2015

Transferrin-Based Therapeutics and Analytical Methods to Characterize Them

Son N. Nguyen

University of Massachusetts - Amherst, snnguyen@chem.umass.edu

Follow this and additional works at: http://scholarworks.umass.edu/dissertations_2

Part of the Analytical Chemistry Commons, and the Biotechnology Commons

Recommended Citation


http://scholarworks.umass.edu/dissertations_2/386

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations May 2014 - current by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
TRANSFERRIN-BASED CONJUGATES AND ANALYTICAL METHODS TO CHARACTERIZE THEM

A Dissertation Presented

by

SON N. NGUYEN

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 2015

Department of Chemistry
TRANSFERRIN-BASED THERAPEUTICS AND ANALYTICAL METHODS TO CHARACTERIZE THEM

A Dissertation Presented
by
SON N. NGUYEN

Approved as to style and content by:

________________________________________________________________________
Igor A. Kaltashov, Committee Chair

________________________________________________________________________
Richard W. Vachet, Committee Member

________________________________________________________________________
Anne B. Mason, Committee Member

________________________________________________________________________
Stephen J. Eyles, Committee Member

________________________________________________________________________
Craig Martin, Head
Department of Chemistry
DEDICATION

To my wife Ly who has been there for me with love and encouragement

To my children Hải Nam and Thu Giang

To Mom and Dad
ACKNOWLEDGEMENTS

I would like first of all to thank my advisor, Professor Igor Kaltashov, for his advice, guidance, and support from day one of my graduate studies. I am grateful for his knowledge instruction and critical questioning of my works. It has been a wonderful experience working and learning with him.

I would also like to express my gratitude to the members of my dissertation committee; Professor Richard Vachet, for his incisive comments and discussions on my work; Professor Anne Mason, who generously agreed to come all the way from University of Vermont and her unlimited source of transferrin receptor for my study; and Dr. Stephan Eyles, for his helpful suggestions on all stages of my work and for helping me learn multiple mass spectrometers.

Special thanks also go to Dr. Cedric Bobst and Dr. Rinat Abzalimov for useful discussions. Their knowledge and humor made the lab a fun place to work in. I also would like to thank my friends, co-authors and lab members, Shunhai, Hanwei, Adriana, Burcu, Guanbo, Yao, Jinyi, Muneer, Jake, Lola, Shengsheng, Chengfeng, Yunlong and Honglin. They all were a great source of support. I also want to thank Gokhan from Vachet’s group for ICP-MS measurements and Dr. Sung Tae Kim from Rotello’s group for the animal works. You guys are simply the best.

This project would have been much more difficult had it not been helped by my understudies, Alex, Thanh Bùi, Stephanie and Matts. Thank you very much for your wonderful works.
My five and a half year study here at UMass, Amherst has been filled with memories. My Vietnamese friends in Amherst have really been a second family for me, A Lộc, Chí Hoa, A Long và Việt Anh, Thanh, A Hùng và chị Thanh, Chị Hướng và cháu Quỳnh Hướng, Vân và Vân, Minh, Quân, Quản và Thu, A Hoàng và chị Hải Anh, Thụy, Hằng, Ngọc Lê, Ngọc Thái, Hòa, Huy, Bách, A Duy... I will also miss my North Village Sunday soccer buddies, Luis, Wawan, Pibe, Rafael, Roberto, Saurav, Ahmed, Hesham, Gonzalo, Ricardo, Javier, Qian, Tsukasa... Thank you all for the good games.

My Acknowledgments would be incomplete without my closest friends in Vietnam on the list, Dũng, Hùng Trần, Hùng Mạc, Soàn, Môn. I am also grateful to my former professors and colleagues, Thầy Luận, Thầy Ri, Thầy Trung, Thầy Thu, Thầy Xuân Anh, Thầy Huyền, Thầy Nghi, Thầy Tín, Thầy Bối, A Sơn, Chí Thảo, Chí Giang, Chí Mai, A Tuấn, Cô Mão, A Thanh, Huệ, Đức. I would not be here without your guidance and help.

I am grateful to my extended family, dad, mom, my mother in-law, my big brother, my sister in law, my little brother in-law, my cute little nephew, they have always loved and supported me. And last but not least, special thanks go to my beloved wife and my children. It is your love, support, encouragement and sacrifice that enables me to stay focused on my study. This is dedicated to you.
ABSTRACT

