TRAVELING WAVE ION MOBILITY SPECTROMETRY STUDIES OF PROTEIN-GLYCOSAMINOGLYCAN INTERACTIONS

by

YUEJIE ZHAO

(Under the Direction of I. Jonathan Amster)

ABSTRACT

Glycosaminoglycans (GAGs) are linear, highly negatively charged polysaccharides expressed by all animal cells. Through interactions with many proteins, GAGs act as regulatory molecules in many physiological and pathological processes. Travelling Wave Ion Mobility Spectrometry (TWIMS), a gas phase biophysical technique that separates ions based on their size, shape and charge is a valuable tool in the structural characterization of biomolecules based on measurement of rotationally average collision cross sections (CCSs). In this dissertation, we explore the application of native mass spectrometry (native MS) coupled with TWIMS to investigate different protein-GAG binding systems. ATIII and Arixtra, a well-studied protein-GAG system with a high binding specificity, was examined to establish the capabilities of this approach. More complicated and less well characterized systems were studied, including FGF1 and Robo1 with heparin-like GAGs. Our results demonstrate the potential for applying native MS-TWIMS to achieve a molecular level understanding of the nature of these protein-GAG interactions.

INDEX WORDS: Traveling Wave Ion Mobility Spectrometry, Native Mass Spectrometry, Glycosaminoglycans, Heparin/Heparan Sulfate-Protein Interactions

TRAVELING WAVE ION MOBILITY SPECTROMETRY STUDIES OF PROTEIN-GLYCOSAMINOGLYCAN INTERACTIONS

by

YUEJIE ZHAO

B.S., Northwest University, China, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2017

TRAVELING WAVE ION MOBILITY SPECTROMETRY STUDIES OF PROTEIN-GLYCOSAMINOGLYCAN INTERACTIONS

by

YUEJIE ZHAO

Major Professor:	I. Jonathan Amster
Committee:	John L. Stickney Joshua Sharp

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

DEDICATION

To my beloved parents Yuli Wu and Zhi Zhao, who always love, support and believe in me.

To the memories of my beloved grandfather Min Zhao, who inspired my love for science and research.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my dearest advisor I. Jonathan Amster, for your support and mentoring throughout these years, and for always believing in me and recognizing the potentials within me. It has been an honor to be your student.

TABLE OF CONTENTS

CHAPTER

APPENDICES

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Overview of glycosaminoglycans

Glycosaminoglycan (GAG) family of polysaccharides

Glycosaminoglycans (GAGs) are biopolymers that are widely found on cell surface, in the extracellular matrix and in basement membranes of many organisms. GAGs occur as free polysaccharides or as constituent glycans of proteoglycans (PGs). Four subgroups of GAGs are known, distinguished from each other by their saccharide composition and glycosidic linkage, including: hyaluronic acid (HA), keratan sulfate (KS), chondroitin/dermatan sulfate (CS/DS) as well as heparin (Hp) and heparan sulfate (HS). GAGs are composed of an unbranched polysaccharide chain with alternating disaccharide units encompassing hexuronic acid (D-glucoronic acidc(GlcA) or L-iduronic acid (IdoA)) and amino sugar (D-glucosamine (GlcN) or D-galactosamine(GalN)), except for KS which contains D-galactose (Gal) instead of hexuronic acid [1, 2].

As a unique member of the GAG family, HA is not sulfated, nor covalently linked to core proteins and is synthesized at the cell membrane instead of in the Golgi apparatus in contrast to other classes of GAG. The backbone of HA is composed of repeating disaccharide units of $(\text{-4-GlcA-}\beta 1\text{-3-GlcNAc-}\beta 1\text{-})_n$ [3, 4]. Being one of the most hydrophilic molecules, HA is abundantly distributed in connective tissues such as skin, synovial and vitreous fluids as well as main organs and muscle tissues, playing important roles in balancing fluid levels in interstitium, lubrication, water transport and microcirculatory exchanges. Recently, HA has been widely used in anticancer drug delivery and tissue scaffolding, due to its biocompatibility, biodegradability, non-toxicity, and non-immunogenicity [5, 6].

The backbone of CS is composed of alternating disaccharide units of (-4-GlcA-β1-3-GalNAc-β1-)$_n$, which can be sulfated on the C4, C6 position of the GalNAc residues and C2, C3 position of GlcA. They are covalently linked to the serine /threonine residues on the core proteins of PGs via a GlcA-β1-3-Gal-β-1-3-Gal-β-1-4-xylose (Xyl) tetrasaccharide linkage. CS is categorized into three main types: CS-A, CS-B (also known as DS) and CS-C, depending on the sulfation pattern of the disaccharide units and uronic acid stereochemistry. There are also many less common CS that have more complicated sulfation patterns. Both CS and DS interact with a

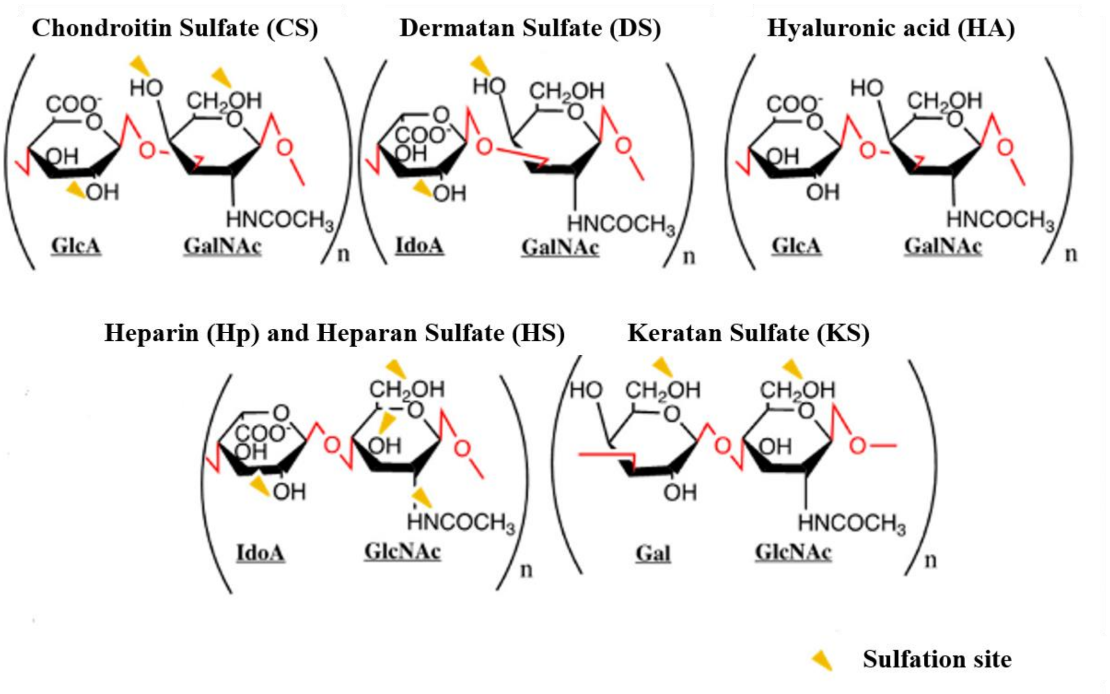


Figure 1.1 Structure of subgroups of GAGs. Glycosidic linkages are shown in red and sites of sulfation are indicated by yellow triangles, modified from [11]

large number of protein ligands, playing major roles in neurite outgrowth during CNS development, tissue morphogenesis, wound repair and growth factors signaling [7-9].

KS is one of the simplest type of GAG, whose backbone is composed of repeating disaccharide units of $(\text{-3-Gal-}\beta\text{1-4-GlcNAc-}\beta\text{1-})_n$. Both Gal and GlcNAc can be sulfated on the C6 position. KS is categorized into two main types based on the way it is attached to core proteins: KSI is conjugated to asparagine residues on core protein through a complex form of N-glycan, whereas KSII is conjugated to serine /threonine residues. The most important function of KSI is to modulate the dynamic level of hydration in cornea, whereas KSII is the functional component of PGs from bone, cartilage, and skin [1, 10].

Heparin (Hp) and heparan sulfate (HS)

Hp/HS are the best known therapeutically active type of GAG, and the most heavily sulfated GAG with the highest structural heterogeneity. The backbone of Hp and HS share the same repeating disaccharide units of $(\text{-4-GlcA/IdoA-}\beta\text{1-4-GlcNAc-}\alpha\text{1-})_n$, on which the N-sulfation or O-sulfation can occur at the C2, C3 and C6 positions of GlcN as well as the C2 position of GlcA/IdoA, but their primary structure is quite different. Hp has a higher density of sulfation, and thus a lower structural diversity than HS: the average frequency of sulfation per disaccharide is 1.8–2.6 for Hp and 0.6–1.8 for HS; Hp has a lower average molecular weight compared with HS: the size range of Hp is 7-20 KDa, whereas the size range is 20-100 KDa for HS [12].

Heparin occurs as a glycoconjugate of many important proteins. It is expressed exclusively as part of Serglycin, which is the major constituent of the cytoplasmic secretory granules of connective tissue-type mast cells and bipotential glial progenitor cells. HS is expressed by virtually all animal cells as the glycan component of three types of heparan sulfate proteoglycans (HSPGs): transmembrane type, glycerophosphatidylinositide (GPI)-anchored type and secreted extracellular matrix (ECM) type [2]. An example of a transmembrane bound HSPG is Syndecan, which is a transmembrane protein composed of a highly conserved intracellular cytoplasmic domain, a transmembrane domain and a structurally distinct extracellular domain that carries up to five HS or CS chains, with CS attached closer to the membrane than HS. An example of a GPI-anchored HSPG is Glypican, which features a globular N-terminus composed of fourteen conserved cysteine residues on which three or four HS chains are attached [13]. The transmembrane HSPG and GPI-anchored HSPG can be cleaved by matrix metalloproteases or heparanases to release the attached HS chains into the ECM. HS can also be directly secreted into the ECM as part of the multidomain basement membrane HSPGs including Perlecan, Collagen XVIII, and Agrin [13-15].

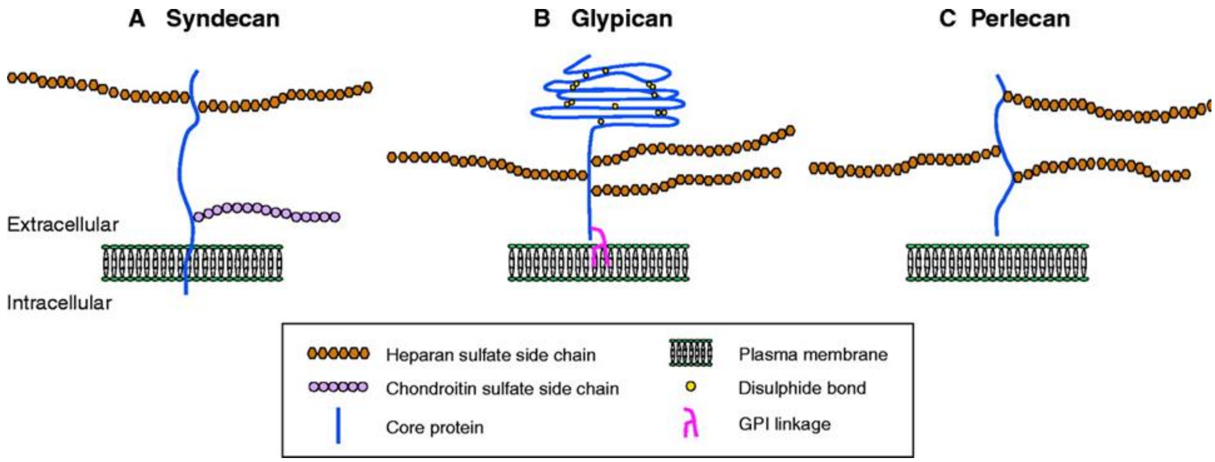


Figure 1.2 Schematic representation of HSPGs distribution at the cell surface and in the basement membranes [17]

Like other types of post-translational modifications (PTMs), the number, size and structure of the Hp/HS chains attached on PGs are highly dependent on different cell types, tissue types and development stages of an organism [16, 17], and these GAG chains dictate the chemical proprieties and biological functions of PGs [15].

Biosynthesis of heparin (Hp) and heparan sulfate (HS)

Heparin (Hp) and Heparan Sulfate (HS) are synthesized via a series of strictly catalyzed, sequential enzymatic interactions in the Golgi apparatus. These interactions do not go to completion, resulting in highly heterogeneous and polydisperse structure of mature Hp/HS chains which differ by their chain size, modification pattern, epimerization pattern and distribution of differently modified domains [18-20]. The first step of the biosynthesis of Hp/HS is the growth of a tetrasaccharide linkage region (GlcA-β1-3-Gal-β-1-3-Gal-β-1-4-Xyl) initiates with conjugating the Xyl from UDP-Xyl to the serine residues on the core protein, followed by a stepwise addition of two Gals and one GlcA to form the tetrasaccharide. This process is catalyzed by a series of specific enzymes including xylosyltransferase 1/2 (XylT1/2), galactosyltransferase 1 (GalT1), galactosyltransferase 2 (GalT2) and glucuronyltransferase 1 (GlcAT1), and occurs in the endoplasmic reticulum as a post-translational modification of serine residues in a protein.

The second step of the biosynthesis of Hp/HS is the chain polymerization, initiated with addition of a GlcNAc residue onto the tetrasaccharide linker, catalyzed by N-acetylglucosaminyl transferases Exostosin 2 and Exostosin 3 (EXTL2 and EXTL3). Then a continuous polymerization of Hp/HS backbone through alternating transferring GlcA and GlcNAc units from their corresponding UDP-sugars takes place, catalyzed by copolymerase exostosin 1 and exostosin 2 (EXT1 and EXT2).

In the last step of the biosynthesis of Hp/HS, post-synthetical modification reactions occur concomitantly and randomly along the Hp/HS backbone [21], including: (a) removal of the N-acetyl group of GlcNAc and substitution of sulfation of the free amine, catalyzed by GlcNAc N-deacetylase/Nsulfotransferases (Ndst1-4); (b) epimerization of GlcA to IdoA, catalyzed by GlcA C5 epimerase (HSepi); (c) addition of 2-O-sulfation on IdoA and less frequently on GlcA, catalyzed by 2-O-sulfotransferase (Hs2st); (e) addition of 6-O-sulfation on N-sulfated or less frequently on N-acetyl GlcN, catalyzed by 6-O-sulfotransferases (Hs6st1-3) and (f) addition of 3-O sulfation to N-sulfated or N-unsubstituted GlcN, catalyzed by 3-O-sulfotransferases (Hs3st1-6). These post-synthetical modifications of HS are typically found in clusters, resulting in domain structure of mature HS chains, consisting of N-sulfated regions (NS domains) which are highly sulfated, interspersed by N-acetylated regions (NA domains) and transition regions N-acetylated and N-sulfated domains (NA/NS domains) [18]. After transfer to the cell surface, remodeling of the sulfation pattern can occur, for example, catalyzed by two endosulfatases (Sulf1 and Sulf2) [2, 22].

Heparin (Hp)/heparan sulfate (HS)-protein interactions

Hp/HS binding proteins (HBPs)

Having the highest density of negative charge of any known biomacromolecules due to the presence of high level of spatially defined carboxylate groups and sulfo groups, Hp/HS participate in many biological and physiological processes through interacting and regulating a diverse pool of protein ligands, referred as Hp/HS binding proteins (HBPs). So far more than 400

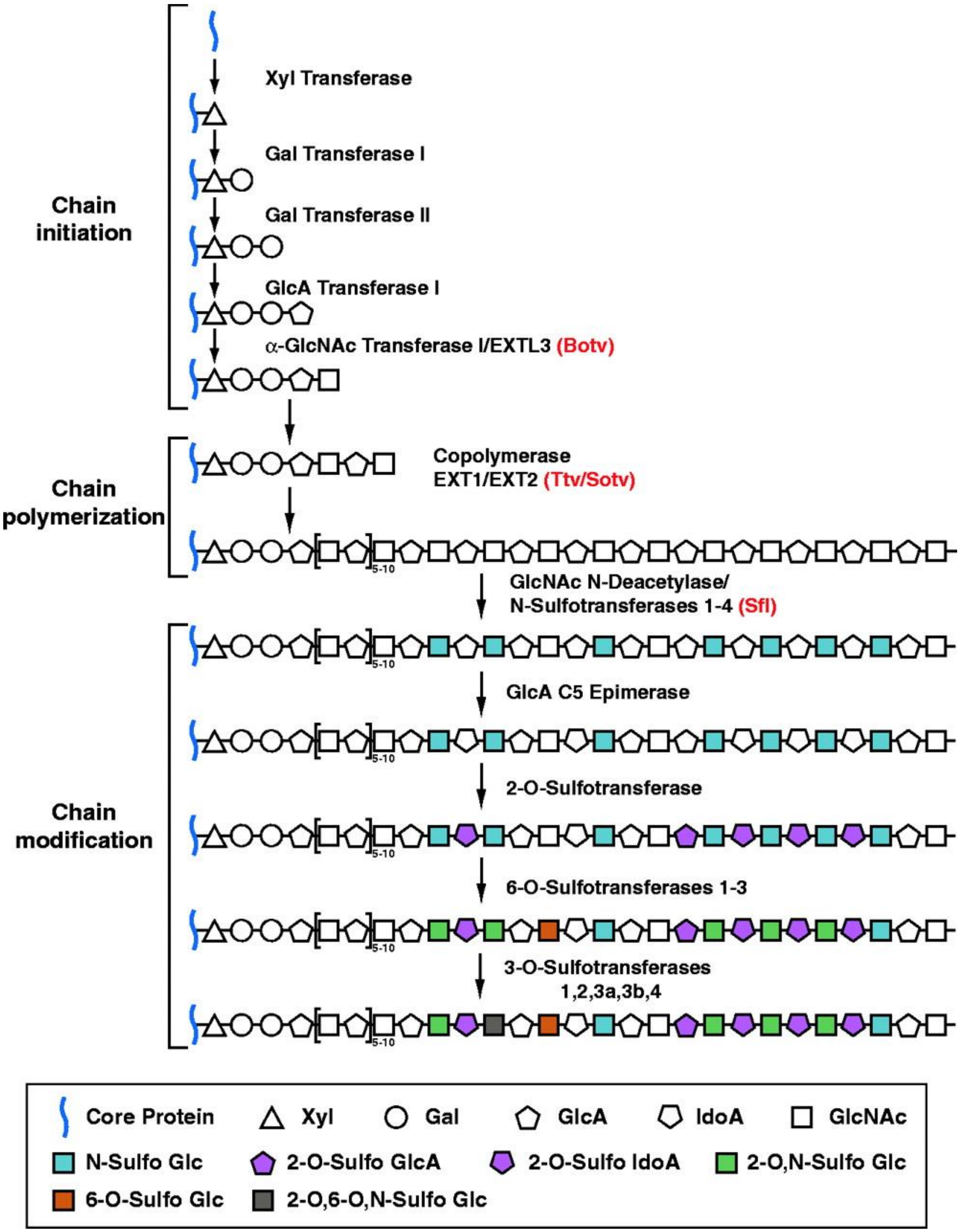


Figure 1.3 Schematic illustration of the Hp/HS biosynthesis [13]

HBPs are known, including growth factors and their receptors, morphogens, chemokines, cytokines, enzymes, extracellular matrix proteins, lipoproteins and various proteins associated with diseases [23].

Nature of Hp/HS-protein interactions

Electrostatic interactions play a predominant role in Hp/HS-protein binding. Hp/HS binding sites are either located on protein external surface or in a shallow groove where the glycan chain can fit in [21]. They are composed of a cluster of positively charged basic amino acid residues, including lysine, arginine and protonated histidine residues at low pH. These amino acids are either intrinsically contiguous in the primary sequence, or more commonly are not adjacent in sequence but are arranged closely in the folded protein [24].

Non-ionic interactions including hydrogen bonds, hydrophobic interactions and van der Waals (VDW) interactions also contribute to Hp/HS-protein binding. Amino acids that form hydrogen bonds including asparagine, glutamine and tyrosine are also commonly found at Hp/HS binding sites. The contribution of non-ionic interaction ranges from 20-70%, depending on each specific binding interaction [18].

The binding mode of Hp/HS-protein interaction is dictated by the primary sequence of Hp/HS as well as the degree and distribution of sulfo and carboxylate groups. The protein binding site on Hp/HS ranges in size from 6 to 18 saccharide units [25]. Larger protein binding sequences on HS are typically composed of multiple NS domains interspersed by NA/NS domains [26], and are either required for presenting multiple identical binding sites for the assembly of homo-oligomer, exemplified by the vascular endothelial growth factor (VEGF)165 [27], interleukin-8 [28] and flatelet factor 4 [29], or are required for bridging several different

HBPs to form complexes or hetero-oligomers, exemplified by antithrombin III (ATIII) and its target proteases.

The binding mode of Hp/HS-protein interaction is also affected by the three-dimensional structure of Hp/HS, which determines the spatial arrangement and orientation of sulfo and carboxylate groups for interacting with proteins. Hp/HS adopts a helical structure with limited flexibility, governed by the partially restricted rotation around the glycosidic linkages [30, 31] and a high conformational flexibility of the pyranose ring of IdoA, which can exist in three equienergetic conformations ($^1C_4$, $^4C_1$ and $^2S_0$). The conformation adopted by IdoA is affected by modifications on IdoA and its adjacent residue. For instance, unsubstituted IdoA adopts predominantly chair conformation, while 2-O-sulfated IdoA exists in equilibrium between $^1C_4$ and $^2S_0$ forms [32-34]. Though switching in between multiple conformations in unbound state, IdoA stays locked in one predominant conformation upon binding to its protein ligand [32].

Binding specificity

The specificity of Hp/HS-protein interactions can be highly sensitive to the sulfation patterns of the GAG. As defined by the review from Linhardt, specificity defines whether the Hp/HS-protein interaction is driven by the net charge of Hp/HS polysaccharides, or requires the presence of a specific structural motif, that is, the defined display of sulfo and carboxylate groups on a minimum number of saccharide residues [32, 35].

For the interactions with highest binding specificity, a distinct structural motif containing a rare structural component is often observed, exemplified by a 3-O-sulfated GlcN present in a specific pentasaccharide binding sequence for ATIII [36] and the binding sequence for Herpes simplex virus (HSV) glycoprotein gD [37]. This rare 3-O-sulfated GlcN is found on about 1/20

Hp disaccharides and less than 1/100 HS disaccharides [21]. Whereas for the majority of HBPs, their Hp/HS binding sites consist of relatively common structure units. Lacking a specific structural motif containing a rare structural component does not imply that this interaction has no selectivity. In some cases, particular types of sulfation are more relevant than the other types, which cannot be explained if the interaction is charge dependent or nonspecific. HBPs that have selectivity for specific types of sulfation include: some CC chemokines [38], some fibroblast growth factors(FGFs) [39, 40], human papillomavirus type 16 L1 capsid protein [41] and VEGF-A_{165} [35].

Previously, HSs were only available in small quantities, so abundant Hp and Hp-derived oligosaccharides from natural sources were used predominantly. These Hps are often heterogeneous and polydisperse mixtures with less composition diversity. Though the specific pentasacchride sequence recognized by ATIII is the only specific structural motif that has been widely acknowledged and used for therapeutics, thanks to recent improvements in chemical and enzymatic synthesis of HS, bulk production of HS of more structural variance and controllable sulfation patterns should become available, which will dramatically facilitate the recognition of Hp/HS-protein interactions with high binding specificity and physiological significance [42, 43].

Biological functions of Hp/HS

The non-covalent interactions with Hp/HS have major effects on HBPs, including their spatial distribution and concentration, conformation or topology, functional activity, stability against proteolysis and degradation, oligomerization state, transportation through basement membranes as well as their interaction with other proteins or receptors. Therefore, Hp/HS play essential roles in many developmental and homeostasis processes, including: blood coagulation

[44], cell growth and differentiation [45], cell adhesion and migration [46], lipid transport and metabolism [47], morphogenesis and organogenesis [45] as well as cell-cell and cell-matrix signaling [48]. Moreover, Hp/HS are also functionally involved in many pathophysiological

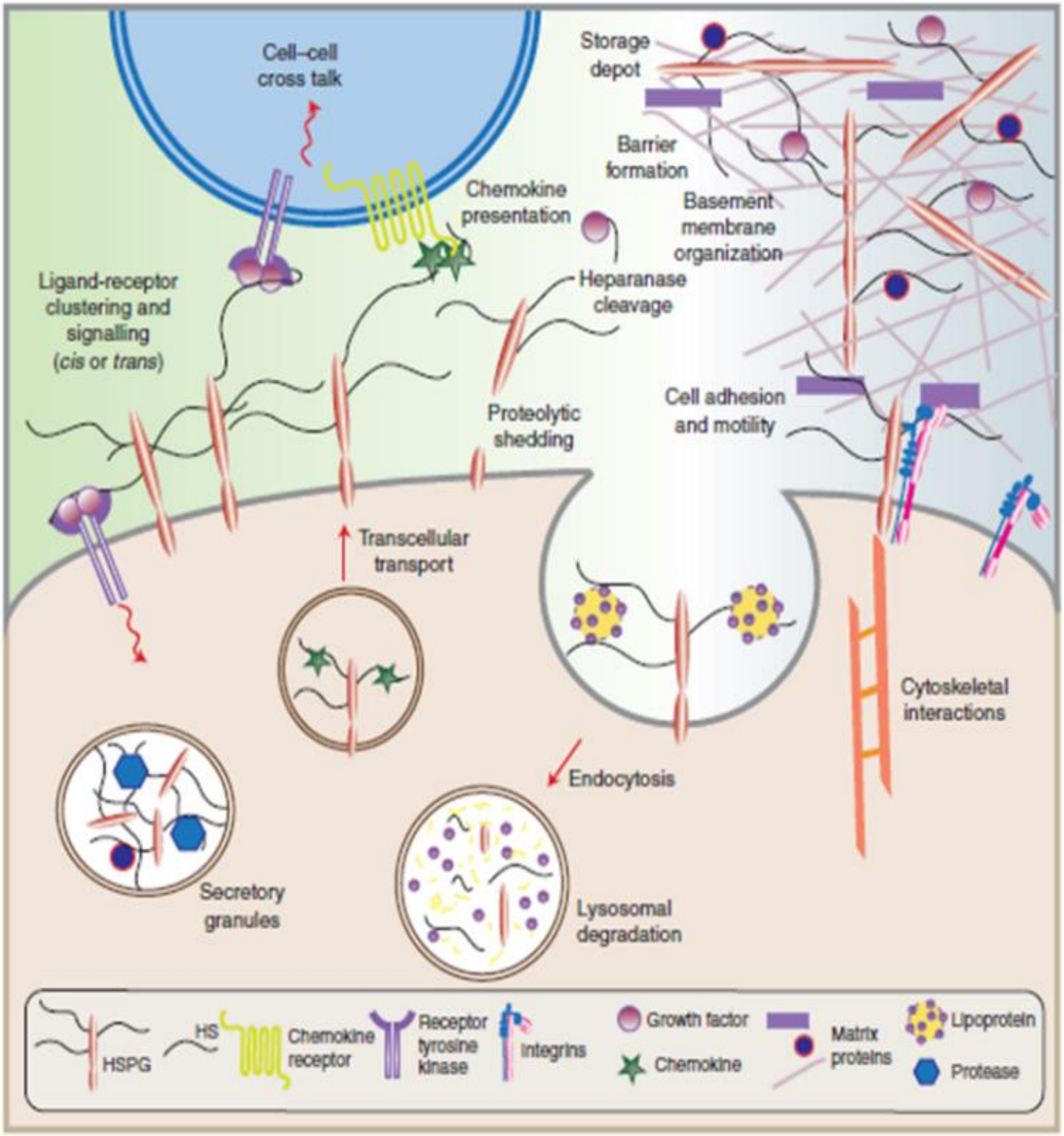


Figure 1.4 Hp/HS–protein interactions in various functional settings [53]

processes, including: tumor development and metastasis [49], host defense and viral infection [50], immune cell migration and inflammation [51] and neurodegenerative disorders [52], making them promising therapeutics. Some important Hp/HS-protein interactions are discussed below.

Allosteric effects on enzyme activity

The best-studied example of the Hp/HS-protein interaction is the high-affinity binding (an affinity of ~50nM at physiological pH) between ATIII and the pentasaccharide binding domain of Hp. This interaction is the basis for the clinical use of Hp as an anticoagulant drug [54]. ATIII is a protease inhibitor which regulates blood coagulation proteases (thrombin, factor Xa and factor IXa) in control of the blood clotting cascade. In the absence of Hp, the anticoagulant activity of ATIII is repressed since the protease binding reactive center loop (RCL) is initially buried in a partially open sheet A of ATIII and shows less affinity towards proteases. Binding of a unique pentasaccharide sequence containing a 3-O-sulfated GlcN on Hp causes a large scale conformational change within ATIII, leading to expulsion of the RCL. Therefore, Hp serves as a catalyst in ATIII-mediated inhibiting reactions and facilitate the formation of an ATIII-proteinase complex [55, 56].

Coreceptor for complex assembly

The study of interactions between growth factors and their corresponding receptors has mostly focused on the FGF family of proteins, which play important roles in tissue patterning, organogenesis as well as proliferation, differentiation, adhesion and apoptosis of cells [57]. HS serves as a coreceptor for the assembly and stabilization of a ternary signaling complex through binding to both the FGF and its tyrosine kinase receptor. The binding of HS induces dimerization of FGF and enhance its binding affinity towards the extracellular domains of FGFR, which further enables the transphosphorylation of FGFR and subsequent signal transduction [58]. Previous studies suggest that different members of FGF interact with HS and FGFR differently, resulting in distinctively different ternary complexes. For instance, two crystal structures of different binding stoichiometry and intermolecular interactions have been resolved for FGF1 and FGF2, including a 2:2:1 FGF1-FGFR2-HS dodecasaccharide model and a 2:2:2 FGF2-FGFR1-HS dodecasaccharide model [59, 60]. HS also serves as a coreceptor for hepatocyte growth factors and VEGFs and facilitates their interaction with their specific tyrosine kinase receptor in a similar mechanism [27, 61].

Deposition and presentation of HBPs

Another important function of Hp/HS is recruiting HBPs to the correct site of activation at the correct time, exemplified by the interactions between HS and CXC chemokines. These interactions are essential for appropriate recruitment and activation of leukocytes during inflammation, angiogenesis and cancer metastasis [62].

Specifically secreted by perivascular cells in response to inflammation, chemokines are then transported to endothelial cell surfaces, where they interact with HSs and become immobilized on cell surface to form a gradient. Formation of this chemokine gradient facilitates the presentation of chemokines to the chemokine receptors on leukocytes in blood, which is essential for leukocyte arrest via integrin activation [63] and correct migration of leukocytes from circulation to sites of inflammation and injury [64-66].

Template for oligomerization

In addition to increasing the local concentration of chemokines near sites of inflammation, HS also induces oligomerization of chemokines, which is directly associated with the presentation and their activity. Chemokine oligomers tend to have larger binding surfaces hence higher binding affinities towards HS compared with monomers. Different members of chemokines adopt different forms of oligomers with dramatically different dynamic stabilities in the presence of HS: some members are most stable in oligomeric forms, some members do not oligomerize at all, and other members convert between monomeric and oligomeric forms [64]. Thus, change of the oligomerization state of chemokines may contribute to the binding specificity of chemokine-HS interactions [67].

Roles of Hp/HS in pathophysiological conditions

Hp/HS have been implicated in multiple amyloidosis diseases, including type-2 diabetes and Alzheimer’s, Parkinson's and prion diseases [52, 68]. Hp/HS are functionally involved in the folding disorder of more than 40 disease-specific proteins, accelerating aggregation and fibrillization of these proteins to form insoluble and rigid deposits in amyloidosis. The interaction between Hp/HS and an amyloid protein is critically dependent on the size of Hp/HS, and longer oligosaccharides (in a size range of dp4 to dp12) have a higher accelerating effect [69]. In addition, Hp/HS also mediate the internalization, cytotoxicity and stability of amyloids by accelerating the lateral aggregation of highly cytotoxic intermediate fibrils into nontoxic and insoluble mature fibrils, inhibiting the amyloid deposits from being proteolytically degraded [70].

Moreover, HSs function as cell surface receptors in many infectious diseases to facilitate pathogen invasion. Bacteria, viruses and protozoa all express HBPs [31]. These HSBs interact with cell surface HS chains, facilitating localization and invasion of host cells [71, 72]. For instance, initial contact and attachment of HSV-1 is mediated by the interaction between HSV-1 glycoproteins gB and gC with HS, and entry of HSV-1 into a cell is mediated by the interaction between glycoprotein gD and a specific HS sequence containing a 3-O-sulfated GlcN. Together, these HSV glycoprotein-HS interactions are responsible for the enhancement of HSV infectivity [73].

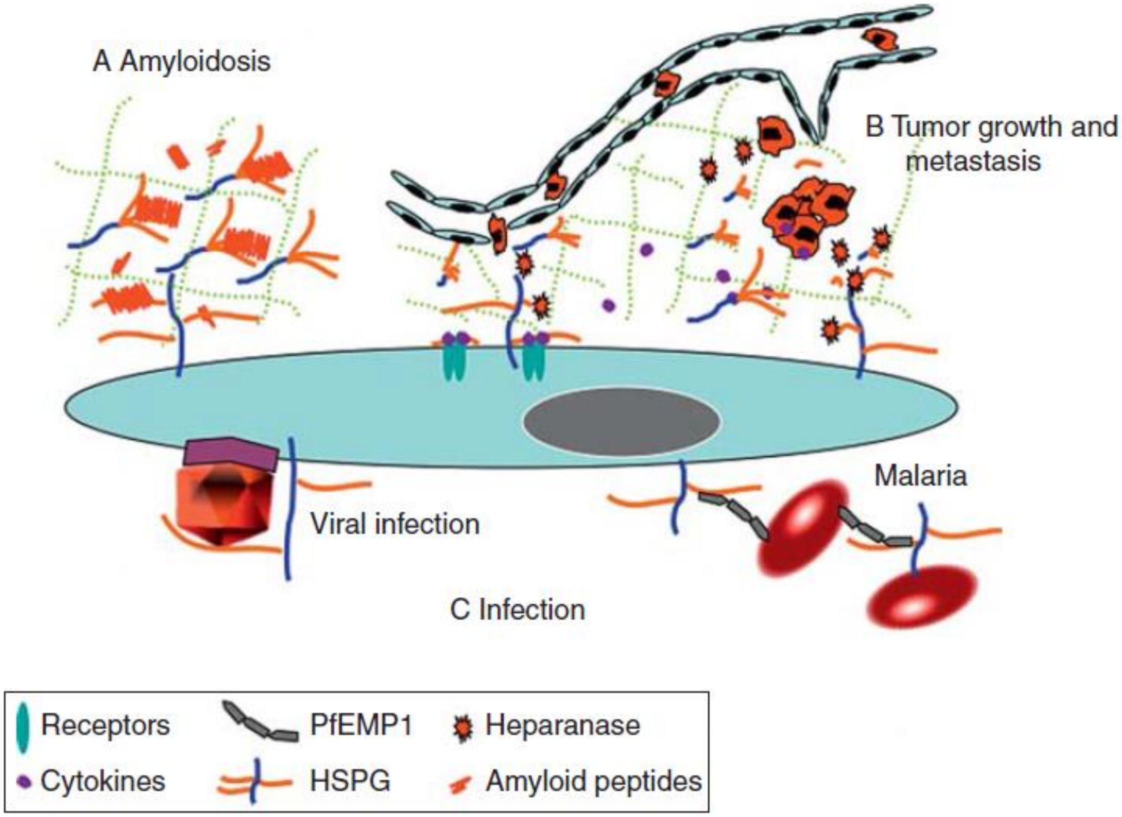


Figure 1.5 HS/Hp–protein interactions in selected diseases [53]

In addition, HS-protein interactions are also responsible for the development, angiogenesis and metastasis of tumor cells [49]. Many tumor-released growth factors are HBPs. Cell surface HSs act as co-receptors to bind and activate soluble tumor growth factors such as VEGFs and FGFs, therefore overexpression of HSPGs in tumor cells promotes tumor angiogenesis and growth [74]. HS also plays important role in tumor invasion. Overexpression of secreted heparanases leads to degradation of the ECM and basement membranes, which facilitates the penetration of tumor cells through vascular walls into circulation. The degradation of ECM also leads to increased mobility of cytokines and growth factors, which also promotes events including receptor activation such as adhesion and angiogenesis of tumor cells [75, 76].

Non-MS approaches for investigating GAG-protein interaction

Many biophysical approaches have been used to derive important structural information about GAG-protein interactions, including gel mobility shift assay [77], affinity chromatography [78], electrophoretic techniques [79], surface plasmon resonance (SPR) [80], isothermal titration calorimetry (ITC) [81, 82], circular dichroism (CD) [83], X-ray crystallography [60], NMR spectroscopy [84, 85] and molecular docking [86, 87], providing important structural information at the molecular or atomic level including: structures of protein-GAG complexes, GAG-binding sequences, interaction stoichiometry, thermodynamic information, kinetics information and binding strength of GAG-protein interactions [16].

High resolution techniques: NMR spectroscopy and X-ray crystallography

NMR spectroscopy and X-ray crystallography are the golden standard for high-resolution protein or protein complex structure characterization [88, 89]. X-ray crystallography has been used to provide atomic-level three-dimensional structural information of many native like proteins and GAG-protein complexes. Electrons of the atoms of the crystallized protein scatter X-rays, and the resulting scattered waves are recorded as a three-dimensional diffraction pattern, which can be processed to provide information about crystal packing symmetry, the repeat unit of the crystal and electron density [90]. Despite its popularity, there is still debate about whether the structures captured in a solid crystalline environment where the protein is trapped in a static state accurately represent the structures in a more physiologically relevant environment. The possible sources of artefacts include incomplete solvation and crystal packing. Also, not all proteins can be crystallized to form ordered or well-resolved structures, especially for those proteins with highly disordered structures and membranes proteins [91].

NMR spectroscopy, on the other hand, can be used to analyze protein structures and dynamics in a solvated environment at native pH values. NMR spectroscopy measures different resonant frequencies of NMR active nuclei upon different electromagnetic radiation, and the measured resonant frequency is sensitive to their local electronic environment. However, application of NMR spectroscopy is limited by protein size (the molecular weight limit is about 50kDa) [92]. The sensitivity of NMR measurements decreases with molecular weight, while the complexity of NMR spectra increases. Broadened lines are typically observed due to overlapped and rapidly decaying NMR signals. Both these two traditional high-resolution methods require relatively large quantities of highly purified sample from overexpression of recombinant proteins [93] and cannot be used on protein samples directly obtained from biological mixtures [94].

Circular dichroism (CD)

Circular dichroism (CD) measures the difference in the absorbance of left-handed and right-handed circularly polarized light by a solution of a chiral biomolecule [95]. Combining information from far-UV CD spectra, which corresponds to the absorption of peptide bond and near-UV CD spectra which corresponds to the absorption of aromatic side chains, quantitative analysis of secondary structures and qualitative characterization of tertiary or even quaternary structures of proteins and GAG-protein complexes can be achieved. Therefore, CD can be used for the detection of subtle conformational changes or folding/unfolding events involved in GAG-protein interactions. However, CD provides only a global measurement of structure and cannot detect the specific locations of secondary structure [95, 96].

Surface plasmon resonance (SPR)

Surface plasmon resonance (SPR) is a commonly used method for examining protein-ligand complexes. It measures the change in refractive index near a biosensor surface as the surface-immobilized binding partner (protein or GAG) forms a complex with the other binding partner introduced as a flowing solution. SPR provides a real-time measurement of the equilibrium bindings over a broad affinity range (pM to mM) as well as the kinetics of association and dissociation of GAG-protein interactions. However, this method is susceptible to errors from nonspecific binding. There are also concerns about whether an immobilized protein or GAG maintains its binding activity [97, 98].

Isothermal calorimetry (ITC)

In isothermal calorimetry (ITC), one binding partner (protein or GAG) is titrated in a thermostatic cell containing the solution of another binding partner, and the heat change required to maintain the temperature of the cell is tracked [99]. Therefore, ITC offers a direct thermodynamic characterization (enthalpy, entropy, free energy and stoichiometry) of GAG-protein interactions through a single experiment. However, ITC requires milligram quantities of sample and can only be used to measure a limited range of association constants [32].

MS approaches for investigating GAG-protein interaction

There is a growing need for new techniques which complement classic structural approaches but are more compatible with heterogeneous protein samples at physiologically relevant concentrations. Using mass spectrometry (MS) based approaches to study GAG-protein interactions has gained much recognition in recent years and become a powerful tool in structural biology, both by itself and in combination with other analytical techniques. Compared with those classic structural approaches, MS-based approaches have many advantages [93, 100]:

1) MS can analyze very small quantities of mixtures or heterogeneous samples due to its high sensitivity and high tolerance of impurities;

2) MS is not limited by protein size, which enables the analysis of large protein complexes up to several megadaltons;

3) MS can be used for simultaneous analysis of multiple ionic species in a mixture;

4) MS is free of the perturbating effect from solvents, counter-ions and other chemicals in solution.

More importantly, MS can be coupled with many other analytical techniques to expand its applications for the elucidation of protein structure, conformational dynamics and interactions with other biomolecules at different levels of structural resolutions. These combined MS approaches can be categorized into "non-native" and "native". For "non-native" MS approaches, proteins and protein complexes are typically modified followed by fragmentation with or without proteolytic digestion (bottom-up or top-down), and the mass increases are monitored on the peptide level to locate the site of modification. Measuring the differences of modification sites

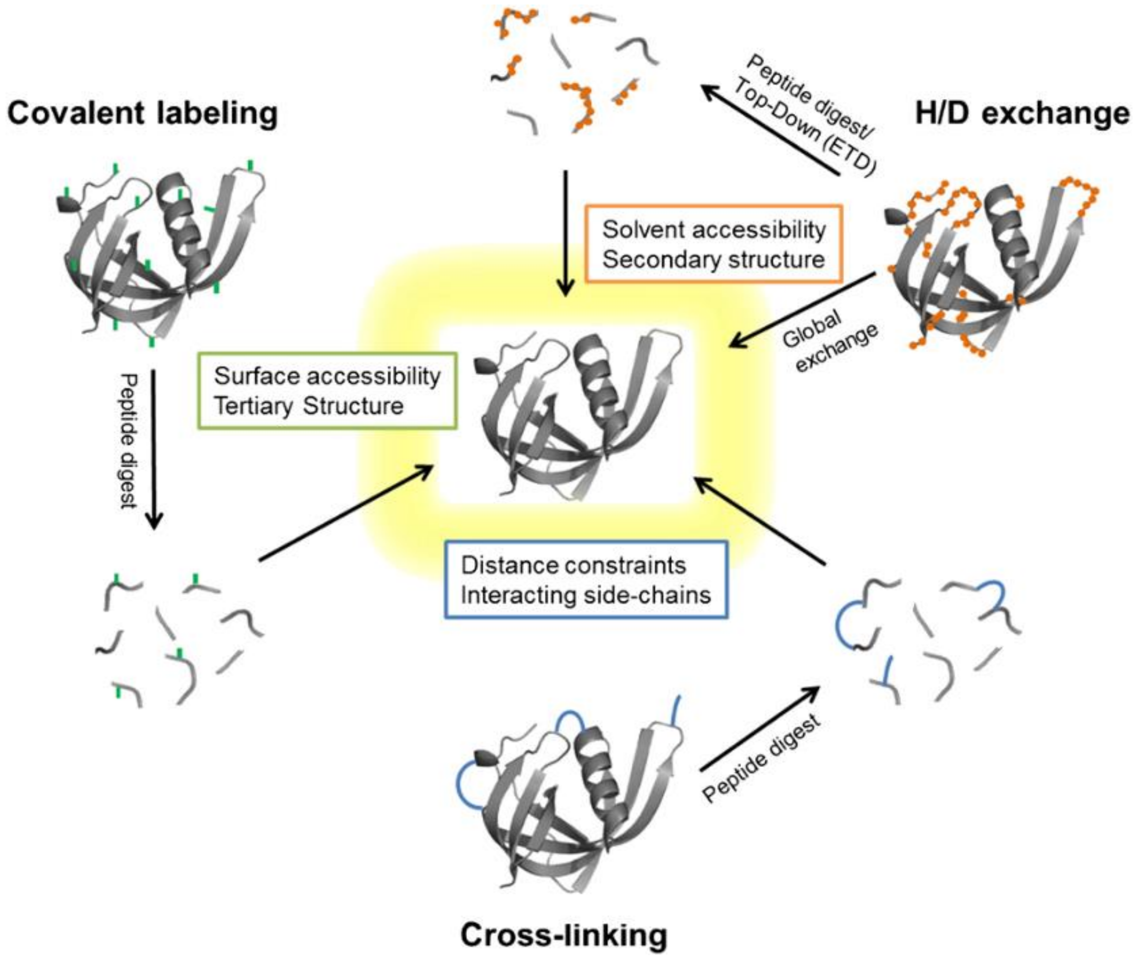


Figure 1.6 Overview of non-native mass spectrometry approaches [101]

and level before and after binding can be used to track the structural alteration induced by ligand binding. Some of the important “non-native” approaches are discussed in detail.

Hydrogen/deuterium exchange (HDX)

Hydrogen/deuterium exchange (HDX) monitors the isotope exchange rate of amide protons on protein backbone when exposed to a deuterated solvent. Amide protons that are located on protein surface or not involved in hydrogen bonds exchange rapidly, in contrast to slowly-exchanged amide protons that are either located in the interior of protein or involved in hydrogen bonds [102, 103]. The level of H/D exchange rate directly reflects the protein structure and dynamics, which is proportional to the flexibility and inversely proportional to protein stability, thus HDX is a useful technique for tracking dynamic protein structural fluctuations and stability-effect induced by GAG-protein interactions [104].

There are two ways to perform HDX experiment, continuous flow labeling or pulse-labeling. For continuous flow labeling, a protein undergoes a slow conformational transition while being continuously exposed to a deuterated solvent, and the resulting level of HDX increases as a function of incubation time and can be used to determine the conformational dynamics of the protein and integrated populations participating in the folding/ unfolding process. For pulse HDX, a protein is exposed only briefly to a pulse of a deuterated solvent after a conformational change or even a denaturing event. Since only the unfolded proteins can be labeled, the measured level of HDX determines the instantaneous population of short-lived folded and unfolded intermediates [105-107].

HDX is the mildest way to label proteins, so it is not likely to induce major changes in protein three-dimensional structure or stability. However, HDX is known to suffer from serious limitations. A primary issue is back-exchanging of deuteria with protons during analysis. Also, randomization of deuterium distribution can occur during analysis, referred as proton scrambling. To minimize or suppress back-exchange, a rapid digestion and liquid chromatography separation at low pH and 0 °C is required immediately after deuterium exchange. To minimize proton scrambling, electron based fragmentation methods (electron transfer dissociation (ETD) or electron capture dissociation (ECD) instead of vibrational excitation method (collision-induced dissociation (CID)) are preferred for determining sites of exchange [108].

Covalent labeling

Covalent labeling is used to non-reversibly modify exposed or unprotected protein surface. The most commonly used covalent labeling method is hydroxyl radical footprinting (HRF), which oxidizes side chains of exposed amino acid by OH radicals [109]. OH radical is one of the least selective covalent labels and can react with a wide variety of amino acids, but with different intrinsic reactivities. The level of covalent labeling using HRF is a function of both solvent accessibility and the chemical property of target sites [110, 111].

Side chain oxidation via HRF features a sub-millisecond time-scale, which allows the detection of short-lived transient structures. This time scale is much faster than most conformational changes [112], therefore the effects that covalent labels have on protein structure are not significant. Moreover, modification via HRF is highly stable, so it is free of proton migration issue as in HDX and single amino acid resolution can be achieved [101].

The biggest weakness of HRF is its susceptibility to “secondary effects”. For instance, factors other than changes in protein conformation or binding of ligands such as altered level of available hydroxyl radicals can also alter the oxidative levels detected by HRF [110]. The influence of secondary effect can be normalized by using unstructured peptides as internal standards [113].

