marker ΔU2,3,0-R) that can be affected by glycol-splitting reaction conditions

| | SEC-TDA | | 2D HSQC NMR (%) | | | | | LC-MS |
|---|---|---|---|---|---|---|---|---|
| | $M_w$, kDa | | $I_{2S}$ | gsI | gsG | $A_{NS/NAC}$-(R) | | Area ratio (%) ΔU2,3,0-R/ΔU2,3,0 |
| gs-PMH (a) | 10.8 | | 89.5 | 2.0 | 8.3 | + | | 4.4 |
| gs-PMH (b) | 16.6 | | 77.5 | 9.6 | 12.9 | trace | | 0.4 |

(a) obtained by adding a 0,1M HCl solution at the step of pH adjustment
(b) obtained by adding a 5% (v/v) $CH_3COOH$ solution at the step of pH adjustment

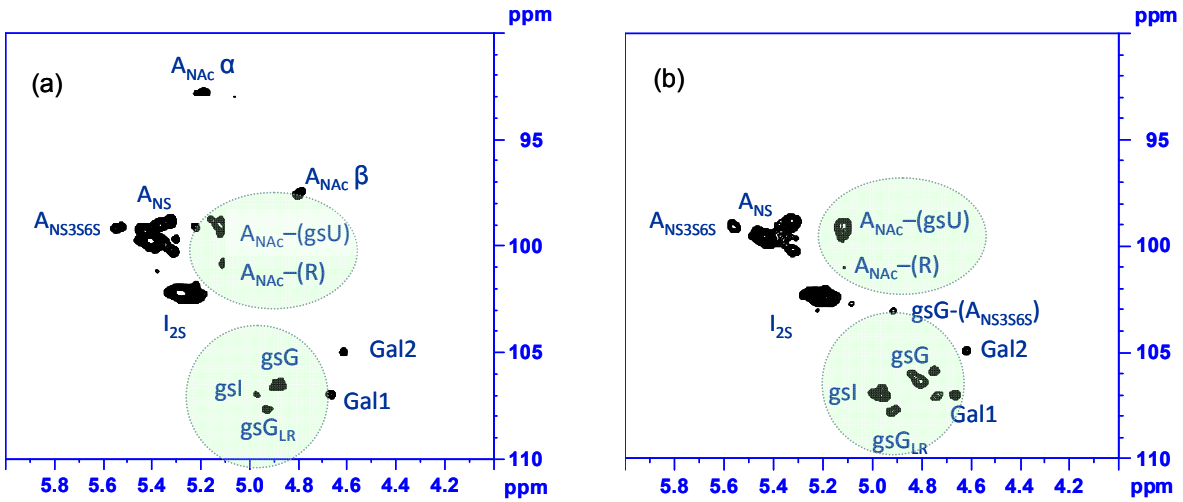



**Fig.3.25.** Comparison of HSQC spectra of two gs-PMH prepared under slightly different conditions: (a) using 0.1 M HCl, (b) using 5% (v/v) $CH_3COOH$ for pH adjustment at the last step

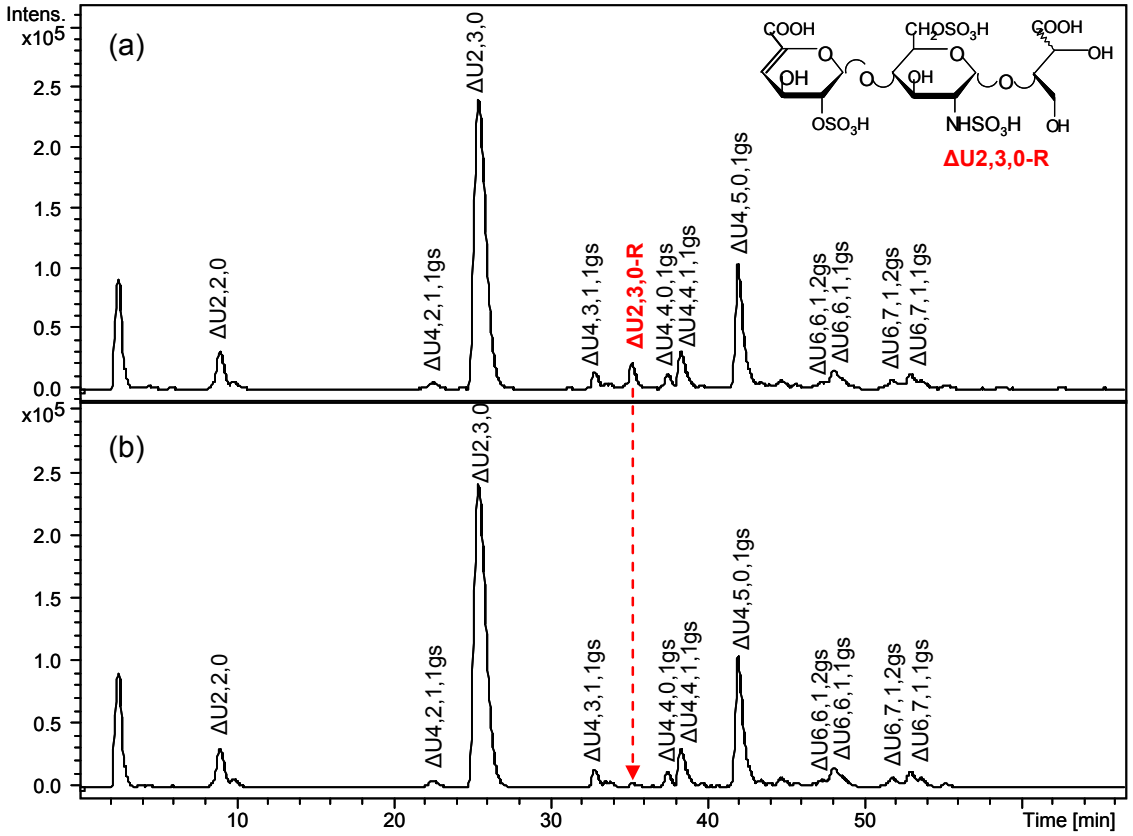



**Fig. 3.26.** LC-MS profiles of heparinase-digested gs-PMH samples prepared under slightly different conditions: (a) using 0.1M HCl, (b) using 5% (v/v) $CH_3COOH$ for pH adjustment at the last step

The main observed difference is the decreased content of hydrolysis marker ΔU2,3,0-R

### III.6. Discussion

Structural characterization of gs heparins is of high interest because their potential anti-cancer and anti-inflammatory activity is likely affected by location of gs-residues along the sulfated polysaccharide chains as well as by any modification in the original heparin sequences. Such a characterization is especially difficult because of problems associated with intrinsic microheterogeneity and polydispersity. In the present work, these problems have been faced by converging utilization of advanced NMR, SEC-TDA and LC-MS methods.

The performed NMR studies have permitted to assign for the first time all the cross peaks associated with gs residues in the 2D HSQC NMR spectra of gs-heparins prepared from heparins of different animal and tissue sources. The obtained NMR spectra, especially their anomeric regions, appeared highly characteristic for gs-heparins (Fig. 3.7). Their NMR

quantitative analysis was also performed relying on the required condition that the coupling constants $^{1}J_{C\text{-}H}$ for gs-residues were almost the same than those of the unmodified uronic acids. It permitted the structural characterization of individual gs-samples as well as to distinguish and to compare them in terms of monosaccharide composition.

Notably, different types of gs-containing sequences, including gs uronic acids followed by a trisulfated glucosamine (the recognized marker of ATBR) can be distinguished by 2D HSQC NMR (Table 3.2, Fig.3.25). A similar signal was observed also in HI-3 spectrum (Fig.3.3a) but with slightly different chemical shift (4.94/103.1 ppm). Since it becomes visible in the spectrum of the gs-product derived from partially 2-*O*-desulfated heparin, we supposed it corresponded to gsI-(A*). This result confirms that there are some $I_{2S}$-$A_{NS3S6S}$ units outside the ATBR of original heparin chains and could explain the fact that the G-(A*) content was found almost always lower than total A* content for unfractionated heparins (*Rudd* et al, 2012). The presence of ΔU2,4,0, corresponding to $\Delta U_{2S}$-$A_{NS3S6S}$, in heparinase-digested heparins (*unpublished observation of our group*), is also in agreement with the observed data.

Also the residues of the linkage region in gs-heparins were identified by NMR. Such a sequence is present in most of the unmodified heparins (*Iacomini* et al., 1999) as well as in some LMWHs (*Bisio* et al., 2009). Two residues of the LR, i.e., the $G_{LR}$ and the terminal Xyl (when not lost in the purification/bleaching treatments of the pharmaceutical grade polysaccharide) have been shown to be susceptible to glycol-splitting. In the HSQC NMR analysis of the anomeric region in gs-heparins (Fig. 3.1, 3.7), their glycol-splitting is clearly reflected by the corresponding signals at 4.92/107.7 and 4.75/105.9 ppm (Table 3.2, Fig.25b).

Additionally, the spectra contain several signals that may be used as markers of both glycol-splitting reaction completeness (by monitoring the anomeric signals of unmodified I and G as well as H2/C2 cross peak of G, Fig. 3.1) and hydrolytic processes (anomeric signals of $A_{NS/NAc}$ bearing a remnant of gsG/gsI, Fig.3.6a, H2/C2 cross peaks of trisulfated glucosamine at the NRE, Fig.3.1d).

The obtained results show that uncontrolled reaction conditions, first of all uncontrolled pH value, may lead to different hydrolysis processes, at the level of either gs-units (Fig.1.19, *Conrad* et al., 1992) or glycosidic bonds (*Jandik* et al, 1996). Moreover, it was observed that gsI are hydrolyzed to a higher degree with respect to the gsG. The characteristic cross peaks of $A_{NS}$ and $A_{NAc}$ bearing the generated remnant (R) in HSQC spectra were also assigned. Another marker of the hydrolysis is the formation of glucosamine at the NRE, as shown in Fig.1.19e. While the cross peaks of *N*-sulfated and *N*-acetylated NRE glucosamines overlap with those of the same units inside the internal sequences, the trisulfated glucosamine formed at the NRE can be easily detected by monitoring the H2/C2 signals of A* and $A^*_{NRE}$ (Fig. 3.1d). The $A^*/A^*_{NRE}$ ratio in the ring region of the HSQC spectrum may be used as a marker for evaluation of hydrolysis degree of A*-containing sequences (including gs-ATBR).

A set of NMR signals, associated with the terminal reducing aminosugars (such as α- and β-anomers of $A_{NAc}$ were observed, Fig.3.7), are clearly evident in the HSQC spectra of gs-heparins, but not significantly present in the parent heparin spectra. We have shown that also hydrolysis of the glycosidic bonds may occur under strong acid conditions, leading to the formation of RE amino sugars (Fig.3.6). The present study indicated that the major terminal reducing glucosamine in gs-heparins is *N*-acetylated, with typical signals at 5.16/92.7 ppm for α anomer and 4.77/97.5 for β one (*Guerrini* et al., 2008). A substantial loss of *N*-acetylated glucosamine-containing sequences upon periodate oxidation, followed by treatment with base, was reported for heparin chains with high affinity for AT and shown to occur preferentially at the level of the active site for AT (-I-$A_{NAc(6S)}$-G-$A_{NS3S6S}$-$I_{2S}$-$A_{NS6S}$-) (*Islam* et al., 2002).

Evidently, these results should be taken into account for monitoring the conditions for the preparation of gs products. Both NMR and SEC-TDA represent attractive tools for this aim. However, minor modifications, including side products, can be unraveled only through careful cleavage of gs-heparin chains and identification of the generated di- and oligosaccharide fragments by LC-MS analysis.

The assumption, at the basis of the present work that digestion of glycol-split heparins with a mixture of heparin lyases I, II and III would generate small oligosaccharides containing glycol-split residues was correct. The combination of glycol-splitting and enzymatic digestion leads to the skipping of the units containing gs residues and generation of longer oligosaccharides with internal gs residues (Fig. 3.11). As exemplified in Fig.3.16 and summarized in Table 3.6 for the gs-derivative of a sample of the most common heparin (gs-PMH) and illustrated in Fig.3.21 for gs-heparins of different origin, the LC-MS analysis indicates that, in addition to "normal" disaccharides as expected major products, heparinase digests of gs-heparins are largely constituted by tetra- and hexasaccharides containing one or two glycol-split residues. Most of fragments in the heparinase digests of gs-heparins are even-numbered and start at their NRE with an unsaturated, 2-*O*-sulfated uronic acid residue. Odd-numbered fragments are only present in very minor amounts.

Analyses based only on mass values do not permit to discriminate among all the different sequences of isomeric oligosaccharides. However, reasonable assumptions on the actual sequences of fragments containing gs residues can be made by comparison with structures established for the major oligosaccharides obtained from heparin lyase digests of unmodified heparins. The structure of gs-containing oligosaccharides expected on the basis of the major PMH oligosaccharides isolated and identified thus far (*Xiao* et al., 2011,*Yamada* et al., 1998) are compared in Table 3.10 with the major fragments identified in the present work for gs-PMH. Clearly, fragments containing gs residues derive from non-sulfated G and/or I of heparin sequences.

Even if the initial aim of the work was to structurally characterize gs-heparins, the developed method appeared a powerful tool for characterizing the starting heparin material. As an example, different types of ATBR sequences, differing by sulfation/*N*-acetylation pattern, were found in gs-heparin digests (Fig. 3.23).

**Table 3.10.** Several major G- and I-containing tetra- and hexasaccharide fragments obtained from PMH (literature data) and the corresponding gs-oligosaccharides found in this work

| Literature data | | Present work | |
|---|---|---|---|
| ΔU4,4,1 | $\Delta U_{2S}$-$A_{NAc6S}$-G-$A_{NS6S}$ | ΔU4,4,1,1gs | $\Delta U_{2S}$-$A_{NAc6S}$-gsG-$A_{NS6S}$ |
| ΔU4,4,0 (1) | $\Delta U_{2S}$-$A_{NS}$-G-$A_{NS6S}$ | ΔU4,4,0,1gs | $\Delta U_{2S}$-$A_{NS}$-gsG-$A_{NS6S}$ |
| ΔU4,4,0 (2) | $\Delta U_{2S}$-$A_{NS6S}$-G-$A_{NS}$ | ΔU4,4,0,1gs | $\Delta U_{2S}$-$A_{NS6S}$-gsG-$A_{NS}$ |
| ΔU4,5,0 | $\Delta U_{2S}$-$A_{NS6S}$-G-$A_{NS6S}$ | ΔU4,5,0,1gs | $\Delta U_{2S}$-$A_{NS6S}$-gsG-$A_{NS6S}$ |
| ΔU4,6,0 | $\Delta U_{2S}$-$A_{NS6S}$-G-$A_{NS3S6S}$ | ΔU4,6,0,1gs | $\Delta U_{2S}$-$A_{NS6S}$-gsG-$A_{NS3S6S}$ |
| ΔU6,6,1 (1) | $\Delta U_{2S}$-$A_{NS6S}$-I-$A_{NAc6S}$-G-$A_{NS6S}$ | ΔU6,6,1,2gs | $\Delta U_{2S}$-$A_{NS6S}$-gsI-$A_{NAc6S}$-gsG-$A_{NS6S}$ |
| ΔU6,6,1 (2) | $\Delta U_{2S}$-$A_{NS6S}$-I-$A_{NAc6S}$-G-$A_{NS3S}$ | ΔU6,6,1,2gs | $\Delta U_{2S}$-$A_{NS6S}$-gsI-$A_{NAc6S}$-gsG-$A_{NS3S}$ |
| ΔU6,7,0 | $\Delta U_{2S}$-$A_{NS6S}$-G-$A_{NS6S}$-G-$A_{NS6S}$ | ΔU6,7,0,2gs | $\Delta U_{2S}$-$A_{NS6S}$-gsG-$A_{NS6S}$-gsG-$A_{NS6S}$ |
| ΔU6,7,1 | $\Delta U_{2S}$-$A_{NS6S}$-I-$A_{NAc6S}$-G-$A_{NS3S6S}$ | ΔU6,7,1,2gs | $\Delta U_{2S}$-$A_{NS6S}$-gsI-$A_{NAc6S}$-gsG-$A_{NS3S6S}$ |

Oligosaccharides from unmodified heparin were identified from Linhardt's group upon exhaustive digestion with heparinase I, II, or III (Xiao et al., 2011) and partial digestion only with heparinase I (*Xiao* et al, 2011). Hexasaccharides were obtained from pig mucosal heparin by Sugahara group (*Yamada* et al, 1998) from a "hexasaccharidic fraction" isolated by gel-filtration on Bio-Gel P-10 and sub-fractionated on a column of amine-bound silica under a sodium phosphate gradient. Tetrasaccharide $\Delta U_{2S}$-$A_{NS6S}$-G-$A_{NS3S6S}$ was isolated from bovine lung heparin (*Loganathan* et al., 1990).

Interestingly, also octasaccharides with three gs units, even if as trace components, were found in the heparinase-digests of gs-heparins (Fig.3.21c), indicating that the sequences with up to three subsequent non-sulfated uronic acids may exist within the original heparin chains. The observed octasaccharides should contain *N*,6-*O*-sulfated glucosamines, suggesting that sequences with subsequent non-sulfated uronic acids are still highly sulfated. This result may suggest that the internal undersulfated domain of heparin chains is very short.

Longer chain oligosaccharides were observed only in the HI-3 enzymatic digest after SEC-fractionation. Several decasaccharides with four gs-residues were observed (Fig.3.24). These compounds differ not only by their sulfation degree, but also by different number of *N*-acetyl-glucosamines (from 0 up to 2) in their structures. In contrast to the found gs-decasaccharides, octasaccharides with three gs-residues containing more than one *N*-acetyl-group were not found (except for the sequences close to the LR, see Chapter IV) in the gs-heparin digests, nor in the HI-3 digest. It is likely that two *N*-acetyl-glucosamines do not exist within the sequences shorter than decasaccharides and they are separated by at least three *N*-sulfated disaccharide units within heparin chains.

Although the main scope of this study was the identification and profiling of the major components of glycol-split heparins, some minor mono- and tri-saccharide components of the enzymatic digests were also identified (Table 3.6). These small fragments may provide information on end-groups in the original heparin and/or on those generated by some fragmentation associated with the glycol splitting reaction. Thus, the finding of monomeric glucosamine fragments 1a (A1,2,0) and 1b (A1,3,0(A*)), where A* indicates that the aminosugar is 3-*O*-sulfated, rises the question whether they can be taken as a proof of concept that heparin molecules essentially as used in clinics are the result of cleavage, by an *endo*-β-D-glucuronidase of consistently longer chains of a proteoglycan biosynthetic precursor ("macromolecular heparin") (*Gong* et al., 2003). That enzyme is essentially a heparanase and would cleave the heparin chains at the level of G residues (*Sandbäck-Pikas* et al., 2008) generating shorter chains terminating with an aminosugar at their non-reducing end. Exhaustive cleavage with heparinases of these chains could then release glucosamine monomers as found in the present work (Fig.3.16) and in a number of other heparin preparations (*unpublished data from our laboratories*). However, an alternative explanation, namely the hydrolysis of the glycosidic bond between G and $A_{NS3S6S}$ within the ATBR (*Islam* et al., 2002) should be considered. In fact, gsG-A* appeared as a low signal in the NMR spectra, along with the generation of A* at the NRE (*discussed above*).

As by NMR, also LC-MS analysis revealed partial Smith degradation, i.e., cleavage of gs-heparins at the level of gs-residues as previously observed (*Conrad* and *Guo*, 1992): several hydrolysis marker fragments (such as ΔU2,3,0-R and ΔU4,5,0,1gs-R, Fig.3.19, Table 3.6) terminating with the remnant (R) of a glycol-split and hydrolyzed uronic acid, were identified for some gs preparations. Such a cleavage may occur under relatively mild acid conditions (*Conrad* and *Guo*, 1992). In addition, borohydride reduction could modify some of the residues at the reducing end, converting them into the corresponding alditols, such as ΔU3,4,0-ol (Fig. 3.19, Table 3.6).

### III. 7. Conclusions

Glycol-split heparins from different animal sources have been characterized for the first time by the combined NMR and LC-MS analysis. While 2D HSQC NMR permits to evaluate the average monosaccharide composition, HPLC/MS analysis of glycol-split heparins, obtained from commercial and chemically modified heparins under the optimized conditions of the present work, permits to profile their di-, tetra-, and hexasaccharide composition in a way that is characteristic of the structure of the starting materials and especially emphasizes the extent of their glycol-splitting. The combined analytical approach is also suitable to distinguish glycol-split heparins obtained from heparins of different animal and tissue origin. It is also attractive to think that, in parallel with an analysis of the parent heparins, analysis of their glycol-split derivatives could permit to extend structural information to the original unmodified heparins, especially regarding sequences containing non-sulfated G and I residues. It has been shown that this method may permit to characterize the sequences with subsequent non-sulfated uronic acids, and their length, as well as the distribution of *N*-acetyl-glucosamines within heparin sequences. As additional analytical tool, size-exclusion chromatography coupled with triple detector array, not requiring the use of oligosaccharide standards, permits both to characterize molecular weight distribution and to determine hydrodynamic parameters for heparin-related samples. In the case of gs-derivatives the modifications in flexibility of heparin chains, that may be important for understanding their biological activity, can be monitored. Combined together NMR, LC-MS and SEC-TDA proved to be a powerful tool for reaction monitoring and controlling both glycol-splitting completeness and hydrolytic side reactions that can contribute to additional microheterogeneity of the final gs-product.

# CHAPTER IV: STRUCTURAL FEATURES OF GLYCOL-SPLIT LOW-MOLECULAR WEIGHT HEPARINS

*Most of the present results have been published in:*

- *Alekseeva A., Casu B., Cassinelli G., Guerrini M., Torri G., Naggi A. Structural features of glycol-split low-molecular-weight heparins and their heparin lyase generated fragments. Analytical and Bioanalytical Chemistry. 406 (2014) 249-265.*
- *Alekseeva A., Elli S., Cosentino C., Torri G., Naggi A. Susceptibility of enoxaparin reducing end amino sugars to periodate oxidation. Carbohydrate Research. 400 (2014) 33-43.*

As mentioned in the Introduction, gs-derivatives are currently being developed mainly starting from LMWHs (*Zhou* et al, 2011; *Leitgeb* et al., 2011; *Ritchie* et al, 2011) due to their better bioavailability, however, their structural features have not been reported yet. Additional microheterogeneity of the starting material induced by specific depolymerization reactions makes the analysis of gs-LMWHs even more complex than for gs-heparins. The present chapter describes the structural features of gs-derivatives prepared from three commercially available LMWHs, tinzaparin, dalteparin and enoxaparin, revealed by the approach developed and discussed in Chapter III.

Based on results reported in section III.5, their gs-derivatives were prepared using mild acidic conditions (diluted acetic acid at the last step of preparation, with pH adjustment after borohydride reduction) for minimizing hydrolytic side reactions. The results on the molecular weight distribution, obtained by SEC-TDA analysis, show that mild conditions do prevent significant hydrolysis, even if slight chain shortening was still observed (Table 4.1). The hydrodynamic radius $R_h$ (Table 4.1) and Mark-Houwink *a* and *logK* parameters (Annex 3) were also determined. As for unfractionated heparins, the presence of the gs-residues favors the random coil formation. This trend is more evident for longer chain oligosaccharides.

**Table 4.1.** Average molecular weight ($M_w$), polydispersity (D), hydrodynamic radius ($R_h$) and fractions weights for fractions with molecular weight < 2 kDa and >8 kDa, obtained for parent LMWHs and their glycol-split derivatives by SEC-TDA

| Sample | | $M_w$, kDa | D | $R_h$*, nm | Fraction weight, % mass | |
|---|---|---|---|---|---|---|
| | | | | | < 2 kDa | > 8 kDa |
| Parent LMWHs | Tinzaparin | 7.9 | 1.5 | 2.2 | 0.4 | 24 |
| | Dalteparin | 7.5 | 1.2 | 2.3 | nd | 18 |
| | Enoxaparin | 5.2 | 1.4 | 1.8 | 18 | 14 |
| | | | | | | |
| Glycol-split LMWHs | gs-Tinzaparin | 7.2 | 1.4 | 2.0 | 7 | 22 |
| | gs-Dalteparin | 6.4 | 1.2 | 2.2 | – * | – * |
| | gs-Enoxaparin | 4.7 | 1.3 | 1.8 | 20 | 9 |

$M_w$ and D were measured using 0.1M $NaNO_3$ as eluent to avoid aggregation
$R_h$ and fraction weight (% mass) were determined using 0.01M $NaNO_3$ as eluent
* Fraction weight for gs-dalteparin fractions with molecular weight <2 kDa and > 8 kDa were not determined because of formation of aggregates, whose response disturbs the chromatographic profile

### IV.1. Two-dimensional HSQC NMR analysis of gs-LMWHs

As in the case of gs-UFHs, for LMWHs (blue spots) and gs-LMWHs (red spots), signals for typical residues are located in subregions of the spectrum (Fig.4.1). The most intense signals are those of the TSD sequences for both starting materials and gs-products. Cross peaks of non-sulfated I/G and those associated with these residues (amino sugars preceding them) disappeared, while new ones corresponding to gs units appeared in the spectrum (Fig.4.1) at positions similar to those obtained for gs-UFHs (Table 3.2).

Table 4.2 reports the relative content of typical monosaccharide residues for the three prepared gs-LMWH samples measured from the signal volumes as described in Chapter III. For comparative purposes also the data for the parent LMWHs are reported. As for gs-heparins, differences in relaxation times may lead to the overestimation of some gs-residues, so that the currently obtained data should be taken in consideration as preliminary.

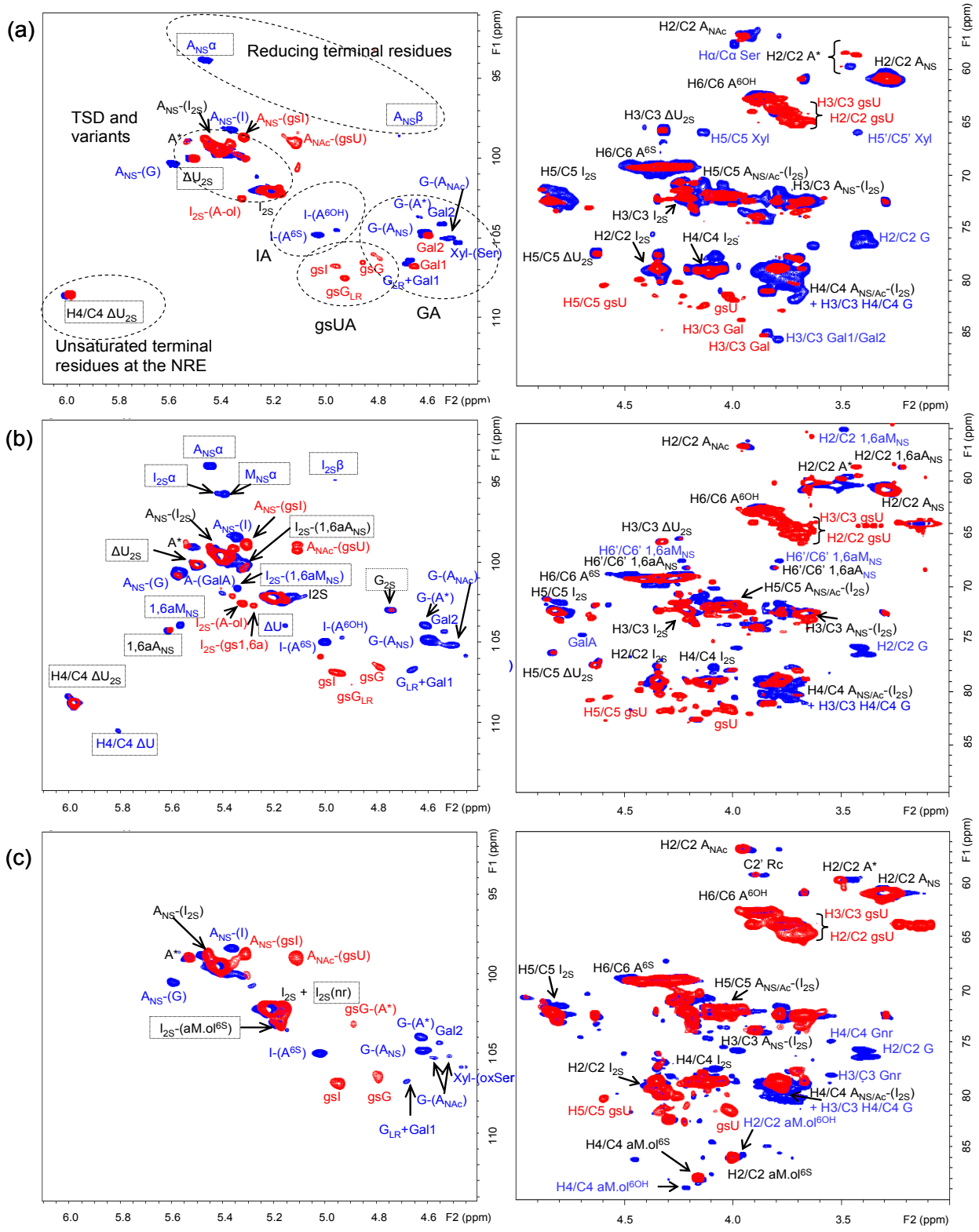



**Fig.4.1.** Anomeric (left panel) and "ring" (right panel) regions of the $^{1}H/^{13}C$ HSQC NMR spectra of gs-tinzaparin (a), gs-enoxaparin (b) and gs-dalteparin (c) (red spots) superimposed on the spectra of the corresponding unmodified LMWH (blue spots)

Symbols for typical end-group signals are framed with rectangles. A – glucosamine, A* – glucosamine-*N*,3-*O*,6-*O*-trisulfate, $M_{NS}$ – mannosamine-*N*-sulfate, $1,6A_{NS}/1,6aM_{NS}$ – 1,6-anhydro-glucosamine-*N*-sulfate and 1,6-anhydro-mannosamine-*N*-sulfate, aM.ol – 2,5-anhydro-mannitol, $\Delta U_{(2S)}$ – 4,5-unsaturated (2-*O*-sulfated) uronic acid, I/G – iduronic and glucuronic acids, $G_{LR}$ – glucuronic acid of the LR sequence, GalA – galacturonic acid, gsI/gsG – gs iduronic and glucuronic acids, gsU – gs uronic acids, Gal – galactose, Xyl – xylose, Ser – serine, nr – non-reducing end, Rc – ring contracted residues (Fig.1.17)

**Table 4.2.** Relative content (molar ratios) of the major individual residues of LMWHs and the corresponding gs-LMWHs

| | Tinzaparin | gs-Tinzaparin | Enoxaparin | gs-Enoxaparin | Dalteparin | gs-Dalteparin |
|---|---|---|---|---|---|---|
| **REPEAT SEQUENCE** | | | | | | |
| **$A_{NS}$(total)**[a] | 88.2 | 86.8 | 91.5 | 89.6 | 91.9 | 93.9 |
| **$A_{NAc}$(total)**[a] | 11.8 | 13.2 | 9.5 | 10.4 | 8.1 | 6.1 |
| **A*** [b] | 2.0 | 1.7 | 3.3 | 4.2 | 4.2 | 6.3 |
| **$I_{2S}$** | 60.0 | 67.5 | 53.3 | 53.2 | 76.8[c] | 79.8 |
| **I** | 8.3 | <LOD | 6.1 | <LOD | 7.3 | <LOD |
| **G** | 14.7 | <LOD | 13.6 | <LOD | 10.3 | <LOD |
| **G-(A*)** [b] | 1.7 | <LOD | 3.1 | <LOD | 4.1 | <LOD |
| **gsI** | – | 7.7 | – | 9.9 | – | 9.2 |
| **gsG** | – | 8.7 | – | 10.1 | – | 7.7 |
| **gsG-(A*)** [b] | – | <LOD | – | 1.6 | – | 3.3 |
| **NON-REDUCING END GROUPS** | | | | | | |
| **$\Delta U_{2S}$** | 14.9 | 11.8 | 16.4 | 18.2 | – | – |
| **ΔU** | – | – | 1.3 | <LOD | – | – |
| **REDUCING END GROUPS** | | | | | | |
| **Ared** | 12.3 | <LOD | 11.6 | <LOD | <LOD | <LOD |
| **1,6aA** | – | – | 4.1 | 2.5 | – | – |
| **aM.ol**[d] | – | – | – | – | 12.7 | 7.3 |
| **LINKAGE REGION (LMWHs)** | | | | | | |
| **$G_{LR}$+Gal1** | 3.0 | <LOD | 2.6 | <LOD | 1.9 | <LOD |
| **Gal2** | 1.6 | <LOD | 1.3 | <LOD | 1.3 | <LOD |
| **Xyl-(Ser)** | 1.6 | <LOD | <LOD | <LOD | <LOD | <LOD |
| **Xyl-(oxSer)** | <LOD | <LOD | 0.8 | <LOD | 0.9 | <LOD |
| **LINKAGE REGION (gs-LMWHs)** | | | | | | |
| **$gsG_{LR}$** | – | 2.6 | – | 1.6 | – | 2.0 |
| **Gal1** | – | 2.2 | – | 1.2 | – | 1.1 |
| **Gal2** | – | 2.1 | – | 0.9 | – | 1.1 |
| **gsXyl** | – | 1.0 | – | 1.9 | – | 2.1 |

LOD – limit of detection; "–" – for residues atypical for the type of LMWH

[a] calculated from the cross peaks of H2/C2 $A_{NS}$, H2/C2 $A_{NAc}$, H2/C2 $A_{NS3S6S}$ and for enoxaparin H2/C2 1,6a$A_{NS}$/1,6a$M_{NS}$

[b] A* = $A_{NS3S(6S)}$

[c] including $I_{2S}$-(aM.ol) and I(nr) residues

[d] calculated as sum of the cross peaks (C-2 aM.ol6S + C-4 aM.ol6S + C-2aM.ol6OH + C-4aM.ol 6OH) divided 2*(total amines)

As in the spectra of the parent LMWHs, the $gsG_{LR}$ signal intensity indicates that the content of LR in gs-tinzaparin is higher than in gs-enoxaparin, while almost absent in gs-dalteparin (Fig.4.1c, Table 4.2). The integration of the signals of the gsLR sequence should be taken in account with a caution, because longer relaxation times $T_2$ for terminal sequences may lead to a notable overestimation. In fact, the quantitative values obtained for the gsLR residues (especially, $gsG_{LR}$ and gsXyl, Table 4.2) seem to be overestimated.

Notably, cross-peak G-(A*) at 4.62/103.9 ppm, taken as a marker of the ATBR (*Bisio* et, 2009), is common to the three unmodified LMWHs but disappears upon glycol-splitting. Signal gsG-(A*) (at 4.90/103.2 ppm) is clearly observed only in the spectrum of gs-dalteparin (Fig.4.1c), but can be also detected in the expanded scale spectrum of the other gs-LMWHs. The signal is at the position similar to gsG-(A*) in the gs-derivative of the synthetic pentasaccharide $A_{NS6S}$-gsG-$A_{NS3S6S}$-$I_{2S}$-$A_{NS6S}$ (4.93/103.5 ppm) (Table 3.2). The assignment has been confirmed by TOCSY experiments, indicating a correlation of the signal at 4.90/103.2 ppm with the primary alcohol group of gsU residues (Fig. 4.2). Different contents of the gsG-(A*) reflect different contents of ATBR sequence in the original LMWHs Table 4.2.