TRANSFERRIN-BASED THERAPEUTICS AND ANALYTICAL METHODS TO CHARACTERIZE THEM

MAY 2015

SON N. NGUYEN, B.A., VIETNAM NATIONAL UNIVERSITY, HANOI

M.A., VIETNAM NATIONAL UNIVERSITY, HANOI

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Igor A. Kaltashov

Modern development of medicine requires detailed characterization by state-of-the-art analytical techniques that can be used to analyze covalent structure, conformations and protein-receptor interaction to quantitatively measure biodistribution of protein therapeutics. Mass spectrometry has already become an indispensable tool facilitating all stages of protein drug development. Particularly, this work has demonstrated the tremendous potential of electrospray ionization (ESI) mass spectrometry (MS) in this arena by providing invaluable information beyond mass measurement that can be used to optimize protein drug conjugate structures during early stages of development, and to further catalyze drug design efforts. Additionally, a new sensitive and selective method that uses metal tracers and inductively coupled plasma (ICP) MS developed in our lab has been successfully applied for quantitating exogenous transferrin (Tf) and Tf-based drugs in biological tissues and fluids. Furthermore, ICP-MS based method using metal tracer in combination with size exclusion chromatography (SEC) method proved to be able to probe into protein stability post-injection and to yield useful data not
accessible by other methods. For the first time a small soluble protein aggregation of injected protein drug was studied in live animals. Finally, a simple and cost-effective $^{18}$O labeling-based method has been developed for quantitating lysine modification sites of protein drug conjugates and has been successfully applied for N-succinimidyl-S-acetyltioacetate (SATA)-Lysozyme (Lz) conjugate.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Biopharmaceuticals</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1. General</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2. Second generation (engineered) biopharmaceuticals</td>
<td>2</td>
</tr>
<tr>
<td>1.1.3. Transferrin and its potential as a drug carrier</td>
<td>4</td>
</tr>
<tr>
<td>1.2. Experimental methods to characterize biologics</td>
<td>5</td>
</tr>
<tr>
<td>1.2.1. Separation-based methods</td>
<td>5</td>
</tr>
<tr>
<td>1.2.2. Mass spectrometry</td>
<td>7</td>
</tr>
<tr>
<td>1.2.3. Experimental methods for pharmacokinetics (PK) studies</td>
<td>9</td>
</tr>
<tr>
<td>1.3. Objectives</td>
<td>11</td>
</tr>
<tr>
<td>2. MASS SPECTROMETRY-GUIDED OPTIMIZATION AND CHARACTERIZATION OF A</td>
<td>15</td>
</tr>
<tr>
<td>BIOLOGICALLY ACTIVE TRANSFERRIN-LYSOZYME MODEL DRUG CONJUGATE</td>
<td></td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>17</td>
</tr>
<tr>
<td>2.2. Experimental</td>
<td>19</td>
</tr>
<tr>
<td>2.3. Results</td>
<td>21</td>
</tr>
<tr>
<td>2.3.1. Production, purification and characterization of the Lz-Tf</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.3.2. Influence of conjugation and chemical modifications on interaction with transferrin receptor (TfR) ................................................................. 26

2.3.3. Influence of conjugation and chemical modification on enzymatic activity ........................................................................................................... 28

2.4. Discussion ........................................................................................................ 30

2.5. Conclusions ..................................................................................................... 34

3. BIODISTRIBUTION STUDIES OF TRANSFERRIN-BASED DRUG IN ANIMAL MODELS BY INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY (ICP-MS) .................................................................................. 47

3.1. Introduction ...................................................................................................... 48

3.2. Experimental .................................................................................................... 50

3.2.1. Preparation of indium (In) loaded Lz-Tf conjugate ..................................... 50

3.2.2. Animal Models ............................................................................................ 50

3.2.3. ICP-MS. ..................................................................................................... 51

3.2.4. Imaging LA-ICP-MS ................................................................................ 52

3.3. Results and discussion .................................................................................... 52

3.3.1. In Vivo Distribution of Tf and Lz-Tf ............................................................ 52

3.3.2. Distribution of Lz-Tf at tissue level by imaging LA-ICP-MS. .................... 54

3.4. Conclusions ..................................................................................................... 56

4. STUDY OF IN-VIVO AGGREGATION OF TRANSFERRIN-BASED DRUG BY SIZE EXCLUSION CHROMATOGRAPHY AND INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY ..... 63

4.1. Introduction ...................................................................................................... 63

4.2. Materials and methods .................................................................................. 65

4.2.1. Preparation of indium (In) loaded Lz-Tf conjugate .................................. 65

4.2.2. Animal Models .......................................................................................... 65

4.2.3. Instrumentation. ....................................................................................... 66

4.3. Results and discussion .................................................................................. 67
4.3.1. Aggregation study of In$_2$Tf-Lz under typical storage conditions .......... 67

4.3.2. In-vitro, Ex-vivo and In Vivo aggregation study of In$_2$Tf-Lz ............... 69

4.4. Conclusions .................................................................................................. 72

5. A NEW $^{18}$O LABELING METHOD FOR MODIFICATION SITES QUANTITATION OF PROTEIN-DRUGS CONJUGATES .................................................................................................. 78

5.1. Introduction .................................................................................................. 78

5.2. Materials and methods ............................................................................ 81

5.2.1. Chemicals and Reagents ........................................................................ 81

5.2.2. Preparation of SATA-Lysozyme conjugate ......................................... 82

5.2.3. Protein Digestion and Post-digestion Labeling .................................... 82

5.2.4. Instrumentation .................................................................................... 83

5.3. Results and discussion ............................................................................ 83

5.3.1. $^{18}$O method for quantitation of modification sites ............................... 83

5.3.2. Comparison of $^{18}$O method and XIC estimation method for modification sites quantitation ................................................................. 87