Crosslinking

Crosslinking measures the distance between specific amino acids, and is therefore sensitive to the three-dimensional structure of a protein. Crosslinking uses an agent with two reactive ends separated by a spacer of a defined length. The reactive ends form covalent bonds to specific functional groups at the sidechains of amino acid, for example, the ε-amino group of lysine.

There are many considerations for selecting a proper crosslinker: chemical properties of the functional group determine the selectivity and reactivity of the crosslinker, whereas the distance between reactive amino acids determines the spacer length [114]. Other properties of the crosslinker including cleavability, ability to be enriched, isotope labeling, homo-functionality or hetero-functionality. Solubility and fragmentability are also important [102]. The most commonly used crosslinkers are N-hydroxysuccinimidyl or sulfosuccinimidyl esters whose target site is the primary amine of lysine [115]. The common target amino acids for crosslinking are highly reactive cysteine, lysine and aromatic amino acids including phenylalanine, tryptophan and tyrosine

Only amino acids that are located within the spacer distance can be cross-linked. Natively folded structures of proteins are not affected by the integration of crosslinkers. Crosslinking experiments can generate different types of products, including: dead-end products in which one reactive site is deactivated and does not form interaction with any amino acid; intra-peptide products in which two amino acid residues within the same peptide are bridged; inter-peptide products in which two amino acid residues from two different peptides are bridged. Dead-end products can help map the solvent accessible residues. Only inter-peptide products are structurally-informative and can be used for monitoring the folded structure. Data analysis to distinguish different products followed by information extraction from the less abundant inter-peptide products can be very challenging [116].

These analytical techniques discussed above are complementary to each other. For instance, HDX modifies backbone amides and provides data on solvent accessibility of residues within proteins, while crosslinking modifies side chains and provides data on the distance between different parts of a protein and protein complexes. Therefore, integration of different methods allows a more detailed overview of the structure and dynamics of proteins and GAG-protein complexes. The structural information obtained from these approaches can provide distance constraints or surface accessibility data to facilitate molecular modelling and docking studies, which can be critical when high-resolution (X-ray crystallography or NMR) protein or GAG-protein structures are not available.

Native mass spectrometry (native MS)

Native mass spectrometry (Native MS) allows a gentle transfer of an intact protein or protein complex from its solution environment into the solvent-free environment of a mass spectrometer without changing its natively folded conformation or disrupting the non-covalent interactions that maintain the structure of a protein complexes [93]. The protein or protein complexes are sprayed from buffer solutions, which is often an aqueous solution of ammonium acetate with a neutral pH to mimic physiological conditions. The capability of native MS to maintain the folded structures of proteins and keep their interactions with ligands intact, and provide the foundation for gas-phase structural biology studies.

Electrospray ionization (ESI)

In native MS experiments, protein ions are generated using electrospray ionization (ESI), one of the softest desorption/ionization techniques, which avoids unwanted fragmentation and conformational changes during ionization. ESI is much more widely used for native MS than the ionization method of matrix-assisted laser desorption/ionization (MALDI). Also, ESI is capable of generating multiple charged ions, which are required for analyzing protein and protein complexes of higher molecular weight.

In ESI, a solution containing the dissolved analyte is directly infused through an electroconductive spraying capillary, on which kilovolts of electric potential is applied under atmospheric pressure. A conventional ESI solution is a mixture of water and organic solvent with a low concentration of acid to aid the protonation process. However, the solution conditions denature the protein. For native ESI, the protein is introduced in an aqueous solution at neutral pH. The potential difference between the capillary tip and entrance aperture of the mass spectrometer results in formation and elongation of a Taylor cone, and subsequent generation of fine nanodroplets with excess charges. These nanodroplets undergo stepwise evaporation aided by nebulizing gas or heating, resulting in the shrinking of their sizes as their charge density and corresponding charge repulsion increase. Once the increased repulsion force overcomes droplet surface tension, successive charge-induced decomposition of droplets occurs. Eventually desolvated gaseous analyte ion are generated and ready to enter the vacuum region of mass spectrometer source [117, 118].

So far two mechanisms have been proposed for analyte ion generation:

1. Charged residue model (CRM), which proposes that gaseous analyte ion emerges from an exhaustive evaporation of the droplet to dryness. The size of droplet keeps decreasing until only one analyte molecule can accommodate, so the charge carried by the analyte ion is directly inherited from droplet surface. This mechanism is accepted to be accurate for the ion generation process for larger protein or protein complexes [119].

2. Ion evaporation model (IEM): which proposes that as the size of the droplet decreases, the repulsion force between charges increases and leads to the ejection of a charged analyte ion

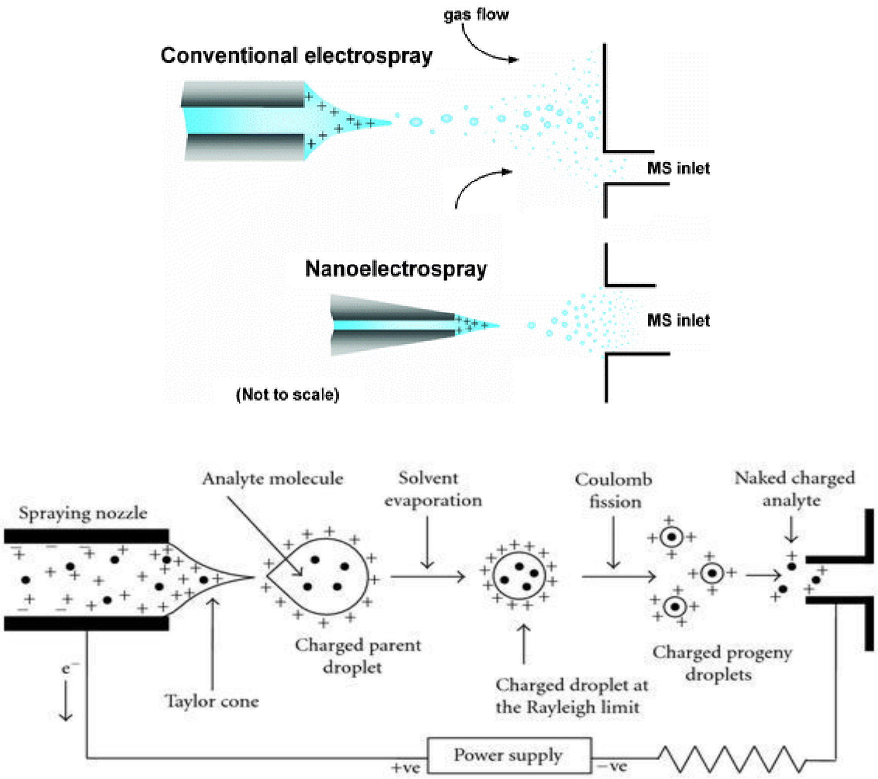


Figure 1.7 (a) Comparison of conventional ESI and nanoESI (modified from [122]); (b) Schematics of the formation of positive ions with conventional ESI [123].

before complete desolvation occurs. This mechanism is accepted to be accurate for the ion generation process for small molecules [120, 121].

ESI produces multiply charged ions. The charge state distribution generated can be directly related to the conformation that protein adopts in gas phase: folded or compact conformations have a smaller surface area with less exposed ionizable sites, thus a narrow distribution of low protonation states is observed. In contrast, for more elongated or unfolded conformations, a wide distribution of higher protonation states is observed. It is believed that the ion with the lowest charge state is most closely related to the native solution phase conformation [105].

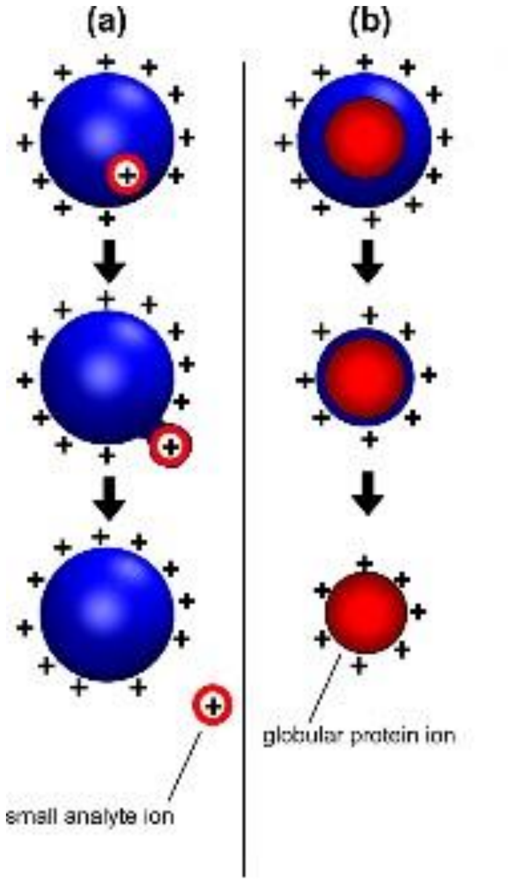


Figure 1.8 Schematic of ESI mechanisms: (a) IEM and (b) CRM [118]

Nanoelectrospray (NanoESI)

Nanoelectrospray (NanoESI), by its name, is a miniaturized version of normal ESI and was invented by Wilm and Mann [124, 125]. In normal ESI, a 50% aqueous, 50% volatile organic solvent at low pH is used with high spray voltages and high source temperatures to aid desolvation and ionization. Thus, ions generated from normal ESI have higher internal energy and are more likely to undergo unfolding or dissociation [126]. Compared with normal ESI, nanoESI uses a spray tip with a smaller diameter (1-10um) that is placed much closer to the entrance aperture of the mass spectrometer. It is softer, more efficient and sensitive (2000 times) than normal ESI in terms of producing finer nanodroplets and subsequent ionization of analyte ions [127]. These advantages of nanoESI are particularly important for native MS, for which aqueous buffer at physiological pH or water is used to preserve the native conformation and non-covalent interactions of protein and protein complexes. Also, due to the smaller dimeter of spray tip, the flow rate during direct infusion of the sample solution using nanoESI can be decreased from microliter per minute to nanoliter per minute, thus dramatically decreasing sample consumption [128].

Gas phase structures of proteins and protein complexes

Biomolecules and their interactions are a function of their aqueous environment, where hydration plays important roles in mediating the natively-folded conformation, stability and activity of proteins. Their structures result from hydrogen bonds with polar residues and the orienting of hydrophobic residues away from solution [129].

The transition of native structure of proteins from solution to gas phase is accompanied with loss of hydration around the protein, resulting in disruption of structure and changes in the relative strength of each interaction (hydrophobic, hydrogen bonding and electrostatic interactions). However, utilizing buffers of neutral pH and gentle experimental conditions, there have been many examples from previous studies suggesting that structures of proteins and their complexes in gas phase are very close to their biologically relevant solution structures [130-132]. A study by Breuker and McLafferty [133] suggest that in the gas phase, charged side chains on protein surface collapse within picosecond after desolvation, and form stronger electrostatic interactions within themselves or heteroatoms on the protein backbone. Another simulation study also showed that in a few nanoseconds after desolvation, the level of hydrogen bonds formed within the proteins increases from 43% to 56% [134]. Therefore, alterations of the equilibrium among these interactions promote a temporary stabilization of the native structure [135]. Moreover, residual solvents or buffer adducts can temporarily protect protein ions from structural rearrangements to form new gaseous structures [136, 137]. Overall, desolvated protein ions can get trapped in a local free energy minima and temporarily preserve most of its natively folded conformation for miliseconds. More importantly, previous studies also suggest that not

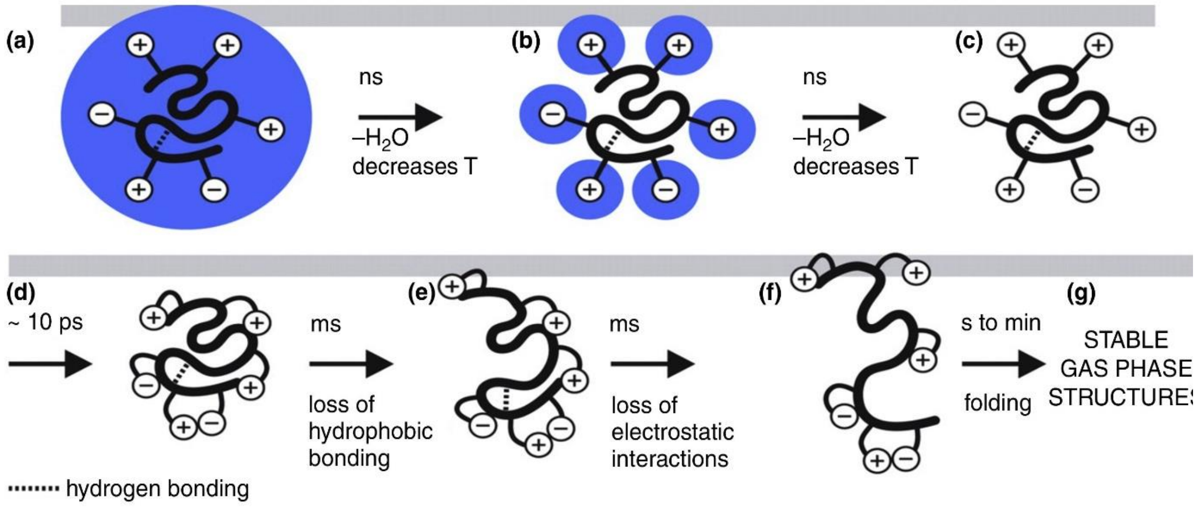


Figure 1.9 Stepwise evolution of protein structure after desolvation [133]

only is the non-covalent binding between protein and ligand maintained in the gas phase, but also ligand binding can provide additional gas phase stability of protein ions and further stabilization of their native structures [138, 139].

The development of nanoESI greatly facilitated the progress of native MS. Though nanoESI is an extremely soft ionization technique, tuning experimental settings is crucial in minimizing extensive activation and avoiding any undesired alteration of protein structures. Generally, native MS spectra exhibit lower intensities than non-native MS spectra and display undissolved clusters of broadened peaks, due to the incomplete dissolvation of salt or solvent adducted peaks that are overlaid with each other [140, 141]

Ion mobility spectrometry (IMS) coupled with native MS to study GAG-protein interactions

Recent years have witnessed enormous growth of using native MS coupled with ion mobility separation (IMS) to elucidate the structures of biomolecules and their relationship to their function. This appraoch is particularly useful for investigating large, dynamic and heterogeneous biomolecule systems that cannot be easily analyzed by classical methods.

Native MS allows protein ions to preserve their natively folded structure and allow protein complex ions to preserve its their noncovalent interactions, therefore information about the composition and binding stoichiometry of a biomolecule system can be achieved. IM provides a highly sensitive and rapid measurement of the size, shape and topology of proteins species as well as their conformational dynamics. Combing the advantages of both native MS and IMS, IMMS has been applied in the investigations of topology and architecture of protein complexes [142], conformational change of protein due to ligand binding [139], conformational dynamics and equilibrium [143, 144], transiently stable structures [145] and gas phase stability of protein and protein complexes [138].

Overview of ion mobility and ion mobility spectrometry (IMS)

The use of ion mobility spectrometry (IMS) as a new separating method was first reported at the beginnings of the 20th century [137]. Shortly after the hybridization of IMS with MS to support a study on gas phase ion-molecule reactions in the 1960s [146] and the launch of the first commercialized IMMS instruments by Waters in 2006 [147], IMS has witnessed tremendous growth and has been widely applied to the field of structural biology.

IMS separates and analyzes ions of interests in the gas phase based on their distinctive mobilities, which is size, shape and charge dependent. Compared with other chromatography based separation techniques, IMS does not have as a high peak capacity and resolving power but

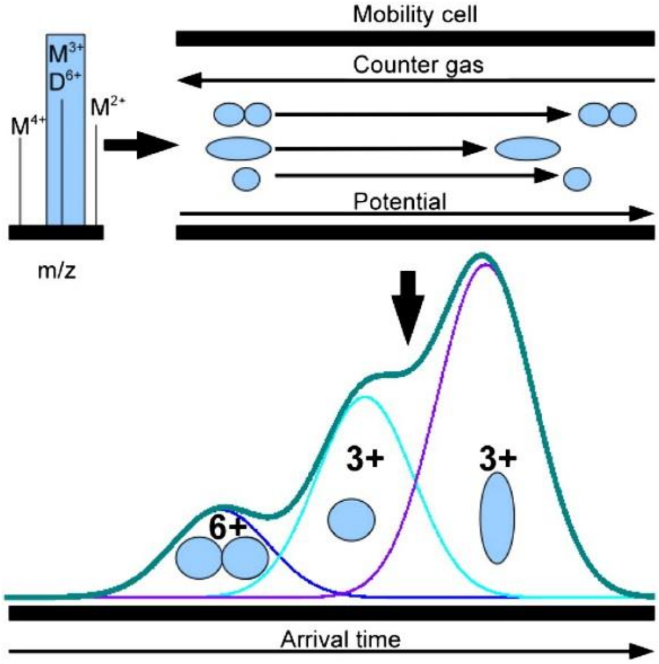


Figure 1.10 Schematics of ion mobility separation [101]

can separate ions in a 4 to 6 orders of magnitude less time [148]. Coupled with MS, IMMS can provide a low-resolution, two-dimensional characterization of both the mass and conformation of all biomolecular species in a heterogeneous assemblies [149].

There are three types of IMS: time-dispersive IMS, space-dispersive IMS and ion trapping and selective release IMS [150]. Temporally dispersive IMS separates ions based on their mobility-dependent arrival times as they flow through a drift tube. Spatially dispersive IMS detects ions based on their mobility-dependent path of motion. In ion confinement and release IMS, the conjugation of an electric field gradient and a gas flow is applied to control the spatial distribution of ions, which is dependent on their mobilities. Examples of each type of IMS will be discussed in the following sections.

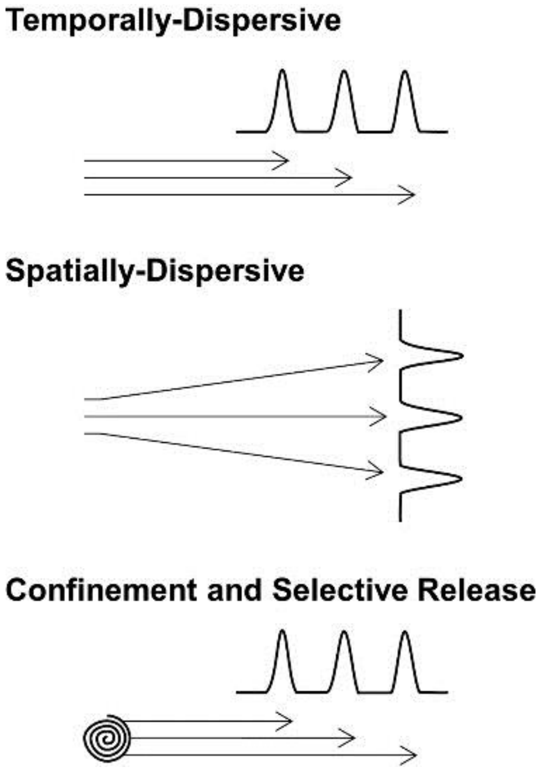


Figure 1.11 Different types of IMS [150]

Drift time ion mobility spectrometry (DTIMS)

Drift time ion mobility spectrometry (DTIMS) is the simplest platform for time-dispersive IMS. Drift time IMS separates gas-phased ions based on their mobilities when they drift through a chamber filled with inert neutral gas under the influence of a static, uniform and weak electric field (a range of 5 to 100 V/cm).

In the drift region of DTIMS, which varies in length from centimeters to meters, ions experience both the electrostatic force from electric field which pulls them through and the countered force from collisions with the neutral gas. As a result, ions quickly reach a thermal equilibration with a constant drift velocity V_d, which is proportional to the strength of the applied electric field:

$$K_d = \frac{\mathrm{D}}{\mathrm{t}} = \mathrm{KE}$$

Where E is strength of electric field, D is drift distance and t is drift time.

The proportionality constant K is defined as the mobility of ion, which, according to Mason−Schamp equation, depends upon the masses of the ion and neutral gas molecule, the charge carried by the ion and collisonal cross sections (CCSs) of both the ion and the neutral molecules [151, 152]. CCS represents the rotationally average shape from all orientations. The reduced ion mobility, K_0, normalized with respect to pressure P and temperature T is defined as:

$$K_0 = \frac{3\mathrm{ze}}{16\mathrm{N}\Omega}\sqrt{\frac{2\pi}{\mathrm{k_B T}\mu}}$$

where z is the ion charge, e is the elementary charge, N is the drift gas number density, k_B is the Boltzmann constant, T is the absolute temperature, Ω is the average CCS and μ is the ion-neutral reduced mass.

The resolving power of DTIMS is defined as:

$$R = \frac{t}{\Delta t} = \left(\frac{DEze}{16K_B T \ln 2}\right)^{1/2}$$

Therefore, increasing the magnitude of static electric field or length of drift tube can both improve its resolving power. Typically, the electric field strength of DTIMS is limited to around 10 to 30 V/cm to avoid the gas discharge at high electric field and ion loss at high gas pressure [153], so increasing the drift length of ions is a more commonly used to improve resolution. The

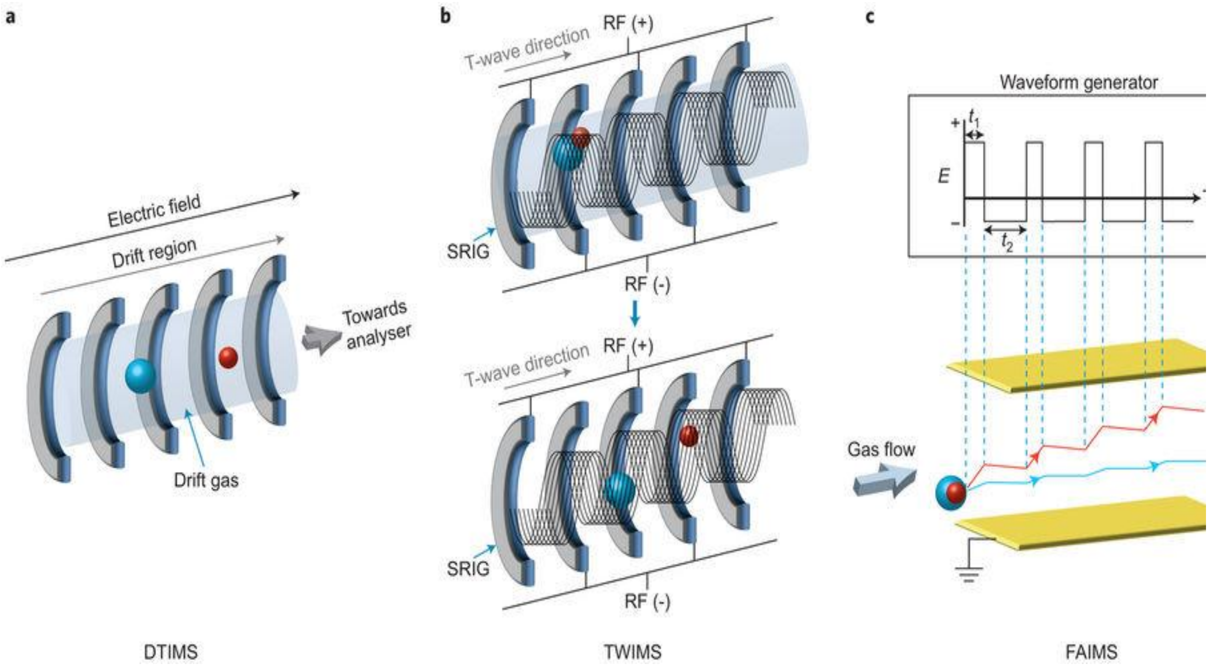


Figure 1.12 Schematics of the separation mechanism of DTIMS, TWIMS and FAIMS [100]

intrinsic limitations of DTIMS is its low duty cycle and low sensitivity because of the pulsed injections of ions.

Traveling wave ion mobility spectrometry (TWIMS)

Traveling wave ion mobility spectrometry (TWIMS) is another widely used time-dispersive platform of IM and the first commercially available IM-MS instrument. TWIMS platform was produced by Waters corporation as the Synapt HDMS system in 2006. Further modifications of the Synapt included incorporation of a helium cell prior to IM separation and alterations in source and ion transfer optics to improve its resolving power and ion transmission efficiency.

The basic design of TWIMS is very close to DTIMS, but their working mechanism are quite different. The IM cell of TWIMS contains three stacked ring ion guides: a trap ion guide for accumulating ions before being subjected to IM separation, a transfer ion guide for conveying mobility-separated ions for mass analysis and a IM ion guide, which is the essential component of the system. The IM ion guide is composed of a series of ring electrodes with radio frequency (RF) voltage in opposite phases being applied to adjacent electrodes, producing a radially confining potential barrier that restricts radial ion diffusion, increasing the sensitivity of TWIMS [154, 155]. To achieve mobility, a transient direct current (DC) voltage is superimposed on RF voltages and is swept to an adjacent electrode downstream at regular time intervals in the direction of ion migration, providing a continuous sequence of travelling waves. Ions are pushed by these waves, in a process that can be described as “surfing”. Ions with low mobility (larger CCS) cannot keep up with the traveling waves and slip behind, so they are transmitted much slower than those ions with high mobility (smaller CCS). Consequently, ions are separated based on their mobility (CCS) when appropriate speed and magnitude of travelling waves are applied. One intrinsic issue of TWIMS is that CCS cannot be directly converted from drift time due to the presence of constant changing electric field [156, 157], and so this value must be obtained by calibration.

High-field asymmetric waveform ion mobility spectrometry (FAIMS)

High-field asymmetric waveform ion mobility spectrometry (FAIMS) is a typical platform of space disperse IMS. Instead of measuring the drift time of an ion as in the time-dispersive IMS, FAIMS measures the voltage required to maintain a stable transmission of an ion [158]. FAIMS has two electrodes. One electrode is held at ground potential while an asymmetric waveform called dispersion voltage is applied to the other electrode. This asymmetric waveform is usually a square wave in which a high positive electric field being applied for a shorter time and a low negative electric field being applied for a longer time, while the product of voltage and time of each polarity of the waveform equals. The collision between ions and neutral gas molecules is more energetic in higher electric field, resulting in an electric field dependent high field mobility, which is different from the constant low field mobility [159].

Being exposed alternately to high and low electric fields, ions in FAIMS device oscillate radially between the two electrodes and tend to drift towards one of the electrode as a result of their different mobilities in high and low fields. Therefore, an compensation voltage (CV) is required to balance the offset and stop the ions from being neutralized upon contacting with the electrode. CV is compound specific and is dependent on the mobility difference of one ion in low and high electric fields. FAIMS cannot be used to calculate absolute CCS of ions, it is typically used as a mobility filter, especially a pre-MS narrowband filter. The specific CV of an ion of interest can be pre-determined by scanning across a range of voltage. Then this CV is constantly applied to maintain the stable path of only ion of interest while other ions are removed [158].

Trapped ion mobility spectrometry (TIMS)

Trapped Ion Mobility Spectrometry (TIMS) utilizes selective release to achieve ion mobility separation. In TIMS, ions are entrained in a flow of neutral gas, and an axial electric field gradient is applied to oppose the gas flow. An RF voltage is applied to ring electrodes to generate a confining quadrupolar potential, which is used to trap ions radially and improve ion transmission efficiency. When the drift velocity of an ion equals the velocity of gas flow, this ion is trapped at an equilibrium position along the slope of the electric field gradient. Since the trapped position is mobility-dependent: $V_d = KE$, and the magnitude of electric field increases

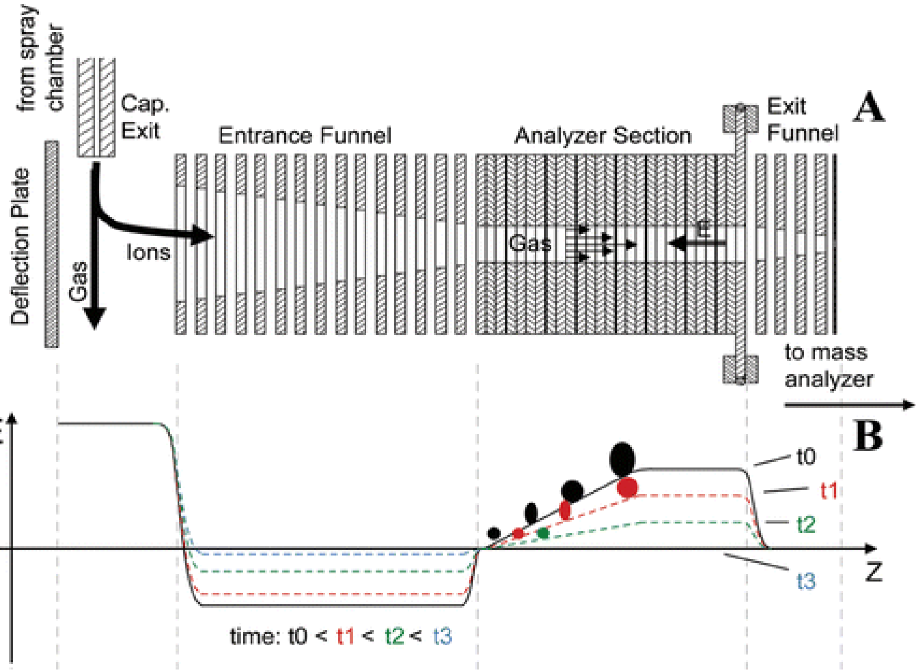


Figure 1.13 Schematics of the separation mechanism of TIMS [162]

along the gradient, ions with lower mobility are trapped near the exit of analyzer while ions with higher mobility being trapped near the entrance. As the magnitude of the ramping voltage is gradually lowered, a sequential ion release is prompted, and ions with lower mobility are eluted and detected first [160-162].

Quadrupole mass analyzer

A quadrupole mass analyzer is composed of two pairs of parallel electrical rods. Different phases of potential containing both DC and RF components are applied to each pair of rods to form a quadrupolar electric field. The Mathieu parameters of each component is defined as:

$$\mathrm{qM} = \frac{4\mathrm{eV}}{(\mathrm{m/z})\omega^2 \mathrm{r}^2}$$

$$\mathrm{aM} = \frac{8\mathrm{eU}}{(\mathrm{m/z})\omega^2 \mathrm{r}^2}$$

where e is the charge of an electron, V and ω are the amplitude and angular frequency of the RF voltage, respectively, U is the value of the DC voltage and r is the radius of the quadrupole.

The applied DC and RF voltage determines the way ions oscillate in the quadrupole. When operated in an isolation mode, a certain combination of V, U, and ω allows only ions within a very narrow range of m/z to travel in stable trajectories and get transmitted. When operated in an RF-only or a wideband scanning mode in which no U is applied, ions below a certain m/z ratio adopt instable trajectories and get filtered out [163].

The upper limit of the m/z that can be transmitted by quadrupole is proportional to the amplitude of RF voltage, and inversely proportional to the frequency of RF voltage and the inner radius of the rods. To avoid the use of high voltage as well as ion loss due to smaller radius, the frequency of the quadruple is uaually modulated, when coupled with native MS, to optimized for the transmission of ions with larger m/z [164, 165].

Time-of-flight (TOF) mass analyzer

The Synapt instrument is equipped with a time-of-flight (TOF) mass analyzer, taking advantage of its rapid data acquisition rate to record millisecond wide ion mobility peaks as well as its effectively unlimited mass range to measure biomolecules of large size.

In TOF, pulses of ions with different mass but same charge state are accelerated to the same kinetic energy through a homogeneous electric field, then they fly though a field-free flight tube before being detected [167].

$$\mathrm{KE} = \frac{1}{2}\mathrm{mv}^2 = \mathrm{PE} = \mathrm{zU}$$

where PE is the potential energy, z is the charge of the ion, U is the magnitude of accelerating electric field, KE is the kinetic energy, m is the mass of the ion and v is the velocity.

Therefore, the flight time that ions take to drift though a flight region of length d is proportional to the square root of m/z and is defined as:

$$\mathrm{t} = \frac{\mathrm{d}}{\sqrt{2\mathrm{U}}}\sqrt{\frac{\mathrm{m}}{\mathrm{z}}}$$

However, it is inevitable that ions of same mass may have different starting location and different initial velocities. Therefore, reflectrons which are composed of a series of electrostatic ion mirrors are placed at the end of the flight region and are used to correct for the distribution of KE while reversing the direction of ion motion, resulting in improved resolution and narrower peaks for TOF measurements [168]. Ion with a higher KE thus higher velocity can penetrate further into the reflectron and spend longer time in the retarding electric field, whereas ion with a lower KE bounces out in a shorter amount of time, therefore the velocity differences can be corrected and ions of same m/z arrive at the detector simultaneously [169, 170].

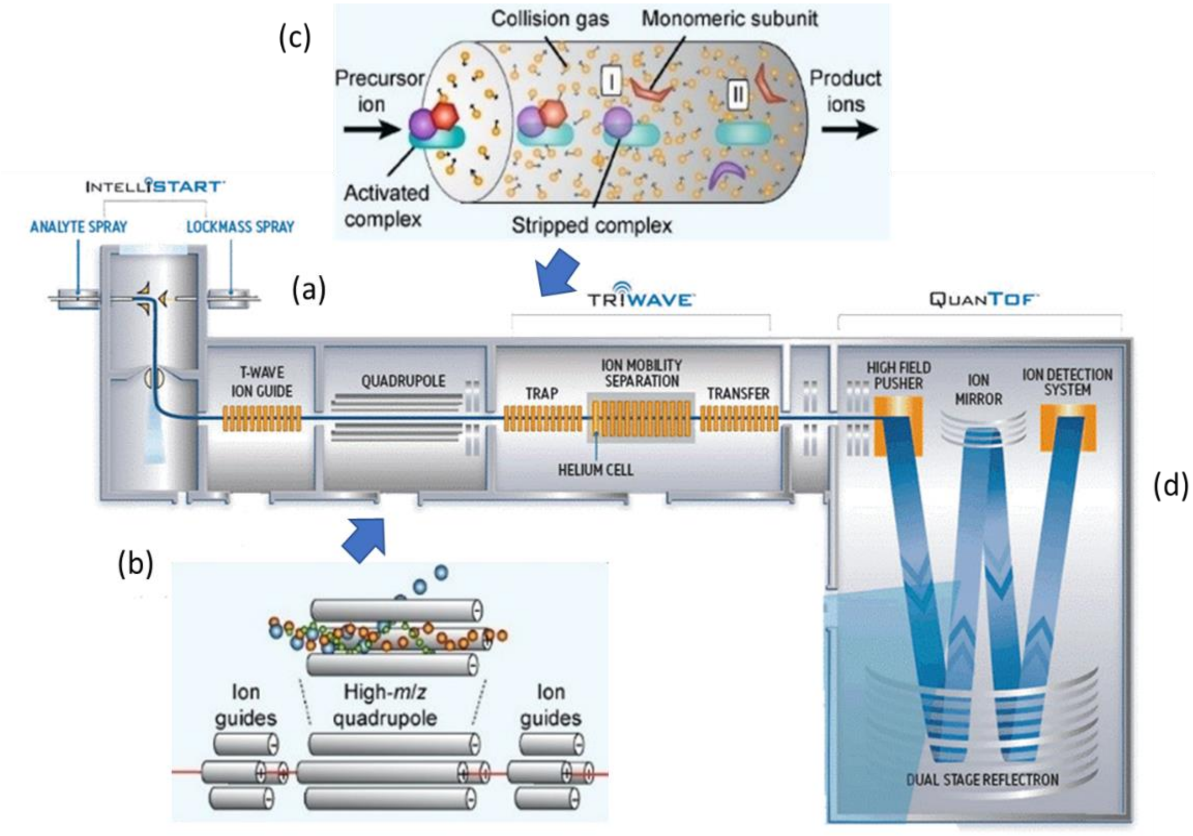


Figure 1.14 (a) Schematics of Synapt G2 mass spectrometer; (b) quadrupole mass filter; (c) mechanism of CIU; (d) TOF mass analyzer (modified from [166])

Collision-induced unfolding (CIU) for analyzing protein stability

The gas phase collision-induced unfolding (CIU) experiment takes advantage of the slow-heating CID approach [171], and has a similar mechanism as solution-based thermal denaturation experiment.

For TWIMS instrument, CIU typically occurs in the trap ion guide, in which protein ions undergo activation through collisions with neutral gas molecules while the applied electric field keeps increasing. During CIU, protein ions gain additional kinetic energy, which accumulates to increase the total internal energy of the protein ions and cause structural distortion or even dissociation. The activated/unfolded protein ions are conveyed to the IM ion guide and MS analyzer, in which appearance and disappearance of each structure/conformation as the collisional energy (CE) increases can be monitored and recorded. The collisional energy required to convert the native, folded conformation to the first intermediate structure is directly indicative of gas-phase stability of the original structure [139]. Recently, the Ruotolo group developed a CIU suite software package to aid analysis and interpretation of CIU results. The CIU fingerprints are generated by plotting drift time (or CCS) of different protein conformations and their intensity against applied collisional energy [172], providing an easy way to visualize the unfolding and dissociation pathway of protein ions.

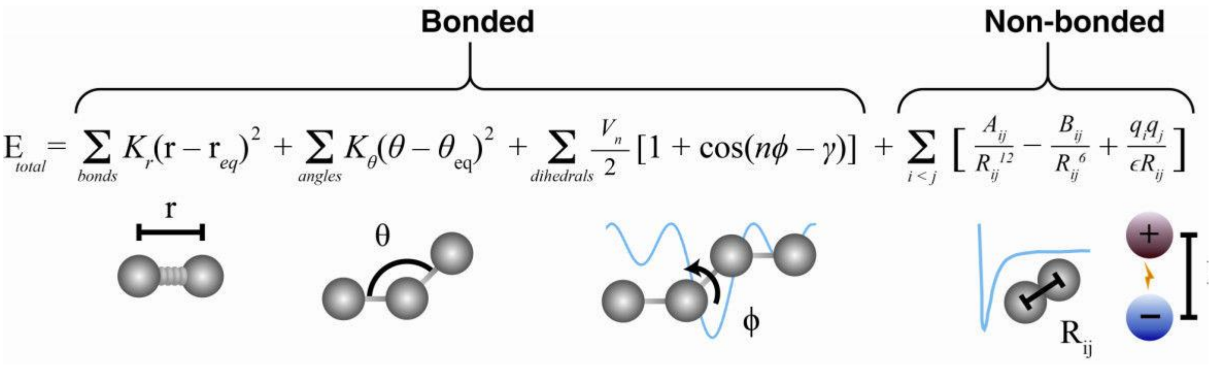
Molecular dynamic (MD) simulation and its application in structural biology

Force fields

Molecular dynamic (MD) simulations use a semi-empirical, molecular mechanics based force field to describe the potential energy of a biological system as a function of its three-dimensional structural coordinates. Since the repetitive calculations of movements and interactions of all the atoms of even the simplest biomolecular system is computationally infeasible using quantum mechanics, molecular mechanics is chosen as a simplified but only applicable solution.

The potential energy of a biological system is represented by energy terms that describe both bonded and nonbonded contributions [173]. Terms for covalently bonded interaction and atomic angle are modeled using harmonic potentials, which describe chemical bonds as spherical atoms connected by springs. Dihedral angle terms are modeled using a sinusoidal function to describe energy differences between eclipsed and staggered conformations. For nonbonded interactions, electrostatic terms are described by coulombic potential, and can der Waals term is described by Lennard-Jones potentials that account for both attractive and repulsive forces [174]. The parameter set for a force field is pre-determined from both quantum mechanical calculations and experimental data of small molecules, including stiffness and lengths of bonds, partial atomic charges to describe electrostatic interactions and atomic VDW radii.

Commonly used force fields include: CHARMM (Chemistry at Harvard Molecular Mechanics) [175], AMBER (assisted model building with energy refinement) [176], GROMOS (Groningen molecular simulation) [177] and CVFF (consistent valence force field) [178].



$$E_{total} = \underbrace{\sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]}_{\text{Bonded}} + \underbrace{\sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right]}_{\text{Non-bonded}}$$

Figure 1.15 The expression of force field [174]

Overview of MD simulation

A MD simulation approximates each atom in a biological system as a single particle with a partial charge and a van der Waals radii. According to classical mechanics, the atomic positions and velocities evolve according to the laws of classical physics, and a MD simulation can be considered as a process of iteratively solving Newton's law of motion to predict the time dependent behavior of the system.

Specifically, using the force field with a pre-defined parameter set, the potential energy function $U(r_1, r_2, \ldots, r_n)$ of a biological system can be described. At a given time point (t), every atom with a point mass (m_i) has a velocity (v_i) at the position indicated by a three-dimensional vector (r_i). The atom is under the influence of external forces (F_i), which can be calculated based on the gradient of potential energy. Then based on Newton's law of motion, acceleration (a_i) of each atom can be calculated [173, 179].

$$F_i = -\frac{dU}{dr_i}$$

$$a_i = \frac{d^2x(t)}{dt^2} = \frac{F_i}{m_i}$$

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2$$

$$v(t + \Delta t) = v(t + 0.5\Delta t) + 0.5a(t + \Delta t)\Delta t$$

The time scale of MD simulation is divided into small steps separated by a short period of Δt. Combining the atomic positions and movements at time (t) and prior step (t - Δt), atomic positions and movements at next step (t + Δt) can be predicted. Therefore, the updated atomic positions and velocities throughout the time scale of a MD simulation can be calculated [179, 180].

MD simulations require a thermodynamic ensemble to simulate multiple experimental conditions. The ensemble defines the bulk thermodynamic properties of a biomolecular system, including number of atoms (N), temperature (T), pressure (P) and mean energy (E). Commonly used ensembles include a microcanonical ensemble (NVE, constant particle number, volume and energy), canonical ensemble (NVT, constant temperature, volume and number of atoms) and isothermal-isobaric ensemble (NPT, constant pressure, temperature and number of atoms).

The output of a MD simulation is recorded in a trajectory, which consists of individual structure from each snapshot of simulation and describes coordinates, motion and energy of each particle of a biological system as a function of time.

The application of MD simulation in structural biology

MD simulation is specifically useful when dealing with interchangeable transient conformations of biomolecules that do not form ordered or well-resolved structures, and cannot be analyzed using the high resolution X-ray crystallography [181]. Recent advances in force field and underlying physical models, as well as directly incorporating experimental data to restrain MD simulations have greatly improve their accuracy [182]. The computational expense and speed of MD simulations have also been greatly enhanced due to breakthroughs in computer hardware and parallelization algorithms [183]. The dramatically improved time scale of MD simulations permit simulation of milliseconds long biological processes such as conformational change of protein, protein-drug binding and membrane transport [184].

With all these improvements, MD simulation has become a main computational approach in structural biology and has been widely applied in the investigations of: protein stability [185], protein allosteric properties [186], molecular recognition [187], ion and small molecule transport [188, 189], formation of protein complexes [190, 191], protein hydration [192] as well as protein design [193] and drug design [194].

Interpretation of IMMS data

IMMS studies are often supported by MD simulation and other molecular modeling studies, which provide computational prediction of the evolution of structure and dynamics of biomolecules over time, binding orientations and binding affinities of complexes as well as their molecular shape, essential for the correct interpretation for IMMS experimental data and confirmation of the biological relevance of gas phase structure measured by IMMS. For a more accurate comparison to experimentally measured, gas phase structures, protein structures obtained using X-ray crystallography needs to be refined through MD simulations and energy minimizations: an energy minimization is performed to remove any possible sidechain clashes or structural artefacts embed in the original protein structure in the presence of crystal-packing environment, followed by a solution phase MD simulation in which protein structure is fully solvated by water to mimic the physiological environment. Lastly, an *in-vacuo* energy minimization is performed to account for possible alterations of protein structure in the gas phase environment of the mass spectrometer [195]. Moreover, the trajectory of a MD simulation is also required as input for theoretical estimation of molecular shape (CCS) using the trajectory method. Vice versa, experimentally determined CCS can be used to discriminate between different predicted models corresponding to similar energy minima.

Calculation of experimental CCS

Based on the Mason−Schamp equation, the experimental CCS measured using IMS is direct proportional to the measured drift time when a low and uniform electric field is applied [196].

$$\mathrm{t} \propto \frac{1}{\mathrm{K}} \propto \frac{\Omega\sqrt{\mu}}{\mathrm{z}}$$

The main limitation of TWIMS is that its electric field is no longer uniform due to the use of transient DC voltage whose electric potential keeps changing over time and position within IM ion guide. As a result, drift time measured from TWIMS cannot be directly converted into CCS. Instead, a calibration curve needs to be constructed by plotting the known CCSs of calibrants, corrected for their charge state and reduced mass with respect to buffer gas, against the drift time of calibrants which are measured under exact same experimental condition as that for analytes. The drift times also need to be corrected for mass-dependent flight time spent in the transfer ion guide and TOF mass analyzer and mass-independent flight time spent in the transfer ion guide. The calibration curve can be used to estimate CCSs of analytes given the drift times of the analytes [197, 198].

Comparing experimental CCSs with theoretical CCSs

The most common way to calculate theoretical CCSs is through the open source program MOBCAL, developed by Shvartsburg and Jarrold [199]. Three methods are available from MOBCAL: projection approximation (PA) [200], exact hard sphere scattering (EHSS) [199] and trajectory method (TM) [201].

As the simplest and least computationally demanding method, the PA method approximates CCS by projecting the molecule onto a plane followed by estimating orientationally averaged projected area from all possible orientations through a Monte Carlo integration [195]. This method is known to underestimate CCS by up to 20%, due to neglecting the influence of long-range interactions as well as scattering and collision between ion and neutral gas. Therefore, the PA method is not appropriate for measuring CCS of a molecule whose mass is greater than 2kDa [202].

A previous study by Ruotolo and coworkers suggest that there is a direct, linear scaling relationship between experimental CCSs and values calculated using the PA method, thus the PA results can be corrected for scattering phenomenon based on a linear scaling function to generate the scaled PA results, which should agree better with experimental data [203, 204].

$$CCS(cal) = 1.14 \times CCS(PA)$$

The EHSS method calculates CCS by estimating the averaged momentum transfer cross section over relative velocity and collision geometry. This method fails to consider the influence of long-range interactions between ion and neural gas and is known to overestimate CCS [199].

In contrast, the TM method treats the biomolecule as a collection of atoms defined by their Lennard-Jones potential, and relates the scattering angle of buffer gas molecules before and after collisions with the geometry of the biomolecule. CCS can be obtained by integrating over all possible collision geometries. The TM provides the most accurate results since it considers both the effects of long range interactions and effects of collisions and scattering. However, the TM is known to be computationally more expensive than other methods, which limits its application for calculating CCS of proteins and protein complexes of larger sizes.

An improved version of the PA method, the projection superposition approximation (PSA) method was proposed by Bleiholder [205]. The PSA uses the same mechanism of the PA method while taking the shape of protein ion and neural gas molecule into consideration by incorporating a tunable shape factor, therefore providing a better description of long-range interactions. The PSA can be used as a replacement of the TM method, with the benefit of lower computational expense.

Typically, experimental CCSs derived from the lowest observed charge state are found to be smaller than TM or PSA results and are closer to scaled PA results. The difference between experimental CCS and TM CCS can be as high as 40%, and is believed to be a result of the partial collapse of protein structure upon desolvation. Previous studies suggest that the CCS difference increases as the size of protein ion increases [202]. Therefore, a gas-phase minimization should be performed to mimic the vacuum environment in mass spectrometer before calculating theoretical CCSs and comparing them to experimental CCSs.