Notably, ***end-groups*** are evident in the NMR spectra of all the LMWHs (Fig.4.1). Chemical shifts of the terminal groups typical for each LMWHs are listed in Annex 4. The NMR spectrum of dalteparin shows its relative structural homogeneity, with $I_{2S}$ at the NRE and aM.ol at the RE, this latter residue well-identifiable in the “ring region” of the NMR spectrum (Fig.4.1c). Enoxaparin appears the most heterogeneous due to the variety of RE amino sugars (Fig.4.1b).

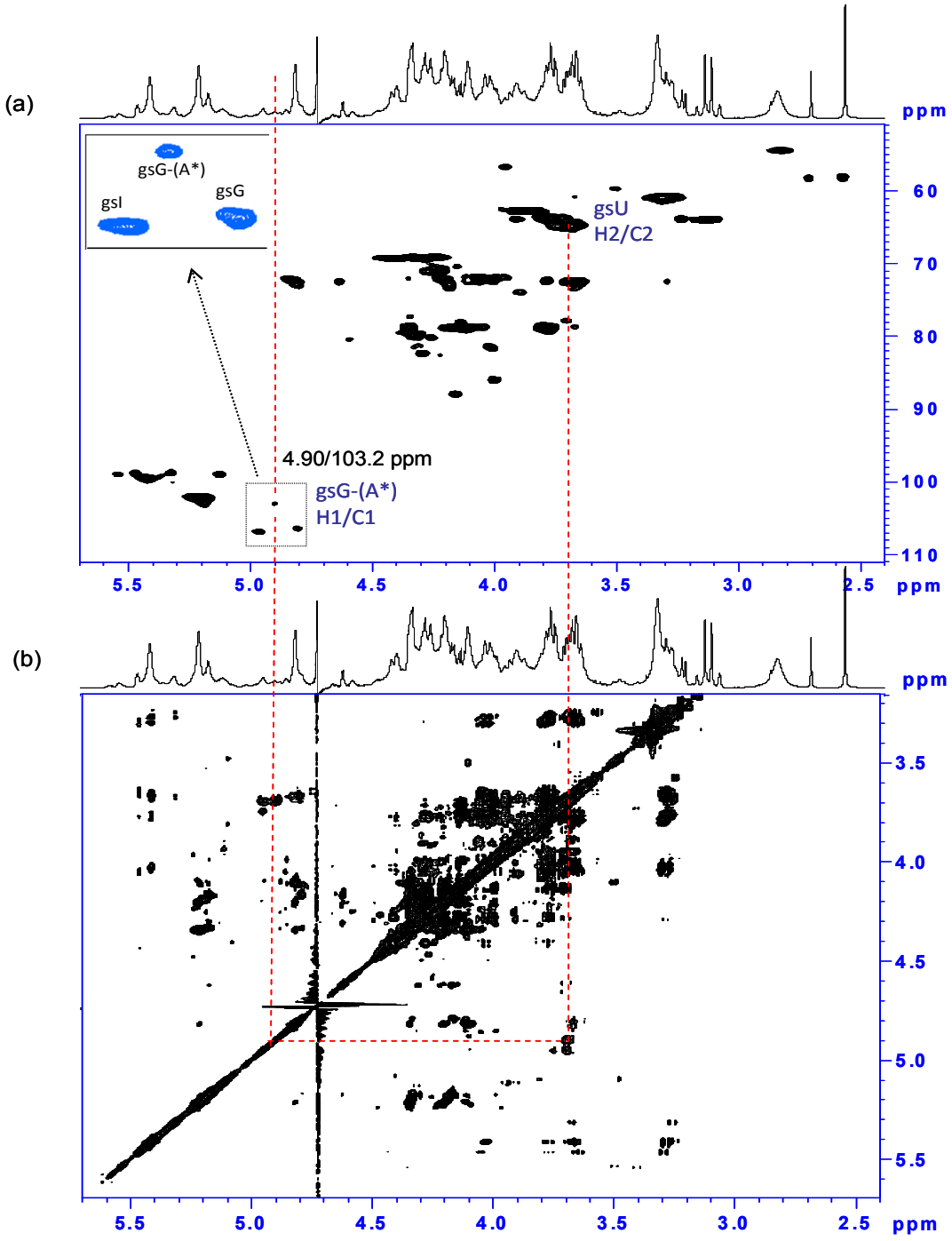



Fig. 4.2. HSQC (a) and 2D-TOCSY (b) spectra of gs-dalteparin

Correlation of H1/C1 and H2/C2 signals of gsG-(A*) are shown by red line

NMR data (Fig.4.1) show that not only internal sequences but also several end-groups are susceptible to glycol-splitting and/or reduction. The enoxaparin minor cross-peaks of ΔU at the NRE are among those that disappear upon periodate oxidation/borohydride reduction because of the presence of vicinal OH groups. The cross-peaks of both α and β anomers of aminosugars and uronic acids at the RE typical for unmodified tinzaparin ($A_{NS}\alpha$, $A_{NS}\beta$, Fig.4.1a) and

enoxaparin ($A_{NS}\alpha$, $A_{NS}\beta$ $M_{NS}\alpha$, $M_{NS}\beta$, $I_{2S}\alpha$, $I_{2S}\beta$, *see* Fig.4.1b) disappeared after the gs-reaction. The resulting RE-alditols (A-ol) are not detectable in the complex "ring region" of the NMR spectra due to signal overlapping with the free primary hydroxyl groups at C-6 of glucosamine and those at C-2 and C-3 of gs-uronic acids (Fig.4.1). However, $I_{2S}$ followed by A-ol can be detected in both gs-tinzparin and gs-enoxaparin (Fig.4.1, 4.3). Its chemical nature was shown by HMBC NMR spectroscopy, where the long-range correlation of the anomeric H1/C1 (5.34/103.1 ppm) signal and H5/C5 was clearly observed (Fig. 4.3). Its absence in gs-dalteparin, with its resistant RE aM.ol residues, independently confirms this interpretation. The direct proof was obtained by borohydride reduction of the enoxaparin octasaccharide fraction: along with disappearance of the RE glucosamine cross-peaks ($A_{NS}\alpha$, $A_{NS}\beta$, $M_{NS}\alpha$), a new signal at 5.34/103.1 ppm appeared in the spectrum (Fig.4.3b).

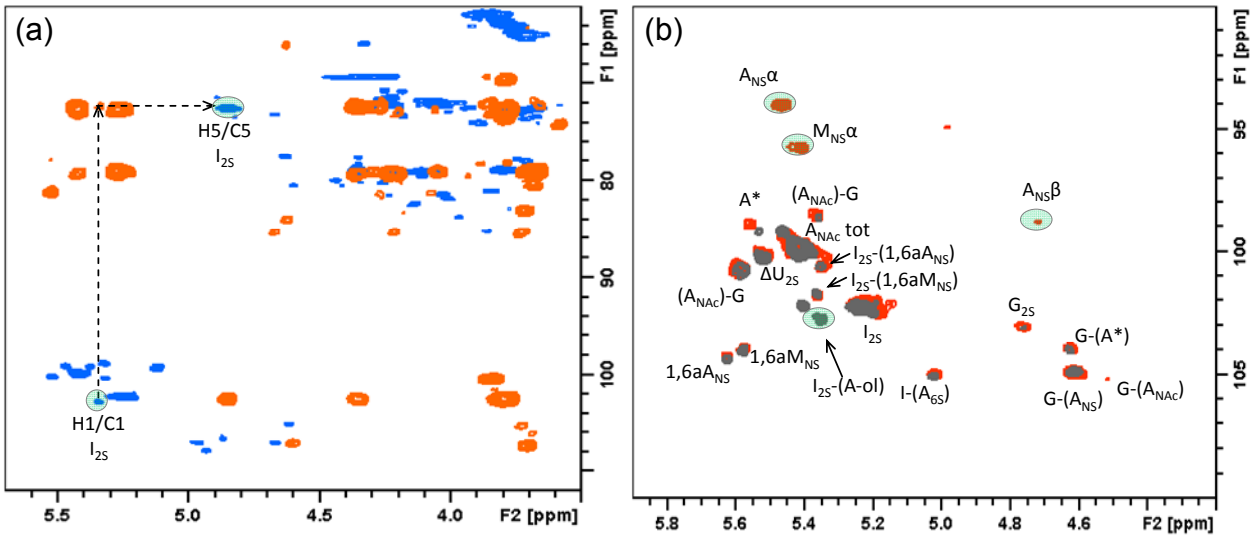



**Fig.4.3.** (a) Superimposed HSQC (blue spots) and HMBC (orange spots) spectra of gs-tinzaparin, (b) superimposed HSQC spectra of enoxaparin octasaccharide fraction before (red spots) and after borohydride reduction (grey spots)
Correlated signals H1/C1 and H5/C5 of $I_{2S}$ in the HMDC spectrum (a) as well as the anomeric HSQC signals susceptible to reduction and a new signal of $I_{2S}$-(A-ol) (b) are circled.

Unexpectedly, also the enoxaparin cross-peaks related to $1,6aM_{NS}$ disappeared upon periodate oxidation/borohydride reduction (Fig.4.1b). This anomalous behavior of $1,6aM_{NS}$ has been studied independently and is discussed in section IV.4.2 of the present chapter.

## IV.2. LC-MS profiling of heparinase-digested gs-LMWHs

*Mass spectra interpretation*

LC-MS analysis was carried out using the gradient and MS parameters optimized for gs-heparin digests (Chapter III). The identification of oligosaccharide components was performed using the exact mass values and characteristic isotope patterns. Enoxaparin 1,6-anhydro- (1,6aA) oligosaccharides can be detected by the characteristic mass loss of 18 Da with respect to the regular oligosaccharides. Dalteparin aM.ol end-group can be easily detected by the difference of 15 Da with respect to the unmodified natural oligosaccharides. Minor dalteparin oligosaccharides containing both aM.ol and Rc residues (Fig.1.17) were detected due to a mass decrease of 15 Da in comparison with the oligosaccharides bearing only aM.ol residues. The chemical nature and provenience of the sample (such as depolymerisation process) were taken into consideration to confirm the identified structures.

The **abbreviation system** was the same as that we used for gs-oligosaccharides generated by gs-heparins digestion and includes the number of monosaccharide residues, sulfate groups, *N*-acetyl groups, and gs residues when present (for example, ΔU4,5,0,1gs). Symbols ΔU, U and A indicate a 4,5-unsaturated uronic acid, a saturated one and a glucosamine unit, respectively, at the NRE. "Remnants" of the hydrolyzed internal gs residues were indicated by the symbol R (Fig.1.19e), and those at NRE as R(Unr). Symbol "ol" indicates the alditol form of terminal units at the RE; symbols aM.ol and Rc indicate 2,5-anhydromannitol and internal contracted ring residues typical for dalteparin. Symbol 1,6aA was used for both epimers (1,6-anhydro-glucosamine and 1,6-anhydro-mannosamine) because it was not possible to distinguish them by MS data. As two particular cases, gs1,6a and $R_{am}$ were used to indicate glycol-split 1,6-anhydro-amino sugar and remnant of an amino sugar. Their generation is discussed in section IV.4.2.

An example of the interpretation of mass spectrum of a peak found in gs-enoxaparin digest is shown in Fig.4.4.

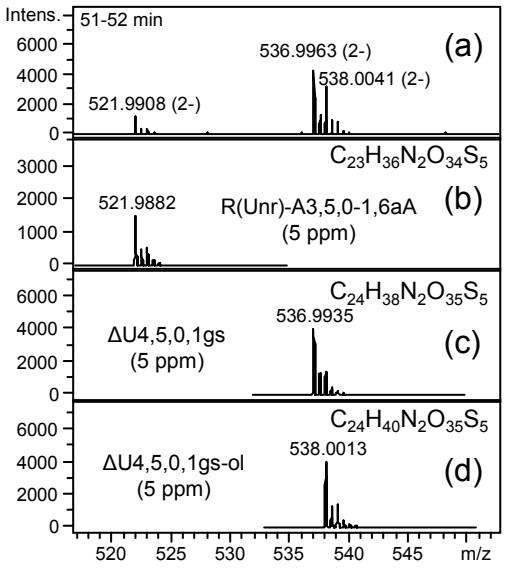



**Fig.4.4.** Experimental MS data (a) and the simulated ones for three overlapping oligosaccharides found in the gs-enoxaparin digest (b-d)

A first-level structure was assigned to each signal. The assignment relies on the exact mass value and isotope pattern. The error values between experimental and *m/z* calculated for the proposed structure are reported in ppm.

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. Symbol 1,6aA indicates the terminal 1,6-anhydro-amino sugars in gs-enoxaparin

Sodium borohydride reduction was used to distinguish between internal gs-unit and alditol end residue. Species susceptible to borohydride reduction (i.e., those converted to alditol forms) were identified as oligosaccharides originally terminating with hexosamines and, consequently, bearing internal gs, while the reduction resistant oligosaccharides were considered to exist already in the alditol form. For example, the double-charged ion with *m/z* 497.02 could correspond to both ΔU4,4,0,1gs and ΔU4,4,0-ol. After sodium borohydride reduction two peaks were observed: a) *m/z* 498.02, indicating that this oligosaccharide contains one internal gs-unit (gsU) and corresponds to the structure $\Delta U_{2S}$-$A_{NS(6S)}$-gsU-$A_{NS(6S)}$, where only one of the two sulfate groups within parenthesis is present; b) unmodified *m/z* 497.02, which should correspond to $\Delta U_{2S}$-$A_{NS}$-$U_{2S}$-$A_{NS}$-ol structure, having an internal 2-*O*-sulfated uronic acid ($U_{2S}$).

This approach was used also to distinguish two isomers ΔU4,6,0-ol (Fig.4.5) and ΔU4,6,0,1gs. The first tetrasaccharide represents the highly sulfated sequences at the RE of the original tinzaparin/enoxaparin chains, while the minor gs-tetrasaccharide is likely to be one of the markers of the *N*-sulfated ATBR (**$I_{2S}$-$A_{NS6S}$**-G-$A_{NS3S6S}$-$I_{2S}$-$A_{NS6S}$), observed in the enzymatic digest of gs-heparin from bovine lung (see Chapter III).

The additional $NaBH_4$ reduction of the gs-enoxaparin digest confirmed also the structure assigned to the minor tetrasaccharides terminating with an uronic acid remnant (ΔU4,4,1,1gs-R, *m/z* 577.03 for $[M-2H]^{2-}$ ion form, and ΔU4,5,0,1gs-R, *m/z* 596.01 for $[M-2H]^{2-}$ ion form), for which the *m/z* values remained constant after the additional reduction.

All the oligosaccharides in hemiacetal forms with internal gs-units had the same retention times as those observed in gs-heparin digests (Chapter III), confirming that they do not bear an alditol group (alditol formation induces a slight increase of the retention time, see Discussion).

*Structural features of gs-LMWHs by analysis of heparinase-generated gs-oligosaccharides*

Like those of similarly digested gs-heparins, the LC-MS profiles of heparinase-digested gs-LMWHs essentially consist of disaccharides, **gs-tetrasaccharide** and **gs-hexasaccharide** fragments and very low amounts of some **gs-octasaccharides** (Fig.4.5).

The major disaccharide ΔU2,3,0 ($\Delta U_{2S}$-$A_{NS6S}$), the minor disaccharide ΔU2,2,0 ($\Delta U_{2S}$-$A_{NS}$) and several tetrasaccharides resistant to the enzymatic cleavage (such as, ΔU4,6,0-ol in gs-tinzaparin, ΔU4,5,0-1,6aA in gs-enoxaparin, ΔU4,5,0-aM.ol in gs-dalteparin) are also present in the corresponding digests of LMWHs (*unpublished observation of our group*). Notably, no disaccharide in alditol form was detected in the gs-tinzaparin digest. The resistance to heparinases of the sequences with LMWH unnatural residues was previously observed for dalteparin and enoxaparin (*Linhardt* et al., 1990; *Ozug* et al., 2012; *Viskov* and *Mourier*, 2004).

LC-MS profile of gs-enoxaparin enzymatic digest (Fig.4.5b) is definitely more heterogeneous, especially as regards end residues, than those of gs-tinzaparin and gs-dalteparin digests, making structural assignments more difficult. LC-MS chromatogram of the heparinase-digest of gs-dalteparin (Fig.4.5c) is less complex than those of the other two gs-

LMWH digests, reflecting in part a rather homogeneous internal structure of the original LMWH and the fact that practically all chains in the original dalteparin terminate at the RE with aM.ol groups, resistant to glycol-splitting. Notably, the gs-dalteparin digests contain also octasaccharides ΔU8,9,1,1gs-aM.ol and ΔU8,9,1,2gs-aM.ol (both in traces), indicating that all the glycosidic bonds of these oligosaccharides are resistant to cleavage by heparinases (*see* Discussion). The heparinase-resistant aM.ol-containing hexasaccharides with the ring contracted residue (Fig.1.17), ΔU6,6,1,Rc-aM.ol and ΔU6,7,0,Rc-aM.ol (Fig.4.5) was observed at the first time in the present study.

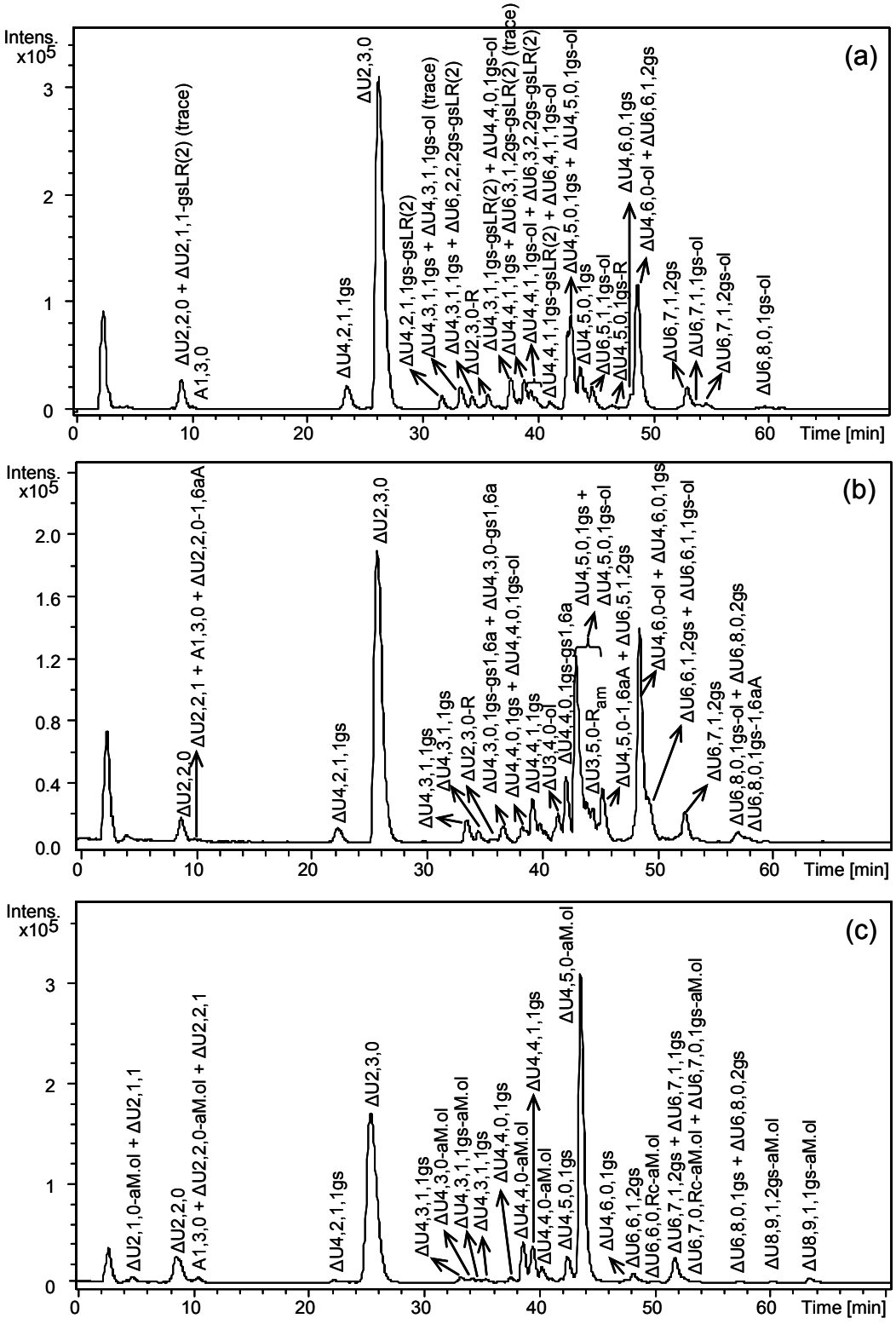



**Fig.4.5.** LC-MS chromatogram of the heparinase-digests of gs-tinzaparin (a), gs-enoxaparin (b), gs-dalteparin (c)

LC-MS conditions: see Fig.3.12

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. Symbols 1,6aA and aM.ol indicate the terminal 1,6-anhydro-amino sugars in gs-enoxaparin and 2,5-anhydro-mannitol in gs-dalteparin, respectively. Rc indicates internal ring-contracted structures in gs-dalteparin.

Among gs-hexasaccharides, ΔU6,7,1,2gs, marker of gs-ATBR, is present in all the samples but in different amounts, correlating with NMR data of A* and G-(A*) contents in the original LMWHs (Table 4.2). Several octasaccharides with three gs uronic acids in contiguous disaccharide units were found in gs-LMWHs in trace amounts. This indicates that sequences with three subsequent non-sulfated uronic acids are present within the original LMWH chains. During the study of the gs-enoxaparin structure, a fraction containing highly sulfated mono-*N*-acetylated octasaccharides with three gs units (ΔU8,7,1,3gs, ΔU8,8,1,3gs, ΔU8,9,1,3gs) was isolated from the gs-enoxaparin digest by SEC fractionation for detailed study (Fig.4.6). In order to distinguish between octasaccharides with two internal gs-residues and reduced terminal glucosamine (alditol form), having the same molecular mass, borohydride reduction of the total octasaccharide fraction was used. The species susceptible to borohydride were those with three gs-residues and hemiacetalic glucosamine at the RE, while unaffected oligosaccharides should bear already reduced glucosamine. The presence of both types of octasaccharides (those with three gs-units ΔU8,7,1,3gs, ΔU8,8,1,3gs, ΔU8,9,1,3gs (Fig.4.6) and those with two gs-units and a terminal alditol ΔU8,7,1,2gs-ol, ΔU8,8,1,2gs-ol, ΔU8,9,1,2gs-ol, ΔU8,10,1,2gs-ol) was revealed. Except for ΔU8,10,1,2gs-ol, alditol forms were found in trace (not shown in Fig.4.6). The cross-peak of the α-anomer of *N*-sulfated glucosamine ($A_{NS}\alpha$), observed in the 2D NMR spectrum of the digest before reductive treatment, also confirms the presence of octasaccharides with three gs units and a hemiacetalic *N*-sulfated glucosamine at the RE. It is worth noting that in the case of the octasaccharides with two gs residues their resistance to the enzymatic cleavage may be explained by the position of the 2-*O*-sulfated uronic acid ($U_S$) that is likely to be close to the alditol unit in the proposed structure $\Delta U_S$-A-gsU-A-gsU-A-**$U_S$-A-ol** (the resistance of the sequences with unnatural residues is noted above).

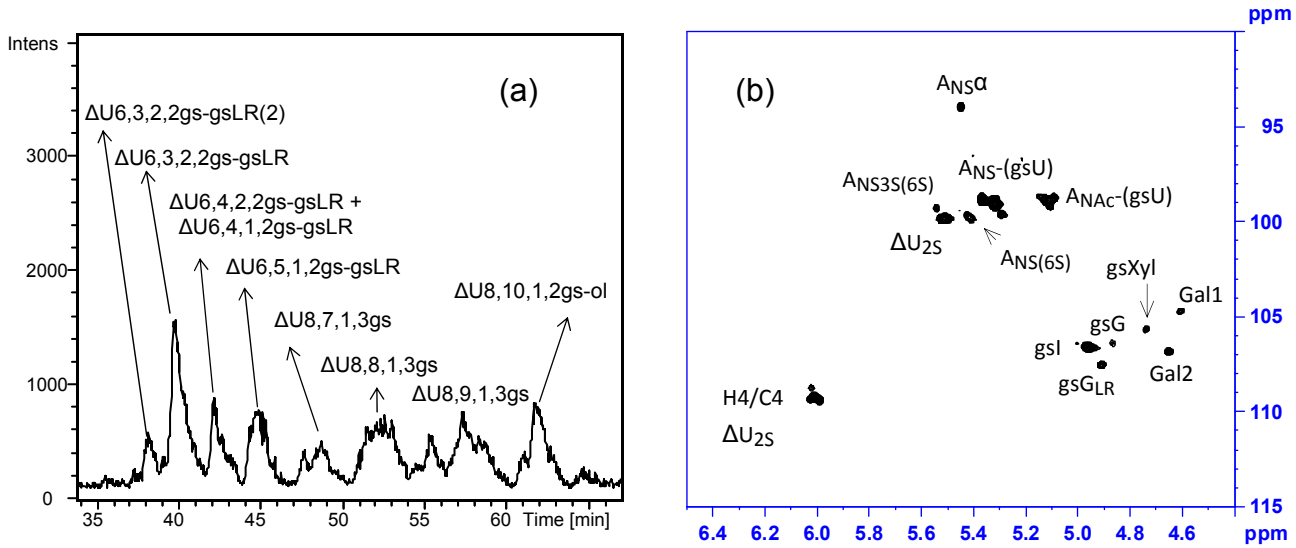



**Fig. 4.6.** LC-MS chromatogram (a) and HSQC NMR spectrum (b) of octasaccharide fraction of gs-enoxaparin digest obtained by SEC fractionation

LC-MS conditions: see Fig.3.12
Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. gsLR – gsG-$(Gal)_2$-gsXyl-$Ser_{ox}$, gsLR(2) – gsG-$(Gal)_2$-$CH(CH_2OH)_2$
The signal at 4.75/105.9 ppm (panel b) was attributed to gsXyl residue. It was the only additional signal to those assigned for gs-heparins in HSQC spectra (see Table 3.2), and its amount was found much higher in the fraction enriched by gsXyl-bearing oligosaccharides, shown in this figure (panel a). Moreover, the ratio of the intensities of $G_{LR}$, Gal1, Gal2 and gsXyl in this fraction was 1:1:1:1.

Longer sequences with more than three gs units were not detected, even after fractionation by SEC. The same results were obtained also for the sequences in the closest proximity to the LR: this undersulfated region, followed by gsLR, was found to vary from disaccharide to octasaccharide with three subsequent non-sulfated uronic acids. These data are reported in detail in section IV.4.1 and in the discussion of the present chapter.

### IV.3. Direct LC-MS analysis of gs-LMWHs

While LC-MS analysis of the digested heparin-related samples allows also the detection of minor components, such as the end-residues in LMWHs, they do not permit to determine the primary composition of the oligosaccharide mixtures. Unlike UFHs, which have molecular weights too high to be directly analyzed by LC-MS, profiling of the intact LMWHs is possible

due to their lower molecular weight (*Langeslay* et al., 2013). LC-MS analyses of gs-LMWHs in comparison with the starting LMW material were expected to provide additional information about individual components of tinzaparin, enoxaparin and dalteparin.

### IV.3.1. Method development

*Gradient optimization*

For profiling gs-LMWHs, the methanol gradient used for the analysis of the enzymatic digests did not permit to achieve simultaneously good resolution and detection of longer oligosaccharides. An attempt to improve the sensitivity by using *n*-pentyl- or *n*-hexyl-amines (*Doneanu* et al., 2009) was done, but it led to the decrease of the isomer resolution. Substitution of methanol in the DBA-containing mobile phases by acetonitrile permitted to improve chromatographic performance in lower total analysis time by decreasing column backpressure and consequent increase of the flow-rate (from 100 up to 250 µl/min). Acetonitrile-based eluents allowed detecting oligosaccharides up to dp 20 in enoxaparin and up to dp 22 in tinzaparin and in dalteparin. Comparable results were reported using pentylamine acetate as ion-pair agent for the UPLC separation of tinzaparin oligosaccharides (*Langeslay* et al., 2013) and in an earlier SEC-MS study of Henriksen (2004) oligosaccharides up to dp 18 were detected in tinzaparin. The optimized gradient conditions are reported in the experimental part (Chapter VI) and the obtained compositional LC-MS profiles are discussed below (section IV.3.2), while the complete chromatograms are reported in Annexes 5-7.

*Mass spectra interpretation*

For each peak a "first-level structure" was assigned using MS data and considering the chemical nature and mode of depolymerization used for LMWH preparation. The mass spectra interpretation was based on the same concepts applied for the digested gs-heparins and gs-LMWHs. High accuracy and resolution of MS permit to accurately measure *m/z* and to obtain

characteristic isotope patterns (Fig.4.7-4.8). Their formula was only accepted if the *m/z* of the candidate analyte matched with the theoretical value within 5 ppm for oligosaccharides up to dp 8. For higher chain length oligosaccharides an error within 10 ppm was accepted.

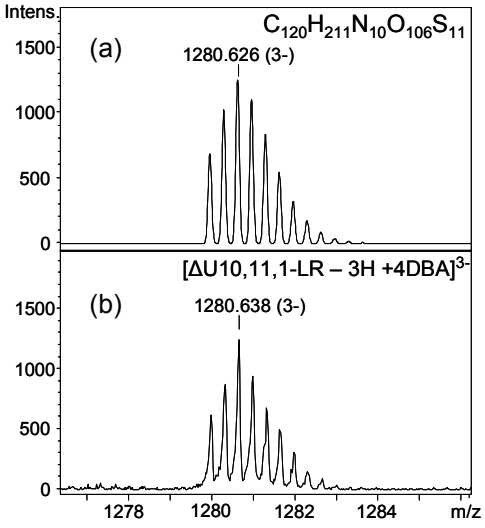



**Fig.4.7.** Example of MS interpretation using isotope pattern and exact mass: (a) simulated pattern for [ΔU10,11,1-LR - 3H + 4DBA]$^{3-}$ ion form, (b) experimental one

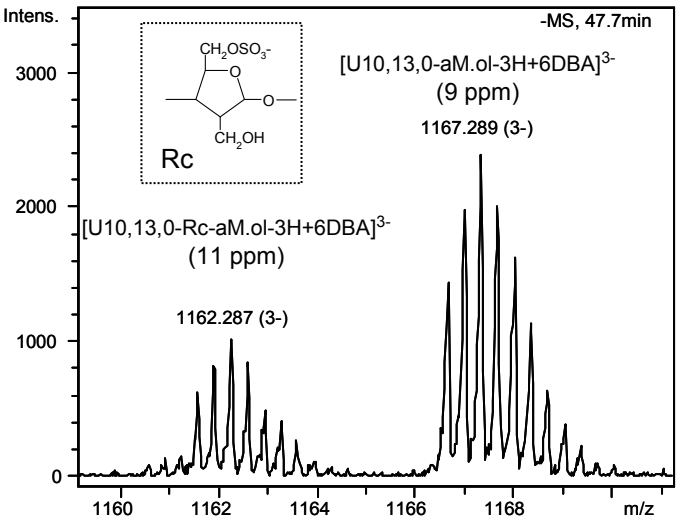



**Fig.4.8.** Mass spectra for aM.ol-oligosaccharides, regular chain decasaccharide in [U10,13,0-aM.ol - 3H + 6DBA]$^{3-}$ ion form and another one containing internal Rc-unit [U10,13,0-Rc-aM.ol -3H+ 6DBA]$^{3-}$. The errors between experimental *m/z* and calculated one are also shown

For gs-oligosaccharides, a 2 Da mass increase is expected for each internal gs residue with respect to the corresponding unmodified oligosaccharides, however, the susceptibility of the RE residues to reduction has to be taken in account for the structural assignment of oligosaccharides in gs-LMWHs. The "first" increment of 2 Da ("2H") was assigned to the reduced RE residue (either glucosamine or uronic acid), and a further increase of the mass value of 2 or 4 Da was assigned to oligosaccharides (in the alditol form) with 1 or 2 gs residues, respectively. Because the dalteparin RE residues are exclusively constituted by aM.ol, which is not susceptible to glycol-splitting nor to borohydride reduction, the gs-oligosaccharides present in gs-dalteparin are easier to be assigned on the MS basis, because each increment of 2 Da should indicate only the presence of one gs-residue.

The abbreviation system was the same that for digested gs-LMWHs samples (*see* IV.2).

*Evaluation of the method sensitivity to compositional differences*

As for UFHs, the method was tested in terms of its "sensitivity" to hydrolytic side reactions. Tinzaparin was chosen due to its relative simplicity among the commercially available LMWHs and its being the most similar to heparin. We prepared gs-tinzaparin using both hydrochloric and acetic acid at the last reaction step, and the differences in the hydrolysis degree were confirmed by 2D NMR, LC-MS and SEC-TDA (Table 4.3). LC-MS chromatograms under the optimized separation conditions of both gs-tinzaparin samples differ significantly (Fig.4.9): the sample prepared with 0,1M HCl acid solution appeared to contain a number of remnant (R) bearing species. Having shown that the present LC-MS method permits to observe and monitor such a hydrolysis process, it was applied for the profiling of LMWHs and their gs-derivatives.