5.4. Conclusions .................................................................................................. 89

6. SUMMARY AND FUTURE OUTLOOK ............................................................ 99

BIBLIOGRAPHY ................................................................................................. 101
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.</td>
<td>Comparison of small molecule and biologic-based new molecular entities (NMEs) (1982–2013)</td>
<td>13</td>
</tr>
<tr>
<td>1.2.</td>
<td>Key parameters for protein conjugates</td>
<td>13</td>
</tr>
<tr>
<td>2.1.</td>
<td>Bacteriolytic activity of Lz-Tf and related proteins</td>
<td>37</td>
</tr>
<tr>
<td>3.1.</td>
<td>Operating parameters of LA-ICP-MS</td>
<td>58</td>
</tr>
<tr>
<td>4.1.</td>
<td>Operating parameters of ICP-MS</td>
<td>73</td>
</tr>
<tr>
<td>5.1.</td>
<td>Predicted Lysine Surface Accessibilities and pKa Values</td>
<td>91</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1.</td>
<td>Chapter 2 table of content graphic</td>
<td>15</td>
</tr>
<tr>
<td>S1.</td>
<td>ESI mass spectra of intact Tf and Tf modified with the Traut’s reagent under denaturing conditions</td>
<td>38</td>
</tr>
<tr>
<td>S2.</td>
<td>ESI mass spectrum of the mixture of Lz-Tf conjugate (produced with SM(PEG)12 as a linker) and TfR acquired under near-native conditions</td>
<td>39</td>
</tr>
<tr>
<td>2.2.</td>
<td>Schematic diagrams illustrating conjugation of Tf to Lz using the Traut’s reagent and SMCC</td>
<td>40</td>
</tr>
<tr>
<td>2.3.</td>
<td>ESI mass spectra of activated Tf, charge state +32</td>
<td>41</td>
</tr>
<tr>
<td>2.4.</td>
<td>ESI mass spectra of crude (A) and IXC-purified (B) conjugation products of Lz and Tf</td>
<td>42</td>
</tr>
<tr>
<td>2.5.</td>
<td>ESI mass spectra of the purified 1:1 Lz-Tf conjugate</td>
<td>43</td>
</tr>
<tr>
<td>2.6.</td>
<td>ESI mass spectra of Tf/TfR (A) and Lz-Tf/TfR (B) mixtures acquired under near-native conditions (3 µM of each protein in 20mM ammonium acetate, pH 7.1)</td>
<td>44</td>
</tr>
<tr>
<td>3.1.</td>
<td>Chapter 3 TOC graphic</td>
<td>47</td>
</tr>
<tr>
<td>3.2.</td>
<td>Schematic of the ICP-MS based method for tracking Tf-based therapeutics distribution in organs and fluids</td>
<td>59</td>
</tr>
<tr>
<td>3.3.</td>
<td>Organ distribution of 115In over a 24 hours period after intravenous injection</td>
<td>60</td>
</tr>
<tr>
<td>3.4.</td>
<td>LA-ICP-MS images for the simultaneous monitoring of 57Fe, 63Cu, and 115In on spleen sections from rats treated with a single dose of In2Lz-Tf</td>
<td>61</td>
</tr>
<tr>
<td>3.5.</td>
<td>LA-ICP-MS images for the simultaneous monitoring of 57Fe, 63Cu, and 115In on brain sections from rats treated with a single dose of In2Lz-Tf</td>
<td>62</td>
</tr>
<tr>
<td>4.1.</td>
<td>Experimental scheme of the Tf-Lz conjugate stability test</td>
<td>74</td>
</tr>
<tr>
<td>4.2.</td>
<td>SEC chromatograms and native ESI-MS spectra of Tf-Lz under storage conditions</td>
<td>75</td>
</tr>
<tr>
<td>4.3.</td>
<td>SEC chromatograms of In2Tf-Lz in different conditions</td>
<td>76</td>
</tr>
<tr>
<td>4.4.</td>
<td>In contents of SEC collected fractions measured by ICP-MS. Data were reported as fraction (percentage) of total In concentration</td>
<td>77</td>
</tr>
</tbody>
</table>
5.1. ESI-MS of the SATA-Lz conjugate (+10 charge state) SATA:Lz (4:1) molar ratio. ...... 92
5.2. Tryptic peptides that are monitored in order to quantitate lysine modification sites ........................................................................................................................................................................ 92
5.3. \(^{18}\)O labeled internal standard strategy used to quantitate modification sites ... 93
5.4. Calibration plots of monitored peptides. .................................................................................................................................................................................. 94
5.5. Normalized \(^{16}\)O/\(^{18}\)O ratios of all peptides against the ratio of CT3 ....................... 95
5.6. A calibration curve of reference peptide CT3 was plotted using the measured \(^{16}\)O/\(^{18}\)O ratios versus the corresponding mixing ratios. ................................................................................................................................. 95
5.7. Modification percentages at all six lysine sites along with measured control peptide percentages (gray bars) for method validation purposes .................................................................................................................. 96
5.8. Typical problems of XIC method when used for quantitating modification sites...... 97
5.9. Results from \(^{18}\)O method and XIC method ......................................................................................................................... 98
CHAPTER 1
INTRODUCTION

1.1. Biopharmaceuticals

1.1.1. General

Biopharmaceuticals, also known as biologics, which are distinct from chemically synthesized pharmaceutical products, generally refer to medicinal products manufactured in or extracted from biological sources.\(^1\) Vaccines, blood and blood products, allergenic extracts, human cells and tissues, gene therapies, cellular therapies and recombinant therapeutic protein are among examples of biopharmaceuticals. Thanks to their excellent features such as high specificity and activity at relatively low concentrations, biopharmaceuticals have become the fastest growing segment among the new pharmaceutical products.\(^2\)

Nowadays, biopharmaceuticals extracted from natural sources continue to be used for diverse diseases. Modern drug manufacturing processes, however, have limited this need thanks to the improvement in production and safety from recombinant DNA technology.\(^3\) There are 91 recombinant-protein-based new molecular entities (NMEs) have been approved by the FDA as therapeutics since the first biologic was approved (Table 1.1).\(^4\) Three major categories of biologics include: monoclonal antibodies, enzymes replacement/modulators and cell surface receptor function replacement/modulators.

