July 2016

Novel Analytical Methods Combining Non-Denaturing Chromatography and Mass Spectrometry to Study Biopolymer Structure and Interactions

Khaja Muneeruddin
University of Massachusetts Amherst

Novel Analytical Methods Combining Non-Denaturing Chromatography and Mass Spectrometry to Study Biopolymer Structure and Interactions

A Dissertation Presented

by

KHAJA MUNEERUDDIN

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2016

Chemistry

Novel Analytical Methods Combining Non-Denaturing Chromatography and Mass Spectrometry to Study Biopolymer Structure and Interactions

A Dissertation Presented

by

KHAJA MUNEERUDDIN

Approved as to style and content by:

_____

Igor A. Katlashov Chair

_____

Richard W. Vachet, Member

_____

Paul L. Dubin, Member

_____

Stephen J. Eyles, Member

_____

Craig T. Martin, Department Head

Chemistry

DEDICATION

To my loving wife, Huma who has been by my side with love, support, and patience

To my daughter, Mizna

To my parents and brother

ACKNOWLEDGMENTS

I would like to express thanks and sincere gratitude to my advisor, Prof. Igor Kaltashov for his unfaltering support and guidance during this research work. I have always felt more confident and focused because of his encouragement and feedback. I highly appreciate the guidance and insightful comments from my committee members Prof. Stephen Eyles, Prof. Richard Vachet, and Prof. Paul Dubin.

My heart full thanks go out to my lab members, Guanbo, Son, Shunhai, Jake, Hanwei, Lola, Shengsheng, Honglin, Yunlong, and Chengfeng. You all have been great colleagues, friends and responsible for creating environment conducive of carrying out great research. Many thanks especially to Dr. Rinat Abzalimov and Prof. Cedric Bobst for invigorating discussions and training me on various laboratory instruments. I would also like to thank my undergraduate researchers, Mark and Sarmad.

I had the opportunity of working with some great collaborators, who were responsible for providing sample and inputs for designing experiments and analyzing data. These include my collaborators from Biogen; Ruth Frenkel, Dr. Iva Turyan, and Dr. Zoran Sosic, collaborators from Shire; Dr. Paul Salinas and Dr. John Thomas (now at Alexion), and my scientific liaison, Anita Szajek from the U.S. Pharmacopeial Convention (USP). I would like to thank all of them for making this a fruitful learning experience. I would also like to acknowledge Dr. Andrei Lipatnikov for his assistance in providing code for mathematical models used in chapter 3.

I also appreciate and highly value the funding support from Chemistry department in the form of teaching assistantships, Biogen (industry Co-Op), and Shire. I am also thankful to the U.S. Pharmacopeial Convention for global USP fellowship award.

Lastly I acknowledge the support and unfailing patience of my wife, family and friends for bearing my long periods of distraction and absence during the course of this study. I owe them a lot too.

ABSTRACT

NOVEL ANALYTICAL METHODS COMBINING NON-DENATURING CHROMATOGRAPHY AND MASS SPECTROMETRY TO STUDY BIOPOLYMER STRUCTURE AND INTERACTIONS

MAY 2016

KHAJA MUNEERUDDIN, B. Pharm., JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Igor A. Kaltashov

Biotherapeutics, an emerging class of medicines containing biopolymers (e.g., proteins, peptide, and polysaccharides) have been developed for a variety of indications including cancer, autoimmune, genetic, and blood diseases. Among these biopolymers protein therapeutics have been the rapidly growing segment in the pharmaceutical industry. This trend combined with the complexity of proteins has necessitated the development of powerful and robust analytical methods to study their structure and interactions with physiological partners.

Mass spectrometry (MS) has become an indispensable tool to analyze various attributes of protein drugs such as profiling of intact mass, amino acid sequencing, and post translational modifications (PTMs). In this work novel analytical methods have been developed by combining non-denaturing chromatographic separations (such as

size exclusion chromatography and ion exchange chromatography) with native ESI MS. In addition to providing information of mass and non-covalent assemblies, native ESI MS has an advantage of probing conformational integrity of proteins by analyzing ionic charge state distributions in the mass spectrum.

Size exclusion chromatography with online native ESI MS detection method developed in this work allowed characterization of higher order structure of proteins and probe aggregation propensity. Using the same SEC/native ESI MS workflow a method to find kinetics and equilibrium binding constant of transient protein interactions have also been developed. Combination of ion exchange chromatography (IXC) with native ESI MS and MS/MS detection was developed and shown to be effective in characterization of positional isomers of protein conjugates and PTMs of a biotherapeutic recombinant human interferon beta 1a. To characterize highly heterogeneous PEGylated glycoprotein (conjugated with a 20kDa PEG) analytical methods using IXC/native ESI MS combined with limited charge reduction and collisional activation were developed.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Biopolymers as pharmaceuticals

Biotherapeutics, also known as biologics are a class of medicines containing biopolymers (e.g., proteins, peptides, polysaccharides, and nucleic acids) either extracted from biological sources or derived by recombinant DNA technology.[1-4] Among the many biopolymers, protein based therapeutics represent the majority of biotherapeutics, and in the past few years has seen a rapid growth with now over 212 products approved for the treatment of a variety of chronic and life-threatening disorders ranging from cancer, metabolic, autoimmune to genetic diseases.[5, 6] The main advantage of biotherapeutics over small molecule drugs is their ability to bind their physiological target in vivo with high specificity resulting in high potency with fewer side effects.[1] However, owing to large size and complexity of biotherapeutics compared to small molecule drugs; biotherapeutics tend to be highly heterogeneous due to various modifications that can happen during expression or upon storage or stress. These modifications inevitably affect the activity and safety profile of biotherapeutics thus validating the need of strong and robust analytical methods to characterize them.

1.1.1. Protein based therapeutics

Protein therapeutics constitute a broad type of biological products, which include recombinant enzymes, monoclonal antibodies, hormones, growth factors, fusion proteins, cytokines, and blood factors.[1] This class of therapeutics usually works by

controlling immune system, decreasing the inflammatory response, or modulating various metabolic processes.

To enhance the existing pharmacokinetic profile of protein therapeutics and to expand the therapeutic activity many strategies have been incorporated resulting in the development of the so-called second generation protein therapeutics. The notable strategies among these are genetically engineering the glycosylation pattern[7] or fusing proteins with Fc region of antibody to increase serum half-life[8], covalent surface attachment of polyethylene glycols (PEG) to decrease renal clearance,[9] and chemical modification of protein with drug payload for site specific activity.[10]

1.1.2. Peptide based therapeutics

Peptide therapeutics are relatively smaller molecules containing fewer than 50 amino acids. Generally peptide molecules act selectively on various cell surface receptors such as G protein-coupled receptors.[11] Though many peptide drugs have been approved, the development of peptide therapeutic is more often hindered due to poor chemical and physical stability, and lack of adequate plasma half-life.[2] To overcome these limitations various chemical strategies such as incorporating specific amino acids, introducing salt bridges or disulfide bonds to increase solubility and stability are being researched.[12, 13]

1.1.3. Polysaccharide based therapeutics (heparin based medicines)

Heparin is a negatively charge linear polysaccharide belonging to the family of molecules known as glycosaminoglycans.[14] Due to involvement of heparin in modulating proteins of coagulation cascade, heparin has been used as an anticoagulant since 1935

for renal dialysis and cardiac surgery. Currently heparin is obtained from porcine mucosa. Owing to safer and effective profile of bovine heparin, it is being re-evaluated to be introduced in market.[15]

1.2. Structural attributes and heterogeneity of polypeptide based drugs

1.2.1. Sequence variants

Sequence variants of polypeptide based drugs (protein and peptide therapeutics) refer to unintentional addition or substitution of amino acids during bio-manufacturing step. This heterogeneity at the primary structure level could be due to DNA mutations, misincorporation of amino acids during translation step, and mis-cleavage during post-translational processing.[16, 17] It is important to analyze the presence of sequence variants as it can have an overall effect on specificity, stability, and immunogenicity of polypeptide drugs.[18] In case the amino acid occurs in the functional domain, any variation of this amino acid can also have an impact on the functional activity.

1.2.2. Enzymatic and non-enzymatic post translational modifications

Post translational modifications (PTMs) add to heterogeneity of primary structure by enzyme mediated covalent attachment of chemical groups to specific amino acids (the so-called enzymatic PTMs). Around 300 of these PTMs have been discovered, some which include phosphorylation, glycosylation, acetylation, methylation, and ubiquitination. These PTMs play an important role in modulating localization, stability, and interaction of proteins. For example it is known that phosphorylation of proteins affects their role in cell cycle, growth, apoptosis and signal transduction pathways. From the context of protein therapeutics glycosylation;

attachment of carbohydrates N-linked to asparagine or O-linked to serine/threonine residues have been shown and indeed used to enhance their stability.[19]

PTMs (e.g., deamidation, oxidation, glycation) can also have adverse effects on proteins and peptides.[20, 21] These types of modifications are known as non-enzymatic PTMs, which occur post-production as a result of storage or stress.

1.2.3. Higher order structure

Correct folding of the polypeptide chain into secondary (alpha helix and beta sheets) and tertiary structures give rise to a unique three dimensional conformation of proteins. This higher order structure is not only important for the functional activity, but also for the stability of protein therapeutics. Unfolding or partial destabilization of the protein conformation (due to storage, stress or non-enzymatic PTMs) can make them prone to aggregation, which has an adverse effect on the safety profile of protein drugs by eliciting immunogenic response in patients. In the case of multimeric proteins, quaternary structures produced by association with identical monomeric units (homo-oligomerization) or with non-identical units (hetero-oligomerization) play a role in function, localization, and stability of protein in vivo by protecting it from proteolytic degradation.[22]

1.2.4. Second-generation protein therapeutics

The extent of heterogeneity is enhanced in the second-generation protein therapeutics especially for products formed by chemical conjugation of drugs (protein – drug conjugates) or PEG chains on proteins post expression. Although the conjugation procedure is optimized carefully; it is very difficult to control the occurrence of

structural heterogeneity (e.g., number of drug molecules attached or presence of isomers due to modification at chemically similar sites), and maintain higher order structure. These heterogeneities affect the properties of such protein therapeutics, and also present great difficulty in their analytical characterization.

1.2.5. Protein-protein interactions

Protein-protein interactions (PPI) can be classified based on various factos.[23] Homo- or hetero-oligomers if interaction occurs between identical or non-identical protein chains. PPI can be distinguished if the complex formed is obligate or non-obilgate. An obligated complex is formed between two proteins that cannot exist as stable structures on their own in vivo, and a non-obligate complex is formed between independently stable proteins. PPI can also be grouped based on binding affinity into permanent and transient interactions. Understanding protein-protein interactions is crucial as they are involved with wide range of biological processes including cell-cell interactions, metabolic, and developmental control. It also plays a key role in drug development.

Traditionally in drug development strong interactions with the equilibrium dissociation constant (K_D) of low nM were favored (e.g., to develop drug molecules binding irreversibly with the target enzyme for complete inhibition). Recently, there has been a shift towards exploiting lower affinity interactions (the so-called transient protein interactions) with K_D in high nM – low μM range, due to their significance in cellular functions[24] and also for the development of targeted drug delivery systems.[25]

1.3. Structural attributes and heterogeneity of polysaccharides (heparin-based medicines)

Heparin is a mixture of linear polysaccharide chains with molecular weight ranging from 5000 – 25000 Da. It contains repeating disaccharide units of N-acetyl D-glucosamine and D-glucuronic acid. The heterogeneity of heparin polysaccharide chain is increased by the presence of a mixture of unsubstituted forms and variable non-template driven sulfated forms (N-and O- sulfation of glucosamine and O-sulfation of glucuronic acid).[26] This heterogeneity has made analytical characterization of this class of medicine very challenging.[27] The structural complexity of heparin-based medicines is reduced in low molecular weight heparins (LMWHs), which are produced from chemical or enzymatic depolymerization of unfractionated heparin. These LMWHs (e.g., tinzaprin, dalteparin, enoxaparin) have a mean molecular weight of 5000 Da and have been demonstrated to have similar anticoagulant activity as heparin.[28] Arixtra is a chemically synthesized anticoagulant containing a fixed length pentasaccharide.

1.4. Commonly used analytical methods to characterize structural heterogeneity

1.4.1. Liquid chromatography methods

Liquid chromatography (LC) is the first line of methods to characterize structural heterogeneity of protein therapeutics by separating protein species using different LC modes. The most common LC techniques for analytical characterization are reverse phase liquid chromatography, size exclusion chromatography, and ion exchange chromatography.

Reverse phase liquid chromatography

Separation in reverse phase liquid chromatography (RPLC) occurs due to interaction of analytes with bonded ligands of the stationary phase.[29] Analytes are eluted in the increasing order of their hydrophobicity using organic solvent gradient. This type of chromatography provides excellent resolution and reproducibility, and is the most common technique used in combination with mass spectrometry detection due to ease of evaporation of the volatile solvent used for elution. RPLC/MS is an excellent choice for characterization of peptides; however for proteins its use is limited to intact mass analysis and to analyze reduced fragments of antibody. Due to the nature of eluent it cannot be used to analyze intact higher order structure of proteins.

Size exclusion chromatography

Size exclusion chromatography (SEC) is the most common technique used in biotechnology industry to analyze the presence of higher molecular weight species or aggregates in protein therapeutics. SEC separates proteins based on their hydrodynamic radius as they diffuse through porous particles of the stationary phase.[30] Isocratic elution conditions (ionic strength: 100 - 150 mM) using physiological buffers in SEC separation maintains higher order structure of proteins, which allows it to analyze non-covalent protein interactions. Though SEC is advantageous due to its simplicity and ease of use, certain limitations do exist such as non-ideality of separation due to interaction with stationary phase and inaccuracy of molecular weight determination from calibration curve due to non-globular shape of certain proteins.

Ion exchange chromatography

Ion exchange chromatography (IXC) is another separation technique in which protein is in its intact state due to non-denaturing conditions employed for separation. IXC works by electrostatic interaction of charged groups on the proteins with the ionic groups of the IXC resin. The pH of the eluent is decided based on the isoelectric point (pI) of protein being analyzed and mode of ion exchange chromatography: for cation exchange chromatography pH of eluent is lower than pI such that acidic groups of cation exchange resin can interact with net positively charged proteins and vice versa for anion exchange chromatography.[31] Elution is carried out using linear salt or pH gradient. Ion exchange chromatography is extremely powerful for characterization of charge heterogeneity of proteins.

1.4.2. Mass spectrometry

Since the advent of soft ionization techniques, electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI), mass spectrometry based techniques have developed at a very fast pace compared to any other analytical technique. Mass spectrometry has now become a powerful and robust tool for characterization of various aspects of proteins including intact mass, amino acid sequencing, PTM profiling, non-covalent complexes, aggregation propensity, and conformational integrity.[32, 33]

In a typical ESI MS experiments proteins are buffer exchanged into a solution containing 1:1 mixture of water and organic solvent. Under these conditions proteins are unfolded and the multiply charged gas phase ions generated by ESI can be used to

measure mass accurately. High throughput hyphenated techniques (RPLC/MS or RPLC/MS/MS) are now being routinely used for detection of intact mass and peptide sequencing. Intact mass analysis with high accuracy allows identification of proteins and the presence of any modification. Peptide sequencing of proteins is carried out by using either bottom-up or top-down approach. In classical bottom-up method protein is digested using a protease, and the proteolytic peptides generated are analyzed by LC/MS or tandem MS/MS. In top-down method protein digestion step is completely eliminated, and the ions for sequencing are produced by fragmenting the intact multiply charged protein ions in gas phase. Both of these techniques (bottom-up and top-down MS) have been successfully applied for sequencing, localization and quantitation of PTMs in protein therapeutics.

Native electrospray mass spectrometry (ESI MS)

Unlike classical scheme native ESI MS can analyze native-like quaternary protein structure and non-covalent protein/receptor complexes by using near native solvents such as ammonium acetate.[34] It has been an emerging technique and has been recently shown to characterize protein assemblies extending to sub-million daltons.[35]

In addition to profiling mass, native ESI MS can also be used to probe large-scale conformational changes in protein by monitoring ionic charge state distribution in mass spectrum. Native proteins are compactly folded and therefore gas phase ions generated by ESI carry relatively fewer charges and show a narrow ionic distribution at high m/z region. However, unlike folded species ions generated from unfolded protein carry a larger number of charges and populate at lower m/z region with a broad ionic charge

state distribution.[36] A bimodal ionic distribution will be detected under conditions when both compact state (highly structured native-like) and non-native states are present in equilibrium.

Hydrogen deuterium exchange mass spectrometry

Hydrogen deuterium exchange mass spectrometry (HDX) is a labelling technique for characterization of structure, dynamics, and the presence of multiple protein conformers in solution. This technique works by monitoring the exchange of backbone amide hydrogen with deuterium when exposed to deuterated water (D_2O).[37] Amide hydrogens of unstructured regions will show a rapid exchange compared to structured region, which are protected from HDX resulting in slow exchange. Mass shifts from this deuterium labelling can be easily measured using MS-based peptide mapping. HDX experiments have been used for mapping epitopes in antigen-antibody interactions,[38] to map protein-receptor interactions,[32, 39] and to compare batch-to-batch variation in protein therapeutics.[40]

1.4.3. Light scattering based methods

Dynamic light scattering (DLS)[41] and multiangle laser light scattering (MALLS)[42] are the two light scattering techniques used for detection of higher molecular weight species/aggregates in protein therapeutics. Hydrodynamic radius is obtained from DLS, which is directly correlated with the measured diffusion coefficient.

MALLS determines weight-average molecular weight (M_w) and radius of gyration (R_g) of proteins from the angular dependence of scattered light. This technique is used over static light scattering, when R_g is larger than angle of incident light. A typical MALLS

instrument consists of multiple detectors placed around the flow cell at different angles with respect to incident beam. Rayleigh ratio (R_θ) measured from the scattered light intensity at different scattering angles (θ) can be used to extrapolate M_w and R_g using Zimm's equation. MALLS is usually used in combination with SEC separation, and can detect protein aggregates even at very low concentration.[43]

1.4.4. Circular dichroism

Circular dichroism (CD) spectroscopy of proteins is based on differential absorption of left- and right- polarized light in 190 – 350 nm range, which is based on structural asymmetry. CD measurements in far UV region 190 – 250 nm are used for analysis of secondary structure of proteins, and measurements in near UV region 250 – 300 nm gives information about arrangement of the aromatic side chains and/or disulfide bridges within the protein. CD experiments are simple to perform, but give low resolution information.[44]

1.4.5. Field flow fractionation

Field flow fractionation (FFF) is a separation based technology with application in characterization of protein aggregates (0.01 - 50μm in size). In this method an external field applied perpendicular to the channel flow control retention of analytes, and separation occurs based on diffusion coefficients.[45] Smaller molecules elute earlier than larger molecules. This technique is advantageous as analysis can be carried out directly in the formulation buffer, but the complexity of instrumentation and data handling have limited the widespread use of this technique.

1.5. Commonly used methods to analyze protein interactions

1.5.1. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a label free, solution based method to get equilibrium dissociation constant (K_D), stoichiometry, enthalpy, entropy, and Gibbs free energy of protein interactions.[46] It is based on measurement of heat evolved or absorbed upon complex formation, when ligand is titrated gradually. Slope and inflection point of the binding isotherm obtained from ITC can be used to derive K_D and stoichiometry. Even though interactions with K_D values of 10^{-9} to 10^{-6} M can usually be analyzed using ITC experiments, concentration of the protein concentration used in titration needs to be optimized to get the sigmoidal curve. For low affinity interactions high concentration of proteins needed for ITC can lead to aggregation. Another limitation of ITC is that it not a high throughput technique.

1.5.2. Surface plasmon resonance

Surface plasmon resonance (SPR) is a widely used technique for qualitative and quantitative characterization of protein interactions.[47] In this technique one of the binding partners is immobilized on the sensor and the binding information is obtained by changes in the refractive index of the solvent near the sensor surface upon association or dissociation of protein complexes. SPR is advantageous as it requires lesser amount of proteins than ITC to measure interactions in the similar K_D range. However, immobilization of protein can lead to structural changes ultimately affecting binding data.

1.5.3. Analytical ultra centrifugation

Analytical ultra centrifugation (AUC) is a powerful technique to characterize protein solution structure and to determine binding constants of protein association.[48] It is based on real-time measurement of sedimentation of protein under centrifugal force along the radial position using either absorbance or interference optical systems. AUC experiments are operated either in sedimentation velocity (higher centrifugal field and completes in 1-2 hours) or sedimentation equilibrium mode (lower centrifugal field and takes 12 – 24 hours to complete). It can be used to derive molecular weight, oligomeric state, presence of aggregates, and equilibrium dissociation constant for protein association. Equilibrium dissociation constants in the range of 10^{-4} – 10^{-8} M can be detected using AUC; however aggregation of sample at higher concentration and loss of sample from sticking to sample chamber has to be carefully monitored. AUC data analysis for single species is straight forward, but complex data handling tools are required to analyze multi component species.[45]

1.5.4. Hummel-Dreyer method

Hummel Dreyer method is commonly used to study protein/small ligand binding, but can also be used for protein/protein interactions. In this method protein is dissolved in the buffer solution containing the same concentration of ligand used for equilibrating the SEC column. If there is binding between the protein and ligand a trough will be detected, and using the area of this trough unbound ligand concentration can be obtained to get equilibrium binding constant.[49] The limitations of this method include

requirement of large amount of sample, inability to obtain kinetic data, and non-applicability to study systems with slower binding kinetics.[50]

1.5.5. Frontal analysis continuous capillary electrophoresis

Capillary electrophoresis can be used to characterize binding of weak affinity (K_D above 10^{-6} M) as long as protein/ligand complex has different electrophoretic mobility compared to unbound protein. Frontal analysis continuous capillary electrophoresis (FACCE) involves continuous sampling and electrophoretic separation. The electropherogram obtained from FAACE analysis shows multiple plateaus depending on the components in the system, and the height of the first plateau is used to determine the concentration of free protein to plot binding isotherms. FAACE have been applied to deduce stoichiometry and binding constants of protein/polyelectrolyte interactions.[50, 51]

1.5.6. Affinity chromatography

Two methods based on affinity chromatography exist for qualitative analysis of protein interactions. The first method involved use of affinity purification and mass spectrometry for mapping protein interactions. This method involves tagging of a protein followed by affinity purification of tagged-proteins and identification of the eluted complexes using LC/MS/MS. This scheme failed to detect transient protein complexes unless additional crosslinking is utilized.[52]

The second affinity chromatography based method was able to detect transient protein complex. In this elution profiles were used to confirm binding (transient binding of protein to the immobilized binding partner resulted in asymmetric elution profile).

This technique was able to detect transfer of retinoic acid to retinoic acid receptor as it formed a transient complex with immobilized holo CRABP II.[53]

1.5.7. Fluorescence based methods

The two most common fluorescence based techniques to quantitate protein interactions in vivo include fluorescence resonance energy transfer (FRET) and bimolecular fluorescence complementation. FRET is a non-radiative energy transfer between an excited fluorophore (donor) and another fluorophore (acceptor), which have overlapping emission/absorption spectra.[54] For interacting proteins FRET occurs when one protein labelled with an acceptor chromophore is in close proximity (10-80 Å) with other protein labelled with donor chromophore. Bimolecular fluorescence complementation is based on association of complementary protein fragments fused to interacting domains of two proteins and give out a measurable fluorescence signal.[55] A major limitation of both of these techniques is it requires attachment of reporter molecule to proteins, which could affect the structure of protein leading to inaccurate binding data.

1.6. Objectives

Biotherapeutics containing biopolymers have emerged to be very successful for the treatment of chronic diseases. However, the increased heterogeneity of biotherapeutics compared to small molecule drugs requires powerful and robust analytical tools to characterize biotherapeutics at various levels. This work reports multiple novel analytical methods that utilize synergy of non-denaturing chromatography separation (IXC, SEC) and native ESI MS detection to characterize

biopolymer structure and interactions. Specifically the objectives include: (i) develop SEC with online native ESI MS detection to characterize higher order structure of proteins (oligomeric states and aggregates), and probe conformational integrity (chapter 2); (ii) develop a label free method to measure kinetics and binding affinity of transient protein interactions using SEC/native ESI MS (chapter 3); (iii) develop IXC with online native ESI MS detection to characterize conformational integrity and positional isomers of protein conjugates, and glycoforms of a biotherapeutic. Implement top-down MS/MS in this online scheme for modifications wherein MS1 analysis could not provide unequivocal confirmation (chapter 4); (iv) develop a IXC/ESI MS method to characterize highly heterogeneous PEGylated glycoprotein (chapter 5).

CHAPTER 2

CHARACTERIZATION OF SMALL PROTEIN AGGREGATES AND OLIGOMERS USING SIZE EXCLUSION CHROMATOGRAPHY WITH ONLINE DETECTION BY NATIVE ELECTROSPRAY IONIZATION MASS SPECTROMETRY

Peer-reviewed article published: Muneeruddin, K.; Thomas, J. J.; Salinas, P. A.; Kaltashov, I.A.; Characterization of Small Protein Aggregates and Oligomers Using Size Exclusion Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry *Analytical Chemistry,* **2014**, 86, 10692 - 10699

2.1. Overview

Protein self-association is a key process in a variety of diverse phenomena ranging from the highly ordered assembly of multiunit proteins into their functional states to mostly chaotic aggregation. The latter gained considerable notoriety in the past several decades due to its obvious role in the etiology of the so-called conformational diseases, such as Alzheimer's and Parkinson's.[56-58] Renewed interest in protein aggregation in the biotechnology area is caused by the advent of protein therapeutics, where aggregation affects not only the economy of the production process but is also a cause of grave concerns vis-à-vis safety of biopharmaceutical products.[59, 60] The attention in this field has been traditionally focused on relatively large subvisible aggregates, while small soluble oligomers were typically considered as transient precursors to catastrophic aggregation. However, recent studies point out that even reversible formation of very small oligomers (such as noncovalent dimers of monoclonal antibodies) may have very negative consequences by interfering with the drug delivery, which is triggered by increased formulation viscosity.[61] The importance of protein aggregation places a premium on developing robust analytical methods capable

of detecting and characterizing protein oligomers across the entire aggregation spectrum. Size exclusion chromatography (SEC) and, to a lesser extent, analytical ultracentrifugation (AUC) are the most popular analytical tools capable of detecting lower molecular weight oligomers, although their utility is often limited by inadequate resolution.[62] Light scattering[42, 43, 63] and, more recently, small-angle X-ray scattering (SAXS)[64] are also enjoying popularity in this field, although data interpretation is frequently model-sensitive, which limits their applicability to complex heterogeneous systems.

Native electrospray ionization mass spectrometry (ESI MS) is another technique that enjoyed rapid growth of popularity in the studies of protein associations in the past two decades,[34, 65, 66] with several recent examples highlighting its utility for the analysis of small molecular weight aggregates of therapeutic proteins. In addition to providing information on the stoichiometry of noncovalent protein complexes in solution, native ESI MS allows the structural heterogeneity of protein therapeutics to be evaluated (by analyzing mass profiles) and conformational integrity to be assessed (by analyzing ionic charge state distributions).[32] However, using native ESI MS to detect and especially quantify noncovalent associations is not always straightforward, as artifacts are known to be frequently produced during the ESI process, e.g., due to forced protein self-association during the final stages of charged droplets life when solvent evaporation may result in a dramatic increase of the protein concentration.[67] Other factors (e.g., nonlinear dependence of the response factors on solution composition) may also complicate the quantitation efforts.[68] Finally, presence of multiple oligomers in solution

usually results in crowded mass spectra, where overlapping peaks representing different oligomers can complicate the data interpretation and make it nearly impossible to distinguish contributions from different species.

The ESI MS analysis of protein oligomers can be aided by SEC separation of small aggregates based on their physical size. Such analyses can be carried out off-line by collecting SEC fractions prior to their analysis,[69, 70] but the results of such measurements can be compromised due to the possibility of the composition changes of collected fractions prior to MS analyses. For example, Heck and co-workers demonstrated that the transient nature of the small aggregates of monoclonal antibodies may result in their dissociation following fractionation by SEC;[69] the opposite process where additional protein oligomerization in collected fractions is triggered by desalting and preconcentration is also possible.[71] These problems can be circumvented by using online MS detection for ESI MS, but because native ESI MS is typically less sensitive compared to conventional ESI MS, acquisition of meaningful data on the chromatographic time scale remained problematic until recently. As a result, the online SEC/ESI MS detection scheme was only used for measuring the mass distribution of heterogeneous single-chain macromolecules, such as synthetic polymers,[72, 73] antibodies[74] and antibody-drug conjugates,[75] while the unique ability of native ESI MS to detect noncovalent biopolymer complexes (such as an antibody–antigen complex) has not been utilized in such analyses. However, the continuous improvement in ESI MS hardware has made it possible to carry out measurements under near-native conditions using only minute quantities of the analyte and relatively short acquisition time windows. We are taking

advantage of these expanded capabilities of native ESI MS to use it as an online detection tool for the analysis of noncovalent protein complexes (including both small soluble aggregates and highly organized self-assemblies). This combination of two powerful analytical techniques not only allows complex and heterogeneous protein mixtures to be profiled vis-à-vis the presence of protein oligomers (based on both retention time and mass information), but also enables assessment of protein conformational integrity (based on ionic charge state distributions in ESI mass spectra). Online detection with native ESI MS also allows meaningful information to be extracted from SEC chromatograms for protein assemblies that undergo rapid dissociation/reassociation on the chromatographic time scale.

2.2. Materials and Methods

Bovine serum albumin (BSA) and human transferrin (Tf) were purchased from Calbiochem (La Jolla, CA) and Sigma-Aldrich Chemical Co. (St. Louis, MO), respectively. A recombinant form of human arylsulfatase A (rhASA) was provided by Shire (Lexington, MA). All other chemicals used in this work were of analytical grade or higher. BSA and Tf solutions for online SEC/MS analysis were prepared by dissolving the protein powder directly in 100 mM ammonium acetate without any desalting step. Analysis of rhASA was carried out by diluting the drug substance sample in 100 mM ammonium acetate to a required concentration (0.29 to 2.3 mg/mL), and pH adjusted to a desired level using formic acid.