**Table 4.3.** Parameters differing between gs-tinzaparin prepared by different methods ($M_w$, uronic acid content and the relative content of hydrolysis marker ΔU2,3,0-R)

| | **SEC-TDA** | | **2D HSQC NMR (%)** | | | | | | **LC/MS** |
|---|---|---|---|---|---|---|---|---|---|
| | $M_w$, kDa | | $\Delta U_{2S}$ | $I_{2S}$ | gsI | gsG | $A_{NS/NAC}$-(R) | | Area ratio (%) ΔU2,3,0-R/ΔU2,3,0 |
| gs-tinzaparin (a) | 6.4 | | 13.2 | 73.5 | 2.2 | 5.9 | + | | 3.9 |
| gs-tinzaparin (b) | 7.2 | | 11.8 | 67.5 | 7.7 | 8.7 | trace | | < LOD |

(a) obtained by adding a 0,1M HCl solution at the step of pH adjustment
(b) obtained by adding a 5% (v/v) $CH_3COOH$ solution at the step of pH adjustment

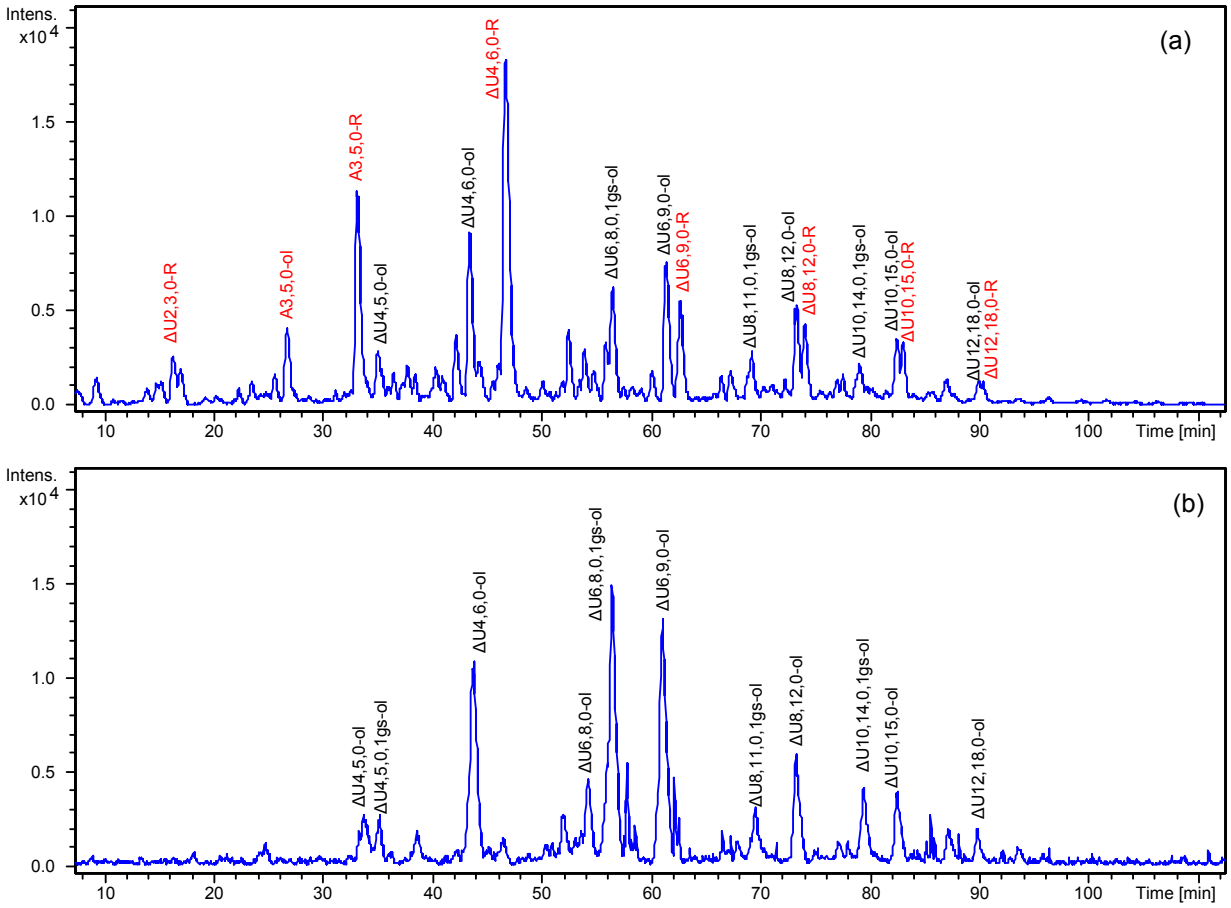



**Fig.4.9.** Total LC-MS chromatograms of two gs-tinzaparin preparations: (a) using 0.1M HCl or (b) 5% (v/v) $CH_3COOH$ for pH netralization

Only major components are shown.

LC conditions: column C18 100 x 2.1 mm, 2.6 µm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; gradient: 0 min – 5%B, 5 min – 5%B, 125 min – 35%B, 130 min – 70%B, 135 min – 70%B, 138 min – 5%B, 160 min – 5%B; flow rate – 0.25 ml/min

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form, while R indicates remnant of uronic acids (hydrolysis marker)

### IV.3.2. LC-MS profiling of gs-LMWHs

To obtain compositional profiles of the gs-LMWHs, their samples were analyzed directly by the LC-MS method in comparison with those of the original unmodified LMWHs. Although a detailed analysis of unmodified LMWHs is outside the scope of this work, some general features of their LC-MS chromatograms are described here for comparison with the corresponding gs-derivatives (Fig.4.10-4.12).

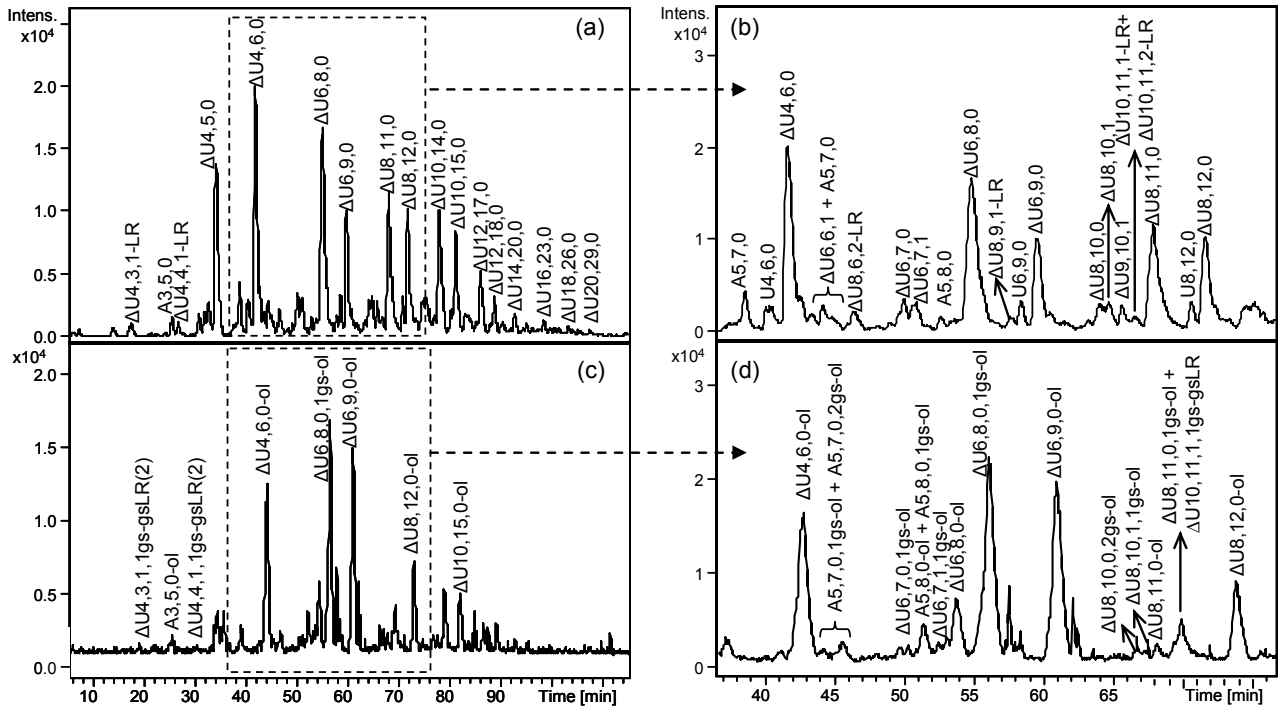



**Fig.4.10.** LC-MS profiles of tinzaparin (a,b) and gs-tinzaparin (c,d)

LC-MS conditions: see Fig.4.9.

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. LR – linkage region, gsLR – gsG-$(Gal)_2$-gsXyl-$Ser_{ox}$, gsLR(2) – gsG-$(Gal)_2$-$CH(CH_2OH)_2$

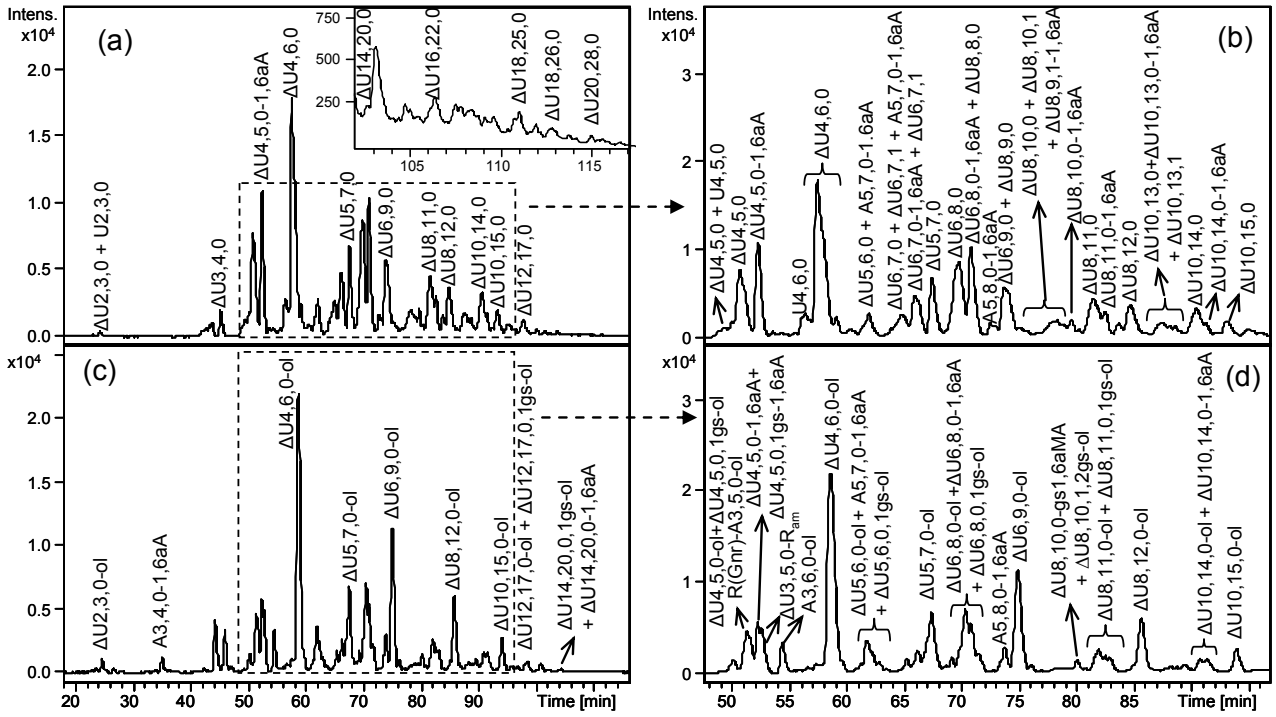



**Fig.4.11.** LC-MS profiles of enoxaparin (a,b) and gs-enoxaparin (c,d)

LC conditions: column C18 100 x 2.1 mm, 2.6 μm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; gradient: 0 min – 0%B, 5 min – 0%B, 130 min – 35%B, 140 min – 70%B, 145 min – 70%B, 148 min – 0%B, 170 min – 0%B; flow rate – 0.25 ml/min

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. 1,6aA indicates both 1,6-anhydro-glucosamine and 1,6-anhydro-mannosamine.

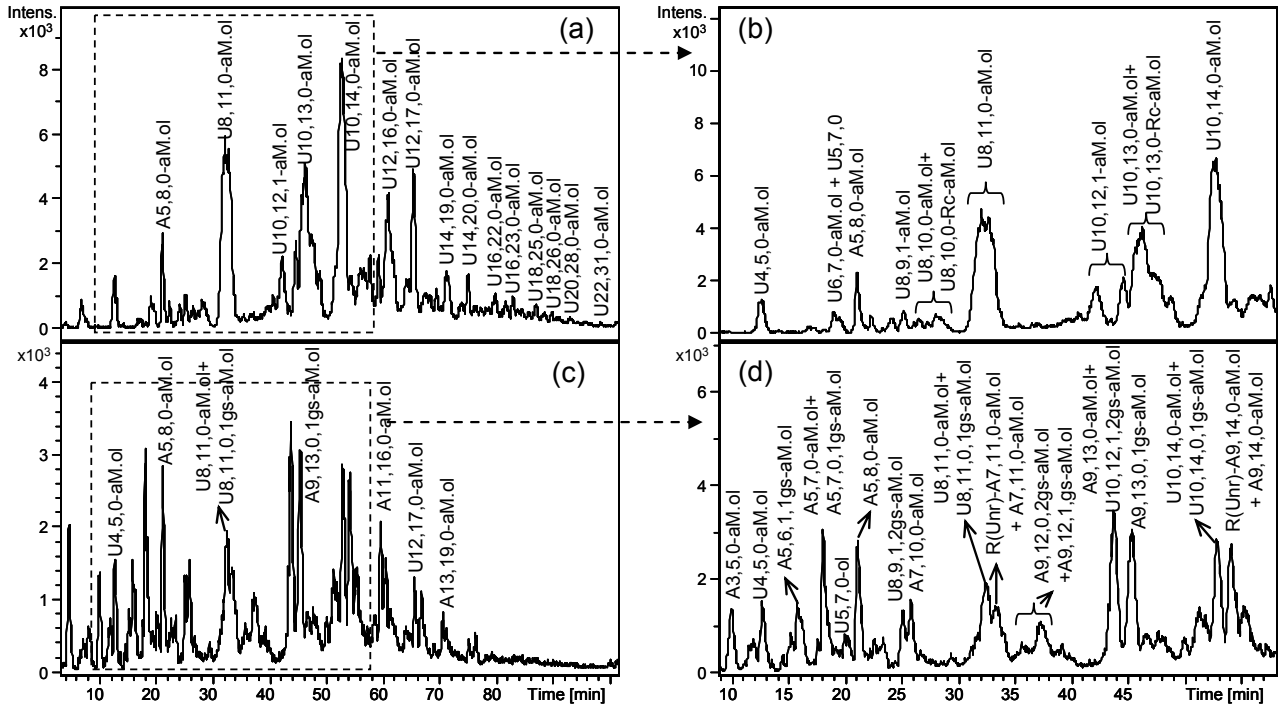



**Fig.4.12.** LC-MS profiles of dalteparin (a,b) and gs-dalteparin (c,d)

LC conditions: column C18 100 x 2.1 mm,.2.6 μm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; gradient: 0 min – 10%B, 4 min – 10%B, 15 min – 18%B, 30 min – 18%B, 115 min – 35%B, 133 min – 70%B, 138 min – 70%B, 140 min – 10%B, 160 min – 10%B; flow rate – 0.25 ml/min

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. Symbols aM.ol and Rc indicate 2,5-anhydro-mannitol and ring-contracted units.

The present LC-MS method highlights both the structural heterogeneity of the analyzed samples and their different molecular weight distributions. For example, while the LC-MS chromatogram of enoxaparin appears highly heterogeneous (Fig.4.11), that of dalteparin (Fig.4.12), a LMWH known to be extensively purified by ion-exchange chromatography (*Bisio* et al, 2009), reflects a lower heterogeneity and a lower content of dp 4 and dp 6 fractions with respect to tinzaparin (Fig.4.10) and enoxaparin (Fig.4.11). The highly sulfated even-numbered oligosaccharides are prevalent in all three LMWHs. *N*-acetyl groups were observed more frequently within longer chains (Annexes 5-7).

Peculiar residues typical for each LMWH are clearly observed in their LC-MS profiles. In tinzaparin and enoxaparin LC-MS chromatograms the major components absorb at 232 nm

due to the presence of 4,5-unsaturated uronic acid residues ($\Delta U_{2S}/\Delta U$) at the NRE (Fig.1.15), formed as a result of the depolymerisation by $\beta$-elimination reaction. The enoxaparin LC-MS profile appears more complex due to the presence of oligosaccharides bearing at the RE either hemiacetalic or 1,6-anhydro (1,6aA) forms (Fig.4.11). Some dalteparin oligosaccharides with an internal contracted ring (Rc) residues were detected for the first time (Fig.4.12, Annex 8).

Among the minor components, oligosaccharides with an odd number of residues and a glucosamine at the NRE are detectable in all three LMWHs (Fig.4.10-4.12), suggesting that some of the original heparin chains terminate with glucosamine units, as discussed in Chapter III. Odd-numbered oligosaccharides starting and terminating with uronic acid residues are minor components of tinzaparin (such as ΔU9,10,1, Fig.4.10) and in dalteparin (U5,6,0 and U5,7,0, Fig.4.12), but are well evident in the enoxaparin (ΔU5,6,0 and ΔU5,7,0, Fig.4.11). These RE residues are likely to be present in original heparin chains, but their content in LMWHs may be increased during the depolymerisation process as a result of a peeling reaction (*Viskov and Mourier*, 2005).

LC-MS profiles of gs-derivatives significantly differ from their parent LMWHs (Fig.4.10-4.12). LC-MS analysis shows that, together with internal gs-residues, glycol-splitting-induced modifications involve several end-groups, confirming the NMR data and LC-MS analysis of the digested gs-LMWHs. The major modifications occurring in the end-groups of LMWHs upon periodate/borohydride treatment are summarized in Fig.4.13.

The most evident effect of the reduction step is the formation of **alditol residues at the RE** (involving a 2 Da mass increase) for tinzaparin and enoxaparin oligosaccharides originally bearing amino sugars in hemiacetalic form (Fig.4.13a). The RE amino sugars may be also oxidized (Fig.4.13a,b), as reported in IV.4.2.

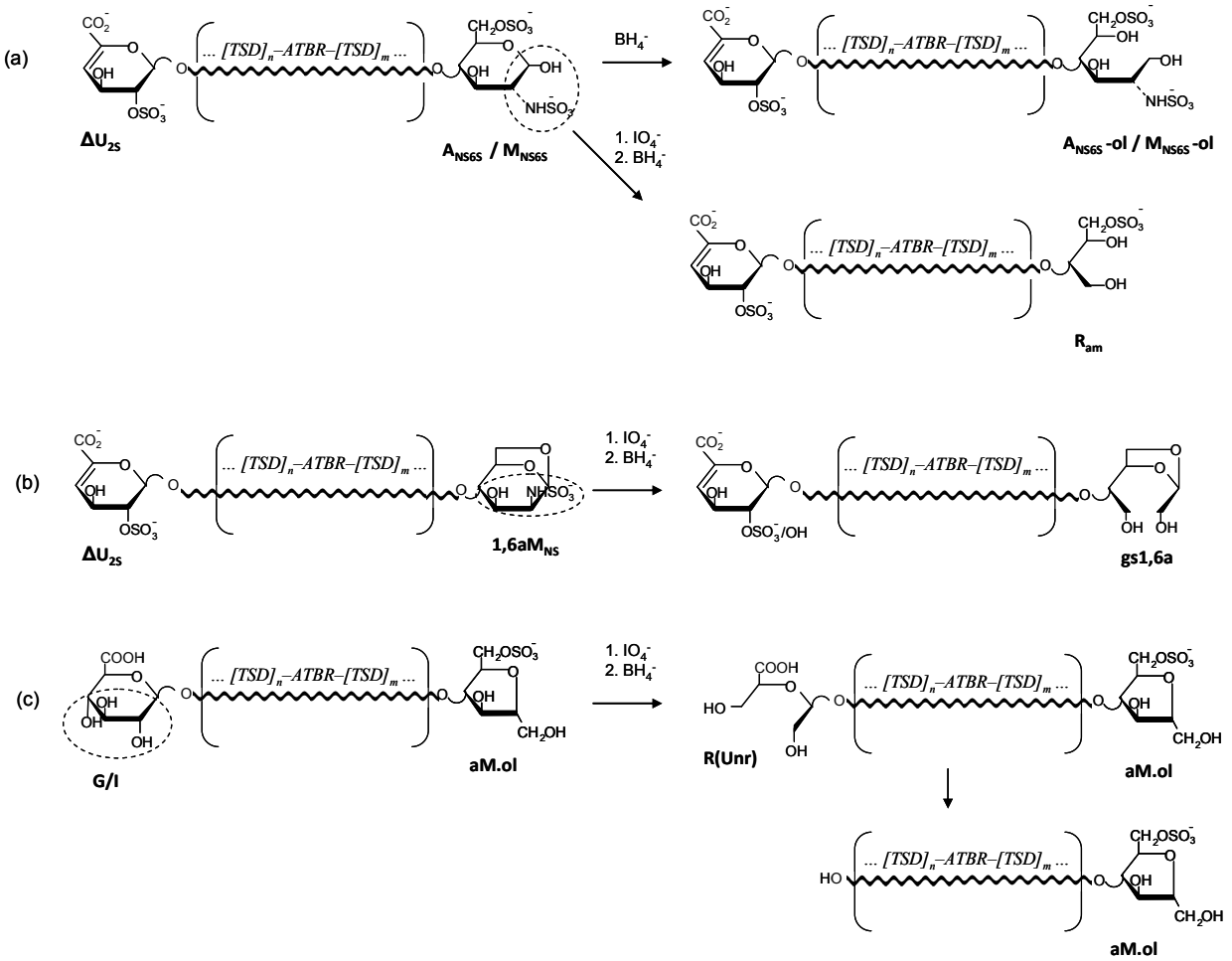



**Fig.4.13.** Major modifications of end groups of tinzaparin (a), enoxaparin (a-c) and dalteparin (c), induced by periodate/borohydride treatment
TSD – trisulfated disaccharide, ATBR – antithrombin binding region

On the other hand, LC-MS showed that several **NRE residues** of gs-dalteparin are susceptible to glycol-splitting and to a further chain cleavage (as illustrated in Fig.4.12). In the case of the non-sulfated uronic acid at the NRE, simultaneous oxidation of the three vicinal OH-groups at C-2, C-3 and C-4 is likely to occur. Further loss of the formed remnants at the NRE (R(Unr), Fig.4.13c) seems to easily occur. In fact, a relatively large number of gs-dalteparin odd-numbered oligosaccharides start with a glucosamine (Fig.4.12c,d). Hydrolytic loss of the remnants formed at the NRE were also observed to lower extents in enoxaparin, due to the low amount of oligosaccharides starting with glucuronic acid at the NRE. For example, R(Unr)-A3,5,0-ol and R(Unr)-A3,5,0-1,6aA (Fig.4.11, Annex 8) were observed in gs-enoxaparin and the last trisaccharide could be structurally correlated with a tetrasaccharide

G-$A_{NS3S6S}$-$U_{2S}$-1,6a$A_{NS}$, isolated from the enoxaparin dp 4 fraction (*unpublished data of our group*). A further remnant loss at the NRE is reflected, especially in gs-enoxaparin, by the presence of the trisaccharides A3,5,0-ol and A3,6,0-ol, this latter bearing an $A_{NS3S6S}$ residue.

For species starting with glucosamine at the NRE, splitting between C(3) and C(4) is also possible. On the other hand, it seems that A5,8,0-1,6aA in enoxaparin (Fig.4.11) and A5,8,0-aM.ol in dalteparin (Fig.4.12) are resistant to the oxidation because the corresponding analytes were found unchanged in the corresponding gs-derivative chromatograms. The experiment, performed also on the synthetic pentasaccharide $A_{NS6S}$-G-$A_{NS3S6S}$-$I_{2S}$-A-$OMe_{NS6S}$ (Arixtra®), showed that an excess of periodate should be used for completing the splitting of both internal glucuronic acid and glucosamine at the NRE (*Ferro M.,* Master degree thesis, 2009). The C(3)-C(4) bond of the NRE saturated $I_{2S}$ was also found to be resistant to periodate oxidation, at least under the applied conditions. It is likely that the prevalent *a-/a-trans*-position of the corresponding OH groups in $^1C_4$ conformation of $I_{2S}$ (Fig.1.8), shown by *Ferro* et al., 1990, does not permit the formation of cyclic intermediate with periodate ion.

The possibility to obtain more structural information by differentiating isomers through gs-modification is also worth noting. For example, after glycol-splitting the enoxaparin tetrasaccharide ΔU4,5,0 gives rise to at least two tetrasaccharides with a terminal alditol residue, one containing internal 2-*O*-sulfated uronic acid ΔU4,5,0-ol and another one containing a gs uronic acid (ΔU4,5,0,1gs-ol) (Fig.4.11). As particular cases, several gs-ATBR containing oligosaccharides were found (Table 4.4). The corresponding unmodified oligosaccharides were further found in the HA fraction of enoxaparin, independently confirming the presence of ATBR within their sequences (*unpublished data of our group*).

**Table 4.4.** LC-MS data (*m/z* value of the prevalent ion form and retention times) of gs-oligosaccharides likely containing gs-ATBR

| Compound | Prevalent ion form | Monoisotopic *m/z* value | Retention time, min |
|---|---|---|---|
| **ΔU8,10,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+2DBA]^{3-}$ | 817.104 | 80.3 |
| **ΔU8,9,1,2gs-1,6aA** *(may contain gs-ATBR)* | $[M-3H+2DBA]^{3-}$ | 783.776 | 75.9 |
| **ΔU10,13,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+5DBA]^{3-}$ | 1138.579 | 89.6 |
| **ΔU12,15,1,2gs-1.6aA** *(may contain gs-ATBR)* | $[M-3H+8DBA]^{3-}$ | 1426.727 | 94.0 |
| **ΔU12,16,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+9DBA]^{3-}$ | 1503.105 | 95.7 |
| **ΔU14,19,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+12DBA]^{3-}$ | 1824.581 | 102.2 |

The formation of R(Unr)-fragments also permitted to distinguish isomers that was not possible to individually identify in unmodified LMWHs. For example, the broad U8,11,0-aM.ol peak in the dalteparin LC-MS chromatogram can be explained by the presence of several isomers. Various sequences may correspond for this number of sulfate groups: either a regular sequence constituted by repeating TSD units ($I_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-$I_{2S}$-$aM.ol_{6S}$) or another one with a G-A* unit. The formation of R(Unr)-A7,11,0-aM.ol in gs-dalteparin (Fig.4.12d) indicates the presence of the structure with the G-A* disaccharide unit at the NRE, which is well-separated from the regular octasaccharide U8,11,0-aM.ol (unaffected by glycol-splitting). This observation correlates with the NMR data indicating a relatively high contents of A*, G-(A*) and Gnr residues in the parent dalteparin (Table 4.2) and its octasaccharide fraction (A* and G-(A*): ~10% of total glucosamine uronic acid content, respectively; Gnr: about 50% of total glucuronic acid amount).

## IV.4. Glycol-splitting-induced modifications of peculiar end-groups (linkage region, 1,6-anhydro-D-mannosamine-*N*-sulfate, D-glucosamine-*N*-sulfate and D-mannosamine *N*-sulfate)

### IV.4.1. Glycol-splitting induced modifications of linkage region sequence

The LC-MS chromatogram of tinzaparin reveals the presence of highly sulfated species (ΔU4,3,1-LR, ΔU4,4,1-LR, ΔU8,6,2-LR, ΔU8,9,1-LR, ΔU10,11,1-LR, ΔU10,11,2-LR (Fig.4.10)), containing the intact **linkage region** $G_{LR}$-Gal-Gal-Xyl-Ser. As expected and confirmed by NMR (Chapter III), the $G_{LR}$ and Xyl units of the LR are susceptible to glycol-splitting. Analysis of gs-LMWHs after enzymatic digestion (Fig.4.5a) showed that, depending on the reaction conditions, the LR can generate two forms, such as the one shown in Fig.1.19d (gsLR) or further hydrolyzed at the level of gsXyl, to give the serine lacking fragment gsG-Gal-Gal-$CH(CH_2OH)_2$ (gsLR(2)).

The LC-MS analysis of the oligosaccharides containing the gsLR/gsLR(2) in the gs-tinzaparin digest allowed the characterization of the environment near the LR. Non-*N*-acetylated gsLR-oligosaccharides were not found, in accord with the knowledge that the first glucosamine is always *N*-acetylated (*Iacomini* et al., 1999). Of note, ΔU2,1,1-gsLR(2) is formed in trace amount, which indicates that the first uronic acid preceding $G_{LR}$ is almost always non-sulfated. Otherwise, it would have been split and then not recognized by the enzymes. A similar concept, that each product of the enzymatic digestion should contain a 2-*O*-sulfated uronic acid at its NRE, was applied for further oligosaccharide identification. Different mono-*N*-acetylated tetrasaccharides containing one gs residue linked to the gsLR(2) $\Delta U_{2S}$-$A_{NS(6S)}$-gsU-$A_{NAc(6S)}$-gsLR(2) were detected (Table 4.5).

**Table 4.5.** First-level structures and hypothesis of the sulfation pattern identified in tinzaparin and gs-tinzaparin, and the corresponding LC-MS data (monoisotopic *m/z* for the prevalent ion form and retention time (RT)) for several linkage region containing oligosaccharides

| Compound | Structure hypothesis * | Prevalent ion form | Monoisotopic *m/z* value | RT, min |
|---|---|---|---|---|
| TINZAPARIN | | | | |
| ΔU4,3,1-LR | $\Delta U_{2S}$-$A_{NS6S}$-G/I-$A_{NAc}$-LR | $[M-2H]^{2-}$ | 836.640 | 17.4 |
| ΔU4,4,1-LR | $\Delta U_{2S}$-$A_{NS6S}$-G/I-$A_{NAc6S}$-LR | $[M-2H]^{2-}$ | 876.619 | 26.5 |
| ΔU8,6,2-LR | $\Delta U_{2S}$-$A_{NS(6S)}$-$I_{2S}$-$A_{NS(6S)}$-G/I-$A_{NAc(6S)}$-G/I-$A_{NAc}$-LR | $[M-3H]^{3-}$ | 876.119 | 46.3 |
| ΔU8,9,1-LR | $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-G/I-$A_{NAc}$-LR | $[M-3H+2DBA]^{3-}$ | 1028.173 | 57.5 |
| ΔU10,11,1-LR | $\Delta U_{2S}$- $A_{NS(6S)}$-$I_{2S}$-$A_{NS(6S)}$-$I_{2S}$-$A_{NS(6S)}$-$I_{2S}$-$A_{NS(6S)}$-G/I-$A_{NAc}$-LR | $[M-3H+4DBA]^{3-}$ | 1279.946 | 66.3 |
| ΔU10,11,2-LR | $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$-G/I-$A_{NAc6S}$-G/I-$A_{NAc6S}$-LR | $[M-3H+4DBA]^{3-}$ | 1293.949 | 66.4 |
| GLYCOL-SPLIT TINZAPARIN | | | | |
| ΔU2,1,1-gsLR(2) *(trace)* | $\Delta U_{2S}$-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 516.622 | 8.5 |
| ΔU4,2,1,1gs-gsLR(2) | $\Delta U_{2S}$-$A_{NS}$-gsU-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 726.159 | 31.6 |
| ΔU4,3,1,1gs-gsLR(2) | $\Delta U_{2S}$-$A_{NS(6S)}$-gsU-$A_{NAc(6S)}$-gsLR(2)/gsLR | $[M-2H]^{2-}$ | 766.137 | 37.6 |
| ΔU4,3,1,1gs-gsLR | | $[M-2H]^{2-}$ | 825.150 | 39.8 |
| ΔU4,4,1,1gs-gsLR(2) | $\Delta U_{2S}$-$A_{NS6S}$-gsU-$A_{NAc6S}$-gsLR(2) | $[M-2H]^{2-}$ | 806.115 | 40.9 |
| ΔU4,1,2,1gs-gsLR(2) *(trace)* | $\Delta U_{2S}$-$A_{NAc}$-gsU-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 707.186 | 20.3 |
| ΔU6,3,1,2gs-gsLR(2) *(trace)* | $\Delta U_{2S}$-$A_{NS}$-gsU-$A_{NS}$-gsU-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 935.695 | 38.5 |
| ΔU6,2,2,2gs-gsLR(2) | $\Delta U_{2S}$-$A_{NS}$-gsU-$A_{NAc}$-gsU-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 916.722 | 35.2 |
| ΔU6,3,2,2gs-gsLR(2) | $\Delta U_{2S}$-$A_{NS(6S)}$-gsU-$A_{NAc(6S)}$-gsU-$A_{NAc(6S)}$-gsLR(2)/gsLR | $[M-2H]^{2-}$ | 956.700 | 39.7 |
| ΔU6,3,2,2gs-gsLR | | $[M-2H+1DBA]^{2-}$ | 1080.290 | 46.2 |
| ΔU6,4,2,2gs-gsLR *(trace)* | $\Delta U_{2S}$-$A_{NS6S}$-gsU-$A_{NAc6S}$-gsU-$A_{NAc}$-gsLR | $[M-2H+1DBA]^{2-}$ | 1120.268 | 49.5 |
| ΔU8,2,3,3gs-gsLR(2) *(trace)* | $\Delta U_{2S}$-$A_{NS(6S)}$-gsU-$A_{NAc(6S)}$-gsU-$A_{NAc(6S)}$-gsU-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 1107.286 | 37.3 |
| ΔU8,3,3,3gs-gsLR(2) *(trace)* | $\Delta U_{2S}$-$A_{NS(6S)}$-gsU-$A_{NAc(6S)}$-gsU-$A_{NAc(6S)}$-gsU-$A_{NAc}$-gsLR(2) | $[M-2H]^{2-}$ | 1147.264 | 41.3 |

The hypothesis of the structures are based on the fact that gs-residues are resistant to the heparinases, while 2-*O*-sulfated iduronic acid residues are susceptible to these enzymes, so that the first residue within the heparinase-generated oligosaccharides is always $\Delta U_{2S}$. Since glucosamines with unsubstituted $NH_2$-group were not found in the tinzaparin sample by NMR, all the glucosamines should be *N*-substituted in the proposed structure. Moreover, the first glucosamine in the closest proximity to the LR is almost always *N*-acetylated and rarely 6-*O*-sulfated, as shown in the present study and by *Iacomini* (1999).

Symbol (6S) indicates possible 6-*O*-sulfation.

The first level structure of ΔU4,3,1,1gs-gsLR(2) corresponds to the previously observed heparin oligosaccharides $\Delta U_{2S}$-$A_{NS6S}$-G/I-$A_{NAc}$-LR (*Iacomini* et al., 1999). Only one species containing two acetyl groups (ΔU4,1,2,1gs-gsLR(2)) was observed, but in very low amounts, indicating that, when the second uronic acid after $G_{LR}$ is 2-*O*-sulfated, the second glucosamine residue is almost always *N*-sulfated and not *N*-acetylated. Moreover, the presence of ΔU4,4,1,1gs-gsLR(2) indicates that the first *N*-acetyl-glucosamine can also be 6-*O*-sulfated, as previously found in unmodified heparin (*Iacomini* et al., 1999).

Among gsLR-bearing hexasaccharides, those incorporating two *N*-acetyl-glucosamines and two gs-residues (ΔU6,3,2,2gs-gsLR(2) and ΔU6,2,2,2gs-gsLR(2)) were the most evident, while a mono-*N*-acetylated analog (ΔU6,3,1,2gs-gsLR(2)) was observed in traces, indicating that when the second uronic acid after $G_{LR}$ is non-sulfated, the following glucosamine is *N*-acetylated. Notably, no species with three *N*-acetyl-glucosamines was found. Octasaccharide chains with three gs units linked to the gsLR(2) were also found, even if in traces (Table 4.5, see Discussion). Several gsLR-bearing gs-oligosaccharides with the sulfation/acetylation pattern similar to that observed in tinzaparin were also found in enoxaparin after SEC-fractionation of its digested gs-derivative (Fig.4.6). No oligosaccharides bearing 2-*O*-phosphorylated xylose (*Tone* et al., 2008) were observed.

### IV.4.2. Susceptibility to periodate of the enoxaparin reducing end amino sugars

Enoxaparin oligosaccharides bear various amino sugars at their reducing end (RE): glucosamine *N*-sulfate ($A_{NS(6S)}$, ~9% of total amino sugar content; *Guerrini* et al., 2007) and unnatural residues formed during depolymerisation of heparin under alkaline conditions (*Mourier*, and *Viskov*. US Pat. 2005/0119477 A1, 2005), namely mannosamine *N*-sulfate ($M_{NS(6S)}$), 1,6-anhydro-glucosamine *N*-sulfate ($1,6aA_{NS}$) and 1,6-anhydro-mannosamine *N*-sulfate ($1,6aM_{NS}$), respectively estimated ~3, 2 and 2.5% of the total amino sugar content;

*Guerrini* et al., 2007). These residues were expected to be resistant to periodate because of *N*-substitution. The oxidation by periodate of amino sugars having vicinal OH- and $NH_2$-groups is a well-known process leading to splitting of the corresponding C-C bond (*Dryhurst, 1970*). However, the presence of different *N*-substituents induces a variety of oxidative behaviors. For example, *N*-alkyl-amino alcohols are also susceptible to periodate oxidation but the reaction rate largely varies for secondary and tertiary amines (*Dryhurst,* 1970; *Perlin,* 2006). However, when the amino-group is *N*-sulfated the nucleophilic properties of nitrogen are expected to be diminished. Unlike RE $A_{NS}/M_{NS}$ groups, whose disappearance from the NMR spectra could be explained by their susceptibility to borohydride reduction, the disappearance of $1,6aM_{NS}$ (see Fig.4.1b) was unexpected.