**Receptor modulator.** These are substances that are identical to human key signaling proteins. Examples of these are the biosynthetic human insulin and its analogues and
growth hormone. Initially, insulin was extracted from animals’ pancreas, however, due to medical need and safety, recombinant insulin proved to be a very good alternative. In 1982, Eli Lilly’s recombinant human insulin produced in E. coli was the first biologic approved. This approval helped open the door for the production of biologics by this technology.

**Enzyme modulator.** Enzyme modulators are used to replace enzymes which are deficient or absent in patients. This medical treatment is currently available for lysosomal diseases such as Gaucher disease, Fabry disease, etc. It is important to note that, although enzyme replacement therapy can greatly modify or reduce the symptoms and severity of the condition as well as disease progression, it does not cure these disorders.

**Monoclonal antibodies.** “Custom-designed" antibodies are similar to those in the human immune system that can target any specific cell type or counteract or block any given substance in the body. The first approved monoclonal antibody, muromonab CD3, was a murine protein used in acute organ transplant rejection. However, minimizing immunogenicity was not possible until the approval of the second monoclonal antibody, abciximab (1993), in which complementarily-determined regions were replaced by the human version. Currently, technologies enhancing drug safety and efficacy continue to advance alongside the fully human antibodies.

1.1.2. **Second generation (engineered) biopharmaceuticals**
Unlike the first generation, often refers as “replacement proteins”, the second generation biologics are engineered in ways to control their activity. These include covalent surface modification with polyethylene glycol (PEG), alteration of glycosylation and amino acid sequence, or protein-protein conjugation.

**PEGylation and fusion protein.** Small biologics that are less than ~60kDa are likely to be cleared from the body via glomerular filtration. In order to extend the circulation half-life of protein-drugs, their hydrodynamic size can be increased by covalently modifying the surface with hydrophilic polymers such as polyethylene glycol, sialic acid, polysaccharides hyaluronic acid and hydroxyl ethyl starch.

Other approaches exploiting the FcRn recycling process by genetically fusing protein drugs with the Fc region of IgG have also been used to prolong drug circulation half-life. The blockbuster etanercept (Enbrel; Amgen) is one of these.

**Protein-drug conjugation.** Protein-drug conjugates are very promising approach for cancer treatment. They combine the carrier proteins’ specificity with the potency of cytotoxic drugs, thus, having the best features of both components. Carrier proteins, such as monoclonal antibodies (mAbs) or other transport proteins such as transferrin (Tf), can be linked to cell-killing drugs via various chemical linkers. Following the guidance, binding of carriers to target cancer cells and the internalization of the protein conjugates, the toxic drugs can be chemically and/or enzymatically released and activated to kill the cells. The success of the targeted delivery approach depends upon: (1) the characteristics of the carrier, (2) the potency of the drug, and (3) the method of
linkage of the carrier to the drug. Table 1.2 shows key parameters for protein conjugates. 17-31

**Conjugation chemistry.** Due to the long exposure of protein-drug conjugates, normally several days post-injection, it is very important that the drug-linker is stable. It is reported that protein-drugs linked via disulfide and hydrazone linkers are not as stable in circulation compare to those using protease-cleavable linkers. 22-24 The three most common methods for protein-drugs conjugating include: lysines and cysteines conjugation (either from reduced interchain disulfides or by addition of genetically engineered cysteines). Due to the abundance of lysine residues (up to 100 on IgG1 mAbs), cysteine-based conjugates are more uniform than that of lysine-based conjugates 29. However, to some extent, the number of drug molecules attached is more important than the location. 29, 31 For example, it was reported that a maximum of four drug molecules, depending on their size, to one carrier protein usually circulated longer and was found to be superior to those more heavily loaded. 28

1.1.3. Transferrin and its potential as a drug carrier

Transferrin (Tf) is a promising drug carrier that has the potential to deliver metals, small molecule medicines and therapeutic proteins to cancer cells 32 and/or across physiological barriers (such as the blood-brain barrier, BBB 33). Despite this promise, very few Tf-based therapeutics have been developed and/or reached clinical trials.

Most of the earlier explorations of Tf as a potential delivery vehicle focused on intracellular drug delivery. Various therapeutic agents such as metal ions,
chemotherapeutics, proteins drugs, genes and nanoparticles have been coupled to Tf for tumor specific targeting.\textsuperscript{34-39} For example, \textsuperscript{67}Ga\textsuperscript{3+} and \textsuperscript{111}In\textsuperscript{3+} were found to be able to replace Fe\textsuperscript{3+} ions in Tf molecules and have been used as a diagnostic technique for many malignancies.\textsuperscript{40} In another example, artemisinin tagged holo-Tf was used as a cancer targeting therapeutics, thanks to the free radicals released when it reacts with iron.\textsuperscript{41} Notably, a diphtheria toxin (DT) point mutant, CRM107, has been conjugated to Tf and tested for the treatment of malignant brain tumors. This is the only Tf-based conjugate that has reached clinical trials before being discontinued at phase III.\textsuperscript{42}

As one of few proteins that have the ability to cross physiological barriers, Tf has also been exploited to target difficult-to-reach intracellular targets.\textsuperscript{43-47} Recently, an animal model study has showed the successful BBB-crossing of a Tf-human serum albumin nanoparticles conjugate. Additionally, in our group, we have explored the properties of Lysozyme (Lz) conjugated to Tf as a model therapeutic that targets the CNS and develop analytical protocols to characterize its structure and interactions with therapeutic targets and physiological partners critical for its successful delivery.\textsuperscript{48}