SEC separations were carried out using Agilent 1100 HPLC (Agilent Technologies, Santa Clara, CA) equipped with a Waters (Milford, MA) Biosuite Ultra high resolution

column (4.6 × 300 mm). Typically, 100 µL of a 1 mg/mL protein solution was injected unless stated otherwise, and a 100 mM ammonium acetate solution (pH adjusted to a desired level using either formic acid or ammonium hydroxide) was used as the mobile phase in all separations at a flow rate of 0.1 mL/min. Upon exiting the column, the eluate was directed to a UV absorption detector (operated at 280 nm) followed by QStar-XL (ABI-Sciex, Toronto, Canada) hybrid quadrupole/time-of-flight mass spectrometer equipped with a standard ESI source. The ESI source conditions were optimized to provide a stable spray and optimal ion desolvation: drying gas, 30 L/min; nebulizing gas, 39 L/min; curtain gas, 22 L/min; declustering potential on the skimmer, 290 V; declustering potential on the orifice, 160 V. The source temperature was maintained in the 200–250 °C interval. Collisional cooling in the ion guide region (a gas flow restricting sleeve in Q0) was used to enhance focusing and stability of noncovalent complexes.

2.3. Results and Discussion

2.3.1. Conformational stability and aggregation propensity of a commercial protein

The ESI mass spectrum of bovine serum albumin (BSA) acquired under near-native conditions (Figure 2.1) features a convoluted charge state distribution for ions whose masses are consistent with BSA monomer (instead of a single mass, a distribution of masses is observed, consistent with extensive nonenzymatic post-translational modifications of the protein). A bimodal character of the charge state distribution of the monomeric ions clearly suggests that conformational integrity of BSA is compromised: either a fraction of the protein molecules is partially unfolded, or all BSA molecules

undergo transient partial unfolding (native ESI MS alone cannot distinguish between these two possibilities). The presence of partially unfolded protein molecules in the BSA sample is not surprising, given the extent of nonenzymatic post-translational modification of this protein (see the convoluted shapes of peaks of BSA ions in the mass spectrum acquired under denaturing conditions, inset in Figure 2.1), which are known to change the physical properties of proteins effecting the energy barrier between folded and unfolded states (a detailed analysis of various forms of BSA from the commercial sample reveals the occurrence of several nonenzymatic post-translational modifications, such as cysteinylation, glycation, formation of pyridoxyl phosphate adducts, and several others, as shown in Figure S1 in the Supporting Information). Furthermore, production of commercial BSA was likely to include a thermal serum pretreatment process, which is usually carried out to remove any bacterial or viral contamination in the sample.

Alternatively, the presence of partially unfolded BSA molecules in the native ESI mass spectrum can be explained by invoking the notion of gas phase dissociation of larger oligomeric ions. Should dissociation of metastable protein aggregates occur in the gas phase, asymmetric charge partitioning is likely to generate highly charged monomeric ions, a phenomenon reported in the past for other aggregation-prone proteins. Ionic peaks corresponding to protein dimers and higher oligomers are indeed observed in the higher m/z range (above 5000 u) of the BSA mass spectrum, and they are also present at lower m/z range (below 3500 u). It is not clear, however, if these ionic species represent dimers and higher oligomers existing in the protein solution at equilibrium with monomers, or if their presence in the mass spectrum is an artifact

related to forced association of proteins in rapidly shrinking electrosprayed droplets documented by Klassen and co-workers. The size exclusion chromatogram of BSA (gray trace in Figure 2.2) contains three peaks whose retention times are consistent with monomeric (31.5 min), dimeric (27.6 min) and trimeric (25.7 min) forms of BSA, although the earlier-eluting peaks could also be interpreted as partially unfolded protein monomers. Therefore, neither native ESI MS nor SEC alone can provide unequivocal characterization of the protein sample which is prone to both partial unfolding and oligomerization.

The uncertainties regarding the data interpretation are readily resolved by combining SEC separation with online native ESI MS detection. The appearance of the total ion chromatogram (TIC) is consistent with the shape of the UV trace in that three partially resolved peaks can be identified (Figure 2.2); however, the ability of MS to identify ionic species giving rise to each TIC peak allows this chromatogram to be readily interpreted. For example, only trimeric protein species are present in the earliest eluting peak, allowing it to be assigned as a BSA trimer. Importantly, both high- and low-charge density trimer ions are observed in the mass spectrum averaged across this peak (see inset in Figure 2.2), indicating that BSA trimers undergo reversible and rapid (on the SEC time scale) loss of compactness. A very similar behavior is exhibited by the protein dimers, where both low-charge density ions (m/z > 5000 u) and high-charge density ones (m/z < 3000 u) are prominent in the mass spectrum averaged across the second SEC peak. No BSA monomers can be observed in either of the spectra averaged across

the early eluting peaks, clearly indicating that the transient loss of higher order structure within both dimers and trimers does not result in their dissociation.

The most prominent ionic species in the mass spectrum averaged across the latest eluting SEC peak corresponds to BSA monomers. The charge state distribution is bimodal, indicating presence of both compact and partially unfolded proteins. Importantly, the extracted ion chromatograms for both high- and low-charge density monomeric ions exhibit virtually indistinguishable elution profiles (right inset in Figure 2.2), which unequivocally demonstrates that the two forms of the protein are in rapid equilibrium with each other, and their separation cannot be achieved by SEC alone. In addition to the BSA monomers, the mass spectrum averaged across the latest eluting peak contains contributions from oligomers. Because the extracted ion chromatogram for these ions closely follows that of the monomeric ions in the 30–35 min time window, these oligomers represent nonspecific association of proteins in rapidly shrinking electrosprayed droplets in the ESI interface, a well-known artifact of ESI MS.[76] Thus, SEC with online native ESI MS detection not only provides a convenient way for detecting and characterizing all protein species present in solution but also allows the ESI MS artifact to be readily identified and eliminated from the analysis.

A very intriguing question that arises in connection with the analysis of BSA oligomers is the nature of the protein/protein association that leads to their formation. Although it is frequently assumed that these species are formed through noncovalent interactions, the possibility of covalent interactions being involved in the early stages of BSA aggregation cannot be excluded. Indeed, BSA has one free cysteine residue, which

might participate in formation of external disulfide bonds (unless protected by cysteinylation, as seen in at least some BSA molecules, see Figure S1 in the Supporting Information). One can also argue that formation of higher oligomers may proceed through disulfide scrambling leading to formation of multiple external disulfide bonds. However, the ESI mass spectrum of BSA acquired under denaturing conditions does not contain signals corresponding to protein dimers and tetramers, indicating the reversible nature of the association process. Furthermore, dissociation under nonreducing conditions is also evident in SEC chromatograms of oligomeric fractions of BSA (see Figure S2 in the Supporting Information).

It is interesting to note that albumin in fresh human plasma does not exhibit many features that are prominent in the mass spectra of commercial BSA. For example, we note that the high-charge density peaks are absent from the ESI mass spectrum acquired under near-native conditions, and the protein mass profile does not exhibit the same high level of heterogeneity as BSA (see Figure 2.3). The two major forms correspond to intact protein and cysteinylated protein,[77] with all other forms being relatively minor contributors. The mass difference between these two forms of albumin is sufficient enough to allow extracted ion chromatograms for each of them to be plotted separately, without the interference from the other form. The overlaid plots of the two extracted ion chromatograms (for intact and cysteinylated human serum albumin) display identical elution profiles, indicating that cysteinylation of the only free cysteine residue in albumin does not compromise its conformational integrity (Figure 2.3). Unfortunately, the multiplicity of nonenzymatic PTMs observed in the commercial

BSA sample does not allow adequate resolution to be achieved under the native conditions to afford selective monitoring of these species in SEC/MS.

We note that the online analysis of BSA aggregates carried out in this work using native ESI MS as a detection tool can also be accomplished using other tools, such as multiangle light scattering (MALLS).[63] However, as can be seen from the following sections, SEC with online detection by native ESI MS has a unique advantage of being able to provide meaningful information when analyzing coeluting proteins of different masses, a task that remains very challenging for MALLS. One potential concern related to the analysis of protein aggregates by SEC/MS is the use of volatile electrolytes (such as ammonium acetate) as solvent systems. Although long-term exposure of proteins to ammonium acetate or similar solvent systems may have the potential to affect their aggregation propensity, the short time required for a single SEC/MS experiment makes such changes unlikely. Indeed, comparison of aggregation profiles extracted from SEC chromatograms acquired with ammonium acetate and a conventional phosphate buffer as mobile phases revealed no differences outside of the experimental error (see Figure S3 in the Supporting Information). The distribution of protein aggregates also appears to be insensitive to the SEC flow rate (see Figure S4 in the Supporting Information), giving indication that the measurements are unlikely to be affected by the shear flow around the stationary phase particles.

2.3.2. Analysis of coeluting proteins

Analysis of protein mixtures where the difference in molecular weight (or, more correctly, hydrodynamic radii) does not allow adequate chromatographic resolution to

be achieved is a frequently encountered problem in SEC. The ability of online SEC/native ESI MS to provide meaningful information for such systems was illustrated by analyzing the equimolar mixture of BSA and human serum transferrin (Tf). The molecular weights of these two proteins (66 and 80 kDa, respectively) are too close to each other to allow satisfactory separation to be achieved using SEC. Indeed, the SEC chromatogram (UV trace) shows a large peak with a convoluted shape consistent with the presence of two abundant proteins having very close elution times (Figure 2.4, left panel). Although the SEC alone fails to provide a clear distinction between BSA and Tf, online detection with native ESI MS allows the elution profiles of both proteins to be determined (see the extracted ion chromatograms for the +16 charge state of Tf and the +14 charge state of BSA in Figure 2.4). The data presented in Figure 2.4 are also notable for another reason: unlike BSA, ionic charge state distribution of Tf does not exhibit bimodal character, with only low-charge density ion peaks present in the mass spectra (e.g., see the lower trace in the right panel of Figure 2.4). This suggests that partially unfolded Tf molecules are not present in the sample; incidentally, the SEC/MS data also fails to reveal the presence of oligomeric forms of Tf (only BSA dimers and trimers could be observed, as shown in Figure 2.4, right panel).

Another important observation from the analysis of extracted ion chromatograms of BSA and Tf is that the elution order is not the same as the molecular weight order (Tf is larger). Furthermore, Tf has a bilobal structure,[78] whereas albumin has a globular fold,[79] which should make the difference in hydrodynamic radii even larger. Nevertheless, the extracted ion chromatograms for BSA and Tf monomers clearly

show shorter elution time for the former. Although it is possible that the longer retention time of Tf molecules may be related to their interactions with the column packing material beyond purely mechanical interactions (filtration), a more likely cause for this anomalous behavior is the increase of the average hydrodynamic radius of BSA monomers due to their transient unfolding during the separation process.

2.3.3. Characterization of protein self-association: arylsulfatase A

Both proteins considered in the previous sections (BSA and Tf) are monomeric, and formation of oligomeric species is a clear sign of degradation. However, ordered noncovalent association of several polypeptide chains is frequently required in order to endow proteins with specific properties. Formation of the requisite quaternary structure is frequently reversible, as is the case with arylsulfatase A, a critical lysosomal enzyme that exits as a homodimer at neutral pH, but must assemble into octamers at mildly acidic pH to avoid degradation by lysosomal proteases.[22, 80] Recently, we demonstrated that SEC correctly identifies homodimeric and homo-octameric arrangements as quaternary structures of the recombinant form of human aryl sulfatase A (rhASA) at pH 7.0 and 5.0, respectively.[81] However, SEC analysis of rhASA at intermediate pH (between 6.1 and 6.9) failed to provide conclusive results, yielding only broad unresolved peaks whose profiles were both pH- and concentration-dependent, and the average elution time did not match that of either dimers or octamers. At the same time, native ESI MS analysis carried out within this intermediate pH region revealed the presence of only dimeric and octameric species of rhASA,[81] suggesting that the two forms of the protein coexist at a dynamic equilibrium, and their interconversion (dissociation/reassociation)

is very fast on the chromatographic scale, which does not allow these species to be distinguished from one another by SEC.

Online SEC/ESI MS analysis of rhASA at the two terminal pH points (5.0 and 7.0) confirms that the protein exists as an octamer and a dimer, respectively (Figures 2.5 and 2.6). Interestingly, the single peak in SEC chromatogram acquired at pH 7.0 (elution time 30 min) contains contributions from both dimeric and octameric ions, although their elution profiles are indistinguishable (compare the extracted ion chromatograms for the two groups of ions in Figure 2.6). This strongly indicates that the octameric ions of rhASA observed at neutral pH do not reflect protein association in solution, but rather represent the artifacts related to the forced association in the ESI interface, similar to what had been observed for BSA (vide supra).

SEC/ESI MS analysis of rhASA carried out at intermediate pH 6.4 yields a total ion chromatogram that also contains contributions from both octameric and dimeric ions (Figure 2.7), although their respective extracted ion chromatograms differ from each other in a very dramatic way. The elution profile of the octameric protein species is very close to that recorded under the mildly acidic conditions (Figure 2.5). In contrast, the elution profile of the dimers has a convoluted form with the elution front coinciding with that of the octameric species, and tailing all the way to the elution window observed for rhASA dimers at neutral pH (Figure 2.6). This significant dissimilarity of the elution profiles of rhASA dimers and octamers observed at intermediate pH indicates that the octameric species dissociate as they move through the column. Indeed, if we denote the linear velocity of stable octamers and dimers as u_8 and u_2 ($u_8 > u_2$), which

relate to the column length L and the observed elution times T_8 and T_2 (e.g., those seen in Figures 2.5 and 2.6) as $u_k = L/T_k$, then the elution time of the dimer t_2 produced upon dissociation of the octamer at time τ_{diss} ($\tau_{diss} < T_8$) can be calculated as

$$t_2 = \tau_{diss} + \frac{L - \tau_{diss} \cdot u_8}{u_2} = T_2 - \tau_{diss} \cdot \left(\frac{T_2}{T_8} - 1\right) \quad (1)$$

In other words, the elution times of the dimers formed on the column will spread between T_8 and T_2. The shape of this distribution can be used to determine the kinetics of the octamer dissociation (e.g., to establish the distribution of τ_{diss}), although no such calculations were performed in this work due to insufficient chromatographic resolution.

Dissociation processes alone do not account for all of the chromatographic features observed in Figure 2.7: even though the extracted ion chromatogram of the octameric species has a sharp frontal edge (similar top that seen in Figure 2.5), there is a small (but measurable) delay in the elution. Furthermore, the shape of this chromatogram is asymmetric, as a slight tailing toward longer elution times is observed (a feature that was absent in the octamer elution profile in Figure 2.5). These two phenomena cannot be explained without invoking the notion of reassociation of the products of dissociation (dimers generated during the chromatographic run) to yield octamers. A simplified mathematical treatment of the reassociation process can be carried out by assuming that a new octamer is formed via association of four dimers, all of which were produced at a single time point τ_{diss} (i.e., by ignoring the possibility of multiple

dissociation/reassociation events throughout a single chromatographic run). In this case, the elution time of the newly formed octamer can be estimated as

$$t_8^* = \tau_{diss} + \tau_{reass} + \frac{L - \tau_{\mathrm{diss}} \cdot u_8 - \tau_{\mathrm{reass}} \cdot u_2}{u_8} = T_8 - \tau_{reass} \cdot \left(1 - \frac{T_8}{T_2}\right) \quad (2)$$

where τ_{reass} is the time interval between formation of the dimers and their reassociation back to an octamer. Of course, a more rigorous mathematical treatment should take into account the possibility of multiple dissociation/reassociation events, which would require numerical simulations. Nevertheless, the simple algebraic expression 2 is useful and informative as a way to rationalize the tailing of the elution profile of rhASA octamer in Figure 2.7, and might be used to obtain a first-order approximation of the distribution of τ_{reass}, especially under conditions when the extent of the dissociation during the chromatographic run is limited.

Because the kinetics of the reassociation process should be concentration dependent, one would expect to observe a shift in the elution front of the octameric species following the decrease of protein concentration. Decrease in rhASA concentration from 2.3 to 1.1, 0.55 and 0.29 mg/mL (while keeping the acidity constant at pH 6.4) resulted in a progressive decrease of the relative abundance of the octameric species and also led to a noticeable change of their elution profiles (shift of the elution front to longer times and enhanced peak tailing, see Figure 2.8), consistent with the expected equilibrium shift toward the dissociation products when the overall protein concentration goes down. A similar trend was observed under more acidic conditions (pH 6.1), where total protein concentration reduction from 1.1 mg/mL to 0.29 mg/mL

resulted in a monotonic decrease of the relative ionic signal of octameric species and a shift of the octamer elution front (see Figure S5 in the Supporting Information).

It is important to note that even though the partial dissociation of rhASA octamers at pH 6.4 might be inferred from SEC alone by observing the trailing of the UV trace of SEC chromatograms to longer elution times (e.g., see the brown trace in Figure 2.7), it is the ability of MS to make a clear distinction between the dimeric and octameric ions that allows the underlying cause of this peak broadening to be readily established. We also note that the extent of peak broadening of the TIC is much more significant compared to the UV trace, as dissociation of an octamer to four dimers does not change the UV absorbance (as the total number of aromatic residues remains unchanged), while the ionic signal increases (as the number of molecules capable of producing ions increases 4-fold).

Online detection with native ESI MS also allows another interesting phenomenon to be observed, which relates to the uneven concentration profiles across chromatographic peaks and is very important vis-à-vis establishing the realistic limits for the accuracy of the SEC-based measurements. As has been already mentioned, kinetics of the reassociation process are concentration-dependent (unlike octamer dissociation, which is a unimolecular process). Therefore, it seems plausible that the octamer/dimer equilibrium will be shifted toward the latter not only at the “far end” of the chromatographic peak (as can be seen by comparing the elution profiles of octamers and dimers in Figure 2.7), but also at the “front end” of the chromatographic peak. Indeed, a careful analysis of the elution profiles of the two oligomeric species at the

elution front provides unequivocal evidence that the dissociation process is strongly favored during a very short time interval at the elution front (Figure 2.9). Indeed, while the extracted ion chromatogram of the octameric species displays a monotonic increase, reaches an apex and begins to decrease monotonically, the elution front of dimers has a very sharp edge followed by a rapid drop in intensity (with the minimum intensity for M_2 coinciding with the apex of the M_8 peak) before the "normal" behavior is observed. The width of the aberrant peak at the front of the elution profile of dimeric species defines the lower resolution limit for kinetic measurements that can be carried with the SEC/native ESI MS.

2.4. Conclusions

The SEC/native ESI MS analysis of proteins and their assemblies presented in this work provides an elegant way to characterize complex mixtures of proteins that include both monomeric species, their low molecular weight soluble aggregates and larger assemblies. A distinction can be readily made among the incompletely resolved proteins based on their mass differences, and the analysis of protein ion charge state distribution allows the conformational integrity to be assessed. The ability to resolve different protein assemblies based on their masses allows a meaningful analysis to be carried out when such species cannot be separated from one another by means of SEC alone due to their rapid interconversion on the chromatographic time scale. Importantly, the new technique does not require the protein samples to be purified/desalted prior to the analyses, which makes it very appealing as a robust and powerful tool for the analysis of biopharmaceutical products.

FIGURES

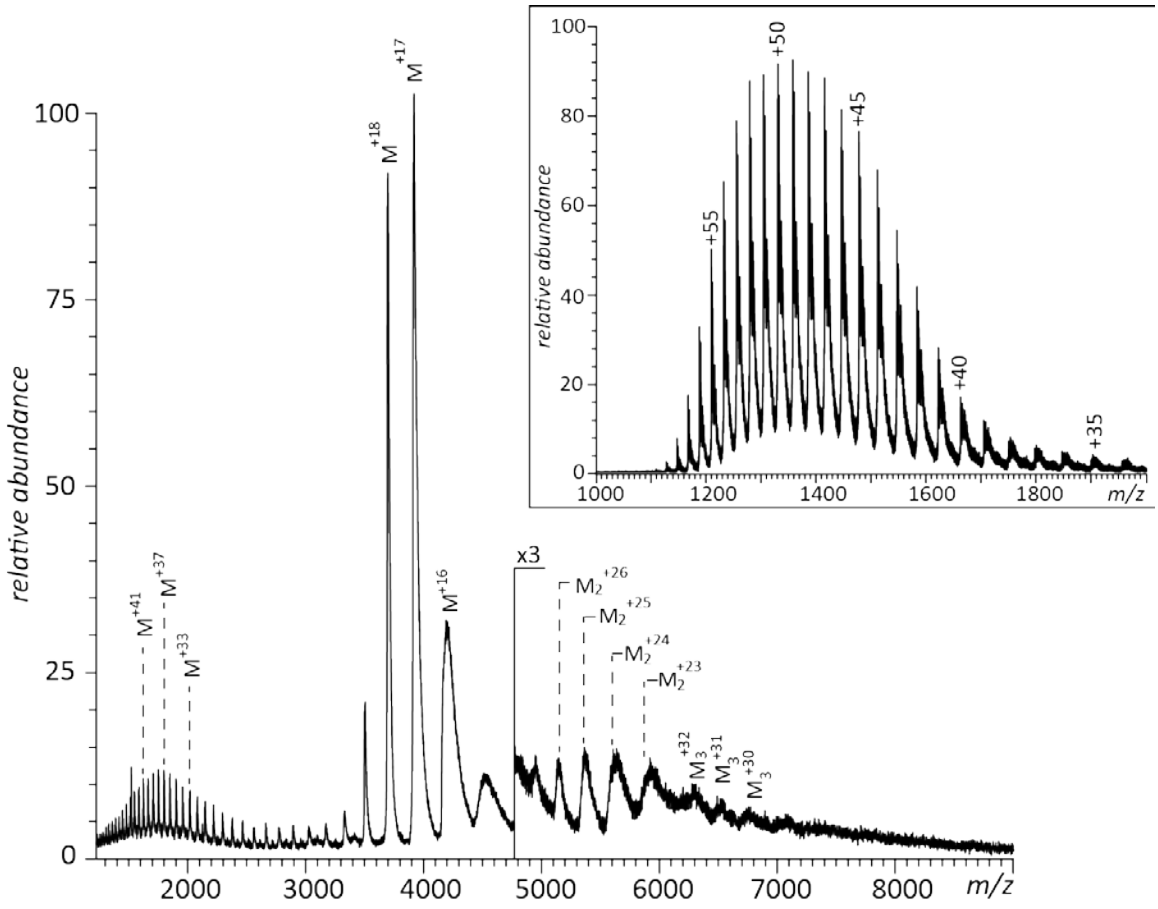

Figure 2.1 ESI mass spectrum of BSA acquired under near-native conditions without SEC separation of protein monomers and oligomers. The mass spectrum acquired under denaturing conditions is shown in the inset (only 1000–2000 m/z range is shown, as no protein ions were detected outside of this range).

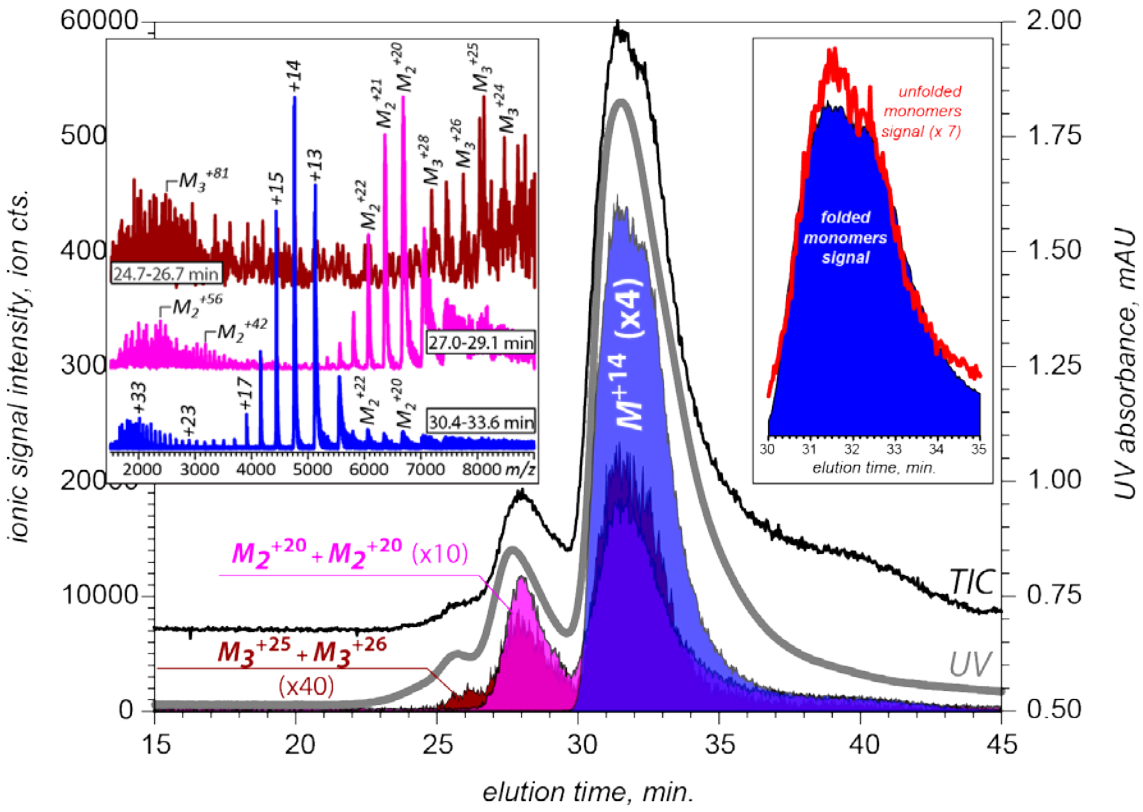

Figure 2.2 Online SEC/MS analysis of BSA: UV chromatogram (gray); TIC (black); extracted ion chromatograms of the BSA monomers (blue, +14 charge state), dimers (magenta, a sum of +20 and +21 charge states) and trimers (brown, a sum of +25 and +26 charge states). Inset on the left shows mass spectra averaged across the three SEC peaks (acquisition times are indicated on each trace). Inset on the right shows extracted ion chromatograms of compact and partially unstructured monomers (represented by low- and high-charge density ions).

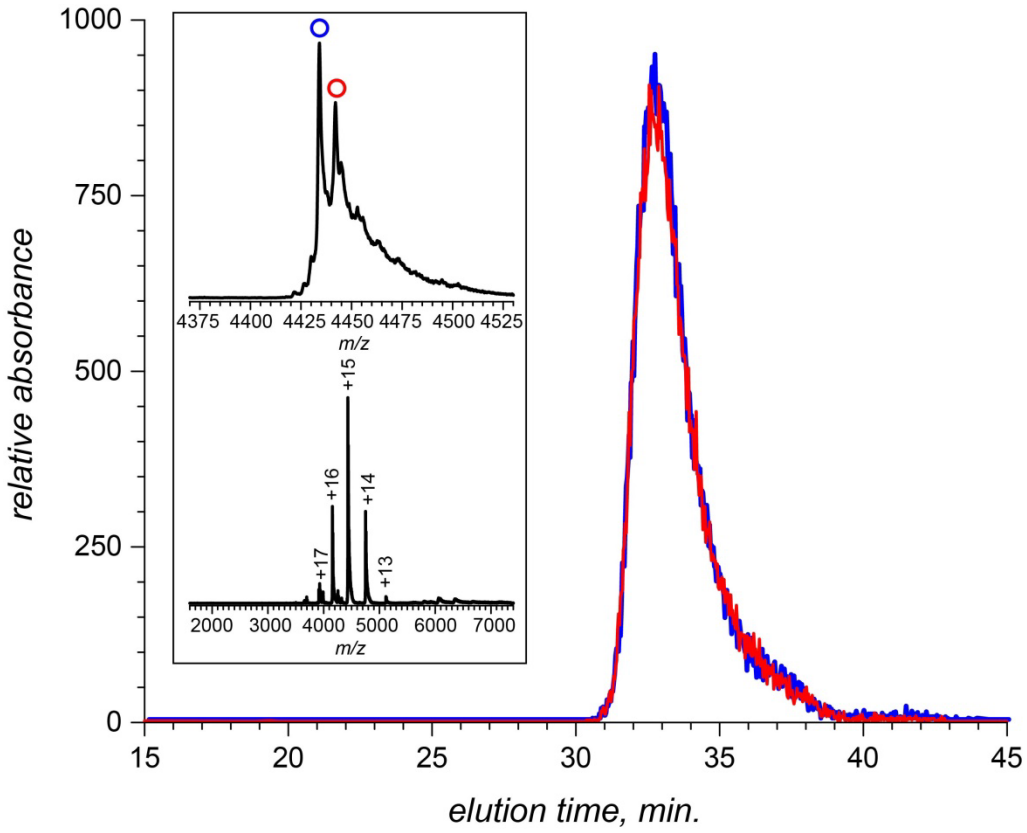

Figure 2.3 Extracted ion chromatograms for two forms of human serum albumin (intact, blue; cysteinylated, red) in blood plasma. The inset shows the spectrum of the protein averaged across the SEC peak and the zoomed view of the ion peak corresponding to +14 charge state showing the protein mass distribution.

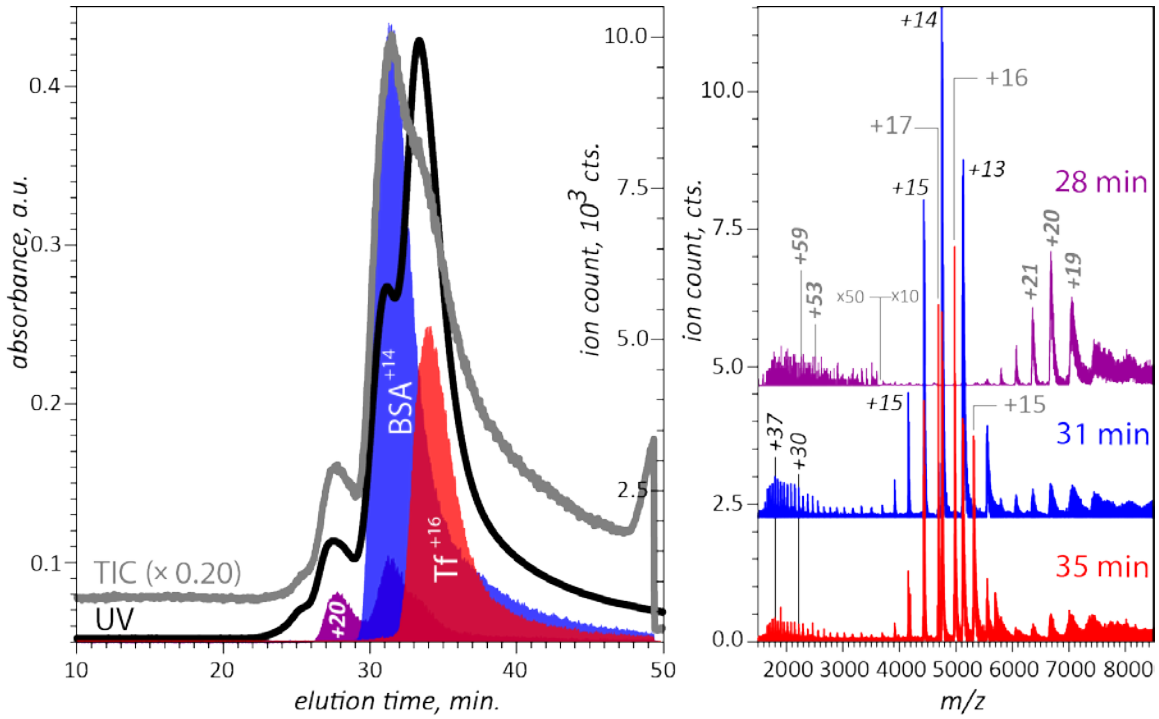

Figure 2.4 Online SEC/MS characterization of an equimolar mixture of BSA and Tf: UV chromatogram (black); TIC (gray); extracted ion chromatograms of the BSA monomers (blue, +14 charge state), BSA dimers (charge state +20, magenta) and Tf monomers (charge state +16, red). Inset shows mass spectra averaged across the three SEC peaks (acquisition times are indicated on each trace).