Interestingly, LC-MS analysis of both intact gs-enoxaparin and its heparinase-digest indicated the generation of unknown structures by glycol-splitting. To better understand whether these structures are associated with the modification of $1,6aM_{NS}$ residues, the less complex enzymatically depolymerized sample was studied with particular attention. The LC-MS data show that the analytes with the unknown structures likely have the same backbone but they differ in the number of sulfate groups, and gs residues (Table 4.6). Odd-numbered mass values indicate the presence of only one nitrogen atom within these molecules.

The fact that all the three species absorb at 232 nm indicates the presence of a $\Delta U_{2S}$ at the NRE. Useful information was obtained also by a subsequent SEC-fractionation (Fig. 4.14a). LC-MS analysis showed that the SEC-fraction "F" eluted between the pools of lowly (fraction G in Fig.4.14) and highly sulfated (fraction E in Fig.4.14) tetrasaccharides is enriched in compounds of the unknown structures (*m/z* 448.5, 449.5 and 488.5, Fig.4.14), that is in agreement with the hypothesized neutral formula (tetrasaccharides with 3-4 sulfate groups, Table 4.6).

**Table 4.6.** Experimental data obtained by LC-UV-MS analysis and proposed neutral formula of some unknown compounds found in gs-enoxaparin after its enzymatic digestion

| EXPERIMENTAL DATA | | | |
|---|---|---|---|
| *m/z* experimental $[M-2H]^{2-}$ | 448.524 | 449.532 | 488.504 |
| Experimental isotope pattern (M)/(M+1)/(M+2) | *Overlapped with 449.532 (due to similar RT)* | 100/31/23 | 100/30/30 |
| Absorbance at 232 nm | + | + | + |
| *m/z* experimental $[M-2H]^{2-}$ after additional reduction with $NaBH_4$ | 448.524 | 449.533 | 488.504 |
| **HYPOTHESIS of STRUCTURE** | | | |
| Neutral molecular formula (tolerance <5 ppm) | $C_{24}H_{37}N_1O_{29}S_3$ | $C_{24}H_{39}N_1O_{29}S_3$ | $C_{24}H_{37}N_1O_{32}S_4$ |
| Structural connections between species | M | M + 2H (1gs) | M + $1SO_3$ |
| Monoisotopic neutral mass (theoretical) | 899.0613 | 901.0770 | 979.0182 |
| *m/z* theoretical $[M-2H]^{2-}$ | 448.523 | 449.531 | 488.502 |
| Theoretical isotope pattern (M)/(M+1)/(M+2) | 100/30/24 | 100/30/24 | 100/31/29 |

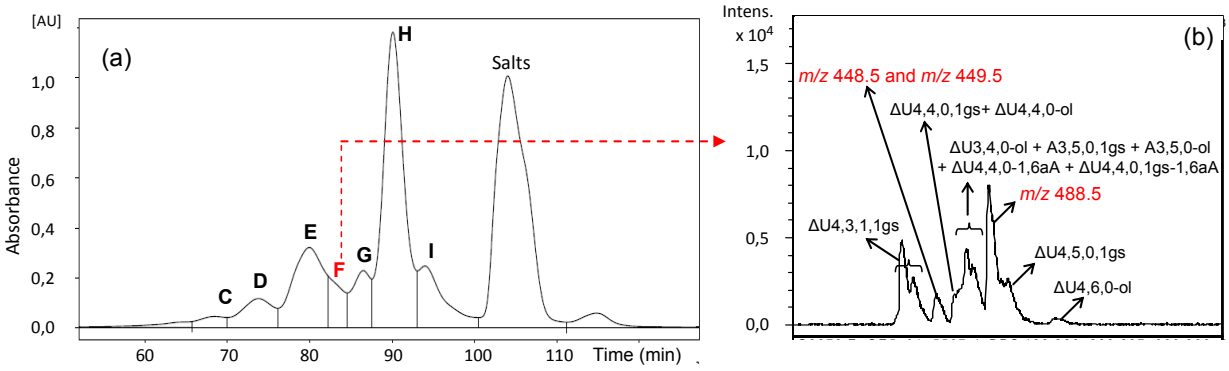


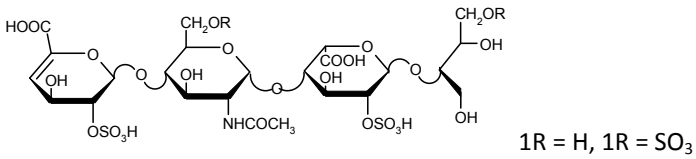
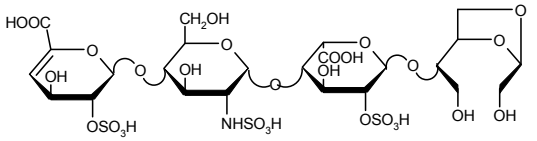
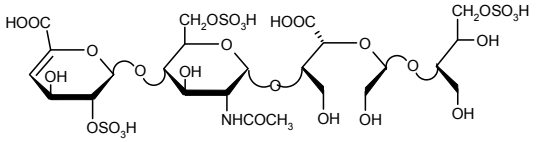
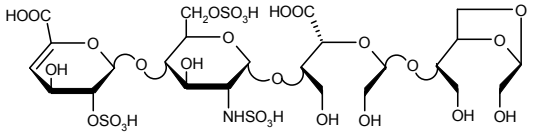
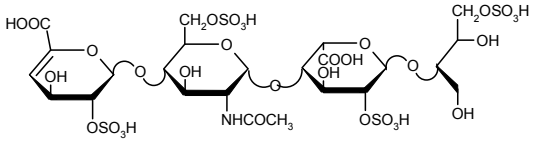
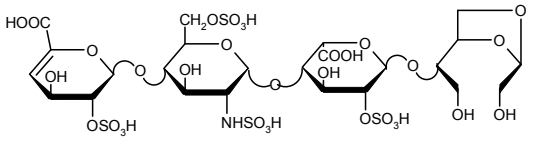

**Fig.4.14.** SEC profile (a) of gs-enoxaparin digested with heparinases and LC-MS chromatogram of the fraction enriched in unknown species with *m/z* 448.5, 449.5 and 488.5 (b)

Components of all the fractions were determined by LC-MS. Major components were: fraction I – ΔU2,2,0, fraction H – ΔU2,3,0, fraction G - ΔU2,3,0-R and ΔU4,2,1,1gs, fraction F – unknown species (see panel b), fraction E - ΔU4,5,0,1gs and ΔU4,6,0-ol, fraction D – ΔU6,6,1,2gs, fraction C – octasaccharides (see Fig.4.6)
LC-MS conditions and abbreviations used: see Fig.4.6.

These species were also found to be resistant to borohydride reduction, suggesting that they do not have a hemiacetalic monosaccharide at the RE. Their presence in the enzymatic digest and, therefore, their resistance to the heparinases suggests the presence of unnatural terminal residue(s) in their molecules.

Two different structures with the same neutral formula were found to correspond to the observed unusual *m/z* values: the first one an unsaturated *N*-acetylated trisaccharide containing the remnant of amino sugar ($R_{am}$) and the second structure bearing a split 1,6-anhydro-amino sugar at the RE (Table 4.7). A subsequent MS/MS experiment indicated that the main fragments could be formed from both the backbone types (Fig.4.15). A minor signal *m/z* 247,494 may be more informative: it corresponds to a disaccharide unit with two sulfate groups, suggesting that the unique glucosamine, present in the molecule, is *N*-sulfated, ruling out the structure with a remnant and a *N*-acetyl substituent.

**Table 4.7.** Possible structures of analytes with *m/z* 448.5, 449.5, 488.5 found in the heparinase-digest of gs-enoxaparin

| *m/z* | Abbreviation | Structure |
|---|---|---|
| **448.52** | ΔU3,3,1-$R_{am}$ |  1R = H, 1R = $SO_3$ |
| | ΔU4,3,0,1(gs1,6a) |  |
| **449.53** | ΔU3,3,1,1gs-$R_{am}$ |  |
| | ΔU4,3,0,1gs,1(gs1,6a) |  |
| **488.50** | ΔU3,4,1-$R_{am}$ |  |
| | ΔU4,4,0,1gs,1(gs1,6a) |  |

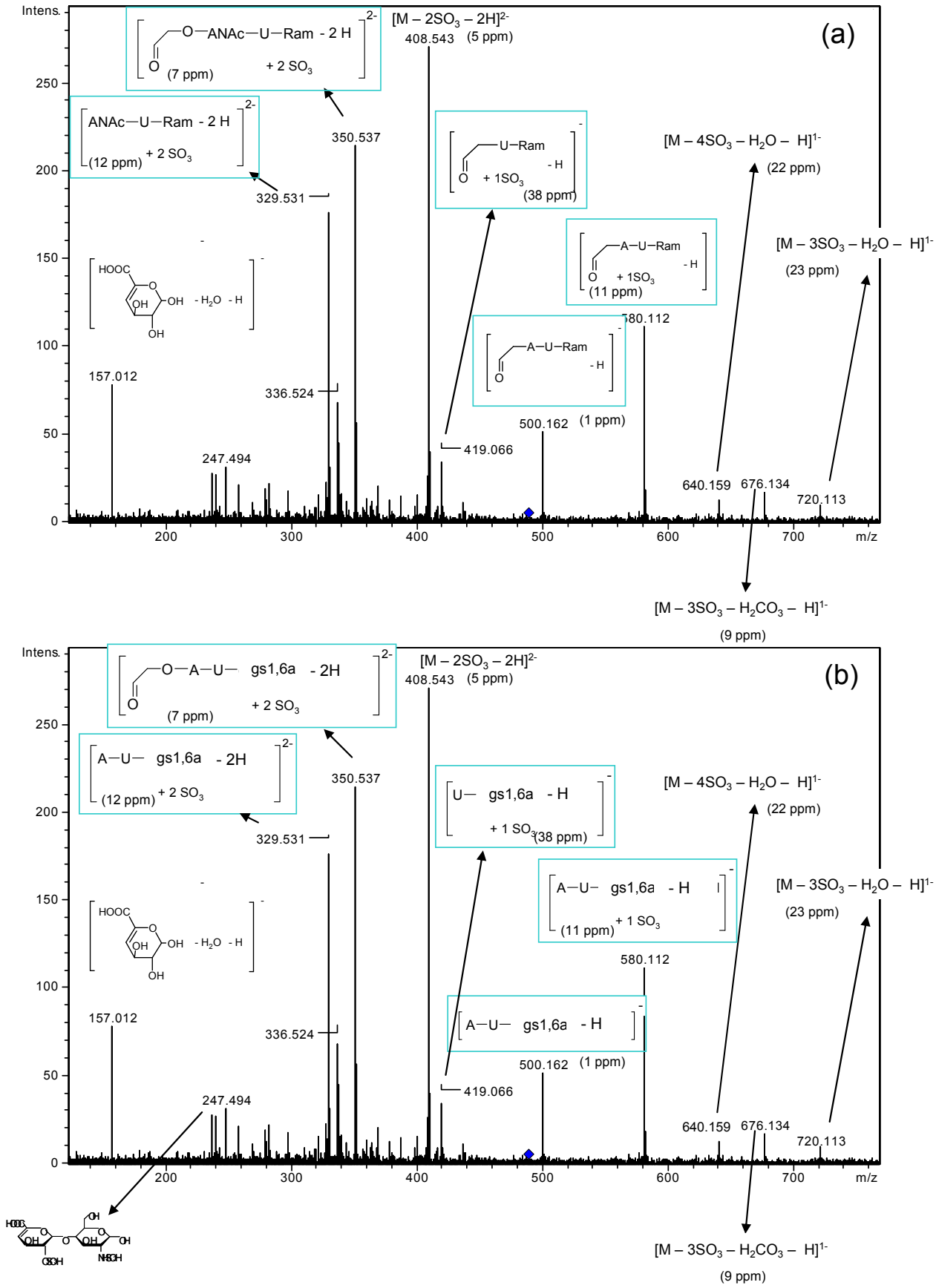



**Fig.4.15.** MS/MS fragmentation of *m/z* 488.5 with two different interpretations (a) trisaccharide bearing a remnant of the oxidized mannosamine ($R_{am}$), (b) tetrasaccharide with the split 1,6-anhydro-aminosugar residue (gs1,6a)

NMR data also suggest that the structure with split 1,6-anhydro-ring is more probable. The HSQC spectrum of the SEC fraction enriched in the structurally unknown compounds did not show cross-peaks corresponding to the $A_{NAc6S}$ linked to $I_{2S}$ (H1/C1 5.15/96.6 ppm, H2/C2 4,03/56.2 ppm, *Yates* et al, 1996). This cross-peak should have been observed if the unknown compound with *m/z* 488.5 had the *N*-acetylated structure $\Delta U_{2S}$-$A_{NAc6S}$-$I_{2S}$-$R_{am}$. Moreover, the $A_{NAc6S}$-($I_{2S}$) sequence is rarely present within heparin chains according to its biosynthesis pathway.

All the experimental data, i.e.: i) absorbance at 232 nm; ii) resistance under the reductive conditions; iii) *m/z* value; iv) correspondence of the experimental isotopic pattern with the theoretic one; v) resistance to the reduction and to the enzymatic cleavage; and vi) NMR data, indicate that the split 1,6-anhydro-ring sugar is the most probable structure. However, it was still difficult to prove that its formation is caused by oxidation of the $1,6aM_{NS}$.

To verify whether periodate can split terminal $1,6aM_{NS}$ residues of some enoxaparin oligosaccharides, model compounds were needed. Since no oligosaccharide standards containing 1,6-anhydro-sugars are available, we performed a multi-step fractionation of enoxaparin components to isolate a fraction containing two isomeric pentasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$1,6aA_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$1,6aM_{NS}$ (Fig.4.16).

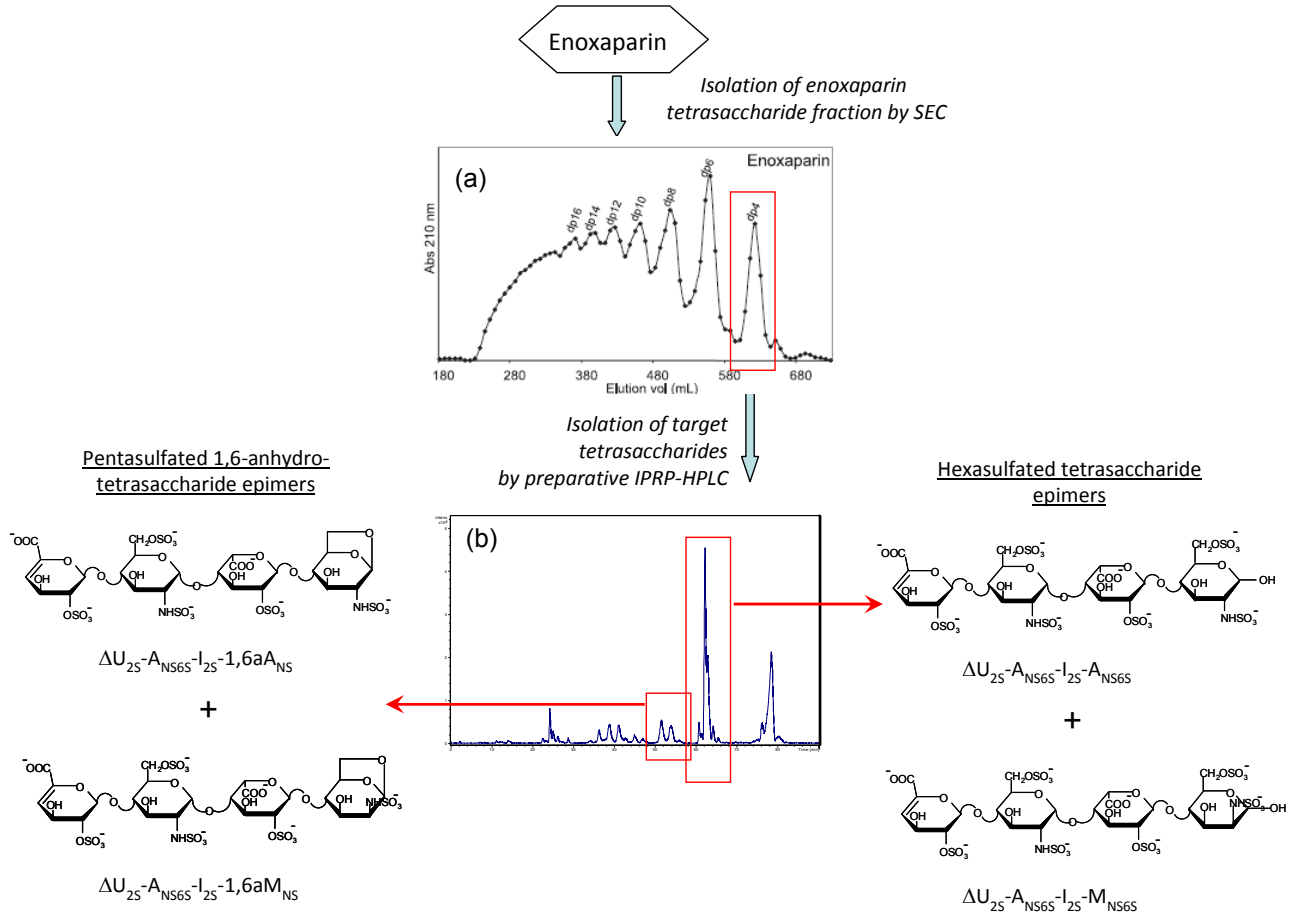



**Fig.4.16.** Multi-step isolation scheme of two epimeric tetrasaccharide pairs: pentasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ and hexasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$
(adapted from *Alekseeva* et al., Carb. Res. 2014)

The tetrasaccharide fraction was chosen because its structural characterization by NMR and LC-MS is easier with respect to longer-chain oligosaccharides. To verify the role of the 1,6-anhydro-bridge in the susceptibility towards oxidation a fraction containing the hexasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ have been also studied in the presence of periodate. As first step, SEC-fractionation of the enoxaparin sample was carried out to obtain a fraction enriched in tetrasaccharides, which was, then, sub-fractionated by IPRP-HPLC to isolate the target tetrasaccharides (Fig.4.16). The obtained fractions were fully characterized by 2D NMR and LC-MS analysis. The corresponding spectral data are reported in Fig.4.17.

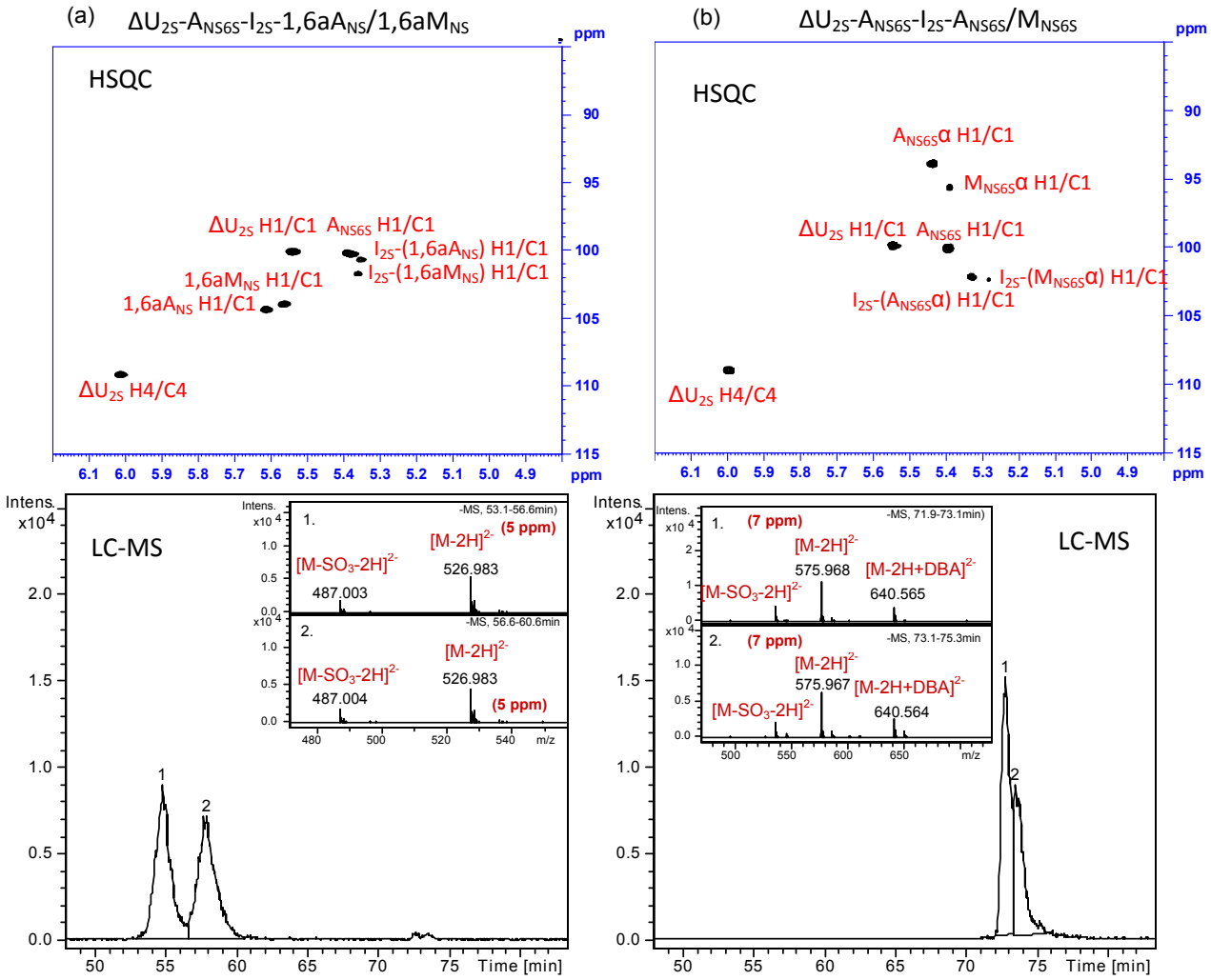



**Fig.4.17.** HSQC NMR spectra (anomeric region) and LC-MS data of (a) enoxaparin fraction containing pentasulfated isomers $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$, (b) fraction containing hexasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ (reported from *Alekseeva* et al., Carb. Res. 2014)

*Behavior of reducing end 1,6-anhydro-amino sugars in the presence of periodate*

The periodate treatment of the isolated mixture of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ was performed directly in the NMR tube in order to monitor the anomeric signals in real time. The absence of $A_{NS}$-G disaccharide unit within the isolated tetrasaccharides permitted to monitor the anomeric signal of 1,6a$M_{NS}$ by monodimensional $^1$H NMR spectroscopy. In the case of more complex fractions or unfractionated enoxaparin, the 1,6a$M_{NS}$ $^1$H NMR signal (at 5.57-5.59 ppm; *Guerrini* et al., 2007; *Mascellani* et al., 2007) overlaps with the anomeric signal of $A_{NS}$ followed by G (at 5.58 ppm; *Guerrini* et al., 2005).

Even if their differentiation is possible by 2D HSQC NMR spectroscopy, it could not permit monitoring reaction because of much longer acquisition times. In Fig.4.18 the $^{1}$H NMR spectra of the pentasulfated tetrasaccharide mixture in the presence of periodate after different time periods are shown.

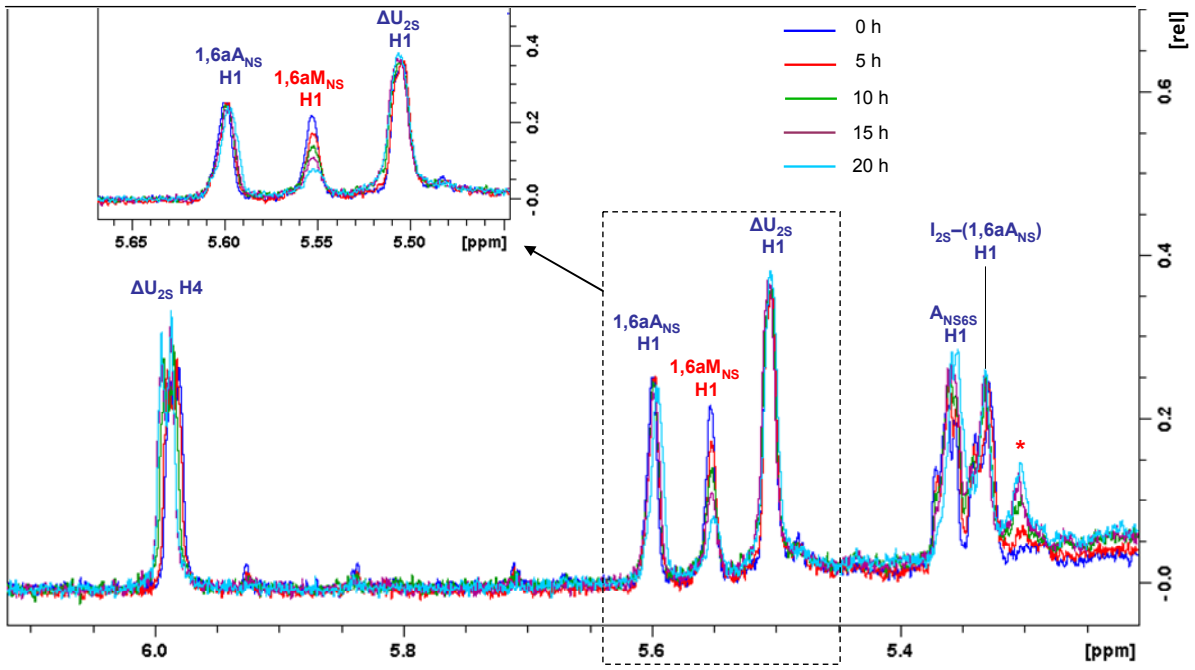



**Fig.4.18.** $^{1}$H NMR real-time monitoring of anomeric H1 signals and H4 signals of unsaturated uronic acids of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ mixture before and after addition of periodate (reported from *Alekseeva* et al., Carb. Res. 2014)
Symbol * indicates the anomeric signal of $I_{2S}$ adjacent the modified 1,6a$M_{NS}$

It can be noted that the chemical shifts of the 1,6-anhydro-residues (1,6a$M_{NS}$ 5.56 ppm and 1,6a$A_{NS}$ 5.60 ppm) are slightly shifted with respect to the published values for enoxaparin (1,6a$M_{NS}$ 5.59 ppm and 1,6a$A_{NS}$ 5.63 ppm; *Guerrini* et al., 2007), probably due to the shorter length of these oligosaccharides. In fact, the observed values are similar to those of some disaccharides with 1,6-anhydro-terminal units (1,6a$M_{NS}$ 5.57 ppm and 1,6a$A_{NS}$ 5.61 ppm; *Mascellani* et al., 2007). The anomeric 1,6a$M_{NS}$ signal showed a notable intensity decrease after 5 h of reaction with periodate, and after 20 h it was present only in trace amount (Fig.4.18). Together with the disappearance of 1,6a$M_{NS}$ peak, a new signal (indicated as * in

the Fig.4.18), likely corresponding to the anomeric signal of $I_{2S}$ adjacent to the modified 1,6a$M_{NS}$, at 5.30 ppm (*see below*) appeared.

LC-MS analysis of the reaction mixture after 40 h showed one of the unaffected isomeric pentasulfated tetrasaccharides (Fig.4.19), identified as 1,6a$A_{NS}$-containing isomer because all its signals were preserved in the NMR spectrum (Fig.4.18). On the contrary, the LC-MS peak of the second isomer, bearing the 1,6a$M_{NS}$ residue, disappeared, and a new peak with *m/z* 486.4881 appeared in the LC-MS chromatogram (Fig.4.19).

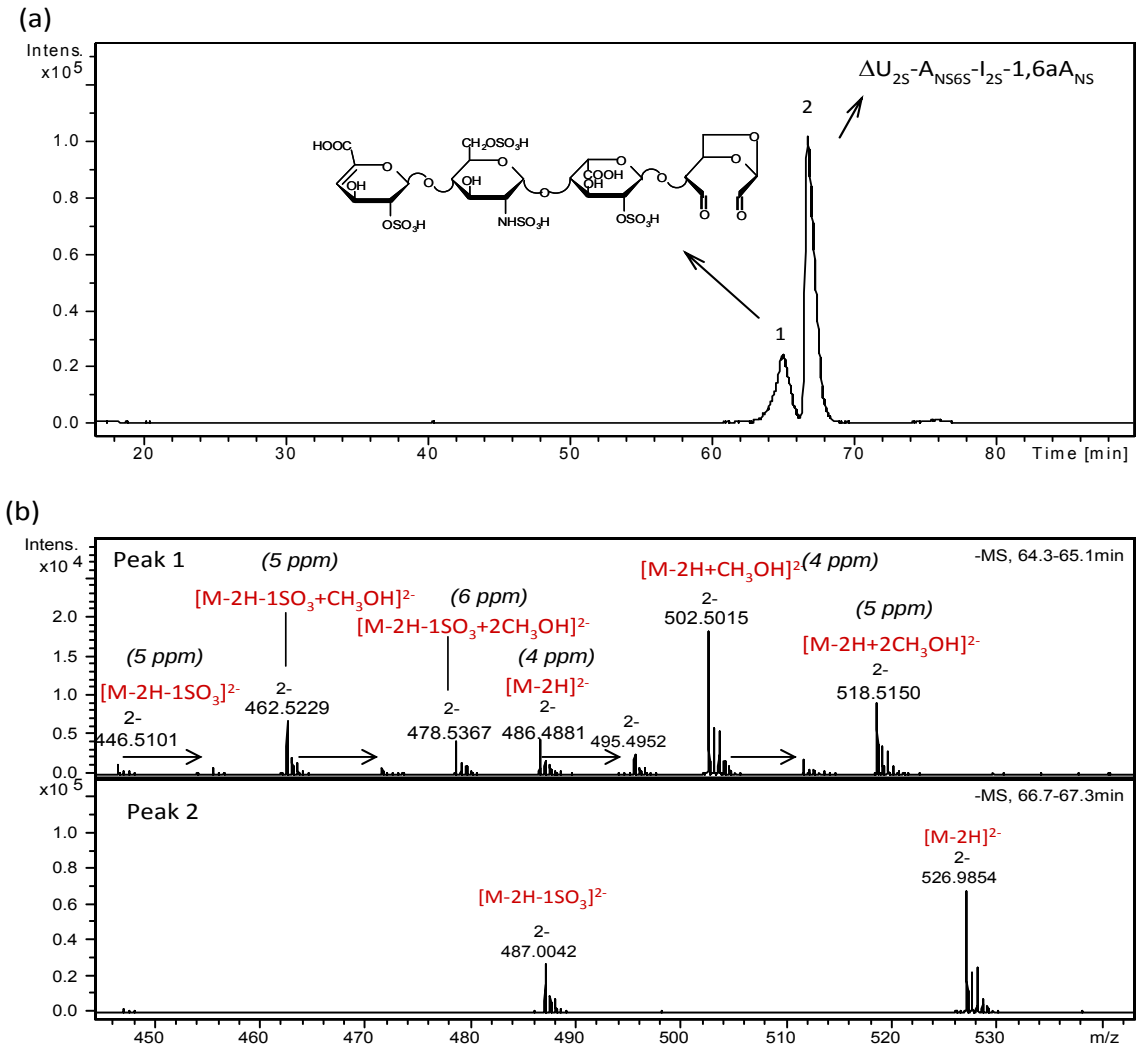



**Fig.4.19.** LC-MS data of the mixture of ΔU2S-ANS6S-I2S-1,6a$M_{NS}$ and ΔU2S-ANS6S-I2S-1,6a$A_{NS}$ after periodate addition (40 h) (adapted from *Alekseeva* et al., Carb. Res. 2014)

(a) – LC-MS chromatogram, (b) – mass spectra of the two peaks observed (errors between experimental and calculated *m/z* values are reported for each ion form of the new compound generated by periodate)

Arrows indicate sodium adducts

The odd value of the nominal mass (975 Da) suggests the presence of a single nitrogen atom. Since the internal $A_{NS6S}$ residue was shown by NMR to be unaffected (Fig.4.18), it is the RE 1,6-anhydro-ring that does not have a nitrogen atom. The hypothesized structure containing split 1,6-anhydro-ring is shown in Fig.4.19. Interestingly, in the MS spectrum it appeared in free (*theoretical m/z* $486.4861^{(2-)}$) as well as hydrated and methyl acetal forms (formed by both methanol and water present in the eluent), proving the presence of two aldehyde groups in its molecule (Fig.4.19).

The solution recovered from the NMR tube was treated with sodium borohydride to stabilize the formed aldehydes and, then, analyzed by NMR and LC-MS. HSQC NMR spectroscopy showed that all the cross-peaks in the anomeric region, except for $1,6aM_{NS}$ and the adjacent $I_{2S}$ preceding it ($I_{2S}$-($1,6aM_{NS}$)), maintained the same positions as in the original mixture of the two pentasulfated tetrasaccharides (Fig.4.20). The anomeric signal of $I_{2S}$ linked to the modified terminal residue shifted from 5.37 (blue spot of $I_{2S}$-($1,6aM_{NS}$) in HSQC spectrum, Fig.4.20a) to 5.31 ppm (red spot of $I_{2S}$-(gs1,6a) in HSQC spectrum) as shown by TOCSY spectroscopy (Fig.4.21).

Notably, while the anomeric $1,6aM_{NS}$ cross-peak disappeared, a new signal at 5.04/105.9 ppm was observed in the region typical for anomeric signals of gs-units (4.8-5.0/105-107 ppm). Its correlation with the cross-peak at 3.72/64.0 ppm, typical for $CH_2OH$ groups (their presence was confirmed by HSQC-DEPT experiment, *not shown*), supports the presence of a gs-residue. The other three signals that appeared in the "ring region" after the periodate/borohydride reaction (Fig.4.20b), namely, at 4.10/69.5 ppm, 3.94/82.1 ppm and 3.75/64.7 ppm (gs1,6a H3/C3), correlate in the TOCSY spectrum (indicated as gs1,6a in Fig.s 4.20, 4.21). The latter signal was assigned by HSQC-DEPT experiment (*not shown*) to the second $CH_2OH$ group of the gs-unit, arising from the C(3) of the starting $1,6aM_{NS}$ residue.