1.2. Experimental methods to characterize biologics

1.2.1. Separation-based methods

Size exclusion (gel filtration) chromatography (SEC) separates biomolecules purely based on their hydrodynamic size differences rather than interaction with stationary phase. Analytes of different size diffuse differently through a SEC column containing pore size controlled particles. Due to the ability to diffuse freely into the pores, small
molecules tend to elute from the column later than the larger one. Thus, analytes are separated in order of decreasing hydrodynamic sizes which, in most cases, are proportional to their molecular weights. SEC is a low resolution separation technique. It is, therefore, normally applied for fractionation (purification or characterization) or group separation (desalting or removing low-molecular-weight contaminants).49  

**Ion-exchange chromatography (IXC)** separates biomolecules on the basis of their charge differences. IEX have the ability to resolve identical proteins that differ by only a single charged group.50 Typical charged groups on biomolecules that form ionic bonds with the stationary phase ligands are \( \text{–NH}^{3+}, =\text{NH}^{2+}, \geq\text{NH}^+, \text{–COO}^-, \text{PO}_4^-, \text{SO}_3^{2–} \). After removing the non-bound species by washing, bound biomolecules are eluted with either a higher ionic strength or altered pH buffer. It is a high resolving power technique with fast separations, high recovery and most importantly it can retain biological active form of proteins even though the samples may contain high salt concentrations (>1 M).51  

**Reversed-phase liquid chromatography (RP-HPLC)** separates different analytes from the mobile phase on the basis of their hydrophobic binding to the ligands attached to the stationary phase. Analytes will be eluted in order of increasing molecular hydrophobicity using either isocratic or gradient elution. For the analysis of peptides and proteins, RP-HPLC is a very powerful technique due to its excellent resolution even for very closely related molecules along with high recoveries and excellent reproducibility.50 Unfortunately, biologically active form of proteins may not be recovered with this technique due to the use of high organic content solvent causing irreversible denaturation.49
Field-flow fractionation (FFF) is a family of chromatography-like elution techniques first described 1966\textsuperscript{52} which separates molecules in a single liquid phase with the help of an external field applied orthogonally to the flow direction.\textsuperscript{53} It can be applied for a wide colloidal size range of analytes from proteins, supramolecular assemblies to colloids and particles.\textsuperscript{54} A broad range of external fields were used including cross-flow stream (FIFFF), temperature gradient (ThFFF), electrical potential (ElFFF), centrifugal force (SdFFF), gravitational force (GrFFF), dielectrophoretic (DiFFF), standing acoustic wave (AcFFF) and magnetic fields (MgFFF).\textsuperscript{55} Of these techniques, flow FFF (FIFFF), has been the most popular FFF method applied for protein separation and analysis due to its ability to separate intact proteins within $\sim 10^5$ mass range in a single run.\textsuperscript{56}

Capillary electrophoresis (CE) is the most efficient family of separation techniques that is available for the analysis of both large and small molecules.\textsuperscript{57} It separates analytes based on the differences in electrophoretic mobility which relates to their charge, size and viscosity when applying a high voltage electric field over a sample in buffer solution held in a silica capillary.\textsuperscript{58} Of all techniques in its family, capillary zone electrophoresis (CZE) that moves analytes via electroosmotic flow (EOF) and separates them based on mass-to-charge ratio is the most commonly used. In most cases, changing the pH will manipulate the separation thanks to the pH dependence of analytes' charge.\textsuperscript{59} For CZE separation of basic proteins, however, the fused silica's wall need to be modified so that silanol groups are no longer exposed, thus, preventing the absorption of the proteins being analyzed.\textsuperscript{60}

1.2.2. Mass spectrometry
Mass spectrometry (MS) has become an indispensable tool for characterizing biological macromolecules. It provides sensitive and accurate measurements for a wide variety of molecules and complex mixtures of biomacromolecules by the separation and characterization based on ions mass-to-charge ratio (m/z). Mass spectrometry (MS) not only allow characterizing the covalent structure of biomolecules but also at a wide range of levels, including conformation, dynamics and interaction with physiological partners.61,62

**Primary structure.** The MS methods for identifying amino acid sequences and post-translation modifications (PTMs) of proteins have been extensively reported.63, 64 Although those MS methods depend on PTMs and the availability of MS instrumentation, in most cases, either intact proteins (“top-down”) or enzymatic digestion approach (“bottom-up” or “peptide mapping”) was involved. Due to the presence of salt and/or others MS incompatible matrix components, HPLC is commonly used with MS for the analysis. Moreover, in order to analyze proteins from cell culture supernatants, an additional “capture” step (e.g., affinity chromatography) may be required.65

**Native electrospray mass spectrometry (ESI-MS)** is the key to preserve protein higher order structures which enables the study of protein conformation and dynamics. In ESI-MS, the use of “volatile buffers” is necessary since most buffer solutions that are used in structural biology are not ESI-MS compatible and may interfere with the protein ionization. Those compatible buffers include ammonium acetate (NH₄Ac) and ammonium bicarbonate (NH₄HCO₃) with neutral pH and concentrations ranges from
5mM to 1M. Under these conditions, it is well known that quaternary protein structures can often be preserved.\textsuperscript{61, 62, 66} It is also worthy to note that, native ESI-MS volatile solvents usually generate less charged ions than those in organic ESI solvents because the surface of the folded species is more compact.\textsuperscript{67}