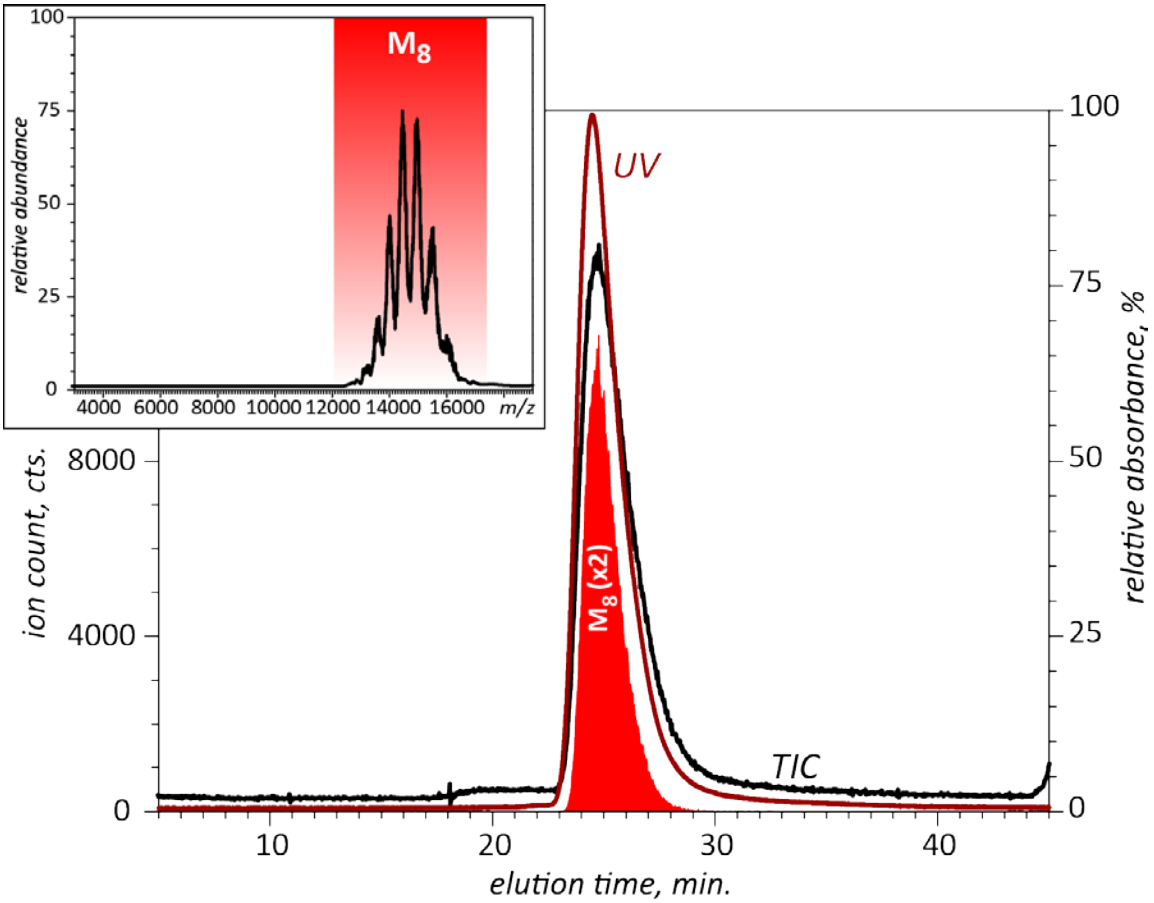

Figure 2.5 Online SEC/MS of rhASA (1.1 mg/mL) at pH 5.0: UV chromatogram (brown), TIC (black) and the cumulative extracted ion chromatogram of the octameric species (red). Inset shows a mass spectrum averaged across the SEC peak.

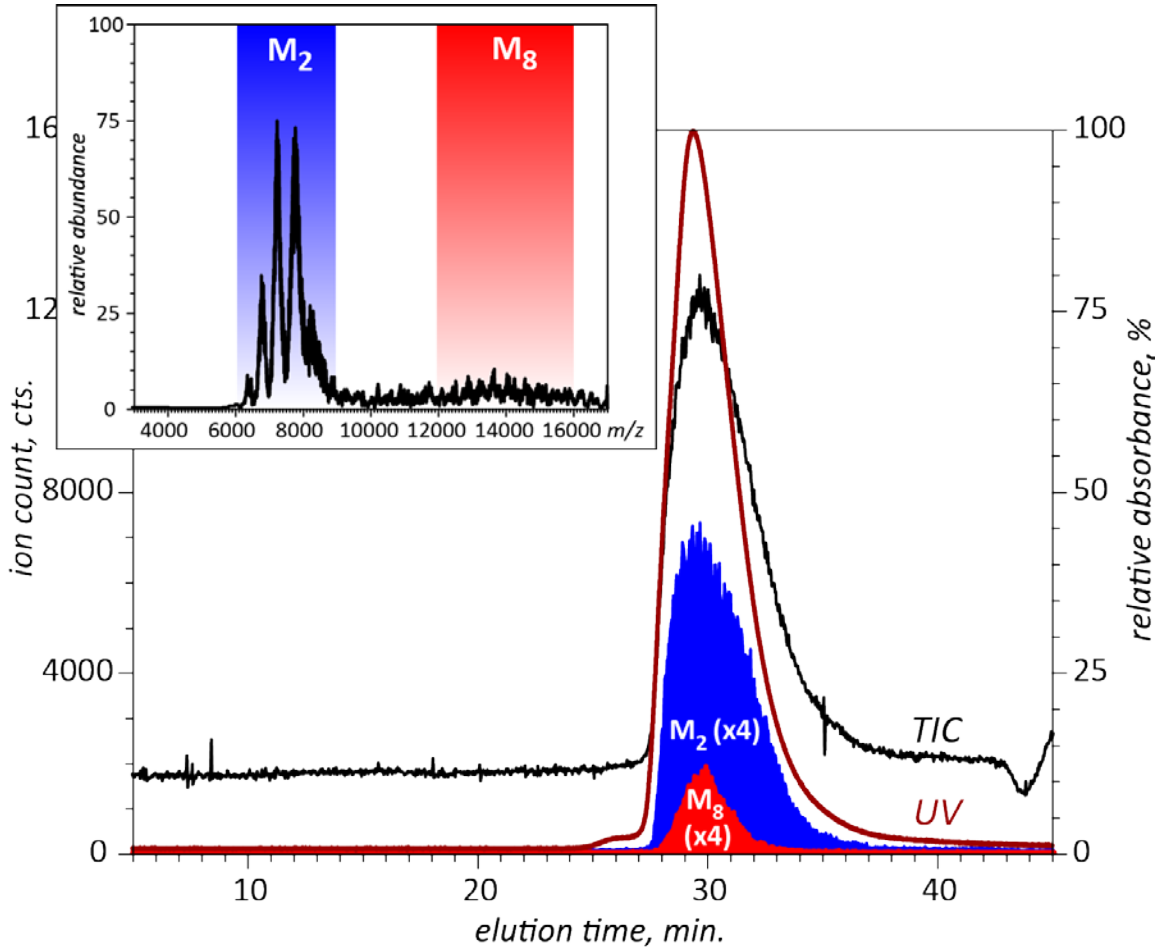

Figure 2.6 Online SEC/MS of rhASA (1.1 mg/mL) at pH 7.0: UV chromatogram (brown), TIC (black) and the cumulative extracted ion chromatograms of the dimeric (blue) and octameric (red) species. Inset shows a mass spectrum averaged across the SEC peak.

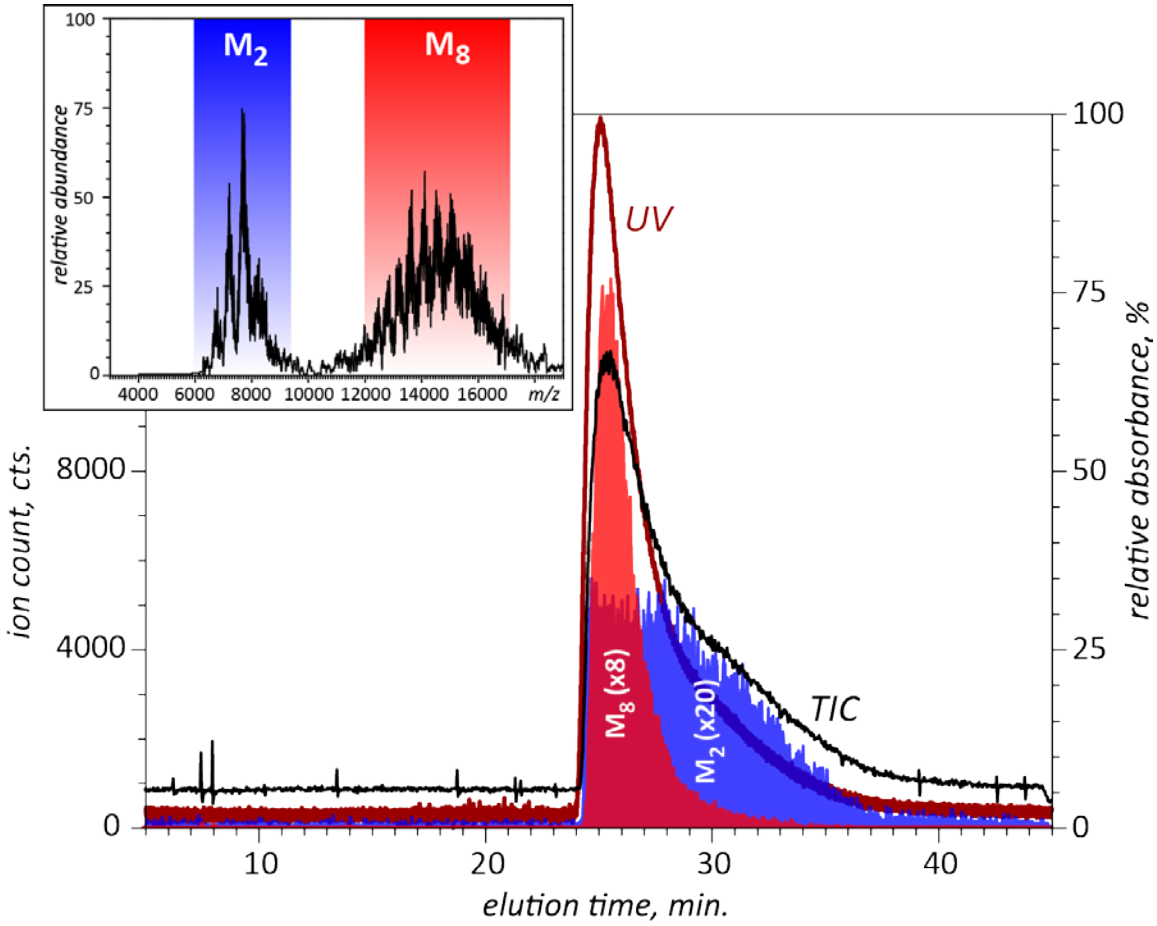

Figure 2.7 Online SEC/MS of rhASA (2.3 mg/mL) at pH 6.4: UV chromatogram (brown), TIC (black) and the cumulative extracted ion chromatograms of the dimeric (blue) and octameric (red) species. Inset shows a mass spectrum averaged across the SEC peak.

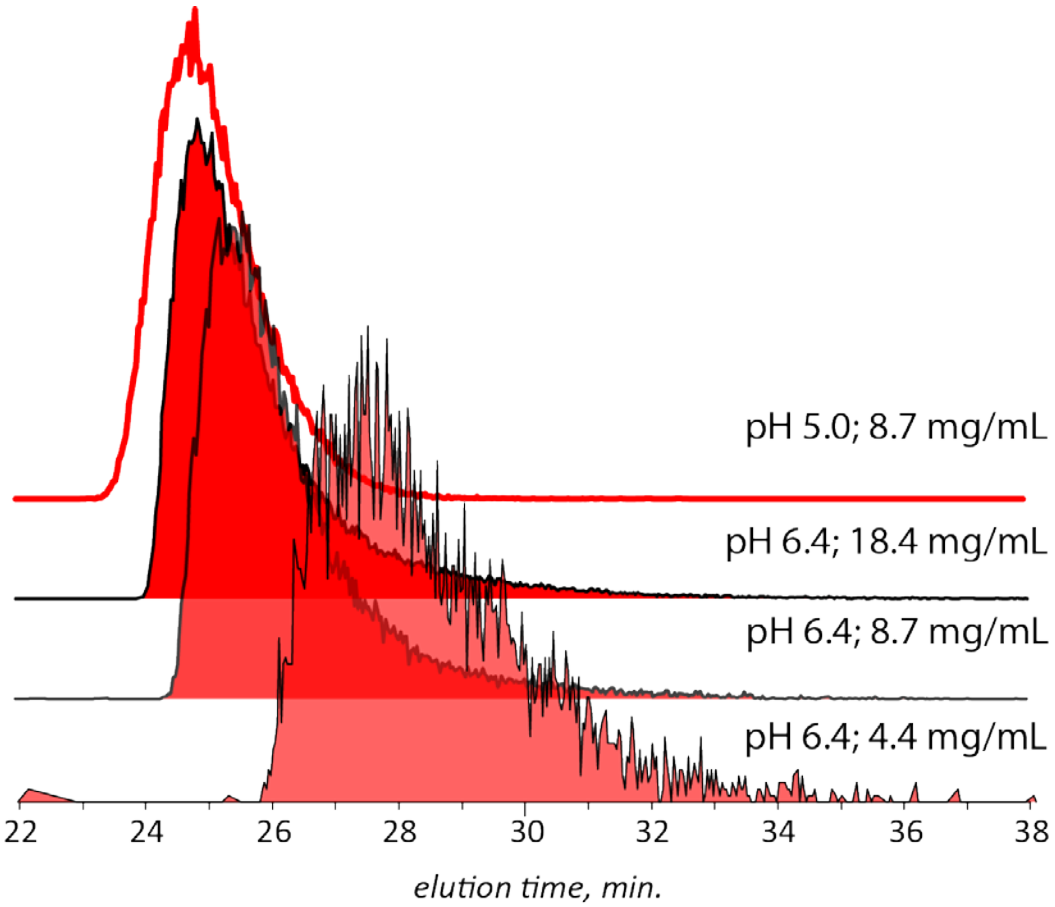

Figure 2.8 Elution profiles of the octameric species of rhASA acquired with online SEC/ESI MS at pH 5.0, protein concentration 1.1 mg/mL (red trace) and pH 6.4, protein concentrations (from top to bottom) 2.3, 1.1 and 0.55 mg/mL.

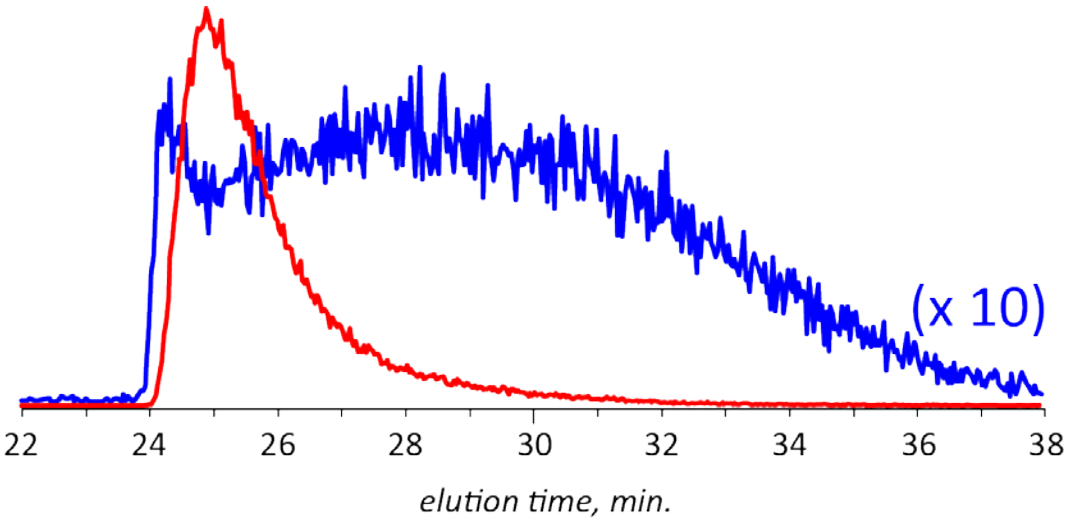

Figure 2.9 Zoomed views of the elution profiles of octameric (red) and dimeric (blue) species of rhASA acquired with online SEC/ESI MS at pH 6.4, protein concentration 2.3 mg/mL.

Supplementary information

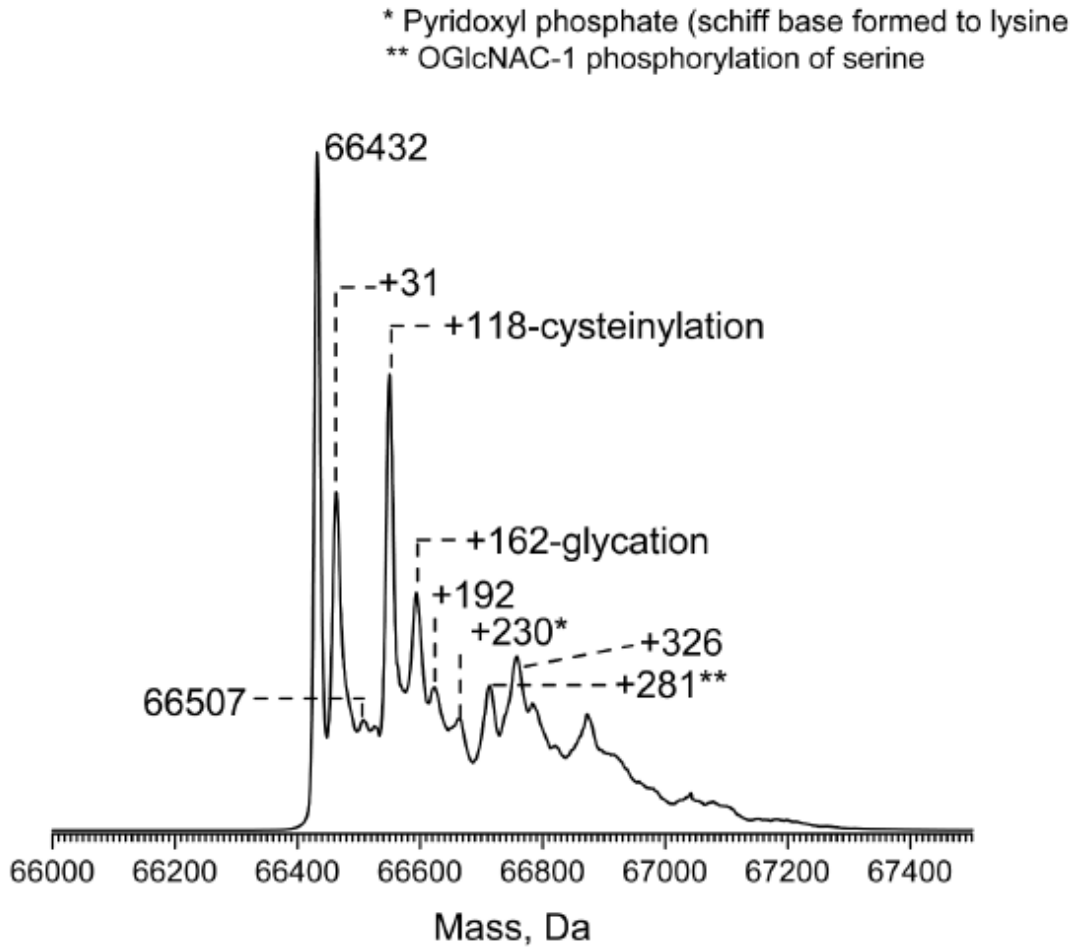

Figure S1. Mass profile of BSA from a commercial source deconvoluted from a mass spectrum acquired under denaturing conditions.

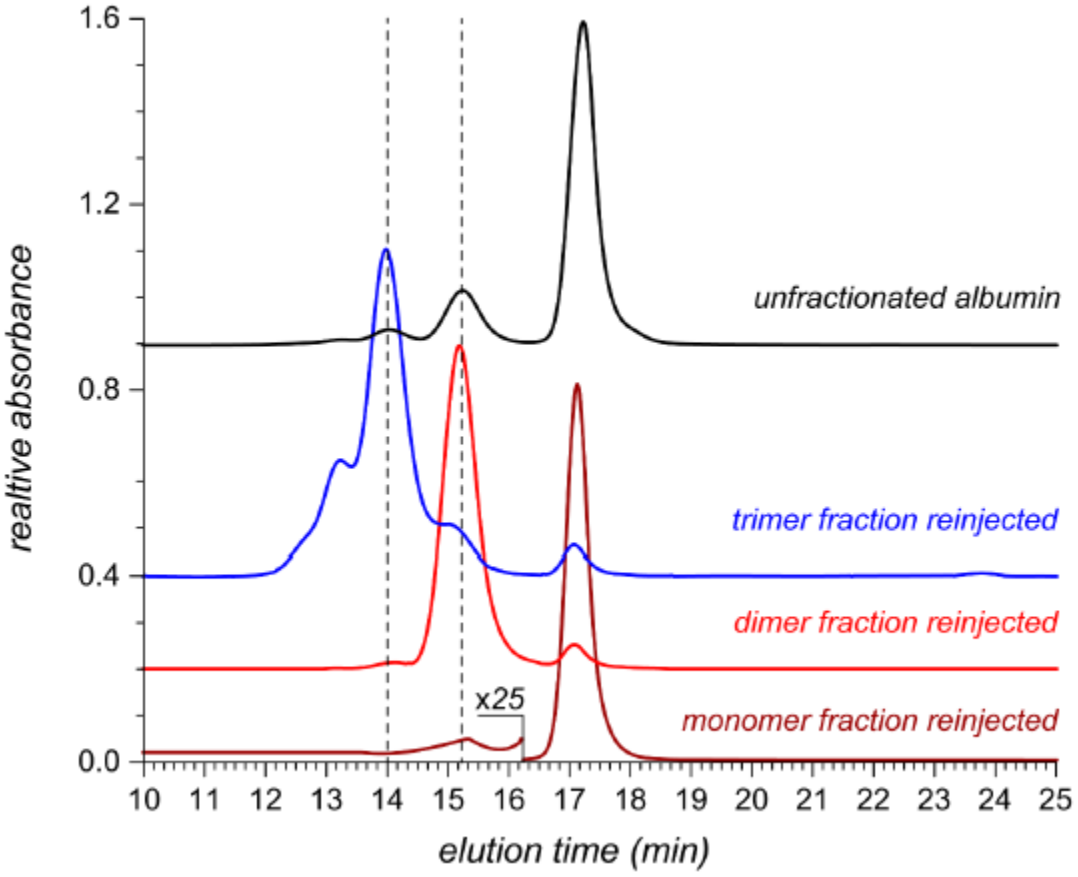

Figure S2. Elution profiles of unfractionated BSA (black trace) and re-injected fractions of trimers (blue), dimer (red) and monomers (brown) showing dynamic character of low molecular weight protein aggregates.

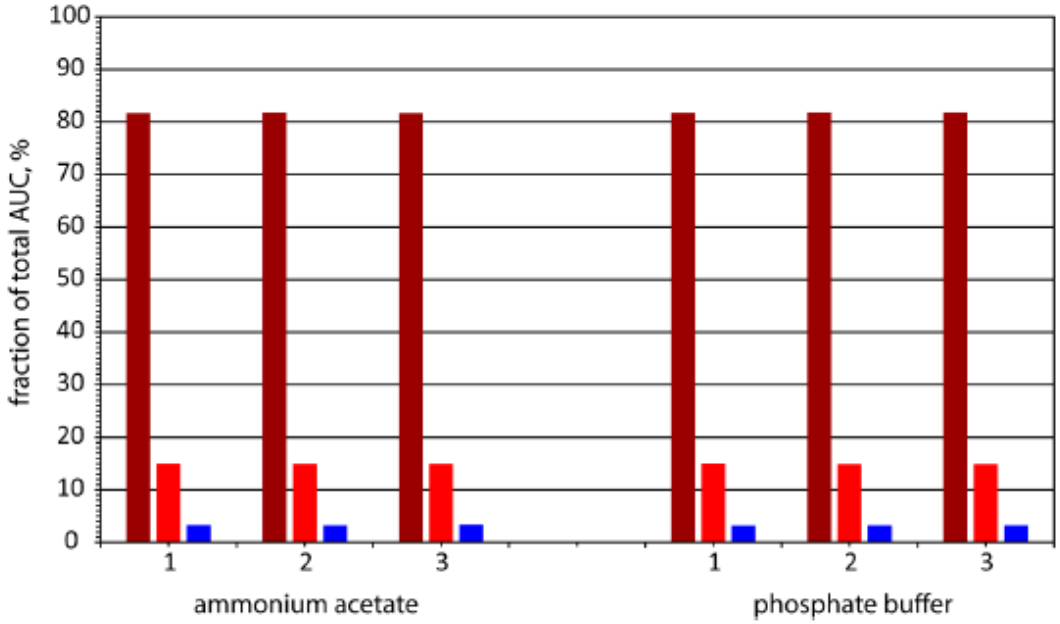

Figure S3. Size distribution of BSA low molecular weight aggregates from SEC chromatograms (UV detection only) acquired with ammonium acetate (left) and phosphate buffer (right) as mobile phases. The colored bars represent monomers (brown), dimers (red) and trimers (blue).

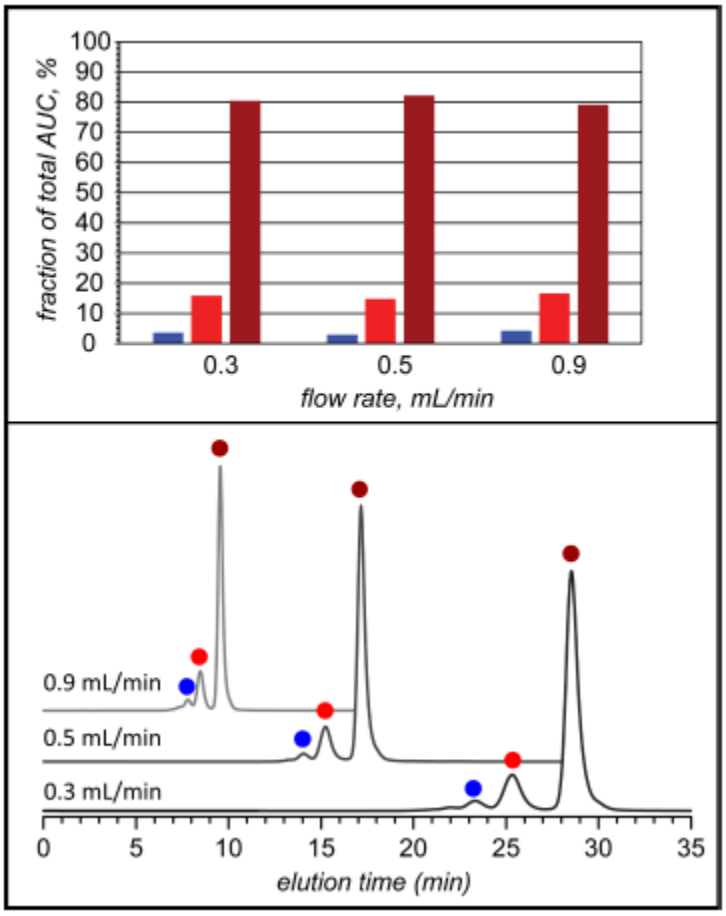

Figure S4. Size distribution of BSA low molecular weight aggregates (top) and representative SEC chromatograms (UV detection only) acquired at different flow rates.