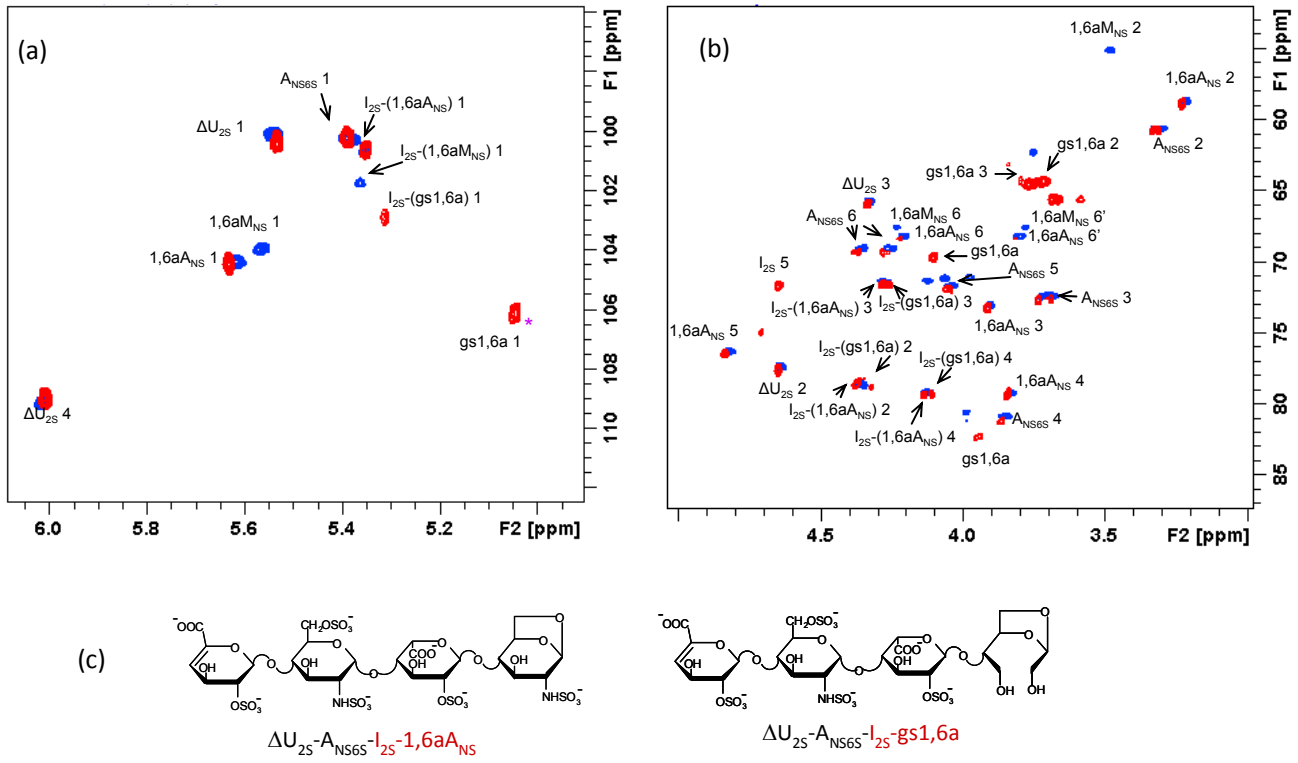



**Fig.4.20.** HSQC spectra of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ mixture before (blue spots) and after (red spots) periodate/borohydride treatment (reported from *Alekseeva* et al., Carb. Res. 2014)
(a) – anomeric region, (b) – "ring" region. In panel (c) the structures of the unaffected tetrasaccharide $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and the gs one $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-gs1,6a are shown.
The first two residues $\Delta U_{2S}$ and $A_{NS6S}$ within their sequence (cross-peaks shown in black) show very similar chemical shifts in the anomeric region of the spectrum, whereas the $I_{2S}$ and terminal 1,6a$M_{NS}$/1,6a$A_{NS}$ (shown in red) differ from each other. Chemical shifts of $I_{2S}$ residues differ from each other due to the so-called "sequence effect"

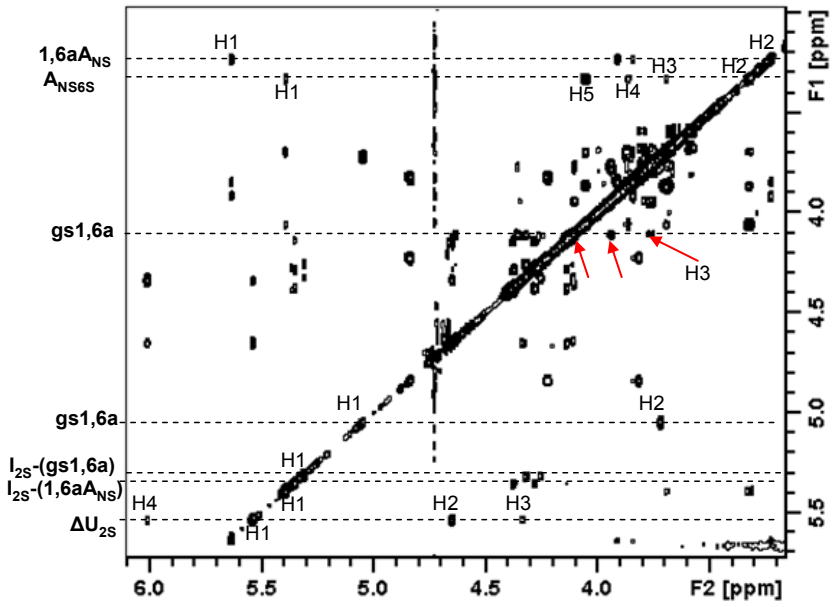



**Fig.4.21.** TOCSY spectrum of the mixture containing the unaffected tetrasaccharide $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and the gs-product $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-gs1,6a
(reported from *Alekseeva* et al., Carb. Res. 2014)

The LC-MS chromatograms of the pentasulfated tetrasaccharide mixture before and after periodate/borohydride treatment are shown in Fig.4.22.

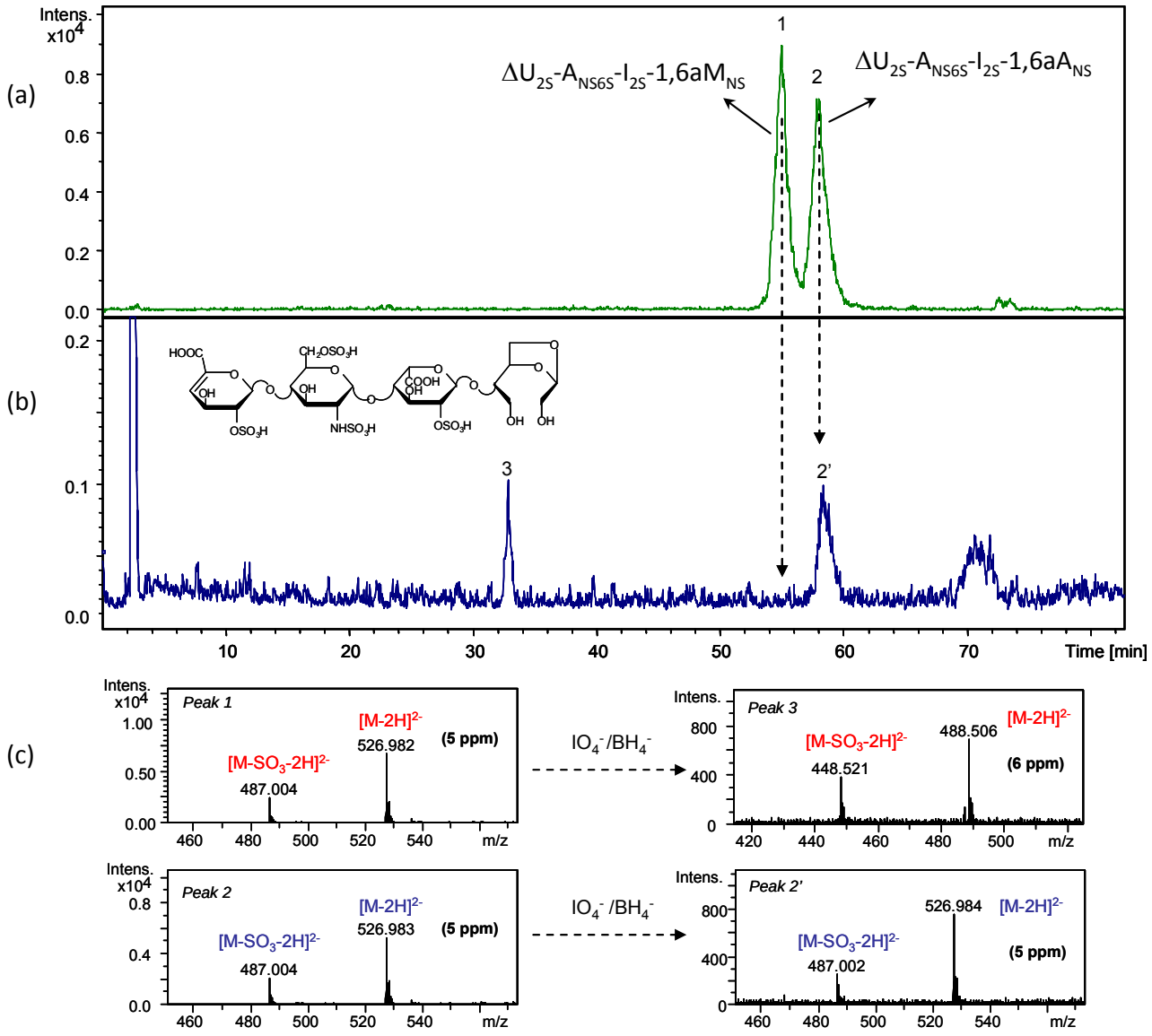



**Fig.4.22.** LC-MS chromatogram of the $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ mixture (a) and the same mixture after periodate/borohydride treatment (b). In panel (c) the corresponding mass spectra are shown (reported from *Alekseeva* et al., Carb. Res. 2014)

Peak at 70 min in panel (b) is a system signal (due to the gradient)

Combining NMR and LC-MS data several conclusions can be drawn. First of all, as in the case of the solution treated only with periodate, the 1,6a$A_{NS}$-containing epimer remains unaffected (*m/z* $526.983^{(2-)}$). Our knowledge about the elution order of heparin oligosaccharides in IPRP mode is in agreement with this result. It has been noted that

1,6aM$_{NS}$-containing oligosaccharides were eluted before their 1,6aA$_{NS}$ isomers (*unpublished data from our group*). In fact, the peak with a shorter retention time (RT 55.0 min) disappeared from the chromatogram, while the second one (RT 58.2 min) remained unaffected (Fig.4.22). The new peak (RT 32.9 min) observed in the chromatogram after periodate/borohydride treatment corresponds to the reduced form (*m/z* 488.506$^{(2-)}$) of the compound generated by periodate oxidation (*m/z* 486.488$^{(2-)}$, Fig.4.19). The increase of the nominal mass value by 4 Da after reduction is in agreement with the proposed structure with two aldehyde groups. The analyte formed after reduction is the same observed in the heparinase-digest of the gs-enoxaparin.

Based on the well-known mechanism of periodate oxidation that includes the formation of a cyclic intermediate (*Dryhurst*, 1970; *Perlin*, 2006) and the predicted conformations of the tetrasaccharides under study (*Alekseeva* et al., Carb. Res. 2014, Fig.4.22), the observed oxidation of the 1,6-anhydro-D-mannosamine-*N*-sulfate may be explained by the *e-/a-cis*-position of $NHSO_3^-$ and OH groups at the C(2) and C(3), respectively (Fig. 4.23b). It is likely that the coordination of nitrogen atom by periodate ion could lead to *N*-desulfation, formation of a cyclic intermediate and further splitting of the C(2)-C(3) bond. In the case of 1,6aA$_{NS}$ the *a-/a-trans*-position of $NHSO_3^-$ and OH groups (Fig.4.23a) should not allow the formation of a periodate cyclic intermediate, and, as a consequence, the 1,6aA$_{NS}$ should not be oxidized.

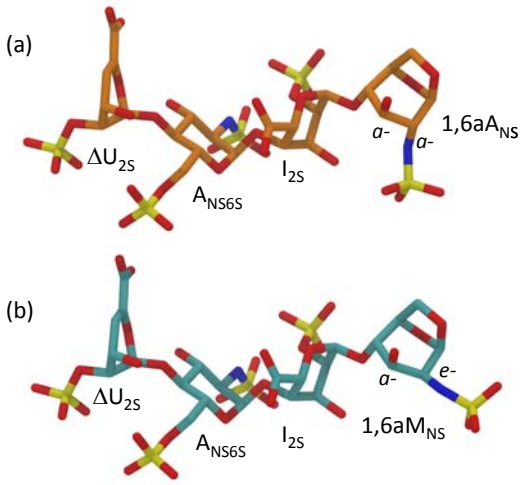



**Fig.4.23.** Conformation predicted for the tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ (a), $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ (b)
(from *Alekseeva et al., Carb. Res. 2014*)

*Behavior of terminal N-sulfated mannosamine and glucosamine residues in the presence of periodate*

Unlike for 1,6-anhydro-tetrasaccharides, monodimensional $^{1}$H NMR could not be used for monitoring the anomeric signals of RE amino sugars of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ because of the overlapping of the H1/C1 of $M_{NS6S}$ (α-anomer) with that of the internal $A_{NS6S}$ (Fig.4.1b). 2D HSQC spectra were acquired after 5, 10 and 20 h. However, even after 40 h and additional portion of periodate, the cross-peak volumes of H1/C1 signals of $A_{NS6S}\alpha$ and $M_{NS6S}\alpha$ did not significantly change, as, at first glance, also the LC-MS profile of the hexasulfated tetrasaccharide mixture (Fig.4.24). Both the isomeric peaks were still present in the chromatogram after 40 h (after borohydride treatment appeared in their alditol forms, *not shown*), along with a new peak with *m/z* $506.47^{(2-)}$ observed only in trace amount (Fig.4.24). The observed *m/z* value $506.47^{(2-)}$ corresponds to $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$R_{am}$, where $R_{am}$ is a remnant formed by simultaneous splitting of C(1)-C(2) and C(2)-C(3) bonds of the RE amino sugar (Fig.4.24c). Borohydride treatment gave rise to its reduced form with *m/z* $507.47^{(2-)}$. A mass increase of 2 Da, together with the presence of only a mono-methyl acetal form of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$R_{am}$ (*m/z* $522.48^{(2-)}$) (Fig.4.24c), confirms the presence of one aldehyde group in its structure.

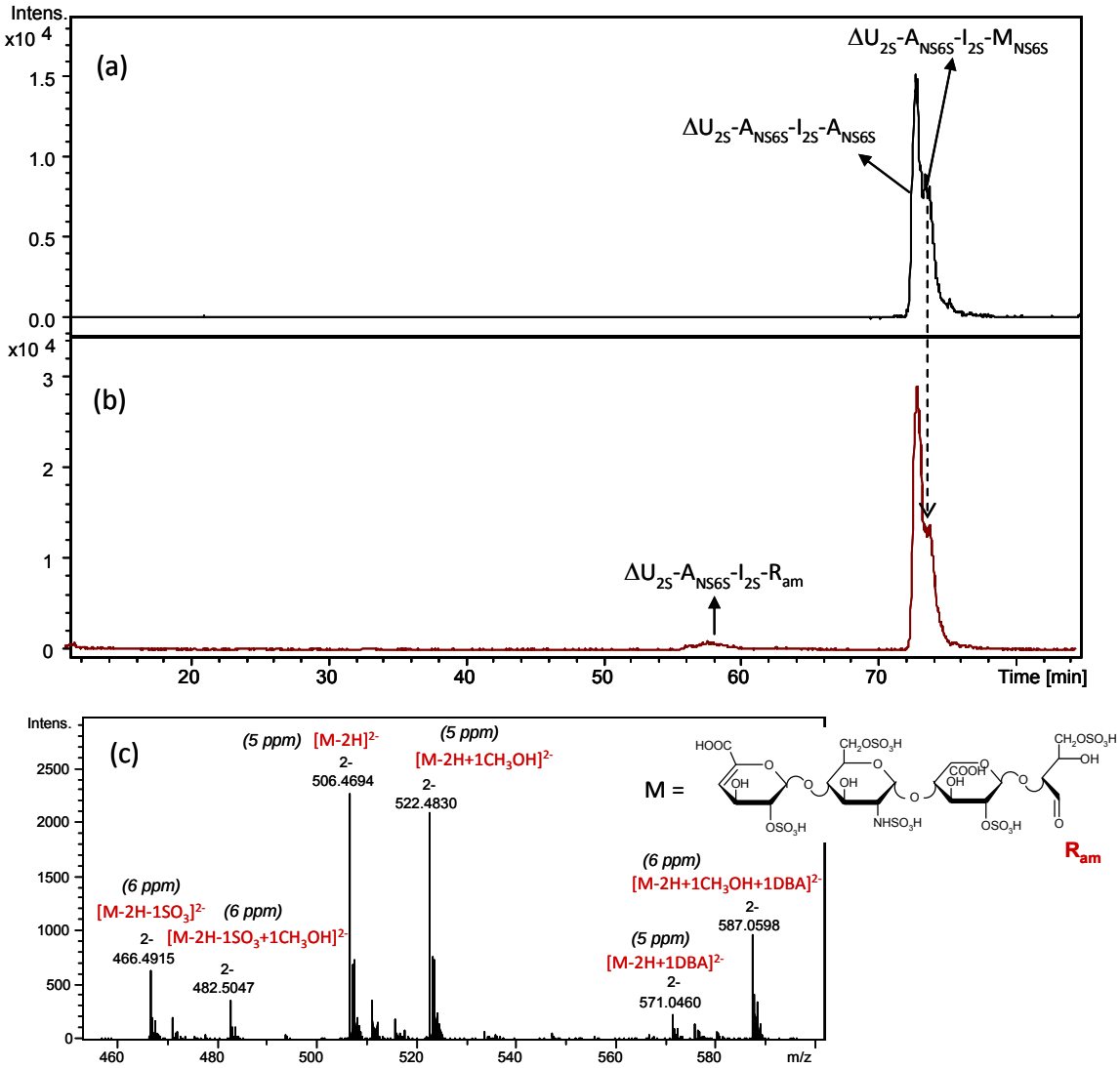



**Fig. 4.24.** LC-MS chromatograms of isomeric hexasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ before (a) and after (b) periodate treatment (40 h). In panel (c) the mass spectrum of the peak generated by periodate oxidation is shown after 5 days but not after 40 h because for the latter case the peak intensity was too low for exact mass determination (reported from *Alekseeva* et al., Carb. Res. 2014)

Peak broadening may be explained by the presence of numerous forms of the oxidized tetrasaccharide because aldehyde groups reversibly react with water and methanol present in the mobile phase. Signals in the mass spectrum without attribution correspond to sodium and potassium adducts of the ion form shown.

$\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$R_{am}$ could be generated by both epimers. However, a change in the peak intensity ratio of the two hexasulfated tetrasaccharides after addition of periodate suggested a prevalent oxidation of the isomer with a longer RT (73.8 min) (Fig.4.24). Based on our knowledge about the elution order of enoxaparin oligosaccharides (*unpublished data of our group*), this component is the $M_{NS6S}$-bearing isomer. Because of a slow reaction rate and insufficient chromatographic resolution it was difficult to establish whether also the terminal

$A_{NS6S}$ residue was affected by periodate. The possibility of its oxidation was independently verified using a disaccharide $\Delta U_{2S}$-$A_{NS6S}$, isolated during SEC fractionation of the heparinase-digest of gs-enoxaparin (fraction H in Fig. 4.14a). The natural $A_{NS6S}$ residue of this disaccharide (formed by enzymatic hydrolysis) was found to be partially oxidized by 0.1 M sodium periodate: a new minor peak, corresponding to the $\Delta U_{2S}$-$R_{am}$ (*m/z* 436,9701), appeared in the LC-MS chromatogram after 20 h of oxidation (*not shown*). We also carried out the oxidation of hexasulfated tetrasaccharides for longer time (5 days). Even if the intensity of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$R_{am}$ was slightly increased with respect to the results shown in Fig.4.24, about 90 % of tetrasaccharides were still unaffected (*not shown*).

Altogether, these data indicate that both terminal residues, $M_{NS6S}$ and $A_{NS6S}$, may be partially affected by periodate under drastic conditions (long reaction time and high periodate concentration). The oxidation of the $M_{NS6S}$-isomer appeared relatively faster probably due to the *cis-a-/e*-position of $NHSO_3^-$ and OH groups (Fig.4.25), but much slower than that of 1,6a$M_{NS}$, suggesting an essential role of the 1,6-anhydro-"bridge" (see Discussion).

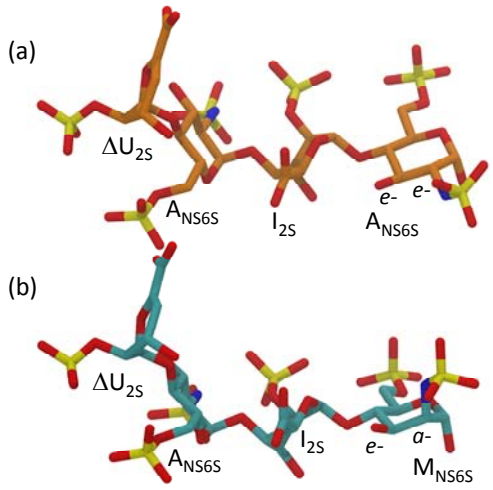



**Fig.4.25.** Conformation predicted for the tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ (a), $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ (b)
(from *Alekseeva et al., Carb. Res. 2014*)

It is worth noting that C(2) $NHSO_3^-$ and C(3) OH groups of the internal $A_{NS6S}$ residue of all the target tetrasaccharides as well as RE residue of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ are in *e-/e-trans*-position that could allow the formation of a cyclic periodate intermediate (Fig.4.25). However, the internal ones were found resistant to the oxidation, probably due to the effect of bulky and electrostatic repulsive groups of the adjacent 1,4-linked residues.

### IV.5. Discussion

Modifications generated in LMWHs by the glycol-splitting periodate oxidation followed by a "stabilizing" borohydride reaction involve not only internal non-sulfated uronic acids but also several end-residues, that was not observed for gs derivatives of the unfractionated heparins (Chapter III). Combination of NMR spectra and LC-MS data revealed all these structural modifications in the LMWHs, providing complementary and characteristic fingerprints of both LMWHs and their gs-derivatives. It is worth noting that under the mild experimental conditions used for gs-LMWHs preparation (i.e. acetic acid used for pH neutralization after borohydride reduction), the conversion yields of the parent LMWHs to the corresponding gs-derivatives were quite high (>85%). However, a slight $M_w$ decrease (9-15% lower than for the parent LMWHs, Table 4.1) was still observed, indicating some chain shortening.

As for gs-heparins, a note of caution is necessary when dealing with gs-derivatives of LMWHs, since the local flexibility of gs residues may affect their relaxation times $T_2$ (*Casu* et al., 2002). Quantification of the internal residues of gs-heparin chains was possible, as shown in Chapter III. However, it should be considered with caution especially for terminal residues, such as gsXyl, because of longer $T_2$ values of the more mobile gs residues. In fact, the highest discordance was observed for gs-LR residues: the gsXyl content was almost always overestimated with respect to the other residues of gsLR ($gsG_{LR}$, Gal1 and Gal2, Table 4.2). The reported quantitative data for gs-LMWHs (Table 4.2), though compatible with our overall conclusions, should accordingly be taken as provisional, until a deeper study is made using well-defined model compounds.

Both HSQC NMR spectra and LC-MS profiles of gs-LMWHs (as well as their digests) also retain some features characteristic of the LMWH production method: partial depolymerisation with heparinase I (tinzaparin), base-catalyzed β-elimination (enoxaparin), and nitrous acid depolymerisation (dalteparin) in terms of typical end groups. All of these

depolymerisation methods cleave only glycosidic bonds between an amino sugar and the adjacent uronic acid, leading to formation of even-numbered oligosaccharides (Fig1.15). Accordingly, all the parent LMWHs as well as gs-tinzaparin and gs-enoxaparin mainly consisted of even-numbered oligosaccharides. Surprisingly, gs-dalteparin contained numerous odd-numbered oligosaccharides in its LC-MS chromatogram (Fig.4.12). Their formation is likely associated with glycol-splitting of the dalteparin terminal uronic acid residues at the NRE bearing 3,4-diol groups and the further hydrolysis of the generated remnants (Fig.4.13c). Together with the observed high content of odd-numbered oligosaccharides, also the occurrence of occasional ring contracted residues Rc diversifies gs-dalteparin from the other gs-LMWHs (Fig. 4.12). In the case of tinzaparin and enoxaparin, 4,5-unsaturated 2-*O*-sulfated uronic acids at the NRE "protect" the chains from shortening during periodate oxidation. Given their susceptibility to periodate, minor ΔU and saturated uronic acid residues of enoxaparin may contribute to some chain shortening from the NRE.

As for RE, tinzaparin and enoxaparin amino sugars are susceptible to borohydride reduction, while dalteparin aM.ol rings are resistant. The apparently anomalous behavior of terminal enoxaparin 1,6-anhydro-D-mannosamine-*N*-sulfate residues has been studied and confirmed by NMR spectroscopy and LC-MS using model tetrasaccharides isolated from the LMWH enoxaparin. Moreover, also RE D-mannosamine-*N*-sulfate and D-glucosamine *N*-sulfate, lacking the 1,6-anhydro-bridge, can be oxidized by periodate but with significantly lower rate. It is likely that the *cis*-position of OH and $NHSO_3^-$ groups of terminal residues ($1,6aM_{NS}$ or $M_{NS6S}$) is not a sufficient requirement for the oxidation. The 1,6-anhydro-bridge present at the RE of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$1,6aM_{NS}$ induces a rigid ring conformation more favorable for oxidation to occur. Also the differences in residue electrostatic charge between $1,6aM_{NS}$ (charge = -1) and $M_{NS6S}$ (charge = -2) may explain the preference for the negatively charged periodate to react with the former substrate.

The partial or minimal slow oxidation of RE $A_{NS6S}$ and $M_{NS6S}$ is likely to be associated with their hemiacetalic or acyclic hydrated aldehyde forms, whose C(1) hydroxyl groups can partially contribute to the formation of cyclic periodate intermediates, involving the adjacent C(2) *N*-sulfated groups. Notably, the generated oligosaccharide $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$R_{am}$ ($\Delta$U3,5,0-$R_{am}$) was observed in relatively high quantity in the digest of gs-enoxaparin (Fig.4.5) and in trace in the heparinase-digests of gs-heparins (Chapter III). Initially, its formation in the gs-heparins digests was hypothesized to be the result of oxidation of non-*N*-substituted amino sugars that could be present in heparins. The present study suggests that also partial oxidation of RE hexosamines of heparin and enoxaparin chains may lead to the formation of $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$R_{am}$. Its relatively higher quantity in the gs-enoxaparin digest is likely to be explained by its lower molecular weight and, consequently, higher amount of RE terminal *N*-sulfated amino sugars (~16% of total glucosamine content; *Guerrini* et al., 2007) with respect to unfractionated heparins (<3%; *Desai* and *Linhardt*, 1995).

In order to compare the behavior of 1,6a$M_{NS}$ residues with that of non-sulfated uronic acids and to evaluate its possible contribution to the gs-product heterogeneity, we also performed the oxidation of enoxaparin for different time periods (Fig.4.26). 2D HSQC NMR analysis has shown that the complete oxidation of non-sulfated iduronic acids occurred in 2 h, while glucuronic acids appeared more resistant under the applied conditions, and both glucuronic acids and 1,6a$M_{NS}$ residues were totally oxidized only after 12 h of reaction (Fig.4.26). The obtained data indicate that *N*-substitution and conformational flexibility may drive periodate oxidation. Despite *N*-substitution, 1,6a$M_{NS}$ is oxidized with almost the same rate than non-sulfated glucuronic acid, probably because it is "blocked" in a conformation (*cis*-*a*-/*e*-position of OH and $NHSO_3^-$ groups) favorable for formation of a cyclic intermediate with periodate. Glucuronic acid is mainly present in $^4C_1$ conformation with *trans*-*e*-/*e*-position of OH and $NHSO_3^-$ groups (Fig.1.8), which is probably less favorable for periodate-cycle formation. $^1C_4$ conformation of iduronic acid with *trans*-*a*-/*a*-position of OH and $NHSO_3^-$

(Fig.1.8) could not form a cyclic intermediate with periodate ion. Consequently, its higher susceptibility to periodate with respect to glucuronic acid may be explained by its higher conformational flexibility (*Ferro* et al., 1990).

The difference in the oxidation kinetics of different units could permit to control the glycol-splitting reaction.

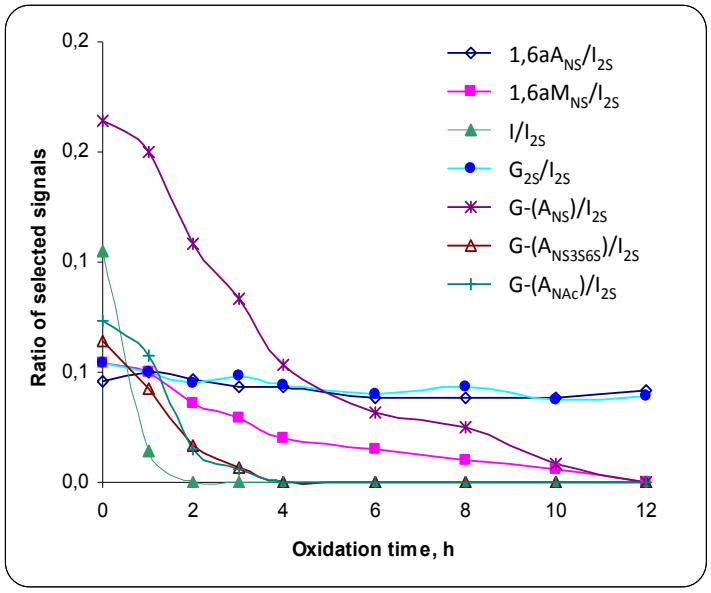



**Fig.4.26.** HSQC volume integrals of selected residues of enoxaparin normalized by the value of $I_{2S}$ integral as a function of periodate oxidation time

$1,6aA_{NS}$ – 1,6-anhydro-D-glucosamine-*N*-sulfate, $1,6aM_{NS}$ – 1,6-anhydro-D-mannosamine-*N*-sulfate, I – L-iduronic acid, $I_{2S}$ – L-iduronic acid 2-*O*-sulfate, G-($A_{NS}$) – D-glucuronic acid followed by D-glucosamine-*N*-sulfate, G-($A_{NS3S6S}$) – D-glucuronic acid followed by D-glucosamine-*N*,3-*O*,6-*O*-trisulfate, G-($A_{NAc}$) – D-glucuronic acid followed by *N*-acetyl-D-glucosamine

LC-MS analysis of heparinase digests provide additional structural information on oligosaccharide components of gs-LMWHs, and, indirectly, on the parent compounds. Their LC-MS profiles are more complex, and probably more informative, than those of gs-heparins (Chapter III) because of interference on the enzyme by several unnatural end-groups. RE terminals, such as alditols (in gs-tinzaparin and gs-enoxaparin), 1,6-anhydro-mannosamine (in gs-enoxaparin), 2,5-anhydro-mannitol (in gs-dalteparin), were not recognized by heparinases, leading to the preferential cleavage of glycosidic bond one disaccharide unit further towards the NRE. In fact, LC-MS peaks of the disaccharides with such unnatural residues at the NRE were never observed in the enzymatic digests. These results are in agreement with previous reports on enoxaparin and dalteparin (*Linhardt* et al., 1990; *Ozug* et al., 2012; *Viskov* and *Mourier*, 2004), showing that tetrasaccharide fragments were generated instead of disaccharides as it would be the case if the end residues were unmodified glucosamines. It may be useful to

characterize the sequences at the RE of the LMWH chains. For example, the presence of ΔU3,4,0-ol in the gs-enoxaparin digest suggests that $I_{2S}$-$A_{NS6S}$-$I_{2S}$ sequence, probably generated during the bleaching manufacture process, was at the RE of several enoxaparin chains.

An additional inhibition exerted by gs-residues, first observed on gs-heparins (Chapter III), permitted to characterize the heparin sequences near the LR by the first time, and to identify the ATBR-containing sequences with different sulfation/*N*-acetylation patterns.

The results obtained for the gs-tinzaprin digest suggest that the sequences in the closest proximity to the LR are shorter (i.e., dp 2-6) than previously assumed. Traditional representation of heparin structure (Fig.1.2) shows a well-defined low sulfated *N*-acetylated domain constituted by repeating $A_{NAc}$-G disaccharide units, however, the present results (see IV.4.1) show that its sulfation degree is not so low. Oligosaccharides ΔU4,3,1-LR, ΔU4,4,1-LR, ΔU8,6,2-LR, ΔU8,9,1-LR, ΔU10,11,1-LR and ΔU10,11,2-LR were the prevalent LR-containing species of tinzaparin, while ΔU4,3,1,1gs-gsLR, ΔU4,3,1,1gs-gsLR, ΔU6,2,2,2gs-gsLR and ΔU6,3,2,2gs-gsLR (Table 4.5) were prevalent in the gs-tinzaparin digest. Octasaccharides bearing gs-LR were observed only in trace, and longer sequences were not detected even after accurate fractionation of gs-enoxaparin digest. Similar conclusions were made for the internal sequences: the found gs-octasaccharides with three gs residues are mono-*N*-acetylated (ΔU8,8,1,3gs, ΔU8,9,1,3gs, etc), suggesting that the sequences with three contiguous non-sulfated uronic acids within the original heparin chains are highly sulfated.

Concerning ATBR-containing oligosaccharides, the hexasaccharide ΔU6,7,1,2gs (in the case of dalteparin it could be also incorporated in the minor octasaccharide ΔU8,9,1,2gs-aM.ol) was observed in the LC-MS profiles of all the gs-LMWH digests. As for gs-heparins, tetrasaccharide ΔU4,6,0,1gs (generated from a highly sulfated ATBR variant) was found in trace amount also in all gs-LMWHs digests.

It is important to note that the glycol-splitting-induced local mobility is likely to exert an influence on retention times in LC-MS profiles of components of gs-LMWHs, allowing separation of some isomeric oligosaccharides otherwise difficult to achieve. Several trends in the elution order in ion-pair reversed phase HPLC, which are in agreement with the separation mechanism (Fig.1.24), were observed. For oligosaccharides with the same dp, the higher the sulfation degree, the higher the retention time RT (for example, RT (ΔU6,7,0) < RT (ΔU6,8,0) < RT (ΔU6,9,0) in enoxaparin Fig.4.11): the greater number of negatively charged sulfate groups induces stronger interaction with the stationary phase, dynamically modified by the positively charged dibutylamine cation. The presence of a *N*-acetyl group causes a slight increase of the retention time (RT (ΔU6,7,0) < RT (ΔU6,7,1)) that indicates that not only electrostatic but also hydrophobic interactions between analytes and stationary phase can take place. Probably for the same reason, enoxaparin oligosaccharides with 1,6aA-residues are eluted after their analogs bearing hemiacetalic aminosugars (RT ΔU6,8,0 < RT ΔU6,8,0-1.6aA). Unexpectedly, the retention time of ΔU16,21,0 is somewhat lower than that of ΔU14,21,0 in tinzaparin (Annex 5) and enoxaparin (Annex 6). This finding may be explained by the fact that the longer the oligosaccharide chains, the less the influence of additional disaccharide units and the more crucial the influence of the number of anionic groups per chain. Unsaturation in the NRE uronic acid increases the RT (RT (U6,9,0) < RT (ΔU6,9,0)), probably due to the higher acidity of the uronic acid carboxy-group when conjugated with a double bond, leading to a stronger interaction with the dibutylammonium cation. The same effect may take place in the case of the odd enoxaparin oligosaccharides starting with $\Delta U_{(2S)}$, which show higher RT than oligosaccharides starting with a glucosamine residue (RT (A3,4,0) < RT (ΔU3,4,0)); moreover, ΔU3,4,0, bearing two carboxyl groups, have more anion centres able to interact with dibutylammonium cation. Oligosaccharides in alditol form are eluted after their unmodified analogs (RT (ΔU6,8,0) < RT (ΔU6,8,0-ol)), and this is probably caused by the higher flexibility of the reduced open form with respect to that of the hemiacetalic one, that

facilitates the interactions between sulfate groups and dibutylamine on the stationary phase. The presence of an uronic acid remnant increases the retention time (e.g., in gs-tinzaparin (RT (ΔU8,12,0-ol) < RT (ΔU8,12,0-R)), probably due to the additional charge of the remnant carboxyl group (R) present at the RE (Fig.1.19e).