**Hydrogen deuterium exchange mass spectrometry (HDX MS).** Protein HDX studies focus on hydrogens that are located at main chain amides since they represent a continuous structural probe covering the entire protein chain. The HDX patterns of the amide hydrogens are affected by protein flexibility and mobility, thus, making HDX an extremely sensitive probe of protein structure and stability. HDX MS is a reliable, robust and sensitive technique as well as being a very promising technique for characterizing the conformation of protein drugs thanks to its ability to detect structurally compromised proteins in the background of the natively folded species in highly complex matrices.\textsuperscript{68}

**1.2.3. Experimental methods for pharmacokinetics (PK) studies**

**Enzyme-linked immunosorbent assay (ELISA)** is a sensitive and versatile method for both detection and quantitation of either antibodies or many antigens including viruses\textsuperscript{69}, hormones\textsuperscript{70}, peptides\textsuperscript{71}, and proteins.\textsuperscript{72} This immunological assay method is one of the most popularly used that can be applied for both single sample analysis and high-throughput screening thanks to its simple protocol. Principally, ELISA involves the specific binding of antibody to antigen with one of them immobilized to a solid support. The enzyme-conjugated antibody followed by a chromogenic substrate is used to amplify and visualize the binding activity.\textsuperscript{73}
Radiotracer. Radioisotopes of elements such as carbon⁷⁴, ⁷⁵, titanium⁷⁶, gallium⁷⁷, iodine⁷⁸ and indium⁷⁹ have been widely used for pharmacokinetic studies as well as tracking the distribution of biopharmaceuticals in cells or tissues. The radiotracer method is based on the idea that all isotopes of the same element will have the same chemical properties. Therefore, instead of detecting non-radioactive isotopes, it is far easier to follow the radioactive one in the complex matrix due to the high sensitivity and selectivity of this approach. However, the routine use of this method is hampered because of the health and safety issues over handling of radioactive compounds as well as the disposal costs.⁸⁰

MS-based methods to study pharmacokinetics of protein therapeutics. Pharmacokinetics of any drug, being traditional small molecule medicines or biopharmaceuticals, is one of the factors that decides its clinical success along with the ability to interact specifically with a desired therapeutic target.⁸¹ MS-based protein quantitation methods have dramatically expanded the scope in pharmacokinetic studies of protein drugs.⁸² For accurate and reliable quantitation of proteins, a stable isotope labeled internal standard is normally introduced in order to compensate for the matrix effect of the LC-MS measurements. The internal standard can be introduced at the protein level or at the peptide level. Undoubtedly, a protein-based internal standard is superior to a peptide-based method due to the fact that the errors associated with downstream sources can effectively be removed or compensated for. Stable isotopes used to label internal standard include ¹⁵N (whole protein labeling)⁸³, ¹³C (partial protein labeling, ¹³C₆-Arginine and ¹³C₆-Lysine)⁸⁴, ¹⁸O (protein labeling: ¹⁸O-cysteine⁸⁵, ICAT⁸⁶;
peptide labeling: $^{18}$O-Arginine and $^{18}$O-Lysine). Among them, $^{18}$O based methods are usually more attractive due to the relative ease and universality of labeling.

**Inductively coupled plasma (ICP) MS and its potential for protein quantitation.** ICP MS is a very powerful technique for elemental analysis that can be applied for the detection of metal- and metalloid-containing compounds. It relies on a hot argon plasma to break down compounds in homogenous solution into atoms of elements and eventually ionize them before introducing to a mass analyzer. This method is extremely sensitive toward metals due to their low ionization potential compare to other elements, i.e. halogen, thus, it is widely used for the quantitation of metals. In the last several years, ICP-MS has been applied in a many different areas, particularly bioanalysis. Indeed, ICP-MS provides an excellent alternative method for absolute protein quantitation via metal labeling strategy due to the high sensitivity and selectivity of the method. The most commonly used metal tags are the chelates of lanthanides due to their absence in biological samples, which mean that there is no spectroscopic interferences and a very low background signal. Furthermore, metal tagged protein drugs also enable the biodistribution study at the cellular level by imaging laser ablation (LA)-ICP-MS. Over the last few years, imaging LA-ICP-MS for the analysis of biological samples has been widely used i.e. cisplatine treatment in kidney, selenoproteins detection in gels, distribution of heavy metals in plants, or using $^{31}$P to indentify tumor boundaries in lymph node biopsies.

**1.3. Objectives**
Modern medicine development and detailed characterization requires state-of-the-art analytical techniques that can be used to analyze covalent structure, conformations and protein-receptor interaction for quantitative measuring biodistribution of protein therapeutics. Mass spectrometry has emerged as a powerful tool for these measurements. My work consists of multiple projects aiming to develop mass spectrometry based analytical protocols to assist the optimization and characterization of transferrin-based conjugates as well as tracking and constructing their distribution maps in vivo. Specific aims include: (i) exploring the properties of an Lz-Tf conjugate as a model therapeutic that targets the CNS and develop analytical protocols to characterize its structure and interactions with therapeutic targets and physiological partners critical for its successful delivery (Chapter 1); (ii) In-vivo studies of the biodistribution of transferrin-based drug in animal models by inductively coupled plasma mass spectrometry (ICP-MS) (Chapter 3); (iii) study of in-vivo aggregation of transferrin-based drug by size exclusion chromatography (SEC) and inductively coupled plasma mass spectrometry (ICP-MS) (Chapter 4); (iv) development of an $^{18}$O labeling based method to quantitatively map the modification sites of protein-drugs conjugates (Chapter 5).
### Table 1.1. Comparison of small molecule and biologic-based new molecular entities (NMEs) (1982–2013).⁴