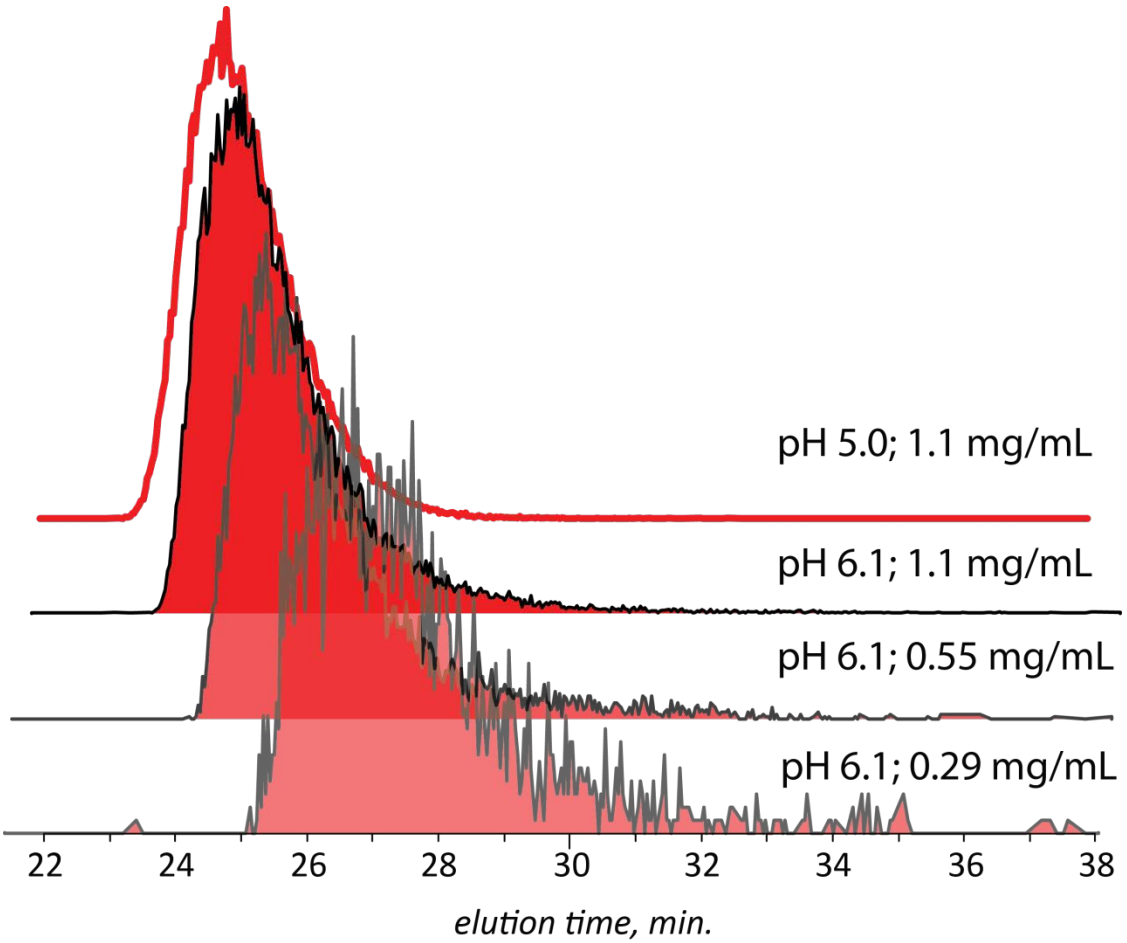

Figure S5. Elution profiles (extracted ion chromatograms) of the octameric species of rhASA acquired with on-line SEC/ESI MS at pH 5.0, protein concentration 1.1 mg/mL (red trace) and pH 6.1, protein concentrations (from top to bottom) 1.1 mg/mL, 0.55 mg/mL, 0.29 mg/mL

CHAPTER 3

MEASURING KINETICS AND AFFINITY OF TRANSIENT PROTEIN-RECEPTOR INTERACTIONS USING SIZE EXCLUSION CHROMATOGRAPHY WITH ONLINE NATIVE ESI-MS DETECTION

3.1. Overview

Protein-protein interactions play an important role in diverse areas of life science research including evaluation of cellular functions and also to aid in the development of therapeutics. Due to this there has always been a tremendous need of analytical methods to characterize and provide accurate information on kinetics and thermodynamics of protein interactions.

Traditionally, strong interactions with the equilibrium dissociation constant (K_D) of low nM were favored (e.g., to develop drug molecules binding irreversibly with the target enzyme for complete inhibition). Recently, there has been a shift towards exploiting lower affinity interactions (the so-called transient protein interactions) with K_D in high nM – low μM range, due to their significance in proper functioning of living cells and also for the development of targeted drug delivery systems.[24, 25, 82] In the case of drug delivery systems transient protein interactions help not only in efficient binding of carrier protein to extracellular receptor, but also in its release after transporting inside the cells. A strong binding in this case will have unfavorable consequences by impeding the release of drug carrier from the protein receptor. Design of efficient targeted drug delivery strategies also critically depends on the availability of the kinetic

data, as the timing of the drug dissociation from its carrier is important for successful delivery.

A wide variety of analytical methods have been developed to characterize protein interactions, however owing to weak binding of transient protein interactions it is technically difficult to detect them compared to stronger interactions. Some of the most commonly used methods include surface plasmon resonance (SPR),[83] analytical ultracentrifugation (AUC),[84] isothermal calorimetry (ITC),[85] and tandem affinity purification-mass spectrometry (TAP-MS).[52] Majority of these methods have disadvantages that invariably affect the binding characteristics such as requirement of protein immobilization in SPR could alter structural properties of proteins, inability to detect transient complexes in TAP-MS without any additional crosslinking, and large concentration of proteins required for ITC and AUC. In vivo analysis of protein interactions can be carried out using fluorescence based assays (e.g., fluorescence resonance energy transfer (FRET) and bimolecular fluorescence complementation (BiFC), however these methods require labeling of proteins with fluorophores, which could again lead to a change in structure of protein.[86] A limitation common to all of these techniques is their inability to deal with highly heterogeneous systems frequently encountered in the field of biotherapeutics.

In the past two decades native electrospray ionization mass spectrometry (ESI MS) has enjoyed growing popularity to characterize structure of non-covalent protein assemblies.[34] In chapter 2, the advantages of combining separation by size exclusion chromatography with native ESI MS detection to characterize aggregation,

conformation, and rapidly equilibrating protein oligomers were presented.[87] Although relating ionic abundance of protein species with fractional concentration in solution is not straightforward due to differences in their ionization and transmission efficiencies, there have also been reports of using native ESI MS to quantify binding affinities.[88]

In this chapter a new label-free analytical method to measure kinetics and binding affinity of transient protein interactions is reported. This method is based on the SEC/native ESI MS method developed in chapter 2. Here we have developed a mathematical model to explain the transfer of dynamic protein species through the column, and by fitting the profile generated from this mathematical solution with the extracted ion chromatograms (XICs) of SEC/MS binding and kinetic parameters of transient protein interactions can be extracted. In this chapter first an explanation for the elution behavior of recombinant arylsufatase A (rhASA) detected in chapter 2 is given and then this is extended to protein/receptor interactions to obtain kinetics and binding information. We have validated this method using known binding of transferrin variants with soluble form of transferrin receptor.[89, 90]

3.2. Materials and Methods

Holo human transferrin (hTf) and holo bovine transferrin (bTf) were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). Soluble ectodomain transferrin receptor (TfR), C-lobe and N-lobe of transferrin were generously provided by Prof. Anne B. Mason (University of Vermont College of Medicine, Burlington, VT). Iron-free forms (Apo) of transferrin were prepared by acid-denaturing the protein in the presence of EDTA followed by centrifugal ultrafiltration (30k molecular-weight-cutoff vivaspin filters)

to 100 mM ammonium acetate. All other chemicals used in this work were of analytical grade or higher. Tf/TfR solutions were mixed in 2:1 molar ratio by diluting from stock solution, and after equilibration were analyzed by SEC/ESI MS without any desalting step.

SEC separations were carried out using Agilent 1100 HPLC (Agilent Technologies, Santa Clara, CA) equipped with a Tosoh (Tokyo, Japan) TSKgel G3000 SWxl column (7.8 × 300 mm) and a 100 µL injection loop. Ammonium acetate solution (100 mM, pH 7) was used as the mobile phase in all separations at a flow rate of 0.5 mL/min. After exiting the UV detector (operated at 280 nm), using a 1:1 flow splitter the flow rate of eluate directed to ESI MS was set at 0.25 ml/min.

ESI MS detection was carried out using QStar-XL (ABI-Sciex, Toronto, Canada) hybrid quadrupole/time-of- flight mass spectrometer equipped with a standard ESI source. The ESI source conditions were optimized to provide a stable spray and optimal ion desolvation: drying gas, 30 L/min; nebulizing gas, 39 L/min; curtain gas, 22 L/min; declustering potential on the skimmer, 290 V; declustering potential on the orifice, 160 V. The source temperature was maintained in the 200 − 250°C interval. Collisional cooling in the ion guide region (a gas flow restricting sleeve in Q0) was used to enhance focusing and stability of noncovalent complexes.

3.3. Theory

In chapter 2 SEC/ESI MS analysis of recombinant arylsulfatase A (rhASA) at pH 6.4 yielded asymmetric peak shape for XICs of ocatmers (M_8) and diffused pattern for dimers (M_2). This observation can explained by taking into account multiple dissociation

events of octamers to dimers followed by re-association of newly formed dimers to octamers over the fixed column length (L). This behavior is illustrated in figure 3.1 using a 2-dimensional plot (distance traveled inside the column vs. the travel time). Taking into consideration these dynamic reactions along with diffusion and mass transfer of analytes a mathematical solution can be constructed for transfer of octamers and dimers through the column as shown in equation 1. This system is a one-dimensional version of the so-called advection-diffusion-reaction (ADR) equation,[91] which is frequently encountered in chemical engineering, the two significant differences being unequal mass transfer coefficients u_{M8} and u_{M2} (the flow velocity is the same for all reagents in a flow reactor), and non-linear reaction terms.

$$\begin{cases} \frac{\partial C_{M8}}{\partial t} = -u_{M8} \cdot \frac{\partial C_{M8}}{\partial x} + D_{M8} \cdot \frac{\partial^2 C_{M8}}{\partial x^2} - k_{-4} C_{M8} + \frac{1}{2} k_{2} C_{M2}^2 \\ \frac{\partial C_{M2}}{\partial t} = -u_{M2} \cdot \frac{\partial C_{M2}}{\partial x} + D_{M2} \cdot \frac{\partial^2 C_{M2}}{\partial x^2} + 4k_{-4} C_{M8} - 2k_2 C_{M2}^2 \end{cases} \quad (1)$$

(assuming a short lived intermediate tetramer (M_4) during $M_2 \rightleftharpoons M_8$ transition of rhASA)[80]

u_{M8}, u_{M2}: Linear velocity of octamers and dimers, respectively.

D_{M8}, D_{M2}: Diffusion coefficient of octamers and dimers, respectively.

Varying k_2 and k_{-4} will generate a family of solutions for equation 1, from which the best match with the experimental data will be selected, followed by generating another family of solutions by varying the parameters k_2 and k_{-4} within a narrower range, etc. This optimization routine will eventually produce kinetic parameters that are best fits for the experimentally measured XIC profiles of M_8 and M_2.

In this chapter the approach presented above to study homo-oligomerization of protein have been applied to protein/receptor interaction systems.

For protein and receptor that interacts to form a 1:1 complex (equation 2).

$$P \cdot R \rightleftharpoons_{k_1}^{k_{-1}} \rightleftharpoons P + R \tag{2}$$

The interacting partners (P, R, and P·R) have different size, and therefore would travel with different linear velocities during trajectory inside the column. Due to this difference the dissociation products generated in a single dissociation event will not be able to re-associate (as shown in figure for a hypothetical protein receptor system). However, physical interaction leading to reassociation of P and R will be possible if their trajectories intercept as they are formed in two different dissociation events (grey circles in figure 3.2). The partial derivative equations describing dynamic P·R complex is a simple extension of equation 1:

$$\begin{cases} \dfrac{\partial C_{PR}}{\partial t} = -u_{PR} \cdot \dfrac{\partial C_{PR}}{\partial x} + D_{PR} \cdot \dfrac{\partial^2 C_{PR}}{\partial x^2} - k_{-1} C_{PR} + k_1 C_P C_R \\ \dfrac{\partial C_P}{\partial t} = -u_P \cdot \dfrac{\partial C_P}{\partial x} + D_P \cdot \dfrac{\partial^2 C_P}{\partial x^2} + k_{-1} C_{PR} - k_1 C_P C_R \\ \dfrac{\partial C_R}{\partial t} = -u_R \cdot \dfrac{\partial C_R}{\partial x} + D_R \cdot \dfrac{\partial^2 C_R}{\partial x^2} + k_{-1} C_{PR} - k_1 C_P C_R \end{cases} \tag{3}$$

Partial derivative equations 5 are a straight forward modification of equation 3 for bivalent system in which the receptor (R) can bind to two protein molecules

$$P_2 R \rightleftharpoons_{k_2}^{k_{-2}} P + P \cdot R; \quad P \cdot R \rightleftharpoons_{k_1}^{k_{-1}} P + R \tag{4}$$

$$\begin{cases} \dfrac{\partial C_P}{\partial t} = -u_P \cdot \dfrac{\partial C_P}{\partial x} + D_P \cdot \dfrac{\partial^2 C_P}{\partial x^2} + k_{-1} C_{PR} + k_{-2} C_{P_2R} - k_1 C_P C_R - k_2 C_P C_{PR} \\ \dfrac{\partial C_{PR}}{\partial t} = -u_{PR} \cdot \dfrac{\partial C_{PR}}{\partial x} + D_{PR} \cdot \dfrac{\partial^2 C_{PR}}{\partial x^2} + k_{-2} C_{P_2R} + k_1 C_P C_R - k_{-1} C_{PR} - k_2 C_P C_{PR} \\ \dfrac{\partial C_{P_2R}}{\partial t} = -u_{P_2R} \cdot \dfrac{\partial C_{P_2R}}{\partial x} + D_{P_2R} \cdot \dfrac{\partial^2 C_{P_2R}}{\partial x^2} + k_2 C_P C_{PR} - k_{-2} C_{P_2R} \\ \dfrac{\partial C_R}{\partial t} = -u_R \cdot \dfrac{\partial C_R}{\partial x} + D_R \cdot \dfrac{\partial^2 C_R}{\partial x^2} + k_{-1} C_{PR} - k_1 C_P C_R \end{cases} \quad (5)$$

3.4. Results and Discussion

SEC/ESI MS analysis of non-covalent protein assemblies stable in solution on the chromatographic time scale give rise to well-defined peaks whose elution time is consistent with the hydrodynamic size (also controlling the linear velocity) of each species. An example is shown in figure 3.3, where incubation of the metalloprotein transferrin (Fe_2Tf) with the soluble ectodomain of transferrin receptor (TfR) resulted in formation of a stable 2:1 complex, detected as an earlier eluting peak. In this case online detection with native ESI MS provides confirmation based on their masses (*i.e.*, an early-eluting 322 kDa $(Fe_2Tf)_2$·TfR complex, and a late-eluting 79 kDa Fe_2Tf, which was present in the mixture in molar excess). A failure to form a protein/receptor complex would be confirmed by the absence of the early-eluting species in the chromatogram, as illustrated in Figure 3.3 for the mixture of TfR and Tf N-lobe (Tf loses its receptor-binding competence upon removal of the C-lobe).[89]

3.4.1. Protein-receptor complex interacting with 1:1 stoichiometry

A model system of 1:1 complex was made by mixing molar excess of C-lobe form of human transferrin ($FeTf_C$) with transferrin receptor (TfR), which has already been

shown using orthogonal approaches to have a weak affinity.[89, 90] As shown in figure 3.4 the XICs obtained from the SEC/ESI MS analysis of this mixture is convoluted. The XIC behavior of free $FeTf_C$ and 1:1 complex ($FeTf_C$·TfR) follows the pattern shown in figure 3.2; $FeTf_C$·TfR complex peak (purple trace in figure 3.4) is asymmetrical with significant tailing and the dissociation products that fail to re-associate appear in chromatograms as early eluting free $FeTf_C$ molecules (16.1 – 17.5 min interval).

Figure 3.5 illustrates a family of solutions generated for each of the interacting species (blue - $FeTf_C$, red - TfR, and purple - $FeTf_C$·TfR complex) using equation 3. Symmetrical peaks shapes with elution time characteristic of their linear velocity were generated, when dissociation channels were turned off. Opening only the dissociation path for the complex (k_{-1} = 0.015) shows distinct features that include decrease in abundance $FeTf_C$·TfR complex peak, increase in abundance and broadening of TfR peak, and monotonic trailing of $FeTf_C$ to early elution time interval. Opening both dissociation and association of complexes (k_1 = $1.3.10^4$ and k_{-1} = 0.020) led to tailing and shift of $FeTf_C$·TfR complex to later elution time as these complexes are formed as a result of reassociation of interacting partners on the column, and a bimodal trailing of $FeTf_C$ was also seen.

Based on these observations a family of solutions were generated to get a close match fit with experimental chromatograms (figure 3.6). In our optimization steps a solution was considered a match if its peak shape and elution time aligned with the experimental chromatograms. As TfR signal was not recorded, we have used chromatograms of $FeTf_C$ and $FeTf_C$·TfR complex to get a match. By not taking absolute

abundance of each of these species into consideration, any artifacts generated due to differences in ionization and transmission efficiency are avoided. Following this optimization routine for $FeTf_C$ and TfR interaction, we detected kinetic rates of k_1 = 3.3 x 10^7 M^{-1} $min^{-1}$and k_{-1} = 3.0 min^{-1}. This gives an equilibrium dissociation constant of 90nM.

3.4.2. Protein-receptor complex interacting with 2:1 stoichiometry

Validation of this method to characterize binding of a 2:1 protein/receptor complex was carried out using transferrin variants. The affinity of these transferrin variants to human TfR has already been reported to cover a broad range using orthogonal approaches.[89] Figure 3.7 illustrates experimental data for binding of three transferrin variants; apo human transferrin (hTf), holo bovine transferrin (Fe_2 bTf), and apo bovine transferrin (bTf) interacting with TfR. These chromatograms are highly convoluted and differ significantly from those detected for Fe_2hTf/TfR interaction (figure 3.3). First, Tf·TfR complexes of both 2:1 and 1:1 stoichiometries are observed despite excess of the protein over the receptor in the injection mixture. Second, the elution profiles of the 2:1 complexes (Tf_2·TfR) are asymmetrical and are shifted compared to the elution time of the $(Fe_2hTf)_2$·TfR complex (shown in Figure 3.7 with a dotted line). This is consistent with the notion of the vast majority of 2:1 complexes not surviving the chromatographic run and dissociating prior to elution from the column; the observed complexes are formed within the column as a result of re-association processes. The second, third… generation of dissociation products that fail to re-associate appear in chromatograms in Figure 3.7 as early eluting free Tf molecules (14.5 – 17.5 min interval) and 1:1 complexes.

Generating a family of solutions for $(Tf)_2$·TfR, Tf·TfR, and Tf species using equation 5 and optimizing them by varying kinetic rate constants (k_1, k_{-1}, k_2, and k_{-2}) yielded a close match profile as shown in figure 3.8. Based on these kinetic rates, equilibrium dissociation constants were obtained for transferrin variants used in this validation study. As expected the equilibrium dissociation constants (K_{D1} and K_{D2}) for hTf and Fe_2 bTf were in higher nano molar range, and very weak binding with K_{D1} in micro molar and K_{D2} in higher micro molar for interaction of bTf with human TfR was detected.

3.5. Conclusions

A label-free method using SEC with online ESI MS to obtain kinetic and thermodynamic parameters of transient protein interactions has been presented. Strong protein complexes that do not dissociate on the chromatographic timescale will be outside of reach of this method. In this study we have validated this new method for transient protein interactions (protein/receptor complexes) with both 1:1 and 2:1 stoichiometries. In a similar way this method can also be extended to transient protein interactions with higher stoichiometries as long as each of the interacting species have distinct linear velocities in chromatography.

FIGURES

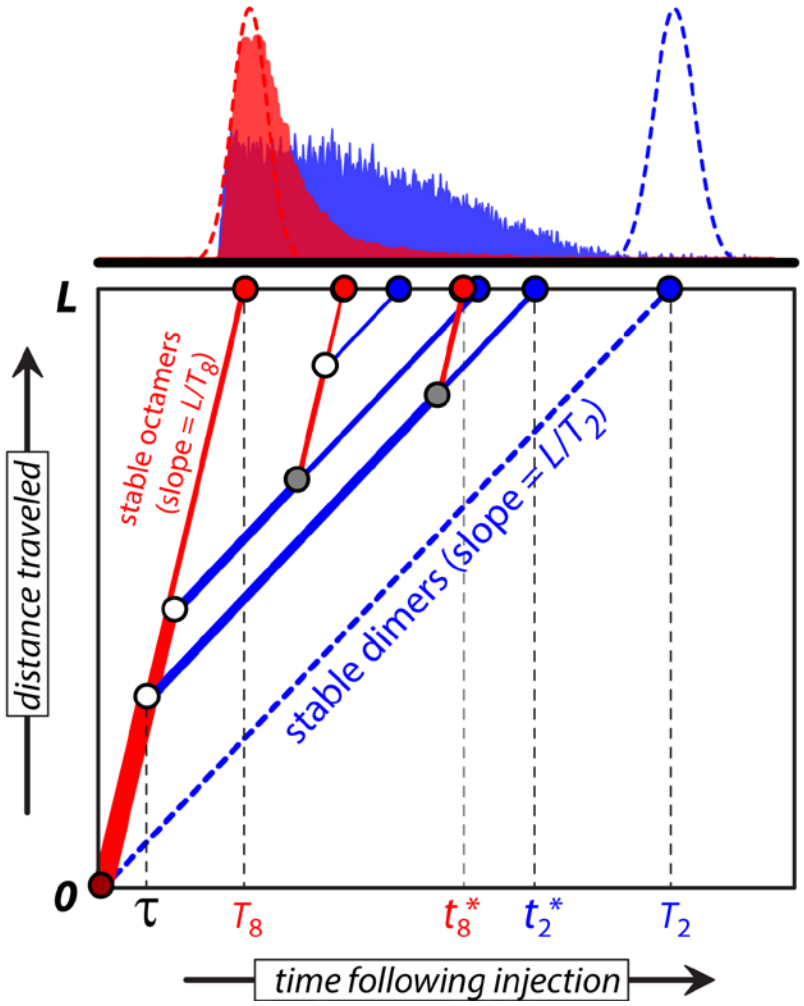

Figure 3.1 Schematic 2-D diagram illustrating behavior of M8 and M2 species during the SEC analysis of metastable rhASA octamers. Stable species would travel only along the trajectories on the far left (M8) and far right (M2), but the actual trajectories branch out as a result of M_8 dissociation (white circles) and M_2 re-association (gray circles). Red and blue circles show elution of M_8 and M_2, respectively.

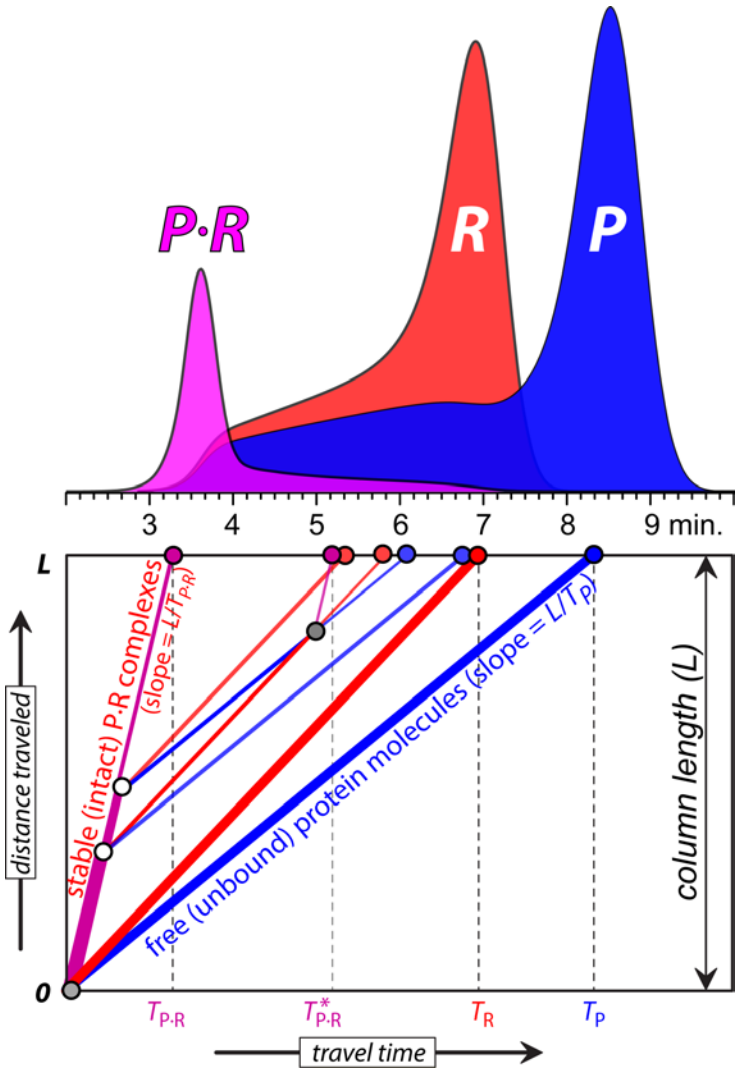

Figure 3.2 Illustration of behavior of two interacting species (P, protein, and R, receptor) during the SEC run. All three species are present at the time of injection. Products of P·R dissociation (white-filled circles) can re-associate with finite probability only if their trajectories intercept (gray circles).

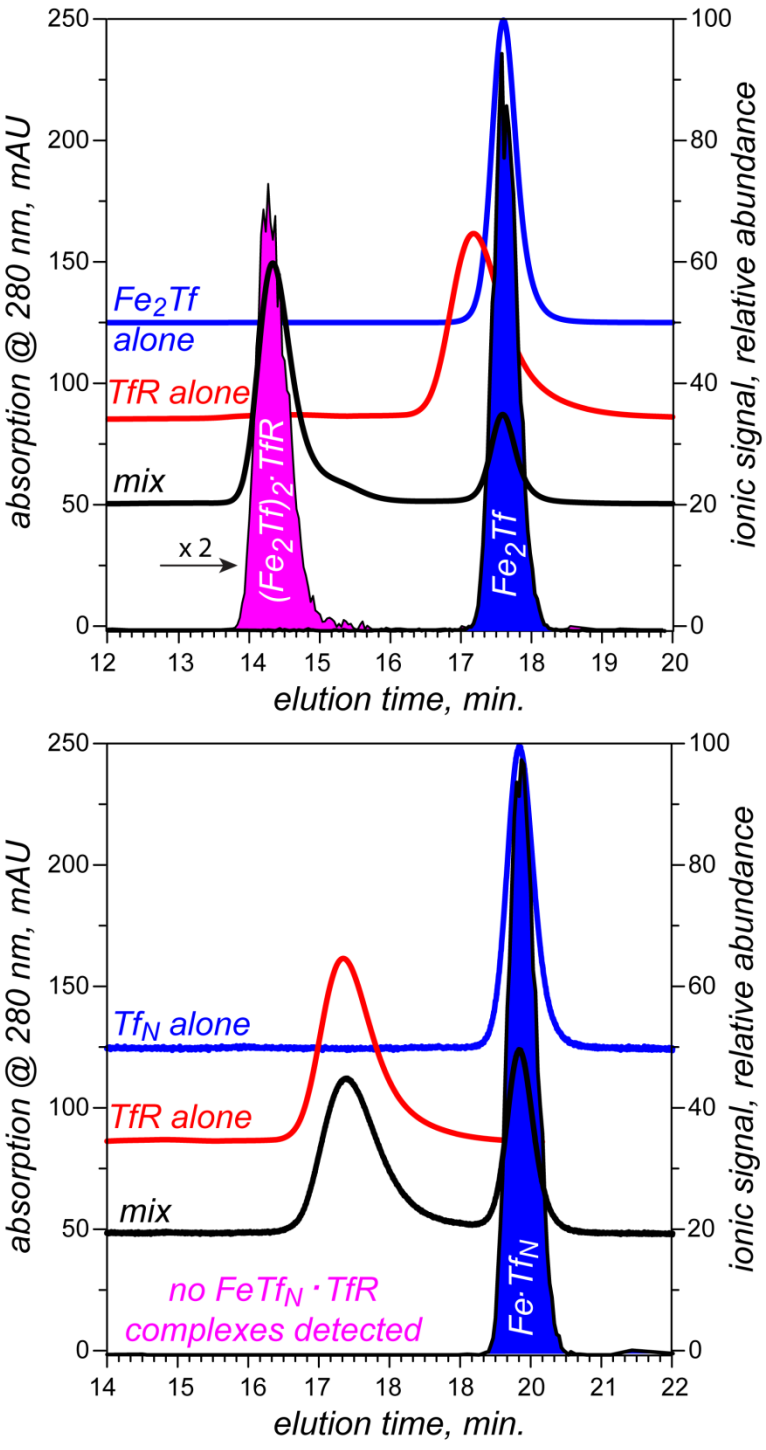

Figure 3.3 SEC and SEC/ESI MS characterization of transferrin binding to its receptor with high affinity (full-length, iron-saturated human Tf, top) and failing to form a complex with the receptor (N-terminal half of human Tf, bottom). The three traces show UV chromatograms of Tf, TfR and their mixtures. Extracted ion chromatograms of the protein/receptor complex (purple) and the protein alone (blue) are also shown.

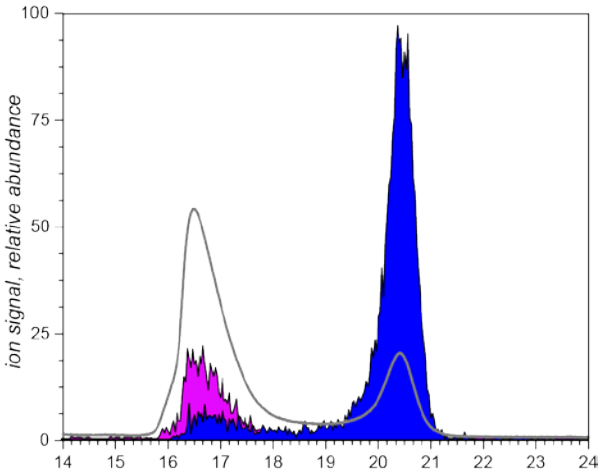

Figure 3.4 SEC/ESI MS characterization of C-lobe form of transferrin ($FeTf_C$) binding to transferrin receptor (TfR). The three traces show UV chromatogram (grey), and extracted ion chromatograms of $FeTf_C$ (blue) and $FeTf_C$·TfR complex (purple).