References

1. Mende, M., Bednarek, C., Wawryszyn, M., Sauter, P., Biskup, M.B., Schepers, U., et al.: Chemical synthesis of glycosaminoglycans. Chemical reviews. **116**, 8193-8255 (2016)

2. Wiederschain, G.Y.: Essentials of glycobiology. Biochemistry (Moscow). **74**, 1056-1056 (2009)

3. Necas, J., Bartosikova, L., Brauner, P., Kolar, J.: Hyaluronic acid (hyaluronan): a review. Veterinarni medicina. **53**, 397-411 (2008)

4. Lever, R., Mulloy, B., Page, C.P. Springer Science & Business Media, (2012)

5. Dosio, F., Arpicco, S., Stella, B., Fattal, E.: Hyaluronic acid for anticancer drug and nucleic acid delivery. Advanced drug delivery reviews. **97**, 204-236 (2016)

6. Collins, M.N., Birkinshaw, C.: Hyaluronic acid based scaffolds for tissue engineering—A review. Carbohydrate polymers. **92**, 1262-1279 (2013)

7. Lamari, F.N., Theocharis, A.D., Asimakopoulou, A.P., Malavaki, C.J., Karamanos, N.K.: Metabolism and biochemical/physiological roles of chondroitin sulfates: analysis of endogenous and supplemental chondroitin sulfates in blood circulation. Biomedical Chromatography. **20**, 539-550 (2006)

8. Volpi, N.: Dermatan sulfate: recent structural and activity data. Carbohydrate Polymers. **82**, 233-239 (2010)

9. Sugahara, K., Mikami, T., Uyama, T., Mizuguchi, S., Nomura, K., Kitagawa, H.: Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. Current opinion in structural biology. **13**, 612-620 (2003)

10. Funderburgh, J.L.: MINI REVIEW Keratan sulfate: structure, biosynthesis, and function. Glycobiology. **10**, 951-958 (2000)

11. Thompson, S., Martínez-Burgo, B., Sepuru, K.M., Rajarathnam, K., Kirby, J.A., Sheerin, N.S., et al.: Regulation of Chemokine Function: The Roles of GAG-Binding and Post-Translational Nitration. International Journal of Molecular Sciences. **18**, 1692 (2017)

12. Xu, D., Esko, J.D.: Demystifying heparan sulfate–protein interactions. Annual review of biochemistry. **83**, 129-157 (2014)

13. Lin, X.: Functions of heparan sulfate proteoglycans in cell signaling during development. Development. **131**, 6009-6021 (2004)

14. Bülow, H.E., Hobert, O.: The molecular diversity of glycosaminoglycans shapes animal development. Annu. Rev. Cell Dev. Biol. **22**, 375-407 (2006)

15. Poulain, F.E., Yost, H.J.: Heparan sulfate proteoglycans: a sugar code for vertebrate development? Development. **142**, 3456-3467 (2015)

16. Monneau, Y., Arenzana-Seisdedos, F., Lortat-Jacob, H.: The sweet spot: how GAGs help chemokines guide migrating cells. Journal of leukocyte biology. **99**, 935-953 (2016)

17. Kramer, K.L., Yost, H.J.: Heparan sulfate core proteins in cell-cell signaling. Annual review of genetics. **37**, 461-484 (2003)

18. Esko, J.D., Lindahl, U.: Molecular diversity of heparan sulfate. Journal of Clinical Investigation. **108**, 169 (2001)

19. Sasisekharan, R., Venkataraman, G.: Heparin and heparan sulfate: biosynthesis, structure and function. Current opinion in chemical biology. **4**, 626-631 (2000)

20. Lindahl, U., Li, J.p.: Interactions between heparan sulfate and proteins—design and functional implications. International review of cell and molecular biology. **276**, 105-159 (2009)

21. Sarrazin, S., Lamanna, W.C., Esko, J.D.: Heparan sulfate proteoglycans. Cold Spring Harbor perspectives in biology. **3**, a004952 (2011)

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

23. Ly, M., Laremore, T.N., Linhardt, R.J.: Proteoglycomics: recent progress and future challenges. Omics: a journal of integrative biology. **14**, 389-399 (2010)

24. Gandhi, N.S., Mancera, R.L.: The structure of glycosaminoglycans and their interactions with proteins. Chemical biology & drug design. **72**, 455-482 (2008)

25. Gallagher, J.T.: Heparan sulfate: growth control with a restricted sequence menu. Journal of Clinical Investigation. **108**, 357 (2001)

26. Kreuger, J., Matsumoto, T., Vanwildemeersch, M., Sasaki, T., Timpl, R., Claesson-Welsh, L., et al.: Role of heparan sulfate domain organization in endostatin inhibition of endothelial cell function. The EMBO journal. **21**, 6303-6311 (2002)

27. Robinson, C.J., Mulloy, B., Gallagher, J.T., Stringer, S.E.: VEGF165-binding sites within heparan sulfate encompass two highly sulfated domains and can be liberated by K5 lyase. Journal of Biological Chemistry. **281**, 1731-1740 (2006)

28. Spillmann, D., Witt, D., Lindahl, U.: Defining the interleukin-8-binding domain of heparan sulfate. Journal of Biological Chemistry. **273**, 15487-15493 (1998)

29. Stringer, S.E., Gallagher, J.T.: Specific binding of the chemokine platelet factor 4 to heparan sulfate. Journal of Biological Chemistry. **272**, 20508-20514 (1997)

30. Sarkar, A., Desai, U.R.: A simple method for discovering druggable, specific glycosaminoglycan-protein systems. Elucidation of key principles from heparin/heparan sulfate-binding proteins. PloS one. **10**, e0141127 (2015)

31. Rostand, K.S., Esko, J.D.: Microbial adherence to and invasion through proteoglycans. Infection and immunity. **65**, 1 (1997)

32. Capila, I., Linhardt, R.J.: Heparin–protein interactions. Angewandte Chemie International Edition. **41**, 390-412 (2002)

33. Sanderson, P.N., Huckerby, T.N., Nieduszynski, I.A.: Conformational equilibria of α-L-iduronate residues in disaccharides derived from heparin. Biochemical Journal. 243, 175-181 (1987)

34. Angulo, J., Hricovíni, M., Gairi, M., Guerrini, M., De Paz, J.L., Ojeda, R., et al.: Dynamic properties of biologically active synthetic heparin-like hexasaccharides. Glycobiology. **15**, 1008-1015 (2005)

35. Teran, M., Nugent, M.A.: Synergistic binding of vascular endothelial growth factor-A and its receptors to heparin selectively modulates complex affinity. Journal of Biological Chemistry. **290**, 16451-16462 (2015)

36. Li, W., Johnson, D.J., Esmon, C.T., Huntington, J.A.: Structure of the antithrombin–thrombin–heparin ternary complex reveals the antithrombotic mechanism of heparin. Nature structural & molecular biology. **11**, 857-862 (2004)

37. Shukla, D., Liu, J., Blaiklock, P., Shworak, N.W., Bai, X., Esko, J.D., et al.: A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. Cell. **99**, 13-22 (1999)

38. Schenauer, M.R., Yu, Y., Sweeney, M.D., Leary, J.A.: CCR2 chemokines bind selectively to acetylated heparan sulfate octasaccharides. Journal of Biological Chemistry. **282**, 25182-25188 (2007)

39. Ashikari-Hada, S., Habuchi, H., Kariya, Y., Itoh, N., Reddi, A.H., Kimata, K.: Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. Journal of Biological Chemistry. **279**, 12346-12354 (2004)

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

41. Knappe, M., Bodevin, S., Selinka, H.-C., Spillmann, D., Streeck, R.E., Chen, X.S., et al.: Surface-exposed amino acid residues of HPV16 L1 protein mediating interaction with cell surface heparan sulfate. Journal of Biological Chemistry. **282**, 27913-27922 (2007)

42. Zong, C., Venot, A., Dhamale, O., Boons, G.-J.: Fluorous supported modular synthesis of heparan sulfate oligosaccharides. Organic letters. 15, 342-345 (2013)

43. Liu, J., Linhardt, R.J.: Chemoenzymatic synthesis of heparan sulfate and heparin. Natural product reports. **31**, 1676-1685 (2014)

44. Petitou, M., Casu, B., Lindahl, U.: 1976–1983, a critical period in the history of heparin: the discoveryof the antithrombin binding site. Biochimie. **85**, 83-89 (2003)

45. Muir, J.M., Andrew, M., Hirsh, J., Weitz, J., Young, E., Deschamps, P., et al.: Histomorphometric analysis of the effects of standard heparin on trabecular bone in vivo. Blood. **88**, 1314-1320 (1996)

46. DeLisser, H.M., Yan, H.C., Newman, P., Muller, W.A., Buck, C., Albelda, S.: Platelet/endothelial cell adhesion molecule-1 (CD31)-mediated cellular aggregation involves cell surface glycosaminoglycans. Journal of Biological Chemistry. **268**, 16037-16046 (1993)

47. Holroyd, E.W., Simari, R.D.: Interdependent biological systems, multi-functional molecules: the evolving role of tissue factor pathway inhibitor beyond anti-coagulation. Thrombosis research. **125**, S57-S59 (2010)

48. Kim, S.-H., Turnbull, J., Guimond, S.: Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. Journal of Endocrinology. **209**, 139-151 (2011)

49. Fuster, M.M., Esko, J.D.: The sweet and sour of cancer: glycans as novel therapeutic targets. Nature Reviews Cancer. **5**, 526-542 (2005)

50. Yoshimura, T., Matsushima, K., Tanaka, S., Robinson, E.A., Appella, E., Oppenheim, J.J., et al.: Purification of a human monocyte-derived neutrophil chemotactic factor that has peptide sequence similarity to other host defense cytokines. Proceedings of the National Academy of Sciences. **84**, 9233-9237 (1987)

51. Mosnier, L.O., Zlokovic, B.V., Griffin, J.H.: The cytoprotective protein C pathway. Blood. **109**, 3161-3172 (2007)

52. Van Horssen, J., Wesseling, P., Van Den Heuvel, L.P., De Waal, R.M., Verbeek, M.M.: Heparan sulphate proteoglycans in Alzheimer's disease and amyloid-related disorders. The Lancet Neurology. **2**, 482-492 (2003)

53. Li, J.-P., Kusche-Gullberg, M.: Chapter Six-Heparan Sulfate: Biosynthesis, Structure, and Function. International review of cell and molecular biology. **325**, 215-273 (2016)

54. Muñoz, E.M., Linhardt, R.J.: Heparin-binding domains in vascular biology. Arteriosclerosis, thrombosis, and vascular biology. **24**, 1549-1557 (2004)

55. Olson, S.T., Richard, B., Izaguirre, G., Schedin-Weiss, S., Gettins, P.G.: Molecular mechanisms of antithrombin–heparin regulation of blood clotting proteinases. A paradigm for

understanding proteinase regulation by serpin family protein proteinase inhibitors. Biochimie. **92**, 1587-1596 (2010)

56. Gettins, P.G., Olson, S.T.: Exosite determinants of serpin specificity. Journal of Biological Chemistry. **284**, 20441-20445 (2009)

57. Belov, A.A., Mohammadi, M.: Molecular mechanisms of fibroblast growth factor signaling in physiology and pathology. Cold Spring Harbor perspectives in biology. **5**, a015958 (2013)

58. Schlessinger, J.: Common and distinct elements in cellular signaling via EGF and FGF receptors. Science. 306, 1506-1507 (2004)

59. Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B., Blundell, T.L.: Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. Nature. **407**, 1029-1034 (2000)

60. Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A.V., Yeh, B.K., Yayon, A., et al.: Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. Molecular cell. **6**, 743-750 (2000)

61. Catlow, K.R., Deakin, J.A., Wei, Z., Delehedde, M., Fernig, D.G., Gherardi, E., et al.: Interactions of hepatocyte growth factor/scatter factor with various glycosaminoglycans reveal an important interplay between the presence of iduronate and sulfate density. Journal of Biological Chemistry. **283**, 5235-5248 (2008)

62. Shute, J.: Glycosaminoglycan and chemokine/growth factor interactions. Springer, (2012)

63. Alon, R., Grabovsky, V., Feigelson, S.: Chemokine induction of integrin adhesiveness on rolling and arrested leukocytes local signaling events or global stepwise activation? Microcirculation. **10**, 297-311 (2003)

64. Salanga, C., Handel, T.: Chemokine oligomerization and interactions with receptors and glycosaminoglycans: the role of structural dynamics in function. Experimental cell research. **317**, 590-601 (2011)

65. Proudfoot, A.E., Johnson, Z., Bonvin, P., Handel, T.M.: Glycosaminoglycan Interactions with Chemokines Add Complexity to a Complex System. Pharmaceuticals. **10**, 70 (2017)

66. Kufareva, I., Salanga, C.L., Handel, T.M.: Chemokine and chemokine receptor structure and interactions: implications for therapeutic strategies. Immunology and cell biology. **93**, 372-383 (2015)

67. Handel, T., Johnson, Z., Crown, S., Lau, E., Sweeney, M., Proudfoot, A.: Regulation of protein function by glycosaminoglycans—as exemplified by chemokines. Annu. Rev. Biochem. **74**, 385-410 (2005)

68. Abedini, A., Tracz, S.M., Cho, J.-H., Raleigh, D.P.: Characterization of the heparin binding site in the N-terminus of human pro-islet amyloid polypeptide: implications for amyloid formation. Biochemistry. **45**, 9228-9237 (2006)

69. Jha, S., Patil, S.M., Gibson, J., Nelson, C.E., Alder, N.N., Alexandrescu, A.T.: Mechanism of amylin fibrillization enhancement by heparin. Journal of Biological Chemistry. **286**, 22894-22904 (2011)

70. Zhang, X., Li, J.-P.: Heparan sulfate proteoglycans in amyloidosis. Progress in molecular biology and translational science. **93**, 309-334 (2010)

71. Spear, P.G., Eisenberg, R.J., Cohen, G.H.: Three classes of cell surface receptors for alphaherpesvirus entry. Virology. **275**, 1-8 (2000)

72. Spear, P.G.: Herpes simplex virus: receptors and ligands for cell entry. Cellular microbiology. **6**, 401-410 (2004)

73. Spear, P.G., Longnecker, R.: Herpesvirus entry: an update. Journal of virology. **77**, 10179-10185 (2003)

74. Lai, J., Chien, J., Staub, J., Avula, R., Greene, E.L., Matthews, T.A., et al.: Loss of HSulf-1 up-regulates heparin-binding growth factor signaling in cancer. Journal of Biological Chemistry. **278**, 23107-23117 (2003)

75. Goldshmidt, O., Zcharia, E., Abramovitch, R., Metzger, S., Aingorn, H., Friedmann, Y., et al.: Cell surface expression and secretion of heparanase markedly promote tumor angiogenesis and metastasis. Proceedings of the National Academy of Sciences. **99**, 10031-10036 (2002)

76. Vlodavsky, I., Elkin, M., Abboud-Jarrous, G., Levi-Adam, F., Fuks, L., Shafat, I., et al.: Heparanase: one molecule with multiple functions in cancer progression. Connective tissue research. **49**, 207-210 (2008)

77. Wu, Z.L., Zhang, L., Beeler, D.L., Kuberan, B., Rosenberg, R.D.: A new strategy for defining critical functional groups on heparan sulfate. The FASEB journal. **16**, 539-545 (2002)

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

79. Lee, M.K., Lander, A.D.: Analysis of affinity and structural selectivity in the binding of proteins to glycosaminoglycans: development of a sensitive electrophoretic approach. Proceedings of the National Academy of Sciences. **88**, 2768-2772 (1991)

80. Yu, H., Munoz, E.M., Edens, R.E., Linhardt, R.J.: Kinetic studies on the interactions of heparin and complement proteins using surface plasmon resonance. Biochimica et Biophysica Acta (BBA)-General Subjects. **1726**, 168-176 (2005)

81. Corredor, M., Bonet, R., Moure, A., Domingo, C., Bujons, J., Alfonso, I., et al.: Cationic peptides and Peptidomimetics bind glycosaminoglycans as potential Sema3A pathway inhibitors. Biophysical journal. **110**, 1291-1303 (2016)

82. Hung, T.-J., Tomiya, N., Chang, T.-H., Cheng, W.-C., Kuo, P.-H., Ng, S.-K., et al.: Functional characterization of ECP-heparin interaction: A novel molecular model. PloS one. **8**, e82585 (2013)

83. Rudd, T.R., Nichols, R.J., Yates, E.A.: Selective Detection of Protein Secondary Structural Changes in Solution Protein− Polysaccharide Complexes Using Vibrational Circular Dichroism (VCD). Journal of the American Chemical Society. **130**, 2138-2139 (2008)

84. Viskov, C., Elli, S., Urso, E., Gaudesi, D., Mourier, P., Herman, F., et al.: Heparin Dodecasaccharide Containing Two Antithrombin-binding Pentasaccharides STRUCTURAL FEATURES AND BIOLOGICAL PROPERTIES. Journal of Biological Chemistry. **288**, 25895-25907 (2013)

85. Blaum, B.r.S., Deakin, J.A., Johansson, C.M., Herbert, A.P., Barlow, P.N., Lyon, M., et al.: Lysine and Arginine Side Chains in Glycosaminoglycan− Protein Complexes Investigated by NMR, Cross-Linking, and Mass Spectrometry: A Case Study of the Factor H− Heparin Interaction. Journal of the American Chemical Society. **132**, 6374-6381 (2010)

86. Pichert, A., Samsonov, S.A., Theisgen, S., Thomas, L., Baumann, L., Schiller, J., et al.: Characterization of the interaction of interleukin-8 with hyaluronan, chondroitin sulfate, dermatan sulfate and their sulfated derivatives by spectroscopy and molecular modeling. Glycobiology. **22**, 134-145 (2011)

87. Singh, A., Kett, W.C., Severin, I.C., Agyekum, I., Duan, J., Amster, I.J., et al.: The interaction of heparin tetrasaccharides with chemokine CCL5 is modulated by sulfation pattern and pH. Journal of Biological Chemistry. **290**, 15421-15436 (2015)

88. Drenth, J. Springer Science & Business Media, (2007)

89. Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., Skelton, N.J. Academic Press, (1995)

90. Smyth, M., Martin, J.: x Ray crystallography. Molecular Pathology. **53**, 8 (2000)

91. Meneghetti, M.C., Hughes, A.J., Rudd, T.R., Nader, H.B., Powell, A.K., Yates, E.A., et al.: Heparan sulfate and heparin interactions with proteins. Journal of The Royal Society Interface. **12**, 20150589 (2015)

92. Lisi, G.P., Loria, J.P.: Solution NMR spectroscopy for the study of enzyme allostery. Chemical reviews. **116**, 6323-6369 (2016)

93. Heck, A.J.: Native mass spectrometry: a bridge between interactomics and structural biology. Nature methods. **5**, 927-933 (2008)

94. Zhong, Y., Hyung, S.-J., Ruotolo, B.T.: Ion mobility–mass spectrometry for structural proteomics. Expert review of proteomics. **9**, 47-58 (2012)

95. Martin, S.R., Schilstra, M.J.: Circular dichroism and its application to the study of biomolecules. Methods in cell biology. **84**, 263-293 (2008)

96. Kelly, S.M., Price, N.C.: The use of circular dichroism in the investigation of protein structure and function. Current protein and peptide science. **1**, 349-384 (2000)

97. Gesslbauer, B., Theuer, M., Schweiger, D., Adage, T., Kungl, A.J.: New targets for glycosaminoglycans and glycosaminoglycans as novel targets. Expert review of proteomics. **10**, 77-95 (2013)

98. Rich, R.L., Myszka, D.G.: Advances in surface plasmon resonance biosensor analysis. Current opinion in biotechnology. **11**, 54-61 (2000)

99. Dutta, A.K., Rösgen, J., Rajarathnam, K.: Using Isothermal Titration Calorimetry to Determine Thermodynamic Parameters of Protein–Glycosaminoglycan Interactions. Glycosaminoglycans: Chemistry and Biology. 315-324 (2015)

100. Lanucara, F., Holman, S.W., Gray, C.J., Eyers, C.E.: The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nature chemistry. **6**, 281-294 (2014)

101. Konijnenberg, A., Butterer, A., Sobott, F.: Native ion mobility-mass spectrometry and related methods in structural biology. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics. **1834**, 1239-1256 (2013)

102. Hyung, S.J., Ruotolo, B.T.: Integrating mass spectrometry of intact protein complexes into structural proteomics. Proteomics. **12**, 1547-1564 (2012)

103. Wei, H., Mo, J., Tao, L., Russell, R.J., Tymiak, A.A., Chen, G., et al.: Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. Drug discovery today. **19**, 95-102 (2014)

104. Iacob, R.E., Engen, J.R.: Hydrogen exchange mass spectrometry: are we out of the quicksand? Journal of the American Society for Mass Spectrometry. **23**, 1003-1010 (2012)

105. Ben-Nissan, G., Sharon, M.: Capturing protein structural kinetics by mass spectrometry. Chemical Society Reviews. **40**, 3627-3637 (2011)

106. Wales, T.E., Engen, J.R.: Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass spectrometry reviews. **25**, 158-170 (2006)

107. Deng, Y., Zhang, Z., Smith, D.L.: Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics. Journal of the American Society for Mass Spectrometry. **10**, 675-684 (1999)

108. Konermann, L., Pan, J., Liu, Y.-H.: Hydrogen exchange mass spectrometry for studying protein structure and dynamics. Chemical Society Reviews. **40**, 1224-1234 (2011)

109. Marcoux, J., Robinson, C.V.: Twenty years of gas phase structural biology. Structure. **21**, 1541-1550 (2013)

110. Konermann, L., Stocks, B.B., Pan, Y., Tong, X.: Mass spectrometry combined with oxidative labeling for exploring protein structure and folding. Mass spectrometry reviews. **29**, 651-667 (2010)

111. Kiselar, J.G., Chance, M.R.: Future directions of structural mass spectrometry using hydroxyl radical footprinting. Journal of mass spectrometry. **45**, 1373-1382 (2010)

112. Gau, B.C., Sharp, J.S., Rempel, D.L., Gross, M.L.: Fast photochemical oxidation of protein footprints faster than protein unfolding. Analytical chemistry. **81**, 6563-6571 (2009)

113. Konermann, L., Tong, X., Pan, Y.: Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. Journal of Mass Spectrometry. **43**, 1021-1036 (2008)

114. Fischer, L., Chen, Z.A., Rappsilber, J.: Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. Journal of proteomics. **88**, 120-128 (2013)

115. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., et al.: Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. Molecular & Cellular Proteomics. **9**, 1634-1649 (2010)

116. Sinz, A.: Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. Journal of mass spectrometry. **38**, 1225-1237 (2003)

117. Kebarle, P., Verkerk, U.H.: Electrospray: from ions in solution to ions in the gas phase, what we know now. Mass spectrometry reviews. **28**, 898-917 (2009)

118. Kebarle, P., Tang, L.: From ions in solution to ions in the gas phase-the mechanism of electrospray mass spectrometry. Analytical chemistry. **65**, 972A-986A (1993)

119. Dole, M., Mack, L., Hines, R., Mobley, R., Ferguson, L., Alice, M.d.: Molecular beams of macroions. The Journal of Chemical Physics. **49**, 2240-2249 (1968)

120. Iribarne, J., Thomson, B.: On the evaporation of small ions from charged droplets. The Journal of Chemical Physics. **64**, 2287-2294 (1976)

121. Kebarle, P.: A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. Journal of Mass Spectrometry. **35**, 804-817 (2000)

122. Smith, R.D., Tang, K., Shen, Y.: Ultra-sensitive and quantitative characterization of proteomes. Molecular BioSystems. **2**, 221-230 (2006)

123. Daniel, J.M., Friess, S.D., Rajagopalan, S., Wendt, S., Zenobi, R.: Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. International Journal of Mass Spectrometry. **216**, 1-27 (2002)

124. Juraschek, R., Dülcks, T., Karas, M.: Nanoelectrospray—more than just a minimized-flow electrospray ionization source. Journal of the American Society for Mass Spectrometry. **10**, 300-308 (1999)

125. Wilm, M.S., Mann, M.: Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? International Journal of Mass Spectrometry and Ion Processes. **136**, 167-180 (1994)

126. Hilton, G.R., Benesch, J.L.: Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. Journal of The Royal Society Interface. rsif20110823 (2012)

127. Wilm, M., Mann, M.: Analytical properties of the nanoelectrospray ion source. Analytical chemistry. **68**, 1-8 (1996)

128. Lorenzen, K., Duijn, E.v.: Native mass spectrometry as a tool in structural biology. Current protocols in protein science. 17.12. 11-17.12. 17 (2010)

129. Raschke, T.M.: Water structure and interactions with protein surfaces. Current opinion in structural biology. **16**, 152-159 (2006)

130. Grabenauer, M., Wu, C., Soto, P., Shea, J.-E., Bowers, M.T.: Oligomers of the prion protein fragment 106− 126 are likely assembled from β-hairpins in solution, and methionine oxidation inhibits assembly without altering the peptide's monomeric conformation. Journal of the American Chemical Society. **132**, 532-539 (2009)

131. Rogniaux, H., Sanglier, S., Strupat, K., Azza, S.d., Roitel, O., Ball, V., et al.: Mass spectrometry as a novel approach to probe cooperativity in multimeric enzymatic systems. Analytical biochemistry. **291**, 48-61 (2001)

132. Zhou, M., Morgner, N., Barrera, N.P., Politis, A., Isaacson, S.C., Matak-Vinković, D., et al.: Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. Science. **334**, 380-385 (2011)

133. Breuker, K., McLafferty, F.W.: Stepwise evolution of protein native structure with electrospray into the gas phase, 10− 12 to 102 s. Proceedings of the National Academy of Sciences. **105**, 18145-18152 (2008)

134. van der Spoel, D., Marklund, E.G., Larsson, D.S., Caleman, C.: Proteins, lipids, and water in the gas phase. Macromolecular bioscience. **11**, 50-59 (2011)

135. Ruotolo, B.T., Robinson, C.V.: Aspects of native proteins are retained in vacuum. Current opinion in chemical biology. **10**, 402-408 (2006)

136. Han, L., Hyung, S.-J., Mayers, J.J., Ruotolo, B.T.: Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent. Journal of the American Chemical Society. **133**, 11358-11367 (2011)

137. López, A., Tarragó, T., Vilaseca, M., Giralt, E.: Applications and future of ion mobility mass spectrometry in structural biology. New Journal of Chemistry. **37**, 1283-1289 (2013)

138. Hyung, S.-J., Robinson, C.V., Ruotolo, B.T.: Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes. Chemistry & biology. **16**, 382-390 (2009)

139. Hopper, J.T., Oldham, N.J.: Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability. Journal of the American Society for Mass Spectrometry. **20**, 1851-1858 (2009)

140. Lössl, P., Snijder, J., Heck, A.J.: Boundaries of mass resolution in native mass spectrometry. Journal of the American Society for Mass Spectrometry. **25**, 906-917 (2014)

141. Loo, J.A.: Studying noncovalent protein complexes by electrospray ionization mass spectrometry. Mass spectrometry reviews. **16**, 1-23 (1997)

142. Uetrecht, C., Versluis, C., Watts, N.R., Wingfield, P.T., Steven, A.C., Heck, A.J.: Stability and shape of hepatitis B virus capsids in vacuo. Angewandte Chemie International Edition. **47**, 6247-6251 (2008)

143. Jenner, M., Ellis, J., Huang, W.C., Lloyd Raven, E., Roberts, G.C., Oldham, N.J.: Detection of a protein conformational equilibrium by electrospray ionisation-ion mobility-mass spectrometry. Angewandte Chemie International Edition. **50**, 8291-8294 (2011)

144. Shi, H., Pierson, N.A., Valentine, S.J., Clemmer, D.E.: Conformation types of ubiquitin [M+ 8H] 8+ ions from water: methanol solutions: evidence for the N and A states in aqueous solution. The Journal of Physical Chemistry B. **116**, 3344-3352 (2012)

145. Nabuchi, Y., Hirose, K., Takayama, M.: Ion mobility and collision-induced dissociation analysis of carbonic anhydrase 2. Analytical chemistry. **82**, 8890-8896 (2010)

146. Barnes, W., Martin, D., McDaniel, E.: Mass spectrographic identification of the ion observed in hydrogen mobility experiments. Physical Review Letters. **6**, 110 (1961)

147. Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinos, K., et al.: An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. International Journal of Mass Spectrometry. **261**, 1-12 (2007)

148. Kliman, M., May, J.C., McLean, J.A.: Lipid analysis and lipidomics by structurally selective ion mobility-mass spectrometry. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids. **1811**, 935-945 (2011)

149. Uetrecht, C., Rose, R.J., van Duijn, E., Lorenzen, K., Heck, A.J.: Ion mobility mass spectrometry of proteins and protein assemblies. Chemical Society Reviews. **39**, 1633-1655 (2010)

150. May, J.C., McLean, J.A.: Ion mobility-mass spectrometry: time-dispersive instrumentation. Analytical chemistry. **87**, 1422-1436 (2015)

151. Borsdorf, H., Mayer, T., Zarejousheghani, M., Eiceman, G.A.: Recent developments in ion mobility spectrometry. Applied Spectroscopy Reviews. **46**, 472-521 (2011)

152. Mason, E.A., McDaniel, E.W. Wiley Online Library, (1988)

153. Bohrer, B.C., Merenbloom, S.I., Koeniger, S.L., Hilderbrand, A.E., Clemmer, D.E.: Biomolecule analysis by ion mobility spectrometry. Annu. Rev. Anal. Chem. **1**, 293-327 (2008)

154. Giles, K., Williams, J.P., Campuzano, I.: Enhancements in travelling wave ion mobility resolution. Rapid Communications in Mass Spectrometry. **25**, 1559-1566 (2011)

155. Zhong, Y., Hyung, S.-J., Ruotolo, B.T.: Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions. Analyst. **136**, 3534-3541 (2011)

156. Cumeras, R., Figueras, E., Davis, C., Baumbach, J.I., Gracia, I.: Review on ion mobility spectrometry. Part 1: current instrumentation. Analyst. **140**, 1376-1390 (2015)

157. Giles, K., Pringle, S.D., Worthington, K.R., Little, D., Wildgoose, J.L., Bateman, R.H.: Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Communications in Mass Spectrometry. **18**, 2401-2414 (2004)

158. Pollard, M.J., Hilton, C.K., Li, H., Kaplan, K., Yost, R.A., Hill, H.H.: Ion mobility spectrometer—field asymmetric ion mobility spectrometer-mass spectrometry. International Journal for Ion Mobility Spectrometry. **14**, 15-22 (2011)

159. Shvartsburg, A.A.: Differential ion mobility spectrometry: nonlinear ion transport and fundamentals of FAIMS CRC Press, (2008)

160. Michelmann, K., Silveira, J.A., Ridgeway, M.E., Park, M.A.: Fundamentals of trapped ion mobility spectrometry. Journal of The American Society for Mass Spectrometry. **26**, 14-24 (2015)

161. Fernandez-Lima, F., Kaplan, D.A., Suetering, J., Park, M.A.: Gas-phase separation using a trapped ion mobility spectrometer. International Journal for Ion Mobility Spectrometry. **14**, 93-98 (2011)

162. Pu, Y., Ridgeway, M.E., Glaskin, R.S., Park, M.A., Costello, C.E., Lin, C.: Separation and identification of isomeric glycans by selected accumulation-trapped ion mobility spectrometry-electron activated dissociation tandem mass spectrometry. Analytical chemistry. **88**, 3440-3443 (2016)

163. Chernushevich, I.V., Loboda, A.V., Thomson, B.A.: An introduction to quadrupole–time-of-flight mass spectrometry. Journal of Mass Spectrometry. **36**, 849-865 (2001)

164. Sobott, F., Hernández, H., McCammon, M.G., Tito, M.A., Robinson, C.V.: A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. Analytical chemistry. **74**, 1402-1407 (2002)

165. van den Heuvel, R.H., van Duijn, E., Mazon, H., Synowsky, S.A., Lorenzen, K., Versluis, C., et al.: Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. Analytical chemistry. **78**, 7473-7483 (2006)

166. Sharon, M., Robinson, C.V.: The role of mass spectrometry in structure elucidation of dynamic protein complexes. Annu. Rev. Biochem. **76**, 167-193 (2007)

167. Boesl, U.: Time-of-flight mass spectrometry: Introduction to the basics. Mass spectrometry reviews. (2016)

168. Wollnik, H., Przewloka, M.: Time-of-flight mass spectrometers with multiply reflected ion trajectories. International journal of mass spectrometry and ion processes. **96**, 267-274 (1990)

169. El-Aneed, A., Cohen, A., Banoub, J.: Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. Applied Spectroscopy Reviews. **44**, 210-230 (2009)

170. Doroshenko, V.M., Cotter, R.J.: Ideal velocity focusing in a reflectron time-of-flight mass spectrometer. Journal of the American Society for Mass Spectrometry. **10**, 992-999 (1999)

171. Mehmood, S., Allison, T.M., Robinson, C.V.: Mass spectrometry of protein complexes: from origins to applications. Annual review of physical chemistry. **66**, 453-474 (2015)

172. Eschweiler, J.D., Rabuck-Gibbons, J.N., Tian, Y., Ruotolo, B.T.: CIUSuite: a quantitative analysis package for collision induced unfolding measurements of gas-phase protein ions. Analytical chemistry. **87**, 11516-11522 (2015)

173. Mackerell, A.D.: Empirical force fields for biological macromolecules: overview and issues. Journal of computational chemistry. **25**, 1584-1604 (2004)

174. Durrant, J.D., McCammon, J.A.: Molecular dynamics simulations and drug discovery. BMC biology. **9**, 71 (2011)

175. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S.a., Karplus, M.: CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. Journal of computational chemistry. **4**, 187-217 (1983)

176. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, T.E., DeBolt, S., et al.: AMBER, a package of computer programs for applying molecular mechanics, normal mode

analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Computer Physics Communications. **91**, 1-41 (1995)

177. Berendsen, H.J., van der Spoel, D., van Drunen, R.: GROMACS: a message-passing parallel molecular dynamics implementation. Computer Physics Communications. **91**, 43-56 (1995)

178. Ewig, C.S., Thacher, T.S., Hagler, A.T.: Derivation of class II force fields. 7. Nonbonded force field parameters for organic compounds. The Journal of Physical Chemistry B. **103**, 6998-7014 (1999)

179. Singh, A., Vanga, S.K., Orsat, V., Raghavan, V.: Application of Molecular Dynamic simulation to study food proteins: A Review. Critical Reviews in Food Science and Nutrition. 1-11 (2017)

180. Mura, C., McAnany, C.E.: An introduction to biomolecular simulations and docking. Molecular Simulation. **40**, 732-764 (2014)

181. Smatanová, I.K.: Crystallization of biological macromolecules. Materials Structure. **9**, (2002)

182. Lindorff-Larsen, K., Best, R.B., DePristo, M.A., Dobson, C.M., Vendruscolo, M.: Simultaneous determination of protein structure and dynamics. Nature. **433**, 128-132 (2005)

183. Gruebele, M., Dave, K., Sukenik, S.: Globular protein folding in vitro and in vivo. Annual review of biophysics. 45, 233-251 (2016)

184. Dror, R.O., Dirks, R.M., Grossman, J., Xu, H., Shaw, D.E.: Biomolecular simulation: a computational microscope for molecular biology. Annual review of biophysics. **41**, 429-452 (2012)

185. Gilquin, B., Guilbert, C., Perahia, D.: Unfolding of hen egg lysozyme by molecular dynamics simulations at 300K: insight into the role of the interdomain interface. Proteins: Structure, Function, and Bioinformatics. **41**, 58-74 (2000)

186. Cui, Q., Karplus, M.: Allostery and cooperativity revisited. Protein science. **17**, 1295-1307 (2008)

187. Wang, W., Donini, O., Reyes, C.M., Kollman, P.A.: Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. Annual review of biophysics and biomolecular structure. **30**, 211-243 (2001)

188. Roux, B.: Computational studies of the gramicidin channel. Accounts of chemical research. **35**, 366-375 (2002)

189. Bond, P.J., Sansom, M.S.: The simulation approach to bacterial outer membrane proteins. Molecular membrane biology. **21**, 151-161 (2004)

190. Elcock, A.H.: Molecular simulations of diffusion and association in multimacromolecular systems. Methods in enzymology. **383**, 166-198 (2004)

191. Sanbonmatsu, K.Y., Joseph, S., Tung, C.-S.: Simulating movement of tRNA into the ribosome during decoding. Proceedings of the National Academy of Sciences of the United States of America. **102**, 15854-15859 (2005)

192. Were, L., Hettiarachchy, N., Kalapathy, U.: Modified soy proteins with improved foaming and water hydration properties. Journal of Food Science. **62**, 821-824 (1997)

193. Koehl, P., Levitt, M.: De novo protein design. I. In search of stability and specificity. Journal of molecular biology. **293**, 1161-1181 (1999)

194. Wong, C.F., McCammon, J.A.: Protein simulation and drug design. Advances in protein chemistry. **66**, 87-121 (2003)

195. Marklund, E.G., Degiacomi, M.T., Robinson, C.V., Baldwin, A.J., Benesch, J.L.: Collision cross sections for structural proteomics. Structure. **23**, 791-799 (2015)

196. Gelb, A.S., Jarratt, R.E., Huang, Y., Dodds, E.D.: A study of calibrant selection in measurement of carbohydrate and peptide ion-neutral collision cross sections by traveling wave ion mobility spectrometry. Analytical chemistry. **86**, 11396-11402 (2014)

197. Ruotolo, B.T., Benesch, J.L., Sandercock, A.M., Hyung, S.-J., Robinson, C.V.: Ion mobility–mass spectrometry analysis of large protein complexes. Nature protocols. **3**, 1139-1152 (2008)

198. Smith, D.P., Knapman, T.W., Campuzano, I., Malham, R.W., Berryman, J.T., Radford, S.E., et al.: Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. European journal of mass spectrometry. **15**, 113-130 (2008)

199. Shvartsburg, A.A., Jarrold, M.F.: An exact hard-spheres scattering model for the mobilities of polyatomic ions. Chemical Physics Letters. **261**, 86-91 (1996)

200. Wyttenbach, T., von Helden, G., Batka, J.J., Carlat, D., Bowers, M.T.: Effect of the long-range potential on ion mobility measurements. Journal of the American Society for Mass Spectrometry. **8**, 275-282 (1997)

201. Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Mobilities of carbon cluster ions: critical importance of the molecular attractive potential. The Journal of chemical physics. **108**, 2416-2423 (1998)

202. Jurneczko, E., Barran, P.E.: How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst. **136**, 20-28 (2011)

203. Politis, A., Park, A.Y., Hyung, S.-J., Barsky, D., Ruotolo, B.T., Robinson, C.V.: Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes. PloS one. **5**, e12080 (2010)

204. Benesch, J.L., Ruotolo, B.T.: Mass spectrometry: come of age for structural and dynamical biology. Current opinion in structural biology. **21**, 641-649 (2011)

205. Bleiholder, C., Contreras, S., Bowers, M.T.: A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (IV). Application to polypeptides. International Journal of Mass Spectrometry. **354**, 275-280 (2013)

CHAPTER 2

EXPERIMENTAL

Preparation of GAG oligosaccharide, protein and protein-GAG complex

Arixtra was purchased from the hospital formulary and desalted on a BioGel P2 column BioRad (Hercules, CA, USA) before use. HS tetrasaccharides were synthesized as previously described based by fluorous supported modular synthesis [1]. HS samples of larger sizes (dp6 to dp12) and modified Arixtra were chemoenzymatically synthesized as previously described [2].

Most proteins and protein calibrant standards were commercially available except for Robo1 and RPTP-LAR. Protein samples were desalted and buffer exchanged into 25 mM ammonium acetate (pH 6.8) with Amicon Ultra-0.5 mL centrifugal filters (molecular weight cutoff, 3000 Da, Millipore, Spain). Protein-GAG complexes were prepared by incubating them at a certain molar ratio at room temperature for multiple reaction times.

IMMS analysis

IMMS experiments were performed on Waters Synapt G2 HDMS quadrupole-TWIM-TOF hybrid mass spectrometer (Waters Corp., Manchester, UK) in positive ionization mode. Protein samples were injected into a nanoESI source through a fused-silica emitter (PicoTip New Objective, Woburn, MA) with a flow rate varying from 0.2-0.5 µl/min. Source parameters were carefully tuned for each sample to prevent the protein and protein complex from unfolding or losing integrity due to extensive activation while keeping substantial ion transmission: For instance, the optimized capillary voltage and cone voltages were reduced to the lowest practical values that are still capable of maintaining a stable electrospray while producing good signal of the sample; The ion source temperature was kept at the lowest possible value to reduce the effect of extra activation.

Proteins ions were collisionally cooled in a helium cell with a helium flow of 180 mL/min, before being injected to the trap ion guide. The helium cell used to adjust to the pressure difference between trap ion guide and IM ion guide, so that higher fields for ion injection can be avoided to decrease the undesired ion scattering and fragmentation. Nitrogen was used for ion mobility separation, and a nitrogen flow of 90 mL/min was used to maintain an ion mobility cell pressure of approximately 3 mbar.

The TWIM DC traveling wave height and wave velocity are the key parameters determining the ion mobility separation and were optimized for each sample. Application of lower wave heights avoid structural distortion whereas wave velocity was adjusted accordingly. The typical range for wave height is 15-20V. The typical range for wave velocity is 200-300 m/s.

Except for CIU experiments, all DC bias values were typically minimized to avoid extra ion activation. For the CIU experiment, precursor ions were isolated in quadrupole before being sent to trap ion guide where they undergo collisions with argon gas at a flow rate of 4 mL/min. Incremental changed trapping CE from 0V to 80V were applied to protein and protein complexes, followed by ion mobility separation.

Data analysis was performed using MassLynx 4.1 and Driftscope (Waters Corp., Manchester, UK). The drift time distribution was fitted with a minimum numbers of Gaussian distributions according to its shape using the peak analyzer function in OriginPro 8.5.0 software (OriginLab Corporation, MA). The CIU data was visualized and analyzed using CIUSuite [3] to generate CIU fingerprints of protein ions, which records the relative ion intensity (normalized and smoothed using Savitsky-Golay filter) as a function of collision energy and drift times. The CIUsuite detect function was applied to identify and extract the most intense CIU features of the fingerprint based on a first derivative analysis, providing centroid drift times, transitional voltage, and stability range of each detected feature.

Converting drift times to experimental CCSs

The measured drift times were converted to CCSs based on calibration due to the presence of non-linear electric field for IM separation. Native calibrants of similar shape and nature as the analytes were included in the calibrant set to improve the accuracy of calibration. A selected set of native and denatured protein calibrants, with a mass range from 12 kDa to 102 kDa and a CCS range from 2303 $Å^2$ to 5550 $Å^2$ were employed, whose mass range and CCS range were broad enough to bracket the masses and drift times of protein or protein complexes of interest to avoid extrapolation of the calibration curve. The known CCS values of these protein calibrants were measured using a drift tube ion mobility instrument.

Experimental CCSs of the unknown protein and protein complexes were calibrated based on a published protocol [4]. Briefly, since the measured drift time is dependent on Ω, μ and z, the known CCS of calibrant was firstly corrected for its charge state and reduced mass with respect to buffer gas to generate normalized CCS Ω'. The drift time was corrected for mass-independent flight time spent in the transfer ion guide and IM ion guide based on the corresponding length of the cell and traveling wave velocity, as well as mass-dependent flight time spent in the TOF mass analyzer, which is dependent on the m/z of each ion of interest and the enhanced duty cycle for each individual instrument. The natural logarithm of corrected CCSs were plotted against the natural logarithm of corrected drift times so that a mathematical formula ($\mathbf{Ln\,\Omega' = A \times Ln\,dt' + B}$)

was derived. The calibration coefficient A was extracted to calculate the effective drift times dt": $\mathbf{dt}''=(\mathbf{dt}')^{A}\frac{z}{\sqrt{\mu}}$. A calibration curve was generated by plotting the literature CCSs as a function of dt". The correlation coefficient R^2 of the calibration curve should be higher than 0.98. The experimental CCS of the analyte ion can be derived from this calibration curve based on experimentally measured drift time.

Preparing input files for MD simulation

Crystal structures of protein and protein complexes were directly obtained from Protein Data Bank. Before subjected to MD simulation, H atoms were added using tLeap program. Missing or unresolved loops or residues were built using the Chimera interface to Modeller [5]. In the case when multiple crystal structures are available, a short implicit solvent minimization followed by a the single-trajectory Molecular Mechanics–Generalized Born Solvent Accessible Surface Area (MM-GBSA) based energy calculation were used for calculating the interaction energy of all the models and selecting the one with the highest binding energy thus highest binding affinity for further studies.

If no protein-GAG complex structure is available, a molecular docking experiment can be performed to predict possible binding modes. A MD simulation followed by free energy calculations can be performed to further refine and examine the binding modes, thus generating candidate starting structures.

MD Simulation

Topology and coordinate files for MD simulation were generated using the tLeap program, employing the parameter from ff99SB [6] force field for protein and GLYCAM06 (version j) [7, 8] force field for GAGs. The net charge of the system was neutralized by addition of an appropriate number of counter ions (Na^+ or Cl^-), followed by the solvation of the system with TIP3P [9] water molecules in a cubic box 12 Å from any atom of the solute.

MD simulations were performed with the GPU implementation of pmemd, pmemd.cuda_SPDP [8] module in Amber14. Both energy minimization of solvent and full system were performed in an NVT ensemble (1000 steps of steepest descent, 24,000 steps of conjugate gradient). The systems were then heated from 5 to 300 K over 60 ps in an NVT ensemble by employing a Berendsen-type thermostat, with a weak cartesian restraints (10 kcal/mol $Å^2$) placed on the Cα atoms of the protein backbone. After the heating step, the restraints were removed from the solute atoms, and the equilibration of 50 ns and production of 150 ns were performed at an NPT ensemble with a constant pressure of 1 atm and a pressure relaxation time of 1 ps. Post-processing of MD simulations was performed using CPPTRAJ [10] module of Amber. The graphical representations were generated using VMD [11].

Calculation of gas-phase binding free energies

Models of the gas-phase protein-GAG complexes were generated from MD simulations by removing water molecules followed by subjecting complexes to an *in-vacuo* minimization (IMIN = 5) using the SANDER module of AMBER 14 package. This was performed for 5000 frames selected at even steps from the last 100 ns of solvated simulation. The resultant gas-phase trajectory was used to calculate the binding free energies with the single-trajectory MM-GBSA method using the MMPBSA.py script from the AMBER 12 package [12].

Calculation of theoretical CCSs

The theoretical CCS values of protein or protein complexes were calculated using the PA method, whose input can be as simple as a single static structure and the TM method, whose input is a MD simulation trajectory. Both two methods are provided by open source software MOBCAL [13]. Both the PSA and scaled PA methods were also used.

References

1. Zong, C., Venot, A., Dhamale, O., Boons, G.-J.: Fluorous supported modular synthesis of heparan sulfate oligosaccharides. Organic letters. **15**, 342-345 (2013)

2. Xu, Y., Cai, C., Chandarajoti, K., Hsieh, P.-H., Li, L., Pham, T.Q., et al.: Homogeneous low-molecular-weight heparins with reversible anticoagulant activity. Nature chemical biology. **10**, 248-250 (2014)

3. Eschweiler, J.D., Rabuck-Gibbons, J.N., Tian, Y., Ruotolo, B.T.: CIUSuite: a quantitative analysis package for collision induced unfolding measurements of gas-phase protein ions. Analytical chemistry. **87**, 11516-11522 (2015)

4. Ruotolo, B.T., Benesch, J.L., Sandercock, A.M., Hyung, S.-J., Robinson, C.V.: Ion mobility–mass spectrometry analysis of large protein complexes. Nature protocols. **3**, 1139-1152 (2008)

5. Šali, A., Blundell, T.L.: Comparative protein modelling by satisfaction of spatial restraints. Journal of molecular biology. **234**, 779-815 (1993)

6. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C.: Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Structure, Function, and Bioinformatics. **65**, 712-725 (2006)

7. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L., et al.: GLYCAM06: a generalizable biomolecular force field. Carbohydrates. Journal of computational chemistry. **29**, 622-655 (2008)

8. Götz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., Walker, R.C.: Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. Journal of chemical theory and computation. **8**, 1542-1555 (2012)

9. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L.: Comparison of simple potential functions for simulating liquid water. The Journal of chemical physics. **79**, 926-935 (1983)

10. Roe, D.R., Cheatham III, T.E.: PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. Journal of chemical theory and computation. **9**, 3084-3095 (2013)

11. Humphrey, W., Dalke, A., Schulten, K.: VMD: visual molecular dynamics. Journal of molecular graphics. **14**, 33-38 (1996)

12. Miller III, B.R., McGee Jr, T.D., Swails, J.M., Homeyer, N., Gohlke, H., Roitberg, A.E.: MMPBSA. py: an efficient program for end-state free energy calculations. Journal of chemical theory and computation. **8**, 3314-3321 (2012)

13. Mesleh, M., Hunter, J., Shvartsburg, A., Schatz, G., Jarrold, M.: Structural information from ion mobility measurements: effects of the long-range potential. The Journal of Physical Chemistry. **100**, 16082-16086 (1996)

CHAPTER 3

INVESTIGATING CHANGES IN THE GAS-PHASE CONFORMATION OF ANTITHROMBIN III UPON BINDING OF ARIXTRA USING TRAVELING WAVE ION MOBILITY SPECTROMETRY (TWIMS) [1]

---

[1] Zhao Y, Singh A, Li L, Linhardt RJ, Xu Y, Liu J, Woods RJ, Amster IJ. 2015. *Analyst*. 140(20): 6980-9. Reprinted here with permission of publisher.