<table>
<thead>
<tr>
<th>NME type</th>
<th>Total</th>
<th>IND to approval (years)</th>
<th>Withdrawn (safety concerns) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small molecule</td>
<td>777</td>
<td>8.5 (n = 200)</td>
<td>26 (3.3%)</td>
</tr>
<tr>
<td>Biologic</td>
<td>91</td>
<td>7.4 (n = 30)</td>
<td>2 (2.2%)</td>
</tr>
<tr>
<td>Monoclonal antibody</td>
<td>34</td>
<td>7.8 (n = 18)</td>
<td>2 (5.9%)</td>
</tr>
<tr>
<td>Enzyme modulator</td>
<td>26</td>
<td>5.9 (n = 8)</td>
<td>0</td>
</tr>
<tr>
<td>Receptor modulator</td>
<td>31</td>
<td>8.3 (n = 4)</td>
<td>0</td>
</tr>
</tbody>
</table>

**IND:** the average times from investigational new drug to approval

### Table 1.2. Key parameters for protein conjugates¹⁷

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Leading examples</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antigen/Receptors</td>
<td>Substantial expression on tumor, limited expression</td>
<td>Her2, CD30, CD33</td>
<td>18-21</td>
</tr>
<tr>
<td></td>
<td>on normal tissues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linker</td>
<td>High stability to avoid systemic release</td>
<td>Dipeptides, direct linkage, some hydrazones and disulfides</td>
<td>22-24</td>
</tr>
<tr>
<td>Drug</td>
<td>High potency since delivery is limited by antigen</td>
<td>Auristatins, maytansines, calicheamicin</td>
<td>25-27</td>
</tr>
<tr>
<td></td>
<td>copy number</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug loading level</td>
<td>Maintain mAb PK by limiting drug loading</td>
<td>2-4 drugs per mAb often optimal</td>
<td>28</td>
</tr>
<tr>
<td>Conjugation site</td>
<td>Homogeneous drug loading by site-specific</td>
<td>Chemical methods, cysteine point mutants</td>
<td>29-31</td>
</tr>
<tr>
<td></td>
<td>specific conjugation to avoid subpopulations with</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>altered PK</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PUBLICATIONS AND MANUSCRIPTS
CHAPTER 2
MASS SPECTROMETRY-GUIDED OPTIMIZATION AND CHARACTERIZATION OF A BIOLOGICALLY ACTIVE TRANSFERRIN-LYSOZYME MODEL DRUG CONJUGATE

This chapter was published on *Molecular Pharmaceutics*, 2013, 10, (5), 1998-2007

Son N. Nguyen, Cedric E. Bobst and Igor A. Kaltashov

*Department of Chemistry, University of Massachusetts Amherst, Amherst, MA*

*Figure 2.1. Chapter 2 table of content graphic*

address correspondence to:
Igor A. Kaltashov
Department of Chemistry
University of Massachusetts-Amherst
710 North Pleasant Street, LGRT 104
Amherst, MA 01003
Tel: (413) 545-1460
Fax: (413) 545-4490
Email: Kaltashov@chem.umass.edu

15
ABSTRACT

Transferrin is a promising drug carrier that has the potential to deliver metals, small organic molecules and therapeutic proteins to cancer cells and/or across physiological barriers (such as the blood-brain barrier). Despite this promise, very few transferrin-based therapeutics have been developed and reached clinical trials. This modest success record can be explained by the complexity and heterogeneity of protein conjugation products, which also pose great challenges to their analytical characterization. In this work, we use lysozyme conjugated to transferrin as a model therapeutic that targets the central nervous system (where its bacteriostatic properties may be exploited to control infection) and develop analytical protocols based on electrospray ionization mass spectrometry to characterize its structure and interactions with therapeutic targets and physiological partners critical for its successful delivery. Mass spectrometry has already become an indispensable tool facilitating all stages of the protein drug development process, and this work demonstrates the enormous potential of this technique in facilitating the development of a range of therapeutically effective protein-drug conjugates.

KEYWORDS

Biopharmaceuticals; drug delivery; protein-drug conjugate; protein cross-linking; ion exchange chromatography; electrospray ionization; mass spectrometry; protein-receptor binding; enzyme-substrate binding
2.1. Introduction

Transferrin (Tf) is a promising drug carrier that has the potential to deliver metals, small molecule medicines and therapeutic proteins to cancer cells\(^\text{32}\) and/or across physiological barriers (such as the blood-brain barrier, BBB\(^\text{33}\)). Despite this promise, very few Tf-based therapeutics have been developed and reached clinical trials. This very modest success record can be explained by the complexity and heterogeneity of protein conjugation products, which also pose great challenges to their analytical characterization. Possible interaction of the therapeutic payload with the carrier protein may have a negative impact on the conformational stability of the latter. Even in the absence of such interactions, the mere presence of the conjugate on the protein surface may sterically interfere with the ability of Tf to associate with its receptor at the cell surface, a critical first step in the drug delivery scenarios. Another complication arises due to the need to either have an effective mechanism of liberating the payload from Tf inside the endosome (in order to allow its routing to the intracellular target as opposed to recycling it back to the cell surface), or ensuring that it still exerts its therapeutic action while being attached to Tf if the latter is intended to ferry the payload across a physiological barrier, such as the BBB.