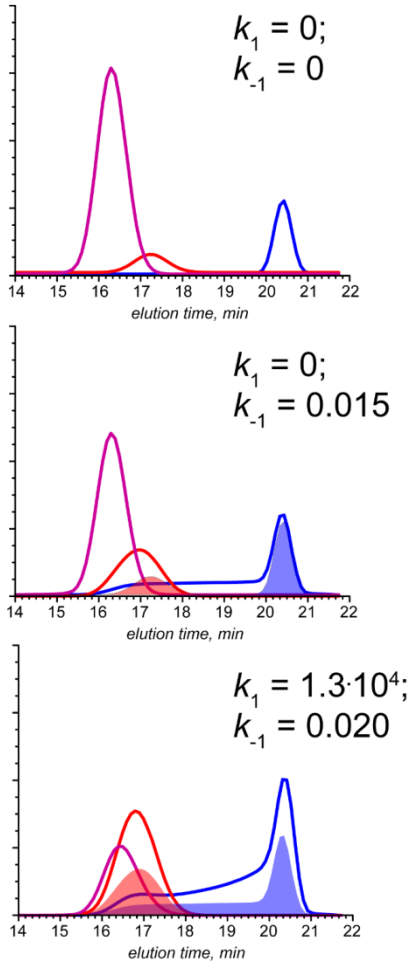

Figure 3.5 Family of solutions generated using initial set of rate constants as shown in each panel. The three traces show extracted ion chromatograms of $FeTf_C$ (blue), TfR (red), and $FeTf_C$·TfR complex (purple). Filled blue and red traces (extracted ion chromatograms of $FeTf_C$ and TfR, respectively at k_1, k_{-1} = 0) are overlaid in middle and bottom panel to illustrate change in peaks under different rate constants.

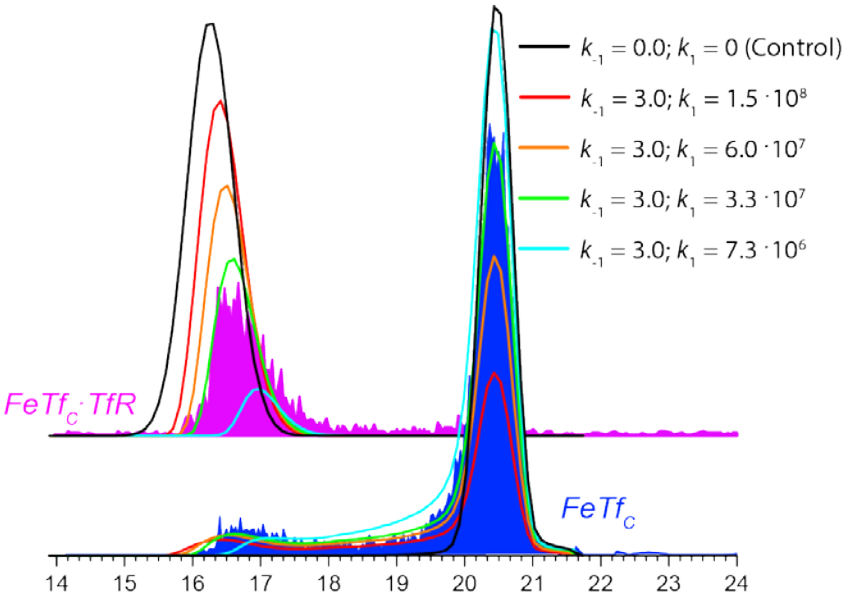

Figure 3.6 Extraction of rate constants for $FeTf_C$/TfR binding and dissociating using experimental SEC/ESI MS data. The calculated elution profiles are overlaid to get the best match. Calculated profile using rate constants of k_1= 3.3 x 107 M^{-1} $min^{-1}$and k_{-1} = 3.0 min^{-1} (green) was considered to be the best match with the experimental SEC/MS data.

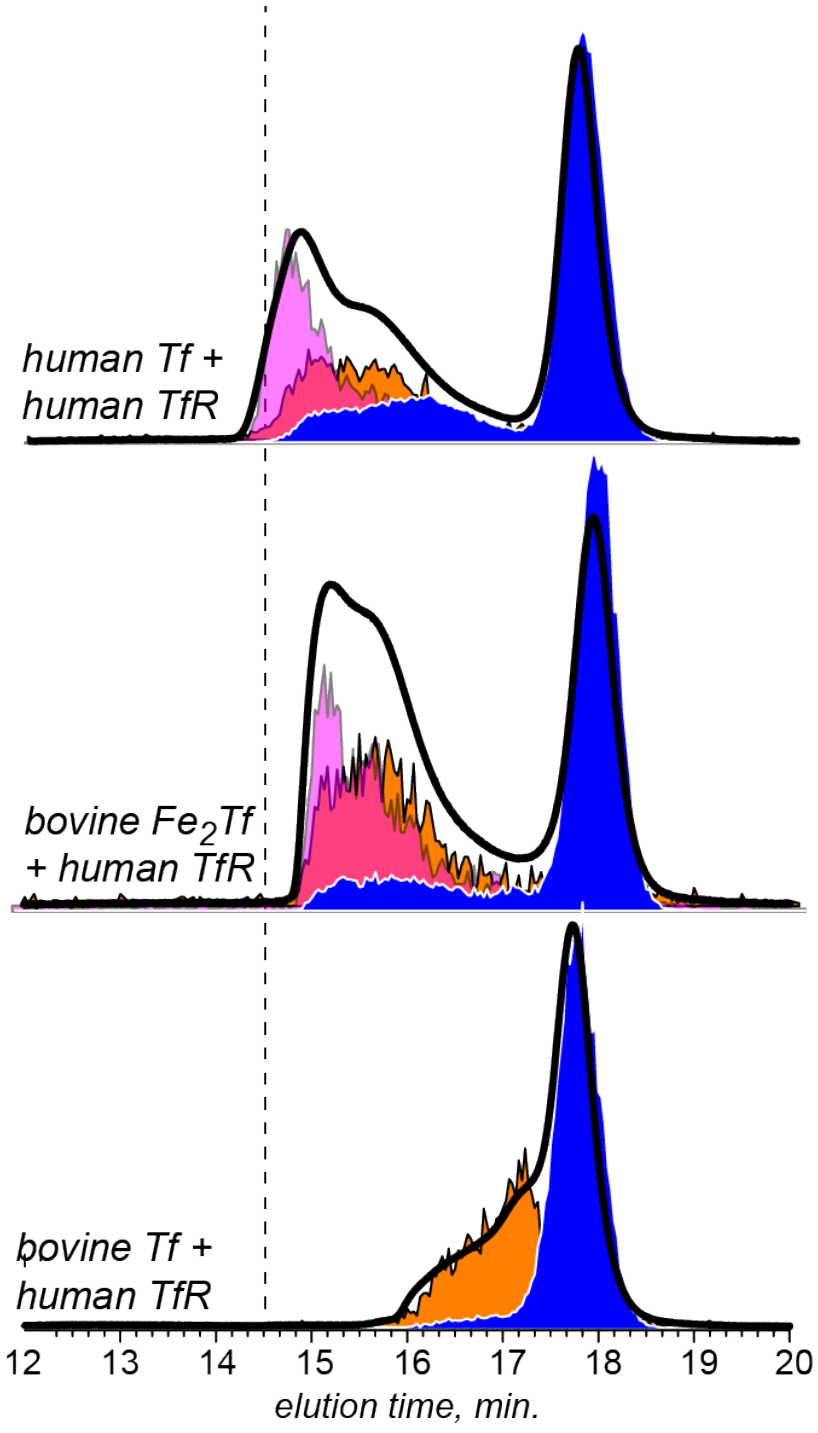

Figure 3.7 SEC/ESI MS analysis of transient Tf/TfR complexes. Tf elution profiles are shown with blue curves, and elution profiles of Tf·TfR complexes are shown with orange (1:1 stoichiometry) and purple (2:1) curves. The receptor affinity order is apo-human Tf > holo-bovine Tf >> apo-bovine Tf.

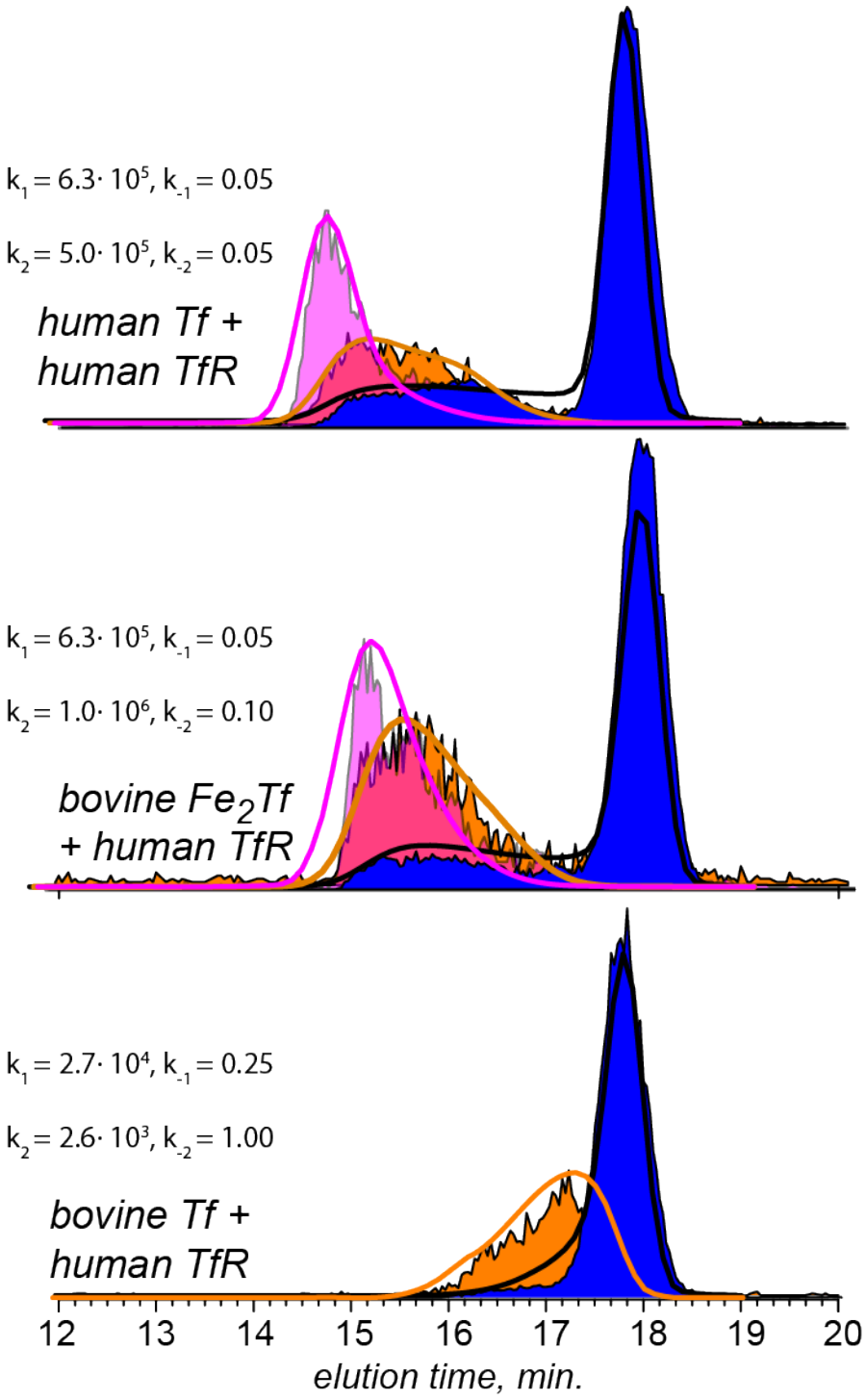

Figure 3.8 Extraction of rate constants for Tf/TfR binding and dissociating using experimental SEC/ESI MS data. The calculated elution profiles that were good match with experimental data are shown. Calculated elution profile of Tf (solid black trace), Tf·TfR complexes with 2:1 stoichiometry (solid purple trace), and 1:1 stoichiometry (solid orange trace) for each experiment are shown. Rates constants; k_1 and k_{-1} are expressed in M^{-1} min^{-1} and min^{-1}, respectively.

CHAPTER 4

CHARACTERIZATION OF INTACT PROTEIN CONJUGATES AND BIOPHARMACEUTICALS USING ION-EXCHANGE CHROMATOGRAPHY WITH ONLINE DETECTION BY NATIVE ELECTROSPRAY IONIZATION MASS SPECTROMETRY AND TOP-DOWN TANDEM MASS SPECTROMETRY

Peer-reviewed article published: Muneeruddin, K.; Nazzaro, M; Kaltashov, I.A.; Characterization of Intact Protein Conjugates and Biopharmaceuticals Using Ion-Exchange Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry and Top-Down Tandem Mass Spectrometry *Analytical Chemistry,* **2015**, 87, 10138 – 10145

4.1. Overview

Protein therapeutics remain the fastest growing segment in the pharmaceutical industry,[5] and as the number of both approved biopharmaceutical products and those still in various development stages continue to grow, so does the demand for powerful and robust analytical technologies capable of comprehensive characterization of these complex molecules. Protein therapeutics are inherently less stable compared to traditional small-molecule medicines, as their large size (and, consequently, a large number of chemically labile sites) makes them prone to a variety of nonenzymatic post-translational modifications (PTMs), many of which impact both therapeutic and safety profiles of the protein drugs. This is usually a result of conformational changes triggered by chemical modifications, although a loss or alteration of the higher order structure may also be precipitated by environmental factors, such as temperature variations and exposure to solvent–air interface. Therefore, complete structural characterization of biopharmaceutical products cannot focus solely on their covalent (primary) structure but must also address the higher order structure, including conformational integrity and

in the case of multiunit proteins correct noncovalent assembly. Demanding and challenging as they are, these tasks are further complicated by structural heterogeneity inherent to many protein-based therapeutics (e.g., due to extensive glycosylation). Lastly, a large number of the second-generation biopharmaceuticals contain "designer" modifications (such as protein–drug or protein–polymer conjugates), which increase the extent of heterogeneity even further due to variation in the extent of conjugation (loading) and the presence of multiple positional isomers within the population of conjugates with fixed stoichiometry.

Obviously, comprehensive characterization of such complex systems is a very tall order, and analytical techniques capable of not only detecting and quantitating nonenzymatic PTMs within highly heterogeneous protein therapeutics but also assessing the impact of these modifications on conformational integrity are at premium. Mass spectrometry (MS) has recently emerged as a very powerful and robust tool to characterize biopharmaceutical products at a variety of levels ranging from mass profiling,[92] amino acid sequencing,[93, 94] and PTM mapping[95-97] to conformational integrity[20, 21, 98] and aggregation propensity.[70, 99, 100] While the benefits of MS for comprehensive characterization of biopharmaceutical products have been well-documented,[32, 33, 101] the growing demands of characterizing ever more complex protein-based therapeutics continue to stimulate development of new analytical methodologies. Use of hyphenated techniques, such as liquid chromatography/mass spectrometry (LC/MS) and liquid chromatography/tandem mass spectrometry (LC/MS/MS), has proved particularly useful for characterization of complex protein

samples both in the field of biopharmaceuticals and beyond;[102-104] however, the vast majority of applications make use of reversed-phase chromatography as the separation step. While the versatility of reversed-phase LC allows it to be applied for separations of both small peptides and large intact proteins, the nature of the mobile phase causes the proteins to be unfolded during the separation step. As a result, any information on the higher order structure (polypeptide chain folding and/or noncovalent assembly of multiunit protein complexes) is lost, and the MS analysis is focused exclusively on covalent structure.

In contrast to reversed-phase LC, ion-exchange chromatography (IXC) and size exclusion chromatography (SEC) can employ nondenaturing conditions during the separation process, and in fact SEC is frequently used in the biopharmaceutical sector for assessment of the higher order structure.[105, 106] IXC has also proved to be a very useful tool for characterization of complex biopharmaceuticals;[107] however, neither of these two techniques is used routinely in combination with online electrospray ionization (ESI) MS detection. As pointed in a recent review, "while the simplicity associated with size exclusion and cation-exchange chromatography make them ideally suited for routine use, the inherent complexity of mass spectrometry restricts its use to characterization, this despite its richness in information."[102] Recently, we demonstrated that SEC/MS can in fact be used for characterization of complex biopharmaceutical products, yielding not only mass profiles, but also information on conformational integrity, aggregation propensity, and quaternary structure;[87] the benefits of using SEC/MS for mass profiling of heterogeneous protein therapeutics have also been

documented.[75, 108] The use of IXC with online MS detection has also been demonstrated recently,[109] although the spray conditions in the ESI interface resulted in protein unfolding prior to entering MS. Consequently, all information on protein conformation was inferred from the chromatographic elution time,[109] rather than from ESI mass spectra.

In this work we present an online IXC/ESI MS in which the ESI source is operated under nondenaturing conditions (the so-called native ESI). Using native ESI MS as a detection tool allows the capability of the detector to be expanded beyond mere mass profiling by allowing the conformational integrity of various species to be assessed based on the analysis of protein ion charge state distributions in the mass spectra.[110] Furthermore, incorporation of top-down fragmentation in the experimental workflow (LC/MS/MS of whole proteins) allows nonenzymatic PTMs to be identified in situations when the mass measurement alone is not sufficient for an unequivocal conclusion to be reached. Although this initial report focuses on proteins of relatively modest size (not exceeding 25 kDa), this technique appears to be ideally suited for the analysis of large systems as well, including antibody–drug conjugates (ADCs) and multiunit protein assemblies.

4.2. Materials and Methods

Chicken egg white lysozyme (Lz), dithiothreitol (DTT), and iodoacetic acid (IAA) were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO), and *N*-succinimidyl-*S*-acetylthioacetate (SATA) was purchased from Pierce Biotechnology (Rockford, IL). Lz was conjugated with SATA using a previously described procedure.[111] Briefly, SATA

(prepared in DMSO) was added to Lz in phosphate buffer (pH 7.0) such that the ratio of SATA to number of primary amine sites in lysozyme was 1:1. The mixture was incubated at 37 °C for 90 min to allow for conjugation reaction to occur and stored at −20 °C prior to characterization. Complete reduction of Lz (all four disulfide bonds) was carried out by following a previously described procedure[112, 113] that was modified to avoid precipitation of the (partially) reduced forms. Lz (1 mg/mL) was incubated in the presence of 50 mM ammonium bicarbonate (pH 8.0), 30% acetonitrile, and 20 mM DTT (freshly prepared in water) at 30 °C for 20 min. Following reduction, free cysteine residues were alkylated by adding IAA to a final concentration of 40 mM and incubating at room temperature for 30 min. Prior to the analyses, the pH of both freshly reduced and alkylated Lz solutions was adjusted to 7.0. Selective reduction and alkylation with IAA of a single disulfide bond, Cys^{6}–Cys^{127}, was carried by following the published protocol[113] using lower DTT content (the so-called CM-lysozyme[113]). A stressed sample of raw interferon-β was provided by Biogen (Cambridge, MA). All other chemicals used in this work were of analytical grade or higher.

Ion-exchange chromatography was performed on an Agilent 1100 HPLC (Agilent Technologies, Santa Clara, CA) equipped with a 2.1 mm × 250 mm weak cation-exchange column PolyCAT A (PolyLC Inc., Columbia, MD), particle size of 5 µm and pore size of 1000 Å. Ammonium acetate at pH 7.0 was used for MS-compatible gradient elution (100 mM in mobile phase A and 1 M mobile phase B). The total flow rate was maintained at 0.1 mL/min (SATA-conjugated Lz and CM-lysozyme) or 0.2 mL/min (fully reduced Lz and stressed interferon-β1a). The specific linear gradient depended on the sample analyzed

(vide infra), but never exceeded 60% mobile phase B. Upon exiting the column, the eluate was directed to a UV absorbance detector (operated at 280 nm) followed by online detection by native ESI MS without any solvent modification or flow splitting using a QStar-XL (ABI-Sciex, Toronto, Canada) hybrid quadrupole/time-of-flight mass spectrometer (for the analysis of SATA-conjugated and reduced Lz) or Synapt G2Si (Waters, Milford, MA) hybrid quadrupole/time-of-flight mass spectrometer (for the analysis of stressed interferon-β). The ESI source conditions were optimized to get a stable spray and efficient desolvation of analyte ions. For Qstar XL the conditions were as follows: drying gas, 37 L/min; nebulizing gas, 30 L/min; curtain gas, 22 L/min; declustering potential on the skimmer, 290 V; declustering potential on the orifice, 160 V. The source temperature was maintained in the 200–250 °C interval. For Synapt G2Si the conditions were optimized as follows: sampling cone, 40 V; source offset, 80 V; desolvation temperature, 400 °C; cone gas flow, 45 L/h; desolvation gas flow, 600 L/h. Online top-down fragmentation of protein ions was carried out with Synapt G2Si using both electron-transfer dissociation (ETD) in the trap cell and collision-induced dissociation (CID) in the transfer cell. For ETD the trap wave height was set at 1.0 V, and trap wave velocity was 300 m/s. CID was carried out by setting collision energy in the transfer cell to 100 and transfer gas flow rate to 4.3 mL/min. Postacquisition processing was carried out with the Origin data analysis software (Origin Lab Corp., Northampton, MA).

4.3. Results and Discussion

4.3.1. Initial evaluation of online IXC/ESI MS: characterization of complex protein conjugate samples

The utility of IXC/ESI MS for characterization of heterogeneous protein samples was initially evaluated using SATA-activated Lz. SATA is a short-chain reagent targeting primary amines, which is used to introduce a sulfhydryl group to proteins. While SATA is frequently used in the studies of protein higher order structure as a cross-linking reagent,[114] our own interest in this reagent stems from its use as a coupling agent to produce transport protein/payload protein conjugates to enable targeted delivery.[111] Since SATA is an amine-targeting reagent, and most proteins contain at least several such groups (i.e., α-amine at the N-terminus and ε-amines of lysine side chains), we frequently observe formation of multiple conjugates, which differ in the extent of conjugation. Low conjugation yields result in poor efficiency of protein cross-linking, while extensive conjugation (placing of multiple sulfhydryls on the protein surface) leads to protein polymerization; hence, controlling the extent of conjugation becomes important, which can be readily achieved by measuring the masses of the conjugation products.[111] However, mass measurements alone fail to characterize another dimension of protein conjugate heterogeneity, namely, the presence of positional isomers within the ensemble of protein species having identical conjugation stoichiometry. Ideally, one aims at producing relatively homogeneous population of conjugates (e.g., by taking advantage of pK_a difference between α- and ε-amines and carrying out the reactions within the pH range where a difference in their reactivity exits), and the ability to

monitor protein heterogeneity due to the presence of positional isomers would obviously be very valuable for optimizing the conjugation protocol.

Previously we used straightforward ESI MS to evaluate the extent of modification of Lz with SATA,[111] while assessing the distribution of conjugation sites among seven possible candidates within this protein required the use of a time-consuming and labor-intensive procedure that involves proteolytic digest and ^{18}O labeling of fragment peptides followed by LC/MS and LC/MS/MS analyses (S. Nguyen et al., in preparation). IXC/ESI MS analysis of SATA-activated Lz at the whole-protein level (Figure 4.1) provides an elegant alternative by allowing both the extent of conjugation and the presence of positional isomers to be determined in a single experiment. While in this particular sample (prepared by using a 1:8 protein/SATA ratio) the extent of conjugation varies from 0 (intact, unmodified Lz) to 4 ($SATA_4$·Lz), only these two species have extracted ion chromatograms with single peaks (52 and 30 min, respectively). The extracted ion chromatogram of SATA·Lz displays two major and two minor peaks (teal-filled curve in Figure 4.1), consistent with the notion that at least four out of seven possible amines participate in conjugation. The numbers of peaks in extracted ion chromatograms of $SATA_2$·Lz and $SATA_3$·Lz (six and four, respectively) are also consistent with the notion of four amines targeted by SATA. This relatively straightforward analysis allows a conclusion to be made that, while the conjugation reaction does not occur at a single site within the protein, the number of participating amines is limited (four out of seven). Importantly, this conclusion could not be made based on either chromatographic or MS data alone. Using ESI MS as a detection tool also allows a distinction to be readily made

among coeluting species with different conjugation stoichiometries (e.g., SATA·Lz and $SATA_2$·Lz at 39.5 min, and $SATA_2$·Lz and $SATA_3$·Lz at 37 min).

While it may be tempting to use the peak heights of various species detected in IXC/ESI MS analysis of the SATA/Lz conjugates for their quantitation, it should be remembered that the intensity of the ionic signal in ESI MS depends on a variety of parameters, including solvent composition[115] (which is a function of time, since a gradual elution is used) and the presence of other coeluting species.[116] Nevertheless, examination of the UV trace and individual extracted ion chromatograms presented in Figure 4.1 suggests that a reasonably good agreement exists between the UV absorbance and the ionic signal throughout the entire chromatogram (based on peaks resolved in the UV chromatogram, the deviation of relative UV and extracted ion chromatogram (XIC) peak intensities does not exceed 15%; see the Supporting Information for more information). Therefore, extracted ion chromatograms appear to be well-suited for at least semiquantitative assessments of relative abundances of various protein species present in the sample.

4.3.2. Online IXC/ESI MS provides information on protein conformational integrity

Above and beyond providing mass information on various protein conjugates, ESI MS detection allows large-scale conformational changes to be detected based on the analysis of protein ion charge state distributions. Mass spectra of all $SATA_n$·Lz species detected in the experiment shown in Figure 4.1 are similar to each other in that only low charge density ions are present in the spectra (charge states ranging from +5 to +8 populating the high m/z regions of the mass spectra). These narrow charge state

distributions are nearly identical to that of the intact (unconjugated) Lz (blue trace in Figure 4.1), suggesting that even multiply conjugated protein molecules maintain compact structure in solution.[110]

In order to prove that large-scale unfolding can in fact be detected by IXC with online detection by native ESI MS, this technique was applied to characterize a mixture of intact Lz and fully reduced and alkylated Lz (Figure 4.2). The mixture contains only two species that can be readily identified based on their masses: reduction of all four disulfide bonds followed by alkylation of free sulfhydryls leads to a mass increase of 464 Da (this species eluted at 15 min); no partially reduced Lz species (expected masses 14 421, 14 537, and 14 653 Da) were observed. An important distinction between the two mass spectra averaged across the peaks representing fully reduced (15 min) and intact (43 min) Lz is the appearance of the ionic charge state distributions (inset in Figure 4.2): the low *m*/*z* region of the mass spectrum collected at 15 min clearly contains abundant ions suggesting that the protein eluted at this time undergoes (partial) unfolding in solution. This, of course, is not surprising, as complete reduction of disulfide bonds in Lz is known to compromise its higher order structure,[112] which is reflected in the ionic charge state distribution in native ESI MS.

In contrast to the completely reduced Lz species, removal of a single disulfide bond Cys^{6}–Cys^{127} (which can be selectively reduced) has little effect on the structure and activity of native Lz.[113] We used this three disulfide containing derivative of Lz (CM-lysozyme) to demonstrate that the ionic charge state distributions do reflect protein conformation in solution, rather than diminished stability in the gas phase due to a loss

of an internal cross-link, and that IXC/ESI MS does not generate false positives regarding loss of protein conformational integrity. The CM-lysozyme sample in which the Cys^{6}–Cys^{127} disulfide bond was selectively reduced based on existing protocols[113] was analyzed by IXC/ESI MS, revealing three components (Figure 4.3). These species were readily identified based on their masses: CM-lysozyme eluting at 36 min (mass 14 421 Da), intact Lz eluting at 52 min (mass 14 305 Da), and a minor component eluting at 40 min whose mass (14 363 Da) is consistent with a single alkylation event (most likely representing a product of a side reaction, where a fully oxidized Lz species is alkylated at a histidine side chain, which is known to be marginally reactive with IAA)[117]. Importantly, the ionic charge state distributions for all three detected Lz species are essentially identical, and no ionic signal is observed in the low *m*/*z* regions of the mass spectra (see inset in Figure 4.3). This is in excellent agreement with the notion of the Cys^{6}–Cys^{127} disulfide bond not being critical for the conformational stability of this protein,[118] providing additional evidence that IXC/ESI MS can be used not only for identification of various protein forms in complex mixtures based on their masses, but also for assessing their conformational integrity based on their ionic charge state distributions.

4.3.3. Characterization of heterogeneous protein therapeutics with IXC/ESI MS and online protein ion top-down fragmentation: identification of stress-induced modifications in Interferon-β

Evaluation of IXC/ESI MS for the analysis of complex biopharmaceutical samples was carried out using a stressed sample of interferon-β, a protein drug that is used for treatment of multiple sclerosis.[119] While historically the first form of interferon-β that became available for clinical use was homogeneous, as it was expressed in bacteria and

was carbohydrate-free,[120] the presence of carbohydrates is known to improve the activity of this protein therapeutic significantly by enhancing its stability.[121, 122] Not surprisingly, just like many other glycoproteins, interferon-β expressed in mammalian cells exhibits a significant extent of heterogeneity due to the presence of multiple glycoforms.[123] While more homogeneous preparations of this protein can be achieved using either extensive fractionation/purification or methods of molecular biology,[124] in this work we intentionally used raw protein material to have the highest degree of heterogeneity in the analytical sample. Furthermore, the sample was stressed to induce nonenzymatic PTMs (deamidation) on the background of extensive enzymatic PTMs (glycosylation). Mass profiling of this sample carried out by ESI MS reveals the presence of six major glycoforms that can be distinguished from one another based on their masses (see inset in Figure 4.4). This number appears to agree with the number of partially resolved peaks observed in both UV and TIC chromatograms of the interferon-β sample analyzed by IXC (Figure 4.4), although the signal intensity distribution of chromatographic peaks does not match that of the ionic signals of different glycoforms in the mass spectrum of the sample prior to separation.

The reason for this discrepancy becomes obvious when XICs are plotted for each of the six major glycoforms (color-fitted curves in Figure 4.4) and overlaid with the TIC. Intriguingly, each glycoform identified by ESI MS gives rise to two peaks in the chromatogram of the protein sample (e.g., the protein species whose mass is consistent with interferon-β bearing a BiNA2 carbohydrate chain elutes at both 17.5 and 19 min, and analogous behavior is exhibited by all other glycoforms). Although there must be a

significant difference between the two species to cause the elution time change, the mass measurement alone in this case is not sufficient for this difference to be identified. There are three nonenzymatic PTMs causing only minimal mass difference that could conceivably occur as a result of protein stress: disulfide reduction (mass difference between the intact and the modified protein 2 Da), deamidation (1 Da), and disulfide scrambling (0 Da). The protein mass changes caused by these PTMs are very small (<0.01% of the total mass) and can easily evade MS detection, while the underlying chemical changes might be sufficient to afford a notable change in IXC elution time.