Abstract

We validate the utility of ion mobility to measure protein conformational changes induced by the binding of glycosaminoglycan ligands, using the well characterized system of Antithrombin III (ATIII) and Arixtra, a pharmaceutical agent with heparin (Hp) activity. Heparin has been used as therapeutic anticoagulant drug for several decades through its interaction with ATIII, a serine protease inhibitor that plays a central role in the blood coagulation cascade. This interaction induces conformational changes within ATIII that dramatically enhance the ATIII-mediated inhibition rate. Arixtra is a smallest synthetic Hp containing the specific pentasaccharide sequence required to bind with ATIII. Here we report the first travelling wave ion mobility mass spectrometry (TWIMS) investigation of the conformational changes in ATIII induced by its interaction with Arixtra. Native electrospray ionization mass spectrometry allowed the gentle transfer of the native topology of ATIII and ATIII-Arixtra complex. IM measurements of ATIII and ATIII-Arixtra complex showed a single structure, with well-defined collisional cross section (CCS) values. An average 3.6% increase in CCS of ATIII occurred as a result of its interaction with Arixtra, which agrees closely with theoretical estimation of the change in CCS, using protein crystal structures. Comparing the binding behavior of ATIII under both denaturing and non-denaturing conditions revealed the significance of a folded tertiary structure of ATIII for its biological activity. A Hp oligosaccaride whose structure is similar to Arixtra but missing the 3-O sulfate group on the central glucosamine residue showed a dramatic decrease in binding affinity towards ATIII, but no change in the mobility behavior of the complex, consistent with prior studies that suggested that 3-O sulfation affects the equilibrium constant for binding to ATIII, but not the mode of interaction. In contrast, nonspecific binding by a Hp tetrasaccharide showed more complex mobility behavior, suggesting more promiscuous interactions with ATIII.

The effect of collisional activation of ATIII and ATIII-Arixtra complex were also assessed, revealing that the binding of Arixtra provided ATIII with additional stability against unfolding. Overall, our results validate the capability of TWIMS to retain the significant features of the solution structure of a protein-carbohydrate complex so that it can be used to study protein conformational changes induced by the binding of glycosaminoglycan liga

Introduction

Heparin (Hp) and heparan sulfate (HS) are highly sulfated, linear polysaccharides, consisting of disaccharide repeat units of 1-4 linked hexuronic acid and N-acetyl-glucosamine, and are members of a class of carbohydrates known as glycosaminoglycans (GAGs) [1].The sequence of Hp and HS features three types of domains: highly sulfated (NS) domains, less or non-sulfated N-acetylated (NA) domains and partially sulfated domains (NA/NS). Some of these domains are selectively recognized by over hundreds of secreted and membrane associated human proteins [2-5]. By regulating the location, stability and activity of these interacting proteins, Hp and HS play crucial role in many important physiological and pathological processes [6].

The Hp/HS induced, allosteric activation of Antithrombin (ATIII) is the most studied and best understood example of a specific GAG-protein interaction. The anticoagulant property of Hp was discovered in 1916, and it has been used for prophylaxis and treatment of venous thrombosis, thrombophlebitis and embolism since 1940s [2]. Antithrombin III (ATIII), a 58.2 kDa N-glycosylated mono-chain protein in the serpin (serine protease inhibitors) family of proteins, serves as a principal regulator of blood coagulation serine and cysteine proteinases including factor IXa, factor Xa, and thrombin. The inhibitory activity of ATIII is repressed until it is activated by Hp/HS cofactor, either from therapeutic Hp or endothelial cells surface HS proteoglycans at the site of a vascular injury [7]. The native tertiary structure of ATIII is centered at a five-stranded β-sheet A, in which the N-terminal hinge of a reactive center loop (RCL) is initially buried with an orientation unfavorable to react with target proteinases [8]. Recent studies also found that the constraint of the RCL also intensifies the repulsive exosite interactions which counteracts the favorable interaction between proteinases and exosite determinants on strand 3 of sheet C surrounding the RCL [9].

The activation of ATIII occurs via its interaction with Hp, and more specifically, with a unique pentasaccharide sequence with a rare 3-O sulfate [2]. Upon binding of Hp/HS, a local conformational change at the Hp-binding site (the N-terminal region, the N-terminal end of helix A and all of helix D) is triggered. This further induces conformational changes on the proteinase binding site on ATIII [10, 11]. As a result of the allosteric activating structure arrangement, the RCL is released and the equilibrium between favorable and repulsive exosite interactions shifts to the favorable side [12]. Therefore, the formation and stabilization of ATIII-proteinase complex are promoted. The inhibitory rate of ATIII can be accelerated up to 150~500-fold by the binding of the specific Hp pentasaccharide domain against factors Xa, IXa and VIIa, and this rate can be accelerated up to 2000~200000-fold by full-length Hp [13].

As shown in the case of the interaction between ATIII and Hp/HS, protein-ligand or protein-protein interaction is often characterized by three dimensional conformational change of protein in response to a specific biological function. Conventional solid-phase method (X-ray crystallography) and solution-phase method (NMR spectroscopy) allow elucidation of structural details of protein and protein complex, and the structures of ATIII and ATIII-Arixtra complex have been characterized in this fashion. However, there are issues including difficult sample preparation, lengthy data analysis and low sample compatibility, that reduce the widespread applicability of these methods to a wide variety of protein-GAG complexes [14].

Ion mobility mass spectrometry (IMMS) is a rapid, sensitive and high-throughput gas-phase technique combing the advantages of both ESI mass spectrometry and ion mobility separation, and has gained much attention and recognition in the field of structural and dynamical biology [15]. IMMS separates gas-phase ions according to their mobility, an intrinsic property determined by size, shape and charge state of ions [16]. Travelling wave ion mobility spectrometry (TWIMS) is a commonly-used approach for IMMS, and is a commercially available product. In TWIMS, ion mobility separation takes place in an ion guide filled with a neutral gas. A radially confining potential barrier stops ion diffusion away from the path of the ion beam, while a continuous series of low voltage pulses called travelling waves push ions through the device [17]. Protein ions with larger collisional cross section (CCS) have lower mobility, undergo more frequent collisions and fall behind the traveling waves, with the result that they are transmitted more slowly than ions with higher mobility. The drift time of a protein ion can be related to its CCS and conformation though calibration with standards. The biological relevance of the measured CCS can be validated by comparison with theoretical CCSs generated from NMR or X-ray crystal structures [18].

Due to the advantage of TWIMS to experimentally estimate the CCS of a gas phase ion in a rapid and sensitive manner, it has been applied by a number of researchers for studies of protein interactions: Leary and coworker have applied TWIMS to detect differences in the conformations of two classes of chemokines as well as the effect of degree of sulfation of Hp-like oligomers on Hp-chemokine interactions [19]; Ruotolo and coworkers have applied TWIMS to study the gas-phase conformational stability of wild-type tetrameric transthyretin and its disease-associated variants with and without ligand binding [20]; Robinson and coworkers have applied TWIMS to measure the quaternary structure of the *trp* RNA binding attenuation protein (TRAP) complex and how the addition of tryptophan or RNA enhances the stability of its ring topology [21]; Heck and coworkers have applied TWIMS to measure the CCSs of oligomeric viral capsid assembly intermediates of norovirus and hepatitis B virus [22]; Jarrold and coworkers have applied TWIMS to study how the binding of metal ions influences the conformation transition of human metallothionein-2A (MT), an intrinsically disordered protein [23]; Bowers and coworkers have used TWIMS to investigate the CCSs and self-assembling pathways of amyloid-β protein assemblies [24]. Though there are concerns regarding the extent to which the structure of a gas-phase ion matches the native structure of a protein in solution, numerous studies have established a good correspondence between the two, particularly for the short time periods of an IMMS measurement, and under carefully controlled experimental conditions [25].

The ATIII-Hp interaction featured by its well-studied conformational change and high binding specificity represents an excellent model to test the applicability of the TWIMS approach to examining conformational changes associated with GAG-protein interaction. In this study, we applied the method of TWIMS to investigate the nature and extent of conformational change within ATIII induced by the binding of Arixtra, a synthetic analogue of the Hp pentasaccharide sequence known to bind this protein with high specificity [26]. TWIMS experiments were performed in order to solve these questions: Can the solution structure of ATIII survive the gas-phase environment in TWIMS and still maintain its activity to bind with Hp? Is TWIMS capable of detecting the small difference in CCS for ATIII that is induced by the binding of Arixtra? Do gas-phase ATIII ions exhibit selective binding with Arixtra compared to other structural-similar compounds? Do the complexes of ATIII and Hp oligosaccharides lacking some of the features of the specific binding motif show similar or different behaviors in their ion mobility? Collectively, the answers to the following questions provide insight into the utility of an IM approach, and more specifically, a TWIMS approach, for studying conformational changes in proteins as a result of their GAG interactions

Experimental

Reagents

All chemicals and solvents (ammonium acetate, methanol, water and fomic acid) were of HPLC grade and purchased from Sigma-Aldrich. AT III was purchased from Aniara/Hyphen Biomed as lyophilized powder (West Chester, OH). Stock solution of ATIII was made by dissolving the lyophilized protein into HPLC-grade water and then stored at -80 °C. Arixtra was purchased from the hospital formulary and desalted on a BioGel P2 column BioRad (Hercules, CA, USA) before use. Modified Arixtra was chemoenzymatic synthesized as previously described [27]. The Hp tetrasaccharide was produced from naturally occurring source as previously described [28]. Protein calibrants (myoglobin from equine heart, cytochrome c from equine heart, avidin from egg white, concanavalin A from *Canavalia ensiformis* and bovine serum albumin) were purchased from Sigma-Aldrich as lyophilized powder.

Sample preparation

For MS analyses under denaturing conditions, ATIII was diluted in a water/methanol/formic acid solution (49.5:49.5:1, v/v/v) to a final concentration of 3 μM. For MS analyses under non-denaturing conditions, ATIII was diluted in 20 mM ammonium acetate buffer, pH 6.8, to a final concentration of 10 μM. ATIII-Hp complex was obtained by incubating ATIII with Arixtra or other Hp oligosaccharides at a molar ratio of 1:1 at room temperature. Protein calibrants were diluted in either denaturing solution or non-denaturing solution to a final concentration of 10 μM.

IMMS measurement

NanoESI-IMMS experiments were performed using a quadrupole-TWIM-TOF hybrid mass spectrometer (Synapt G2 HDMS, Waters Corp., Manchester, UK) in positive ionization mode. Protein samples were injected into the nanoESI source through a fused-silica emitter (PicoTip New Objective, Woburn, MA) with a flow rate varying from 0.2-0.5 ul/min. Experimental parameters were carefully tuned to prevent the protein and protein complex from unfolding or losing integrity due to extensive activation while keeping substantial ion transmission. The applied experimental parameters were: capillary voltage, 1.5 kV; sampling cone voltage, 30 V; extraction cone voltage, 5 V; source temperature, 90 °C; flow rate of nitrogen in the IM ion guide, 50 ml/min; flow rate of helium in the helium cell, 180 mL/min; Trap collision energy, 0 V; transfer collision energy, 0 V. Different sets of wave height and corresponding wave velocity were used to optimize the mobility separation. The drift times of the calibrants and ATIII samples were measured when identical experimental conditions were stringently applied. Data analysis was performed with MassLynx 4.1(Waters Corp., Manchester, UK).

CCS calibration

The biggest challenge of IMMS using TWIMS is proper calibration. Mobility is a unique property for a given ion, and depends only on the features of a protein ion and neutral gas (mass, charge, size and shape). The measured drift time of an ion is under the influence of many other factors, including experimental conditions (magnitude and velocity of travelling wave, pressure of neutral gas, temperature of the source and ion guide). Since the electric field in TWIMS is not uniform due to the existence of travelling wave voltages whose electric potential is changing over time and position within the IM ion guide, drift time acquired using TWIMS cannot be converted directly into a CCS. Instead, the CCS of an analyte needs to be determined based on an empirical relationship between the drift times of protein calibrants and their known CCSs obtained previously by conventional drift time IMS [29].

Experimental CCSs were calibrated, as described in several published protocols [18, 29]. Briefly, the CCSs of calibrants were corrected for their charge state and reduced mass with respect to the buffer gas. The drift times were corrected for mass-dependent flight time spent in the transfer ion guide and TOF mass analyzer and mass-independent flight time spent in the transfer ion guide. The natural logarithm of corrected CCSs were plotted against the natural logarithm of corrected drift times and a mathematical formula ($\mathrm{Ln}\ \Omega' = \mathrm{A} \times \mathrm{Ln}\ \mathrm{dt}' + \mathrm{B}$) was derived. The calibration coefficient A was extracted to calculate the effective drift times dt'': $\mathrm{dt}''=(\mathrm{dt}')^{A}\frac{z}{\sqrt{\mu}}$ [18]. A calibration curve was generated by plotting the literature CCSs as a function of dt''. The correlation coefficient R^2 of the calibration curve should be higher than 0.98. The experimental CCS of the analyte ion can be derived from this calibration curve based on the measured drift time.

Another intrinsic problem for this experiment is that the calibrant library uses CCSs which were measured in helium while the TWIMS measurement uses an IM ion guide filled with nitrogen. However, the absolute error of CCS calibration derived from measuring CCS in different gases can be minimized when appropriate separation parameters are applied and when protein calibrants used to construct calibration curves are carefully selected [30, 31].

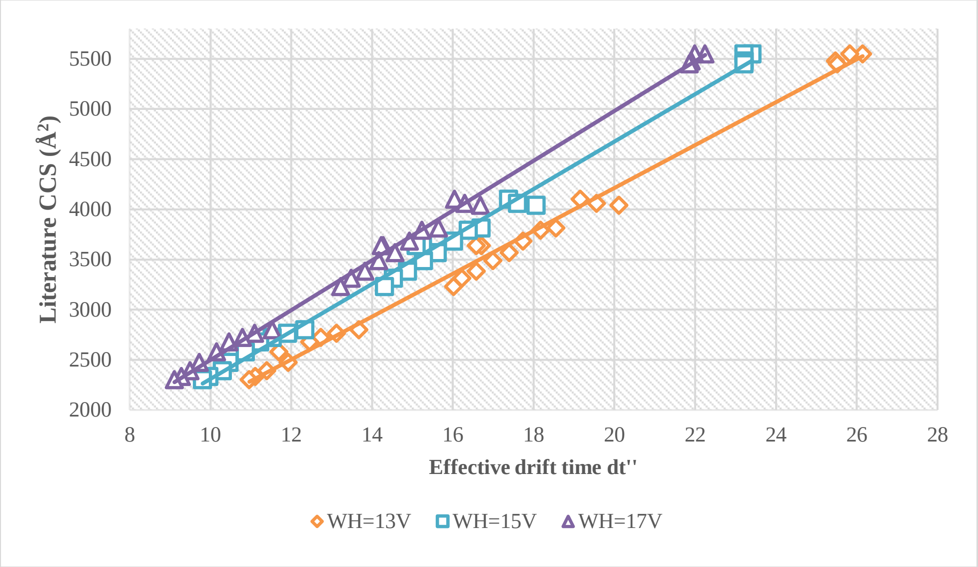


Figure 3.1 The calibration curves were constructed at three wave heights, combining data from cytochrome c, myoglobin, avidin, BSA and concanavalin A, displayed as literature CCS vs effective drift time for each charge state.

We have noted that CCS estimation was substantially improved by using native protein calibrants in the calibration, compared with using only denatured myoglobin (data not shown). Previous studies showed that CCS calibrated using only denatured calibrants were strongly influenced by separation conditions in TWIMS experiment. Therefore, including native calibrants of similar shape and nature as the analytes in the calibrant set greatly increased the accuracy of calibrated CCS of native ATIII and ATIII-Hp complex [32].

A selected set of native and denatured protein calibrants, with a mass range from 12kDa to 102kDa and a CCS range from 2303 Å^2 to 5550 Å^2 were employed. The charge states and literature CCSs of these protein calibrants used to construct the calibration curve are selected from the Collision Cross Section Database, Bush Lab [33]. The mass range and CCS range chosen were broad enough to bracket the masses and drift times of ATIII and ATIII-Hp complex ions (Figure 3.1), so no extrapolation of the calibration curve is necessary.

The separation parameters controlling the ion mobility separation were optimized, so that they fit both the analytes and calibrants. Multiple sets of separation parameters were applied to exclude the effect of electric field on the drift time measurement. The calibration curves constructed at wave height of 13, 15 and 17V are shown in Figure 3.1. R^2 values of 0.9919, 0.9944 and 0.9966 were observed for each trend line.

Theoretical calculation of CCSs

The missing residues in the PDB files for the ATIII-Arixtra complex (PDBid:1E03) and ATIII (PDBid: 1E05) [34] were built using MODELLER [35]. In the complex, the protein was missing 5 residues from the N-terminus and 8 residues from a disordered portion of the N-terminus. The free protein lacked 2 residues from the N terminus and 12 residues from the disordered region that was missing in the protein in complex with Arixtra. Both forms also lacked one residue at the C-terminus.

Topology and coordinate files for the ATIII-Arixtra complex and free ATIII were generated with the tLeap program, employing the Protein ff99SB [36] and GLYCAM06 (version

j) [37] parameters for the protein and GAGs, respectively. The net charge on each system was neutralized with the addition of appropriate number of Na^+ ion. The systems were solvated with TIP3P water[38] in a octahedron box extending to at least 12 Å from any atom of the solute.

All MD simulations were performed with the GPU implementation of pmemd, pmemd.cuda_SPDP [39] in Amber14 [40]. Energy minimization of the solvent was performed in an NVT ensemble (1000 steps of steepest descent, 24000 steps of conjugate gradient), followed by a full system energy minimization (1000 steps of steepest descent, 24000 steps of conjugate gradient). The systems were heated from 5 K to 300 K over 60 ps in an NVT ensemble, with a weak positional restraint (10 kcal/mol-$Å^2$) on the atoms in the solute. A Berendsen-type thermostat [41] with a time coupling constant of 1 ps, was utilized for temperature regulation. Equilibration and production was performed at constant pressure (NPT ensemble; 1 atm), with a pressure relaxation time of 1 ps. After the heating step the restraints were removed from the solute atoms, and the entire system was allowed to equilibrate at 300K for 1 ns. All covalent bonds involving hydrogen atoms were constrained using the SHAKE [42] algorithm, allowing a simulation time step of 2 fs. Scaling factors for 1–4 interactions were set to the recommended values of 1.0 and 1.2 for the GAG [37] and protein [36], respectively, and a non-bonded interaction cutoff of 8.0 Å was employed. Long-range electrostatics were computed with the particle mesh Ewald (PME) method. Data were collected for 20 ns for both the systems. Post processing of the MD simulations was performed using ptraj [43] module of Amber and graphical representations of the results were generated with VMD [44].

Theoretical CCSs of ATIII and ATIII-Arixtra complex were calculated on 30 frames from the last 5 ns of the simulation using MOBCAL[45]. Both the projection approximation (PA) [45], and trajectory method (TM) [46] were employed for the calculations.

Results and discussion

Native Mass Spectrometry of unbound ATIII and Arixtra-bound ATIII

ATIII was first analyzed under non-denaturing condition without the addition of any Hp (Figure 3.2 a). The protein was represented by a single charge state envelope, indicating that only protein monomers existed under this experimental condition. Five charge states were observed from +17 to +13 over a range of m/z 3300~4700. Since the charge states detected by using native MS depend on the three dimensional conformation of the protein ion, a narrow distribution of lower charge states indicates a folded and compact conformation with fewer basic sites exposed for protonation [47]. The three charge states, +16 to +14, were selected for further investigation since they were the most dominant charge states with relatively higher S/N. An average molecular weight of 57876±6 Da was obtained, in agreement with the literature molecular weight of ATIII (57875 Da) [48].

ATIII was next incubated with a small molar excess of Arixtra. Two incubating times were examined to monitor the progress of the binding reaction: 1h (Figure 3.2 b) and 12h (Figure 3.2 c). The binding of Arixtra by ATIII did not change the charge state distribution observed for native protein. An additional peak was observed for each charge state, corresponding to the formation of Arixtra-ATIII complex. An average mass increase of 1512 Da was measured for the new peaks, in good correspondence to the molecular weight of a single Arixtra molecule (1505 Da). To confirm the binding stoichiometry, titration experiments with different molar ratios were examined, but increasing the amount of Arixtra did not result in the formation of any new peaks. After 12h of incubation, only the peaks corresponding to Arixtra bound to ATIII in a 1:1 ratio were present (Figure 3.2 c). Since longer incubation time allowed the binding reaction to go to completion, this result indicates that the ATIII only presents one binding site for Arixtra, in accordance with the 1:1 binding stoichiometry between ATIII and Arixtra reported previously [49]. All of these observations suggest that the solution structure of ATIII survives the ionization process and the environment of the TWIMS instrument.

Interestingly, the intensity ratio between ATIII with and without Arixtra varied with charge state. The ratio for +17, +16, +15 and +14 were 0.69, 1.02, 3.07 and 5.75, respectively. It is evident that the intensity ratio increases with the decrease of charge state, suggesting weakening in the binding affinity with the increase of charge state. One possible explanation is

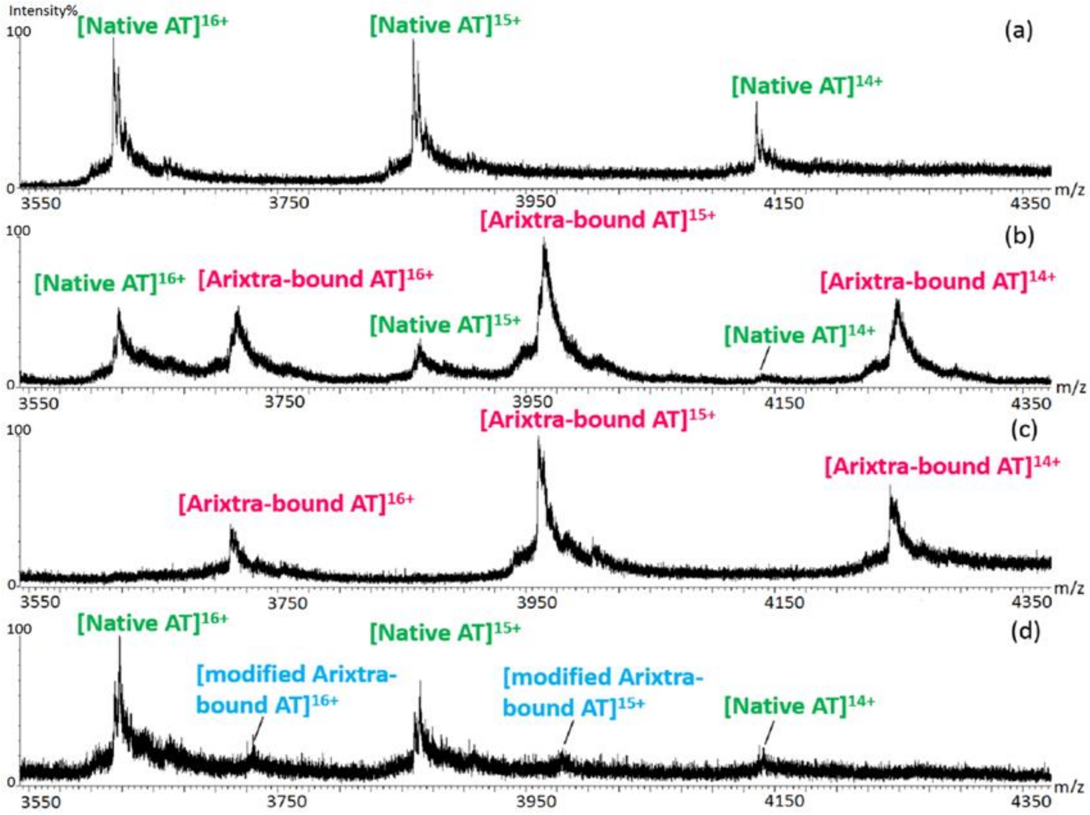


Figure 3.2 Electrospray ionization mass spectra, obtained under nondenaturing conditions, of (a) ATIII; (b) ATIII incubated with Arixtra for 1h; (c) ATIII incubated with Arixtra for 12h; (b) ATIII incubated with the Arixtra-like hexasaccharide, minus 3-O sulfation, for 12h.

that the lower charge states are more representative of the native structure of ATIII due to a lowering of charge-charge repulsion.

TWIMS of ATIII and its complex with Arixtra

To investigate whether the protein conformational change caused by the interaction between ATIII and Arixtra can be detected by TWIMS, ion mobility measurements were performed on the gas-phase ATIII and Arixtra-bound ATIII ions. The measured drift times of each charge state of ATIII and Arixtra-bound ATIII are shown in Figure 3.3, at a wave height of 17 V. One narrow drift time distribution was observed for each charge state of ATIII, indicating the presence of a single compact and folded conformation for the ATIII ions. For Arixtra-ATIII complex, a single narrow peak at higher drift time range was observed for each charge state, implying that binding of Arixtra causes a change in the folded conformation of ATIII, leading to a specific structure with a larger CCS. These observations suggest that the tertiary structures of ATIII and the non-covalent Arixtra-ATIII complex are stable in the TWIMS experiment.

After calibration, the measured drift times for ATIII and its complex with Arixtra have been converted to CCSs, shown in Table 3.1. For each charge state, the reported CCS was an average from triplicate experiments, and was found to have a 0.6% relative standard deviation. The binding of Arixtra induces a 3.6% increase in CCS of ATIII, several times larger than the standard deviation in the measurements. The conformational change was highly reproducible, and was observed for each charge state, as well as for a variety of wave heights and wave velocities. The drift times and peak shape measured for residual bound ATIII, which is present at moderate abundance in the mass spectrum of the 1hour mixture, Figure 3.2 b, were the same as the drift times measurements for the ATIII peaks in the mass spectrum of the pure protein sample, Figure 3.2 a, evidence of the stability of the drift time measurements over different times and different samples.

Specificity and selectivity of ATIII-Arixtra interaction

To test the specificity of the interaction between ATIII and Hp, two control experiments were applied by altering the structure and conformation of both ATIII and its binding partner.

The first control experiment was to test the binding capability of ATIII after being denatured due to the presence of organic solvents under harsh pH condition (49.5 water: 49.5

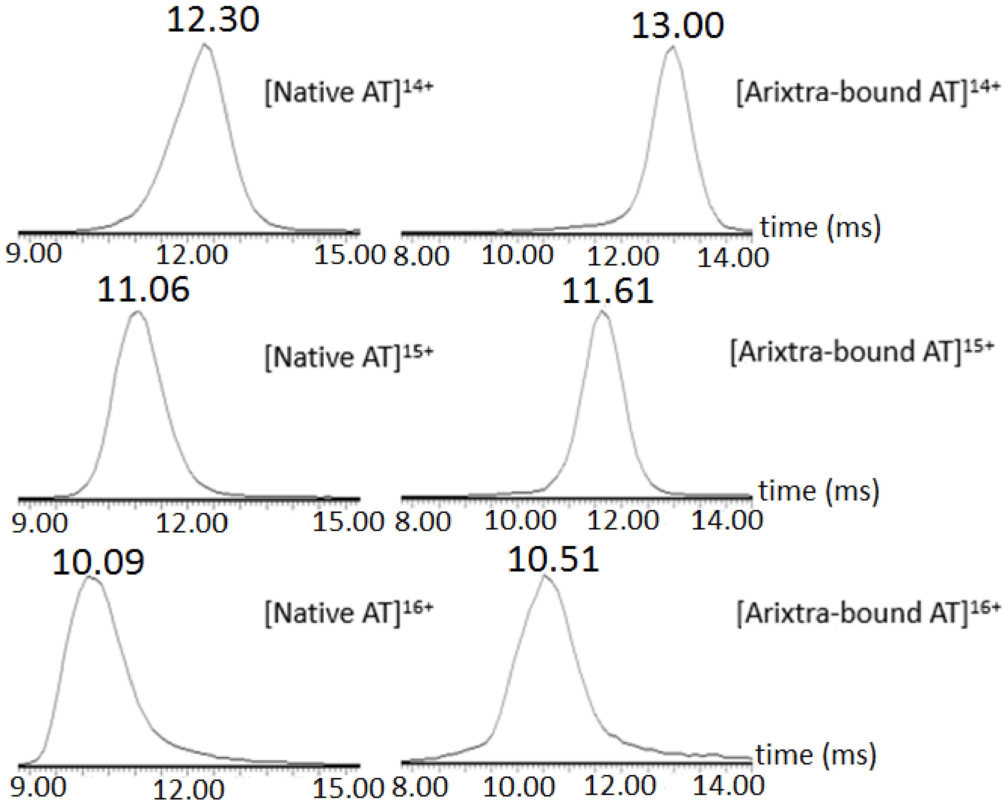


Figure 3.3 Drift timed distributions (ms) of the dominant charge states (+14~+16) of native ATIII and Arixtra-bound ATIII, at a wave height of 17V.

methanol: 1 formic acid). Denatured ATIII was incubated with Arixtra at a molar ratio of 1:3 for 12h and sprayed in denaturing solution (Figure 3.4). A wide distribution of higher charge states from + 47 to +23 were observed, in contrast to the narrow distribution of lower charge states under non-denaturing conditions. No peaks corresponding to the formation of ATIII-Arixtra complex were observed for these higher charge states, indicating that the fully denatured ATIII loses its capability to bind with Arixtra. This control eliminates the concern that the complexation of ATIII and Arixtra is non-specific, and results from the columbic attraction of a positively charged protein with an anionic carbohydrate. A mass scale expansion of the higher mass range of the denatured ATIII saw Arixtra-bound ATIII peaks for charge states lower than +20 (Figure 3.4, inset). It is possible that these relatively low charge states may represent some partially folded structures of ATIII.

A broad drift time distribution with multiple features was observed for each charge state of denatured ATIII (data not shown) while a single sharp peak was observed for native ATIII. This observation indicates that the fully extended or partially unfolded structures of ATIII have higher flexibility than the folded structures, and adopt multiple conformations in the gas phase.

Previous studies have collected evidence that Hp binding sites on proteins are more than clustered basic amino acid residues. Instead, several structural elements (loops, α-helices and β-strands) have been found in Hp binding sites, implying that specific spacing and spatial patterns of the Hp-binding basic residues are essential in Hp-protein interactions [3, 5, 50]. As in the case of ATIII-Hp binding, the folded, compact conformation of ATIII allows a spatial alignment of the Hp-binding basic residues which were widely distributed in the sequence [11, 51].This alignment facilitates the formation of a basic patch over several helices and the N-terminal region of ATIII, which provides a binding site for Hp. In this control experiment, denaturing the protein led to a substantial alternation of the topology of the Hp binding sites [4], therefore disrupting the high affinity binding interaction between ATIII and Hp. Our observation revealed the importance of a folded, compact structure of ATIII to maintain its biological activity.

A second control experiment was performed to test the binding specificity between ATIII and Arixtra induced by the pattern of sulfation. Arixtra has eight sites of sulfation, with a sequence of: $GlcNSO_3$, $6SO_3$-GlcA-$GlcNSO_3$, $3SO_3$, $6SO_3$*-$IdoA2SO_3$-$GlcNSO_3$, $6SO_3$, as shown in Figure 3.5. The ATIII binding behavior of a Hp hexasaccharide and a Hp tetrasaccharide, with the structures shown in Figure 3.5, were examined. The sulfation pattern of the Hp hexasaccharide closely resembles that of Arixtra, minus the 3-O sulfate group on the central glucosamine. The Hp tetrasaccharide lacks two sulfate groups and a monosaccharide on the non-reducing end, compared to the structure of Arixtra.

After incubating ATIII and the hexasaccharide, peaks corresponding to the hexasaccharide-bound ATIII complex, Figure 3.2 d, were observed, but in much lower abundance compared to Arixtra, for the same incubation time and Hp concentration. Increasing the Hp concentration produced only a small increase in the intensity of the hexasaccharide-bound ATIII complex peaks. This reaction never went to completion, as can be inferred from the presence of unbound ATIII ions which dominated the mass spectra, even after longer incubation time. All of these observations indicate that the binding affinity of the hexasaccharide towards ATIII is much lower than Arixtra, consistent with previous studies [52]. Similar results were observed for the binding between ATIII and the tetrasaccharide, with the peaks corresponding to ATIII-tetrasaccharide complex in even lower abundance (results not shown). Our results suggest that the hexasaccharide and tetrasaccharide both have reduced binding affinities for ATIII. The data confirms the significant contribution of the rare 3-O sulfation in Hp-protein interaction.

The drift time distributions of ATIII-negative control complexes were measured and only the results for +15 charge state were shown in Figure 3.6. For the complex of ATIII and hexasaccharide, a narrow drift time distribution was observed (Figure 3.6 b), almost identical to that measured for ATIII-Arixtra complex (Figure 3.6 c). The drift time measured for hexasaccharide-bound ATIII was the same as that measured for Arixtra-bound ATIII, which agrees with a previous study that concluded that removing the 3-O sulfate group affected the equilibrium of native ATIII and activated ATIII but not the conformational change associated with the equilibrium [53]. In other words, this control Hp oligosaccharide retains the specificity of binding, but has a reduced binding affinity.

In strong contrast, the Hp tetrasaccharide-ATIII complex revealed a very different behavior, exhibiting a broad peak with many features in the drift time spectrum, Figure 3.6 a, suggesting several structures for the complex. We interpret these as reduced specificity in the binding of the Hp ligand. The drift times measured for Hp tetrasaccharide-bound ATIII were smaller than that measured for Arixtra-bound ATIII, suggesting that the binding of the Hp tetrasaccharide may induce less of a conformational change within ATIII, due to the loss of several essential components from the specific pentasaccharide sequence [54].

The results of these control experiments are consistent with the known high-affinity, specific interaction between ATIII and Hp. It requires more than the existence of negatively charged functional groups on the Hp and positively charged residues on the protein. Instead, a well-defined three-dimensional presentation of the negatively charged groups in the Hp and a well-defined tertiary structure of the protein are critical to the specificity of ATIII-Arixtra binding.

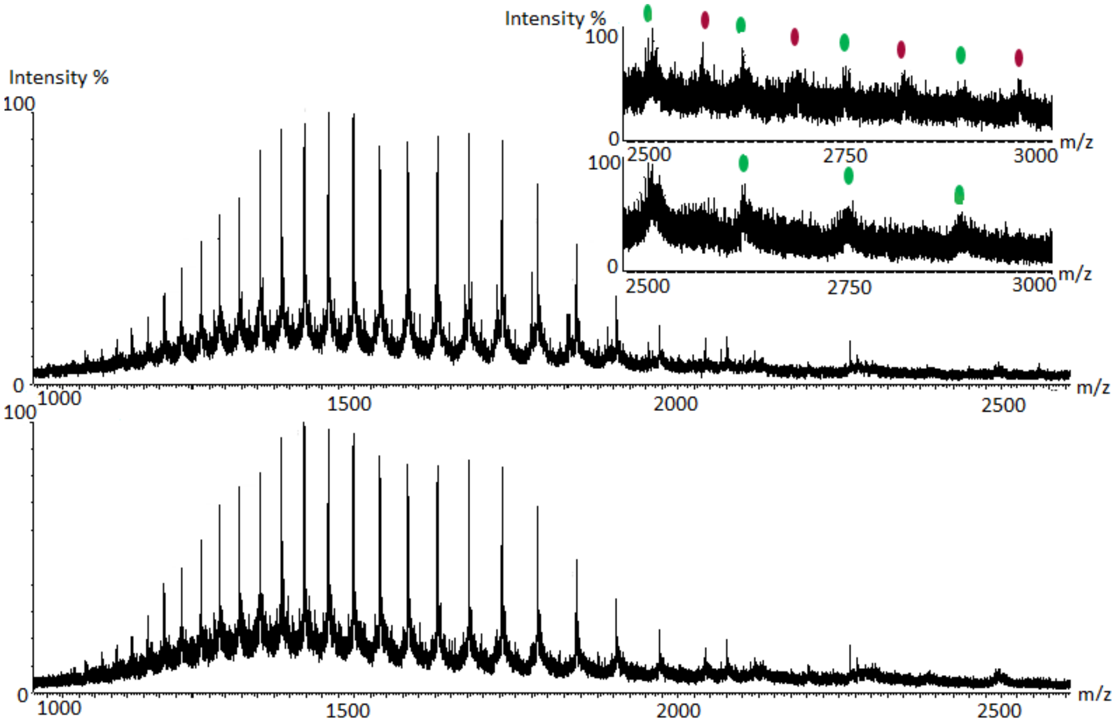


Figure 3.4 Electrospray ionization mass spectra of denatured ATIII with (upper) and without (lower) addition of Arixtra; The inset shows an expansion of the mass range (2500 m/z-3000 m/z) of the two spectra. Unbound ATIII peaks were labeled with green and Arixtra-bound peaks were labeled with purple.

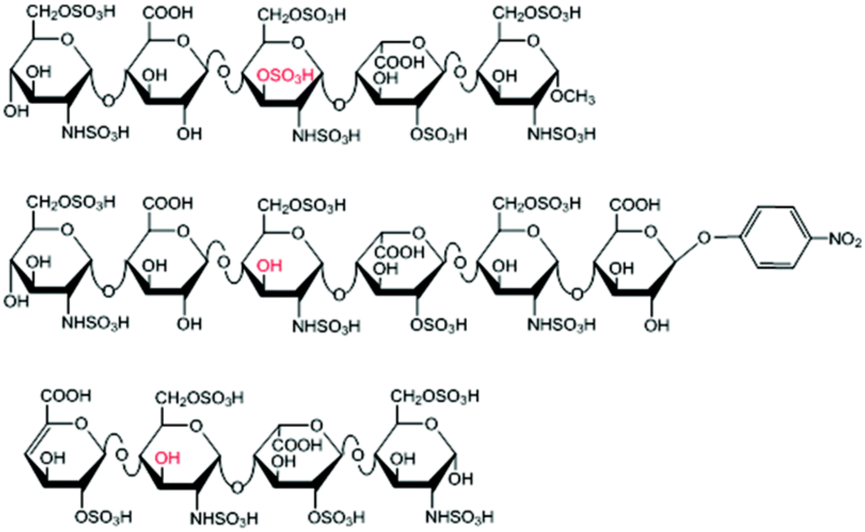


Figure 3.5 Structures of Arixtra (dp5 with 8 SO3) (upper), modified Arixtra (dp6 with 7 SO3) (middle) and the Hp tetrasaccharide (dp4 with 6 SO3) (below).

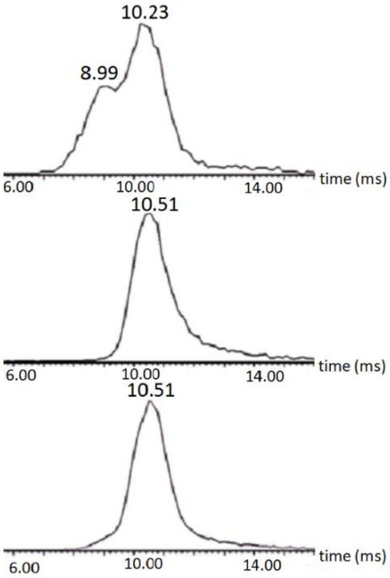


Figure 3.6 Drift time distributions (ms) of the +15 charge state of ATIII in complex with compounds, shown in Figure 3.5, the Hp tetrasaccharide (a), modified Arixtra (b) and Arixtra (c), at a wave height of 17V.

Gas phase stability of ATIII and ATIII in complex with Arixtra

Previous studies have shown that the addition of a ligand or counter ions brings in additional conformational stability of protein ions against collisional activation [55, 56]. To test whether the specific binding of Arixtra has altered the conformational stability of ATIII, we performed a series of experiments with incremental changes in trapping collisional energy (CE) from 5V to 30V being applied to ATIII with and without Arixtra in the trap ion guide prior to ion mobility separation. This probes the susceptibility to collision-induced unfolding for native ATIII and the ATIII-Arixtra complex.

The resulting drift time distributions of +15 charge state of both ATIII and ATIII-Arixtra complex are shown in Figure 3.7. Each drift time distribution was fitted with a minimum numbers of Gaussian distributions according to its shape using the peak analyzer in OriginPro 8.5.0 software (OriginLab Corporation, MA), as described previously [57]. Each Gaussian distribution generated as part of the fitness represents a potentially stabilized gas-phase conformation.

At a trapping CE of less than 20V (data not shown) or equal to 20V, a single, narrow peak was observed for both ATIII and ATIII-Arixtra complex, suggesting a single conformation of the protein and its Hp-bound complex.

The behavior of ATIII and ATIII-Arixtra complex started to deviate at elevated activation energy. For ATIII, a shoulder appeared on the high-drift-time side of the original peak when the trapping CE was raised to 22.5V. Four Gaussian distributions fit this drift time distribution, indicating that several partially unfolded or unfolded conformations of gas-phase ATIII ions were present at this collision energy. At a CE of 25 V, the abundance of the shoulder peak clearly increased. Two distinct populations of higher mobility (from 10.37 to 14.10 ms) and lower mobility (from 14.24 to 18.11 ms) conformations were observed, each modeled by two or three Gaussian distributions. At a CE of 27.5 V, a broad drift time distribution resolved by eight Gaussian distributions was observed, indicating the presence of a series of protein conformations of different degree of unfolding.

In contrast, for a range of trapping CE from 0 to 25V, the drift time distributions of ATIII-Arixtra complex remained unaffected, evidenced by a single, narrow peak in the drift time spectrum. A lower mobility peak corresponding to partially unfolded conformation of ATIII complex was not formed until a trapping CE as high as 27.5V was applied, while the higher mobility peak corresponding to the folded structure was still the dominant peak. Clearly, ATIII, when complexed with Arixtra, is more stable than ATIII itself, also consistent with solution-phase behavior of this system [58].

Previous studies showed that the binding of Hp/HS improves conformational stability of some proteins against heat and proteolysis [59]. The additional stability is also related to the activity of proteins [58]. Our results showed that ATIII in complex with Arixtra evidently possesses higher conformational stability than ATIII itself, indicating that the stability of Hp-activated ATIII may be a driving factor for its biological activity, which is responsible for shifting the conformational equilibrium towards greater stability.

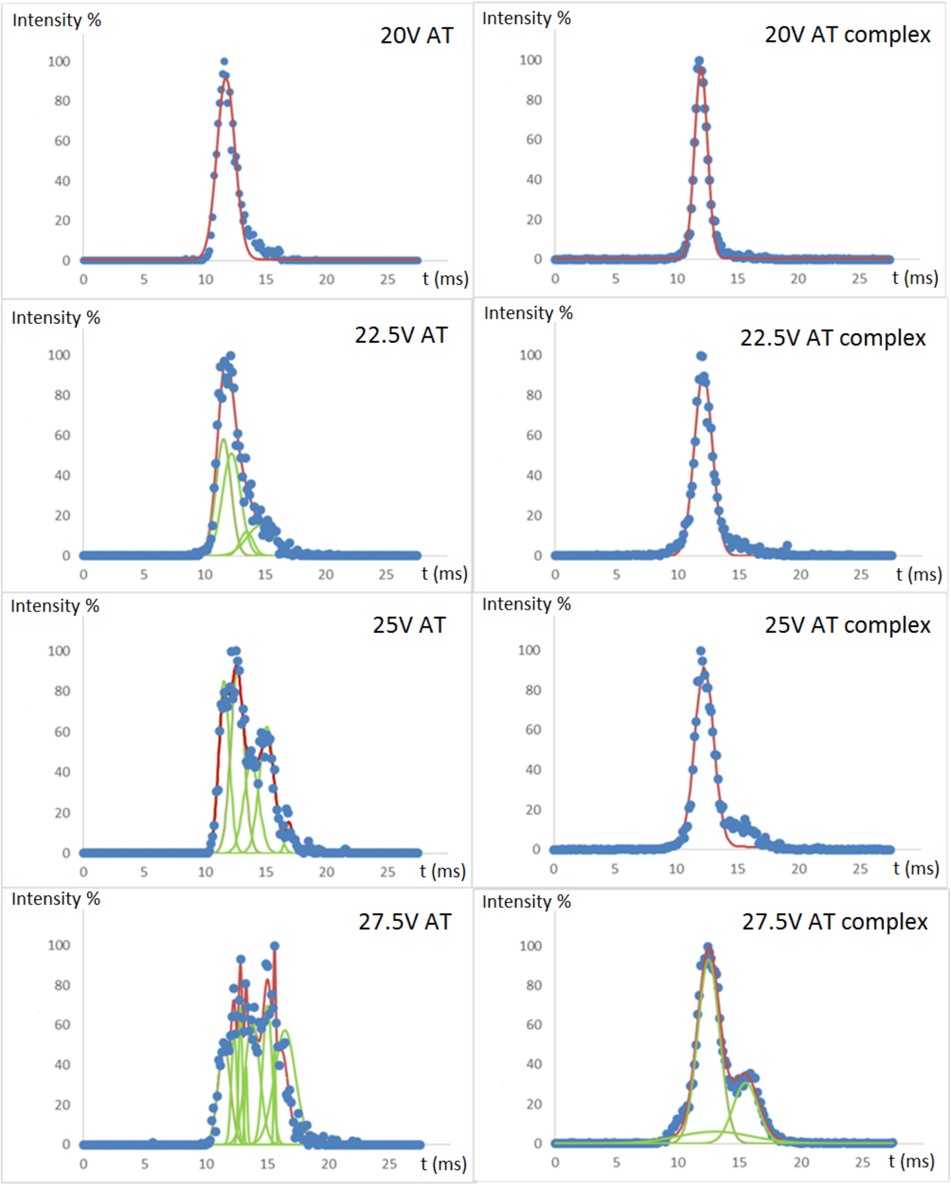


Figure 3.7 Experimental drift time distributions (ms) of +15 charge state of ATIII (left panels) and ATIII-Arixtra complex (right panels) as a function of trapping CE varying from 20V to 27.5V are shown in 2.5 V increments. The normalized experimental drift time distribution is shown by the blue dots. The fit of the data to a series of Gaussian distributions is shown in green, and the sum of the component Gaussians is shown in red.

Comparison of experimental CCS with theoretical CCS

Theoretical CCSs estimated from X-ray crystal structures of ATIII and Arixtra-bound ATIII were compared with experimentally measured CCSs. The PDB files of ATIII and ATIII-Arixtra complex display two ATIII molecules, one in the inhibitory form and the other in the latent form [34]. Only the inhibitory structure was used for the CCS calculation. In addition, N-linked oligosaccharide chains are present in the structures in both PDF files. We chose to leave the oligosaccharides out of the CCS calculation. The N-linked oligosaccharides are not thought to participate in Hp binding, so they are expected to contribute equally to the CCS of ATIII and ATIII-Arixtra complex. As we are interested in the conformation change within the protein rather than the absolute measurement of CCSs, the glycosylation-free estimation of CCS should allow us a direct and simple observation of the conformational change. We added the missing H-atoms to the crystal structures since their effect may not be ignored in the accurate estimation of the subtle conformational change.