Not only the total charge but its distribution within the analyte molecules plays a role in the IPRP-HPLC separation. For the tetrasaccharides epimeric pairs, $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$/ $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$/$\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$, the value of the electric dipole appeared to influence the epimer capacity to interact with the stationary phase, driving their separation. In this case the higher dipole moment magnitude of the molecule correlates with the stronger interaction with the stationary phase. The 1,6a$A_{NS}$-bearing tetrasaccharide with a dipole moment length (11.9 eÅ, reported in *Alekseeva* et al., 2014), greater than that of 1,6a$M_{NS}$-bearing isomer (9.0 eÅ, reported in *Alekseeva* et al., 2014), is characterized by a longer RT (Fig.4.22a). The same observation can be done for the isomeric hexasulfated tetrasaccharides: $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$, characterized by a lower dipole moment (8.5 eÅ, reported in *Alekseeva* et al., 2014), is eluted with a shorter RT (73.5 min,) than its isomer $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ (dipole moment 9.3 eÅ, reported in *Alekseeva* et al., 2014; RT 73.8 min, Fig.4.24a). These data underline a possible correlation between the molecule electric dipole moment, calculated using molecular mechanics techniques, and retention time in the ion-pair reversed phase mode of LC separation.

### IV.6. Conclusions

The combined NMR/LC-MS method developed for gs-heparins was successfully applied for gs-LMWH characterization. The obtained data show that also the direct LC-MS analysis without cleavage with heparinases, even if does not give detailed structural information because of the high complexity of the sample, is sensitive to the structural differences and composition of gs-products and can be a useful tool for gs-LMWH development. When used

together, these techniques permit the detection of the major internal modifications introduced in LMWH chains, including those caused by side hydrolytic processes, as well as modifications of several end residues susceptible to periodate/borohydride. As for gs-heparins, the combination of glycol-splitting and enzymatic digestion provides more informative data also on the starting material (chain distribution of non-sulfated uronic acids, structure of terminal sequences in gs-LMWHs, gs-ATBR containing sequences). In particular, the *N*-acetylated domain in the closest proximity to the LR was characterized using this approach. Additional heparinase-resistance of the oligosaccharides with unnatural residues, typical for each type of LMWHs/gs-LMWHs, may provide even more detailed structural information about the LMWH sequences. Ring contracted residues were observed in dalteparin and its heparinase-digest by LC-MS for the first time. Different oxidation rate of isomeric iduronic and glucuronic acid has been shown in the present study, which can be useful for further optimization of the reaction conditions in order to obtain more homogeneous gs-products.

The data obtained by LC-MS analysis suggest that numerous interactions, such as electrostatic, hydrophobic etc., take place in the ion-pair reversed phase separation mechanism. Total charge and sulfate distribution were shown to drive the separation efficacy. Using model tetrasaccharides it has been shown that the higher the electric dipole moment magnitude value, the longer is the retention time. These data may give a possibility to predict the chromatographic behavior of heparin-related oligosaccharides and to distinguish epimers in the ion-pair reversed phase mode of separation. Glycol-splitting as well as end group modifications (such as alditol formation) definitely improve the separation of isomers.

# CHAPTER V: GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

## General conclusions

In the present study a method, combining advanced analytical techniques NMR, LC-MS and SEC-TDA, has been developed, providing previously unavailable criteria for structural analysis of gs-heparins and gs-LMWHs as well as a complement to current methods for characterization of parent unfractionated and low-molecular weight heparin preparations. It is worth noting that even if NMR, LC-MS and SEC-TDA give characteristic sample fingerprints, they provide much more structural information when used in combination, in particular for complex polysaccharides such as heparins/gs-heparins.

It has been shown that numerous reactions may occur during glycol-splitting, contributing significantly to the composition of the final gs-product. Major internal modifications introduced by periodate oxidation/borohydride reduction in heparin/LMWH chains, as well as in the end residues, can be detected by the developed method, allowing both characterization of individual gs-heparins/gs-LMWHs and comparison of gs-samples from different preparations. Our results show that the present analytical approach can be used for monitoring the reaction completeness and hydrolytic side reactions. In view of a further optimization of reaction conditions that could limit the impact of side-reactions, this approach represents a useful and reproducible tool for fingerprinting gs-samples.

**2D HSQC NMR** turned out to be a useful technique, not requiring previous depolymerisation, to determine the average monosaccharide composition of gs-samples. As for unmodified heparins and LMWHs, the anomeric region is the most informative. All the HSQC NMR anomeric signals characteristic for gs-heparins and gs-LMWHs, including those of the linkage region and signals markers of hydrolytic chain cleavage, have been assigned in the present work. The performed study permits to distinguish gs-iduronic and gs-glucuronic acids (as well as associated glucosamine residues) in various sequences with different

sulfation/acetylation pattern, a differentiation that is impossible to achieve by LC-MS analysis. Since the coupling constants measured for gs-uronic acids were negligibly different from those of unmodified uronic acids, the HSQC NMR spectra permit quantitative analysis of gs-heparins. However, these analytical results should be considered with caution until more detailed investigation will be performed, because the longer relaxation time $T_2$ values characteristic for gs residues may lead to overestimation of their quantities.

**SEC-TDA** data may turn out useful not only for characterizing molecular weight distribution, but also for determining hydrodynamic properties of heparins/gs-LMWHs and their gs-derivatives, which may be of interest for understanding interaction with proteins. The hydrodynamic parameters were sensitive to molecular weight, sulfation degree, and the introduced gs-residues. Interestingly, glycol-splitting favors the formation of random coil structures in solution, suggesting higher chain flexibility with respect to unmodified heparins.

The developed **ion-pair reversed phase HPLC** separation, coupled with **ESI-Q-TOF** mass spectrometer, are reproducible and effective for profiling both digested gs-heparin/gs-LMWH samples and for fingerprinting intact gs-LMWHs. Unlike heparins/gs-heparins, characterized by molecular weights too high to be directly analyzed by LC-MS, profiling of intact LMWHs/gs-LMWHs is possible due to their lower molecular weight. It should be noted that the mass spectrometer used in this work provided a mass accuracy and signal resolution sufficient to identify oligosaccharides up to dp 22 in complex mixtures. Direct LC-MS analysis permitted to determine prevalently the major components, however, it provided also information on the primary composition of the oligosaccharide mixture.

LC-MS analysis of **enzymatically digested gs-heparins/gs-LMWHs** provides an informative profiling complementary to that currently obtained for the starting heparin. The additional information is mainly due to the resistance of gs-residues to heparinases leading to the generation of longer sequences with internal gs-units. Along with a better resolution of isomeric oligosaccharides, the LC-MS chromatograms of heparinase-digested gs-LMWHs

permitted the detection of minor components, such as end-residues in LMWHs. It should be highlighted that the method, initially developed for characterizing gs-samples only, appeared a potentially powerful technique for sequencing the original heparins. Combination of glycol-splitting and enzymatic digestion permitted to characterize for the first time: i) heparin chains in the closest proximity to the linkage region, which appeared shorter and with higher sulfation degree than currently reported, ii) variants of the antithrombin binding region, including its closest environment, iii) sequences of contiguous disaccharides containing non-sulfated uronic acids. It has been also shown that the exhaustive glycol-splitting (as in the case of HI-3 samples) followed by the enzymatic digestion may be used for characterizing *N*-acetylation pattern of heparin-related samples.

For both intact LMWHs and digested gs-heparin/LMWHs samples, glycol-splitting has been shown to improve the resolution of isomeric oligosaccharides, for example, dalteparin oligosaccharides with or without non-sulfated glucuronic acid at the non-reducing end enoxaparin epimers containing reducing end 1,6-anhydro-D-glucosamine-*N*-sulfate and 1,6-anhydro-D-mannosamine-*N*-sulfate.

This study revealed also a possible correlation between the molecule electric dipole moment, calculated using molecular mechanics techniques, and retention time in the ion-pair reversed phase mode of LC separation. For each pair of isomeric tetrasaccharides, the higher the electric dipole moment magnitude value, the longer appears the retention time. These data provide additional insights about the ion-pair reversed phase separation mechanism and may give a possibility to predict the chromatographic behavior of heparin-related oligosaccharides and to distinguish epimers.

**Future directions**

The results obtained in the present study, namely a better understanding of modifications induced by glycol-splitting in internal and terminal residues of heparin/LMWH samples as

well as resistance of gs-residues to heparinases, were used for rationalizing the preparation of gs-oligosaccharides of different length that can also be used for studying interactions with heparanase or other target proteins. The initial idea to fractionate a LMWH and treat the fraction of interest with periodate/borohydride was discarded. The susceptibility of terminal groups to oxidation and/or reduction could generate fragments with unnatural terminals, both NRE and RE, residues that could influence their interaction with heparanase. For this reason, we prepared an extensively glycol-split heparin (HI-3), starting from a partially 2-*O*-desulfated porcine mucosal heparin, using mild conditions in order to avoid hydrolytic chain cleavage. The obtained gs-residue-enriched heparin was digested with heparinases I and II, and the obtained oligosaccharide mixture was fractionated by SEC. As shown by NMR and LC-MS analysis, the isolated fractions are characterized by high glycol-splitting extent (all the internal uronic acids are split) and different chain length (dp from 4 up to 10) and various sulfation degree, representing a good oligosaccharide library for structure – activity relationship studies by affinity chromatography, NMR and MS (Fig.5.1). The gs-tetrasaccharide fraction is planned to be used as a model for the signal assignment in the complex “ring” region of the HSQC NMR spectra.

Another issue raised by the present study is concerning the dependence of periodate oxidation kinetics on various factors. We have observed that glucuronic acid is more resistant to periodate with respect to its isomer, iduronic acid, that may be associated with the conformational differences of these isomer residues. Based on preliminary results, the reaction rate may depend also on the pH value. The on depth investigation of these aspects by NMR, LC-MS and molecular modelling could be useful for further reaction optimization and control. The pH value may also drive hydrolytic side reactions at various steps of the gs-sample preparation, so that further study of the susceptibility of gs-residues to acidic and basic hydrolysis should be investigated in detail.

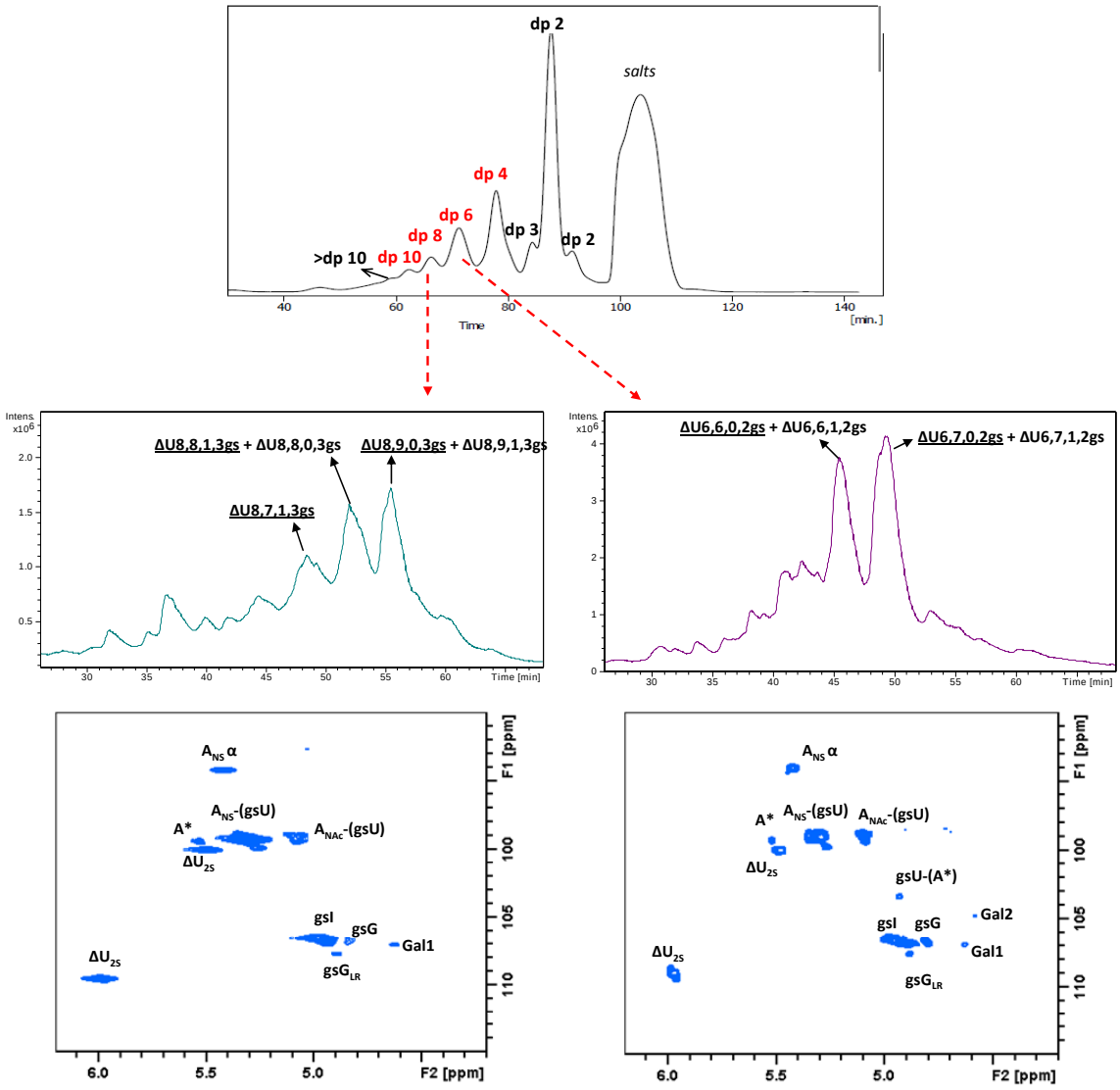



**Fig.5.1.** SEC profile of the heparinase-digest of a HI-3 sample (containing~50% of gs uronic acids) and LC-MS profiles and HSQC NMR spectra of the hexa- and octasaccharide fractions enriched in gs-containing compounds

The work is in progress. Only major components are shown in these preliminary results.

# CHAPTER VI: EXPERIMENTAL

## VI.1. Reagents and starting materials

Heparins were commercial and experimental preparations from: porcine mucosal heparin, PMH, bovine mucosal heparin, BMH (both from Laboratori Derivati Organici, Italy), bovine lung heparin, BLH (Upjohn), ovine mucosal heparin, OMH (Sigma-Aldrich). Low-molecular weight heparins were samples of commercially formulated or bulk materials of: tinzaparin sodium (Innohep, LEO Pharma, Denmark), enoxaparin sodium (Clexane, Sanofi Aventis, Italy) and dalteparin sodium (Fragmin, Pfizer, Italy). *N*-acetyl-heparosan was a gift of K. Jann), and the synthetic pentasaccharide Arixtra® (fondaparinux) was from GSK (France). Uronate unsaturated disaccharides standards $\Delta U_{2S}$-$A_{NS6S}$ (> 95%), $\Delta U_{2S}$-$A_{NS}$ (> 95%) $\Delta U$-$A_{NS6S}$ (> 95%), $\Delta U$-$A_{NS}$ (> 95%), $\Delta U$-$A_{NAc6S}$ (> 95%) and $\Delta U_{2S\text{-}ANAc}$ (> 95%) were from Iduron, UK. Heparin lyases I (EC4.2.2.7), II, III (EC4,2,2,8) were from Iduron or Grampian Enzymes, UK.

Sodium chloride (>99.5%), formic acid (98-100%), hydrochloric acid standard solution 0.1 N, Tris (tris(hydroxymethyl)aminomethane)), hydrochloric acid >37% (p.a.), sodium periodate (> 99%), dibutylamine (>99.5%), pentylamine (99%), hexylamine (99%), methanol (LC-MS grade), acetonitrile (LC-MS grade), acetic acid (glacial, 99.9%), formic acid (98-100%), and ammonium chloride (> 99.5%) were purchased from Sigma-Aldrich. Volumetric solutions of sodium hydroxide 0.1 N and sodium borohydride (95%) were from Riedel-de Haën. Solid sodium hydroxide (>99%) and sodium acetate were from Merck, calcium acetate (>97%), from BDH, ethylenediaminetetraacetic acid (EDTA, >98%) from Fluka. Deionized water was prepared with an osmosis inverse system (conductivity less than 0.06 μS), filtered (Millipore filter 0.22 μm) and used for sample dilution and mobile phases.

*Preparation of an antithrombin high affinity fraction of PMH*

A column of AT-Sepharose gel (10 x 1.6 cm) prepared as previously reported (*Bisio* et al., 2009) was used for a semi-preparative affinity fractionation of a 12 mg PMH by a two steps

elution. The fraction with no AT affinity (NA) was eluted at lower ionic strength (Tris-HCl 0.05 M pH 7.4, containing 0.05 M NaCl) while the high affinity fraction (HA) was eluted at higher salt concentration (Tris-HCl 0.05 M pH 7.4, containing 2.5 M NaCl). Fractions HA and NA were desalted first by dialysis with 3500 Da cut-off membranes (Spectrapor), then by size exclusion chromatography with Toyopearl HW-40 resin (TOSOH) column (58 x 2.6 cm). To remove the last traces of Tris buffer a cation-exchange chromatography was finally performed on an Amberlite IR-120 [$H^+$] column (12.5 x 2 cm). Eluents were neutralized with 0.1 N NaOH, and samples recovered after freeze drying.

*Preparation of partially 2-O-desulfated heparin*

Partially and selectively 2-*O*-desulfated heparins were prepared by alkaline treatment as described previously (*Casu* et al., 2002). PMH (~1g) was dissolved in 6.4 ml of water; after addition of 6.4 ml of 2M NaOH the solution was heated at 60°C for 30 min under stirring. The reaction was stopped by cooling, and the solution was neutralized to pH 7 with 0,1M HCl. The obtained intermediate heparin epoxide was freeze-dried, dissolved in 75 mL of water and heated for 48 h at 70°C. After cooling, desalting, dialysis, and freeze drying the partially 2-*O*-desulfated heparin derivative was obtained in a 80% yield.

**VI.2. Preparation of glycol-split derivatives of heparins and LMWHs**

Glycol-split heparins were prepared by periodate oxidation followed by borohydride reduction of parent heparin/LMWH samples (40 – 50 mg). All the material was dissolved in 3 ml of 0.1M $NaIO_4$ and, then the reaction mixture was stirred at 4°C overnight in the dark. The excess of periodate was neutralized by adding 150 µl of ethylene glycol maintaining the stirring for further 2–3 h at 4°C. Solid sodium borohydride (~70 mg) was added in portions to the reaction mixture, which was then kept for 16 h at 4°C in the dark under stirring. After 3h the pH was adjusted to 4 with 0,1M HCl, and the solution was then neutralized with 1M NaOH. For gs-LMWH preparation the pH value at this step was adjusted under stirring at 4°C to 7

with 5% (v/v) acetic acid. Prepared gs-derivatives were desalted on a Sephadex G-10 column (80 x 2.5 cm) using water – ethanol 9:1 (v/v) as a mobile phase at 2.6 ml/min, monitoring the absorbance at 210 nm (Cary 50 UV-Vis spectrophotometer, Varian). After freeze drying, gs-samples were obtained with yields in the range of 80-92 %. The gs-derivative of HA PMH was prepared starting from 4 mg of HA PMH prepared as described above.

The *Smith degradation* of a *gs-PMH* sample was performed by adding 1M HCl to lower to 3 the pH value of the sample solution. After 3h of incubation at room temperature, the pH was increased to 7 by addition of 5% NaOH. The sample was desalted by dialysis using a membrane with 100-500 Da cut-off, freeze-dried, and then analyzed by SEC-TDA and NMR.

### VI.3. Molecular weight determination

Molecular weight determinations were performed by HP-SEC-TDA on a Viscotek (Houston, Texas) instrument equipped with a VE1121 pump, Rheodyne valve (100 µl), and triple detector array 302 equipped with refraction index (RI), viscometer, and light-scattering (90° and 7°) systems. A separation system, including G2500 and G3000 (7.8 mm x 30 cm TSK GMPWxl) Tosoh columns, was used with 0.1M $NaNO_3$ as eluent (flow, 0.6 ml/min). Samples were dissolved in the eluent to obtain 5-10 mg/ml solution. Peak integration and data processing were performed by using OMNISEC 4,1 software. All molecular weight parameters (number-average molecular mass ($M_n$), weight-average molecular mass ($M_w$), and polydispersity ($M_w/M_n$, D)) were determined for each sample.

Hydrodynamic parameters were determined using the same HP-SEC-TDA system and 0,01M $NaNO_3$ as eluent in order to decrease the ionic strength and to better discriminate between different heparin samples.

### VI.4. NMR analysis of parent heparin/LMWH samples and their gs-derivatives

The samples (5 – 8 mg of parent heparin/LMWHs and their gs-derivatives) were dissolved in 600 µl of $D_2O$, containing 2 mM solution of EDTA, then freeze dried. 500 µl of

$D_2O$ were added to each lyophilized sample and lyophilization was repeated. Spectra were recorded at 25°C on a Bruker Avance 500 MHz or on a Bruker Avance 600 MHz spectrometer (Karlsruhe, Germany). Both instruments were equipped with 5-mm TCI cryoprobe. Integration of peak volumes in the HSQC spectra was made using standard Bruker TOPSPIN 3.0 software. The relative content of monosaccharide residues was calculated from the corresponding anomeric cross-peaks (*Guerrini* et al., 2005) and to LMWHs (*Guerrini* et al., 2007).

HSQC spectra were obtained in phase-sensitive, sensitivity pure-absorption mode with decoupling in the acquisition period (Bruker pulse program hsqcetgpsisp.2). The $J_{C-H}$ coupling constant was set to 150 Hz. Spectra were recorded using a spectral width of 8 ppm and 80 ppm in the proton and carbon dimensions, respectively. The carrier frequencies for proton and carbon were 4.7 ppm and 80 ppm. Spectra were acquired into a time domain of 1024 complex points, using from 16 to 32 scans for each of 320 increments. The repetition time and the acquisition time were set at 2s and 0.178s, respectively. The matrix size of 1024 × 320 data points was zero filled to 4096 × 2048 and a squared cosine function was applied prior to Fourier transformation.

One-bond heteronuclear coupling constants $^1J_{C-H}$, measured for gsI and gsG residues present in a gs-PMH sample, were determined by removing the carbon decoupling during acquisition of the HSQC sequence.

2D-TOCSY spectra was acquired using 32 scans per series of 2048 × 512 data points and a mixing time of 80ms. A zero filling in F1 (4096 × 2048) and a shifted ($\pi$/3) squared cosine function was applied prior to Fourier transformation. HMBC spectra were acquired in phase-sensitivity enhanced pure-absorption mode (Bruker pulse program hmbcetgpl2nd). Spectra were recorded using a long range $J_{C-H}$ coupling of 8 Hz. Spectral widths of 8 ppm and 190 ppm were set for the proton and carbon dimensions, respectively. The carrier frequencies for proton and carbon were 4.7 ppm and 95 ppm. Spectra were acquired into a time domain of 1024 complex points, using 64 scans for each of 360 increments. The repetition time and the

acquisition time were set at 2.5 s and 0.107 s, respectively. The matrix size of 1024 × 360 data points was zero filled to 4096 × 2048 prior to Fourier transformation.

### VI.5. LC-MS analysis of the heparinase-digests of gs-heparins and gs-LMWHs

*Enzymatic cleavage of original heparins and their gs-derivatives*

The original heparin and gs-heparin samples (100 μg) were enzymatically depolymerised by heparin lyases I, II and III. The substrate (5-10 mg) was dissolved in water to obtain a 20 mg/ml solution. Then 149 μl of solution containing 50 mM sodium acetate and 5 mM calcium acetate, and 3 μl of heparin lyases mixture (1 μl of each lyase, 2 mU/μl enzyme solution) were added to 5 μl of the substrate 20 mg/ml solution. The reaction was stirred at 37°C (Termo shaker TS-100 Biosan) for 24 h, then, stopped by adding 3 μl of 3% formic acid. Each sample was diluted two times with water and analyzed by LC-MS.

*IPRP-HPLC/ESI-TOF-MS analysis*

LC-MS analysis was performed on a LC system (Dionex Ultimate 3000, Dionex) equipped with degasing system (model LPG-3400), pump (model LPG-3400A), autosampler (model WPS-3000TSL) and UV-detector (model VWD-3100) and coupled with an ESI-QTOS mass-spectrometer (microqTOF, Bruker Daltonics).

The chromatographic separation was performed using a Kinetex C18 analytical column (100 × 2.1 mm I.D., 2.6 μm particle size, Phenomenex) with Security Guard Cartridges Gemini C18 (4 x 2.0 mm, Phenomenex). A binary solvent system was used for gradient elution. Solvent A (10 mM DBA, 10 mM $CH_3COOH$ in water:methanol (9:1)) and solvent B (10 mM DBA and 10 mM $CH_3COOH$ in methanol) were delivered at 0.1 ml/min. Oligosaccharides were separated using a multi-step gradient: the solvent composition was held at 17% B for the first ten minutes, then increased to 42% B over 20 min, and to 50 % B over another 20 min; afterwards, the content of the eluent B was increased to 90 % where it was held for 10 min; finally, it was returned to 17% B over 1 min, and held for the last 19 min for equilibrating the

chromatographic column before the injection of the next sample. The injection volume was 5 µl. The MS spectrometric conditions were as follows: ESI in negative ion mode, drying gas temperature +180°C, drying gas flow-rate 7.0 l/min, nebulizer pressure 0.9 bar; and capillary voltage +3.2 kV. The mass spectra of the oligosaccharides were acquired in a scan mode (*m/z* scan range 200 – 2000). An external calibration method using sodium formate clusters as calibrants was used to achieve an accuracy of better than 5 ppm.

The quantitative analysis of the unsaturated disaccharides has been performed using calibration curves obtained with the standard solutions that were prepared by diluting the disaccharides stock solutions (10 mg/ml in $H_2O$). The regression equations, correlation coefficients, limit of detection (LOD), limit of quantification (LOQ) and area variability (RSD) are given in Table 3.7. LOD and LOQ values were calculated as the concentration of the injected sample (with constant injection volume, 5 µl) which gave signal-to-noise ratio equal to 3 and 10, respectively, using the calibration curves based on the linear regression between signal-to-noise ratio (S/N) value and analyte concentration. Despite of the lower LOQ values provided by MS, for quantitative analyses the calibration curves obtained with UV detection was used because the area variability is lower than for the MS detection mode (Table 3.7).

For the digest of the parent PMH $\Delta U\text{-}A_{NS}$ and $\Delta U\text{-}A_{NAc6S}$ content were also evaluated. Since their extinction coefficients at 232 nm were shown to be equal under the applied analytical conditions, it was possible to quantify their total content, even if eluted together, using the same calibration curve. The content of $\Delta U\text{-}A_{NS6S}$, present in the PMH digest, was determined using the calibration curve of its positional isomer $\Delta U_{2S}\text{-}A_{NS}$.

### VI.6. Direct LC-MS analysis of LMWHs and gs-LMWHs

For profiling the intact unmodified LMWHs and gs-LMWHs, methanol in the mobile phases, used for the analysis of the enzymatic digests (see VI.5), was substituted by acetonitrile (IV.3.1). Thus, for profiling the non-digested original and gs-LMWHs, mobile phases A (10

mM dibutylamine, 10 mM acetic acid in water – acetonitrile (9:1)) and B (10 mM dibutylamine and 10 mM acetic acid in acetonitrile), delivered at 0.25 ml/min, were used. Since enoxaparin, tinzaparin and dalteparin are characterized by a different molecular weight distribution, slightly different multistep gradients were used for their profiling (Table 6.1). Under the applied conditions the retention times vary in a range 1.0 – 1.5 %. The MS conditions were the same as for LC-MS analysis of the heparinase-digests (see VI.5).

**Table 6.1.** Gradient used and the LMWH quantity loaded for LC-MS profiling of the LMWHs and their gs-derivatives

| **Tinzaparin and gs-tinzaparin** | | **Enoxaparin and gs-enoxaparin** | | **Dalteparin and gs-dalteparin** | |
|---|---|---|---|---|---|
| Multi-step gradient | | | | | |
| Time (min) | % Eluent B | Time (min) | % Eluent B | Time (min) | % Eluent B |
| | | | | 0 | 10 |
| 0 | 5 | 0 | 0 | 4 | 10 |
| 5 | 5 | 5 | 0 | 15 | 18 |
| 125 | 35 | 130 | 35 | 30 | 18 |
| 130 | 70 | 140 | 70 | 115 | 35 |
| 135 | 70 | 145 | 70 | 133 | 70 |
| 138 | 5 | 148 | 0 | 138 | 70 |
| 160 | 5 | 170 | 0 | 140 | 10 |
| | | | | 160 | 10 |
| Sample quantity loaded, µg | | | | | |
| 15 | 15 | 30 | 30 | 30 | 30 |
| Corresponding Figure | | | | | |
| Fig. 4.10 and Annex 5 | | Fig. 4.11 and Annex 6 | | Fig. 4.12 and Annex 7 | |

Column C18 100 x 2.1 mm, 2.6 µm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; flow rate – 0.25 ml/min

### VI.7. Study of susceptibility of enoxaparin RE groups to periodate oxidation

Gs-enoxaparin was digested with heparinases I, II and III also on preparative scale for further fractionation and detailed studies by LC-MS and NMR: 175 µl of 2 mU/µl solutions of each enzyme were added to a gs-enoxaparin solution (37 mg in 50 ml of a buffer containing 50 mM sodium acetate buffer and 5 mM calcium acetate). The digestion was carried out for 24 h at 37°C, and after the second addition of an equivalent aliquot the enzymes, the reaction was maintained for further 24 h at 37°C, then, stopped by adding 1 ml of 3% formic acid. The

reaction mixture was filtered using a Millipore filter 3 μm and concentrated up to 2 ml for SEC fractionation (*see next section*) and further detailed LC-MS and NMR analysis.

*Preparative size-exclusion chromatography*

HPLC system (Bioline Knauer S1050) was used for the size fractionation of the heparinase-digested gs-enoxaparin. Oligosaccharide mixture (2 ml of the concentrated solution, *see the previous section*) was separated using two subsequent columns (Superdex S30, 48×3 and 83×2 cm) and 0.25M ammonium chloride as eluent, at a flow-rate 5.0 ml/min. Based on the detection at 210 nm, seven fractions were obtained and, then, desalted on the Biofox 40/100 SEC Agarose column (89 x 3 cm) using water – ethanol 9:1 (v/v) as eluent delivered at a flow-rate 15 ml/min. After concentration and freeze-drying each gs-oligosaccharide fraction was analyzed by NMR and LC-MS.

*NMR spectra* were acquired on a Bruker Avance 600 MHz spectrometer (Karlsruhe, Germany) using 3 mm NMR tubes (each fraction was prepared in 220 μl $D_2O$).

*LC-MS analysis* was performed as described in VI.5 for total enzymatic digests. The fragmentation of the ion with *m/z* 488.5 (Fig.4.15), present in the tetrasaccharide fraction of digested gs-enoxaparin, was performed by applying the LC-ESI-$MS_2$ method. Collision-induced dissociation (CID) was used, the isolation window was set at 5 Da and the collision energy was -20 eV.

*Sodium borohydride reduction of the digested gs-enoxaparin fractions*

The reduction of the several fractions obtained by SEC-fractionation of the digested gs-enoxaparin (as described above) were performed by adding of 5 μl of 30 mg/ml solution of sodium borohydride into an aliquot (50 μl) of 0.5 mg/ml solution of each fraction. The mixture was left at room temperature for 4 h and, then, analyzed by LC-MS.

*Multi-step isolation of enoxaparin tetrasaccharide fractions: pentasulfated 1,6-anhydro- and hexasulfated regular tetrasaccharides*

The first-step fractionation was performed using size exclusion chromatography (Fig.4.16). 2 ml of 120 mg/ml solution were loaded onto the column packed with Biogel P6 and eluted with 0.25 M ammonium chloride at a flow-rate 1.8 ml/min. Based on the detection at 210 nm, the tetrasaccharide fraction was isolated and then desalted on a TSK gel HW40S column (100 x 5 cm) using water – ethanol 9:1 (v/v) as a mobile phase delivered at a flow-rate 4.8 ml/min. After desalting, the sample was concentrated under reduced pressure at 35°C and freeze-dried. The obtained fraction (~6 mg) was then sub-fractionated by IPRP-HPLC (Fig.4.16) in order to isolate the isomeric tetrasaccharides of interest. The separation was carried out on a preparative HPLC/UV system "Knauer Smartline 1000", using conditions described in Table 6.2. Three runs were performed.

**Table 6.2.** IPRP-HPLC/UV conditions for isolation of pentasulfated 1,6-anhydro- and hexasulfated tetrasaccharides from the total tetrasaccharide fraction of enoxaparin

| | |
|---|---|
| Column | C18 Knauer 250 x 8 mm, 5 µm |
| Eluents | Eluent A: 10 mM DBA, 10 mM $CH_3COOH$ in $H_2O$-$CH_3OH$ = 43:57 (v/v)<br>Eluent B: 10 mM DBA, 10 mM $CH_3COOH$ in $H_2O$-$CH_3OH$ = 50:50 (v/v)<br>Eluent C: 10 mM DBA, 10 mM $CH_3COOH$ in $H_2O$-$CH_3OH$ = 80:20 (v/v) |
| Gradient | 0 – 45 min isocratic step 100 % A<br>45 – 45.2 min linear gradient from 100%A to 100%B<br>45.2 – 60 min isocratic step 100 %B<br>60 – 60.2 min linear gradient from 100%B to 100%C<br>60.2 – 65 min isocratic step 100 %C<br>65 – 65.2 min linear gradient from 100%C to 100%A<br>65.2 – 90 min isocratic step 100 % A (column equilibrating) |
| Flow | 2,5 ml/min |
| UV detection | 232, 210 nm |
| Sample amount | 1 mg |

A fraction eluted in the range of RT 50 – 56 min (Fig.4.16) was concentrated; then, the solution was dialyzed using a 100-500 Da cut-off membrane in order to eliminate the excess of the DBA acetate. The same purification procedure was performed for the fraction eluted in the

range 60 – 65 min (Fig.4.16). After desalting, the samples was freeze-dried and characterized by NMR and LC-MS (Fig.4.17) before further study. The sample from the first fraction contained pentasulfated 1,6-anhydro-tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$A_{NS}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-1,6a$M_{NS}$ (~1 mg), while from the second one a mixture of the hexasulfated tetrasaccharides $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$A_{NS6S}$ and $\Delta U_{2S}$-$A_{NS6S}$-$I_{2S}$-$M_{NS6S}$ (~4 mg) was obtained.

*Periodate oxidation/borohydride reduction of the enoxaparin tetrasaccharide mixtures*

All the content (~1 mg) of the first fraction containing 1,6-anhydro-tetrasaccharides was dissolved in 230 µl of $D_2O$, and 5 mg of sodium periodate were added directly into the NMR tube. The final concentration of periodate ion corresponded to 0.1 M. The reaction was monitored by monodimensional NMR as described below. After 40 h, 2 µl of the solution containing the tetrasaccharide mixture and periodate was diluted with 50 µl of water and directly analyzed by LC-MS (*see below*). 12 mg of sodium borohydride were added to the rest of the solution to reduce the formed aldehyde groups. The reaction was carried out for 20 h in the dark, at 4°C. The sample was then desalted using a 100-500 Da cut-off membrane and freeze-dried for further structural characterization by NMR and LC-MS.