Since the disulfide bridge in interferon-β is important for maintaining its tertiary fold, its complete reduction would be expected to cause partial unfolding; a putative scrambling event (formation of a non-native disulfide bridge involving free cysteine residue at position 17) would also be expected to alter the protein conformation. At the same time, deamidation events[125] frequently do not induce large-scale conformational changes.[126] The charge state distributions for BiNA2 species of interferon-β eluting at 17.5 and 19 min were nearly identical to each other, both featuring only low charge density ions (in the high *m*/*z* region of the mass spectra). Therefore, a conclusion was made that both forms of the protein maintained compact structure in solution, and the stress-induced nonenzymatic PTMs failed to result in large-scale conformational changes. While this seems to suggest that the modification in question is in fact deamidation, a more definitive proof was needed to unequivocally confirm the presence of the deamidated species in the stressed interferon-β sample.

In order to establish the specific nature of the stress-induced nonenzymatic PTM of interferon-β, an online IXC/ESI MS/MS experiment was carried out by selecting precursor ions within a 2480–2490 *m*/*z* window followed by ETD in the trap cell and collisional activation of the products in the transfer cell (see the Experimental Section for more detail). The ETD process was supplemented with collisional activation in order to facilitate physical separation of protein ion fragments in the gas phase: since the protein ions were generated under near-native conditions, they were expected to retain some higher order structure in the gas phase. This phenomenon is known to prevent fragment dissociation (following cleavage of a chemical bond connecting them) in many cases due to the presence of residual noncovalent bonds (e.g., hydrogen bonds) holding the two fragments together.[127] Collisional heating eliminates these residual bonds and facilitates the dissociation process.[127] The fragmentation patterns obtained for both species of interferon-β (see the Supporting Information) are nearly identical to each other and feature abundant fragments within the Met^{1}–Tyr^{30} and Ala^{142}–Asn^{156} segments of the protein. The near absence of fragments due to backbone cleavages in the protein segment flanked by Cys^{31} and Cys^{141} suggests that the disulfide bond connecting them is intact in both forms of the protein and rules out both disulfide scrambling and complete reduction as stress outcomes.

There is only one additional ion that is present in the fragmentation spectrum averaged across the earlier-eluting peak of the BiNA2 species of interferon-β, namely, b_{25}^{2+}; the only other difference between the two sets of fragment ions is the 1 Da mass shift for the c_{30}^{2+} ion recorded at earlier elution time (see insets in Figure 4.5). The

protein segment covered by the latter fragment incorporates an asparagine residue (Asn^{25}), and deamidation at that position would indeed result in a mass increase of 1 Da, which is reflected in the mass shift of c_{30}^{2+}. Furthermore, deamidation at this position would create a proximal acid–basic residue pair ($Asp^{25}Gly^{26}Arg^{27}$), a motif known to facilitate a direct bond cleavage of the amide bond adjacent to the aspartic acid residue,[128] hence the appearance of the b_{25}^{2+} fragment in the spectrum averaged across the earlier-eluting peak. We also note that the monoisotopic mass of the b_{25}^{2+} fragment is 1 Da higher compared to the theoretical mass of this fragment calculated based on the sequence of intact interferon-β (and, therefore, consistent with a single deamidation event occurring within the N-terminal segment of the protein). Taken together, these observations not only provide unequivocal evidence that the nonenzymatic PTM triggered by stress is indeed deamidation, but also allow the affected site to be localized within the protein sequence. While there is a variety of mass spectrometry-based methods to detect protein deamidation and localize the affected sites,[129, 130] a top-down approach to this task offers multiple advantages, such as elimination of additional sample preparation steps (enzymatic digestion and separation of peptide fragments).[130] In addition to decreasing the analysis time to less than 1 h, it also eliminates artifacts that could occur during sample preparation required for bottom-up analysis.[129] However, until recently the top-down approach was applied only to small carbohydrate-free proteins. Combining top-down fragmentation with IXC and online native ESI MS expands the capability of this technique dramatically by enabling its application to significantly larger and more complex (glyco)protein samples.

4.4. Conclusions

Online LC/MS and LC/MS/MS analysis of proteins and related macromolecules has long been a workhorse in many fields of bioanalysis, including the biopharmaceutical sector. Although the majority of applications make use of reversed-phase chromatography, other modes of separation are also actively evaluated. Of particular interest are methods that do not trigger protein denaturation during the separation step, as they enable analysis of not only covalent (primary) structure, but also higher order structure (including conformation and quaternary assembly). IXC is shown in this work to be a very attractive candidate for these analyses, as it can be applied to a wide range of challenging targets of interest to biopharmaceutical and biotechnology sectors, can be operated under conditions that do not disrupt protein higher order structure, and can be interfaced with native ESI MS. The latter provides information on both identity of components of protein-based products (by measuring masses) and their conformational integrity (via analysis of ionic charge state distributions), while the former allows isomers (as well as isoforms with nearly indistinguishable masses) to be separated. Incorporation of the top-down fragmentation step in the IXC/ESI MS workflow allows stress-induced nonenzymatic PTMs to be detected and localized in very complex and heterogeneous samples on the background of extensive enzymatic PTMs, such as glycosylation. Although the size of the largest protein therapeutic interrogated using the new IXC/ESI MS method in this work is relatively modest (23 kDa), a work is currently underway in our laboratory to expand the

reach of this new methodology to significantly larger and more heterogeneous targets, including PEGylated glycoproteins.

FIGURES

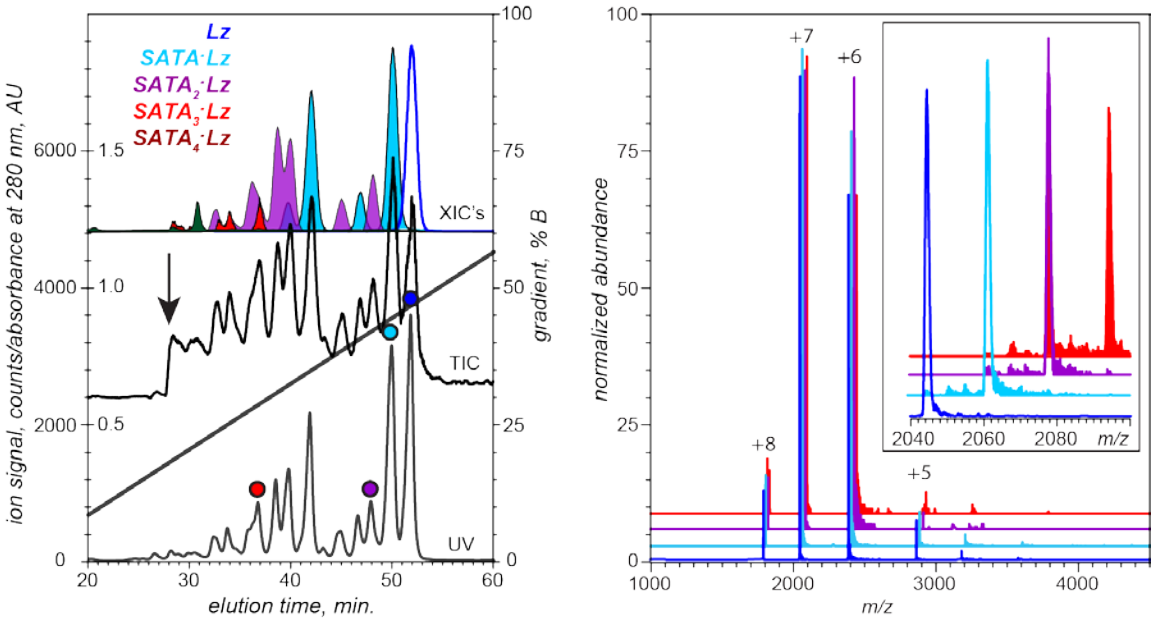

Figure 4.1 IXC/ESI MS analysis of a mixture of SATA-conjugated Lz: a UV chromatogram, a total ion chromatogram, and extracted ion chromatograms for unmodified Lz and conjugates with different loadings (left panel) and ESI mass spectra (right panel) averaged across chromatographic peaks as indicated with colored dots on the chromatogram. The inset shows zoomed views of charge states +7 (note that the top spectrum contains ionic contributions from both $SATA_2$·Lz and $SATA_3$·Lz). The total amount of injected protein was 0.1 mg. The arrow indicates a point in time when the eluate was directed to the ESI source.

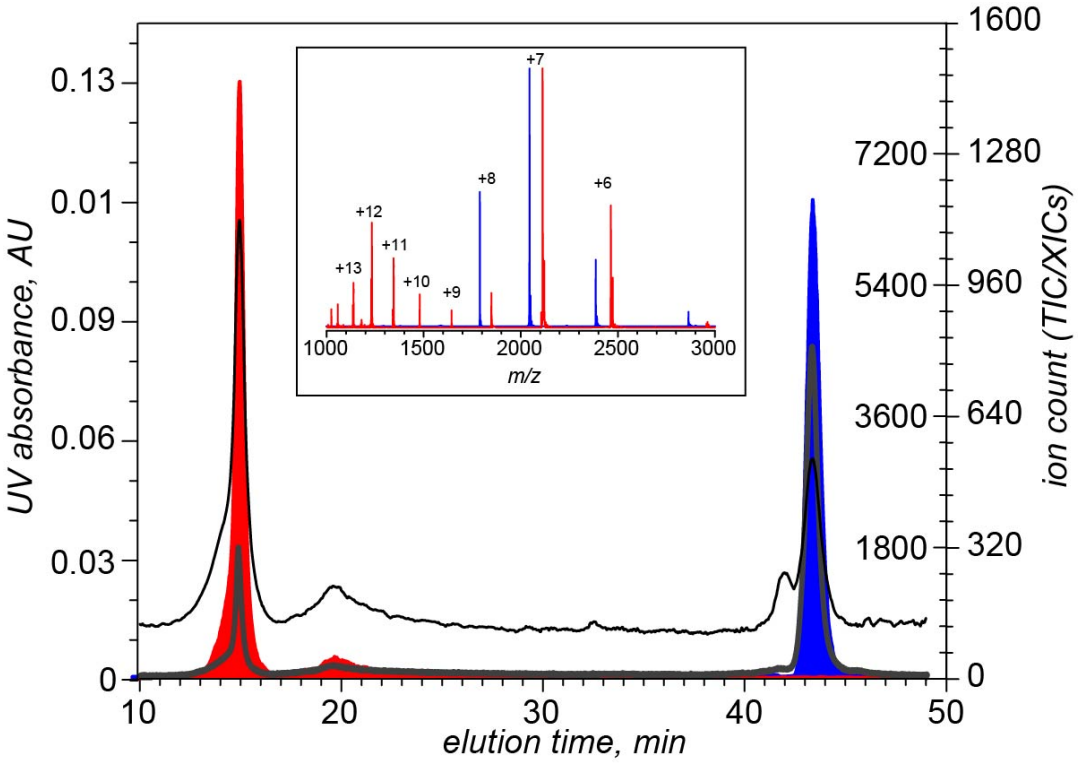

Figure 4.2 IXC/ESI MS analysis of a mixture of intact (fully oxidized) Lz and a fully reduced and alkylated Lz: a UV chromatogram (gray trace), a TIC chromatogram (black trace), and extracted ion chromatograms for ions representing the fully reduced (red) and intact Lz (blue). The inset shows mass spectra averaged across the two chromatographic peaks.

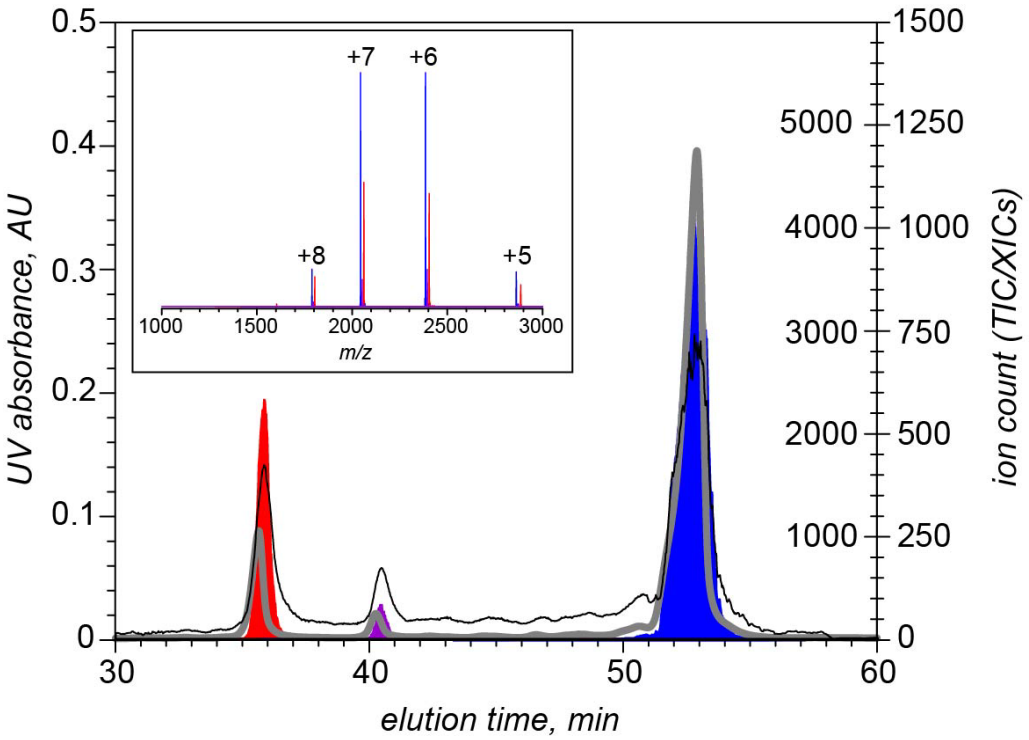

Figure 4.3 IXC/ESI MS analysis of Lz following selective reduction of a single disulfide bond: a UV chromatogram (gray trace), a TIC chromatogram (black trace), and extracted ion chromatograms for ions representing the species with three disulfide bonds (red), oxidized intact (blue), and oxidized singly alkylated Lz (purple). The inset shows mass spectra averaged across the three chromatographic peaks.

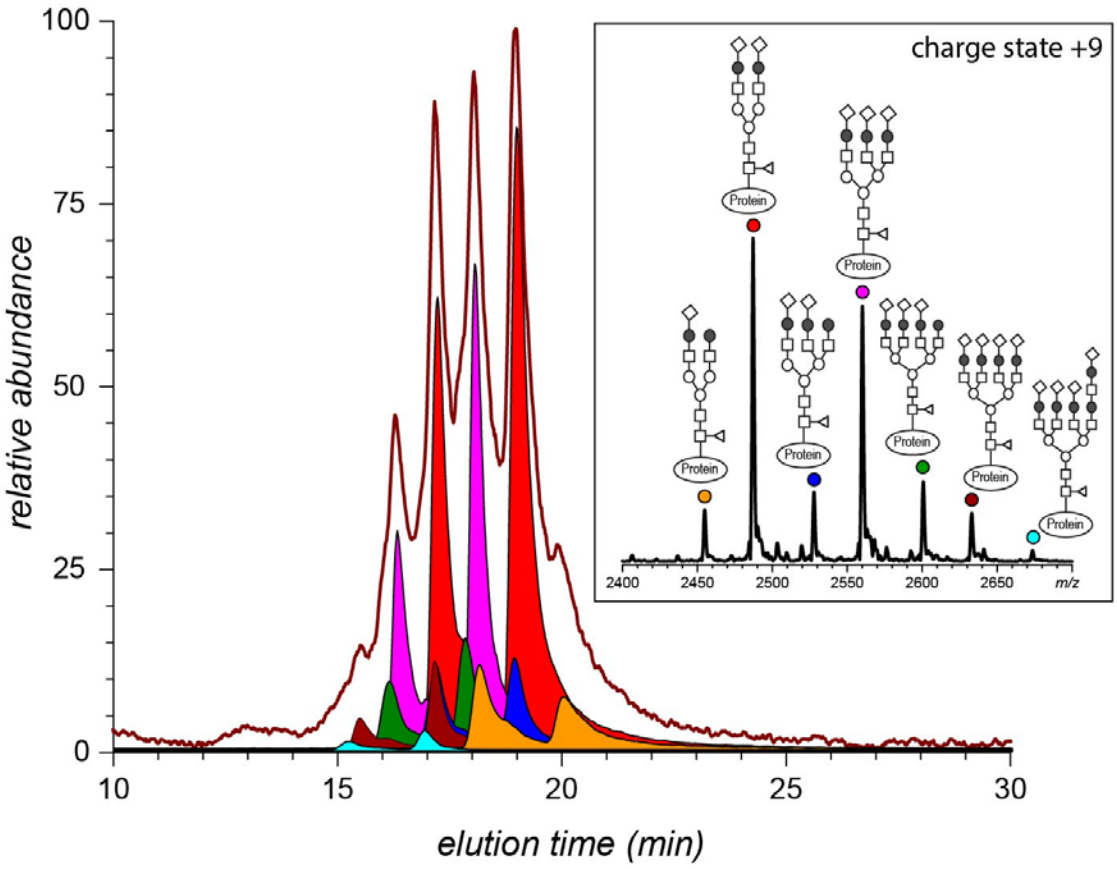

Figure 4.4 IXC/ESI MS analysis of stressed interferon-β sample: TIC chromatogram and extracted ion chromatograms of seven different glycoforms identified in the mass spectrum of the protein sample not subjected to separation (shown in inset). The carbohydrate chain structures were assigned based on measured mass and are also shown in the inset (structure key: squares, GlcNac residues; triangles, fucose residues; open circles, mannose residues; filled circles, galactose residues; diamonds, sialic acid residues). The total amount of protein injected was 4 µg.

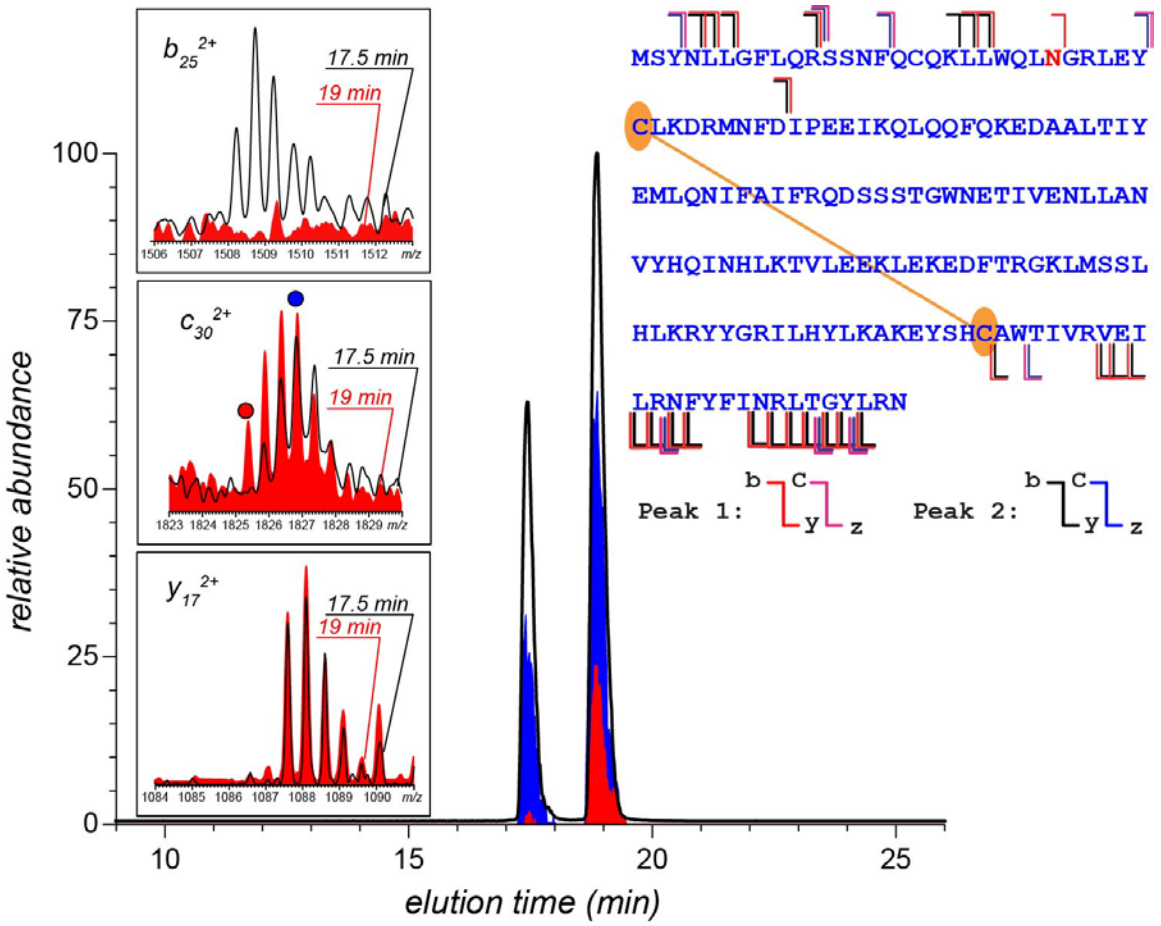

Figure 4.5 Identification and localization of a stress-induced nonenzymatic PTM within the BiNA2 species of interferon-β by IXC/ESI MS/MS: TIC for all fragment ions produced by ETD and collisional activation of precursor ions corresponding to the +9 charge state of the BiNA2 species (black trace) and extracted ion chromatograms corresponding to two different isotopic peaks within the c_{30}^{2+} fragment ion (blue, *m*/*z* 1826.9; red, *m*/*z* 1825.4). The insets on the left show isotopic distributions of fragment ions that are observed in both chromatographic peaks (c_{30}^{2+} and y_{17}^{2+}) and unique to the earlier-eluting species (b_{25}^{2+}). The inset on the right shows overall fragmentation patterns for the two chromatographic peaks.

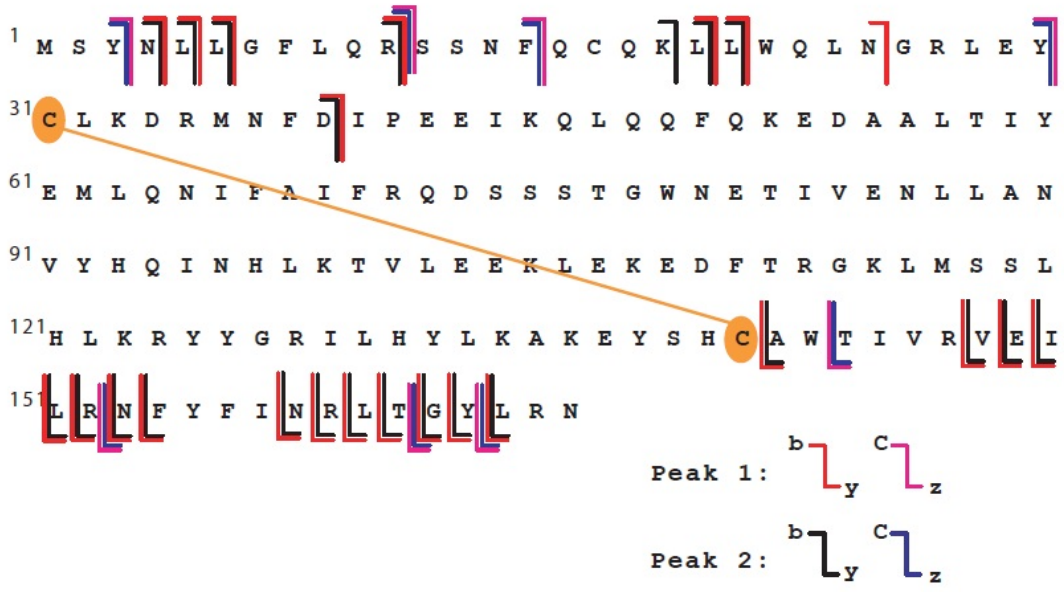

Figure S6. Top-down fragmentation patterns observed for interferon-β species eluting at 17.5 min (peak 1) and 19 min (peak 2) in the chromatogram shown in Figure 4.5.

Table S.1. Relative signal intensities of SATA$_n$Lz conjugates from UV chromatogram and XIC's using the heights of peaks resolved in the UV chromatogram

Elution time of peaks (min.)	Relative abundances, %	
	UV chromatogram	Extracted ion chromatogram (XIC)
52.0	100	100
50.1	99.7	88.2
48.1	28.6	29.4
46.9	23.1	23.5
45.1	15.3	17.6
42.1	73.0	63.5

CHAPTER 5

CHARACTERIZATION OF PEGYLATED-PROTEIN CONJUGATES BY WEAK CATION EXCHANGE CHROMATOGRAPHY WITH ONLINE DETECTION BY NATIVE ESI MS SUPPLEMENTED WITH GAS PHASE REACTIONS

5.1. Overview

Covalent attachment of poly ethylene glycol (PEG) to proteins has demonstrated to be the most successful approach to improve pharmacokinetic and pharmacodynamics profile of protein drugs.[131] PEGylation increases the efficacy of protein drugs by improving solubility, stability over wide pH and temperature range, decreasing renal clearance, and limiting toxicity and immunogenicity.[132] All of these eventually contribute to less frequent administration of PEGylated protein drug to achieve the same therapeutic outcome as its non-PEGylated form. However, due to polydispersity of PEG molecule, PEGylation of protein increases complexity of the sample making its analytical characterization very challenging.

Structural heterogeneity of PEGylated proteins occurs due to the presence of complex species differing in the number of PEG chains conjugated, sites of conjugation, and lengths of PEG chains. Many analytical methods have been reported to analyze the presence of positional isomers in PEGylated proteins, some of these involved enzymatic treatment to leave behind chemical tags to identify site of conjugation,[133] use of new PEG derivatives,[134] and using chromatographic separation followed by top-down MS analysis.[135] Strategies have also been reported for site specific conjugation of proteins with PEG.[9, 136] Apart from heterogeneity due to PEG chain, an additional level of

complexity could arise in PEGylated proteins due to the presence of enzymatic post translational modifications (e.g., glycosylation) or non-enzymatic post translational modifications (e.g., oxidation, deamidation).

Electrospray ionization mass spectrometry (ESI MS) has demonstrated to be very useful for structural characterization of biopharmaceutical products.[32] While native ESI MS has been particularly useful as a tool of providing information on mass, conformational integrity and interactions of protein drugs, it often fails for highly heterogeneous samples (e.g., PEGylated proteins for reasons mentioned in the preceding paragraphs). To analyze such samples ESI MS supplemented with limited charge reduction introduced by our lab has shown a lot of potential, and has indeed been applied to analyze highly glycosylated proteins and their complexes, and soluble protein aggregates.[70, 137] In chapters 2 and 4 native ESI MS based methods have also been presented in which chromatographic separation of various protein species by size exclusion chromatography or ion exchange chromatography (IXC) helped in the analysis of protein heterogeneity.

IXC is also helpful in the analysis of PEGylated proteins; a method was reported from our lab by combining separation capabilities of IXC and top-down tandem MS to monitor positional isomers and map PEGylation sites on a 8.6 kDa ubiquitin protein conjugated with 5kDa PEG chain.[135] It also presented important fundamental aspects to reduce the complexity of mass spectrum (related to polydispersity of PEG) using collisional activated dissociation (CAD). In this work a new analytical strategy by combining ion exchange chromatography with online native ESI MS detection is

presented to analyze proteins conjugated with larger PEG chain. Limited charge reduction (LCR) or CAD is implemented in online LC/MS mode for characterization of PEGylation stoichiometry and different post translational modifications. This strategy have been applied to a complex mixture containing 22 kDa glycoprotein conjugated with a 20kDa PEG chain.

5.2. Materials and Methods

Stressed (partially deamidated) sample of PEGylated glycoprotein was provided by Biogen (Cambridge, MA). All other chemicals used in this work were of analytical grade or higher. Online IXC/MS analysis of stressed PEGylated glycoprotein was carried out directly by diluting the sample in 100 mM ammonium acetate, pH 5.5 to a required concentration.

Ion-exchange chromatography was performed on an Agilent 1100 HPLC (Agilent Technologies, Santa Clara, CA) equipped with a 2 mm × 250 mm weak cation-exchange column ProPac WCX-10 (Thermo Scientific Waltham, MA). Ammonium acetate at pH 5.5 was used for MS-compatible gradient elution (100 mM in mobile phase A and 1 M mobile phase B). The total flow rate was maintained at 0.2 mL/minute. Linear gradient reaching 25% B over 50 minutes after equilibration with mobile phase A for 3 minutes was used. Concentration of PEGylated glycoprotein used was 1mg/ml and was injected using a 20 µl sample loop. Upon exiting the column, the eluate was directed to online detection by native ESI MS without any solvent modification or flow splitting using Synapt G2Si (Waters, Milford, MA) hybrid quadrupole/time-of-flight mass spectrometer. LCR was carried using electron transfer dissociation (ETD) on narrow population of ions

isolated at m/z 1150 with 10 m/z isolation window. The conditions were optimized as follows: capillary voltage, 2.5 kV, sampling cone, 5.0 V; source offset, 80 V; desolvation temperature, 400 °C; source temperature, 120 °C, cone gas flow, 45 L/h; desolvation gas flow, 600 L/h; make up gas flow, 35L/h; trap gas, 33.0 ml/min; transfer gas, 0.8 ml/min; trap wave height, 0.08 V. CAD was carried out on ions isolated at m/z 1050 using 10 m/z window in transfer cell by setting the collision energy at 25.0V and transfer gas to 2.0 ml/min.

5.3. Results and Discussion

Separation of this stressed PEGylated glycoprotein by WCX yielded four partially resolved peaks along with an inflection point at 27-28 minutes interval (figure 5.1). The species resolved by WCX could represent positional isomers and species with different number of PEG chains conjugated to protein. In the case of this protein the presence of positional isomers can be ruled out as PEGylation chemistry was optimized to conjugate specifically at N terminal. However extensive heterogeneity due to the presence of various glycoforms and deamidated forms could also lead to occurrence of different PEGylated species. All these factors contributing to heterogeneity along with multiple ESI MS charge states resulted in completely unresolved mass spectra detected for all of the WCX peaks as shown in the upper right panel of figure 5.1. These mass spectra imply occurrence of ion at each m/z in 800 – 1400 m/z range.