The calculated CCSs of native ATIII were 3010.1 Å^2 estimated using PA and 3900.5 Å^2 estimated using TM. The calculated CCSs of Arixtra-bound ATIII were 3120.2 Å^2 estimated using PA and 4033.3 Å^2 estimated using TM. The experimental CCSs for the lowest charge state of native ATIII and ATIII-Arixtra complex were 3467.9 Å^2 and 3607.1 Å^2, respectively, comparable with those theoretical measurements.

The experimental CCS is approximately 12.1% less than the theoretical CCSs estimated using TM and 15.4 % higher than that estimated using PA. These observations are consistent with previous studies on other proteins, which found that reasonable experimental CCSs should lie in between the theoretical estimates made using PA and TM. The experimental CCSs should be smaller than TM results, due to the collapsed structure of protein ions in the gas-phase in response to desolvation [14, 60, 61]. Ignoring the long-range interaction with the neutral gas, collision effect and scattering process, PA results are on average 15% smaller than the experimental CCSs [62], in agreement with our observation.

It is noteworthy that there is close agreement between the relative changes in the CCS measured experimentally versus calculated theoretically. The averaged experimental conformational change was 3.6%, while the calculated conformational change was 3.7% for PA results and 3.4% for TM results, respectively. These results show that the TWIMS experiment generates gas-phase ATIII and Arixtra-bound ATIII ions with a conformation that matches the solution structure close enough, so that the conformational change due to the binding of Arixtra can be accurately measured.

Table 3.1 CCSs measured experimentally

Charge State	Free ATIII CCS (Å^2)	Arixtra-bound ATIII CCS (Å^2)	CCS increase (Å^2)	CCS increase%
14	3467.9±14.5	3607.1±16.6	123.5±6.6	3.4
15	3439.8±8.1	3573.1±7.0	133.2±2.3	3.7
16	3462.3±17.6	3591.4±15.9	133.0±8.4	3.7

Table 3.2 CCSs calculated using MOBCAL

System	PA method (Å2)	TM method (Å2)
AT III – Arixtra complex	3120.2±17.7	4033.3±79.5
AT III	3010.1±15.0	3900.5±53.0

Conclusions

TWIMS was found to provide data that is consistent with the known details of ATIII-Arixtra binding. Our mass spectra showed a 1:1 complex of ATIII with Arixtra, and the corresponding IM spectra were consistent with a single, folded gas-phase conformation of ATIII in its free and Arixtra-bound form, evidenced by a single narrow drift time distribution in both cases. CCSs derived from these data showed that the binding of Arixtra to ATIII caused a 3.6 % increase of ATIII's CCS. Both the absolute CCSs of ATIII and Arixtra-bound ATIII ions as well as the degree of CCS change were in agreement with theoretical CCSs calculated based on their X-ray crystal structures.

Furthermore, the selectivity and specificity of Hp-ATIII binding known from solution measurements appeared to survive the translation of the ions into the gas-phase, as revealed by the control experiments. Disrupting the folded structure of ATIII caused the loss of its binding affinity towards Arixtra. Removing the 3-O sulfate group from the known binding sequence resulted in substantial reduction of the abundance of ATIII-Hp complex but did not change the drift time behavior, as expected based on known solution behavior. These two control experiments not only confirmed the biological relevance of the native ATIII and Arixtra-bound ATIII ions that we observed, also provided evidence to support the existence and significance of the specificity of ATIII-Hp interaction. A control experiment using a Hp tetrasaccharide with a different pattern of sulfation than the known consensus sequence also showed greatly reduced affinity for ATIII, but also showed a reduction in specificity of binding, evidenced by a broad drift time distribution, consistent with a variety of structures for the complex. Moreover, the TWIMS measurement showed the stabilizing effect of Arixtra binding on the ATIII folded structure. This result also matched solution studies of ATIII and its Hp complexes.

Collectively, these results highlight the advantages of TWIMS for investigating GAG–protein interactions at molecular level. Future applications of this approach will surely contribute to achieving a better understanding of the biological processes mediated by GAG-protein interactions.

References

1. Rabenstein, D.L.: Heparin and heparan sulfate: structure and function. Natural product reports. **19**, 312-331 (2002)

2. Gandhi, N.S., Mancera, R.L.: The structure of glycosaminoglycans and their interactions with proteins. Chemical biology & drug design. **72**, 455-482 (2008)

3. Hileman, R.E., Fromm, J.R., Weiler, J.M., Linhardt, R.J.: Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. Bioessays. **20**, 156-167 (1998)

4. Raman, R., Sasisekharan, V., Sasisekharan, R.: Structural insights into biological roles of protein-glycosaminoglycan interactions. Chemistry & biology. **12**, 267-277 (2005)

5. Xu, D., Esko, J.D.: Demystifying heparan sulfate-protein interactions. Annual review of biochemistry. **83**, 129-157 (2014)

6. Capila, I., Linhardt, R.J.: Heparin–protein interactions. Angewandte Chemie International Edition. **41**, 390-412 (2002)

7. Koide, T.: Antithrombin and Heparin Cofactor II: Structure and Functions. Springer, (2008)

8. Langdown, J., Belzar, K.J., Savory, W.J., Baglin, T.P., Huntington, J.A.: The critical role of hinge-region expulsion in the induced-fit heparin binding mechanism of antithrombin. Journal of molecular biology. **386**, 1278-1289 (2009)

9. Gettins, P.G., Olson, S.T.: Exosite determinants of serpin specificity. Journal of Biological Chemistry. **284**, 20441-20445 (2009)

10. Whisstock, J.C., Pike, R.N., Jin, L., Skinner, R., Pei, X.Y., Carrell, R.W., et al.: Conformational changes in serpins: II. The mechanism of activation of antithrombin by heparin. Journal of molecular biology. **301**, 1287-1305 (2000)

11. Jin, L., Abrahams, J.P., Skinner, R., Petitou, M., Pike, R.N., Carrell, R.W.: The anticoagulant activation of antithrombin by heparin. Proceedings of the National Academy of Sciences. **94**, 14683-14688 (1997)

12. Olson, S.T., Richard, B., Izaguirre, G., Schedin-Weiss, S., Gettins, P.G.: Molecular mechanisms of antithrombin–heparin regulation of blood clotting proteinases. A paradigm for understanding proteinase regulation by serpin family protein proteinase inhibitors. Biochimie. **92**, 1587-1596 (2010)

13. Olson, S.T., Swanson, R., Raub-Segall, E., Bedsted, T., Sadri, M., Petitou, M., et al.: Accelerating ability of synthetic oligosaccharides on antithrombin inhibition of proteinases of the clotting and fibrinolytic systems Comparison with heparin and low-molecular-weight heparin. Thromb Haemost. **92**, 929-939 (2004)

14. Lanucara, F., Holman, S.W., Gray, C.J., Eyers, C.E.: The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nature chemistry. **6**, 281-294 (2014)

15. Uetrecht, C., Rose, R.J., van Duijn, E., Lorenzen, K., Heck, A.J.: Ion mobility mass spectrometry of proteins and protein assemblies. Chemical Society Reviews. **39**, 1633-1655 (2010)

16. Kanu, A.B., Dwivedi, P., Tam, M., Matz, L., Hill, H.H.: Ion mobility–mass spectrometry. Journal of Mass Spectrometry. **43**, 1-22 (2008)

17. Giles, K., Pringle, S.D., Worthington, K.R., Little, D., Wildgoose, J.L., Bateman, R.H.: Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. Rapid Communications in Mass Spectrometry. **18**, 2401-2414 (2004)

18. Ruotolo, B.T., Benesch, J.L., Sandercock, A.M., Hyung, S.-J., Robinson, C.V.: Ion mobility–mass spectrometry analysis of large protein complexes. Nature protocols. **3**, 1139-1152 (2008)

19. Seo, Y., Andaya, A., Bleiholder, C., Leary, J.A.: Differentiation of CC vs CXC chemokine dimers with GAG octasaccharide binding partners: an ion mobility mass spectrometry approach. Journal of the American Chemical Society. **135**, 4325-4332 (2013)

20. Hyung, S.-J., Robinson, C.V., Ruotolo, B.T.: Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes. Chemistry & biology. **16**, 382-390 (2009)

21. Ruotolo, B.T., Giles, K., Campuzano, I., Sandercock, A.M., Bateman, R.H., Robinson, C.V.: Evidence for macromolecular protein rings in the absence of bulk water. Science. **310**, 1658-1661 (2005)

22. Uetrecht, C., Barbu, I.M., Shoemaker, G.K., van Duijn, E., Heck, A.J.: Interrogating viral capsid assembly with ion mobility–mass spectrometry. Nature chemistry. **3**, 126-132 (2011)

23. Chen, S.-H., Chen, L., Russell, D.H.: Metal-Induced Conformational Changes of Human Metallothionein-2A: A Combined Theoretical and Experimental Study of Metal-Free and Partially Metalated Intermediates. Journal of the American Chemical Society. **136**, 9499-9508 (2014)

24. Bernstein, S.L., Dupuis, N.F., Lazo, N.D., Wyttenbach, T., Condron, M.M., Bitan, G., et al.: Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. Nature Chemistry. **1**, 326-331 (2009)

25. Williams, D.M., Pukala, T.L.: Novel insights into protein misfolding diseases revealed by ion mobility-mass spectrometry. Mass spectrometry reviews. **32**, 169-187 (2013)

26. Petitou, M., van Boeckel, C.A.: A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? Angewandte Chemie International Edition. **43**, 3118-3133 (2004)

27. Xu, Y., Cai, C., Chandarajoti, K., Hsieh, P.-H., Li, L., Pham, T.Q., et al.: Homogeneous low-molecular-weight heparins with reversible anticoagulant activity. Nature chemical biology. **10**, 248-250 (2014)

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

29. Smith, D.P., Knapman, T.W., Campuzano, I., Malham, R.W., Berryman, J.T., Radforda, S.E., et al.: Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. cell [Equation (1)]. **12**, 13 (2009)

30. Salbo, R., Bush, M.F., Naver, H., Campuzano, I., Robinson, C.V., Pettersson, I., et al.: Traveling-wave ion mobility mass spectrometry of protein complexes: accurate calibrated collision cross-sections of human insulin oligomers. Rapid Communications in Mass Spectrometry. **26**, 1181-1193 (2012)

31. Jurneczko, E., Kalapothakis, J., Campuzano, I.D., Morris, M., Barran, P.E.: Effects of drift gas on collision cross sections of a protein standard in linear drift tube and traveling wave ion mobility mass spectrometry. Analytical chemistry. **84**, 8524-8531 (2012)

32. Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C.V., Ruotolo, B.T.: Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. Analytical chemistry. **82**, 9557-9565 (2010)

33. http://depts.washington.edu/bushlab/ccsdatabase .

34. McCoy, A.J., Pei, X.Y., Skinner, R., Abrahams, J.-P., Carrell, R.W.: Structure of β-Antithrombin and the Effect of Glycosylation on Antithrombin's Heparin Affinity and Activity. Journal of Molecular Biology. **326**, 823-833 (2003)

35. Šali, A., Blundell, T.L.: Comparative protein modelling by satisfaction of spatial restraints. Journal of molecular biology. **234**, 779-815 (1993)

36. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C.: Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Structure, Function, and Bioinformatics. **65**, 712-725 (2006)

37. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L., et al.: GLYCAM06: a generalizable biomolecular force field. Carbohydrates. Journal of computational chemistry. **29**, 622-655 (2008)

38. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L.: Comparison of simple potential functions for simulating liquid water. The Journal of chemical physics. **79**, 926-935 (1983)

39. Götz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., Walker, R.C.: Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. Journal of chemical theory and computation. **8**, 1542-1555 (2012)

40. D. A. Case, J. T. Berryman, R. M. Betz, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A.

Kovalenko, T. S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York and P. A. Kollman, AMBER 2015, University of California, San Francisco, 2015.

41. Berendsen, H.J., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.: Molecular dynamics with coupling to an external bath. The Journal of chemical physics. **81**, 3684-3690 (1984)

42. Ryckaert, J.-P., Ciccotti, G., Berendsen, H.J.: Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. Journal of Computational Physics. **23**, 327-341 (1977)

43. Roe, D.R., Cheatham III, T.E.: PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. Journal of Chemical Theory and Computation. **9**, 3084-3095 (2013)

44. Humphrey, W., Dalke, A., Schulten, K.: VMD: visual molecular dynamics. Journal of molecular graphics. **14**, 33-38 (1996)

45. Mesleh, M., Hunter, J., Shvartsburg, A., Schatz, G., Jarrold, M.: Structural information from ion mobility measurements: effects of the long-range potential. The Journal of Physical Chemistry. **100**, 16082-16086 (1996)

46. Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Mobilities of carbon cluster ions: critical importance of the molecular attractive potential. The Journal of chemical physics. **108**, 2416-2423 (1998)

47. Chowdhury, S.K., Katta, V., Chait, B.T.: Probing conformational changes in proteins by mass spectrometry. Journal of the American Chemical Society. **112**, 9012-9013 (1990)

48. Fermas, S., Gonnet, F., Varenne, A., Gareil, P., Daniel, R.: Frontal analysis capillary electrophoresis hyphenated to electrospray ionization mass spectrometry for the characterization of the antithrombin/heparin pentasaccharide complex. Analytical chemistry. **79**, 4987-4993 (2007)

49. Abzalimov, R.R., Dubin, P.L., Kaltashov, I.A.: Glycosaminoglycans as naturally occurring combinatorial libraries: developing a mass spectrometry-based strategy for characterization of anti-thrombin interaction with low molecular weight heparin and heparin oligomers. Analytical chemistry. **79**, 6055-6063 (2007)

50. Lindahl, U., Li, J.p.: Interactions between heparan sulfate and proteins—design and functional implications. International review of cell and molecular biology. **276**, 105-159 (2009)

51. Huntington, J.A., Olson, S.T., Fan, B., Gettins, P.G.: Mechanism of heparin activation of antithrombin. Evidence for reactive center loop preinsertion with expulsion upon heparin binding. Biochemistry. **35**, 8495-8503 (1996)

52. Johnson, D.J., Huntington, J.A.: Crystal structure of antithrombin in a heparin-bound intermediate state. Biochemistry. **42**, 8712-8719 (2003)

53. Richard, B., Swanson, R., Olson, S.T.: The signature 3-O-sulfo group of the anticoagulant heparin sequence is critical for heparin binding to antithrombin but is not required for allosteric activation. Journal of Biological Chemistry. **284**, 27054-27064 (2009)

54. Desai, U.R., Petitou, M., Björk, I., Olson, S.T.: Mechanism of Heparin Activation of Antithrombin ROLE OF INDIVIDUAL RESIDUES OF THE PENTASACCHARIDE

ACTIVATING SEQUENCE IN THE RECOGNITION OF NATIVE AND ACTIVATED STATES OF ANTITHROMBIN. Journal of Biological Chemistry. **273**, 7478-7487 (1998)

55. Hopper, J.T., Oldham, N.J.: Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability. Journal of the American Society for Mass Spectrometry. **20**, 1851-1858 (2009)

56. Han, L., Hyung, S.-J., Mayers, J.J., Ruotolo, B.T.: Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent. Journal of the American Chemical Society. **133**, 11358-11367 (2011)

57. Shi, H., Pierson, N.A., Valentine, S.J., Clemmer, D.E.: Conformation types of ubiquitin [M+ 8H] 8+ ions from water: methanol solutions: evidence for the N and A states in aqueous solution. The Journal of Physical Chemistry B. **116**, 3344-3352 (2012)

58. Lima, M.A., Hughes, A.J., Veraldi, N., Rudd, T.R., Hussain, R., Brito, A.S., et al.: Antithrombin stabilisation by sulfated carbohydrates correlates with anticoagulant activity. MedChemComm. **4**, 870-873 (2013)

59. Gallagher, J.T.: Heparan sulfate: growth control with a restricted sequence menu. Journal of Clinical Investigation. **108**, 357 (2001)

60. Jurneczko, E., Barran, P.E.: How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst. **136**, 20-28 (2011)

61. Scarff, C.A., Thalassinos, K., Hilton, G.R., Scrivens, J.H.: Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. Rapid Communications in Mass Spectrometry. **22**, 3297-3304 (2008)

62. Hall, Z., Politis, A., Bush, M.F., Smith, L.J., Robinson, C.V.: Charge-state dependent compaction and dissociation of protein complexes: insights from ion mobility and molecular dynamics. Journal of the American Chemical Society. **134**, 3429-3438 (2012)

CHAPTER 4

GAS-PHASE ANALYSIS OF THE COMPLEX OF FIBROBLAST GROWTH FACTOR 1 WITH HEPARAN SULFATE: A TRAVELING WAVE ION MOBILITY SPECTROMETRY (TWIMS) AND MOLECULAR MODELING STUDY[2]

---

[2] Zhao Y, Singh A, Xu Y, Zong C, Zhang F, Boons GJ, Liu J, Linhardt RJ, Woods RJ, Amster IJ. 2017. *Journal of The American Society for Mass Spectrometry*. 28(1): 96-109. Reprinted here with permission of publisher.

Abstract

Fibroblast growth factors (FGFs) regulate several cellular developmental processes by interacting with cell surface heparan proteoglycans and transmembrane cell surface receptors (FGFR). The interaction of FGF with heparan sulfate (HS) is known to induce protein oligomerization, increase the affinity of FGF towards its receptor FGFR, promoting the formation of the HS-FGF-FGFR signaling complex. Although the role of HS in the signaling pathways is well recognized, the details of FGF oligomerization and formation of the ternary signaling complex are still not clear, with several conflicting models proposed in literature. Here, we examine the effect of size and sulfation pattern of HS upon FGF1 oligomerization, binding stoichiometry and conformational stability, through a combination of ion mobility (IM) and theoretical modeling approaches.

Ion mobility-mass spectrometry (IMMS) of FGF1 in the presence of several HS fragments ranging from tetrasaccharide (dp4) to dodecasaccharide (dp12) in length was performed. A comparison of the binding stoichiometry of variably sulfated dp4 HS to FGF1 confirmed the significance of the previously known high-affinity binding motif in FGF1 dimerization, and demonstrated that certain tetrasaccharide length fragments are also capable of inducing dimerization of FGF1. The degree of oligomerization was found to increase in the presence of dp12 HS, and a general lack of specificity for longer HS was observed. Additionally, collision cross sections of several FGF1-HS complexes were calculated, and were found to be in close agreement with experimental results. Based on the CCS, a number of plausible binding modes of 2:1 and 3:1 FGF1-HS are proposed.

Introduction

Fibroblast growth factors (FGFs) are mitogenic polypeptide growth factors expressed in all multicellular organisms [1]. Eighteen human FGFs have been identified so far, which share 13–71% sequence similarity and a homology core domain that consists of three copies of a basic four-stranded antiparallel beta sheet. These FGFs differ in their size, their interaction with receptors and their biological functions [2, 3]. FGFs are potent mitogens and mediate a multitude of biological processes including cell proliferation, cell differentiation, cell migration, morphogenesis, and angiogenesis [4]. Aberrant FGF signaling promotes tumor development, and tumor resistance to anti-angiogenic and other chemotherapies by directly driving cancer cell proliferation and survival [5]. FGF1 (along with FGF2) were the first FGFs to be identified and extensively studied [6].

The growth stimulatory activity of FGFs is achieved through direct interaction with two types of receptors on the cell surface [7]. The high affinity receptor FGFR are transmembrane proteins, four isoforms of which have been identified in mammals. Three extracellular immunoglobulin (Ig)-like domains and an intracellular cytoplasmic tyrosine kinase domain are linked though a single transmembrane helix to form the structure of FGFRs [8]. The ligand binding activity and specificity of FGFRs manifest through the Ig- domain II and III and the linker connecting the two domains [9]. Alternative splicing in Ig domain III dramatically alters the affinity of the FGFRs for FGFs and introduces the first level of specificity in FGF signaling for all FGF ligands except for FGF1, which serves as a universal ligand and activates all of the FGFRs [10].

The low affinity receptor heparan sulfate (HS) are linear sulfated glycosaminoglycan (GAG) polysaccharides comprising of alternating N-acetyl glucosamine and uronic acid disaccharide units linked through a 1–4 glycosidic linkage. HS are abundant components found on cell surfaces, for example, syndecans and glypicans, or in the extracellular matrix, such as perlecan, agrin or collagen XVIII [11].l. The structural polydispersity and heterogeneity of HS is a result of incomplete enzymatic modifications of HS in Golgi apparatus, including: (i) replacing N-acetyl with N-sulfation of glucosamine by N-deacetylase/N-sulfotranferases (NDSTs), (ii) epimerization of glucuronic acid to iduronic acid by C5-epimerase, (iii) addition of sulfo group to the 2-O position of iduronic acid by 2-O sulfotransferase, (iv) addition of sulfo group to the 6-O, and rarely, the 3-O positions of glucosamine by 6-O and 3-O sulfotranferases [12]. These non-template enzymatic modifications produce domains that are either highly N-sulfated, highly N-acetylated, or mixed domains of HS, which serve as potential binding site for a large array of HS binding proteins [13].

It has been generally appreciated that HS plays an essential role in transmitting extracellular signals to intracellular signaling pathways through direct interactions with both FGF and FGFR [14]. The binding of FGF to HS induces FGF oligomerization, increases the affinity of FGF to the extracellular part of FGFR, and further enables activation of FGFR and the formation of the highly stable HS-FGF-FGFR trimeric signaling complex [15]. The kinase domains from activated FGFRs are activated through trans-phosphorylation, followed by recruiting adapter protein to the activated receptors and triggering the downstream intracellular signaling cascades [16].

While the role of HS in FGF signaling has been well elucidated, the structural details of how the FGFs oligomerize in the presence of HS, and how the trimeric signaling complex of FGF, FGFR, and HS are integrated is still a matter of debate. Two physiologically relevant FGF dimerization configurations with HS have been proposed [17]: The *trans* oriented form features two FGF molecules arranged on opposite sides of one HS domain, and is characterized by a complete absence of FGF-FGF interactions. The *cis* oriented form features two FGF molecules present on the same side of the HS domain, requiring a HS with sufficient length. Many studies suggest that FGF1 favors the formation of *trans* form dimeric complex, and so far the crystallographic structure of only the *trans* form has been observed for FGF1 [18]. In contrast, for FGF2 a *cis* and a *trans* dimeric form may coexist [19]. In the presence of FGFR, two crystallographic ternary complex structures have been resolved, the 2:2:1 FGF1-FGFR2-HS dodecasaccharide model (or Pellegrini model [20]) and the 2:2:2 FGF2-FGFR1-HS dodecasaccharide model (or Schlessinger model [21]), differing in their preparation, binding stoichiometry, intermolecular interactions as well as the protein species used.

It has been long disputed whether FGF1 recognizes a consensus sequence with a specific sulfation pattern and minimum length, similar to the specific pentasaccharide sequences recognized by Antithrombin III (ATIII) [22]. HS octasaccharide and decasaccharide have been suggested to be the minimum length required for signaling [23], and a trisaccharide HS motif of IdoA2S-α-(1-4)-GlcNS6S-α-(1-4)-IdoA2S on HS oligosaccharides has been recognized to show strong binding affinity for FGF1 [24], but the minimum length of HS required to activate FGF signaling is still uncertain. There is considerable inconsistency in the literature about the binding specificity of FGF1-HS interaction, where some results suggest that overall degree of sulfation rather than a distinct sulfation pattern mediates this interactions [25, 26], while others propose binding selectivity which is highly susceptible to subtle changes to the fine structure of HS regarding particular type of sulfations [27-30]. Studies of the binding sequence for FGFs show that they each have different preferences for the sulfation pattern of HS [29]. Additionally, some other studies indicate that specific FGF-FGFR pairs, rather than FGF itself, dictate binding specificity and affinity [31]. For example, HS deficient in 2-O- and 6-O- sulfation still maintains its activating effect on FGF1-FGFR2 signaling but not FGF1-FGFR1 or FGF7-FGFR2b signaling [32].

A wide array of biophysical techniques including X-ray crystallography [18], NMR [33], size exclusion chromatography [34], size exclusion chromatography followed by hydrophobic trapping[35], isothermal calorimetry [36], surface plasmon resonance [37], affinity co-electrophoresis [38] have been applied to elucidate the nature of these molecular interactions. Recently, ion mobility-mass spectrometry (IMMS), a gas-phased biophysical technique that separates biomolecules based on their mobilities has emerged and found numerous applications such as measuring conformational change in proteins upon ligand-binding [39], determining binding affinity and specificity [40], monitoring real-time conformational dynamics [41], detecting intermediate structures [42], and elucidating the architecture of protein assemblies [43]. Furthermore, with the commercialization of Synapt instruments from Waters Corporation, travelling wave ion mobility spectrometry (TWIMS) has become a widely used approach for the study of biomolecular structure and dynamics [44]. IMMS studies are often supported by molecular modeling studies, which provide computational predictions of molecular structure, binding orientations, binding affinities as well as theoretical estimation of the molecular shape (collision cross sections (CCS), and serve a critical role for the correct interpretation of experimental findings.

We recently applied TWIMS to examine the well-characterized, highly specific interaction between ATIII and a synthetic heparin (Hp) pentasaccharide [40]. We successfully measured and theoretically validated the conformational change within ATIII induced by the binding of Hp while maintaining the solution structure of ATIII and its complexes as well as the binding selectivity and specificity of Hp-ATIII interaction. In this study, we extended this method to a more complicated system of FGF1 signaling, in order to investigate the conformational and stoichiometric details of the binding of FGF1 and HS. TWIMS experiments and molecular modeling techniques were combined in order to address: the minimum length of HS required to induce FGF1 dimerization, the binding stoichiometry for the interaction of FGF1 and HS of varying sizes and sulfation patterns, the selectivity of the interaction between FGF1 and HS, and the conformational stability of FGF1 and its HS bound complexes. This work provides more details to the model of the manner in which HS interacts with FGF1, as well as additional validation for the approach of using gas-phase measurements of GAG-protein interactions to derive biologically relevant structural details.

Experimental

Reagents

All chemicals and solvents (ammonium acetate, methanol, water and formic acid) were of HPLC grade and purchased from Sigma-Aldrich. Human recombinant FGF 1 expressed in *Escherichia coli* was a gift from Amgen (Thousand Oaks, CA). Protein calibrants (myoglobin from equine heart, cytochrome c from equine heart, avidin from egg white, concanavalin A from *Canavalia ensiformis* and bovine serum albumin) were purchased from Sigma-Aldrich. HS dodecasaccharides were chemoenzymatically synthesized as previously described [45]. HS tetrasaccharides were synthesized as previously described based by fluorous supported modular synthesis [46].

Sample preparation

For MS analyses under non-denaturing conditions, FGF1 was diluted in 20 mM ammonium acetate buffer, pH 6.8, to a final concentration of 10 μM. FGF1-HS complex was obtained by incubating FGF1 with HS oligosaccharides at room temperature for 60 min. Protein calibrants were diluted in either denaturing solution or non-denaturing solution to a final concentration of 10 μM.

IMMS measurement and data analysis

NanoESI-IMMS experiments were performed using a quadrupole-TWIMS-TOF hybrid mass spectrometer (Synapt G2 HDMS, Waters Corp., Manchester, UK) in positive ionization mode. Protein samples were injected into the nanoESI source through a fused-silica emitter (PicoTip New Objective, Woburn, MA) with a flow rate varying from 0.2-0.5 ul/min. Experimental parameters were carefully tuned to prevent the protein and protein complex from unfolding or losing integrity due to extensive activation while keeping substantial ion transmission, including controlling the collisional energy in the ion guides and keeping the source temperature and desolvation energy low. The applied experimental parameters were: capillary voltage, 1.4 kV; sampling cone voltage, 20 V; extraction cone voltage, 5 V; source temperature, 30 °C; flow rate of nitrogen in the IM ion guide, 50 ml/min; flow rate of helium in the helium cell, 180 mL/min; transfer collision energy, 0 V. Different sets of wave height and corresponding wave velocity were examined to optimize the mobility separation. The drift times of the calibrants and FGF1 samples were measured when identical experimental conditions were stringently applied.

For monitoring the collisional induced unfolding (CIU) of FGF1 and FGF1-HS complexes, protein ions at selected charge state were isolated in quadrupole mass filter and activated in the trap ion guide where subsequent collisional induced activation of ions takes place, followed by separation in the TWIMS ion mobility stage, and detection by TOF-MS.

Data analysis was performed using MassLynx 4.1 and Driftscope (Waters Corp., Manchester, UK). The CIU data was analyzed using CIUSuite [47]: CIU fingerprints of protein ions, which records the relative ion intensity (normalized and smoothed using Savitsky-Golay filter) as a function of collision energy and drift times, were shown in a 2-D contour plot using the CIUSuite_plot function; The CIUsuite detect function was applied to identify and extract the most intense CIU features of the fingerprint based on a first derivative analysis, providing centroid drift times, transitional voltage and stability range of each detected feature.

Converting drift times into CCSs

The CCS measurements were calibrated based on an empirical relationship between the drift times of protein calibrants and their known CCS obtained previously by DTIMS [48]. Briefly, a selected set of native and denatured protein calibrants, with a mass range from 12KDa to 102KDa and a CCS range from 2303 Å2 to 5550 Å2 were employed. The drift times of these calibrates were corrected for mass-dependent flight time spent in the transfer ion guide and TOF mass analyzer and mass-independent flight time spent in the transfer ion guide. The CCSs of calibrants were corrected for their charge state and reduced mass with respect to the buffer gas. The natural logarithm of corrected CCSs were plotted against the natural logarithm of corrected drift times and a mathematical formula ($\mathbf{Ln\ \Omega' = A \times Ln\ dt' + B}$) was derived. The calibration coefficient A was extracted to calculate the effective drift times dt": $\mathbf{dt''}=(\mathbf{dt'})^{A}\frac{z}{\sqrt{\mu}}$. A calibration curve was generated by plotting the literature CCSs as a function of dt". The experimental CCS of the analyte ion can be derived from this calibration curve based on the measured drift time [49, 50].

Preparation of theoretical models

FGF1– dp4 complex - The Protein Data Bank (PDB) contains several models of wild type human FGF1 in complex with heparin oligosaccharide, and each of them displays slightly differing contacts with the protein. The interaction energy between each protein-HS tetrasaccharide complex was calculated using a short minimization in implicit solvent followed by single frame binding free energy calculation to select one model for further theoretical investigation (see details in Supplementary Data and Supplementary Table 1). The complex that shows the strongest interaction (2AXM chain B) [18] was used for all further calculations.

FGF1-HS complexes for d6, dp8, dp10 and dp12 - 2AXM chain B bound to Hp hexasaccharide was used to build the model for 1:1 and 2:1 FGF1-HS hexasaccharide complex. In order to avoid disrupting the contacts that the hexasaccharide made with the two FGF1 molecules, the heparin sequence from the crystal structure was kept, which is different from the experimentally used dp6 fragment.

The chosen dp4 complex was used and the heparin sequence was extended on either side using tleap o build the FGF1–dp8 complexes. The FGF1-dp6 complex was used and additional residues were added on either side of the existing dp6, according to the experimentally used sequences (Table 1), using tleap to build the FGF1–dp10, and the FGF1-dp12 complexes.

2:1 FGF1–dp12 cis models and 3:1 FGF-HS-dp12 model - Two models were built for the 2:1 FGF1-fully sulfated dp12 in the cis conformation. Cis-model-1 was modeled using two

FGF1-dp6 molecules connected through the GAG chain using tleap. First chain D of the crystal structure of apo-FGF1 dimer (PDB id: 2AFG [51]) was superimposed on chain A of FGF1-dp6 crystal structure (PDB id: 2AXM) and the coordinates for the dp6 were transferred to make 2AFG-dp6 complex to build the second 2:1 *cis* complex (Cis-model-2). This dp6 structure was extended to match the experimentally used fully sulfated dp12 by adding 3 residues on either side, using tleap.

3:1 FGF1-fully sulfated dp12 complex was built by connecting the available crystal structure for 2:1 FGF1-dp6 dimeric complex with a 1:1 FGF1-dp6 monomeric complex through the HS chain.

Molecular dynamics simulations

Topology and coordinate files for each system were generated using the tleap program, employing the ff99SB [52] and GLYCAM06 (version j) [53, 54] parameters for the protein and GAGs, respectively. Each system underwent energy minimization (1000 steps) in implicit solvent (IGB = 2). The net charge on each system, after energy minimization, was neutralized by addition of an appropriate number of counter ions (Na^+ or Cl^-). This was followed by solvation with TIP3P [55] water molecules in a cubic box extending at least 12 Å from any atom of the solute.

All MD simulations were performed with the GPU implementation of pmemd, pmemd.cuda_SPDP [56] in Amber14 [57]. Energy minimization of the solvent was performed in an NVT ensemble (1000 steps of steepest descent, 24000 steps of conjugate gradient), followed by a full system energy minimization (1000 steps of steepest descent, 24000 steps of conjugate gradient). The systems were heated from 5 K to 300 K over 60 ps in an NVT ensemble, with a weak positional restraint (10 kcal/mol-Å^2) on the atoms in the solute. A Berendsen-type thermostat [58] with a time coupling constant of 1 ps was utilized for temperature regulation. Equilibration and production was performed at constant pressure (NPT ensemble; 1 atm) with a pressure relaxation time of 1 ps. After the heating step, the restraints were removed from the solute atoms, and the entire system was allowed to equilibrate at 300K for 1 ns. All covalent bonds involving hydrogen atoms were constrained using the SHAKE [59] algorithm, allowing a simulation time step of 2 fs. Scaling factors for 1–4 interactions were set to the recommended values of 1.0 and 1.2 for the GAG [53] and protein [52], respectively, and a non-bonded interaction cutoff of 8.0 Å was employed. Long-range electrostatics were computed with the particle mesh Ewald (PME) method. Data were collected for 200 ns for each system.

Post processing of the MD simulations was performed using CPPTRAJ [60] module of Amber. The graphical representations were generated using VMD [61].

Binding energy calculations

Models of the gas-phase complexes were generated from the MD simulations by removing the water molecules and subjecting the biomolecules to *in vacuo* minimization (*IMIN* = 5) with the SANDER module of AMBER14. This was performed for 5000 frames selected at even steps from the last 100 ns of the solvated simulation, and the resultant gas-phase trajectory was used to calculate the binding free energy in the gas phase using the MMPBSA.py script [62].

Collision cross section calculations

Collision cross sections (CCS) were calculated using 50 frames from the last 100 ns of the simulation using MOBCAL [63]. Each of these frames was minimized *in vacuo*, followed by the calculation of CCS using the projection approximation and trajectory methods.

Results and discussion

IMMS of unbound FGF1 and HS tetrasaccharide-bound FGF1

Under non-denaturing conditions, in the absence of HS, FGF1 exists mainly in the form of a monomer with a charge state distribution ranging from +9 to +7, accompanied by a lower abundance by the form of dimer with a charge state distribution ranging from +13 to +11 (Figure 4.1). This self-association of FGF1 has been reported before [51, 64], but the biological relevance of the self-association that we observed is not clear due to its low abundance. The narrow distribution of lower charge states indicates that FGF1 ions adapt a folded and compact conformation with fewer basic sites exposed for protonation [65]. Each observed charge state displays three peaks, corresponding to the sequence mass of FGF1 and masses corresponding to variants with one less amino acid from either the N-terminus or C-terminus. Due to the compositional complexity, only the peaks of lowest mass were selected for further ion mobility analysis.

In an effort to examine the binding stoichiometry for the interaction of FGF1 and HS, a series of tetrasaccharides (dp4) with different sulfation patterns were incubated with FGF1 in a molar ratio of 1:2. The HS dp4a contains the high affinity binding motif for FGF1 IdoA2S-α-(1-4)-GlcNS6S-α-(1-4)-IdoA2S [24]. HS dp4b has the same overall level of sulfation as the dp4a, but the 2-O sulfation on the first residue from the non-reducing end was shifted to 3-O sulfation on the second residue, accompanied by the change of uronic acid stereochemistry (GlcA instead of IdoA) at the non-reducing end. The HS dp4c lacks two 2-O sulfo groups compared with dp4-1. In the presence of HS, the formation of the 1:1 FGF1-HS complex was observed (Figure 4.1) without altering the charge state distributions of apo-FGF1. It is noteworthy that two peaks corresponding to a complex of two FGF1s with one HS (2:1 FGF1-HS complex) were detected only for the HS dp4a containing the high affinity binding motif (Figure 4.1 d).

A majority of previous studies suggest that dimerization of FGF1 or complexation of FGF1 with HS and FGFR requires a long-chain HS oligosaccharides (at least a hexasaccharide or octasaccharide) [23, 66]. Other studies showed that a fully sulfated dp4 is sufficient for a high-affinity interaction with FGF [67, 68], for stabilizing FGF1 against thermal unfolding or digestion [69], and for initiating mitogenic activity [70]. A dimeric complex of FGF1 and HS as well as trimeric complex of HS, FGF1 and FGFR1 produced in the presence of tetrasaccharides containing the critical binding motif have been detected using MALDI mass spectrometry [71] and gel mobility shift assay [68]. Our observations highlight the importance of this critical binding motif, which imparts upon HS oligosaccharides as short as a tetrasaccharide the potential to induce FGF1 dimerization.

The arrival time distributions of the +7 charge state of FGF1 alone, the +7 charge state of 1:1 FGF1-HS dp4a complex, and the +12 charge state of 2:1 FGF1-HS dp4a complex are shown in Figure 4.2 a, recorded at a wave height of 17 V. One narrow arrival time distribution was observed for each species, indicating the presence of one compact and folded conformation for the protein ion and the non-covalently associated protein-HS assemblies at both binding stoichiometries. These observations suggest that the solution structure of FGF1 survived in the gas phase environment of TWIMS, maintaining its capability to bind to HS.

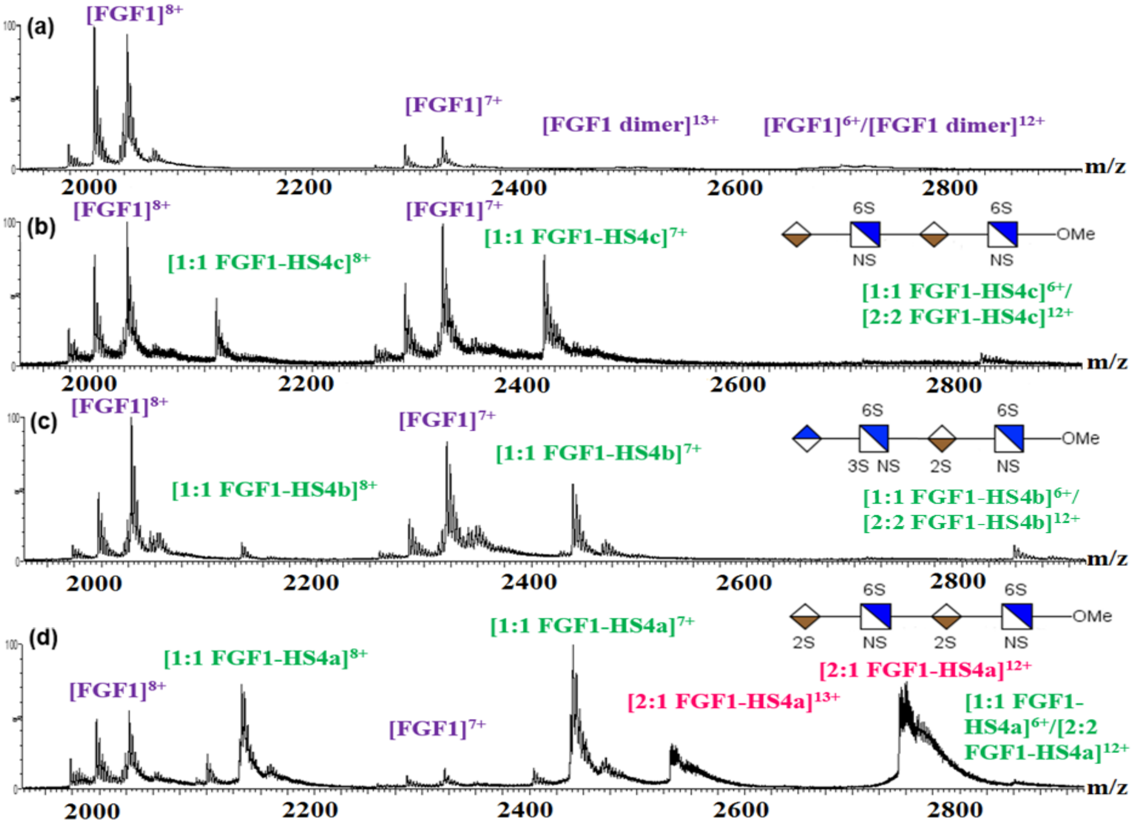


Figure 4.1 ESI mass spectra, obtained under non-denaturing conditions, of (a) FGF1 itself; (b) FGF1 incubated with HS dp4c at 2:1 ratio; (c) FGF1 incubated with HS dp4b at 2:1 ratio; (d) FGF1 incubated with HS dp4a at 2:1 ratio. Associating of two FGF1s by one HS (2:1 FGF1-HS complex) were observed only for the HS dp4a containing the high affinity binding motif.

The peak corresponding to the even charge state of the 1:1 FGF1-HS complex overlaps in the mass spectrum with that corresponding to the 2:2 FGF1-HS complex with a doubled charge. However, the dimeric complex has a decreased charge-normalized cross section [72], so that it

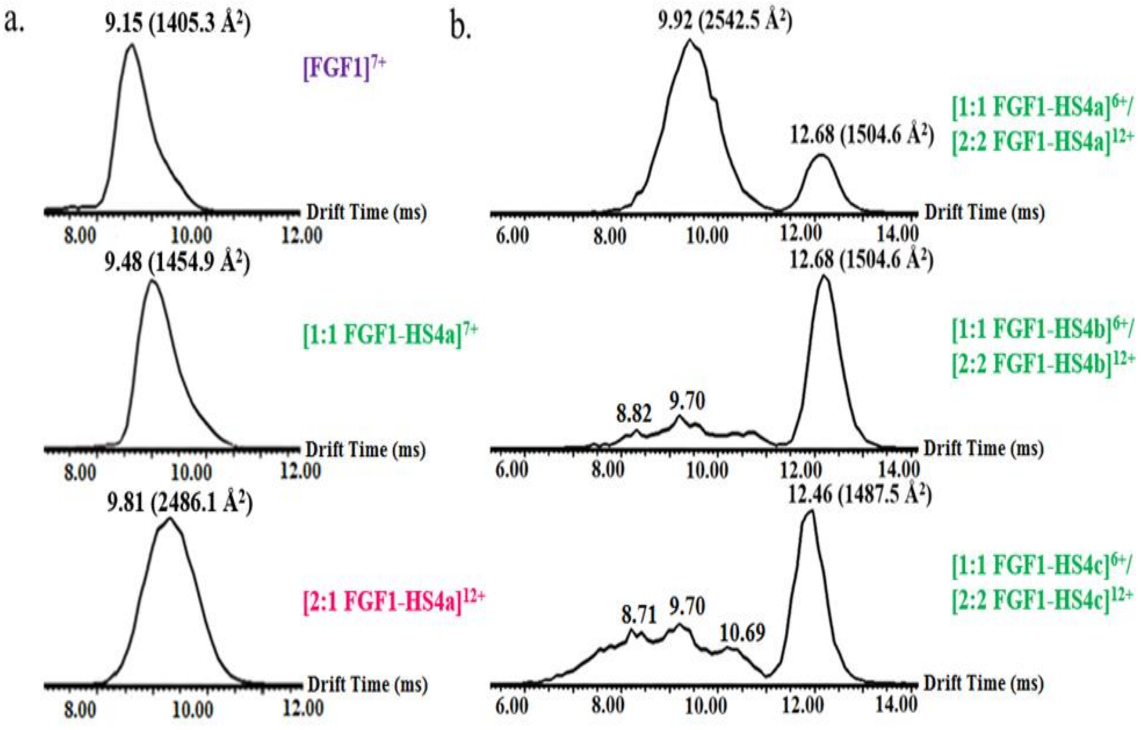


Figure 4.2 (a) Arrival time distributions (ms) of +7 charge state of FGF1 (upper), +7 charge state of 1:1 FGF1-HS dp4a complex (middle) and +12 charge state of 2:1 FGF1-HS dp4a complex (lower). The narrow arrival time distribution observed for each species indicates that the native structure of FGF1 and the non-covalent complex of FGF1 with HS are stable in the TWIMS experiment. (b) Arrival time distributions (ms) of +6 charge state of 1:1 FGF1-HS complex and +12 charge state of 2:2 FGF1-HS complex, with HS4a (upper), HS4b (middle) and HS4c (lower). The comparison between the narrow, well-defined arrival distribution of 2:1 FGF1-HS dp4a complex and the broad arrival time distribution of 2:1 FGF1-HS dp4b complex or 2:1 FGF1-HS dp4c complex, indicates that the high affinity binding motif is responsible for the level of specificity in the dimerization of FGF1 upon the binding of HS.

can be separated from the monomeric complex on the basis of a difference in drift times in a TWIMS experiment (Figure 4.2 b). The feature with the longer drift time corresponds to the 1:1 FGF1-HS dp4 complexes, each of which exhibits a narrow drift time distribution as observed for the same species in a different charge state (Figure 4.2 a). The features with shorter drift times correspond to 2:2 FGF1-HS complexes. One narrow arrival time distribution was detected for the FGF1-HS dp4a mixture, implying the presence of a single compact, ordered gas phase conformation, in strong contrast to the broader peak with many features detected for the FGF1-HS dp4b mixture or the FGF1-HS dp4b mixture, implying the existence of a range of less specific binding modes.

The fact that subtle alterations of the HS sequence with regard to the sulfation pattern or epimerization lead to noticeable changes in both the binding stoichiometry and the binding mode of the FGF1-HS interaction, suggests that the interactions between FGF1 and HS has certain degree of selectivity. These observations strengthen the importance of the critical binding motif in modulating the assembly of FGF1 and HS through a specific dimerization mode.

IMMS of HS dodecasaccharide bound FGF1 complexes

A series of dodecasaccharides (dp12) with varying sulfation patterns were incubated with FGF1 at a molar ratio of 1:4 (GAG to protein) (Figure 4.3). An excess of FGF1was used to saturate the FGF1 binding epitopes on the dp12s. In the presence of the fully sulfated HS dp12, the binding of one HS dp12 to three FGF1 (3:1 FGF1-HS complex) were detected with a charge state distribution ranging from +15 to +13, accompanied by the binding of one HS dp12 to two FGF1 (2:1 FGF1-HS complex) with a charge state distribution ranging from +13 to +11. Peaks corresponding to apo-FGF1 were also detected, but in very low abundance compared to that for the complexes. Previous studies showed that the interaction between FGF1 and HS is driven by positive cooperativity [34, 73], so that FGF1 oligomerizes in preference to staying in monomeric state when bound with long-chain HS and the formation of higher order of FGF1 oligomers was observed as the length of HS increases [74]. Our observations are consistent with these previous studies and highlight the biological function of HS as a storage reservoir for FGFs on the cell surface [23].

The presence of de-2-O sulfated or de-6-O sulfated dp12 also resulted in two types of binding stereochemistry to FGF1, but their abundance was much lower than that the fully sulfated dp12 for FGF1. In the presence of de-2-O and 6-O sulfated dp12, only a small fraction of FGF1 formed 2:1 FGF1-HS complex whereas most of FGF1 stayed in apo form. The marked difference in binding stoichiometry implies that the absence of 2-O and 6-O sulfo groups dramatically reduces the binding affinity.

Although the modified dp12s exhibit varying binding stoichiometry to FGF1, the arrival time distribution of the 3:1 FGF1-HS dp12 complexes all exhibited a single, narrow peak. The same observation has been reported for a different GAG-protein binding system, chemokine CCL2 and HS [75]. The fact that the ion mobility profiles of FGF1-modified HS dp12 complexes are similar implies that changes in HS structure lead to altered interactions with FGFs, but these alterations may not dramatically affect the compact, gas phase conformation adapted by these complexes.