The oxidation of the hexasulfated tetrasaccharide-containing mixture was also performed directly in the NMR tube (3 mm). About 1 mg of the material was dissolved in 230 µl of $D_2O$ and 5 mg of sodium periodate (0,1 M final concentration) were added. Since no oxidation was observed after 20 h, an additional portion of periodate (5 mg of sodium periodate) was added (the final concentration of periodate ion was 0.2 M). The final $^1H$ and 2D HSQC spectra as well as LC-MS chromatograms were acquired after 40 h. LC-MS analysis was repeated after 5 days. The further reduction was carried out as described above for 1,6-anhydro-tetrasaccharides.

The periodate oxidation of $\Delta U_{2S}$-$A_{NS6S}$ (isolated from the gs-enoxaparin digest, fraction H in Fig.4.14; see VI.7) was performed by adding different volumes (230 and 460 µl) of 0.1 M sodium periodate to 1 mg of the disaccharide. 230 µl of 0.1 M correspond to 5 mg of periodate,

used for the oxidation of tetrasaccharides. After 20 h the LC-MS analysis was performed as described below.

The oxidation of the total enoxaparin sample was performed in batch by adding 325 µl of 0.1 $NaIO_4$ to a portion of the solid material (~5.2 mg) and left at 4°C in the dark for different time periods (1, 2, 3, 4, 6, 8, 10 and 12 h). The oxidation was stopped by adding 25 µl of ethylene glycol, and after further 3 h the solution was treated by sodium borohydride. Each sample was desalted using a 100-500 Da cut-off membrane and freeze-dried for further HSQC NMR analysis.

*NMR study*

$^1H$ monodimensional NMR spectra of tetrasaccharide mixture were recorded at 30°C on a Bruker Avance 600 MHz spectrometer equipped with 5 mm TCI cryoprobe (Karlsruhe, Germany), the $^1H$ spectra were acquired with 128 scans. Water presaturation was applied during each 12 s of relaxation delay. To follow the kinetic of the reaction, $^1H$ spectra were performed as pseudo-2D NMR experiments, with 36 increments and 33 min of interval between experiments. Two-dimensional NMR was performed as described in VI.4.

*LC-MS analysis*

LC-MS operating conditions used for the analysis of the isolated tetrasaccharides before and after oxidation/reduction (Fig.4.17, 4.22, 4.24) are reported in Table 6.3.

Analysis of the undesalted tetrasaccharide solution after the periodate oxidation, carried out directly in the NMR tube (Fig. 4.19), was performed using a modified gradient in order to selectively elute salt clusters (mainly periodate salt) and avoid the suppression of the signals. Initially, 0 % of eluent B was held for the first 15 min to eliminate the excess of the salt, then, a linear gradient from 0 to 40 %B was applied to elute the oligosaccharides. LC/MS conditions optimized for the heparinase-digested gs-heparins (VI.5) were used for the analysis of the disaccharide before and after its oxidation.

**Table 6.3.** LC-MS operating conditions for analysis of pentasulfated 1,6-anhydro- and hexasulfated tetrasaccharides of enoxaparin

| | |
|---|---|
| Column | C18 (Eurospher II 100-5 C18 (250 x 4,6 mm, 5 µm, Knauer) |
| Eluents | Eluent A 10 mM DBA, 10 mM $CH_3COOH$ in $H_2O$-$CH_3OH$ = 9:1 (v/v)<br>Eluent B 10 mM DBA, 10 mM $CH_3COOH$ in $CH_3OH$ |
| Gradient | 0 – 18 min isocratic step 33 % B<br>18 – 19 min linear gradient from 33 to 35%B<br>19 – 65 min isocratic step 33 %B<br>65 – 66 min linear gradient from 35 to 45%B<br>66 – 80 min isocratic step 45 %B<br>80 – 82 min linear gradient 80%B<br>82 – 84 min isocratic step 80%B<br>84 – 85 min linear gradient from 80 to 33 %B<br>85 – 105 min isocratic step 33%B |
| Flow | 1 ml/min, splitted 1:2 before ESI-source |
| UV detection | 232 nm |
| ESI-MS conditions | negative ion mode<br>capillary voltage +3.2 kV<br>drying gas 180°C, 7.0 l/min<br>nebulizer pressure 0.9 bar<br>*m/z* scan range 200 – 2000 |

### VI.8. Preparation of oligosaccharides enriched in internal glycol-split residues

The starting HI-3 sample was prepared by glycol-splitting of a partially 2-*O*-desulfated PMH as described in VI.1 and VI.2. Its enzymatic depolymerisation was performed using ~60 mg of substrate maintaining the same proportions substrate–heparinases as used for gs-enoxaparin digestion (VI.7). The digestion mixture was further fractionated by SEC as described in VI.7 for gs-enoxaparin digest. The isolated fractions, analyzed by NMR and LC-MS, contained oligosaccharides of dp 4, 6, 8, 10 with internal gs-residues (Fig.5.1).

**Acknowledgements**

I would like to gratefully and sincerely thank many people accompanying me along the long way that led me where I am today. I would like to, first and foremost, thank my mom *Elena Alekseeva* who helped me to choose this way and to follow it, who has always encouraged me and been near me in both difficult and happy moments. I am also deeply grateful to my sister *Masha* for her love and faith. Many sincere thanks to my partner *Marco* for his patience and unfailing support, so priceless especially during the final stages of the Ph.D thesis.

I would like to give a sincere thank you to my school teacher and undergraduate and Ph.D supervisor Prof. *Anna Kartsova* for making chemistry attractive and for guiding me during my undergraduate study and further Ph.D school at the Saint-Petersburg State University. Today I am aware that the skills, I had developed throughout that period, helped me in carrying out and accomplishing the present Ph.D research.

I would like to sincerely thank my supervisors, Dr. Annamaria Naggi and Prof. Luigi Lay, for scientific advisement and always useful and valuable comments and interesting result discussions. I would like to give a special thank you to Dr. Giangiacomo Torri for giving me the opportunity to carry out my Ph.D research at the “Ronzoni” Institute.

My Ph.D research would have been much more difficult without the contributions of several people. I would like to particularly thank Prof. Benito Casu e Dr. Giuseppe Cassinelli for their passionate participation, valuable suggestions and critical reading of my Ph.D dissertation. I would like also to thank Antonella Bisio, Marco Guerrini, Sabrina Bertini for continuous encouragement and constructive comments on my Ph.D thesis and presentations. Many grateful thanks to Cesare Cosentino, Eleonora Macchi, Lucio Mauri, Laura Muzi for performing NMR measurements, Stefano Elli for molecular modelling experiments and helpful result discussion, Elena Urso for technical and moral support. Many sincere thanks to my lab-neighbors for providing a great work environment. In particular, I would like to sincerely thank all of them for their friendship and unending support.

Last but not least, I would like to thank all the staff of Università degli Studi di Milano, especially Prof. Silvia Ardizzone and Prof. Emanuela Licandro, for organizing interesting courses and workshops, their assistance, availability and support, during the doctorate school.

This accomplishment would not have been possible without all of you. Thank you.

## REFERENCES


*Afratis* N., Gialeli C., Nikitovic D., Tsegenidis T., Karousou E., Theocharis A. D., Pavão M.S., Tzanakakis G. N., Karamanos N. K. Glycosaminoglycans: key players in cancer cell biology and treatment. *FEBS J.* 279 (2012) 1177-1197.

*Babor* K., Kaláč V., Tinlárik K. Periodate oxidation of saccharides. Comparison of the methods for determining the consumption of sodium periodate and the amount of formic acid formed. *Chem.zvesti*. 27 (1973) 676-680.

*Baer* H.H., Astles D.J., Chin H-C., Siemsen L. The formation of branched-chain deoxypentofuranosides by ring contraction in the reductive desulfonyloxylation of hexopyranoside p-toluenesulfonates. *Can. J. Chem.* 63 (1985) 432-439.

*Behr* J.R.; Matsumoto Y.; White F.M.; Sasisekharan R., Quantification of isomers from a mixture of twelve heparin and heparan sulfate disaccharides using tandem mass spectrometry. *Rapid Commun. Mass Spectrom*. 19 (2005) 2553-2562.

*Bernfield* M., Götte M., Park P.W., Reizes O., Fitzgerald M.L., Lincecum J., Zako M. Functions of cell surface heparan sulphate proteoglcans. *Annu. Rev. Biochem.* 68 (1999) 729-777.

*Bertini* S., Bisio A., Torri G., Bensi D., Terbojevich M. Molecular weight determination of heparin and dermatan sulphate by size exclusion chromatography with a triple detector array. *Biomacromolecules.* 6 (2005) 168-173.

*Bisio* A., Mantegazza A., Urso E., Naggi A., Torri G., Viscov C., Casu B. High-performance liquid chromatographic/mass spectrometric studies on the susceptibility of heparin species to cleavage by heparanase. *Semin. Thromb. Hemost*. 33 (2007) 488-495.

*Bisio* A, Vecchietti D., Citterio L., Guerrini M., Raman R., Bertini S., Eisele G., Naggi A., Sasisekharan R., Torri G. Structural features of low-molecular weight heparins affecting their affinity to antithrombin. *Thromb. Haemost*. 102 (2009) 865-873.

*Buszewski* B. and Noga S. Hydrophilic interaction liquid chromatography (HILIC) – a powerful separation technique. *Anal. Bioanal. Chem*. 402 (2012) 231-247.

*Camara* J.E., Satterfield M.B., Nelson B.C. Quantitative determination of disaccharide content in digested unfragmented heparin and low molecular weight heparin by direct-infusion electrospray mass spectrometry. *J. Pharm. Biomed. Anal*. 43 (2007) 1706-1714.

*Capila I*, Linhardt R. Heparin – protein interactions. *Angew. Chem*. Int.Ed. 41 (2002) 390-412.

*Carlsson P.,* Kjellén L. Heparin Biosynthesis In Heparin: a century of progress. Springer: New York. (2012) 23-41.

*Casu* B., Diamantini G., Fedeli G., Mantovani M., Oreste P., Pescador R., Porta R., Prino G., Torri G., Zoppetti G. Retention of antilipemic activity by periodate-oxidized non-anticoagulant heparins. *Arzneimittelforschung.* 36 (1986) 637-642.

*Casu* B. In Heparin: Chemical and biological properties, Clinical Applications; Lane D.A., Lindahl U., Eds; Edward Arnold: London (1989) 25-49.

Casu B., Guerrini M., Naggi A., Torri G., De Ambrosi L., Boveri G., Gonella S., Ferro L., Lanzarotti E., Paternò M., Attolini M., Valle M. G. Characterization of sulfation patterns of beef and pig mucosal heparins by nuclear magnetic resonance spectroscopy. *Arzneim.-Forsch. Drug Res*. 46 (1996) 472-477.

*Casu* B., Lindahl U, Structure and biological interactions of heparin and heparan sulfate. *Adv. Carbohydr. Chem. Biochem*. 57 (2001) 159-206.

*Casu* B., Guerrini M., Naggi A., Perez M., Torri G., Ribatti D., Carminati P., Giannini G., Penco S., Pisano C., Belleri M., Resnati M., Presta M. Short heparin sequences spaced by glycol-split uronate residues are antagonists of fibroblast growth factor 2 and angiogenesis inhibitors. *Biochemistry*. 41 (2002) 10519-10528.

*Casu* B., Naggi A. Antiangiogenic heparin-derived heparan sulfate mimics. *Pure Appl. Chem*. 75 (2003) 157–166.

*Casu* B., Guerrini M., Guglieri S., Naggi A., Perez M., Torri G., Cassinelli G., Ribatti D., Carminati P., Giannini G., Penco S., Pisano C., Belleri M., Resnati M., Presta M. Undersulfated and glycol-split heparins endowed with antiangiogenic activity. *J. Med. Chem.* 47 (2004) 838-848.

*Casu* B. Structure and active domains of heparin. In: H.G. Garg, R.J. Linhardt, C.A. Hales (ed.s) Chemistry and biology of heparin and heparan sulfate, Elsevier, Amsterdam (2005) 1-28.

*Casu* B., Naggi A., Torri G. Heparin-derived heparan sulfate mimics to modulate heparan sulfate-protein interaction in inflammation and cancer. *Matrix Biol.* 29 (2010) 442-452.

*Conrad* H.E., Guo Y. Structural analysis of periodate-oxidized heparin, in: D.A. Lane, I. Björk, U. Lindahl (Eds.), Advances in Experimental Medicine and Biology, 313, Plenum, New York (1992) 31–36.

*Desai* U.R., Linhardt R.J. Molecular weight of heparin using $^{13}C$ nuclear magnetic resonance spectroscopy. *J. Pharm. Sci*. 84 (1995) 212-215.

*Doneanu* C.E., Chen W., Geber J.C. Analysis of oligosaccharides derived from heparin by ion-pair reverse-phase chromatography/mass spectrometry. *Anal. Chem.* 81 (2009) 3485-3499.

*Dryhurst* G. Periodate oxidation of diol and other functional groups. Analytical and structural applications, 1st ed., Pergamon press: London (1970).

*Ekman-Ordeberg* G., Hellgren M., Åkerud A., Andersson E., Dubicke A., Sennström M., Byström B., Tzortzatos G., Gomez M.F., Edlund1 M., Lindahl U., Malmström A. Low molecular weight heparin stimulates myometrial contractility and cervical remodeling in vitro. *Acta Obstetr. Gynecol. Scand.* 88 (2009) 984-989.

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

*Fairbanks M.B.,* Mildner A.M., Leone J.W., Cavey G.S., Mathews W. R., Drong R. F., Slightom J. L., Bienkowski M. J., Smith C. W., Bannow C. A., Heinrikson R. L. Processing of the human heparanase precursor and evidence that the active enzyme is a heterodimer. *J. Biol. Chem*. 274 (1999) 29587-29590

*Ferro* D.R., Provasoli A., Ragazzi M., Torri G., Casu B., Gatti G., Jacquinet J.-C., Sinay P., Petitou M., Choay J. Evidence for conformational equilibrium of the sulfated L-iduronate residue in heparin and in synthetic heparin mono- and oligosaccharides: NMR and force-field studies. *J. Am. Chem. SOC*. 108 (1986) 6773-6778.

*Ferro* D.R., Provasoli A., Ragazzi M., Casu B., Torri G., Bossennec V., Perly B., Sinay P., Petitou M., Choay J. Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences. *Carbohydr. Res*. 195 (1990) 157-167.

*Ferro* V., Dredge K., Liu L., Hammond E., Bytheway I., Li C., Johnstone K., Karoli T., Davis K., Copeman E., Gautam A. PI-88 and novel heparan sulphate mimetics inhibit angiogenesis. *Seminars Thromb. Hemost*. 33 (2007) 557-568.

*Ferro* V., Liu L., Johnstone K.D., Wimmer N., Karoli T., Handley P., Rowley J., Dredge K., Li C.P., Hammond E., Davis K., Sarimaa L., Harenberg J., Bytheway J. Discovery of PG545: a highly potent and simultaneous inhibitor of angiogenesis, tumor growth, and metastasis. *J. Med. Chem*. 55 (2012) 3804-3813.

*Ferro* M. Master degree thesis "Pentasaccaridi sintetici a struttura glicosamminoglicanica, potenziali inibitori dell'attività enzimatica dell'eparanasi", Università degli Studi di Milano (2009).

*Fu* L., Li G., Yang B., Onishi A., Li L., Sun P., Zhang F., Linhardt R.J. Structural characterization of pharmaceutical heparins prepared from different animal tissues. *J. Pharm. Sci.* 102 (2013) 1447-1457.

*Fuster* M.M., Esko J.D. The sweet and sour of cancer: glycans as novel therapeutic agents. *Nat. Rev. Cancer*. 5 (2005) 526-542.

*Galliher* P.M., Cooney C.L., Langer R., Linhardt R.J. Heparinase production by *Flavobacterium* heparinum. *Appl. Environ. Microbiol*. 41 (1981) 360–365.

*Gong* F., Jemth P., Escobar Galvis M.L., Vlodavsky I., Horner A., Lindahl U., Li J.P., Processing of macromolecular heparin by heparanase. *J. Biol. Chem*. 278 (2003) 35152–35158.

*Gray* E., Mulloy B., Barrowcliffe T.W. Heparin and low-molecular weight heparin. *Thromb. Haemost*. 99 (2008) 807-818.

*Guerrini* M, Bisio A., Torri G. Combined quantitative $^{1}H$ and $^{13}C$-NMR spectroscopy for characterization of heparin preparations. *Semin. Thromb. Hemost*. 27 (2001) 473-482.

*Guerrini* M., Naggi A., Guglieri S., Santarsiero R., Torri G. Complex glycosaminoglycans: profiling substitution patterns by two-dimensional nuclear magnetic resonance spectroscopy. *Anal. Biochem.* 337 (2005) 35-47.

*Guerrini* M., Guglieri S., Beccati D., Torri G., Viskov C., Mourier P. Conformational transitions induced in heparin octasaccharides by binding with antithrombin III. *Biochem J.* 399 (2006) 191-198.

*Guerrini* M., Guglieri S., Naggi A., Sasisekharan R., Torri G. Low molecular weight heparins: structural differentiation by bidimentional nuclear magnetic resonance spectroscopy. *Semin. Thromb. Hemost*. 33 (2007) 478-487.

*Guerrini* M., Guglieri S., Casu B., Torri G., Mourier P., Boudier C., Viskov C. Antithrombin-binding octasaccharides and role of extensions of the active pentasaccharide sequence in the specificity and strength of interaction. Evidence for very high affinity induced by an unusual glucuronic acid residue. *J. Biol. Chem*. 283 (2008) 26662-26675.

*Guglieri* S., Hricovini M., Rahul R., Polito L., Torri G., Casu B., Sasisekharan R., Guerrini M. Minimum FGF2 binding structural requirements of heparin and heparan sulfate oligosaccharides as determined by NMR spectroscopy. *Biochemistry.* 47 (2008) 13862–13869.

*Guo* Y.C. and Conrad H.E. The disaccharide composition of heparins and heparan sulfates *Anal. Biochem*. 176 (1989) 96-104.

*Jandik* K.A., Kruep D., Cartier M., Linhardt R.J. Accelerated stability studies of heparin. *J. Pharm. Sci.* 85 (1996) 45-51.

*Jones* C.J., Membreno N., Larive C.K. Insights into the mechanism of separation of heparin and heparan sulphate disaccharides by reverse-phase ion-pair chromatography. *J. Chromatogr. A*. 1217 (2010) 479–488.

*Henriksen* J., Ringborg L. H., Roepstorff P.J. On-line size-exclusion chromatography/mass spectrometry of low molecular mass heparin. *J. Mass spectrom*. 39 (2004) 1305.

*Höök* M., Björk I., Hopwood J., Lindahl U. Anticoagulant activity of heparin: separation of high-activity and low-activity heparin species by affinity chromatography on immobilized antithrombin. FEBS Lett. 66 (1976) 90-93.

*Horton* D. and Philips K.D. The nitrous acid deamination of glycosides and acetates of 2-amino-2-deoxy-D-glucose. *Carbohydr. Res.* 30 (1973) 367-374.

*Hostetter* N., Naggi A., Torri G., Ishai-Michaeli R., Casu B., Vlodavsky I., Borsig L. P-selectin and heparanase-dependent antimetastatic activity of non-anticoagulant heparins. *FASEB J.* 21 (2007) 3562-3572.

*Huber* C.G and Krajete A. Sheath liquid effects in capillary high-performance liquid chromatography – electrospray mass spectrometry of oligonucleotides. *J. Chromatogr. A*. 870 (2000) 413-424.

*Huckerby* T.N., Sanderson P.N., Nieduszynski I.A. NMR studies of oligosaccharides obtained by degradation of bovine lung heparin with nitrous acid. *Carbohydr. Res*. 154 (1986) 15-27.

*Iacomini* M., Casu B., Guerrini M., Naggi A., Pirola A., Torri G. "Linkage region" sequence of heparin and heparan sulfate: Detection and quantification by Nuclear Magnetic Resonance spectroscopy. *Anal. Biochem*. 274 (1999) 50-58.

*Islam* T., Butler M., Sikkander S.A., Toida T., Linhardt R.J. Further evidence that periodate cleavage of heparin occurs primarily through the antithrombin binding site. *Carbohydr. Res*. 33 (2002) 2239-2243.

*Keire* D.A., Trehy M.L., Reepmeyer J.C., Kolinski R.E., Ye W., Dunn J., Westenberger B.J., Buhse L.F. Analysis of crude heparirn by 1H NMR, capillary electrophoresis, and strong-anion-exchange-HPLC for contamination by oversulfated chondroitin sulfate. *J. Pharmaceut. Biomed. Anal.* 51 (2010) 921-926.

*Khan S.,* Rodriguez E., Patel R., Gor J., Mulloy B., Perkins S.J. The solution structure of heparan sulphate differs from that of heparin: implications for function. *J. Biol. Chem*. 286 (2012) 24842-24854.

*Kjellén* L., Lindahl U. Biosynthesis of heparin and heparan sulfate proteoglycans. *Carbohydr. Chem. Biol.* 3 (2000) 395-405.

*Kreuger* J., Prydz K., Pettersson R.F., Lindahl U., Salmivirta M. Characterization of fibroblast growth factor 1 binding heparan sulfate domain. *Glycobiology*. 9 (1999) 723-729.

*Kuberan* B., Lech M., Zhang L., Wu L.Z., Beeler D.L., Rosenberg R. Analysis of heparan sulfate oligosaccharides with ion pair-reverse phase capillary high performance liquid chromatography-microelectrospray ionization time-of-flight mass spectrometry. *J. Am. Chem. Soc.* 124 (2002) 8707–8718.

*Langeslay* D.J., Urso E., Gardini C., Naggi A., Torri G., Larive C.K. Reversed-phase ion-pair ultra-high-performance-liquid chromatography-mass spectrometry for fingerprinting low-molecular-weight heparins. *J. Chromatogr. A.* 1292 (2013) 201-210.

*Lapierre* F., Holme K., Lam L., Tressler R.J., Storm N., Wee J., Stack R. J., Castellot J., Tyrrell D.J. Chemical modifications of heparin that diminish its anticoagulant but preserve its heparanase-inhibitory, angiostatic, anti-tumor and anti-metastatic properties. *Glycobiology.* 6 (1996) 355-366.

*Lawrence* R., Brown J.R., Al-Mafraji K., Lamanna W.C., Beitel J.R., Boons G.-J., Esko J.D. Crawford B.E. Disease-specific non-reducing end carbohydrate biomarkers for mucopolysaccharidoses. *Nature Chem. Biol*. 8 (2011) 197-2049.

*Lawrence* R., Kuperan B., Lech M., Beeler D. L., Rosenberg R.D. Mapping critical biological motifs and biosynthetic pathways of heparan sulfate*. Glycobiology*. 14 (2004) 467-479.

*Lee* S., Raw A., Yu L., Lionberger R., Ya N., Verthelyi D., Rosenberg A., Kozlowski S., Webber K., Woodcock J. Scientific considerations in the review and approval of generic enoxaparin in the United States. *Nat. Biotechnol*. 31 (2013) 220-226.

*Leitgeb* A.M*.,* Blomqvist K., Cho-Ngwa F., Samje M., Nde P., Titanji V., Wahlgren M. Low anticoagulant heparin disrupts *Plasmodium falciparum* rosettes in fresh clinical isolates. *Am. J. Trop. Med. Hyg*. 84 (2011) 390-396.

*Lever* R., Page C.P. Novel drug development opportunities for heparin. *Nat. Rev. Drug. Discov.* 1 (2002) 140-148.

*Levy-Adam* F., Miao H., Heinrikson R.L., Vlodavsky I., Ilan N. Heterodimer formation is essential for heparanase enzymatic activity. *Biochem. Biophys. Res. Commun*. 308 (2003) 885-891.

*Lindahl* U., Kusche-Gullberg M., Kjellén L. Regulated diversity of heparan sulfate. *J Biol. Chem*. 273 (1998) 24979–24982.

*Linhardt* R.J., Rice K.G., Kim Y.S. Lohse D.L., Wang H.M., Loganathan D. Mapping and quantification of the major oligosaccharide components of heparin. *Biochem. J*. 254 (1988) 781-787.

*Linhardt* R.J., Loganathan D., Al-Hakim A., Wang H.-M., Walenga J.M., Hoppensteadt D., Fareed J. Oligosaccharide mapping of low molecular weight heparins: structure and activity differences. *J. Med. Chem*. 33 (1990) 1639–1645.

*Loganathan* D., Wang H.M., Mallis L.M., Linhardt R.J. Structural variation in the antithrombin III binding site region and its occurrence in heparin from different sources. *Biochemistry*. 29 (1990) 4362–4368.

*Mascellani* G., Guerrini M., Torri G., Liverani L., Spelta F., Bianchini P. Characterization of di- and monosulfated, unsaturated heparin disaccharides with terminal N-sulfated 1,6-anhydro-ß-D-glucosamine or N-sulfated 1,6-anhydro-ß-D-mannosamine residues. *Carbohydr. Res*. 342 (2007) 835-842.

*Mourier* P., Viskov C. Process for oxidizing unfractionated heparins and detecting presence or absence of glycoserine in heparin and heparin products. *European patent application*. EP1582531A1 (2005).

*Mourier* P., Viskov C. Chromatographic analysis and sequencing approach of heparin oligosaccharides using cetyltrimethylammonium dynamically coated stationary phases. *Anal. Biochem*. 332 (2004) 299-313.

*Mourier* P., Guichard O.Y., Herman F., Viskov C. Isolation of a pure octadecasaccharide with antithrombin activity from an ultra-low-molecular-weight heparin. *Anal. Biochem*. 453 (2014) 7–15.

*Mousa* S.A., Linhardt R., Francis J.L., Amirkhosravi A. Anti-metastatic effect of a non-anticoagulant low-molecular-weight heparin versus the standard low-molecular-weight heparin, enoxaparin. *Thromb. Haemost*. 96 (2006) 816–821.

*Mulloy* B., Gee C., Wheeler S.F., Wait R., Gray E., Barrowcliffe T. W. Molecular weight measurements of low molecular weight heparins by gel permeation chromatography. *Thromb. Haemost*. 77 (1997) 668-674.

*Naggi* A. In: Garg H.G., Linhardt R.J., Hales C.A. (ed.s) Chemistry and biology of heparin and heparan sulfate, Elsevier, Amsterdam (2005), *and references therein*.

*Naggi* A., Casu B., Perez M., Torri G., Cassinelli G., Penco S., Pisano C., Giannini G., Ishai-Michaeli R., Vlodavsky I. Modulation of the heparanase-inhibiting activity of heparin through selective desulfation, graded N-acetylation, and glycol-splitting. *J. Biol. Chem*. 280 (2005) 12103-12113.

*Naimy* H., Leumarie N., Bowman M.J., Costello C.E., Zaia J. Characterization of heparin oligosaccharides binding specifically to antithrombin III using mass spectrometry. *Biochemistry.* 47 (2008) 3155-3161.

*Okada* Y., Yamada S., Toyoshima M., Dong J., Nakajima M., Sugahara K. Structural recognition by recombinant human heparanase that plays critical roles in tumor metastasis. *J. Biol. Chem*. 277 (2002) 42488–42495.

*Ozug* J., Wudyka S., Gunay N.S., Beccati D., Lansing J., Wang J., Capila I., Shriver Z., Kaundinya G.V. Structural elucidation of the tetrasaccharide pool in enoxaparin sodium. *Anal. Bioanal. Chem*. 403 (2012) 2733–2744.

*Perlin* A.S. Glycol-cleavage oxidation. Adv. *Carbohydr. Chem. Biochem*. 60 (2006) 183-250.

*Pervin,* A., Gallo, C., Jandik K.A.. Han X.J.. Linhardt R.J. Preparation and structural characterization of large heparin-derived oligosaccharides. *Glycobiology.* 5 *(*1995) 83-95.

*Peterson S.B. and Liu J.* Deciphering mode of action of heparanase using structurally defined oligosaccharides. *J. Biol. Chem*. 287 (2012) 34836-34843.

*Peterson S.B. and Liu J.* Multi-faceted substrate specificity of heparanase. *Matrix Biol*. 32 (2013) 223-227.

*Petitou* M. and van Boeckel C.A. A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? *Angew. Chem. Int. Ed*. 42 (2004) 3118-3133.

*Phillips* P.G., Yalcin M., Cui H., Abdel-Nabi H., Sajjad M., Bernacki R., Veith J., Mousa S.A. Increased tumor uptake of chemotherapeutics and improved chemoresponse by novel non-anticoagulant low molecular weight heparin. *Anticancer Res*. 31 (2011) 411-420.

*Piani* S., Casu B., Marchi E., Torri G., Ungarelli F., Barbanti M. Alkali-induced optical rotation changes in heparins and heparan sulfates, and their relation to iduronic acid-containing sequences. *J. Carbohydr. Chem*. 12 (1993) 507–521.

*Pisano* C., Aulicino C., Vesci L., Casu B., Naggi A., Torri G., Ribatti D., Belleri M., Rusnati M., Presta M. Undersulfated, low-molecular-weight glycol-split heparin as an antiangiogenic VEGF antagonist*. Glycobiology*. 15 (2005) 1C-6C.

*Pisano* C., Vlodavsky I., Ilan N., Zunino F. The potential of heparanase as a therapeutic target in cancer. *Biochem Pharmacol*. 89 (2014) 12-19.

*Powell* A.K., Yates E.A., Fernig D.G., Turnbull J.E. Interactions of heparin/heparan sulfate with proteins: Appraisal of structural factors and experimental approaches. *Glycobiology.* 14 (2004) 17R-30R.

*Razi* N., Kreuger J., Lay L., Russo G., Panza L., Lindahl B., Lindahl U. Identification of O-sulphate substituents on D-glucuronic acid units in heparin-related glycosaminoglycans using novel synthetic disaccharide standards. *Glycobiology*. 5 (1995) 807-811.

*Ritchie* J.P., Ramani V.C., Ren Y., Naggi A., Torri G., Casu B., Penco S., Pisano C., Carminati P., Tortoreto M., Zunino F., Vlodavsky I., Sanderson R.D, Yang Y. SST0001, a chemically modified heparin, inhibits myeloma growth and angiogenesis via disruption of the heparanase/syndecan-1 axis. *Clin Cancer Res*. 17 (2011) 1382–1393.

*Rudd T.R.,* Macchi E., Gardini C., Muzi L., Guerrini M., Yates E.A., Torri G. How to find a Needle (or anything else) in a haystack: two-dimensional correlation spectroscopy-filtering with iterative random sampling applied to pharmaceutical heparin. *Anal. Chem*. 84 (2012) 6841-6847.

*Rudd* T.R., Macchi E., Muzi L., Ferro M., Gaudesi D., Torri G., Casu B., Guerrini M., Yates E.A. Unravelling structural information from complex mixtures utilizing correlation spectroscopy applied to HSQC spectra. *Anal. Chem*. 85 (2013) 7487-7493.

*Sandbäck-Pikas* D., Li J.P., Vlodavsky I., Lindahl U. Substrate specificity of heparanase from human hepatoma and platelets. *J. Biol. Chem*. 273 (1998) 18770–18777.

*Shafat* I., Ben-Arush M.W., Issakov J., Meller I., Naroditsky I., Tortoreto M., Cassinelli G., Lanzi C., Pisano C., Ilan N., Vlodavsky I., Zunino F. Pre-clinical and clinical significance of heparanase in Ewing's sarcoma. *J. Cell. Mol. Med*. 15 (2011) 1857-1864.

*Shively* J.E., Conrad H.E. Stoichiometry of the nitrous acid deaminative cleavage of model amino sugar glycosides and glycosaminoglycuronans. *Biochemistry.* 9 (1970) 33-43.

*Shriver* Z., Sundaram M., Venkataraman G., Fareed J., Linhardt R., Biemann K., Sasisekharan R. Cleavage of the antithrombin III binding site in heparin by heparinases and its implication in the generation of low molecular weight heparin. *Proc. Natl. Acad. Sci. U.S.A*. 97 (2000) 10365-10370.

*Shriver* Z., Capila I., Venkataraman G., Sasisekharan R. Heparin and heparan sulfate: analyzing structure and microheterogeneity. In Heparin: a century of progress. Springer: New York. (2012) 159-176.

*Sommers* D.C., Ye H., Kolinski R.E., Nasr M., Buhse L.F., Al-Halim A., Keire D. Characterization of currently marketed heparin products: analysis of molecular weight and heparinase-I digest patterns. *Anal. Bioanal. Chem*. 401 (2011) 2445-2454.

*Stringer* S.E., Kandola B.S., Pye D.A., Gallagher J.T. Heparin sequencing. *Glycobiology.* 13 (2003) 97-107.

*Sturiale* L., Naggi A., Torri G. MALDI mass spectrometry as a tool for characterizing glycosaminoglycan oligosaccharides and their interaction with proteins. *Semin. Thromb. Hemost*. 27 (2001) 465-472.

*Thanawiroon* C., Rice K.G., Toida T., Linhardt R. Liquid chromatography/mass spectrometry sequencing approach for highly sulfated heparin-derived oligosaccharides. *J. Biol. Chem.* 279 (2004) 2608-2615.

*Toida* T., Vlahov I.R., Smith A.E., Hileman R.E., Linhardt R.J. C-2 epimerization of N-acetylglucosamine in an oligosaccharide derived from heparan sulfate. *J. Carbohydr. Chem*. 15 (1996) 351-360.

*Tone* Y., Pedersen L.C., Yamamoto T., Izumikawa T., Kitagawa J., Tamura J., Negishi M., Sugahaara K. 2-O-Phosphorylation of xylose and 6-O-sulfation of galactose in the protein linkage region of glycosaminoglycans influence the glucuronyltransferase-I activity involved in the linkage region synthesis. *J. Biol. Chem*. 283 (2008) 16801-16807.

*Turnbull* J.E., Hopwood J.J., Gallagher J.T. A strategy for rapid sequencing of heparan sulphate and heparin saccharides. *Proc. Natl. Acad. Sci*. 96 (1999) 2698-2703.

*Vann* W.F., Schmidt M.A., Jann B., Jann K. The structure of the capsular polysaccharide (K5 antigen) of urinary-tract-infective *Escherichia coli* 010:K5:H4. *Eur. J. Biochem*. 116 (1981) 359-364.

*Viskov* C., Mourier P. Method for quantitatively determining specific groups constituting heparins or low molecular weight heparins. *US Pat Appl US 2004/0265943 A1* (2004).

*Viskov* C., Elli S., Urso E., Gaudesi D., Mourier P., Herman F., Boudier C., Casu B., Torri G., Guerrini M. Heparin Dodecasaccharide Containing Two Antithrombin-binding Pentasaccharides. Structural reatures and biological properties *J. Biol. Chem.* 228 (2013) 25895-25907.

*Vlodavsky* I., Friedmann Y. Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J. Clin. Invest*. 108 (2001) 341-347.

*Vlodavsky* I., Ilan N., Nadir Y., Brenner B., Katz B.-Z., Naggi A., Torri G., Casu B., Sasisekharan R. Heparanase, heparin and the coagulation system in cancer progression. Thromb. Res. 120 (2007) S112-S120.

*Vogt* A.M., Pettersson F., Moll K., Johnsson C., Normark J., Ribacke U., Egwang T.G., Ekre H-P., Spilmann D., Chen Q., Wahlgren M. Release of sequestered malaria parasites upon injection of a glycosaminoglycan. *PLoS Pathogens.* 2 (2006) 853-863.