To analyze this sample LCR using ETD in the online WCX/ESI MS mode was implemented. For the entire duration of LC, LCR was carried out on a narrow population of ions isolated at m/z 1150. Charge state ladder resulting from this experiment enabled

detection of the mass of isolated ions. In the bottom right panel of figure 5.1, a mass spectrum obtained by averaging ESI MS LCR scans for ions (m/z 1150) eluting at 30 – 32 minutes is shown. Deconvolution of this mass spectrum gave a mass of ~45 kDa, which is in agreement with 1:1 PEGylation of glycoprotein taking into account molecular weight distribution of 20kDa PEG chain and molecular weight of this glycoprotein (22kDa). ESI MS LCR data of m/z 1150 ions detected from other WCX peaks also gave identical charge ladder deconvoluting to a mass consistent with 1:1 PEGylation stoichiometry. This way by using LCR in conjunction with WCX/ESI MS, detection of PEGylation stoichiometry was made possible, and all the stressed PEGylated glycoprotein species resolved partially by WCX (figure 5.1) were found to be conjugated with one chain of PEG. Unfortunately due to continuum of mass and broad LCR peaks we were not able to differentiate various glycoforms.

For further analysis of this sample, a WCX/online ESI MS method was developed in which CAD was applied to a narrow population of ions at m/z 1050 (isolated over entire LC timescale). As seen in figure 5.2 the complexity of mass spectrum is decreased significantly leading to detection of clusters of distinguishable ionic peaks. Charge state assignment for each of these ionic clusters was carried out easily by dividing the adjacent peak spacing with the repeating mass of PEG unit (44 *u).* The mass of these ions correspond to cleavage of PEG chains following the route that produces B and C ions as proposed by Lattimer and his colleagues.[138] Similar ions were also produced, when insource CAD was applied to PEGylated Ubiquitin as reported in the work carried out previously in our lab.[135] CAD conditions were optimized to cleave most of the PEG

chain keeping the polypeptide backbone of the glycoprotein intact (figure 5.2). Multiply charged ion clusters corresponding to six PEG molecules bound to intact glycoprotein (B_6^{+14} - B_6^{+10} ions) were detected. These features led to simplified mass spectra enabling identification of different stressed PEGylated glycoforms present in the sample as they are separated by WCX (see right panel of figure 5.2).

Careful analysis of this data revealed the presence of PEGylated species bearing BiNA2, TriNA3, BiNA1, and TriLACNA3 carbohydrates (see figures 5.3 and 5.4). Mass spectrum correlating with PEG BiNA2 glycoforms was recorded in two chromatographic peaks eluting at 28.2 – 29.3 and 31.9 – 33.5 minutes interval. Similarly based on ESI MS mass measurements TriNA3 was found to elute at 25.0 – 28.0 and 30.0 – 31.4 minutes interval. Identification of these glycoforms based on mass was further confirmed by detection of fragment ions produced from cleavage of terminal carbohydrate chain residues (sialic acid). As shown in figure 5.3, in addition to intact BiNA2 ion mass spectra detected at 28.2 – 29.3 and 31.9 – 33.5 minutes also showed ions whose mass was consistent with loss of one and two sialic acids. BiNA1 was confirmed to elute in the last chromatographic peak based on mass and detection of fragment ions from loss of only one sialic acid residue. Similarly TriNA3 glycoform (figure 5.4) showed ions from fragmentation of up to 3 sialic acids. Though BiNA1 ion seen in figure 5.4 could be due to fragmentation of GlcNAc and galactose units from TriNA2 form, it most likely appears in this mass spectrum due to tailing of chromatographic peak (see extracted ion chromatogram, XIC in figure 5.5).

XIC plotted in a straightforward way using the corresponding ions led to an artificial increase in XIC peak signal due to absence of baseline resolved mass spectra (see mass spectra in figure 5.2). To eliminate this noise, XICs were produced by subtracting ionic intensities of the two most abundant peaks belonging to the same cluster of glycoform ions (e.g., XIC of BiNA2 glycoform using the BiNA1 fragment ions was made by subtracting XIC's abundance of 2033.0 and 2037.6 m/z ions). It is not surprising to see the presence of two chromatographic peaks for each glycoform (BiNA2 and TriNA3) of this stressed PEGylated glycoprotein; a similar observation was made in chapter 3 when interferon was analyzed by WCX/ESI MS. Therefore the earlier eluting chromatographic peaks of each glycoform most likely represent PEGylated deamidated fraction of TriNA3 and BiNA2.

5.4. Conclusions

Growing trend of heterogeneity in biopharmaceuticals occurring either naturally or induced by chemical modification (PEGylation in this work) presents a serious challenge to analytical characterization. In this chapter analytical strategy has been presented, which combined chromatographic separation along with various gas phase chemistry techniques (ETD, CAD) in ESI MS. WCX separation has proved to be very powerful in delineating the presence of non-enzymatic PTM (deamidation) on intact protein glycoforms, which is not possible using bottom-up MS approach or standard RPLC/MS. The feasibility of carrying out these kinds of gas phase measurements in online mode helps by decreasing the analysis time, and also illustrates a novel method to analyze such highly heterogeneous samples. Interfacing WCX/ESI MS with LCR has

been demonstrated to be suitable for characterization of stoichiometry of PEGylated sample. This WCX/MS LCR workflow can also be used for characterization of non-covalent protein complexes involving such PEGylated proteins. For a thorough characterization of PEGylated proteins implementing CAD just enough to cleave most of the PEG chain has shown to be a promising approach to reduce complexity of mass spectrum to identify various PTMs.

FIGURES

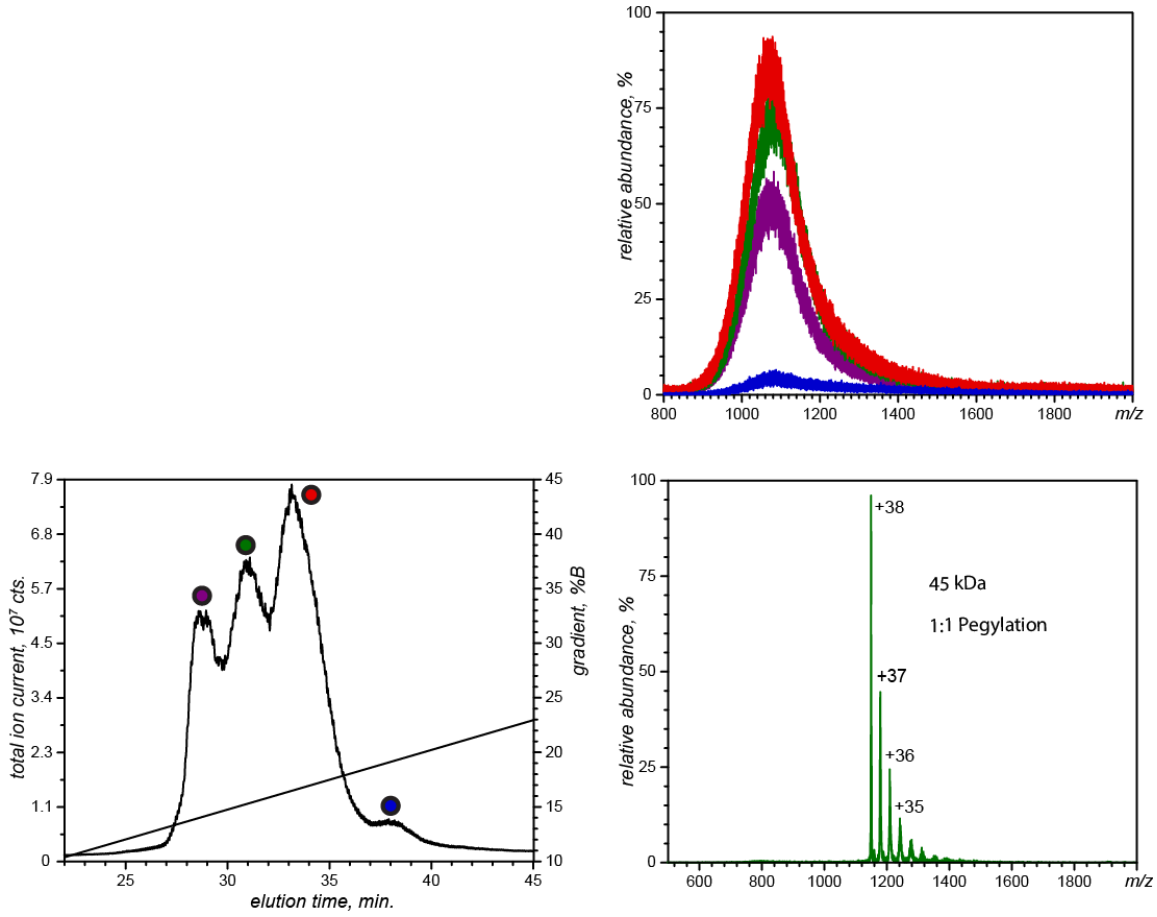

Figure 5.1 IXC/ESI MS analysis of stressed PEGylated glycoprotein: total ion chromatogram (TIC) shown in black trace (lower left panel) and the upper right panel shows mass spectra averaged across the correspnding chromatographic peaks marked with same colored circles. Lower right panel is from IXC ESI MS LCR analysis; charge state ladder detected for the isolated ions (m/z 1150) of chromatographic peak marked with green circle. LCR charge state ladders for the ions detected in other chromatographic peaks were similar giving a mass of 45kDa.

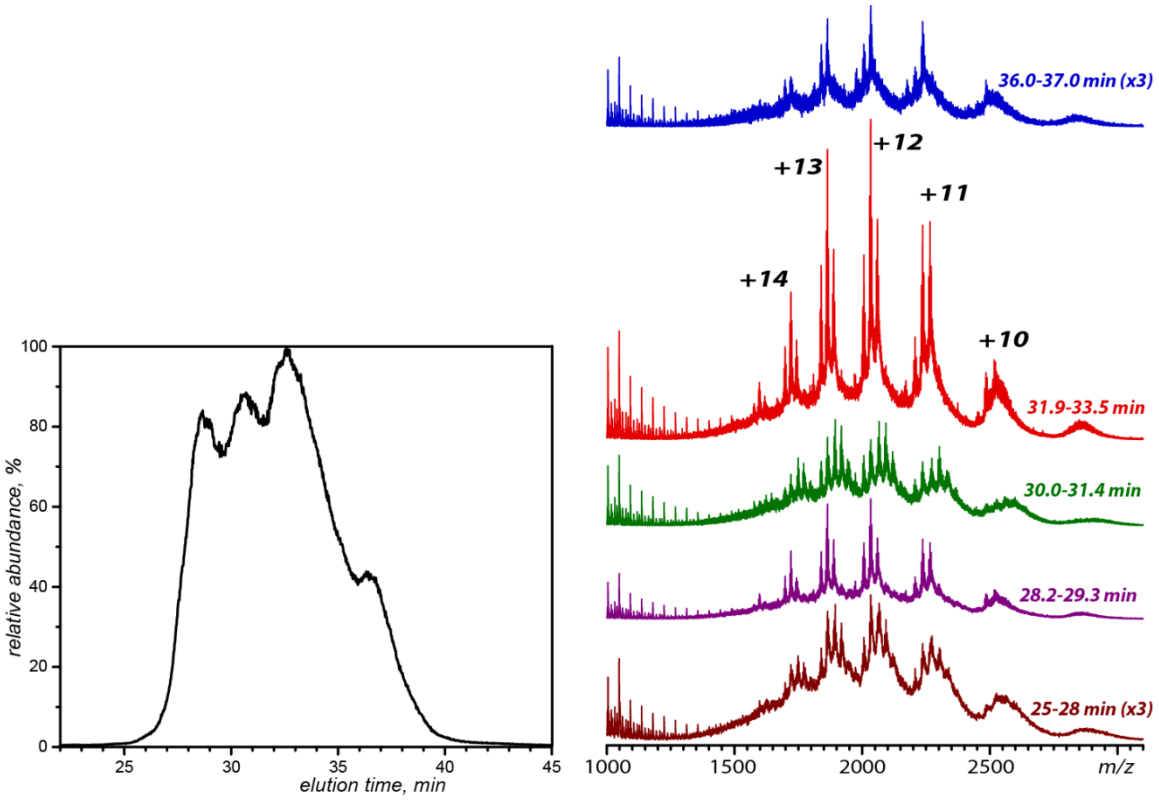

Figure 5.2 IXC/ESI MS CAD analysis of stressed PEGylated glycoprotein: a total ion chromatogram (TIC) obtained from CAD of ions isolated at m/z 1050 is shown in black trace (lower left panel) and the right panel shows mass spectra averaged across each chromatographic peak. BiNA2 glycoform eluted at 28.2 - 29.3 and 31.9 - 33.5 minutes; TriNA2 and TriLACNA2 glycoforms at 25.0 - 28.0 and 30 .0 – 31.4 minutes; BiNA1 glycoform at 36.0 - 37.0 minutes interval.

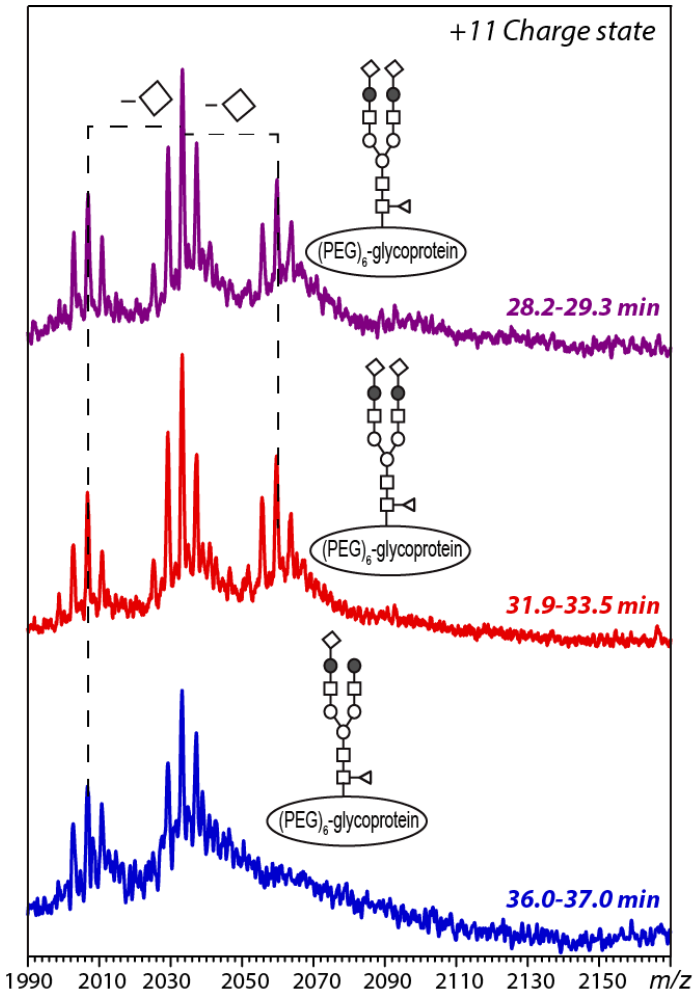

Figure 5.3 Zoom of +11 charge state of the mass spectra detected from IXC/ESI MS CAD of stressed PEGylated glycoprotein. The carbohydrate ions were assigned based on measured mass after subtraction of 6 PEG units' mass. (structure key: squares, GlcNac residues; triangles, fucose residues; open circles, mannose residues; filled circles, galactose residues; diamonds, sialic acid residues). BiNA2 glycoform was detected in peaks eluting at 28.2 – 29.3 and 31.9 – 33.5 minutes range. BiNA1 eluted at 36.0 – 37.0 minutes interval.

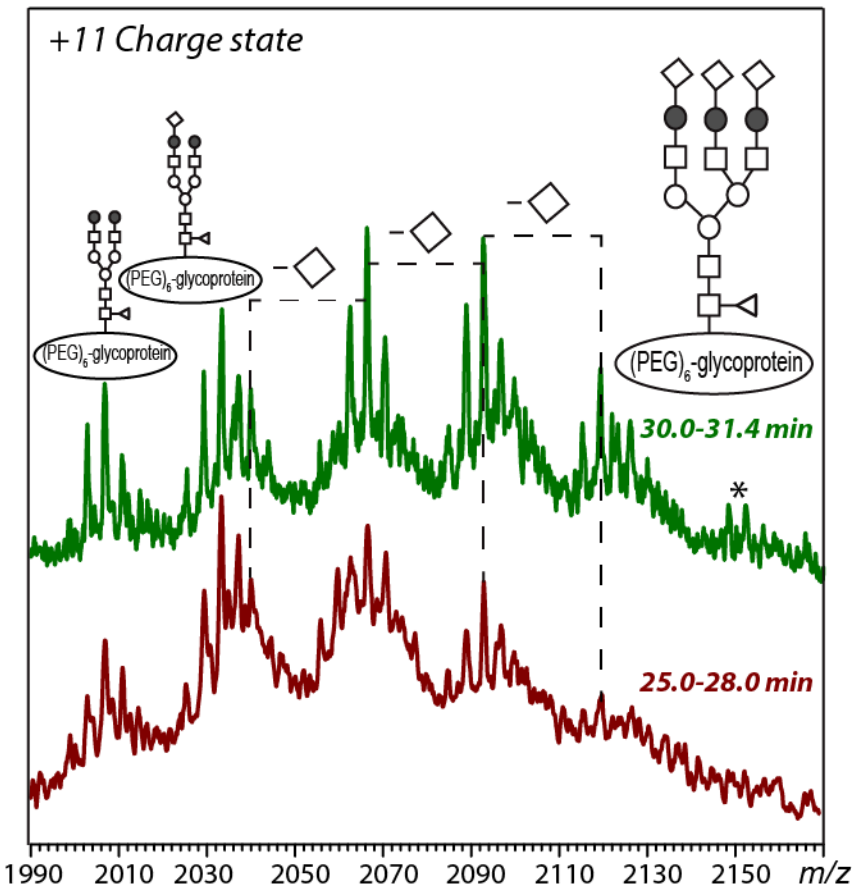

Figure 5.4 Zoom of +11 charge state of the mass spectra detected from IXC/ESI MS CAD of stressed PEGylated glycoprotein. These mass spectra correspond to ions detected in TriNA3 glycoform. Ions marked with asterisk correspond to TriLACNA3. BiNA1 and BiNA0 glycoform fragment ions are observed due to tailing of BiNA2 glycoform. The carbohydrate ions were assigned based on measured mass after subtraction of 6 PEG units' mass. (structure key: squares, GlcNac residues; triangles, fucose residues; open circles, mannose residues; filled circles, galactose residues; diamonds, sialic acid residues).

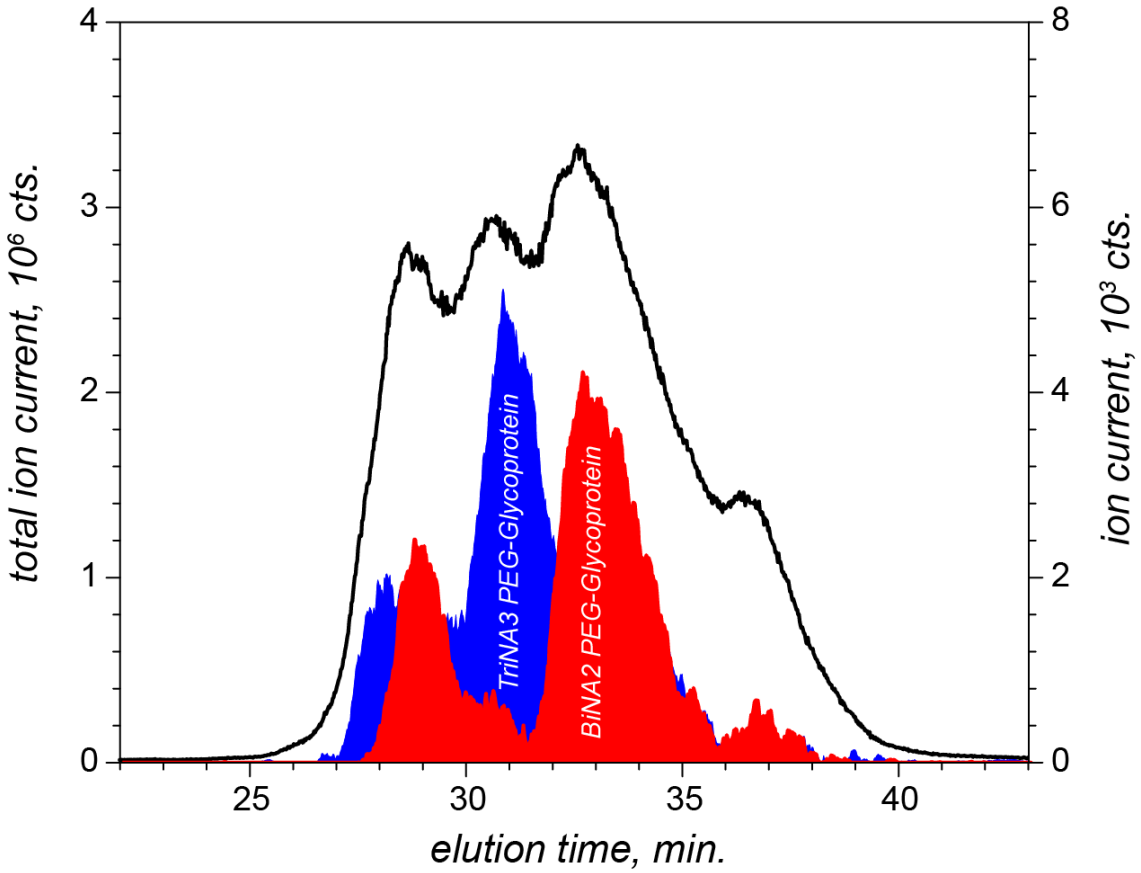

Figure 5.5 Extracted ion chromatograms (XICs) of BiNA2 and TriNA3 PEGylated glycoforms detected from WCX/ESI MS CAD analysis. XIC of BiNA2 glycoform was made using BiNA1 fragment ions detected from fragmentation of sialic acid residue from intact BiNA2 form.

CHAPTER 6

CONCLUSIONS AND FUTURE OUTLOOK

Novel analytical methods combining non-denaturing liquid chromatography separation with native ESI MS detection were developed in this work. These methods have applications in characterization of biopolymer structure and interaction to help in the development and quality control of biotherapeutics.

SEC with online native ESI MS detection method developed in chapter 2 allowed characterization of higher order structure of proteins. It was shown to be useful in probing conformational integrity and detection of soluble protein aggregates of a model protein (bovine serum albumin), and characterize rapidly interconverting oligomeric states of a recombinant glycoprotein human arylsulfatase A.

Using the same SEC/ESI MS workflow a method to measure kinetics and binding affinity of transient protein interactions have also been developed. This method involved fitting of experimental XICs with the profile generated using mathematical solutions. Validation of this new method for analysis of protein interactions with 1:1 and 2:1 binding stoichiometry was carried out using the known binding of C-lobe of human transferrin and transferrin variants, respectively with human transferrin receptor. Further validation of this method will be carried out by changing the concentration of interacting proteins, which will change binding kinetics, but should give the same equilibrium dissociation constant. This method has significant potential as it does not

require immobilization or labelling of proteins, works with small amount of sample, and could be applied for complexes interacting with higher stoichiometries.

Combination of IXC with online native ESI MS detection was developed and shown to be effective in characterization of conformational integrity and positional isomers of protein conjugates (using chemically modified lysozyme as a model) and PTMs of a biotherapeutic recombinant human interferon beta 1a. Top-down fragmentation in this workflow (IXC/MS/MS of whole proteins) allowed non-enzymatic PTMs to be identified in situations when the mass measurement alone was not sufficient for an unequivocal conclusion to be reached. To characterize highly heterogeneous PEGylated glycoprotein IXC/online ESI MS methods supplemented with gas phase reactions (limited charge reduction and collisional activation) were developed. These methods allowed detection of PEGylation stoichiometry and glycoform distribution of intact PEGylated protein. A prerequisite for methods developed in chapter 2 – 4 is that separation needs to be carried out in MS friendly volatile mobile phase.

Separation by SEC applies for systems containing globular proteins, however it is most likely to fail for complexes containing non-globular biopolymers (e.g., protein/DNA, protein/glycosaminoglycans), where differences in hydrodynamic radii is not significant to allow separation based on size. To characterize transient complexes containing such species, other modes of non-denaturing chromatography will be implemented. IXC with online ESI MS detection will be particularly useful in this

situation. Fitting XICs detected in IXC/ESI MS with the ADR equations similar to those developed in chapter 3, kinetics and binding affinity of the complexes can be measured. Unlike size filtration in SEC, linear velocities in IXC will depend on the interaction of the species with the charged resin. This method will be applied for characterization of protein/DNA complexes, which plays an important role in gene regulation.

The feasibility of carrying out limited charge reduction in online LC/MS opens exciting opportunities. Implementing this feature in the SEC/ESI MS method developed in chapter 3 and in IXC/ESI MS will enable qualitative analysis of interaction of highly heterogeneous biopolymers and to measure kinetics and binding affinity. Transient complexes that will be characterized using these methods include highly glycosylated proteins (e.g., hemoglobin-drug conjugate binding to haptoglobin), protein-polymer conjugates, and linear polysaccharides.

APPENDIX

APPENDIX TO CHAPTER 3

Derivation of partial derivative equation for rhASA (chapter 3)

Assuming a short lived intermediate tetramer (M_4); $M_2 \rightleftharpoons M_8$ transition of rhASA can be understood by the following rate equation[80]

$$2M_2 \rightleftharpoons_{k_{-2}\,(fast)}^{k_2\,(slow)} M_4;\ 2M_4 \rightleftharpoons_{k_{-4}\,(slow)}^{k_4\,(fast)} M_8 \quad (1)$$

The cumulative rates of octamer production and loss under steady state condition will be

$$R_{dissoc} = k_{-4}C_{M8};\ R_{assoc} = \frac{1}{2}k_2 C_{M2}^2 \quad (2)$$

Concentration of M_8 at any point *x* across the column length ($0 < x < L$) will also be affected by longitudinal transfer and diffusion.

Mass transfer will be characterized by hydrodynamic flux through column cross-section *S.* Hydrodynamic flux (J_{M8}) of M_8 depends on the linear velocity (u_{M8}) and concentration gradient (C_{M8})

$$J_{M8}(x,t) = u_{M8} \cdot C_{M8}(x,t) \quad (3)$$

Diffusion is determined by diffusion coefficient (D_{M8}) and concentration gradient for M_8

$$J_{M8}(x,t) = -D_{M8} \cdot \frac{\partial C_{M8}(x,t)}{\partial x} \quad (4)$$

Total change of the number of M_8 molecules within the infinitely small slice of the column volume ($S.\delta x$) over an infinitely small period of time δt will reflect a difference

between the rates of association and dissociation, as well as the difference between the incoming flux of M_8 and the outgoing flux

$$\delta C_{M8} \cdot S\delta x = -S \cdot u_{M8} \cdot \delta C_{M8} \cdot \delta t + S \cdot D_{M8} \cdot \delta\left(\frac{\partial C_{M8}(x,t)}{\partial x}\right) \cdot \delta t - S\delta x \cdot R_{dissoc} \cdot \delta t + S\delta x \cdot R_{assoc} \cdot \delta t \quad (5)$$

Association and dissociation terms are defined by eq.2. Since *x* and *t* are independent variables, eq. 5 can be re-written in the form of a partial differential equation, and a similar equation can be constructed for the dimer concentration.

$$\begin{cases} \frac{\partial C_{M8}}{\partial t} = -u_{M8} \cdot \frac{\partial C_{M8}}{\partial x} + D_{M8} \cdot \frac{\partial^2 C_{M8}}{\partial x^2} - k_{-4} C_{M8} + \frac{1}{2} k_2 C_{M2}^2 \\ \frac{\partial C_{M2}}{\partial t} = -u_{M2} \cdot \frac{\partial C_{M2}}{\partial x} + D_{M2} \cdot \frac{\partial^2 C_{M2}}{\partial x^2} + 4k_{-4} C_{M8} - 2k_2 C_{M2}^2 \end{cases} \quad (6)$$

List of parameters used in the calculations of partial differential equations for $FeTf_C$ and TfR interaction

u_P = 1.465 cm/min

u_R = 1.740 cm/min

u_{PR} = 1.845 cm/min

D_P = 0.0027cm^2/min

D_R = 0.0134 cm^2/min

D_{PR} = 0.0122cm^2/min

List of parameters used in the calculations of partial differential equations for hTf and TfR interaction

u_P = 1.71 cm/min

u_R = 1.74 cm/min

u_{PR} = 1.87 cm/min

u_{P2R} = 2.08 cm/min

D_P = 0.0026 cm^2/min

D_R = 0.0134 cm^2/min

D_{PR} = 0.0110 cm^2/min

D_{P2R} = 0.0090 cm^2/min

List of parameters used in the calculations of partial differential equations for Fe_2 bTf and TfR interaction

u_P = 1.67 cm/min

u_R = 1.74 cm/min

u_{PR} = 1.80 cm/min

u_{P2R} = 2.00 cm/min

D_P = 0.0026 cm^2/min

D_R = 0.0134 cm^2/min

D_{PR} = 0.0110 cm^2/min

D_{P2R} = 0.0090 cm^2/min

List of parameters used in the calculations of partial differential equations for bTf and TfR interaction

u_P = 1.70 cm/min

u_R = 1.74 cm/min

u_{PR} = 1.87 cm/min

u_{P2R} = 2.08 cm/min

D_P = 0.0026 cm^2/min

D_R = 0.0134 cm^2/min

D_{PR} = 0.0110 cm^2/min

D_{P2R} = 0.0090 cm^2/min

Initial (injection) concentrations of all the components were represented as boxcar functions with a width σ = 0.21 cm

REFERENCES

[1] Boehncke, W.-H., and Radeke, H. H. (2007) Biologics in general medicine, Springer.

[2] Fosgerau, K., and Hoffmann, T. (2015) Peptide therapeutics: current status and future directions, *Drug Discovery Today 20*, 122-128.

[3] Page, C. (2013) Heparin and Related Drugs: Beyond Anticoagulant Activity, *ISRN Pharmacology 2013*, 910743.

[4] Alvarez-Salas, L. M. (2008) Nucleic acids as therapeutic agents, *Current topics in medicinal chemistry 8*, 1379-1404.

[5] Walsh, G. (2014) Biopharmaceutical benchmarks 2014, *Nat Biotech 32*, 992-1000.

[6] Kinch, M. S. (2015) An overview of FDA-approved biologics medicines, *Drug Discovery Today 20*, 393-398.

[7] Kiss, Z., Elliott, S., Jedynasty, K., Tesar, V., and Szegedi, J. (2010) Discovery and basic pharmacology of erythropoiesis-stimulating agents (ESAs), including the hyperglycosylated ESA, darbepoetin alfa: an update of the rationale and clinical impact, *European journal of clinical pharmacology 66*, 331-340.