Our observations suggest that long-chain HS may not place as stringent requirement on the fine structure of HS sequence as short-chain HS, indicating a relative lack of specificity of binding for long-chain HS. Instead of presenting a specific binding motif for FGF1, the sequence of HS may be composed of many domains with different sulfation patterns and binding affinities toward FGF1, forming an affinity gradient on the cell surface that regulates FGFs signaling by directing them to the site of interaction [76].

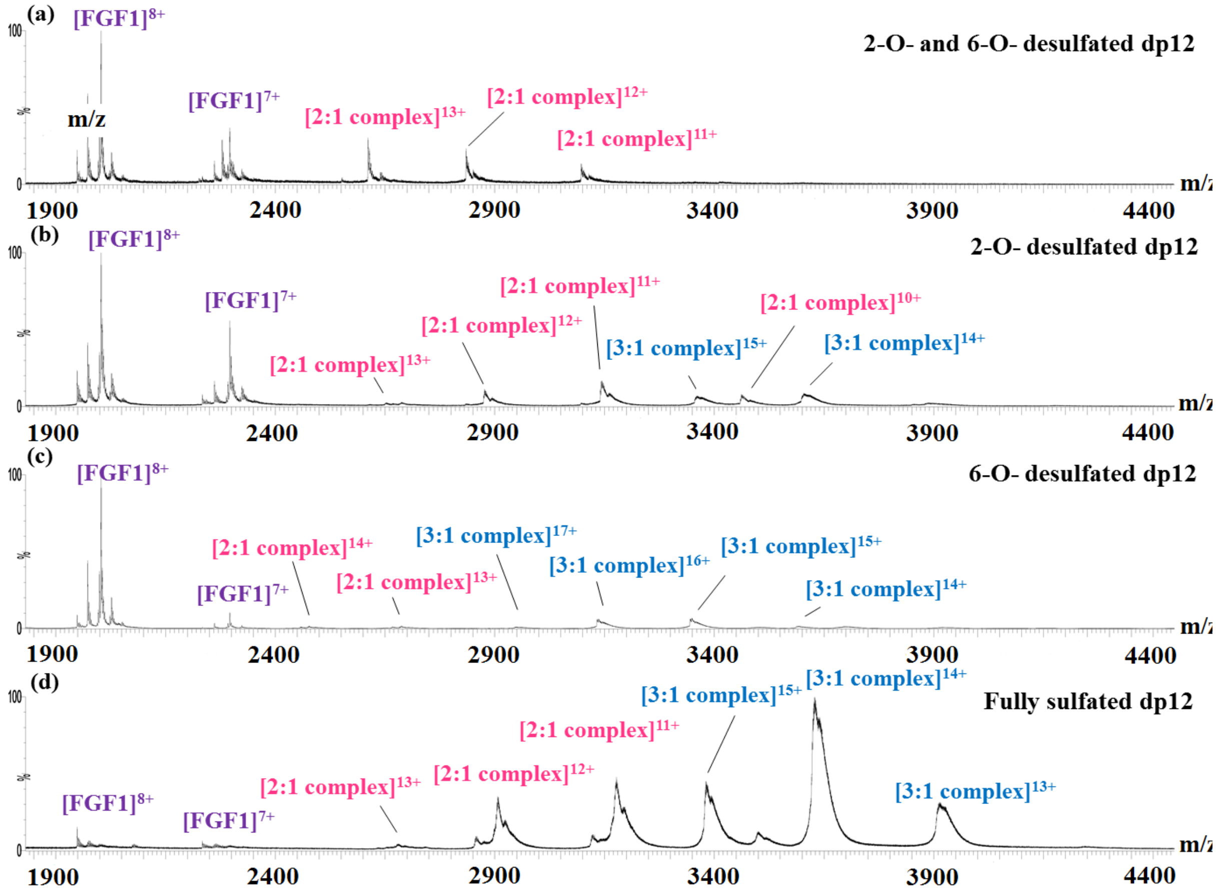


Figure 4.3 Native ESI mass spectra of (a) FGF1 with de-2-O- and de-6-O- sulfated dp12; (b) FGF1 with de-2-O- sulfated dp12; (c) FGF1 with de-6-O- sulfated dp12; (d) FGF1 with the dp12 that contains the critical binding motif. Structures of the dp12s are shown in supplementary data.

Collision cross sections of FGF1-HS complexes

A comparison of the experimental and theoretical collision cross sections (CCSs) for the HS bound FGF1 complexes were carried out to determine their binding mode as well as the higher order oligomerization behavior of FGF1. The experimental CCSs for the complexes (Table 1) were calibrated based on an empirical relationship between the drift times of protein calibrants and their known CCSs obtained previously by conventional drift time IMS. For theoretical estimation we used both the projection approximation (PA) [77] and the trajectory (TJ) [63] method. PA estimates the orientationally-averaged CCS by projecting the molecule onto a randomly chosen plane in space, drawing a circle with the corresponding collision radius around the projection of each atom, and estimating the area of the "shadow" from all faces [78]. This method is known to underestimate CCS since it fails to consider the influence of long-range interactions and the scattering between the ion and neutral gas. TJ represent the ions as a collection of atoms defined by their Lennard-Jones (12-6-4) potential, and relates the scattering angle of the buffer gas molecules before and after collision with the molecule's geometry. An orientationally-averaged CCS can be obtained by integrating over all possible collision geometries of the ion, however, this method is computationally much more expensive than the PA method [63]. The experimental CCSs are generally found to be larger than those computed using the PA method, and smaller than the TM results [79]. A previous study [80] have shown that there is a direct scaling relationship between experimental CCSs and CCS measured using PA methods, so calibrating the PA CCSs though this scaling relationship should provide good standard for correct interpretation of experimental data.

Each model of the 2:1 FGF1-HS complex (Figure 4.4 a) was based on the crystal structure 2AXM, which contains a heparin hexasaccharide bound to two molecules of FGF1 in *trans* orientation. The FGF1-dp4 complex was built by removing two residues from the hexasaccharide. For the complexes formed with oligosaccharides whose chain length is longer than dp6, we extended the crystal hexasaccharide sequence to build models for dp8 and higher in a relatively straightforward manner. However, considering the fact that changing the interactions at the binding interface between the two FGF1 molecules to accommodate GlcA in place of IdoA would affect the interaction dramatically, we chose to use the crystal structure for the FGF1-dp6 modeling.

Molecular dynamics (MD) simulations of these complexes were carried out in explicit solvent, followed by *in vacuo* minimization of a subset of the frames from the MD trajectories, and CCS calculations. The theoretical CCSs for the FGF-dp6 complex do not follow the same trends as the other complexes, and the PA and TM numbers offer a slightly larger range due to the difference between the experimental and theoretically modeled dp6. A comparison of the experimental and theoretical CCSs for the five complexes shows that the experimental numbers lie between the PA and TM numbers, as expected. The experimental CCSs are on an average 12% lower than the TM numbers, which is as expected due to the partial collapse of protein structure as a result of desolvation [81, 82]. The experimental CCSs are on an average 13% higher than the PA numbers, which is in agreement with previous studies [40, 83], which report experimental CCS to be about 15% higher than PA estimates. We have also reported the scaled PA numbers, which showed a close agreement with the experimental numbers. Moreover, we have compared the representative structure of the dimeric protein complex during the MD simulation and after the *in vacuo* minimization (as shown in supplementary data). The protein backbones RMSD between the two structures is 2.07 Å, suggesting that there is no dramatic change of the overall shape of the protein, and the most relevant features of the protein structure are still retained. These results imply that the delsolvation process in the TWIMS experiment did not significantly alter the solvated structure of FGF1 and its complexes, and their biological function to bind with HS was still retained.

In terms of the gas-phase stability of these complexes, it is noteworthy that the experimentally derived CCSs for the five complexes formed with HS of varying lengths lie very

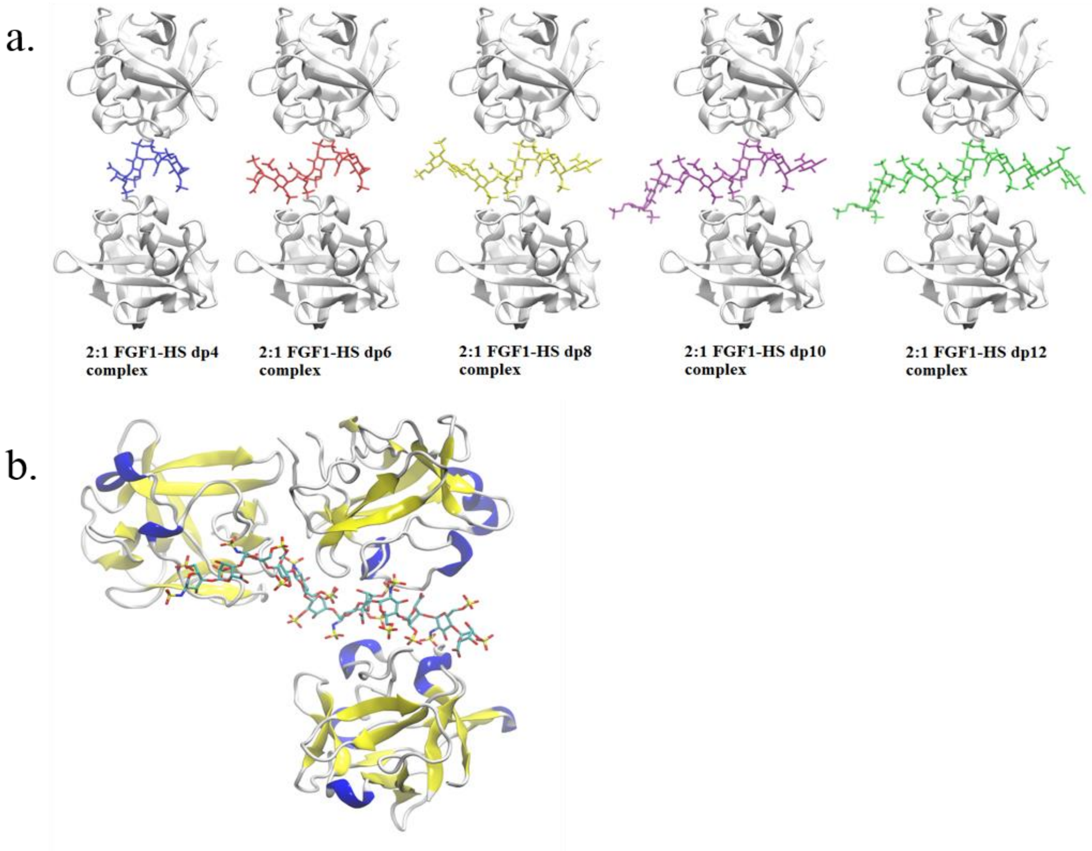


Figure 4.4 (a.) Structural representation of 2:1 FGF1-HS complexes, displaying the binding of dp4-dp12 (left to right). The protein and HS are shown in ribbon and stick representation, respectively. (b.) Structural representation of the hypothesized 3:1 FGF1-HS dodecasaccharide complex, obtained by connecting the available crystal structure for 2:1 FGF1-dp6 dimeric complex with a 1:1 FGF1-dp6 monomeric complex through the HS chain.

close to each other, with the difference between FGF1-dp4 and FGF1-dp12 being 100 Å^2. This difference in the CCSs appears to arise only due to the size of the HS fragment, implying that the compact and folded conformations of the FGF1 complexes were not disturbed, and remained stable during the TWIMS experiments.

A comparison of the CCS of the protein from the crystal structure (PDB id: 2AFG) and the protein after removal of the HS tetrasaccharide from the 1:1 FGF1-dp4 complex was made to determine any conformational changes induced in the protein upon HS binding. The conformational difference in the CCS of the two protein-forms was calculated to be less than 1% using both the TM and PA methods (data in Supplementary Table 2). We can therefore conclude that there is no major conformational change brought about in the protein as a result of HS binding, and the overall increase in the CCS can be attributed to the presence of HS itself, which is in agreement with previous studies [76, 84].

Two different dimerization modes for FGF2 have been previously proposed [17]. In the presence of long-chain HS fragments, FGF2 molecules oligomerize to form side-by-side *cis* dimers, while with short-chain HS disaccharides they form head-to-head *trans* dimers [17]. Similar observations have not been made for FGF1, and the crystal structures available in the PDB for 2:1 FGF1-HS complex only indicate a *trans* binding mode. We generated two models for the plausible *cis* forms of 2:1 FGF1-HS dp12 complexes (Supplementary Figure 1). Cis-model-1 used two FGF1-dp6 molecules joined through the GAG chain, with a *cis*-like starting conformation, to examine whether formation of *cis* form for FGF1 may be feasible, and if it could be differentiated from the *trans* model on the basis of differences in the CCS. However, during the course of MD simulations it adopted a more *trans*-like conformation, likely due to the lack of stable protein-protein interactions between the two FGF1 molecules. In contrast, the Cis-model-2, based on the crystal structure of apo-FGF1 dimer (PDB id: 2AFG) with stable protein-protein interactions, was able to maintain the *cis* conformation during MD simulations. The models differed significantly in their starting and end conformations; however, both formed compact structures that lead to very similar theoretical CCS (data in Supplementary Table 3), making them indistinguishable from each other, and also from the *trans* form (Table 1). On the basis of this data, therefore, it is difficult to determine whether the *cis* form exists in solution, but if it does, it may be difficult to distinguish the *cis* and the *trans* binding conformations solely based on the CCS.

Table 4.1. Experimental and theoretical CCS (Å^2) for FGF1-HS complexes in 2:1 stoichiometry.

		Collision cross section (Å^2)			
ID	HS sequence[a]	Experimental	PA	Scaled PA	TM
FGF1-dp4	IdoA2S-GlcNS6S-IdoA2S-GlcNS6S	2486.1 ± 6.3	2117.2 ± 5.2	2413.6 ± 6.0	2682.2 ± 7.1
FGF1-dp6	GlcNS6S-GlcA-GlcNS6S-IdoA2S-GlcNS6S-GlcA[b]	2520.2 ± 10.2	2350.0 ± 2.5	2679.0 ± 2.8	2990.0 ± 6.6
FGF1-dp8	GlcNS6S-GlcA-[GlcNS6S-IdoA2S]$_2$-GlcNS6S-GlcA	2536.0 ± 8.0	2191.9 ± 2.5	2498.8 ± 2.9	2784.8 ± 6.2
FGF1-dp10	GlcNS6S-GlcA-[GlcNS6S-IdoA2S]$_3$-GlcNS6S-GlcA	2543.9 ± 8.0	2253.6 ± 3.8	2569.2 ± 4.4	2871.9 ± 7.1
FGF1-dp12	GlcNS6S-GlcA-[GlcNS6S-IdoA2S]$_4$-GlcNS6S-GlcA	2581.5 ± 7.9	2271.7 ± 5.3	2589.8 ± 6.0	2885.4 ± 8.8

[a] All but dp4 HS oligosaccharides contained *p*-Nitrophenol linker at the reducing end (dp4 contained OMe instead). Theoretical modeling employed OMe as the terminus for all HS.

[b] For modeling, the dp6 present in the FGF1-HS crystal structure (PDB id: 2AXM) was employed.

The TWIMS experiments also detected a 3:1 FGF1-HS complex in the presence of select sulfated HS dp12. A plausible model, with three molecules of FGF1 in contact with HS-dp12, was computationally examined, using a model that melds the available crystal structure for 2:1 FGF1-dp6 dimeric complex (Figure 4.4 b) with a 1:1 FGF1-dp6 monomeric complex through the

HS chain. For this study, it was ensured that each of the three FGF1 molecules interacts directly with the HS since the formation of this higher order oligomer seems to be induced only in the presence of HS. The MD simulation of the complex was followed by calculation of CCS, and the experimentally calculated CCS was found to lie between the theoretically obtained CCS, using the PA and TJ methods. Specifically, the experimental number was 15.9% higher than the PA and 11% lower than the TJ methods. Although this is only one of the plausible 3:1 FGF1-HS models, the agreement with experimental CCS values provides reasonable confidence in believing that a 3:1 complex would adopt a similar oligomeric conformation.

Gas phase stability of FGF1 and its complexes

The impact of GAG binding upon the stability of the three dimensional structure of FGF1 was examined using collision induced unfolding (CIU) experiments, an approach developed by Ruotolo and co-workers [47]. We performed a series of collision induced activation experiments with incremental changes in trapping collisional energy (CE) applied to selected protein and protein complex ions. Their conformational transition from a fully folded to a fully unfolded state was tracked, and was used to assess the ligand-induced influence on the gas phase stability of these ions. A range of HS oligosaccharides of increasing lengths, containing the critical binding motif of IdoA2S-α-(1-4)-GlcNS6S-α-(1-4)-IdoA2S, were used to form dimeric and monomeric complexes with FGF1. The +8 charge state of monomeric protein complex ions was selected for activation, since this charge state provides higher ion intensity, sufficiently stability against low-energy dissociation, and distinguishable unfolding features within a relatively narrow collision energy window. The resulting CIU fingerprints of FGF1 and FGF1-HS complexes are shown in Figure 4.5, which record the drift time and intensity of each conformation feature that emerged during the unfolding process as the trap CE was ramped from 0V to 60V in a 2.5V increment.

The CIU fingerprint of unbound, apo-FGF1 exhibits three features. The first feature, which has the lowest drift time at 7.9 ms, corresponded to the most compact conformation of the protein ion. As the CE was raised to 10V, this compact conformation diminished while the second feature appeared at 10.2 ms, corresponding to the intermediate state with a more extended conformation. Beyond the CE of 12.5V, the third feature started to emerge and stabilized at 14.3 ms, indicating that the three-dimensional protein structure had unfolded completely.

The CIU fingerprints of 1:1 FGF1-HSdp4/HSdp6 complexes showed similar unfolding pathways as the apo-FGF1, for which the most extended conformation was dominant across the investigated voltage range. However, for both complexes, not only their most compact feature but also their partially unfolded feature showed more elongated shape than those for the apo-FGF1. An increase of 2.5V in the voltage at which the most compact feature starts to unfold was observed. A similar trend was observed for the voltage at which the intermediate feature further extends, shifting the voltages by 5 V and 10 V, for the dp4-bound FGF1 and dp6-bound FGF1, respectively.

The CIU fingerprints of 1:1 FGF1-HSdp8 and HSdp10 complexes show an even more substantial elevation of the voltage at which the most compact feature starts to unfold, shifting the voltage by 5 V and 10 V, respectively. More importantly, the intermediate conformation gained more stability and became dominant for over 66.7% of the investigated voltage range. It is notable that the intermediate state and the extended state stabilize simultaneously for a longer voltage range, indicating that they have a different unfolding pathway from the species discussed above.

In significant contrast to the 1:1 short-chain HS complex bound FGF1 complex, the CIU fingerprint of the 1:1 FGF1-HSdp12 complex exhibited only two principal conformational features, corresponding to the fully folded conformation at 8.7 ms and the intermediate conformation at 11.7 ms, respectively. The survival of the relatively compact conformations at high voltages and the absence of the most extended conformation clearly suggest an even higher resistance towards collision induced unfolding than the apo-FGF1 or short HS bound FGF1 complexes.

The effect of collisional activation was also measured for 2:1 FGF1-HS complexes with HS oligosaccharides of different lengths, and a similar trend as the 1:1 FGF1-HS complexes was observed (Supplementary Figure 2). Comparing the fingerprints clearly shows a significant enhancement of the gas phase stability of FGF1 upon the binding of HS oligosaccharides, regardless of the stoichiometry of the complexes formed. These observations are consistent with the solution-phase behavior of the FGF1 signaling system, which show that HS improves the conformational stability of FGF-FGFR complex against thermal, enzymatic, or pH-dependent inactivation and degradation[85, 86].

CIU fingerprints of the +12 charge state of the 2:1 FGF1-selected sulfated HS dp12s were also compared (Supplementary Figure 3). At low collision energy, their ion mobility profiles all exhibited single, narrow-distribution peaks, indicating that they exist in their compact conformation in the gas phase. As the trap CE was ramped from 0V to 40V, the 2:1 FGF1- fully sulfated dp12 complex showed a slightly higher stability. Nevertheless all four complexes exhibited similar behavior in their unfolding mode, with one compact and one unfolded state with similar centroid drift times and stability ranges. These observations indicate that length of the HS oligomer may be more important that its sulfation pattern in determining the gas phase stability of these protein complexes.

Previous studies proposed that the stability of FGF complexes depends more on the overall degree of sulfation rather than a specific arrangement of the sulfation groups [25]. Out observations suggest a more complicated mechanism than that: four selectively sulfated dp12s display similar conformational stability and gas phase unfolding behavior despite their different overall extent of sulfation; The de 2-*O* and 6-*O* sulfated dp12 with only 6 sulfo groups have a much higher gas-phase stability than the fully sulfated dp4 to dp8 (which have 6 or more sulfo groups). Apparently, both the length of the HS oligosaccharide and the distribution of sulfo groups contribute to stabilizing the gas phase conformation of HS-bound FGF1 complexes.

Since the most compact feature should resemble the native structure of a protein complex, the CE required for its transition to a more extended state should be an experimentally quantitative estimation of the binding affinity of the HS oligosaccharide towards FGF1 [87]. Thus, we measured and plotted the voltage at which the most compact feature starts to unfold for each HS-bound FGF1 complexes, as shown in Figure 4.6. The degree of stabilization is observed to correlate with the length of the HS oligosaccharide, with long-chain HS oligosaccharides interacting with FGF1 and forming complexes that slowed down the initiation and the process of collision induced unfolding more effectively than short-chain HS.

Theoretically, the stability of a complex can be measured as the strength of interaction between the binding partners, and it would be reasonable to believe that the greater the binding affinity, the higher will be the stability, and consequently a higher CE would be required to induce unfolding of the complex. Binding free energies of FGF1 in complex with HS of increasing lengths (dp4-dp12), in both 1:1 and 2:1 binding stoichiometry (Figure 4.4 a) were calculated to provide theoretical quantification. *In vacuo* minimization of a subset of the frames from the MD trajectories was carried out, followed by energetic post processing to obtain the gas-phase interaction energies to mimic the binding modes representative of these complexes during mass spectrometry experiments. We observed that the interaction energies for the monomeric and dimeric complexes both increase as the length of the HS chain increases (Figure

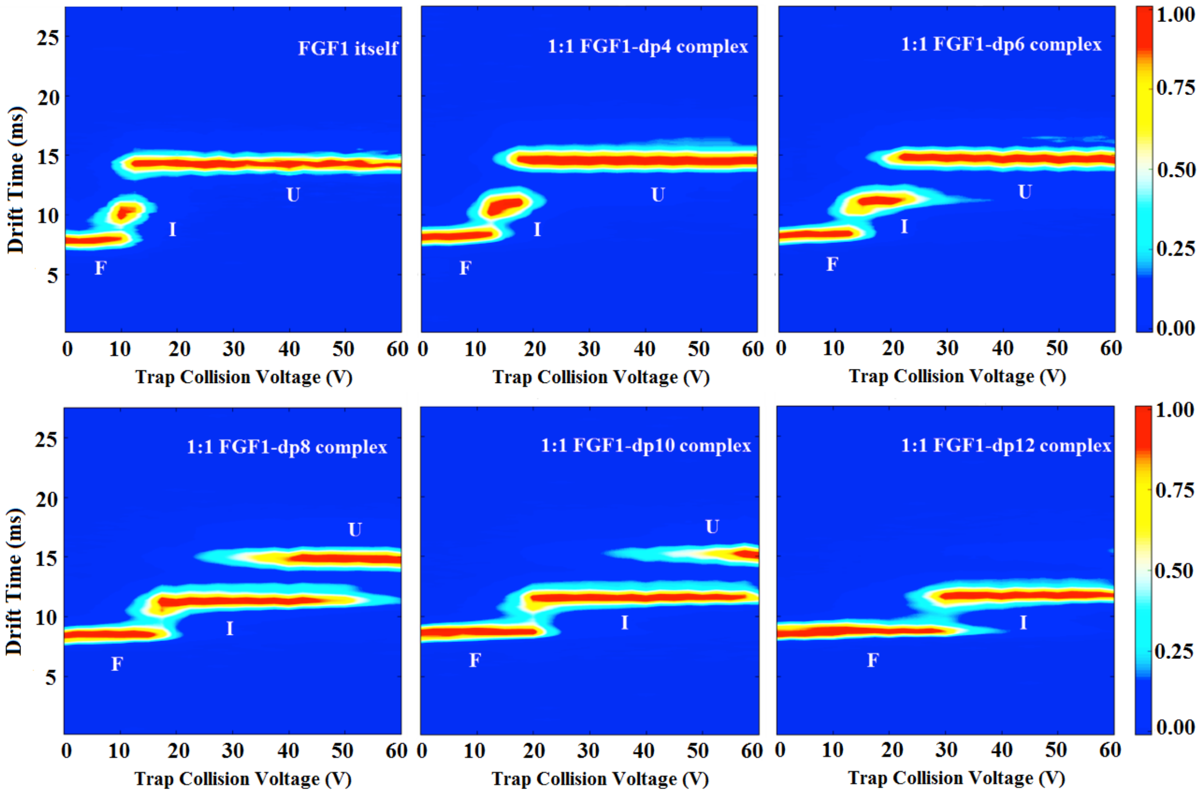


Figure 4.5 The 2D-contour CIU fingerprint of the monomeric apo-FGF1 and the 1:1 FGF1-HS, with HS of increasing length from dp4-dp12. The conformations were labeled as F, I and U, indicating folded, intermediate and unfolded states, respectively. The ion intensities are indicated by a color axis. As the length of the HS increases, they form complexes with FGF1 and increase the conformational stability of the complexes accordingly.

4.6), as expected. These results are in direct agreement with the observed experimental stability trends of the FGF1-HS complexes.

The increase in the stability of the complex with an increase in the length of the HS chain can be attributed to the additional number of saccharide residues which present more sulfo and carboxyl groups to interact with the protein [73]. It was also observed during the MD simulations that the longer HS chains fold onto the proteins, forming more interactions with proteins, further restraining the shape of these structures, and providing them more resistance to collision induced unfolding.

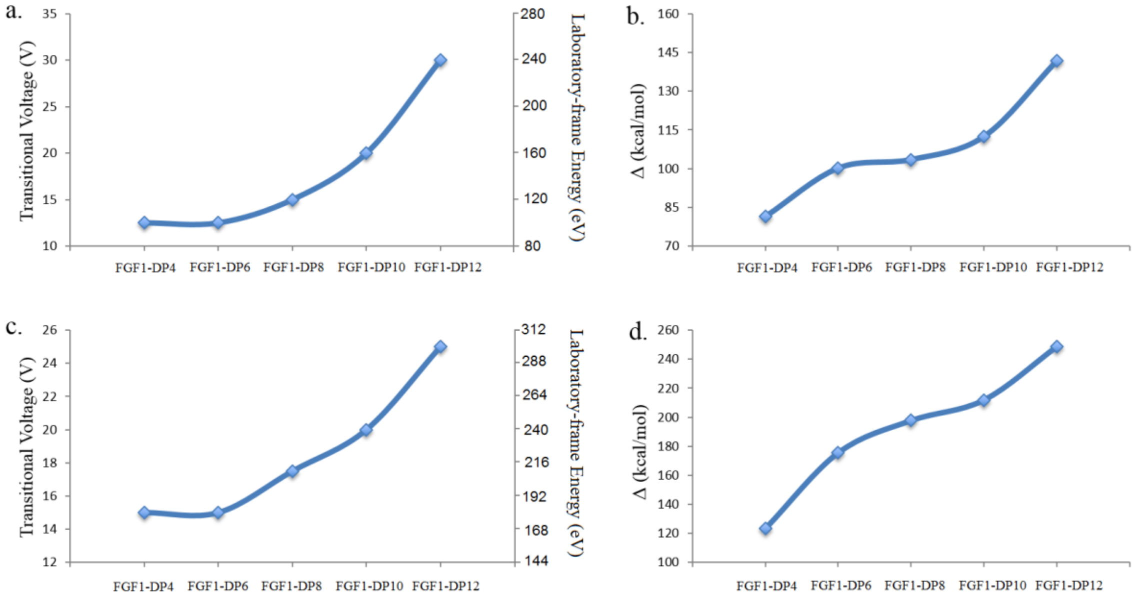


Figure 4.6 Trends for the experimental transitional voltages required for inducing unfolding for the stability of the FGF1-HS complexes in the 1:1 (a) and 2:1 (c) binding stoichiometry as compared to the trends for theoretical binding free energies for the 1:1 (b) and 2:1 (d) complexes. The two trends are comparable, indicating that the binding of longer heparin oligosaccharides imparts greater stability to the protein against unfolding.

These observations demonstrate that while a HS oligosaccharide as short as a tetramer is capable of promoting the formation of a FGF1 dimer complex (2:1 protein :tetramer stoichiometry), its stability is generally low due to its low efficiency to counteract the charge repulsion between the two protein surfaces [18, 19]. This is consistent with the observation that lower concentrations of long-chain HS oligosaccharides display tighter binding and higher biological activity whereas HS dp4 support FGF signaling only when present at a high concentration [67].

Conclusions

Although previous studies of FGF1-FGFR2-HS complexes have been reported using mass spectrometry, the reliance on size-exclusion chromatography to determine that the dominant configuration of the trimeric complex [88], less structural information was extracted about the influence of the variation of HS chain length and sulfation pattern on the FGF1-HS interactions. In the present work, we have applied IMMS to the FGF1-HS signaling complexes, and used molecular modeling to interpret the results. This approach has led to a number of important new conclusions. Although some previous results suggest that HS octasaccharide is the minimal length required for FGF1 dimerization and activation, our results showed that a short tetrasaccharide containing the high affinity binding motif is capable of dimerizing FGF1, forming complexes with a compact, well-defined gas phase conformation. These finding demonstrate the importance of high-affinity binding motif for promoting the correct assembly of FGF1, as well as dictating the binding selectivity of FGF1-HS binding. Moreover, we found that FGF1 can participate in high order oligomers when interacting with long-chain HSs (dodecasaccharides), and the degree of oligomerization increases as the chain length of HS increases. However, the ion mobility profiles of FGF1 in complex with dodecasaccharides with different patterns of sulfate modification did not show much difference in stability, indicating a relative lack of specificity of binding with long-chain HS.

Furthermore, the experiments of collisional activation of FGF1 and FGF1-HS complexes confirmed the stabilizing effect of HS binding on the folded structure of FGF1. The degree of stabilization correlates with the length of the HS oligosaccharide, with longer HS oligosaccharides forming a complex that has higher gas phase stability against collision induced unfolding. The stability of the FGF1-HS complexes correlate with the overall degree of sulfation rather than on the precise location of sulfo groups. Finally, collisional cross sections of FGF1-HS complexes have been examined using theoretical models built using the available X-ray crystal structures of FGF1 in complex with HS. MD simulations with these models are in agreement with the experimental results, indicating that the folded gas-phase conformation of FGF1 and its binding activity survive the transition of the ions into the gas-phase. We also proposed a plausible structure for the 3:1 FGF1-HS complex. The binding affinity of HS of varying lengths has also been studied using molecular modeling, which gave results consistent with experiments.

Collectively, these results extended the application of TWIMS for investigating GAG–protein interactions at the molecular level. Characterization of the interactions between FGF1 and HS will benefit the understanding of the full magnitude FGF signaling pathways and eventually, the development of new medicines. The extension of the methodology will add further dimensions to the drug-development studies of a large array of other GAG-binding proteins.

Acknowledgements

We thank Jessica N. Rabuck, Joseph D. Eschweiler and Dr. Brandon T. Ruotolo for sharing the CIUSuite and for their considerable advice with the application of this software. The authors are grateful for funding from National Institutes of Health grant P41- GM103390 in support of this work.

References

1. Itoh, N., Ornitz, D.M.: Evolution of the Fgf and Fgfr gene families. Trends in Genetics. **20**, 563-569 (2004)

2. Beenken, A., Mohammadi, M.: The FGF family: biology, pathophysiology and therapy. Nature reviews Drug discovery. **8**, 235-253 (2009)

3. Zhu, X., Komiya, H., Chirino, A., Faham, S., Fox, G., Arakawa, T., Hsu, B., Rees, D.: Three-dimensional structures of acidic and basic fibroblast growth factors. Science. **251**, 90-93 (1991)

4. Presta, M., Dell'Era, P., Mitola, S., Moroni, E., Ronca, R., Rusnati, M.: Fibroblast growth factor/fibroblast growth factor receptor system in angiogenesis. Cytokine & growth factor reviews. **16**, 159-178 (2005)

5. Bedussi, F., Bottini, A., Memo, M., Fox, S.B., Sigala, S., Generali, D.: Targeting fibroblast growth factor receptor in breast cancer: a promise or a pitfall? Expert opinion on therapeutic targets. **18**, 665-678 (2014)

6. Böhlen, P., Esch, F., Baird, A., Gospodarowicz, D.: Acidic fibroblast growth factor (FGF) from bovine brain: amino-terminal sequence and comparison with basic FGF. The EMBO journal. **4**, 1951 (1985)

7. Laestander, C., Engström, W.: Role of fibroblast growth factors in elicitation of cell responses. Cell proliferation. **47**, 3-11 (2014)

8. Eswarakumar, V., Lax, I., Schlessinger, J.: Cellular signaling by fibroblast growth factor receptors. Cytokine & growth factor reviews. **16**, 139-149 (2005)

9. Hatch, N.E.: FGF signaling in craniofacial biological control and pathological craniofacial development. Critical Reviews™ in Eukaryotic Gene Expression. **20**, 295-311 (2010)

10. Ornitz, D.M., Itoh, N.: Fibroblast growth factors. Genome Biol. **2**, 1-12 (2001)

11. Esko, J.D., Lindahl, U.: Molecular diversity of heparan sulfate. The Journal of clinical investigation. **108**, 169-173 (2001)

12. Rabenstein, D.L.: Heparin and heparan sulfate: structure and function. Natural product reports. **19**, 312-331 (2002)

13. Garg, H.G., Linhardt, R.J., Hales, C.A.: Chemistry and biology of heparin and heparan sulfate. Elsevier, Oxford, UK, (2011)

14. Shimokawa, K., Kimura-Yoshida, C., Nagai, N., Mukai, K., Matsubara, K., Watanabe, H., Matsuda, Y., Mochida, K., Matsuo, I.: Cell surface heparan sulfate chains regulate local reception of FGF signaling in the mouse embryo. Developmental cell. **21**, 257-272 (2011)

15. Mohammadi, M., Olsen, S.K., Ibrahimi, O.A.: Structural basis for fibroblast growth factor receptor activation. Cytokine & growth factor reviews. **16**, 107-137 (2005)

16. Mohammadi, M., Dikic, I., Sorokin, A., Burgess, W., Jaye, M., Schlessinger, J.: Identification of six novel autophosphorylation sites on fibroblast growth factor receptor 1 and elucidation of their importance in receptor activation and signal transduction. Molecular and Cellular Biology. **16**, 977-989 (1996)

17. Herr, A.B., Ornitz, D.M., Sasisekharan, R., Venkataraman, G., Waksman, G.: Heparin-induced Self-association of Fibroblast Growth Factor-2 EVIDENCE FOR TWO OLIGOMERIZATION PROCESSES. Journal of Biological Chemistry. **272**, 16382-16389 (1997)

18. DiGabriele, A.D., Lax, I., Chen, D.I., Svahn, C.M., Jaye, M., Schlessinger, J., Hendrickson, W.A.: Structure of a heparin-linked biologically active dimer of fibroblast growth factor. Nature. **393**, 812-817 (1998)

19. Waksman, G., Herr, A.B.: New insights into heparin-induced FGF oligomerization. Nature structural biology. **5**, 527-530 (1998)

20. Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B., Blundell, T.L.: Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. Nature. **407**, 1029-1034 (2000)

21. Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A.V., Yeh, B.K., Yayon, A., Linhardt, R.J., Mohammadi, M.: Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. Molecular cell. **6**, 743-750 (2000)

22. Petitou, M., Casu, B., Lindahl, U.: 1976–1983, a critical period in the history of heparin: the discoveryof the antithrombin binding site. Biochimie. **85**, 83-89 (2003)

23. Plotnikov, A.N., Schlessinger, J., Hubbard, S.R., Mohammadi, M.: Structural basis for FGF receptor dimerization and activation. Cell. **98**, 641-650 (1999)

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

25. Kreuger, J., Spillmann, D., Li, J.-p., Lindahl, U.: Interactions between heparan sulfate and proteins: the concept of specificity. The Journal of cell biology. **174**, 323-327 (2006)

26. Kreuger, J., Jemth, P., Sanders-Lindberg, E., Eliahu, L., Dina, R., Basilico, C., Salmivirta, M., Lindahl, U.: Fibroblast growth factors share binding sites in heparan sulphate. Biochemical Journal. **389**, 145-150 (2005)

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

28. Guimond, S.E., Turnbull, J.E.: Fibroblast growth factor receptor signalling is dictated by specific heparan sulphate saccharides. Current Biology. **9**, 1343-1346 (1999)

29. Ashikari-Hada, S., Habuchi, H., Kariya, Y., Itoh, N., Reddi, A.H., Kimata, K.: Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. Journal of Biological Chemistry. **279**, 12346-12354 (2004)

30. Matsuo, I., Kimura-Yoshida, C., Shimokawa, K.: Divergent Roles of Heparan Sulfate in Regulation of FGF Signaling During Mammalian Embryogenesis. In: Kondoh H, Kuroiwa A (eds.). Springer Japan, Tokyo, (2014)

31. Zhang, F., Zhang, Z., Lin, X., Beenken, A., Eliseenkova, A.V., Mohammadi, M., Linhardt, R.J.: Compositional analysis of heparin/heparan sulfate interacting with fibroblast growth factor· fibroblast growth factor receptor complexes. Biochemistry. **48**, 8379-8386 (2009)

32. Ostrovsky, O., Berman, B., Gallagher, J., Mulloy, B., Fernig, D.G., Delehedde, M., Ron, D.: Differential effects of heparin saccharides on the formation of specific fibroblast growth factor (FGF) and FGF receptor complexes. Journal of Biological Chemistry. **277**, 2444-2453 (2002)

33. Moy, F.J., Seddon, A.P., Campbell, E.B., Böhlen, P., Powers, R.: 1H, 15N, 13C and 13CO assignments and secondary structure determination of basic fibroblast growth factor using 3D heteronuclear NMR spectroscopy. Journal of biomolecular NMR. **6**, 245-254 (1995)

34. Robinson, C.J., Harmer, N.J., Goodger, S.J., Blundell, T.L., Gallagher, J.T.: Cooperative dimerization of fibroblast growth factor 1 (FGF1) upon a single heparin saccharide may drive the formation of 2: 2: 1 FGF1· FGFR2c· heparin ternary complexes. Journal of Biological Chemistry. **280**, 42274-42282 (2005)

35. Naimy, H., Buczek-Thomas, J.A., Nugent, M.A., Leymarie, N., Zaia, J.: Highly sulfated nonreducing end-derived heparan sulfate domains bind fibroblast growth factor-2 with high affinity and are enriched in biologically active fractions. Journal of Biological Chemistry. **286**, 19311-19319 (2011)

36. Spivak-Kroizman, T., Lemmon, M., Dikic, I., Ladbury, J., Pinchasi, D., Huang, J., Jaye, M., Crumley, G., Schlessinger, J., Lax, I.: Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation. Cell. **79**, 1015-1024 (1994)

37. Ibrahimi, O.A., Zhang, F., Eliseenkova, A.V., Linhardt, R.J., Mohammadi, M.: Proline to arginine mutations in FGF receptors 1 and 3 result in Pfeiffer and Muenke craniosynostosis syndromes through enhancement of FGF binding affinity. Human molecular genetics. **13**, 69-78 (2004)

38. Lee, M.K., Lander, A.D.: Analysis of affinity and structural selectivity in the binding of proteins to glycosaminoglycans: development of a sensitive electrophoretic approach. Proceedings of the National Academy of Sciences. **88**, 2768-2772 (1991)

39. Atmanene, C.d., Petiot-Bécard, S.p., Zeyer, D., Van Dorsselaer, A., Vivat Hannah, V.r., Sanglier-Cianférani, S.: Exploring key parameters to detect subtle ligand-induced protein conformational changes using traveling wave ion mobility mass spectrometry. Analytical chemistry. **84**, 4703-4710 (2012)

40. Zhao, Y., Singh, A., Li, L., Linhardt, R.J., Xu, Y., Liu, J., Woods, R.J., Amster, I.J.: Investigating changes in the gas-phase conformation of Antithrombin III upon binding of Arixtra using traveling wave ion mobility spectrometry (TWIMS). Analyst. **140**, 6980-6989 (2015)

41. Hyung, S.-J., Robinson, C.V., Ruotolo, B.T.: Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes. Chemistry & biology. **16**, 382-390 (2009)

42. Duijn, E.v., Barendregt, A., Synowsky, S., Versluis, C., Heck, A.J.: Chaperonin complexes monitored by ion mobility mass spectrometry. Journal of the American Chemical Society. **131**, 1452-1459 (2009)

43. Politis, A., Park, A.Y., Hyung, S.-J., Barsky, D., Ruotolo, B.T., Robinson, C.V.: Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes. PloS one. **5**, e12080 (2010)

44. Uetrecht, C., Rose, R.J., van Duijn, E., Lorenzen, K., Heck, A.J.: Ion mobility mass spectrometry of proteins and protein assemblies. Chemical Society Reviews. **39**, 1633-1655 (2010)

45. Xu, Y., Cai, C., Chandarajoti, K., Hsieh, P.-H., Li, L., Pham, T.Q., Sparkenbaugh, E.M., Sheng, J., Key, N.S., Pawlinski, R.: Homogeneous low-molecular-weight heparins with reversible anticoagulant activity. Nature chemical biology. **10**, 248 (2014)

46. Zong, C., Venot, A., Dhamale, O., Boons, G.-J.: Fluorous supported modular synthesis of heparan sulfate oligosaccharides. Organic letters. **15**, 342-345 (2013)

47. Eschweiler, J.D., Rabuck-Gibbons, J.N., Tian, Y., Ruotolo, B.T.: CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-phase Protein Ions. Analytical chemistry. **87**, 11516-11522 (2015)

48. Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C.V., Ruotolo, B.T.: Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. Analytical Chemistry. **82**, 9557-9565 (2010)

49. Smith, D.P., Knapman, T.W., Campuzano, I., Malham, R.W., Berryman, J.T., Radforda, S.E., Ashcrofta, A.E.: Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. cell [Equation (1)]. **12**, 13 (2009)

50. Ruotolo, B.T., Benesch, J.L., Sandercock, A.M., Hyung, S.-J., Robinson, C.V.: Ion mobility–mass spectrometry analysis of large protein complexes. Nature protocols. **3**, 1139-1152 (2008)

51. Blaber, M., DiSalvo, J., Thomas, K.A.: X-ray crystal structure of human acidic fibroblast growth factor. Biochemistry. **35**, 2086-2094 (1996)

52. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C.: Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Structure, Function, and Bioinformatics. **65**, 712-725 (2006)

53. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L., Woods, R.J.: GLYCAM06: a generalizable biomolecular force field. Carbohydrates. Journal of computational chemistry. **29**, 622-655 (2008)

54. Singh, A., Tessier, M.B., Pederson, K., Wang, X., Venot, A.P., Boons, G.-J., Prestegard, J.H., Woods, R.J.: Extension and Validation of the GLYCAM Force Field Parameters for modeling Glycosaminoglycans. Canadian Journal of Chemistry. **94**, 1-9 (2016)

55. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L.: Comparison of simple potential functions for simulating liquid water. The Journal of chemical physics. **79**, 926-935 (1983)

56. Götz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., Walker, R.C.: Routine microsecond molecular dynamics
simulations with AMBER on GPUs. 1. Generalized born. Journal of chemical theory and computation. **8**, 1542-1555 (2012)

57. D.A. Case, V.B., J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman: AMBER14. (2014)

58. Berendsen, H.J., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.: Molecular dynamics with coupling to an external bath. The Journal of chemical physics. **81**, 3684-3690 (1984)

59. Ryckaert, J.-P., Ciccotti, G., Berendsen, H.J.: Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. Journal of Computational Physics. **23**, 327-341 (1977)

60. Roe, D.R., Cheatham III, T.E.: PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. Journal of chemical theory and computation. **9**, 3084-3095 (2013)

61. Humphrey, W., Dalke, A., Schulten, K.: VMD: visual molecular dynamics. Journal of molecular graphics. **14**, 33-38 (1996)

62. Miller III, B.R., McGee Jr, T.D., Swails, J.M., Homeyer, N., Gohlke, H., Roitberg, A.E.: MMPBSA. py: an efficient program for end-state free energy calculations. Journal of Chemical Theory and Computation. **8**, 3314-3321 (2012)

63. Mesleh, M., Hunter, J., Shvartsburg, A., Schatz, G.C., Jarrold, M.: Structural information from ion mobility measurements: effects of the long-range potential. The Journal of Physical Chemistry. **100**, 16082-16086 (1996)

64. Venkataraman, G., Shriver, Z., Davis, J.C., Sasisekharan, R.: Fibroblast growth factors 1 and 2 are distinct in oligomerization in the presence of heparin-like glycosaminoglycans. Proceedings of the National Academy of Sciences. **96**, 1892-1897 (1999)

65. Chowdhury, S.K., Katta, V., Chait, B.T.: Probing conformational changes in proteins by mass spectrometry. Journal of the American Chemical Society. **112**, 9012-9013 (1990)

66. Gambarini, A.G., Miyamoto, C.A., Lima, G.A., Nader, H.B., Dietrich, C.P.: Mitogenic activity of acidic fibroblast growth factor is enhanced by highly sulfated oligosaccharides derived from heparin and heparan sulfate. Molecular and cellular biochemistry. **124**, 121-129 (1993)

67. Delehedde, M., Malcolm, L., Gallagher, J.T., Rudland, P.S., Fernig, D.G.: Fibroblast growth factor-2 binds to small heparin-derived oligosaccharides and stimulates a sustained phosphorylation of p42/44 mitogen-activated protein kinase and proliferation of rat mammary fibroblasts. Biochemical Journal. **366**, 235-244 (2002)

68. Wu, Z.L., Zhang, L., Yabe, T., Kuberan, B., Beeler, D.L., Love, A., Rosenberg, R.D.: The involvement of heparan sulfate (HS) in FGF1/HS/FGFR1 signaling complex. Journal of Biological Chemistry. **278**, 17121-17129 (2003)

69. Mach, H., Volkin, D.B., Burke, C.J., Middaugh, C.R., Linhardt, R.J., Fromm, J.R., Loganathan, D., Mattsson, L.: Nature of the interaction of heparin with acidic fibroblast growth factor. Biochemistry. **32**, 5480-5489 (1993)

70. Ornitz, D.M., Herr, A.B., Nilsson, M., Westman, J., Svahn, C.-M., Waksman, G.: FGF binding and FGF receptor activation by synthetic heparan-derived di-and trisaccharides. Science. **268**, 432-436 (1995)

71. Guerrini, M., Agulles, T., Bisio, A., Hricovini, M., Lay, L., Naggi, A., Poletti, L., Sturiale, L., Torri, G., Casu, B.: Minimal heparin/heparan sulfate sequences for binding to fibroblast growth factor-1. Biochemical and biophysical research communications. **292**, 222-230 (2002)

72. Bernstein, S.L., Wyttenbach, T., Baumketner, A., Shea, J.-E., Bitan, G., Teplow, D.B., Bowers, M.T.: Amyloid β-protein: monomer structure and early aggregation states of Aβ42 and its Pro19 alloform. Journal of the American Chemical Society. **127**, 2075-2084 (2005)

73. Brown, A., Robinson, C.J., Gallagher, J.T., Blundell, T.L.: Cooperative heparin-mediated oligomerization of fibroblast growth factor-1 (FGF1) precedes recruitment of FGFR2 to ternary complexes. Biophysical journal. **104**, 1720-1730 (2013)

74. Faham, S., Linhardt, R.J., Rees, D.C.: Diversity does make a difference: fibroblast growth factor-heparin interactions. Current opinion in structural biology. **8**, 578-586 (1998)

75. Seo, Y., Andaya, A., Bleiholder, C., Leary, J.A.: Differentiation of CC vs CXC chemokine dimers with GAG octasaccharide binding partners: an ion mobility mass spectrometry approach. Journal of the American Chemical Society. **135**, 4325-4332 (2013)

76. Faham, S., Hileman, R., Fromm, J., Linhardt, R., Rees, D.: Heparin structure and interactions with basic fibroblast growth factor. Science. **271**, 1116-1120 (1996)

77. Mack Jr, E.: Average cross-sectional areas of molecules by gaseous diffusion methods. Journal of the American Chemical Society. **47**, 2468-2482 (1925)

78. Wyttenbach, T., von Helden, G., Batka, J.J., Carlat, D., Bowers, M.T.: Effect of the long-range potential on ion mobility measurements. Journal of the American Society for Mass Spectrometry. **8**, 275-282 (1997)

79. Lanucara, F., Holman, S.W., Gray, C.J., Eyers, C.E.: The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nat. Chem. **6**, 281-294 (2014)

80. Benesch, J.L., Ruotolo, B.T.: Mass spectrometry: come of age for structural and dynamical biology. Current opinion in structural biology. **21**, 641-649 (2011)

81. Jurneczko, E., Barran, P.E.: How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst. **136**, 20-28 (2011)

82. Breuker, K., McLafferty, F.W.: Stepwise evolution of protein native structure with electrospray into the gas phase, 10− 12 to 102 s. Proceedings of the National Academy of Sciences. **105**, 18145-18152 (2008)

83. Hall, Z., Politis, A., Bush, M.F., Smith, L.J., Robinson, C.V.: Charge-state dependent compaction and dissociation of protein complexes: insights from ion mobility and molecular dynamics. Journal of the American Chemical Society. **134**, 3429-3438 (2012)

84. Zhu, X., Hsu, B.T., Rees, D.C.: Structural studies of the binding of the anti-ulcer drug sucrose octasulfate to acidic fibroblast growth factor. Structure. **1**, 27-34 (1993)

85. Sommer, A., Rifkin, D.B.: Interaction of heparin with human basic fibroblast growth factor: protection of the angiogenic protein from proteolytic degradation by a glycosaminoglycan. Journal of cellular physiology. **138**, 215-220 (1989)

86. Zhoulli, F.-Y., Kanb, M., Owensa, R.T., McKeehanl, W.L., Thompson, J.A., Linhardtd, R.J., Hook, M.: Heparin-dependent fibroblast growth factor activities: effects of defined heparin oligosaccharides. European journal of cell biology. **73**, 71-78D (1997)

87. Niu, S., Rabuck, J.N., Ruotolo, B.T.: Ion mobility-mass spectrometry of intact protein–ligand complexes for pharmaceutical drug discovery and development. Current opinion in chemical biology. **17**, 809-817 (2013)

88. Harmer, N.J., Ilag, L.L., Mulloy, B., Pellegrini, L., Robinson, C.V., Blundell, T.L.: Towards a resolution of the stoichiometry of the fibroblast growth factor (FGF)–FGF receptor–heparin complex. Journal of molecular biology. **339**, 821-834 (2004)

CHAPTER 5

A TRAVELING WAVE ION MOBILITY SPECTROMETRY AND MOLECULAR MODELING STUDY OF THE HEPARAN SULFATE-ROBO1 INTERACTION[3]

---

[3] Zhao Y, Yang Y, Thieker DF, Xu Y, Zong C, Boons GJ, Liu J, Woods RJ, Moremen KW, Amster IJ. Submitted to *Journal of the American Society for Mass Spectrometry*, 12/02/17.