*Vold* I.M.N., Kristianses K.A., Christensen B.E. A study of the chain stiffness and extention of alginates, in vitro epimerized alginates, and periodate-oxidized alginates using size-exclusion chromatography combined with light scattering and viscosity detectors. *Biomacromolecules*. 7 (2006) 2136-2146.

*Wang B*., Buhse L., Al-Hakim A., Boyne Ii M.T., Keire D.A. Characterization of currently marketed heparin products: Analysis of heparin digests by RPIP-UHPLC-QTOF-MS. *J. Pharm. Biomed. Anal*. 67-68 (2012) 42-50.

*Weitz* J.I. Low-molecular weight heparins. *New Engl. J. Med*. 337 (1997) 688-698.

*Weitz* J.I., Young E., Johnston M., Stafford A.R., Fredenburgh J.C., Hirsh J. Vasoflux, a new anticoagulant with a novel mechanism of action circulation. *Circulation*. 99 (1999) 682-689.

*Wolff* J.J., Chi L.L., Linhardt R.J., Amster I.J. Distinguishing glucuronic from iduronic acid in glucosaminoglycan tetrasaccharides by using electron detachment dissociation. *Anal. Chem*. 79 (2007) 2015-2022.

*Wu* Z.L., M. Lech, Characterizing the non-reducing end structure of heparan sulphate. *J. Biol. Chem*. 280 (2005) 33749-37755.

*Xiao* Z., Tappen B.R., Ly M., Zhao W., Canova L.P., Guan H., Linhardt R. Heparin mapping using heparin lyases and the generation of a novel low molecular weight heparin. *J. Med. Chem*. 54 (2011) 603-610.

*Xiao Z., Zhao W.,* Yang B., Zhang Z., Guam H., Linhardt R.J. Heparinase 1 selectivity for the 3,6-di-O-sulfo-2-deoxy-2-sulfoamido-α-D-glucopyranose (1,4) 2-O-sulfo-α-L-idopyranosyluronic acid (GlcNS3S6S-IdoA2S) linkages. *Glycobiology*, 21 (2011) 13-22.

*Yates* E.A., Santini F., Guerrini M., Naggi A., Torri G., Casu B. 1H and 13C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydr. Res*. 294 (1996) 15-27.

*Yates* E.A., Santini F., De Cristofano B., Payre N., Cosentino C., Guerrini M., Naggi A., Torri G., Hricovini M. Effect of substitution pattern on 1H, 13C NMR chemical shifts and $^{1}J(CH)$ coupling constants in heparin derivatives. *Carbohydr. Res*. 329 (2000) 239-247.

*Yamada S.,* Yamane Y., Tsuda H., Yoshida K., Sugahara K. A major common trisulfated hexasaccharide core sequence, hexuronic acid (2-sulfate)-glucosamine (N-sulfate)-iduronic acid-N-acetylglucosamine-glucuronic acid-glucosamine (N-sulfate) isolated from the low sulfated irregular region of porcine intestinal heparin. *J. Biol. Chem*. 273 (1998) 1863-1871.

*Yu* G., LeBrun L., Gunay N.S., Hoppensteadt D., Walenga J., Fareed J., Linhardt R.J. Heparinase I acts on a synthetic heparin pentasaccharide corresponding to the antithrombin III binding site. *Thrombos Res*. 100 (2000) 549-556.

*Zaia* J., *Costello* C.E. Tandem mass spectrometry of sulphated heparin-like glucosaminoglycan oligosaccharides. *J. Am. Soc. Mass Spectrom*. 18 (2007) 952-960.

*Zhang* F., Yang B., Ly M., Solakyildirim K., Xiao Z., Wang Z., Beaudet J.M., Torelli A.Y., Dordick J.S., Linhardt R.J. Structural characterization of heparins from different commercial sources. *Anal. Bioanal. Chem*. 401 (2011) 2793-2803.

*Zhou* H., Roy S., Cochran E., Zouaoui R., Chu C.L., Duffner J., Zhao G., Smith S., Galcheva-Gargova Z., Karlgren J., Dussault N., Kwan R.Y.Q., Moy E., Barnes M., Long A., Hohan C., Qi Y.W., Shriver Z., Ganguly T., Schultes B., Venkataraman G., Kishimoto T.K. M402, a novel heparan sulfate mimetic, targets multiple pathways in tumor progression and metastasis. *PLosONE 6, e21106*. (2011) doi: 10.1371/journal.pone.0021106

**Annex 1.** Some structural features of heparinases I, II and III and their substrate specificity

| | Heparin lyase I | Heparin lyase II | Heparin lyase III | Ref |
|---|---|---|---|---|
| Basic sites<br>(resembling heparin-binding sequences) | 1 | 3 | 2 | (*Capila* et al., 2002) |
| Calcium-coordinating motifs | 2 | – | 2 | (*Capila* et al., 2002) |
| | Calcium enhances the enzymatic activity of heparin lyases I and III, but can inhibit the enzymatic activity of heparin lyase II | | | |
| Catalytic domain | The active site contains a cystein and a histidine residues | It is thought to contain two active sites, the first acts on heparin, the other on HS.<br>One cysteine and three histidine residues are essentical for its enzymatic activity | Unlike the other two enzymes, its active site contains two histidine residues, and no cysteine residue | (*Capila* et al., 2002) |
| Specificity | Strong specificity for the sulfated domains of both heparin and HS | Very broad | Strong specificity for HS | (*Capila* et al., 2002) |
| Major products of heparin partial digestion by each enzyme *<br>(shown in the order of decreasing cleavage) | $\Delta U_{2S}$-$A_{NS6S}$ (79-81%)<br>$\Delta U_{2S}$-$A_{NS}$<br>$\Delta$U-$A_{NS6S}$ | $\Delta U_{2S}$-$A_{NS6S}$ (60-66%)<br>$\Delta$U-$A_{NS6S}$<br>$\Delta U_{2S}$-$A_{NS}$<br>$\Delta$U-$A_{NS}$<br>$\Delta$U-$A_{NAc6S}$<br>$\Delta U_{2S}$-$A_{Nac}$<br>$\Delta U_{2S}$-$A_{NS6S}$ | $\Delta$U-$A_{NAc}$ (39-52%)<br>$\Delta$U-$A_{NS}$<br>$\Delta$U-$A_{NAc6S}$<br>$\Delta$U-$A_{NS6S}$ | (*Zhang* et al., 2011) |

** Symbol Δ is used to indicate the formation of double bond at the uronic acid as result of β-eliminative enzymatic cleavage*

**Annex 2.** Heparanase sensitive and resistant oligosaccharides (*Okada* et al., 2002)

| Model oligosaccharide | Percentage cleavage by heparanase, % |
|---|---|
| $\Delta U_{2S} - A_{NS6S} - \mathbf{G - A_{NS6S}} - G - \mathbf{A_{NS6S}}$ | >95 |
| $\Delta U_{2S} - A_{NS6S} - I - A_{NAc6S} - \mathbf{G - A_{NS6S}}$ | 85 |
| $\Delta U_{2S} - A_{NS6S} - I - A_{NAc6S} - \mathbf{G - A_{NS}}$ | 59 |
| $\Delta U_{2S} - A_{NS} - \mathbf{G - A_{NS6S}}$ | 44 |
| $\Delta U_{2S} - A_{NS6S} - I - A_{NAc} - \mathbf{G - A_{NS6S}}$ | 30 |
| $\Delta U - A_{NS} - \mathbf{G - A_{NAc}}$ | 0 |
| $\Delta U - A_{NS6S} - \mathbf{G - A_{NAc}}$ | 0 |
| $\Delta U - A_{NS} - I_{2S} - A_{NAc} - \mathbf{G - A_{NS}}$ | 0 |
| $\Delta U - A_{NAc} - \mathbf{G - A_{NAc}}$ | 0 |

The substrate specificity of heparanase was studied using structurally defined oligosaccharides isolated form a heparin digested with heparinases (*Okada* et al., 2002)

$\Delta U_{(2S)}$ – 4,5-unsaturated uronic acid with or without 2-O-sulfation, $A_{NS/NAc(6S)}$ – N-sulfated/N-acetylated glucosamine with or without 6-O-sulfation, G – glucuronic acid, I – iduronic acid

**Annex 3.** Hydrodynamic radius ($R_h$) and Mark-Houwink parameters *a* and *logK* measured by SEC-TDA for heparins/LMWHs and their gs-derivatives

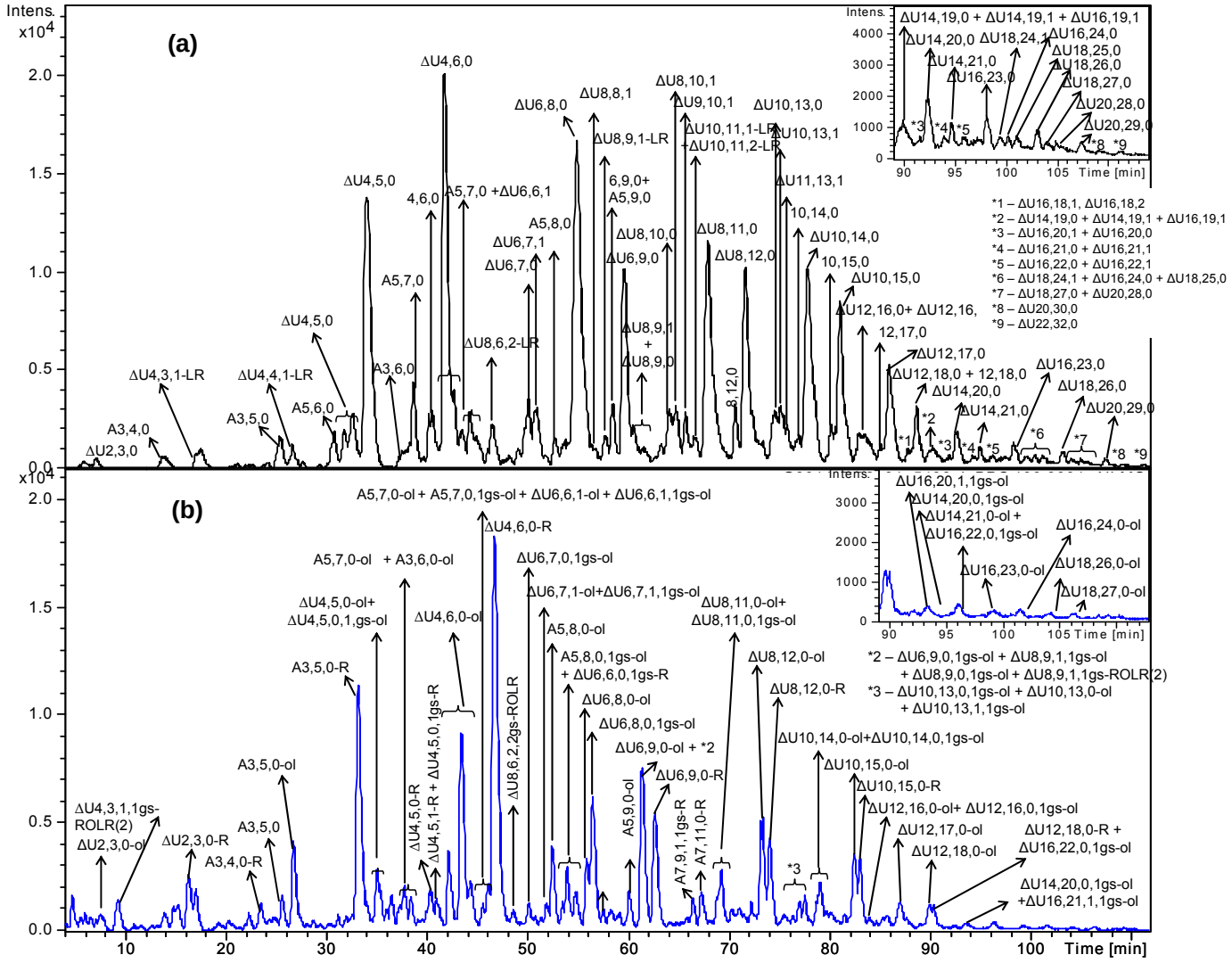
| | Sample | Rh, nm | Range, kDa | % | a | logK | Range, kDa | % | a | logK | Range, kDa | % | a | logK | Range, kDa | % | a | logK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Heparins | PMH | 4.2 | 9 – 13 | 28 | 1.37 | -6.3 | 13 – 16 | 21 | 0.94 | -4.5 | 16 – 21 | 17 | 0.66 | -3.3 | 21 – 36 | 17 | 0.57 | -3.0 |
| | BLH | 3.6 | 8 – 10 | 24 | 1.32 | -6.1 | 10 – 12 | 15 | 0.99 | -4.7 | 12 – 23 | 28 | 0.74 | -3.7 | 23 – 37 | 10 | 0.63 | -3.2 |
| | OMH | 3.5 | 6 – 7 | 12 | 1.41 | -6.4 | 7 – 14 | 50 | 0.99 | -4.7 | 14 – 32 | 31 | 0.66 | -3.4 | > 32 | nd | nd | nd |
| | BMH | 4.0 | < 8 | nd | nd | nd | 8 – 17 | 60 | 1.08 | -5.1 | 17 – 22 | 13 | 0.74 | -3.7 | 22 – 35 | 14 | 0.63 | -3.2 |
| | | | | | | | | | | | | | | | | | | |
| Gs-heparins | Gs-PMH | 2.7 | 3.5 – 4.5 | 15 | 1.44 | -6.3 | 4.5-8.5 | 38 | 0.90 | -4.4 | 8.5 – 22 | 36 | 0.53 | -2.9 | > 22 | nd | nd | nd |
| | Gs-BLH | 2.8 | < 4.5 | nd | nd | nd | 4.5 – 8 | 43 | 1.20 | -5.4 | 8 – 23 | 44 | 0.55 | -3.0 | > 23 | nd | nd | nd |
| | Gs-OMH | 2.1 | 2.5 – 3 | 20 | 1.33 | -5.8 | 3 – 6.5 | 42 | 0.71 | -3.7 | 6.5 - 17 | 29 | 0.55 | -3.0 | > 17 | nd | nd | nd |
| | Gs-BMH | 2.5 | < 3 | nd | nd | nd | 3 – 7.5 | 51 | 0.97 | -4.6 | 7.5 – 21 | 38 | 0.56 | -3.0 | > 21 | nd | nd | nd |
| | | | | | | | | | | | | | | | | | | |
| | **Sample** | **Rh, nm** | **Range, kDa** | **%** | **a** | **logK** | **Range, kDa** | **%** | **a** | **logK** | **Range, kDa** | **%** | **a** | **logK** | **Range, kDa** | **%** | **a** | **logK** |
| LMWHs | Dalte | 2.2 | 2.5 – 3.5 | 30 | nd | nd | 3.5 – 3.8 | 7 | 1.07 | -4.8 | 3.8 – 4.5 | 12 | 0.61 | -3.2 | 4.5 – 14 | 44 | 0.37 | -2.3 |
| | Tinza | 2.3 | 2.5 – 3.2 | 18 | nd | nd | 3.2 – 4.5 | 25 | 0.98 | -4.5 | 4.5 – 6.0 | 18 | 0.66 | -3.4 | 6.0 – 17 | 35 | 0.50 | -2.8 |
| | Enoxa | 1.8 | 1.7 – 2.0 | 10 | nd | nd | – | – | – | – | 2.0 – 4.5 | 46 | 0.67 | -3.5 | 4.5 – 13 | 30 | 0.37 | -2.4 |
| | | | | | | | | | | | | | | | | | | |
| Gs-LMWHs | Gs-Dalteparin | 2.0 | 2.5 – 3.5 | 19 | nd | nd | 3.5 – 5 | 25 | 0.53 | -2.9 | 5 – 11 | 33 | 0.26 | -2.0 | > 11 | nd | nd | nd |
| | Gs-Tinzaparin | 2.2 | < 2.5 | nd | nd | nd | 2.5 – 5 | 45 | 1.13 | -5.2 | 5 – 11 | 40 | 0.48 | -2.7 | > 11 | nd | nd | nd |
| | Gs-Enoxaparin | 1.8 | 1.5 – 2 | 20 | nd | nd | 2 – 4 | 44 | 0.58 | -3.2 | 4 – 12 | 31 | 0.36 | -2.4 | > 12 | nd | nd | nd |

0.01M $NaNO_3$ was used as eluent

**Annex 4.** Chemical shifts of some residues typical for LMWHs (tinzaparin, enoxaparin, dalteparin)

| LMWH | Residue | Signal | Chemical shift $^{1}H/^{13}C$ (ppm) | | |
|---|---|---|---|---|---|
| | | | ***Present work*** | ***Previously published data*** | |
| **Tinzaparin Enoxaparin** | $\Delta U_{2S}$ | H1/C1 | 5.53/100.1 | 5.53/100.1 | (*Guerrini* 2007) |
| | $\Delta U_{2S}$-($A_{NS6S}$) | H4/C4 | 5.98/108.8 | 6.01/108.7 | (*Guerrini* 2007) |
| | $A_{NS}$redα | H1/C1 | 5.47/93.9 | 5.47/93.9 | (*Guerrini* 2007) |
| **Enoxaparin** | ΔU | H1/C1 | 5.16/103.9 | 5.18/103.9 | (*Guerrini* 2007) |
| | ΔU | H4/C4 | 5.81/110.5 | 5.83/110.4 | (*Guerrini* 2007) |
| | $M_{NS}$red | H1/C1 | 5.40/95.8 | 5.41/95.6 | (*Guerrini* 2007) |
| | 1.6a$A_{NS}$ | H1/C1 | 5.61/104.3 | 5.63/104.3 | (*Guerrini* 2007) |
| | 1.6a$M_{NS}$ | H1/C1 | 5.57/103.9 | 5.59/103.9 | (*Guerrini* 2007) |
| | $G_{2S}$ | H1/C1 | 4.75/102.9 | 4.75/102.8 | (*Guerrini* 2007) |
| | $I_{2S}$β | H1/C1 | 4.97/94.8 | 4.99/94.8 | (*Guerrini* 2007) |
| **Dalteparin** | $I_{2S}$-(aM.$ol_{6S}$) | H1/C1 | 5.18/102.4 (overlapped with $I_{2S}$ of the repeat sequences and $I_{2S}$ at the NRE) | 5.19/102.6 (decamer) | (*Huckerby* 1986) |
| | $I_{2S}$(nr) | H1/C1 | 5.18/101.9 ppm (overlapped with $I_{2S}$ of the repeat sequences and $I_{2S}$-(aM.ol)) | 5.19/101.8 (decamer) | (*Huckerby* 1986) |

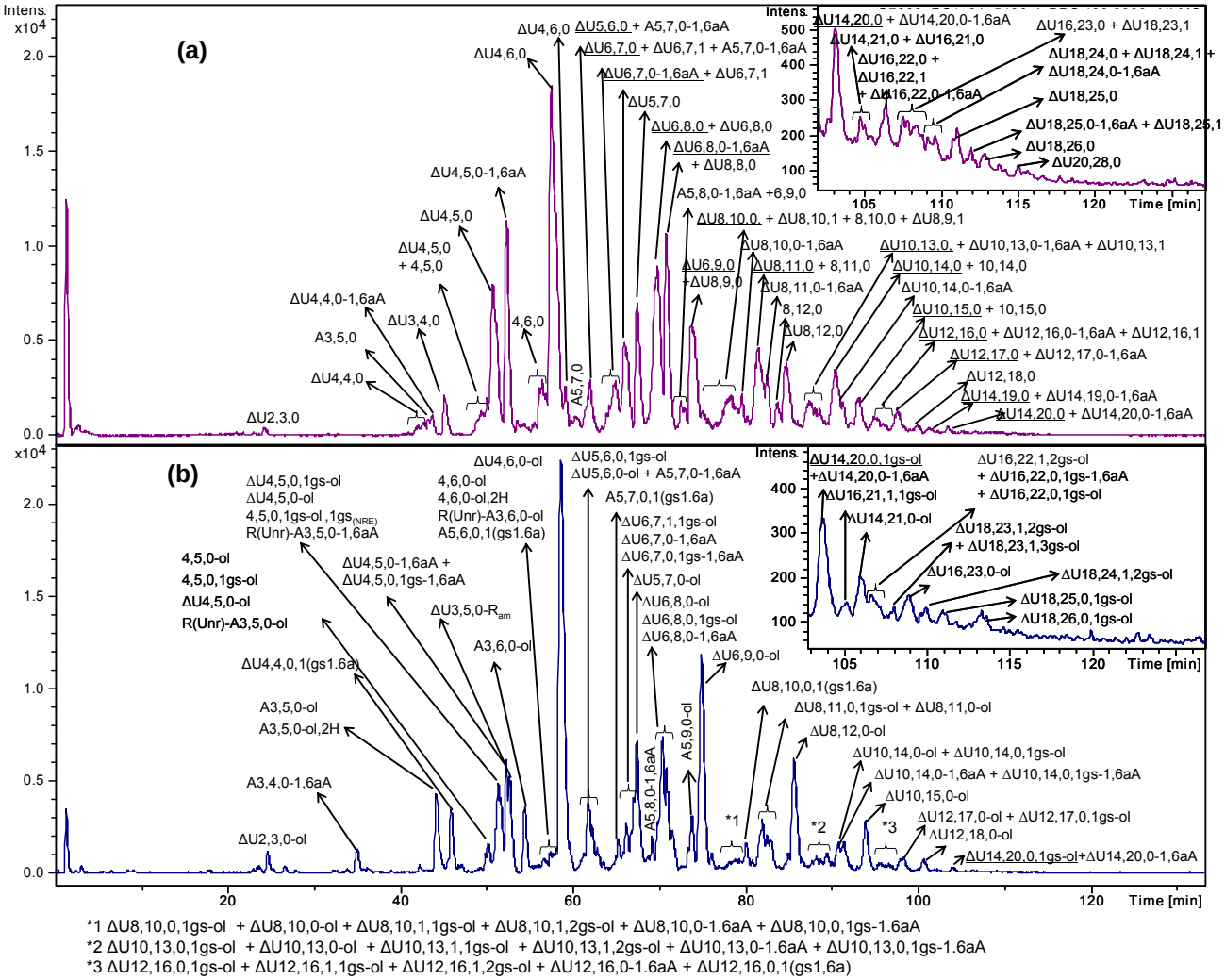
**Annex 5.** Total LC-MS profile of tinzaparin (a) and gs-tinzaparin (b)




LC conditions: column C18 100 x 2.1 mm, 2.6 µm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; gradient: 0 min – 5%B, 5 min – 5%B, 125 min – 35%B, 130 min – 70%B, 135 min – 70%B, 138 min – 5%B, 160 min – 5%B; flow rate – 0.25 ml/min

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form, while R indicates remnant of uronic acids (hydrolysis marker)

**Annex 6.** Total LC-MS profile of enoxaparin (a) and gs-enoxaparin (b)

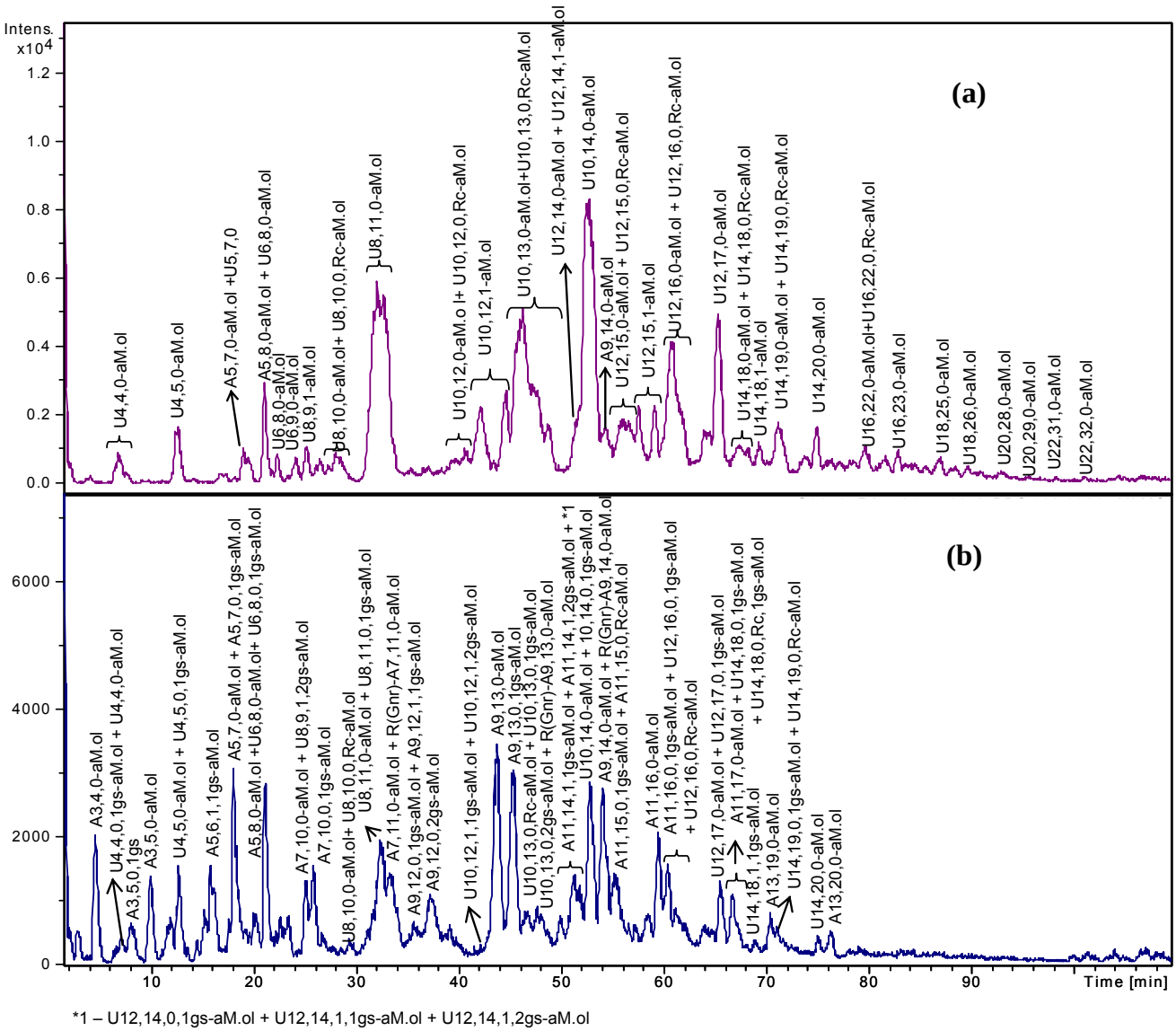



*1 ΔU8,10,0,1gs-ol + ΔU8,10,0-ol + ΔU8,10,1,1gs-ol + ΔU8,10,1,2gs-ol + ΔU8,10,0-1.6aA + ΔU8,10,0,1gs-1.6aA
*2 ΔU10,13,0,1gs-ol + ΔU10,13,0-ol + ΔU10,13,1,1gs-ol + ΔU10,13,1,2gs-ol + ΔU10,13,0-1.6aA + ΔU10,13,0,1gs-1.6aA
*3 ΔU12,16,0,1gs-ol + ΔU12,16,1,1gs-ol + ΔU12,16,1,2gs-ol + ΔU12,16,0-1.6aA + ΔU12,16,0,1(gs1,6a)

LC conditions: column C18 100 x 2.1 mm, 2.6 μm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; gradient: 0 min – 0%B, 5 min – 0%B, 130 min – 35%B, 140 min – 70%B, 145 min – 70%B, 148 min – 0%B, 170 min – 0%B; flow rate – 0.25 ml/min

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. 1,6aA indicates both 1,6-anhydro-glucosamine and 1,6-anhydro-mannosamine.

**Annex 7.** Total LC-MS profile of dalteparin (a) and gs-dalteparin (b)




LC conditions: column C18 100 x 2.1 mm,.2.6 µm; eluents A and B – 10 mM DBA and 10 mM $CH_3COOH$ in $H_2O$-$CH_3CN$ = 9:1 (v/v) and $CH_3CN$, respectively; gradient: 0 min – 10%B, 4 min – 10%B, 15 min – 18%B, 30 min – 18%B, 115 min – 35%B, 133 min – 70%B, 138 min – 70%B, 140 min – 10%B, 160 min – 10%B; flow rate – 0.25 ml/min

Abbreviation system includes the number of monosaccharide residues, sulfate and *N*-acetyl groups, and the number of gs-residues. Symbol ΔU indicates 4,5-unsaturated uronic acid at the NRE. Symbol "ol" indicates the terminal sugars in alditol (reduced) form. Symbols aM.ol and Rc indicate 2,5-anhydro-mannitol and ring-contracted units.

**Annex 8.** LC-MS data for several compounds discussed in Chapter IV

| Compound | Prevalent ion form | Monoisotopic mass-to-charge value (*m/z*) | Retention time, min | Sample |
|---|---|---|---|---|
| **ΔU22,32,0** | $[M-4H+22DBA]^{4-}$ | 2276.009 | 111.0 | Tinzaparin |
| **ΔU9,10,1** | $[M-3H+3DBA]^{3-}$ | 916.816 | 65.5 | |
| **ΔU4,3,1-LR** | $[M-2H]^{2-}$ | 836.640 | 17.4 | |
| **ΔU4,4,1-LR** | $[M-2H]^{2-}$ | 876.619 | 26.5 | |
| **ΔU8,6,2-LR** | $[M-3H]^{3-}$ | 876.119 | 46.3 | |
| **ΔU8,9,1-LR** | $[M-3H+2DBA]^{3-}$ | 1028.173 | 57.5 | |
| **ΔU10,11,1-LR** | $[M-3H+4DBA]^{3-}$ | 1279.946 | 66.3 | |
| **ΔU10,11,2-LR** | $[M-3H+4DBA]^{3-}$ | 1293.949 | 66.4 | |
| | | | | |
| **ΔU9,10,1-ol** | $[M-3H+3DBA]^{3-}$ | 917.488 | 66.0 | gs-Tinzaparin |
| **ΔU4,5,0-R** | $[M-2H]^{2-}$ | 594.999 | 40.2 | |
| **ΔU4,5,0,1gs-R** | $[M-2H]^{2-}$ | 596.007 | 41.0 | |
| **ΔU4,6,0-R** | $[M-2H+1DBA]^{2-}$ | 699.553 | 46.6 | |
| **ΔU6,6,0,1gs-R** | $[M-2H+3DBA]^{2-}$ | 998.263 | 53.8 | |
| **ΔU6,9,0-R** | $[M-2H+4DBA]^{2-}$ | 1181.767 | 62.5 | |
| **ΔU8,12,0-R** | $[M-3H+4DBA]^{3-}$ | 979.832 | 73.9 | |
| **ΔU10,15,0-R** | $[M-3H+7DBA]^{3-}$ | 1301.308 | 82.9 | |
| | | | | |
| **ΔU20,28,0** | $[M-4H+17DBA]^{4-}$ | 1950.34 | 115.3 | Enoxaparin |
| **ΔU5,6,0** | $[M-2H+1DBA]^{2-}$ | 728.556 | 61.8 | |
| **ΔU5,7,0** | $[M-2H+2DBA]^{2-}$ | 833.110 | 67.3 | |
| **ΔU2,3,0** | $[M-H]^{1-}$ | 575.964 | 24.1 | |
| **ΔU3,4,0** | $[M-2H]^{2-}$ | 415.473 | 45.0 | |
| **ΔU4,5,0** | $[M-2H]^{2-}$ | 535.986 | 49.3<br>49.9<br>50.6<br>51.0 | |
| **A5,8,0-1,6aA** | $[M-2H+2DBA]^{2-}$ | 865.607 | 68.9 | |
| | | | | |
| **ΔU2,3,0-ol** | $[M-H]^{1-}$ | 577.982 | 24.5 | gs-Enoxaparin |
| **ΔU3,4,0-ol** | $[M-2H]^{2-}$ | 416.481 | 45.5 | |
| **ΔU4,5,0-ol** | $[M-2H]^{2-}$ | 536.993 | 50.4<br>51.2 | |
| **ΔU4,5,0,1gs-ol** | $[M-2H]^{2-}$ | 538.001 | 51.4 | |

| | | | | |
|---|---|---|---|---|
| **ΔU8,10,1,2gs-ol** *(contains gs-ATBR)* | $[M-3H+2DBA]^{3-}$ | 817.104 | 80.3 | |
| **ΔU8,9,1,2gs-1,6aA** *(contains gs-ATBR)* | $[M-3H+2DBA]^{3-}$ | 783.776 | 75.9 | |
| **ΔU10,13,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+5DBA]^{3-}$ | 1138.579 | 89.6 | |
| **ΔU12,15,1,2gs-1.6aA** *(may contain gs-ATBR)* | $[M-3H+8DBA]^{3-}$ | 1426.727 | 94.0 | |
| **ΔU12,16,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+9DBA]^{3-}$ | 1503.105 | 95.7 | |
| **ΔU14,19,1,2gs-ol** *(may contain gs-ATBR)* | $[M-3H+12DBA]^{3-}$ | 1824.581 | 102.2 | |
| **R(Unr)-A3,5,0-ol** | $[M-2H]^{2-}$ | 532.001 | 50.6 | |
| **R(Unr)-A3,5,0-1,6aA** | $[M-2H]^{2-}$ | 521.988 | 51.4 | |
| **A3,5,0-ol** | $[M-2H]^{2-}$ | 457.983 | 44.0 | |
| **A3,6,0-ol** | $[M-2H]^{2-}$ | 497.961 | 54.3 | |
| | | | | |
| **U22,31,0-aM.ol** | $[M-4H+21DBA]^{4-}$ | 2224.48 | 98.0 | Dalteparin |
| **U22,32,0-aM.ol** | $[M-4H+22DBA]^{4-}$ | 2276.76 | 101.0 | |
| **U8,10,0,Rc-aM.ol** | $[M-3H+3DBA]^{3-}$ | 840.131 | 27.2 | |
| **U10,13,0,Rc-aM.ol** | $[M-3H+6DBA]^{3-}$ | 1161.607 | 45.4<br>47.6 | |
| **U12,16,0,Rc-aM.ol** | $[M-3H+9DBA]^{3-}$ | 1483.082 | 60.3<br>61.9 | |
| **U14,19,0,Rc-aM.ol** | $[M-3H+12DBA]^{3-}$ | 1804.558 | 70.8<br>72.3 | |
| **U16,22,0,Rc-aM.ol** | $[M-4H+12DBA]^{4-}$ | 1497.661 | 79.6<br>80.5 | |
| **U5,6,0** | $[M-2H+1DBA]^{2-}$ | 737.561 | 17.3 | |
| **U5,7,0** | $[M-2H+2DBA]^{2-}$ | 842.136 | 19.6 | |
| **U8,11,0-aM.ol** | $[M-3H+3DBA]^{3-}$ | 871.787 | 32.0 | |
| **A5,8,0-aM.ol** | $[M-2H+3DBA]^{2-}$ | 931.683 | 21.0 | |
| | | | | |
| **U8,11,0,1gs-aM.ol** | $[M-3H+3DBA]^{3-}$ | 872.459 | 32.3 | gs-Dalteparin |
| **A7,11,0-aM.ol** | $[M-2H+7DBA]^{2-}$ | 1478.472 | 33.2 | |
| **R(Unr)-A7,11,0-aM.ol** | $[M-3H+4DBA]^{3-}$ | 905.506 | 33.2 | |
| **A7,10,0-aM.ol** | $[M-2H+5DBA]^{2-}$ | 1309.342 | 24.9 | |
| **A7,10,0,1gs-aM.ol** | $[M-2H+5DBA]^{2-}$ | 1310.350 | 25.7 | |