[8] Huang, C. (2009) Receptor-Fc fusion therapeutics, traps, and MIMETIBODY technology, *Curr Opin Biotechnol 20*, 692-699.

[9] Baker, D. P., Lin, E. Y., Lin, K., Pellegrini, M., Petter, R. C., Chen, L. L., Arduini, R. M., Brickelmaier, M., Wen, D., Hess, D. M., Chen, L., Grant, D., Whitty, A., Gill, A., Lindner, D. J., and Pepinsky, R. B. (2006) N-Terminally PEGylated Human Interferon-β-1a with Improved Pharmacokinetic Properties and in Vivo Efficacy in a Melanoma Angiogenesis Models, *Bioconjugate Chemistry 17*, 179-188.

[10] Sassoon, I., and Blanc, V. (2013) Antibody-drug conjugate (ADC) clinical pipeline: a review, *Methods in molecular biology (Clifton, N.J.) 1045*, 1-27.

[11] Hruby, V. J. (2002) Designing peptide receptor agonists and antagonists, *Nat Rev Drug Discov 1*, 847-858.

[12] Vlieghe, P., Lisowski, V., Martinez, J., and Khrestchatisky, M. (2010) Synthetic therapeutic peptides: science and market, *Drug Discovery Today 15*, 40-56.

[13] Craik, D. J., Fairlie, D. P., Liras, S., and Price, D. (2013) The Future of Peptide-based Drugs, *Chemical Biology & Drug Design 81*, 136-147.

[14] Sasisekharan, R., and Venkataraman, G. (2000) Heparin and heparan sulfate: biosynthesis, structure and function, *Current Opinion in Chemical Biology 4*, 626-631.

[15] Keire, D., Mulloy, B., Chase, C., Al-Hakim, A., Cairatti, D., Gray, E., Hogwood, J., Morris, T., Mourão, P., Soares, M. D. L. C., and Szajek, A. (2015) Diversifying the Global Heparin Supply Chain: Reintroduction of Bovine Heparin in the United States?, *Pharmaceutical Technology 39*.

[16] Zeck, A., Regula, J. T., Larraillet, V., Mautz, B., Popp, O., Göpfert, U., Wiegeshoff, F., Vollertsen, U. E. E., Gorr, I. H., Koll, H., and Papadimitriou, A. (2012) Low Level Sequence Variant Analysis of Recombinant Proteins: An Optimized Approach, *PLoS ONE 7*, e40328.

[17] Fu, J., Bongers, J., Tao, L., Huang, D., Ludwig, R., Huang, Y., Qian, Y., Basch, J., Goldstein, J., Krishnan, R., You, L., Li, Z. J., and Russell, R. J. (2012) Characterization and identification of alanine to serine sequence variants in an IgG4 monoclonal antibody produced in mammalian cell lines, *Journal of Chromatography B 908*, 1-8.

[18] Harris, R. J. (2005) Heterogeneity of recombinant antibodies: linking structure to function, *Developments in biologicals 122*, 117-127.

[19] Solá, R. J., and Griebenow, K. A. I. (2009) Effects of Glycosylation on the Stability of Protein Pharmaceuticals, *Journal of pharmaceutical sciences 98*, 1223-1245.

[20] Bobst, C. E., Thomas, J. J., Salinas, P. A., Savickas, P., and Kaltashov, I. A. (2010) Impact of oxidation on protein therapeutics: conformational dynamics of intact and oxidized acid-beta-glucocerebrosidase at near-physiological pH, *Protein science : a publication of the Protein Society 19*, 2366-2378.

[21] Houde, D., Peng, Y., Berkowitz, S. A., and Engen, J. R. (2010) Post-translational modifications differentially affect IgG1 conformation and receptor binding, *Molecular & cellular proteomics : MCP 9*, 1716-1728.

[22] von Bulow, R., Schmidt, B., Dierks, T., Schwabauer, N., Schilling, K., Weber, E., Uson, I., and von Figura, K. (2002) Defective oligomerization of arylsulfatase a as a cause of its instability in lysosomes and metachromatic leukodystrophy, *The Journal of biological chemistry 277*, 9455-9461.

[23] Nooren, I. M., and Thornton, J. M. (2003) Diversity of protein-protein interactions, *The EMBO journal 22*, 3486-3492.

[24] Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G., and Orengo, C. (2010) Transient Protein-Protein Interactions: Structural, Functional, and Network Properties, *Structure 18*, 1233-1243.

[25] Yu, Y. J., Zhang, Y., Kenrick, M., Hoyte, K., Luk, W., Lu, Y., Atwal, J., Elliott, J. M., Prabhu, S., Watts, R. J., and Dennis, M. S. (2011) Boosting Brain Uptake of a Therapeutic Antibody by Reducing Its Affinity for a Transcytosis Target, *Science Translational Medicine 3*, 84ra44-84ra44.

[26] Rabenstein, D. L. (2002) Heparin and heparan sulfate: structure and function, *Natural Product Reports 19*, 312-331.

[27] Jones, C. J., Beni, S., Limtiaco, J. F., Langeslay, D. J., and Larive, C. K. (2011) Heparin characterization: challenges and solutions, *Annual review of analytical chemistry (Palo Alto, Calif.) 4*, 439-465.

[28] Weitz, J. I. (1997) Low-Molecular-Weight Heparins, *New England Journal of Medicine 337*, 688-699.

[29] Fekete, S., Veuthey, J.-L., and Guillarme, D. (2012) New trends in reversed-phase liquid chromatographic separations of therapeutic peptides and proteins: Theory and applications, *Journal of Pharmaceutical and Biomedical Analysis 69*, 9-27.

[30] Fekete, S., Beck, A., Veuthey, J.-L., and Guillarme, D. (2014) Theory and practice of size exclusion chromatography for the analysis of protein aggregates, *Journal of Pharmaceutical and Biomedical Analysis 101*, 161-173.

[31] Fekete, S., Beck, A., Veuthey, J.-L., and Guillarme, D. (2015) Ion-exchange chromatography for the characterization of biopharmaceuticals, *Journal of Pharmaceutical and Biomedical Analysis 113*, 43-55.

[32] Kaltashov, I. A., Bobst, C. E., Abzalimov, R. R., Wang, G., Baykal, B., and Wang, S. (2012) Advances and challenges in analytical characterization of biotechnology products: mass spectrometry-based approaches to study properties and behavior of protein therapeutics, *Biotechnology Advances 30*, 210-222.

[33] Leurs, U., Mistarz, U. H., and Rand, K. D. (2015) Getting to the core of protein pharmaceuticals--Comprehensive structure analysis by mass spectrometry, *European journal of pharmaceutics and biopharmaceutics : official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik e.V 93*, 95-109.

[34] Heck, A. J. R. (2008) Native mass spectrometry: a bridge between interactomics and structural biology, *Nat Meth 5*, 927-933.

[35] Dyachenko, A., Wang, G., Belov, M., Makarov, A., de Jong, R. N., van den Bremer, E. T. J., Parren, P. W. H. I., and Heck, A. J. R. (2015) Tandem Native Mass-Spectrometry on Antibody–Drug Conjugates and Submillion Da Antibody–Antigen Protein Assemblies on an Orbitrap EMR Equipped with a High-Mass Quadrupole Mass Selector, *Analytical Chemistry 87*, 6095-6102.

[36] Konermann, L., and Douglas, D. J. (1997) Acid-Induced Unfolding of Cytochrome c at Different Methanol Concentrations: Electrospray Ionization Mass Spectrometry Specifically Monitors Changes in the Tertiary Structure, *Biochemistry 36*, 12296-12302.

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

[38] Bloem, E., van den Biggelaar, M., Wroblewska, A., Voorberg, J., Faber, J. H., Kjalke, M., Stennicke, H. R., Mertens, K., and Meijer, A. B. (2013) Factor VIII C1 domain spikes 2092-2093 and 2158-2159 comprise regions that modulate cofactor function and cellular uptake, *The Journal of biological chemistry 288*, 29670-29679.

[39] Shukla, A. K., Westfield, G. H., Xiao, K., Reis, R. I., Huang, L.-Y., Tripathi-Shukla, P., Qian, J., Li, S., Blanc, A., Oleskie, A. N., Dosey, A. M., Su, M., Liang, C.-R., Gu, L.-L., Shan, J.-M., Chen, X., Hanna, R., Choi, M., Yao, X. J., Klink, B. U., Kahsai, A. W., Sidhu, S. S., Koide, S., Penczek, P. A., Kossiakoff, A. A., Woods Jr, V. L., Kobilka, B. K., Skiniotis, G., and Lefkowitz, R. J. (2014) Visualization of arrestin recruitment by a G-protein-coupled receptor, *Nature 512*, 218-222.

[40] Houde, D., Berkowitz, S. A., and Engen, J. R. (2011) The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies, *J Pharm Sci 100*, 2071-2086.

[41] Schmidt, R. (2006) Dynamic Light Scattering for Protein Characterization, In *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Ltd.

[42] Sahin, E., and Roberts, C. J. (2012) Size-exclusion chromatography with multi-angle light scattering for elucidating protein aggregation mechanisms, *Methods in molecular biology (Clifton, N.J.) 899*, 403-423.

[43] Li, Y., Weiss, W. F. t., and Roberts, C. J. (2009) Characterization of high-molecular-weight nonnative aggregates and aggregation kinetics by size exclusion chromatography with inline multi-angle laser light scattering, *J Pharm Sci 98*, 3997-4016.

[44] Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism, *Biochimica et biophysica acta 1751*, 119-139.

[45] Liu, J., Andya, J. D., and Shire, S. J. (2006) A critical review of analytical ultracentrifugation and field flow fractionation methods for measuring protein aggregation, *The AAPS Journal 8*, E580-E589.

[46] Freyer, M. W., and Lewis, E. A. (2008) Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions, *Methods in cell biology 84*, 79-113.

[47] Thillaivinayagalingam, P., Gommeaux, J., McLoughlin, M., Collins, D., and Newcombe, A. R. (2010) Biopharmaceutical production: Applications of surface plasmon resonance biosensors, *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences 878*, 149-153.

[48] Cole, J. L., Lary, J. W., T, P. M., and Laue, T. M. (2008) Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium, *Methods in cell biology 84*, 143-179.

[49] Beeckmans, S. (1999) Chromatographic methods to study protein-protein interactions, *Methods (San Diego, Calif.) 19*, 278-305.

[50] Gao, J. Y., Dubin, P. L., and Muhoberac, B. B. (1997) Measurement of the Binding of Proteins to Polyelectrolytes by Frontal Analysis Continuous Capillary Electrophoresis, *Analytical Chemistry 69*, 2945-2951.

[51] Seyrek, E., Hattori, T., and Dubin, P. L. (2004) Frontal analysis continuous capillary electrophoresis for protein-polyelectrolyte binding studies, *Methods in molecular biology (Clifton, N.J.) 276*, 217-228.

[52] Collins, M. O., and Choudhary, J. S. (2008) Mapping multiprotein complexes by affinity purification and mass spectrometry, *Current Opinion in Biotechnology 19*, 324-330.

[53] Sjoelund, V., and Kaltashov, I. A. (2012) Modification of the Zonal Elution Method for Detection of Transient Protein–Protein Interactions Involving Ligand Exchange, *Analytical Chemistry 84*, 4608-4612.

[54] Truong, K., and Ikura, M. (2001) The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo, *Curr Opin Struct Biol 11*, 573-578.

[55] Michnick, S. W. (2003) Protein fragment complementation strategies for biochemical network mapping, *Curr Opin Biotechnol 14*, 610-617.

[56] Koo, E. H., Lansbury, P. T., Jr., and Kelly, J. W. (1999) Amyloid diseases: abnormal protein aggregation in neurodegeneration, *Proceedings of the National Academy of Sciences of the United States of America 96*, 9989-9990.

[57] Meredith, S. C. (2006) Protein Denaturation and Aggregation, *Annals of the New York Academy of Sciences 1066*, 181-221.

[58] Bellotti, V., and Chiti, F. (2008) Amyloidogenesis in its biological environment: challenging a fundamental issue in protein misfolding diseases, *Curr Opin Struct Biol 18*, 771-779.

[59] van Beers, M. M., and Bardor, M. (2012) Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins, *Biotechnology journal 7*, 1473-1484.

[60] Weiss, W. F. t., Young, T. M., and Roberts, C. J. (2009) Principles, approaches, and challenges for predicting protein aggregation rates and shelf life, *J Pharm Sci 98*, 1246-1277.

[61] Yearley, E. J., Godfrin, P. D., Perevozchikova, T., Zhang, H., Falus, P., Porcar, L., Nagao, M., Curtis, J. E., Gawande, P., Taing, R., Zarraga, I. E., Wagner, N. J., and Liu, Y. (2014) Observation of small cluster formation in concentrated monoclonal antibody solutions and its implications to solution viscosity, *Biophys J 106*, 1763-1770.

[62] Philo, J. S. (2009) A critical review of methods for size characterization of non-particulate protein aggregates, *Current pharmaceutical biotechnology 10*, 359-372.

[63] Ye, H. (2006) Simultaneous determination of protein aggregation, degradation, and absolute molecular weight by size exclusion chromatography-multiangle laser light scattering, *Analytical biochemistry 356*, 76-85.

[64] Tian, X., Langkilde, A. E., Thorolfsson, M., Rasmussen, H. B., and Vestergaard, B. (2014) Small-angle x-ray scattering screening complements conventional biophysical analysis: comparative structural and biophysical analysis of monoclonal antibodies IgG1, IgG2, and IgG4, *J Pharm Sci 103*, 1701-1710.

[65] Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrometry Reviews 16*, 1-23.

[66] Sharon, M., and Robinson, C. V. (2007) The role of mass spectrometry in structure elucidation of dynamic protein complexes, *Annual review of biochemistry 76*, 167-193.

[67] Wang, W., Kitova, E. N., and Klassen, J. S. (2005) Nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization. Factors influencing their formation and stability, *Anal Chem 77*, 3060-3071.

[68] Lin, H., Kitova, E. N., and Klassen, J. S. (2013) Quantifying protein-ligand interactions by direct electrospray ionization-MS analysis: evidence of nonuniform response factors induced by high molecular weight molecules and complexes, *Anal Chem 85*, 8919-8922.

[69] Kukrer, B., Filipe, V., van Duijn, E., Kasper, P. T., Vreeken, R. J., Heck, A. J., and Jiskoot, W. (2010) Mass spectrometric analysis of intact human monoclonal antibody aggregates fractionated by size-exclusion chromatography, *Pharmaceutical research 27*, 2197-2204.

[70] Wang, G., Johnson, A. J., and Kaltashov, I. A. (2012) Evaluation of Electrospray Ionization Mass Spectrometry as a Tool for Characterization of Small Soluble Protein Aggregates, *Analytical Chemistry 84*, 1718-1724.

[71] Wang, W., Nema, S., and Teagarden, D. (2010) Protein aggregation--pathways and influencing factors, *International journal of pharmaceutics 390*, 89-99.

[72] Gruendling, T., Guilhaus, M., and Barner-Kowollik, C. (2008) Quantitative LC-MS of polymers: determining accurate molecular weight distributions by combined size exclusion chromatography and electrospray mass spectrometry with maximum entropy data processing, *Anal Chem 80*, 6915-6927.

[73] Aaserud, D. J., Prokai, L., and Simonsick, W. J. (1999) Gel permeation chromatography coupled to fourier transform mass spectrometry for polymer characterization, *Anal Chem 71*, 4793-4799.

[74] Brady, L. J., Valliere-Douglass, J., Martinez, T., and Balland, A. (2008) Molecular mass analysis of antibodies by on-line SEC-MS, *Journal of the American Society for Mass Spectrometry 19*, 502-509.

[75] Lazar, A. C., Wang, L., Blattler, W. A., Amphlett, G., Lambert, J. M., and Zhang, W. (2005) Analysis of the composition of immunoconjugates using size-exclusion chromatography coupled to mass spectrometry, *Rapid communications in mass spectrometry : RCM 19*, 1806-1814.

[76] Sun, N., Soya, N., Kitova, E. N., and Klassen, J. S. (2010) Nonspecific interactions between proteins and charged biomolecules in electrospray ionization mass spectrometry, *Journal of the American Society for Mass Spectrometry 21*, 472-481.

[77] Li, H., Grigoryan, H., Funk, W. E., Lu, S. S., Rose, S., Williams, E. R., and Rappaport, S. M. (2011) Profiling Cys34 adducts of human serum albumin by fixed-step selected reaction monitoring, *Molecular & cellular proteomics : MCP 10*, M110.004606.

[78] Wally, J., Halbrooks, P. J., Vonrhein, C., Rould, M. A., Everse, S. J., Mason, A. B., and Buchanan, S. K. (2006) The crystal structure of iron-free human serum transferrin provides insight into inter-lobe communication and receptor binding, *The Journal of biological chemistry 281*, 24934-24944.

[79] Majorek, K. A., Porebski, P. J., Dayal, A., Zimmerman, M. D., Jablonska, K., Stewart, A. J., Chruszcz, M., and Minor, W. (2012) Structural and immunologic characterization of bovine, horse, and rabbit serum albumins, *Molecular immunology 52*, 174-182.

[80] Vagedes, P., Saenger, W., and Knapp, E.-W. (2002) Driving Forces of Protein Association: The Dimer-Octamer Equilibrium in Arylsulfatase A, *Biophysical Journal 83*, 3066-3078.

[81] Abzalimov, R. R., Bobst, C. E., Salinas, P. A., Savickas, P., Thomas, J. J., and Kaltashov, I. A. (2013) Studies of pH-dependent self-association of a recombinant form of arylsulfatase A with electrospray ionization mass spectrometry and size-exclusion chromatography, *Anal Chem 85*, 1591-1596.

[82] Mitragotri, S., Burke, P. A., and Langer, R. (2014) Overcoming the challenges in administering biopharmaceuticals: formulation and delivery strategies, *Nat Rev Drug Discov 13*, 655-672.

[83] Homola, J., Yee, S. S., and Gauglitz, G. (1999) Surface plasmon resonance sensors: review, *Sensors and Actuators B: Chemical 54*, 3-15.

[84] Cole, J. L., Lary, J. W., P. Moody, T., and Laue, T. M. (2008) Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium, In *Methods in cell biology*, pp 143-179, Academic Press.

[85] Weber, P. C., and Salemme, F. R. (2003) Applications of calorimetric methods to drug discovery and the study of protein interactions, *Current Opinion in Structural Biology 13*, 115-121.

[86] Piehler, J. (2005) New methodologies for measuring protein interactions in vivo and in vitro, *Current Opinion in Structural Biology 15*, 4-14.

[87] Muneeruddin, K., Thomas, J. J., Salinas, P. A., and Kaltashov, I. A. (2014) Characterization of Small Protein Aggregates and Oligomers Using Size Exclusion Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry, *Analytical Chemistry 86*, 10692-10699.

[88] Boeri Erba, E., Barylyuk, K., Yang, Y., and Zenobi, R. (2011) Quantifying Protein–Protein Interactions Within Noncovalent Complexes Using Electrospray Ionization Mass Spectrometry, *Analytical Chemistry 83*, 9251-9259.

[89] Leverence, R., Mason, A. B., and Kaltashov, I. A. (2010) Noncanonical interactions between serum transferrin and transferrin receptor evaluated with electrospray ionization mass spectrometry, *Proceedings of the National Academy of Sciences 107*, 8123-8128.

[90] Giannetti, A. M., Snow, P. M., Zak, O., and Björkman, P. J. (2003) Mechanism for Multiple Ligand Recognition by the Human Transferrin Receptor, *PLoS Biol 1*, e51.

[91] Socolofsky, S. A., and Jirka, G. H. (2005) Special Topics in Mixing and Transport Processes in the Environment. Engineering – Lectures, 5 ed., Coastal and Ocean Engineering Division. Texas A&M University.

[92] Carini, M., Regazzoni, L., and Aldini, G. (2011) Mass spectrometric strategies and their applications for molecular mass determination of recombinant therapeutic proteins, *Current pharmaceutical biotechnology 12*, 1548-1557.

[93] Bondarenko, P. V., Second, T. P., Zabrouskov, V., Makarov, A. A., and Zhang, Z. (2009) Mass measurement and top-down HPLC/MS analysis of intact monoclonal antibodies on a hybrid linear quadrupole ion trap-Orbitrap mass spectrometer, *Journal of the American Society for Mass Spectrometry 20*, 1415-1424.

[94] Ren, D., Pipes, G. D., Hambly, D., Bondarenko, P. V., Treuheit, M. J., and Gadgil, H. S. (2009) Top-down N-terminal sequencing of Immunoglobulin subunits with electrospray ionization time of flight mass spectrometry, *Analytical biochemistry 384*, 42-48.

[95] Houde, D., Kauppinen, P., Mhatre, R., and Lyubarskaya, Y. (2006) Determination of protein oxidation by mass spectrometry and method transfer to quality control, *Journal of chromatography. A 1123*, 189-198.

[96] Reid, C. Q., Tait, A., Baldascini, H., Mohindra, A., Racher, A., Bilsborough, S., Smales, C. M., and Hoare, M. (2010) Rapid whole monoclonal antibody analysis by mass spectrometry: An ultra scale-down study of the effect of harvesting by centrifugation on the post-translational modification profile, *Biotechnology and bioengineering 107*, 85-95.

[97] Wang, S., and Kaltashov, I. A. (2013) An 18O-labeling assisted LC/MS method for assignment of aspartyl/isoaspartyl products from Asn deamidation and Asp isomerization in proteins, *Anal Chem 85*, 6446-6452.

[98] Bobst, C. E., Abzalimov, R. R., Houde, D., Kloczewiak, M., Mhatre, R., Berkowitz, S. A., and Kaltashov, I. A. (2008) Detection and characterization of altered conformations of protein pharmaceuticals using complementary mass spectrometry-based approaches, *Anal Chem 80*, 7473-7481.

[99] Wang, G., Abzalimov, R. R., and Kaltashov, I. A. (2011) Direct monitoring of heat-stressed biopolymers with temperature-controlled electrospray ionization mass spectrometry, *Anal Chem 83*, 2870-2876.

[100] Iacob, R. E., Bou-Assaf, G. M., Makowski, L., Engen, J. R., Berkowitz, S. A., and Houde, D. (2013) Investigating Monoclonal Antibody Aggregation Using a Combination of H/DX-MS and Other Biophysical Measurements, *Journal of Pharmaceutical Sciences 102*, 4315-4329.

[101] Zhang, H., Cui, W., and Gross, M. L. (2014) Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies, *FEBS letters 588*, 308-317.

[102] Sandra, K., Vandenheede, I., and Sandra, P. (2014) Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization, *Journal of chromatography. A 1335*, 81-103.

[103] An, B., Zhang, M., and Qu, J. (2014) Toward sensitive and accurate analysis of antibody biotherapeutics by liquid chromatography coupled with mass spectrometry, *Drug metabolism and disposition: the biological fate of chemicals 42*, 1858-1866.

[104] van den Broek, I., Niessen, W. M., and van Dongen, W. D. (2013) Bioanalytical LC-MS/MS of protein-based biopharmaceuticals, *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences 929*, 161-179.

[105] Lundblad, R. L. (2010) *Approaches to the Conformational Analysis of Biopharmaceuticals*, CRC Press/Taylor & Francis.

[106] Razinkov, V. I., Treuheit, M. J., and Becker, G. W. (2013) Methods of high throughput biophysical characterization in biopharmaceutical development, *Current drug discovery technologies 10*, 59-70.

[107] Fekete, S., Beck, A., Veuthey, J. L., and Guillarme, D. (2014) Theory and practice of size exclusion chromatography for the analysis of protein aggregates, *J Pharm Biomed Anal 101*, 161-173.

[108] Goldmacher, V. S., Amphlett, G., Wang, L., and Lazar, A. C. (2015) Statistics of the distribution of the abundance of molecules with various drug loads in maytansinoid antibody-drug conjugates, *Molecular pharmaceutics 12*, 1738-1744.

[109] Bertoletti, L., Regazzoni, L., Aldini, G., Colombo, R., Abballe, F., Caccialanza, G., and De Lorenzi, E. (2013) Separation and characterisation of beta2-microglobulin folding conformers by ion-exchange liquid chromatography and ion-exchange liquid chromatography-mass spectrometry, *Analytica chimica acta 771*, 108-114.

[110] Kaltashov, I. A., and Abzalimov, R. R. (2008) Do ionic charges in ESI MS provide useful information on macromolecular structure?, *Journal of the American Society for Mass Spectrometry 19*, 1239-1246.

[111] Nguyen, S. N., Bobst, C. E., and Kaltashov, I. A. (2013) Mass spectrometry-guided optimization and characterization of a biologically active transferrin-lysozyme model drug conjugate, *Molecular pharmaceutics 10*, 1998-2007.

[112] Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways, *Nature 358*, 302-307.

[113] Radford, S. E., Woolfson, D. N., Martin, S. R., Lowe, G., and Dobson, C. M. (1991) A three-disulphide derivative of hen lysozyme. Structure, dynamics and stability, *The Biochemical journal 273(Pt 1)*, 211-217.

[114] Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions, *Mass Spectrom Rev 25*, 663-682.

[115] Cech, N. B., and Enke, C. G. (2001) Practical implications of some recent studies in electrospray ionization fundamentals, *Mass Spectrom Rev 20*, 362-387.

[116] Kuprowski, M. C., and Konermann, L. (2007) Signal response of coexisting protein conformers in electrospray mass spectrometry, *Anal Chem 79*, 2499-2506.

[117] Boja, E. S., and Fales, H. M. (2001) Overalkylation of a protein digest with iodoacetamide, *Anal Chem 73*, 3576-3582.

[118] Eyles, S. J., Radford, S. E., Robinson, C. V., and Dobson, C. M. (1994) Kinetic consequences of the removal of a disulfide bridge on the folding of hen lysozyme, *Biochemistry 33*, 13038-13048.

[119] Bermel, R. A., and Rudick, R. A. (2007) Interferon-beta treatment for multiple sclerosis, *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics 4*, 633-646.

[120] Lin, L., Brown, F., Lubiniecki, A., Murano, G., and Eds, K. (1995) In *Symposium on the Characterization of Biotechnology Pharmaceutical Products*, pp 97-104, Washington, D.C.

[121] Runkel, L., Meier, W., Pepinsky, R. B., Karpusas, M., Whitty, A., Kimball, K., Brickelmaier, M., Muldowney, C., Jones, W., and Goelz, S. E. (1998) Structural and functional differences between glycosylated and non-glycosylated forms of human interferon-beta (IFN-beta), *Pharmaceutical research 15*, 641-649.

[122] Dissing-Olesen, L., Thaysen-Andersen, M., Meldgaard, M., Hojrup, P., and Finsen, B. (2008) The function of the human interferon-beta 1a glycan determined in vivo, *The Journal of pharmacology and experimental therapeutics 326*, 338-347.

[123] Mastrangeli, R., Rossi, M., Mascia, M., Palinsky, W., Datola, A., Terlizzese, M., and Bierau, H. (2015) In vitro biological characterization of IFN-beta-1a major glycoforms, *Glycobiology 25*, 21-29.

[124] Sburlati, A. R., Umana, P., Prati, E. G., and Bailey, J. E. (1998) Synthesis of bisected glycoforms of recombinant IFN-beta by overexpression of beta-1,4-N-acetylglucosaminyltransferase III in Chinese hamster ovary cells, *Biotechnology progress 14*, 189-192.

[125] Wakankar, A. A., and Borchardt, R. T. (2006) Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization, *J Pharm Sci 95*, 2321-2336.

[126] Robinson, N. E., and Robinson, A. B. (2004) *Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins*, Althouse Press: Cave Junction, OR.

[127] Zubarev, R. A. (2004) Electron-capture dissociation tandem mass spectrometry, *Curr Opin Biotechnol 15*, 12-16.

[128] Paizs, B., and Suhai, S. (2005) Fragmentation pathways of protonated peptides, *Mass Spectrom Rev 24*, 508-548.

[129] Yang, H., and Zubarev, R. A. (2010) Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides, *Electrophoresis 31*, 1764-1772.

[130] Li, X., Yu, X., Costello, C. E., Lin, C., and O'Connor, P. B. (2012) Top-down study of beta2-microglobulin deamidation, *Anal Chem 84*, 6150-6157.

[131] Bailon, P., and Won, C.-Y. (2009) PEG-modified biopharmaceuticals, *Expert Opin Drug Deliv. 6*, 1-16.

[132] Harris, J. M., and Chess, R. B. (2003) Effect of pegylation on pharmaceuticals, *Nat Rev Drug Discov 2*, 214-221.

[133] Vestling, M. M., Murphy, C. M., Keller, D. A., Fenselau, C., Dedinas, J., Ladd, D. L., and Olsen, M. A. (1993) A strategy for characterization of polyethylene glycol-derivatized proteins. A mass spectrometric analysis of the attachment sites in polyethylene glycol-derivatized superoxide dismutase, *Drug Metabolism and Disposition 21*, 911-917.

[134] Veronese, F. M., Saccà, B., Polverino de Laureto, P., Sergi, M., Caliceti, P., Schiavon, O., and Orsolini, P. (2001) New PEGs for Peptide and Protein Modification, Suitable for Identification of the PEGylation Site, *Bioconjugate Chemistry 12*, 62-70.

[135] Abzalimov, R. R., Frimpong, A., and Kaltashov, I. A. (2012) Structural characterization of protein–polymer conjugates. I. Assessing heterogeneity of a small PEGylated protein and mapping conjugation sites using ion exchange chromatography and top-down tandem mass spectrometry, *International Journal of Mass Spectrometry 312*, 135-143.

[136] Gilmore, J. M., Scheck, R. A., Esser-Kahn, A. P., Joshi, N. S., and Francis, M. B. (2006) N-Terminal Protein Modification through a Biomimetic Transamination Reaction, *Angewandte Chemie International Edition 45*, 5307-5311.

[137] Abzalimov, R. R., and Kaltashov, I. A. (2010) Electrospray Ionization Mass Spectrometry of Highly Heterogeneous Protein Systems: Protein Ion Charge State Assignment via Incomplete Charge Reduction, *Analytical Chemistry 82*, 7523-7526.

[138] Selby, T. L., Wesdemiotis, C., and Lattimer, R. P. (1994) Dissociation characteristics of $[M+X]^+$ ions (X = H, Li, K) from linear and cyclic polyglycols, *J. Am. Soc. Mass Spectrom. 5*, 1081-1092.