Abstract

Roundabout 1 (Robo1) interacts with its receptor Slit to regulate axon guidance, axon branching and dendritic development in the nervous system and to regulate morphogenesis and many cell functions in the nonneuronal tissues. This interaction is known to be critically regulated by heparan sulfate (HS). Previous studies suggest that HS is required to promote the binding of Robo1 to Slit to form the minimal signaling complex, but the molecular details and the structural requirements of HS for this interaction are still unclear. Here we describe the application of traveling wave ion mobility spectrometry (TWIMS) to study the conformational details of the Robo1-HS interaction. The results suggest that Robo1 exists in two conformations that differ by their compactness and capability to interact with HS. The results also suggest that the highly flexible interdomain hinge region connecting the Ig1 and Ig2 domains of Robo1 plays an important functional role in promoting the Robo1-Slit interaction. Moreover, variations in the sulfation pattern and size of HS were found to affect its binding affinity and selectivity to interact with different conformations of Robo1. Both MS measurements and CIU experiments show that the Robo1-HS interaction requires the presence of a specific size and pattern of modification of HS. Furthermore, the effect of N-glycosylation on the conformation of Robo1 and its binding modes with HS is reported.

Introduction

Robo-Slit signaling is one of the best studied axon guidance pathways and their interaction is known to be critically modulated by heparan sulfate (HS). For instance, during the assembly of the central nervous system (CNS), commissural axon migration at the CNS midline is regulated by the interactions between axon guidance molecules and their receptors, such as the interaction between attractive cues Netrin and its receptor Deleted in colorectal carcinoma, which stimulates pre-crossing of axons to the contralateral side [1]. Interaction between repelling cues Slit and its cellular roundabout receptor (Robo) impedes aberrant post-crossing of ipsilateral axons [2, 3]. Though best known for its roles in neuron development and axon pathfinding, Robo-Slit signaling is not confined to the nervous tissue and has been implicated in multiple nonneuronal tissues, contributing to a variety of development and physiological processes through mediating cell proliferation [4], cell adhesion [5], cell migration and mobility [6, 7], angiogenesis [8] as well as dendritic development and branching [9]. Robo-Slit signaling has also been associated with progression [10, 11], metastasis [12] and angiogenesis of tumor [13]. Deletion, epigenetic modification and abnormal expression of Robo and Slit genes occurred in different types and different stages of cancers [14, 15], making them attractive candidates for cancer diagnostics and therapeutics.

Robos are a family of transmembrane proteins that are mainly expressed in CNS neurons. Sharing structural motifs with the Ig superfamily of cell adhesion molecules (CAM) [16], Robos are composed of five Ig-like domains and three fibronectin type III motifs (FN3) in its extracellular region (except for Robo4), followed by a transmembrane region and an intracellular cytoplasmic domain. Vertebrates have four types of Robo (Robo1-4). Though they have distinct cytoplasmic domains, their Ig1 and Ig2 domains are highly conserved and are believed to be responsible for mediating the Robo-Slit interaction [17].

Two crystal structures are available for the two N-terminal Ig domains of human Robo1 [18]. In the first crystal structure, Ig1 and Ig2 are aligned in a linear format, with the final strand of Ig1 extending and forming a 100 Å long interdomain region to connect the two domains. Compared with the first crystal form, the Ig2 of the second crystal form is rotated by 39.9° at the hinge region, resulting in a slightly bent arrangement. Two similar crystal structures have been observed for the first two Ig domains of Drosophila Robo1 [19]. One tetragonal crystal form and one monoclinic crystal form were observed, corresponding to the slightly bent and linear crystal structures for human Robo1, respectively. Both of the previous studies on human and Drosophila Robo1 observed substantial flexibility embedded in the interdomain hinge region and raised questions about whether this flexibility is related to the biological function of Robo1.

Both Robo1 and Slit are HS binding proteins. HSs are linear, highly sulfated glycosaminoglycan (GAG) polysaccharides. The backbone structure of HS is a product of polymerization of disaccharide units containing glucosamine (GlcN) and uronic acid residues (glucuronic acid (GlcA) or iduronic acid (IdoA)) linked through a 1-4 glycosidic linkage. Modification of HS backbone including C5 epimerization of uronic acid, deacetylation of glucosamine and addition of various types of N- or O- sulfations generate enormous sequence heterogeneity. HSs interact with many proteins and play important roles in a variety of physiological and pathological processes by regulating the structure, conformation, distribution, stability and activity of these proteins [20, 21].

Substantial in vivo and in vitro studies suggest that HS plays an indispensable role in Robo-Slit mediated axon guidance: null mutations in C. elegans enzymes that are responsible for HS modifications including sulfotransferases and epimerases caused defected axonal and cellular guidance [22]; mutations in Drosophila syndecan altered the Slit distribution in developing axons [23] and interfered with the slit signal transmission in target cells [24]; enzymatic digestion of endogenous HS using heparinase III destabilized Robo1-Slit interaction and abolished the repulsive activity of Slit [25]. Collectively, HS has dual roles in Robo1-Slit signaling: the main role of HS is to act as co-receptors of Slit, simultaneously interacting with the LRR2 domain of Slit and Ig1-2 domains of Robo1, enhancing and stabilizing the Robo1-Slit interaction by increasing their binding affinity by 10-fold and promoting the formation of trimeric Robo1-Slit-HS signaling complexes [26, 27]. An additional function of HS is to mediate the distribution of Slit through mechanism of either active transport, buffered diffusion or selective stabilization [23], therefore defining the concentration and accessibility of Slit on the extracellular surface [28]. However, more experimental evidence is required to determine whether the HS-dependence of Robo-Slit interaction is required in all cell types and tissue types in mammals [29].

Though the significance of HS in Robo-Slit interaction has been well established, a clear elucidation of all the molecular details of this interaction is still missing. Previous mutagenesis and X-ray crystallography experiments identified the binding site of HS as a contiguous basic patch spanning both the surfaces of Robo1 and Slit, but the structural requirements of the bound HS have not been well defined. Also, more experiments are required to elucidate the molecular mechanism of how exactly the interaction between Robo and Slit is programmed into signal transduction through interacting with HS.

Previously we developed and applied travelling wave ion mobility spectrometry (TWIMS) to examine the interaction between antithrombin III (ATIII) and heparin [30] as well as the interaction between fibroblast growth factor 1 (FGF1) and HS [31]. TWIMS is capable of providing conformational and stoichiometric details that are consistent with the known behaviors of these two well-characterized interactions. We also observed that the IM profiles of GAG-protein complexes provide useful data regarding the binding specificity of their interactions. In this current study, TWIMS is used to investigate the less well characterized Robo1-HS interaction. Only the first two Ig domains were used as a representative of intact Robo1 since they are the most conserved regions of Robo1 and participate in the binding of Robo1 and Slit with HS [32, 33]. The results provide new insights into the functional significance of the conformational flexibility of the Ig1-2 domains of Robo1, and how HS implements its biological function through selectively recognizing and interacting with different conformations of the Ig1-2 domains of Robo1. The results also provide an in-depth understanding of the way HS modulates its functional interaction with Robo1 through its fine structure. Moreover, a new mechanism for the activation of Robo1 by HS and the promotion of the Robo1-HS interaction is proposed.

Experimental

Reagents

All chemicals and solvents (ammonium acetate, methanol, water, and formic acid) were of HPLC grade and purchased from Sigma-Alldrich, (St. Louis, MO). Protein calibrants (myoglobin from equine heart, cytochrome c from equine heart, avidin from egg white, concanavalin A from *Canavalia ensiformis*, and bovine serum albumin) were purchased from Sigma-Aldrich. HS hexasaccharides (dp6) to dodecasaccharides (dp12) were chemo-enzymatically synthesized as previously described [32]. HS tetrasaccharide (dp4) was chemically synthesized as previously described by fluorous supported modular synthesis [33].

Protein expression

An expression vector encoding the human roundabout homolog 1 precursor (Robo1, NP_002932, Uniprot Q9Y6N7) Ig domains 1 and 2 (residues 58–266) in the pGEn2 vector was previously described [34]. The vector encodes a fusion protein with an NH2-terminal signal sequence followed by an 8xHis tag, AviTag [35], “super-folder” GFP [36], the TEV protease recognition site [37], and the Ig1-2 domains of Robo1. Expression of the construct in mammalian cells leads to translocation of the fusion protein into the secretory pathway, glycosylation, and secretion as a GFP-Robo1 Ig1-2 fusion into the conditioned media.

A separate construct encoding leukocyte common antigen-related receptor protein tyrosine phosphatases (LAR-RPTP, NM_002840, Uniprot P10586) NH2-terminal signal sequence and Ig domains 1 and 2 (residues 1-231) in the pGEc2 vector. This construct encodes a fusion protein containing the LAR signal sequence and Ig1-2 domains followed by the TEV protease recognition site, "super-folder" GFP, AviTag, and 8xHis tag. Expression of the construct in mammalian cells leads to translocation of fusion protein into the secretory pathway and secretion as a Robo1 Ig1-2-GFP fusion into the conditioned media. Mass spectrometry data indicated that the single Asn117-Thr-Ser119 N-glycosylation sequon in the LAR coding region is not occupied with an N-glycan structure.

Both fusion protein expression constructs were transiently transfected into mammalian cells using either FreeStyleTM 293-F cells (Thermo Fisher Scientific) for protein production containing wild type glycans (complex-type glycans) or HEK293S (GnTI-) cells [38] (ATCC) for expression with high mannose glycan structures (predominately Man5GlcNAc2-Asn glycans). Both cell lines were maintained in culture media comprised of 9 volumes FreeStyleTM 293 expression media (Thermo Fisher Scientific) and 1 volume EX-CELL® 293 serum-free medium (Sigma) (9:1 media) as previously described. Suspension cultures (1 L) of either HEK293S (GnTI-) cells or FreeStyleTM 293-F cells were transfected with the respective plasmid DNAs using polyethyleneimine (linear 25 kDa PEI, Polysciences) as transfection reagent as previously described [39]. The cultures were diluted 1:1 with culture medium containing 4.4 mM valproic acid (2.2 mM final concentration) 24 h after transfection, and protein production was continued for a further 5 days at 37°C.

Protein purification

Protein purification, deglycosylation, and tag removal employed workflows similar to prior structural studies on rat ST6GAL1 [39]. Briefly, the conditioned culture medium was loaded on a Ni^{2+}-NTA Superflow (Qiagen) column equilibrated with 20 mM HEPES, 300 mM NaCl, 20 mM imidazole, pH 7.4, washed with column buffer, and eluted successively with column buffers containing stepwise increasing imidazole concentrations (40-300 mM). The eluted fusion protein was pooled, concentrated, and mixed with recombinant tobacco etch virus (TEV) protease (1:10 ratio relative to GFP-Robo1 or LAR-GFP) and incubated at 4°C for 24 h. The samples were diluted 15-fold in elution buffer without imidazole and re-applied to the Ni^{2+}-NTA column to allow collection of the Robo1 or LAR products in the flow-through fractions. To produce Robo1 with minimal glycosylation, the purified protein expressed in HEK293S(GnTI-) cells was treated with endoglycosidase F1 (EndoF1), which cleaves the glycan structure between the two core GlcNAc residues and truncates glycans to a single GlcNAc residue [39]. EndoF1 was added concurrently with TEV protease at a 1:10 ratio relative to the GFP-Robo1 and incubated at 4°C for 24 h. Dilution to lower the imidazole concentration was followed by passing through a Ni^{2+}-NTA column to remove the fusion tag and His-tagged TEV protease and EndoF1 to allow collection of the Robo1 product in the flow-through fractions. Each of the protein forms (Robo1 containing wild type glycans, Man5GlcNAc2-Asn glycans, or a single GlcNAc, or LAR containing no glycans) were further purified on a Superdex 75 gel filtration column (GE Healthcare) and peak fractions were collected and concentrated by ultrafiltration to 30 mg/ml for mass spectrometry analysis.

Sample preparation

For MS analyses under nondenaturing conditions, Robo1 was diluted in 20 mM ammonium acetate buffer, pH 6.8, to a final concentration of 20 μM. Robo1-HS complex was obtained by incubating Robo1 with HS oligosaccharides at room temperature for 60 min. Protein calibrants were diluted in either denaturing solution or nondenaturing solution to a final concentration of 10 μM.

IMMS measurement and data analysis

NanoESI-IMMS experiments were performed using a quadrupole-TWIMS-TOF hybrid mass spectrometer (Synapt G2 HDMS; Waters Corp., Manchester, UK) in positive ionization mode. Protein samples were injected into the nanoESI source through a fused-silica emitter (PicoTip; New Objective, Woburn, MA, USA) with a flow rate varying from 0.2 to 0.5 uL/min. The applied experimental parameters were capillary voltage, 1.4 kV; sampling cone voltage, 20 V; extraction cone voltage, 5 V; source temperature, 30 °C; flow rate of nitrogen in the IM ion guide, 50 mL/min; flow rate of helium in the helium cell, 180 mL/min; transfer collision energy, 0 V. Different sets of wave height and corresponding wave velocity were examined to optimize the mobility separation. The drift times of the calibrants and Robo1 samples were measured and identical experimental conditions were stringently applied.

For monitoring the collisional induced unfolding (CIU) of Robo1 and its complexes, protein ions of a selected charge state were isolated in a quadrupole mass filter and activated in the trap ion guide where subsequent collisional induced activation of ions takes place, followed by separation in the TWIMS ion mobility stage, and detection by TOF-MS.

Data analysis was performed using MassLynx 4.1 and Driftscope (Waters Corp., Manchester, UK). The CIU data was analyzed using CIUSuite [40]. CIU fingerprints of protein ions, which records the relative ion intensity (normalized and smoothed using Savitsky-Golay filter) as a function of collision energy and drift times, were shown in a 2-D contour plot using the CIUSuite plot function. The CIUsuite detect function was applied to identify and extract the most intense CIU features of the fingerprint based on a first derivative analysis, providing centroid drift times, transitional voltage, and stability range of each detected feature.

Converting drift times into collision cross sections (CCSs)

The CCSs of Robo1 and Robo1-HS complexes were calibrated based on an empirical relationship between the drift times of protein calibrants and their known CCSs obtained previously by DTIMS. Briefly, a selected set of native and denatured protein calibrants, with a mass range from 12 to 102 KDa and a CCS range from 2303 to 5550 Å^2 were employed. The drift times of these calibrates were corrected for mass-dependent flight time spent in the transfer ion guide and TOF mass analyzer and mass-independent flight time spent in the transfer ion guide. The CCSs of these calibrants were corrected for their charge state and reduced mass with respect to the buffer gas. The natural logarithm of corrected CCSs were plotted against the natural logarithm of corrected drift times and a mathematical formula ($\mathbf{Ln\,\Omega' = A \times Ln\,dt' + B}$) was derived. The coefficient A was extracted to calculate the effective drift times $\mathbf{dt''}$: $\mathbf{dt''}$=$(\mathbf{dt'})^{\boldsymbol{A}} \frac{\boldsymbol{z}}{\sqrt{\boldsymbol{\mu}}}$. A calibration curve was generated by plotting the literature CCSs as a function of $\mathbf{dt''}$. The experimental CCS of the analyte ion were derived from this calibration curve based on the measured drift time [41].

Molecular dynamics (MD) simulations and theoretical CCS calculations

Topology and coordinate files for Robo1 and Robo1-HS complexes were generated using the tLeap program, employing the ff99SB and GLYCAM06 (version j) parameters for the protein and GAGs, respectively. Each system underwent energy minimization (1000 steps) in implicit solvent (IGB = 2). The net charge on each system, after energy minimization, was neutralized by addition of an appropriate number of counter ions (Na^+ or Cl^-). This was followed by solvation with TIP3P water molecules in a cubic box extending at least 12 Å from any atom of the solute.

All MD simulations were performed with the GPU implementation of pmemd, pmemd.cuda_SPDP in Amber14. Energy minimization of the solvent was performed in an NVT ensemble (1000 steps of steepest descent, 24,000 steps of conjugate gradient), followed by a full system energy minimization (1000 steps of steepest descent, 24,000 steps of conjugate gradient). The systems were heated from 5 to 300 K over 60 ps in an NVT ensemble, with a weak positional restraint (10 kcal/mol-Å^2) on the atoms in the solute. A Berendsen-type thermostat with a time coupling constant of 1 ps was utilized for temperature regulation. Equilibration and production were performed at constant pressure (NPT ensemble; 1 atm) with a pressure relaxation time of 1 ps. After the heating step, the restraints were removed from the solute atoms, and the entire system was allowed to equilibrate at 300 K for 1 ns. All covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm, allowing a simulation time step of 2 fs. Scaling factors for 1–4 interactions were set to the recommended values of 1.0 and 1.2 for the GAG and protein, respectively, and a non-bonded interaction cutoff of 8.0 Å was employed. Long-range electrostatics were computed with the particle mesh Ewald (PME) method. Data were collected for 200 ns for each system. Post-processing of the MD simulations was performed using CPPTRAJ module of Amber. The graphical representations were generated using VMD. CCSs were calculated using 50 frames extracted from the last 100 ns of the simulation using MOBCAL. Each of these frames was minimized *in vacuo*, followed by the calculation of CCS using the projection approximation and trajectory methods.

Results and discussion

Native Mass Spectrometry of unbound and HS-bound Ig1-2 domains of Robo1

The native mass spectrum of the N-terminal Ig1-Ig2 domains of unbound Robo1 was first acquired under non-denaturing conditions without the addition of any HS (Figure 5.1 a). A dominant monomeric form of Robo1 with a charge state distribution ranging from +13 to +9 over a m/z range of 1800-2700 was observed. The narrow distribution of these lower charge states indicates that Robo1 ions adapt a natively folded and compact conformation with fewer basic sites exposed for protonation, and the solution structure of Robo1 survived the ionization process and the environment of the TWIMS instrument. An average molecular weight of approximately 23.5KDa was obtained, in agreement with the literature molecular weight of Robo1 with a single N-Acetylglucosamine (GlcNAc) residues attached at asparagine (210). A dimeric form of Robo1 with a charge state distribution ranging from +16 to +13 was also observed, but its biological relevance is unclear due to its low abundance.

Robo1 was next incubated with an equi-molar mixture of fully sulfated HS ranging in size from dp4 to dp12 to reach a final molar ratio of 1:3 (protein to GAG). These HS oligosacchrides consist of repeating disaccharide units of (IdoA2S-α-(1-4)-GlcNS6S)n. The overall charge state distribution remained unchanged, while additional peaks corresponding to the formation of Robo1-HS complexes in a binding stoichiometry of 1:1 were observed for each charge state (Figure 5.1 b). The fact that HS as short as a dp4 can be able to form 1:1 complex with Robo1, is in accordance with a previous SPR competition experiments which showed that HSs longer than dp2 is required to efficiently compete with chip immobilized heparin for interacting with Robo1[44]. The absence of complexes in other binding stoichiometries agreed with previous findings determined by size-exclusion chromatography [26].

Size dependent specificity of Robo1-HS interactions

Since HS oligosaccharides of different sizes competed with each other for Robo1, the abundance of the complex peaks should provide an experimental estimation of their binding

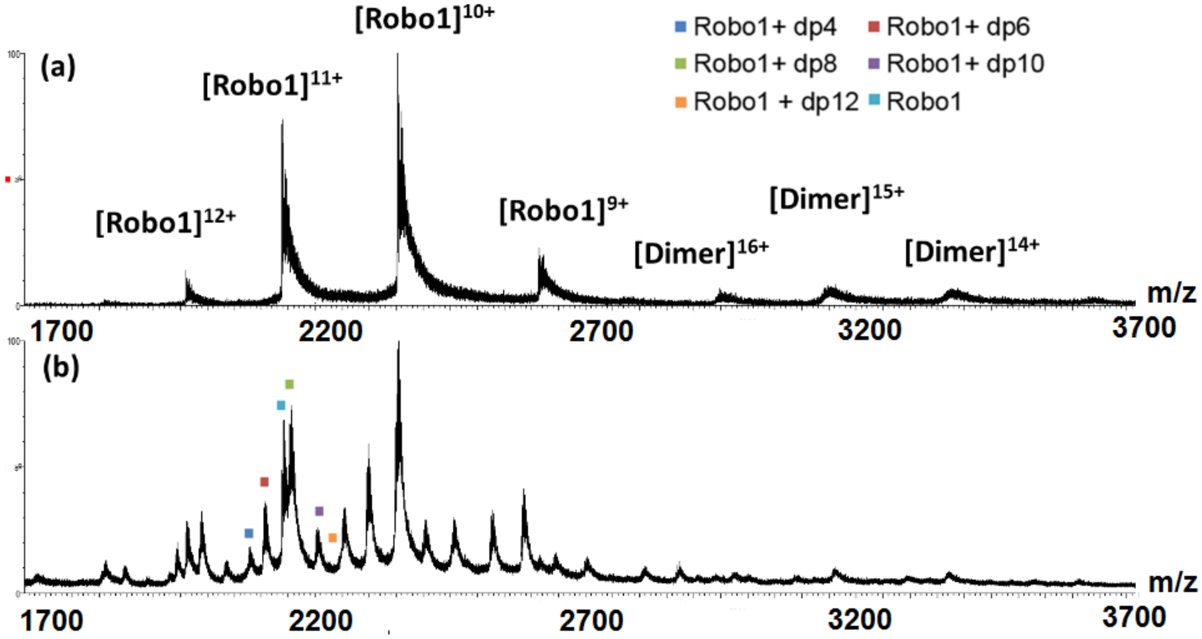


Figure 5.1 (a) native ESI mass spectra of unbound Robo1, which shows a dominant monomeric form of Robo1 with a narrow charge state distribution; (b) Robo1 incubated with an equi-molar mixture of HSs ranging in size from dp4 to dp12. The +12 charge state of each Robo1-HS complex is color coded based on the HS oligosaccharide.

affinity towards Robo1. Despite a few overlapping peaks between +n charge state of unbound Robo1 and +(n+1) charge state of Robo1-HS complex, the overall trend for all the charge states is that the Robo1-dp8 complex displayed the most abundant peak among all the complexes, and its abundance was even higher than those complexes with longer HSs that carry more charged modifications. These observations clearly indicated that Robo1-HS interaction is highly sensitive to the size of HS instead of driven by overall level of charge/sulfation. The facts that a high affinity interaction between Robo1 and HS requires a structural motif of a distinctive size, and adding more saccharide units or sulfations inhibit the interaction agree with the binding behaviors of a highly specific interaction.

Coincidently, a previous experiment on exposing Xenopus embryonic brains to selectively desulfated heparins followed by assessing their mistargeting-inducing activity suggest that the minimum size of HS required for mistargeting is dp8 [45]. Another previous study based on separation and sequencing of Robo1-bound, natural derived HSs also identified a dp8 as a potential ligand for this interaction, despite that the detailed sulfation patterns are different from the dp8 used in this study [46]. Taken together with our observation that a HS dp8 has the highest binding affinity towards Robo1 when competing with longer, more sulfated HSs suggest that the binding epitope on Robo1 could be a shallow grove which prefer to accommodate a HS dp8.

Conformational change involved in Robo1-HS interaction

Evidence of a possible conformational change within Robo1 was observed upon the binding of HS/Hp in previous studies: using surface plasmon resonance (SPR), a much better fit for the Robo1-Heparin/HS binding data was achieved with a two-state (conformational change) binding model than a 1:1 Langmuir model [44]; using hydroxyl radical protein footprinting (HRPF) mass spectrometry (MS), residues on Ig2 domains was observed to be less solvent-exposed upon the binding of heparin, suggesting a dynamic movement of Ig2 domain or a conformational change of the first two Ig domains [47]. Therefore, in the current study, TWIMS was applied to investigate whether Robo1 undergoes a conformational change due to the binding of HS and to directly measure the conformational change if there is any.

The arrival time distribution of unbound Robo1 displayed a distribution featuring two separate peaks. Two Gaussian distributions were used to fit the measured arrival time distribution according to its shape, with each Gaussian distribution representing a potentially stabilized gas-phase conformation (Figure 5.2 a). The conformation with a lower drift time corresponds to a more compact conformation of Robo1, with a relative peak abundance of 20.1± 0.2 %. The more dominant conformation with a higher drift times corresponds to a more open or extended conformation.

The arrival time distribution of Robo1 in complex with a fully sulfated HS dp6 (Figure 5.2 b) also observed two peaks, corresponding to two stabilized gas-phase conformations with different shape and compactness. The relative peak area of the conformation with a lower drift time is 51.7± 1.1%, indicating a higher population of folded conformation in the complex form of Robo1. We attributed the dramatically shifted conformation equilibrium upon the binding of HS to a more favorable interaction between HS and Robo1 in the more compact conformation.

The experimental collision cross sections (CCSs) for the two conformations observed for both unbound and HS-bound Robo1 were calibrated based on the measured drift times and are summarized in Table 5.1. The experimental CCS of the compact conformation is approximately

17.3% smaller compared with the other conformation in the unbound form, which is in close agreement with the experimental CCS difference between the two conformations in the complexed form (approximately 16.8%).

Previously two distinctive crystal forms including one linear shape and one slightly bent shape were resolved for the Ig1-2 region of human Robo1 [18]. Both two crystal forms were used for the calculation of theoretical CCSs of Robo1, which were then compared with the experimental CCSs. Though the crystal structure of HS-bound Robo1 is not available, a constrained docking model proposed by a previous NMR study was used to estimate the theoretical CCS of the Robo1-HS complex [48]. Three methods were used for calculating theoretical CCSs, as shown in Table 5.2. Previous studies showed that projection approximation (PA) and trajectory (TM) method set the lower and upper limit of the range of acceptable experimental CCSs, while calibrating the PA CCSs to generate scaled PA CCSs allow a better comparison of experimental and theoretical CCSs [49, 50].

To our surprise, though the elbow angle of the two crystal forms varies from 140° to 180°, their theoretical CCSs are quite comparable. It is also worth noting that the theoretical CCSs evaluated from the crystals structures are in close agreement with the experimental CCS of the more extended conformation. These observations indicated that in addition to the conformation that matches with the crystal structures, a new conformation was observed in TWIMS experiment, corresponding to a more compact structure of Robo1 with a much smaller CCS. The fact that HS dp6 preferentially interacts with this new and more compact conformation suggest a higher biological relevance of the new conformation, though solution phase evidence is necessary to confirm this suggested postulation.

A putative model for the newly identified conformation was proposed, based on the structural similarities of the extracellular Ig1-Ig2 domains of leukocyte common antigen-related receptor protein tyrosine phosphatases (LAR-RPTP). LAR-RPTP also belongs to the cell adhesion family of prtoein, it shows HS binding affinity and is also functionally involved in the development of vertebrate nervous system [51]. In its crystal structure, the first two Ig domains of LAR-RPTP align parallel with each other and form a much more dramatically bent, V-shaped structure [52], which was used as the template to build a LAR-Robo1 hybrid model (Figure 5.3). The theoretical CCS of the proposed LAR-Robo1 hybrid model was measured, and is in close agreement with the experimental CCS of the more compact conformation of Robo1 (Table 5.2), which indicate that this LAR-Robo1 hybrid model may represent a similar compact structure that Robo1 adopts in TWIMS experiment.

We have also analyzed the Ig1-Ig2 domains of LAR-RPTP. In contrast to Robo1, only one compact conformation was observed for LAR-RPTP. The observance of a single conformation is expected, since the V-shaped structure of LAR-RPTP is greatly stabilized though the presence of tight interaction including hydrogen bonds, salt bridges and hydrophobic interactions between its Ig1 and Ig2 domains. Theoretical CCSs of LAR-RPTP also agreed with its experimental CCSs (data not shown).

Although the role of HS in the Robo1-Slit signaling is well recognized, the molecular details of how they interact to form the ternary signaling complex are still unknown. Previous experiment that selectively mutated amino acids from different domains of Robo1 followed by evaluating the effect using solid phase assay identified that the major Slit binding site is primarily conserved in Ig1 domain of Robo1 [19]. Using SPR, a 10 times lower Kd was measured for the Ig2 domain, indicating a negligble binding affinity compared with Ig1 [18].

However, a previous domain deletion mutagenesis experiment showed that removing the Ig2 domain completely abolished the biological activity of Robo1 to interact with Slit in vitro [32]. A recent domain deletion study suggests that Ig2 domain of Robo1 is the only Ig domain other than Ig1 that contributed to its binding to Slit as well as its function of axonal localization guidance and repulsive signaling in vivo [53]. Therefore, it seems that both the Ig1-Ig2 domains of Robo1 are required for interacting with Slit, but this also contradicts with the crystallographic findings of Robo1-Slit complex, which suggest that a conserved motif on Slit forms a concave binding site for Robo1 that can only accommodate a single Ig domain [18]. These observations raised questions about the contribution of the Ig2 domain in the Robo1-Slit interaction. Moreover, studies on the crystal structures of Ig1-Ig2 domains of human and Drosophila Robo1 both reported substantial flexibility of the hinge region, and have both raised questions about the biological relevance of this flexibility.

Taking our and other's observations together, we proposed that HS plays an essential role in activating Robo1. Native Robo1 prefers to stay in a more extended conformation, as shown in our TWIMS measurement and identified by previous X-ray crystallography experiments. The function of HS is to either selectively recognize and interact with the less abundant compact conformation of Robo1, or induce a conformational change of the more extended conformation as they interact with each other, resulting in the increased population of the more compact conformation in the complexed form of Robo1. It is likely that the compact conformation of Robo1 has higher biological activity than the extended conformation and is more efficient in stabilizing the interactions between Slit and Robo1, further facilitating the formation of the ternary Robo1-Slit-HS complex and the downstream signaling processes.

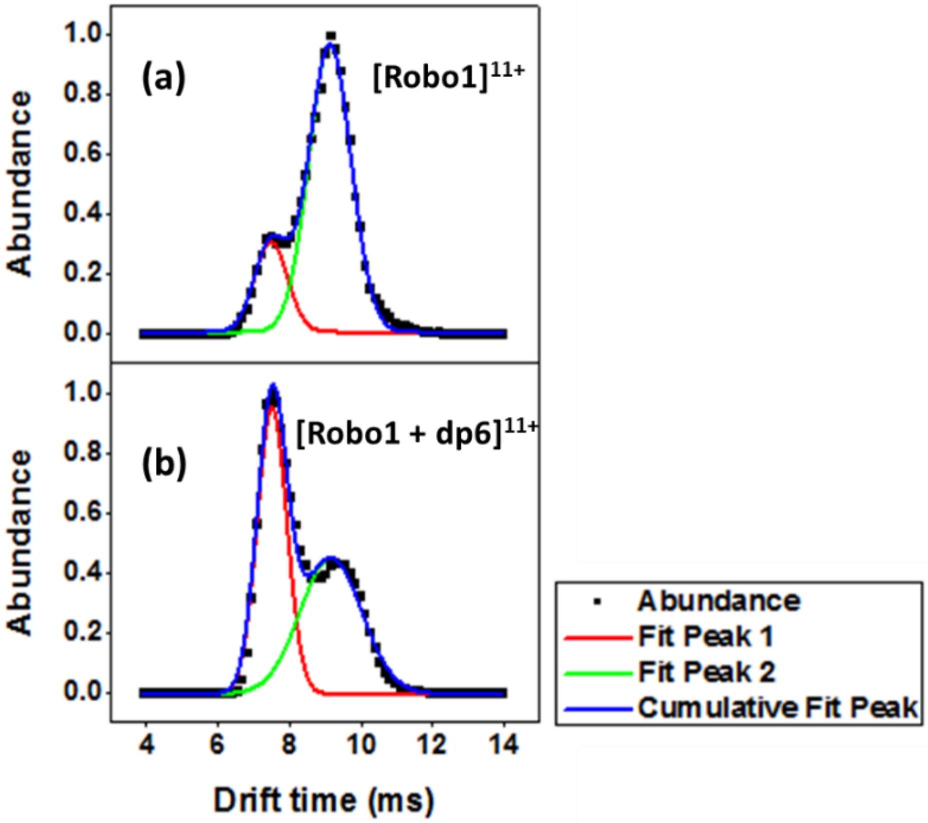


Figure 5.2 Arrival time distributions (ms) of +11 charge state of unbound Robo1 (a) and Robo1 bound with a fully sulfated HS dp6 (b) both display one compact conformation and one extended conformation. The relative ion population areas of the more compact conformation for unbound and complexed Robo1 are 20.1± 0.2 % and 51.7± 1.1%, respectively, indicating that HS dp6 preferentially interacts with the more compact conformation.

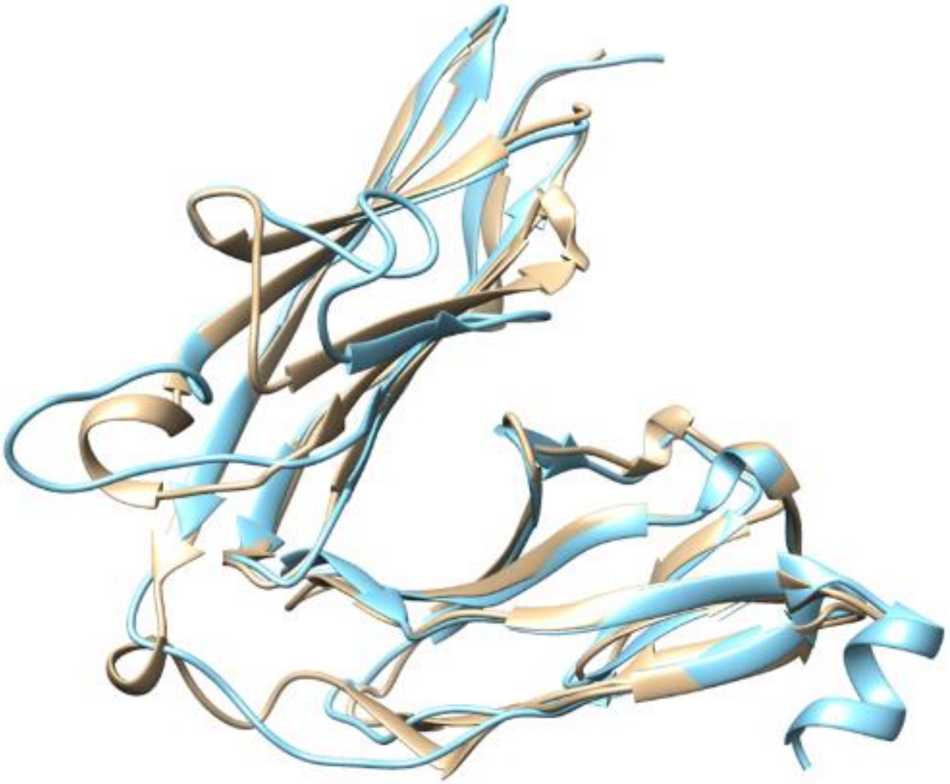


Figure 5.3 The proposed structure for the compact conformation of Robo1 is shown in blue. The structure of LAR-RPTP which was used as the template is shown in yellow.

Table 5.1 Experimental CCSs of the two conformations for Robo1 and Robo1-HS dp6 complex

Experimental CCS (Å^2)	Apo Robo1	Robo1-HSdp6 complex
Compact Conformation	2026.1±8.8	2076.4±8.1
Extended Conformation	2376.7±8.1	2425.1±14.0

Table 2. Theoretical CCSs of Robo1 and its complex

	PA CCS (Å^2)	TM CCS (Å^2)	Scaled PA CCS (Å^2)
Robo1 (linear crystal)	2042.9±4.3	2528.1±9.2	2328.9±7.1
Robo1 (slightly bent crystal)	2038.4±3.4	2524.9±9.8	2323.8±3.9
Robo1-dp6 complex (Docked structure)	2166.6±3.5	2682.0±9.8	2469.0±4.0
Robo1 (Proposed model)	1824.5±3.0	2285.7±6.7	2079.9±3.4

Sulfation pattern dependent specificity of Robo1-HS interaction

In an effort to examine the effect of structural diversity on Robo1-HS interactions, an equi-molar mixture of Hp dp12 with different sulfation patterns were incubated with Robo1 to reach a final molar ratio of 1:2 (protein to GAG). Peaks corresponding to the formation of 1:1 Robo1-HS dp12 complexes were observed for the fully sulfated, N, 6-O- sulfated and N, 2-O-sulfated whereas no complex corresponding to the de-2-O and de-6-Osulfated dp12 bound Robo1 was observed (Figure 5.4). It is evident that the peak abundance for Robo1-fully sulfated dp12 is much lower than those complexes with HS missing 6-O or 2-O sulfations. The fact that absence of 2-O or 6-O sulfations improved binding is in great contrast to the less specific FGF1-HS binding system that we investigated before [31]. These observations indicated that Robo1-HS interaction is sensitive to the sulfation pattern of HS, rather than driven by the amount and density of charged groups. However, since the peak abundance of the N, 6-O- sulfated dp12 bound complex and the N, 2-O- sulfated sulfated dp12 bound complex are comparable, we cannot compare the contribution of 2-O and 6-O sulfation based on MS measurements alone.

We then checked the ion mobility profile of Robo1 in complex with differently modified HS dp12s (Figure 5.5). The IM profiles of different Robo1 complexes were quite distinct from each other, indicating that sulfation pattern of HS dictates different binding modes with Robo1. Though two peaks corresponding to two conformations with different shape and compactness were observed for each complex, the abundance of the more compact conformation varies for each complex, indicating that alteration of the sulfation pattern of HS affected its capability to recognize the compact conformation.

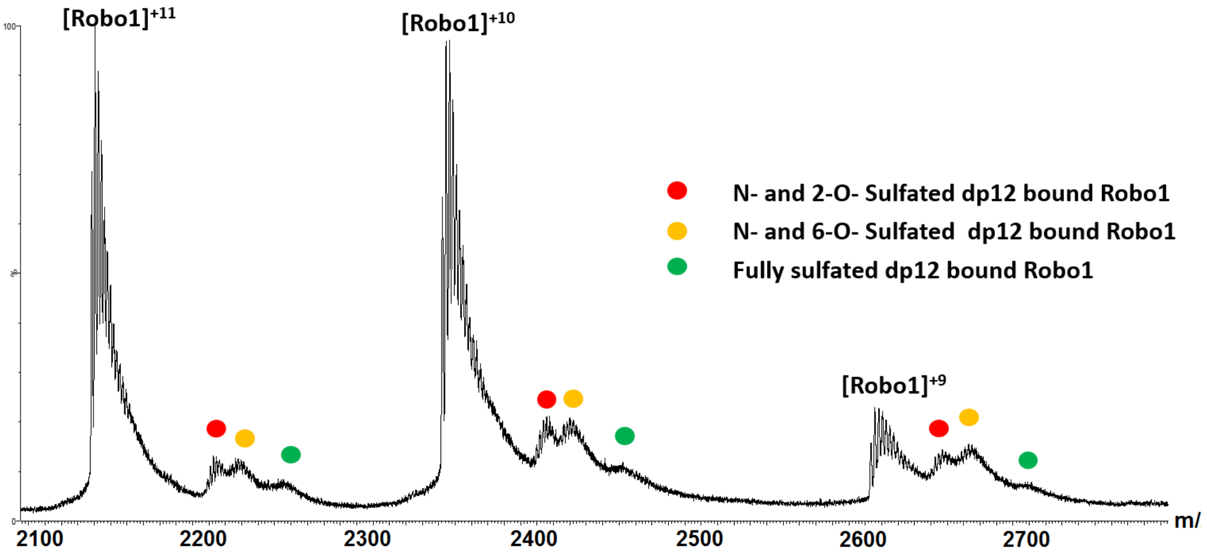


Figure 5.4 Native ESI mass spectra of Robo1 incubated with an equal molar mixture of HS dp12 with different sulftion patterns. Peaks corresponding to Robo1 in complex with fully sulfated and selectively desulfated dp12 are color coded according to the HS oligosaccharide.

A previous SPR study which measured the competition between chip immobilize heparin and selectively desulfated heparin-derived oligosaccharides in solution demonstrate that N-sulfo and 6-O-sulfo sulfations contribute more to the interaction than 2-O sulfations [44]. Another study which measured the binding affinity of chip immobilized Robo1 and a series of chemically modified HS dp4 suggest that 2-O sulfation has a negative effect on the interaction [46]. Comparing the IM profiles of fully sulfated dp12 bound and de-2-O sulfated HS dp12 bound

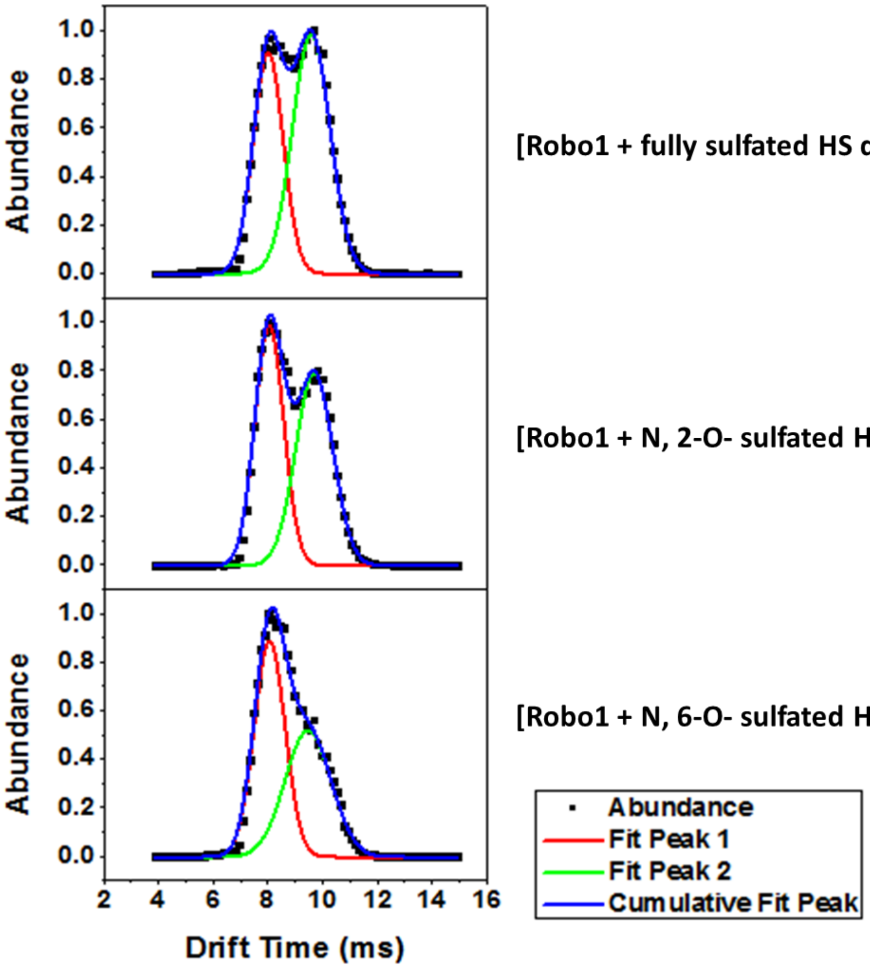


Figure 5.5 Arrival time distributions (ms) of +11 charge state of Robo1 in complex with fully sulfated dp12 (upper), Robo1 with N, 2-O- sulfated dp12 (middle) and Robo1 with N, 6-O- sulfated dp12 (bottom). The relative ion population areas of the compact conformation for each Robo1-dp12 complex are 43.5± 0.7 %, 45.7± 0.7 % and 54.1± 0.4 %, respectively. The observation that HSs with different sulfation patterns have different preference for these two conformations indicates HS sequence specificity.

complex, the relative ion population area of the more compact conformation of Robo1 increase by over 10% after removing 2-O sulfation. This observation indicates that removing 2-O sulfation leads to a more effective binding with the compact conformation, consistent with those previous results. Taken together, the Robo1-HS interaction requires a specific combination of different patterns of sulfation for recognizing the more functionally relevant conformation of the protein.

This observation is in great contrast to the behavior of the chemokine MCP-1 dimer-HS interaction [54]. Though two conformations were also observed for MCP-1 dimer, only fully sulfated HS demonstrated a preferential binding towards the compact conformation. Selectively desulfated HS showed decreased binding affinity towards the more compact conformation. The relative ion population areas of the two conformations of MCP-1 dimer in complex with differently modified HSs were quite comparable. The different behavior of Robo1 and MCP-1 indicates that Robo1 has a more stringent and more specific requirement for its HS ligand.

The effect of glycosylation on conformation and binding affinity of Robo1

Native Robo1 has a glycosylation site at the asparagine residue located near the C-terminus of its Ig1 domain, with a full size, complex type of N-glycan attached. The results above were detained with a more homogeneous form of Robo1 with a single GlcNAc residue attached. Additional experiments were performed to investigate whether different size and composition of glycosylation may affect the form, population/equilibration and stability of the gas-phase conformation of Robo1. Three glycoforms of Robo1 from two batches of expressed Robo1 were investigated. The first batch was expressed in FreeStyleTM 293-F cells to produce wild-type Robo1 which exhibits heterogeneous mixture of N-glycans. The second batch was expressed in HEK293S (GnTI-) cells to produce a high-mannose form of Robo1. For comparison purposes, some of these products were treated with EndoF1, which specifically cleave in between two GlcNAc residues of the core structure of N-glycan to generate some monosaccharide form of Robo1.

The native MS of different glycoforms of Robo1 are shown in Figure 5.6. Three peaks that correspond to +11 charge state of Robo1 with one GlcNAc attached, Robo1 with a full length, high mannose form of N-glycan attached (Man5GlcNAc2) and Robo1 with heterogeneous, full length N-glycans attached (Gal3GlcNAc3Man3GlcNAc2Fuc1), were selected for the further IM experiments. Regardless of glycosylation forms, two IM peaks corresponding to two conformations were observed. The compact form of Robo1 was even more abundant for the glycoforms with full-length glycans than for the monosaccharide form. These observations indicated that increasing the size of N-glycan shifted the conformational equilibrium of Robo1 and preferentially stabilized the more compact conformation. Robo1 used to generate the crystal structures did not have any N-glycan attached, which may explain why the more compact conformation was not resolved by X-ray crystallography [18].

We then incubated Robo1 of different glycoforms with a fully sulfated HS dp6 and measured the ion mobility profiles of formed complexes. As shown in Figure 5.6 b, the compact conformation of Robo1in complex form is still the dominant peak regardless of the attached N-glycan, indicating that HS still maintained its preference for the compact conformation of Robo1, and this preference was not dramatically affected by altering the size and composition of the attached N-glycan.

Altogether, we observed that Robo1 stayed in two conformations regardless of the variations of the size and composition of the N-glycan component, but the relative abundance of each conformation varied. Though a full-length N-glycan contributed to a higher stabilizing effect of the more compact conformation, it did not have a substantial effect on the binding preference of HS for the compact form of Robo1.

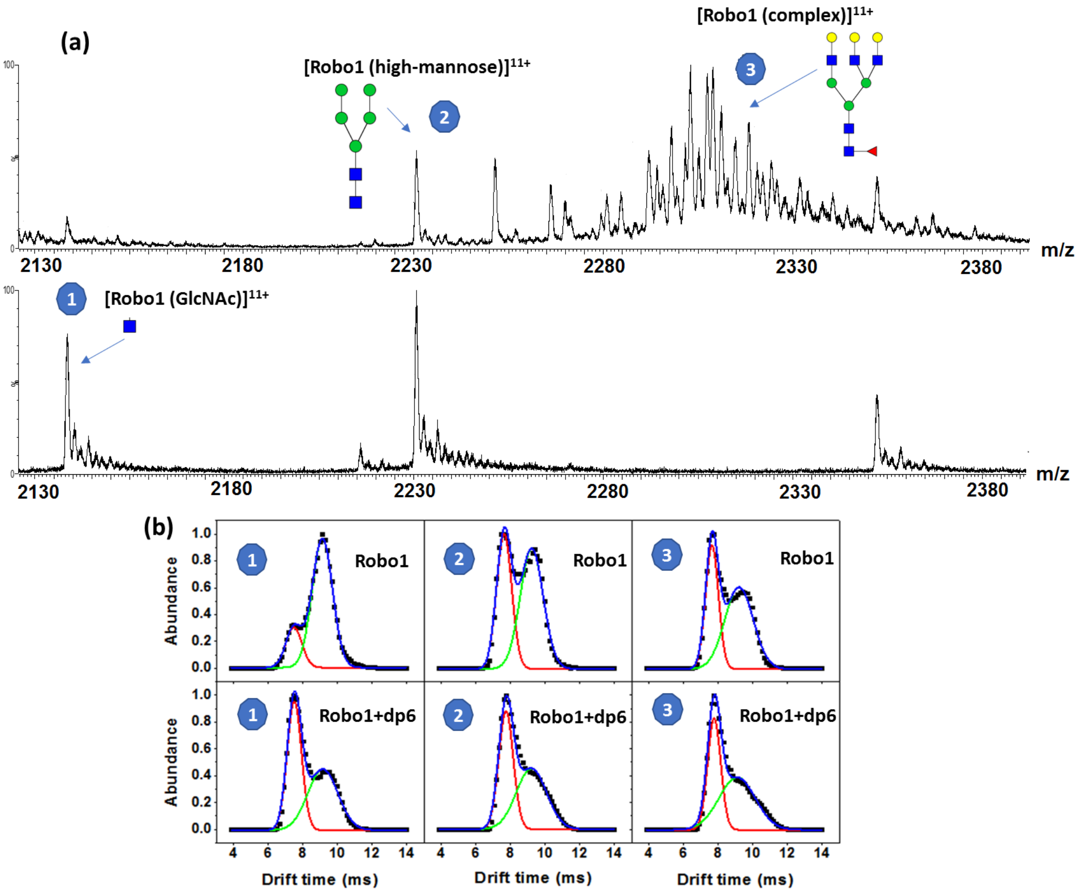


Figure 5.6 (a) Native ESI mass spectra of different glycosylation forms of Robo1; (b) Arrival time distributions (ms) of +11 charge state of different glycosylation forms of Robo1 with and without the binding of a fully sulfated HS dp6. The relative ion population areas of the compact conformation for each glycosylated form of Robo1 are 20.1± 0.2 %, 41.3± 0.9% and 40.3± 0.8 %, respectively. The relative ion population areas of the compact conformation for each glycosylated form of Robo1-HS dp6 complex are 51.7± 1.1%, 46.7± 0.4% and 45.3± 0.9%, respectively.

Gas phase stability of Robo1 and Robo1-HS complexes

The impact of HS binding on the stability of the three-dimensional structures of Robo1 and Robo1 complexes was examined with collision induced unfolding (CIU) experiments. In these experiments, trapping collisional energy (CE) were applied to selected Robo1 and Robo1-HS complex ions to gradually increase their internal energy, while the resulting conformational transition from a fully folded state to a fully unfolded state due to collision activation was tracked. The resulting CIU fingerprints of the 10+ charge state of Robo1 and Robo1-HS complexes are shown in Figure 5.7 as contour plots, in which the drift time and intensity of each feature that emerged during the unfolding process was plotted against the trap CE, which was ramped from 0V to 80V in 2.5V increments.

The CIU fingerprint of unbound Robo1 exhibits three features. The first feature, which has the lowest drift time, corresponds to a natively folded, compact conformation. A second feature appeared at a CE of 10V and stayed stabilized up to 30V, which corresponds to an intermediate state or a more extended, partially unfolded conformation. As the CE was increased to 80V, the third feature started to emerge, indicating that Robo1 has completely unfolded.

The CIU fingerprints of 1:1 Robo1-HS dp4/dp6 complexes showed similar unfolding behavior. Compared with unbound Robo1, the most compact and the partially unfolded features showed more elongated shape while the fully unfolded feature started to evolve at much higher voltages, indicating that all three features were dramatically stabilized upon the binding of these short HS oligosaccharides. In contrast to the complexes formed with HS dp4/dp6, the CIU fingerprints of Robo1-HS dp8/dp10/dp12 complexes exhibited only two principal conformational features which evolved at similar CE. The absence of a fully unfolded feature indicated that interacting with HS oligosaccharide larger than dp8 induced higher stability against collision induced activation and protected the protein from completely unfolding.

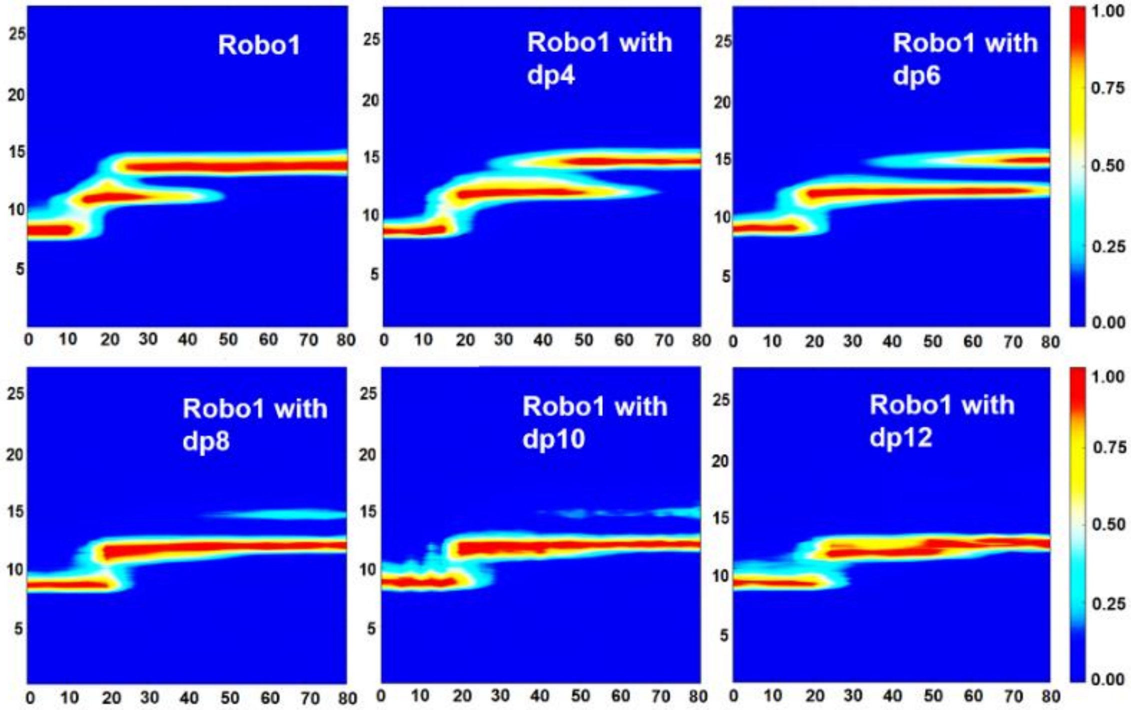


Figure 5.7 The 2D-contour CIU fingerprints of unbound Robo1 and Robo1-HS complexes, with HS of increasing size from dp4-dp12. The ion intensities are indicated by a color axis. As the size of HS increases, the conformational stability of the complex increases accordingly, but HS longer than dp8 prevented the complete unfolding of Robo1 upon activation.

In our previous study on the less specific FGF1-HS interactions, we observed that the degree of gas phase stabilization of FGF1-HS complexes directly correlated with the size of the HS: the larger the HS oligosaccharide, the higher the gas-phase stability of HS bound complexes. In contrast, for the current case of Robo1-HS interaction, we observed comparable unfolding features for Robo1-dp8/dp10/dp12 complexes, indicating that there was no dramatic increase of gas-phase stability as the size of HS becomes larger than dp8. Taken together with our observations that HS dp8 has the highest HS binding affinity in the competition experiment and adding extra saccharide units failed to improve the binding, the collective evidence supports the hypothesis that the binding pocket/motif on Robo1 is about the size of a dp8, and no larger HS oligosaccharide is required for a high affinity interaction with Robo1 to stabilize Robo1 against unfolding.

Conclusions

This study focused on the application of TWIMS to investigate the conformational details of the interactions between Robo1 and a series of HS oligosaccharides of different primary sequence and size. Robo1 exhibits two interchangeable conformations in the gas phase, one with a cross section that matches the structure previously reported using X-ray crystallography and a previously unreported, more compact structure which exhibits a higher binding affinity towards HS. The data suggests that HS mediates the assembly of Robo1-Slit-HS complexes based on stabilizing or promoting Robo1-Slit binding through a conformational change mechanism. Studies using different glycoforms of Robo1 suggest that this mechanism is independent of the composition and size of N-glycan attached to Robo1.

The results also suggest that the Robo1-HS interaction specifically recognizes a structural motif on HS. Alternating the pattern of sulfation on the sequence had a measurable effect on the degree of binding. Though the actual sequence cannot be defined yet, the observations suggest that this structural motif is about the size of a dp8 and is preferentially modified with N- and 6-O- sulfations of GlcN, and weakened by 2-O sulfation of IdoA. Not only did this structural motif mediate the conformation change involved in Robo1-HS interaction, it also showed the highest binding affinity towards Robo1 and was capable of protecting Robo1 from collision-activated unfolding. Collectively, these observations provide new insights into the way HS mediates Robo1-Slit interactions. These results also highlighted the potential of using heparin/HS based agents to regulate biological processes mediated by Robo1-Slit interactions.

References

1. Kennedy, T.E., Serafini, T., de la Torre, J., Tessier-Lavigne, M.: Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell. **78**, 425-435 (1994)

2. Long, H., Sabatier, C., Ma, L., Plump, A., Yuan, W., Ornitz, D.M., et al.: Conserved roles for Slit and Robo proteins in midline commissural axon guidance. Neuron. **42**, 213-223 (2004)

3. Lartillot, N., Philippe, H.: Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. Philosophical Transactions of the Royal Society of London B: Biological Sciences. **363**, 1463-1472 (2008)

4. Borrell, V., Cárdenas, A., Ciceri, G., Galcerán, J., Flames, N., Pla, R., et al.: Slit/Robo signaling modulates the proliferation of central nervous system progenitors. Neuron. **76**, 338-352 (2012)

5. Englund, C., Steneberg, P., Falileeva, L., Xylourgidis, N., Samakovlis, C.: Attractive and repulsive functions of Slit are mediated by different receptors in the Drosophila trachea. Development. **129**, 4941-4951 (2002)

6. Andrews, W., Barber, M., Hernadez-Miranda, L.R., Xian, J., Rakic, S., Sundaresan, V., et al.: The role of Slit-Robo signaling in the generation, migration and morphological differentiation of cortical interneurons. Developmental biology. **313**, 648-658 (2008)

7. Di Meglio, T., Nguyen-Ba-Charvet, K.T., Tessier-Lavigne, M., Sotelo, C., Chédotal, A.: Molecular mechanisms controlling midline crossing by precerebellar neurons. Journal of Neuroscience. **28**, 6285-6294 (2008)

8. Rama, N., Dubrac, A., Mathivet, T., Chárthaigh, R.-A.N., Genet, G., Cristofaro, B., et al.: Slit2 signaling through Robo1 and Robo2 is required for retinal neovascularization. Nature medicine. **21**, 483-491 (2015)

9. Whitford, K.L., Marillat, V., Stein, E., Goodman, C.S., Tessier-Lavigne, M., Chédotal, A., et al.: Regulation of cortical dendrite development by Slit-Robo interactions. Neuron. **33**, 47-61 (2002)

10. Wang, L.J., Zhao, Y., Han, B., Ma, Y.G., Zhang, J., Yang, D.M., et al.: Targeting Slit–Roundabout signaling inhibits tumor angiogenesis in chemical-induced squamous cell carcinogenesis. Cancer Science. **99**, 510-517 (2008)

11. Tseng, R.-C., Lee, S.-H., Hsu, H.-S., Chen, B.-H., Tsai, W.-C., Tzao, C., et al.: SLIT2 attenuation during lung cancer progression deregulates β-catenin and E-cadherin and associates with poor prognosis. Cancer research. **70**, 543-551 (2010)

12. Schmid, B.C., Rezniczek, G.A., Fabjani, G., Yoneda, T., Leodolter, S., Zeillinger, R.: The neuronal guidance cue Slit2 induces targeted migration and may play a role in brain metastasis of breast cancer cells. Breast cancer research and treatment. **106**, 333-342 (2007)

13. Wang, B., Xiao, Y., Ding, B.-B., Zhang, N., Yuan, X.-b., Gui, L., et al.: Induction of tumor angiogenesis by Slit-Robo signaling and inhibition of cancer growth by blocking Robo activity. Cancer cell. **4**, 19-29 (2003)

14. Xian, J., Clark, K.J., Fordham, R., Pannell, R., Rabbitts, T.H., Rabbitts, P.H.: Inadequate lung development and bronchial hyperplasia in mice with a targeted deletion in the Dutt1/Robo1 gene. Proceedings of the National Academy of Sciences. **98**, 15062-15066 (2001)

15. Dallol, A., Krex, D., Hesson, L., Eng, C., Maher, E.R., Latif, F.: Frequent epigenetic inactivation of the SLIT2 gene in gliomas. Oncogene. **22**, 4611-4616 (2003)

16. Walmod, P.S., Pedersen, M.V., Berezin, V., Bock, E.: Cell adhesion molecules of the immunoglobulin superfamily in the nervous system. Springer, (2007)

17. Evans, T.A., Bashaw, G.J.: Functional diversity of Robo receptor immunoglobulin domains promotes distinct axon guidance decisions. Current Biology. **20**, 567-572 (2010)

18. Morlot, C., Thielens, N.M., Ravelli, R.B., Hemrika, W., Romijn, R.A., Gros, P., et al.: Structural insights into the Slit-Robo complex. Proceedings of the National Academy of Sciences. **104**, 14923-14928 (2007)

19. Fukuhara, N., Howitt, J.A., Hussain, S.-A., Hohenester, E.: Structural and functional analysis of slit and heparin binding to immunoglobulin-like domains 1 and 2 of Drosophila Robo. Journal of Biological Chemistry. **283**, 16226-16234 (2008)

20. Lindahl, U., Li, J.p.: Interactions between heparan sulfate and proteins—design and functional implications. International review of cell and molecular biology. **276**, 105-159 (2009)

21. Xu, D., Esko, J.D.: Demystifying heparan sulfate–protein interactions. Annual review of biochemistry. **83**, 129-157 (2014)

22. Bülow, H.E., Hobert, O.: Differential sulfations and epimerization define heparan sulfate specificity in nervous system development. Neuron. **41**, 723-736 (2004)

23. Johnson, K.G., Ghose, A., Epstein, E., Lincecum, J., O'Connor, M.B., Van Vactor, D.: Axonal heparan sulfate proteoglycans regulate the distribution and efficiency of the repellent slit during midline axon guidance. Current Biology. **14**, 499-504 (2004)

24. Steigemann, P., Molitor, A., Fellert, S., Jäckle, H., Vorbrüggen, G.: Heparan sulfate proteoglycan syndecan promotes axonal and myotube guidance by slit/robo signaling. Current biology. **14**, 225-230 (2004)

25. Hu, H.: Cell-surface heparan sulfate is involved in the repulsive guidance activities of Slit2 protein. Nature neuroscience. **4**, 695-701 (2001)

26. Hussain, S.-A., Piper, M., Fukuhara, N., Strochlic, L., Cho, G., Howitt, J.A., et al.: A molecular mechanism for the heparan sulfate dependence of slit-robo signaling. Journal of Biological Chemistry. **281**, 39693-39698 (2006)

27. Ypsilanti, A.R., Chedotal, A.: Roundabout receptors. Springer, (2014)

28. Hou, S.T., Jiang, S.X., Smith, R.A.: Permissive and repulsive cues and signalling pathways of axonal outgrowth and regeneration. International review of cell and molecular biology. **267**, 125-181 (2008)

29. Ballard, M.S., Hinck, L.: A roundabout way to cancer. Advances in cancer research. **114**, 187 (2012)

30. Zhao, Y., Singh, A., Li, L., Linhardt, R.J., Xu, Y., Liu, J., et al.: Investigating changes in the gas-phase conformation of Antithrombin III upon binding of Arixtra using traveling wave ion mobility spectrometry (TWIMS). Analyst. **140**, 6980-6989 (2015)

31. Zhao, Y., Singh, A., Xu, Y., Zong, C., Zhang, F., Boons, G.-J., et al.: Gas-Phase Analysis of the Complex of Fibroblast GrowthFactor 1 with Heparan Sulfate: A Traveling Wave Ion Mobility Spectrometry (TWIMS) and Molecular Modeling Study. Journal of The American Society for Mass Spectrometry. **28**, 96-109 (2017)

32. Xu, Y., Cai, C., Chandarajoti, K., Hsieh, P.-H., Li, L., Pham, T.Q., et al.: Homogeneous low-molecular-weight heparins with reversible anticoagulant activity. Nature chemical biology. **10**, 248-250 (2014)

33. Zong, C., Venot, A., Dhamale, O., Boons, G.-J.: Fluorous supported modular synthesis of heparan sulfate oligosaccharides. Organic letters. **15**, 342-345 (2013)

34. Zhang, F., Moniz, H.A., Walcott, B., Moremen, K.W., Wang, L., Linhardt, R.J.: Probing the impact of GFP tagging on Robo1-heparin interaction. Glycoconjugate journal. **31**, 299-307 (2014)

35. Beckett, D., Kovaleva, E., Schatz, P.J.: A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. Protein Science. **8**, 921-929 (1999)

36. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., Waldo, G.S.: Engineering and characterization of a superfolder green fluorescent protein. Nature biotechnology. **24**, 79-88 (2006)

37. Carrington, J.C., Dougherty, W.G.: A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proceedings of the National Academy of Sciences. **85**, 3391-3395 (1988)

38. Reeves, P.J., Callewaert, N., Contreras, R., Khorana, H.G.: Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proceedings of the National Academy of Sciences. **99**, 13419-13424 (2002)

39. Meng, L., Forouhar, F., Thieker, D., Gao, Z., Ramiah, A., Moniz, H., et al.: Enzymatic Basis for N-Glycan Sialylation STRUCTURE OF RAT α2, 6-SIALYLTRANSFERASE (ST6GAL1) REVEALS CONSERVED AND UNIQUE FEATURES FOR GLYCAN SIALYLATION. Journal of Biological Chemistry. **288**, 34680-34698 (2013)

40. Eschweiler, J.D., Rabuck-Gibbons, J.N., Tian, Y., Ruotolo, B.T.: CIUSuite: a quantitative analysis package for collision induced unfolding measurements of gas-phase protein ions. Analytical chemistry. **87**, 11516-11522 (2015)

41. Ruotolo, B.T., Benesch, J.L., Sandercock, A.M., Hyung, S.-J., Robinson, C.V.: Ion mobility–mass spectrometry analysis of large protein complexes. Nature protocols. **3**, 1139-1152 (2008)

42. Liu, Z., Patel, K., Schmidt, H., Andrews, W., Pini, A., Sundaresan, V.: Extracellular Ig domains 1 and 2 of Robo are important for ligand (Slit) binding. Molecular and Cellular Neuroscience. **26**, 232-240 (2004)

43. Dickson, B.J., Gilestro, G.F.: Regulation of commissural axon pathfinding by slit and its Robo receptors. Annu. Rev. Cell Dev. Biol. **22**, 651-675 (2006)

44. Zhang, F., Moniz, H.A., Walcott, B., Moremen, K.W., Linhardt, R.J., Wang, L.: Characterization of the interaction between Robo1 and heparin and other glycosaminoglycans. Biochimie. **95**, 2345-2353 (2013)

45. Irie, A., Yates, E.A., Turnbull, J.E., Holt, C.E.: Specific heparan sulfate structures involved in retinal axon targeting. Development. **129**, 61-70 (2002)

46. Zong, C., Huang, R., Condac, E., Chiu, Y., Xiao, W., Li, X., et al.: Integrated Approach to Identify Heparan Sulfate Ligand Requirements of Robo1. Journal of the American Chemical Society. **138**, 13059-13067 (2016)

47. Li, Z., Moniz, H., Wang, S., Ramiah, A., Zhang, F., Moremen, K.W., et al.: High structural resolution hydroxyl radical protein footprinting reveals an extended Robo1-heparin binding interface. Journal of Biological Chemistry. **290**, 10729-10740 (2015)

48. Gao, Q., Chen, C.-Y., Zong, C., Wang, S., Ramiah, A., Prabhakar, P., et al.: Structural Aspects of Heparan Sulfate Binding to Robo1–Ig1–2. ACS chemical biology. **11**, 3106-3113 (2016)

49. Jurneczko, E., Barran, P.E.: How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst. **136**, 20-28 (2011)

50. Benesch, J.L., Ruotolo, B.T.: Mass spectrometry: come of age for structural and dynamical biology. Current opinion in structural biology. **21**, 641-649 (2011)

51. Johnson, K.G., Van Vactor, D.: Receptor protein tyrosine phosphatases in nervous system development. Physiological reviews. **83**, 1-24 (2003)

52. Coles, C.H., Shen, Y., Tenney, A.P., Siebold, C., Sutton, G.C., Lu, W., et al.: Proteoglycan-specific molecular switch for RPTPσ clustering and neuronal extension. Science. **332**, 484-488 (2011)

53. Reichert, M.C., Brown, H.E., Evans, T.A.: In vivo functional analysis of Drosophila Robo1 immunoglobulin-like domains. Neural development. **11**, 15 (2016)

54. Seo, Y., Andaya, A., Bleiholder, C., Leary, J.A.: Differentiation of CC vs CXC chemokine dimers with GAG octasaccharide binding partners: an ion mobility mass spectrometry approach. Journal of the American Chemical Society. **135**, 4325-4332 (2013)

CHAPTER 6

CONCLUSIONS

Highly sulfated GAGs, especially Hp/HS interact with many proteins to exert their essential regulatory roles. Though many classical structural biology approaches including NMR spectroscopy, X-ray crystallography and SPR have been used to study these interactions, there is a growing need for new techniques which can complement traditional methods, are compatible with broader range of molecular weight and can analyze heterogeneous biomolecules samples at physiologically relevant concentrations. In the meantime, despite the widely use of Hp as an anticoagulant drug, there is only a few of GAG–protein interactions that feature a defined structure motif or possess specific structural requirements, so new approaches of high speed and sensitivity is required for identifying more GAG-protein binding pairs of specificity and therapeutic potential.

MS techniques use only nanogram to microgram of samples, and can be easily combined with other analytical techniques to probe the structure and dynamics of HBPs and their complexes with GAGs in a fast, sensitive, solvent effect free and high throughput fashion. More importantly, coupled with IM separation, MS techniques allow an easy and straightforward measurement of change in absolute protein conformation and their conformation dynamics, which has been difficult for many classical methods but is crucial for elucidating the nature and function of GAG-protein interactions.

The MS method that is extensively applied in this dissertation is the approach of native MS coupled with TWIMS, an emerging method that has been widely used in the field of structural biology. In **chapter three**, the capability of TWIMS to measure protein conformational changes in Hp/HS-protein binding events was validated using the well characterized HBP ATIII, a serine protease inhibitor whose activity is dependent on a pentasaccharide sequence occurs on a third of Hp chain. The highly specific binding partner for ATIII was chosen as Arixtra, which is the smallest synthetic Hp containing the required pentasaccharide sequence.

Our TWIMS results suggest that not only the natively folded conformations of ATIII in its free and Arixtra bound form were maintained in TWIMS experiments, the gas phase experimental data acquired using TWIMS were consistent with the known molecular details of ATIII-Arixtra binding: TWIMS experiemnt reproduced the 1:1 binding stoichiometry for Arixtra-bound ATIII interaction. Experimental CCSs derived from IM measurements suggest that binding of Arixtra to ATIII caused a 3.6 % increase in its rotationally average conformation. Both the absolute CCSs of ATIII and its complexes as well as the degree of CCS change agreed with theoretical CCSs measurements based on their known crystal structures.

Furthermore, the binging affinity and specificity of ATIII-Arixtra interaction known from previous solution phase measurements survived the translation into gas phase, as revealed by control experiments, in which HS oligosaccharides with different sequence and sulfation modifications were used to verify the importance of an intact pentasacchride sequence in maintaining the anticoagulant activity of ATIII. We also observed that IM profiles of ATIII complexes can be used to determine the corresponding binding specificity.

Moreover, the influence of collisional activation on the gas phase conformation of ATIII and ATIII-Arixtra complex revealed that the binding of Arixtra provided ATIII with additional stability and slowed down the process of collision induced unfolding.

With the successful application of TWIMS in the ATIII-Arixtra binding system, TWIMS was then used to investigate a more complicated GAG-protein interaction, the FGF signaling system, as described in **chapter four**. It is known that HS mediates FGF signaling by dimerizing FGF and associating FGF and FGFR on the surface of target cells to form trimeric signaling complexes, but there are still controversies in terms of binding specificity of their interactions and architecture of the formed trimeric complexes. TWIMS, coupled with molecular modeling were used to investigate how HS oligosaccharides of various lengths and fine structures affect the binding stoichiometry, binding specificity and gas phase conformational stability of FGF1-HS interaction.

Though previous results suggest that HS octasaccharide is the minimal length required for FGF1 dimerization and activation, our TWIMS results observed a specific binding between FGF1 and a HS tetrasaccharide containing the high affinity binding motif as well as the dimerization of FGF1 induced by this interaction. However, a general lack of specificity was observed for the interactions between FGF1 and longer HS oligosaccharides. We also observed that the degree of FGF1 oligomerization increases as the chain length of HS increases. Based on experimental CCSs measurements, plausible models of 2:1 and 3:1 FGF1-HS complexes were proposed.

In addition, the conformational stability of different FGF1 complexes was assessed, and the observed trend suggest that longer HS oligosaccharides forming complex with higher gas phase stability against collision induced unfolding, which was further confirmed through calculating and comparing the binding free energies of each FGF1 complexes.

In **chapter five**, TWIMS was used to study a much less characterized GAG-protein binding system, the interaction between Robo1 and HS. Robo1 is the receptor for secreted axon guidance molecule, Slit. HSs act as essential co-receptor in Robo-Slit signaling by promoting the formation of trimeric signaling complexes. However, since the crystal structure of Robo1-HS or Robo1-Slit-HS complexes are still unavailable, whether a defined HS structure motif is involved in this interaction and how HS modulates the assembly of the signaling complexes are still unresolved questions.

Our TWIMS results provided answers for these questions: HS oligosaccharides of different sizes or modified with different type of sulfations has distinctive binding affinity towards Robo1, suggesting that Robo1-HS interaction is of high selectivity and HSs play important roles in tuning the Robo1-Slit signaling through their fine structures. Interestingly, a conformational change of Robo1 upon the binding of HS was observed. Taking the previously known molecular details including consideration, we proposed that this conformational change is the basis of the HS mediated, high affinity binding of Robo1 to its Slit receptor. Even though most of our conclusions were based on measurements of Robo1 with a single GlcNAc attached, we also tested different glycoforms of Robo1 and our results suggest that the size and composition of the attached N-glycan affect the conformational stability of Robo1 but hardly its interactions with HSs.

In addition, different from previous binding systems in which gas phase stability of GAG-protein complex dramatically increases as the size of HS oligosaccharide increases, Robo1 stopped gaining additional stability as the size of HS oligosaccharide became longer than dp8, suggesting that dp8 could be the appropriate size to fit the binding groove on Robo1.

Collectively, studies covered in this dissertation demonstrate the advantages of native MS coupled with TWIMS for characterizing the nature and structure-function relationship of different GAG–protein interactions. Future applications of this approach will be useful for elucidating the role of GAGs in physiological and pathological processes mediated by their interactions with all types of HSPs as well as the structural requirements of GAGs for these interactions, facilitating the design and development of novel GAG-mimetic therapeutic agents.

APPENDIX

SUPPLEMENTARY DATA FOR CHAPTER 4

SUPPLEMENTARY DATA FOR CHAPTER 4

Crystal structure 1AXM comprises of three 2:1 FGF1-heparin-pentasaccharide complexes, while 2AXM comprises of one 2:1 FGF1-heparin-hexasaccharide complex. The interactions between each heparin-FGF1 pair are slightly different, and in order to choose one structure for modeling, we split the models into pairs of FGF1 monomer - heparin dp4 with the heparin sequence of IdoA2OS-GlcNS(6S)-IdoA2S-GlcNS(6S). 1AXM was split into 6 FGF1-heparin tetrasaccharide pairs, and 2AXM was split into 4 pairs (by removing two residues from reducing or the non-reducing end). To calculate the interaction energies, each model was subjected to a short implicit solvent minimization followed by single frame MM-GBSA energy calculation. The model with the highest interaction energy, chain B of 2AXM with the tetrasaccharide from the reducing end (Supplementary Table 1) was chosen for further simulations and CCS calculations.

The same tetrasaccharide with FGF1 dimer was used to model the FGF1 dimer – DP4 complex.

Supplementary Table 2

FGF1 monomer – heparin tetrasaccharide complex	MM-GBSA interaction energy[a]
1AXM chain A	-47.3
1AXM chain B	-49.9
1AXM chain C	-40.4
1AXM chain D	-40.6
1AXM chain E	-44.4
1AXM chain F	-48.9
2AXM chain A non-reducing end	-31.3
2AXM chain A reducing end	-52.3
2AXM chain B non-reducing end	-57.6
2AXM chain B reducing end	-58.1

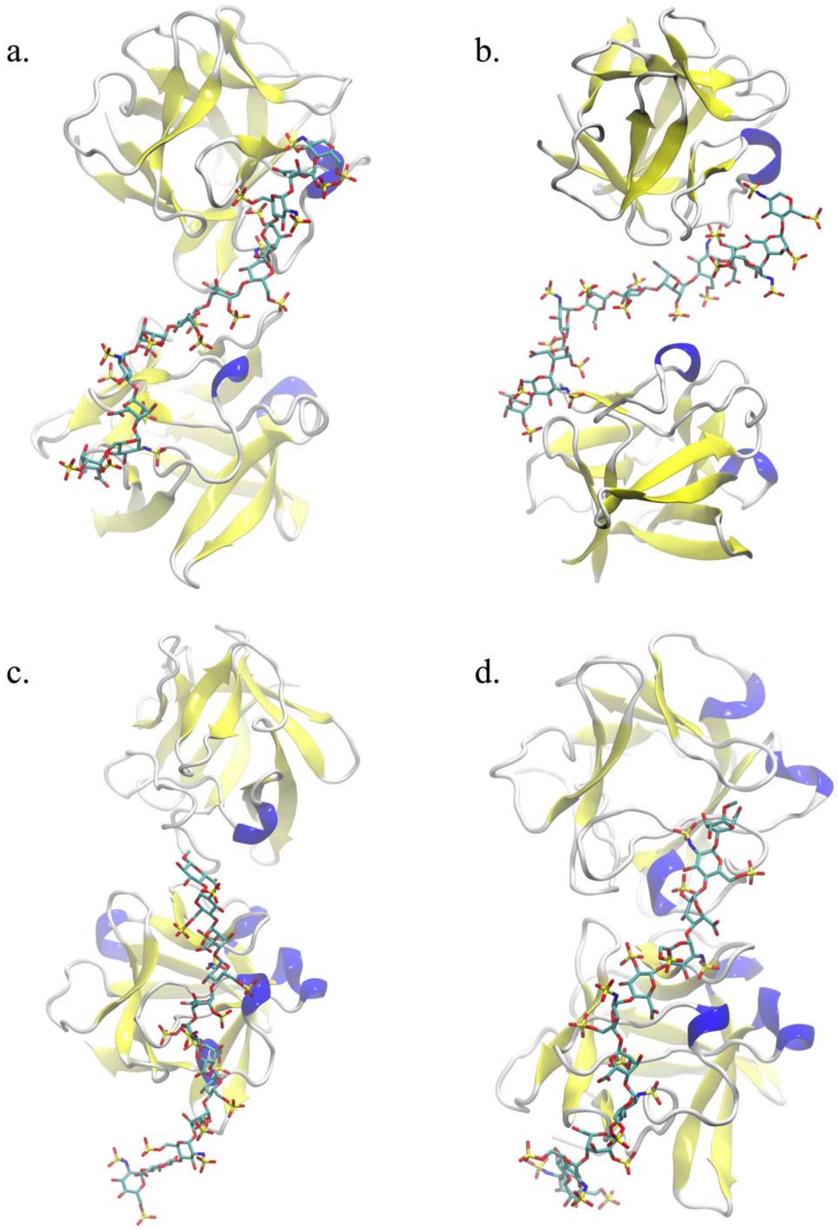
[a] kcal/mol

Supplementary Table 2. Quantification of the conformational difference in the CCS of the apo and complex forms of FGF1.

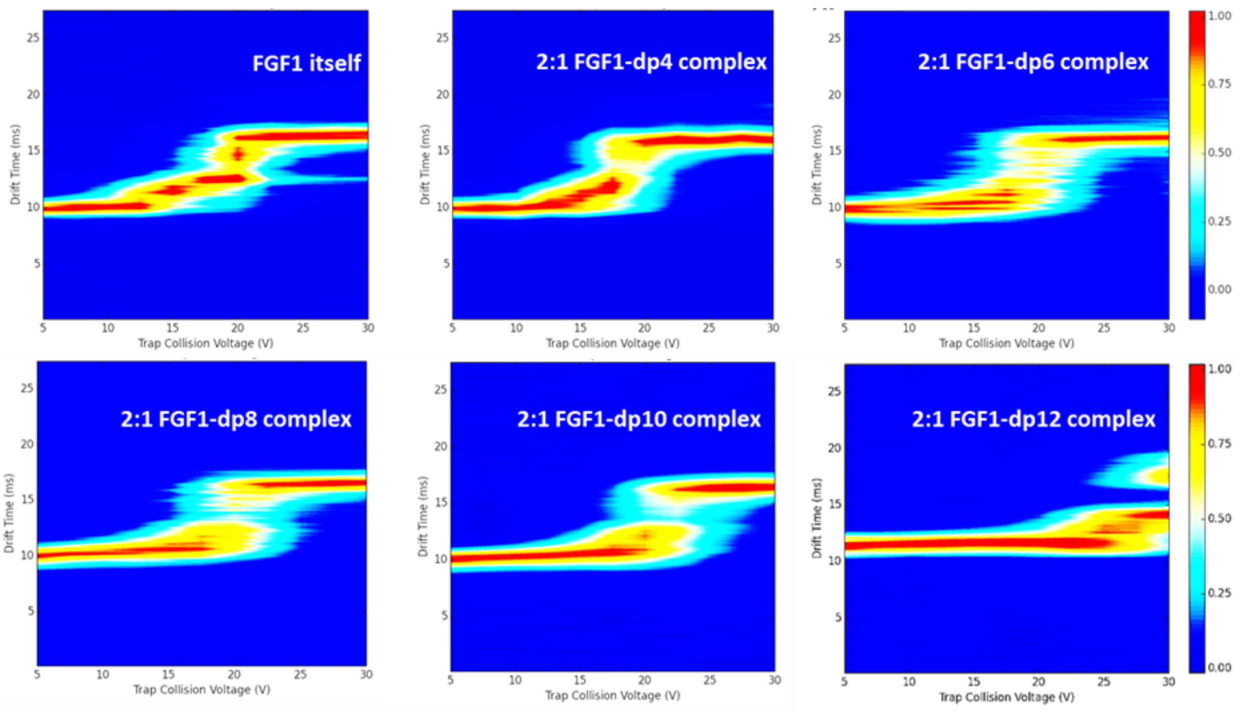
	Collision cross section (Å^2)	
	PA	TM
FGF1-DP4 1:1 complex	1353.1	1699.7
FGF1 after removing HS DP4	1275.2	1604.0
Apo-FGF1 (PDB id: 2AFG)	1288.9	1616.4
CCS increase % due to GAG	4.9	5.2
CCS increase % due to protein conformation	1.0	0.8

Supplementary Table 3. Theoretical CCS (Å^2) for the modeled *Cis* FGF1-dp12 complexes in 2:1 stoichiometry

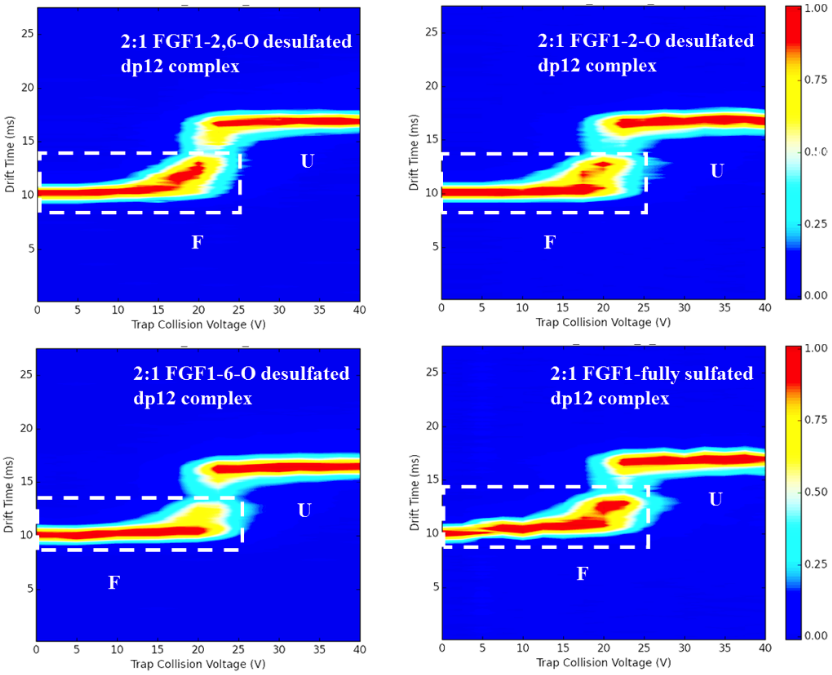
	Collision cross section (Å^2)		
ID	PA	Scaled PA	TM
Cis-model-1 for 2:1 FGF1-dp12 complexes	2149.6 ± 3.6	2450.6 ± 4.1	2742.1 ± 8.3
Cis-model-2 for 2:1 FGF1-dp12 complexes	2192.0 ± 4.2	2498.9 ± 4.8	2788.4 ± 6.9



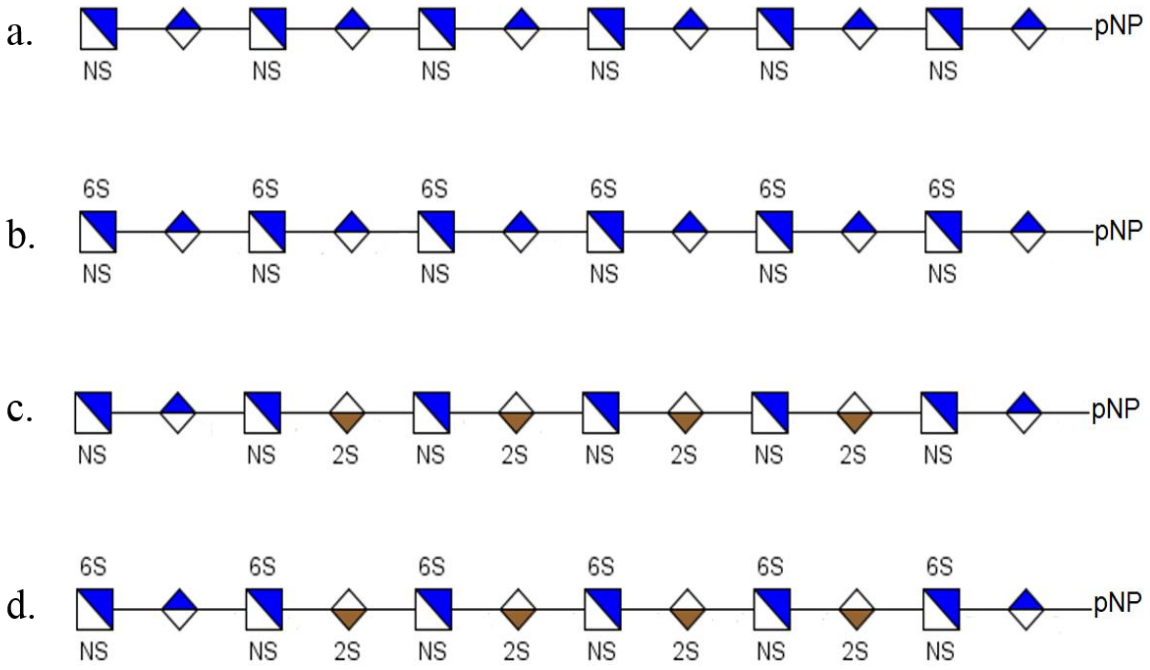
Supplementary Figure 1. Structural representation of the two hypothetical *cis* forms that were modeled for the 2:1 FGF1-HS dp12 complex. Shown here are Cis-model-1 before (a) and after (b) and Cis-model-2 before (c) and after (d) the MD simulations. The protein and HS are represented as ribbons and stick, respectively.



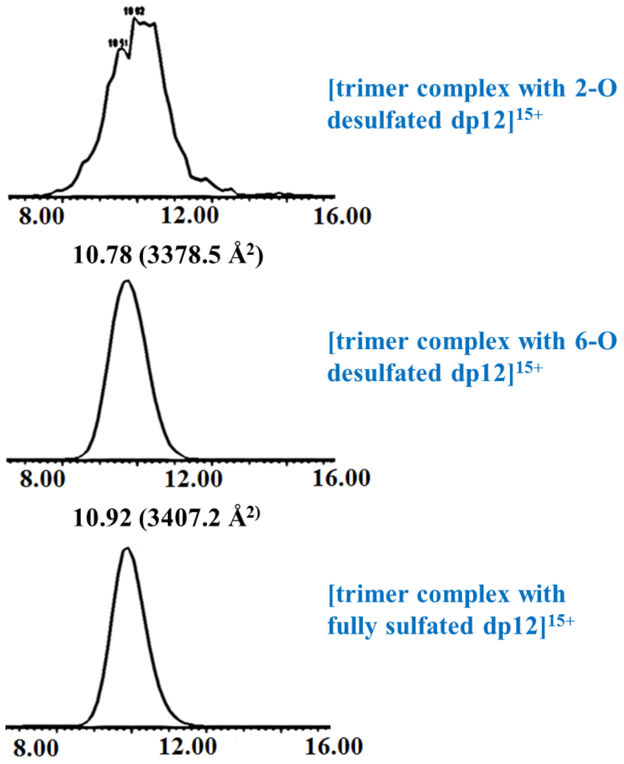
Supplementary Figure 2. The 2D-contour CIU fingerprint of the dimeric apo-FGF1 and the 2:1 FGF1-HS, with HS of increasing length from dp4-dp12. The ion intensities are indicated by a color axis. As the length of the HS increases, they form complexes with FGF1 and increase the conformational stability of the complexes accordingly.



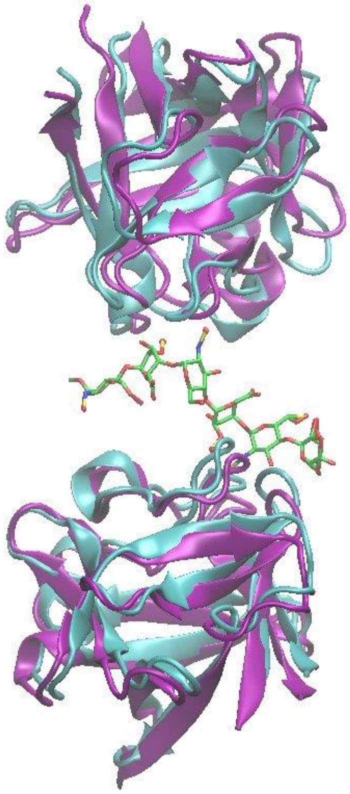
Supplementary Figure 3. The 2D-contour CIU fingerprint of the dimeric FGF1 bound with selected sulfated dodecasaccharides. The conformations were labeled as F and U, indicating folded and unfolded states, respectively. The ion intensities are indicated by a color axis. There is no substantial difference among the unfolding behavior of these four complexes.



Supplementary Figure 4. The structures of the (a) de-2-O- and de-6-O- sulfated dp12; (b) de-2-O- sulfated dp12; (c) de-6-O- sulfated dp12; (d) fully sulfated dp12 that contains the critical binding motif.



Supplementary Figure 5. Arrival time distributions (ms) of +15 charge state of 3:1 FGF1-HS complex, with 2-O desulfated dp12 (upper), 6-O desulfated dp12 (middle) and fully sulfated dp12 (lower).



Supplementary Figure 6. Structural representation of 2:1 FGF1-dp6 dimeric complex, before (purple) and after (cyan) the *in vacuo* minimization. The protein and HS are shown in ribbon and stick representation, respectively. The protein backbones RMSD between the two structures is 2.07 Å, suggest only modest change of the protein structure.