While most of the earlier explorations of Tf as a potential delivery vehicle focused on intracellular drug delivery,\(^\text{34-39}\) its ability to cross physiological barriers may also be exploited to target difficult-to-reach intracellular targets.\(^\text{43-47}\) For example, neuroanatomical obstacles frequently limit the effectiveness of antimicrobial therapy in the central nervous system (CNS) by preventing a large number of effective
antimicrobials from reaching sufficient concentration levels at the infection site.\textsuperscript{104} Furthermore, the ever increasing number of bacterial pathogens resistant to common antibiotics has brought to the fore the question of whether the repertoire of antimicrobials should be expanded beyond classical small molecule drugs (be they natural or synthetic products) by considering larger bio-inspired host defense systems, such as amphiphilic peptides\textsuperscript{105} and other bacteriostatic macromolecules. One particularly attractive class of such bacteriostatic agents is a group of enzymes that compromise the integrity of bacterial cell walls.\textsuperscript{106} Lysozyme (Lz) is an antibacterial enzyme present in a variety of organisms, which exerts its bacteriostatic function by hydrolyzing the $\beta$-1,4-glycosidic bond between the $N$-acetylmuramic acid (NAM) and $N$-acetylglucosamine (NAG) residues of peptidoglycans, resulting in lysis of bacterial cell walls. Although Lz primarily attacks Gram-positive bacteria, where the peptidoglycan layer is not protected by the outer membrane (as it is in Gram-negative bacteria), certain structural modifications can make it effective against Gram-negative bacteria as well.\textsuperscript{107}

While Lz is widely distributed throughout the human body, it is not present in the cerebrospinal fluid (CSF) of healthy subjects\textsuperscript{108} (detectable levels of Lz in CSF is usually associated with various CNS pathologies\textsuperscript{109, 110} and are likely to reflect increased permeability of the BBB\textsuperscript{110}). Therefore, the ability to deliver Lz “on demand” across the BBB might lead to development of novel effective therapeutic strategies aimed at the eradication of Gram-positive infections in the CNS, whose carriers gain access to the brain via a variety of routes.\textsuperscript{104}
In this work we explore the properties of Lz conjugated to Tf as a model therapeutic that targets the CNS and develop analytical protocols to characterize its structure and interactions with therapeutic targets and physiological partners critical for its successful delivery. We demonstrate that electrospray ionization (ESI) mass spectrometry (MS) provides a convenient and effective way to probe both the structure of the conjugation products and their ability to interact with physiologically and therapeutically relevant partners, thereby providing important and valuable feedback that can be used to refine and optimize the conjugation protocols and greatly facilitate the early stages of the drug development process.

2.2. Experimental

Preparation of Lz-Tf conjugate. For each reaction the optimal final conditions are given in the text while various reaction parameters tested during optimization are listed in parenthesis. Lz from chicken egg white (Sigma-Aldrich, St. Louis, MO) was activated (decorated with free thiol groups) using either Traut’s reagent (2-iminothiolane hydrochloride; Sigma-Aldrich, St. Louis, MO) or N-succinimidyl-S-acetylthioacetate (SATA; Pierce Biotechnology, Rockford, IL), which target primary amine groups. The reaction was carried out by incubating 12 h (0.5, 0.75, 1, 2, 12, 24 h) at 0 °C (0, 4, 25, 37 °C) in 50 mM phosphate buffer with 100 mM NaCl pH 8.0 (7.0, 7.5, 8.0, 8.5, 9.0) and consisted of 500 µM (50, 100, 250, 500 µM) Lz with a 1:1 ratio (2:1, 1:1, 1:2, 1:4) of primary amines relative to the thiolating reagent (freshly prepared in H2O). Incorporation of thiol-reactive maleimide groups into human Tf (provided by Prof. Anne B. Mason, University of Vermont College of Medicine) was carried out by reacting with
either sulfo-succinimidyl-4-[N-maleimidomethyl] cyclohexane-1-carboxylate (sulfo-SMCC; Pierce Biotechnology, Rockford, IL) or succinimidyl-[(N-maleimidopropionamido)-dodecaethyleneglycol] ester (SM(PEG)_{12}; Pierce Biotechnology, Rockford, IL) for 90 min (30, 60, 90, 120 min) at 25 °C (4, 25, 37 °C) in 50 mM phosphate buffer with 100 mM NaCl pH 7.0 (6.0, 7.0, 8.0, 9.0) and consisted of 250 µM (50, 100, 250 µM) Tf with a 1:2 (1:2, 1:4, 1:20) ratio of Tf relative to the activating reagent (freshly prepared in DMSO). Excess activation reagents were removed by centrifugal filtration through a 10 kDa Vivaspin molecular weight filter (Sartorius Stedim Biotech SA, Bohemia, NY) and the non-cleavable thio-ether linkage between activated Lz and Tf was formed by incubating 50 µM (25, 50, 100 µM) each of the two modified proteins together for 12h (1, 2, 4h, 12, 24h) at 4°C (4, 25, 37°C) in 50 mM phosphate buffer with 100 mM NaCl, pH 7.0 (7.0, 8.0, 9.0) at 1:1 molar ratio. The 1:1 Lz-Tf conjugate was isolated by cation exchange chromatography on a 4.6 x 100 mm PolyCATA™ column (5 µM, 1000 Å, PolyLC Inc., Columbia, MD) using an Agilent 1100 (Agilent Technologies, Palo Alto, CA) HPLC system. All relevant reaction diagrams are shown in Figure 2.2.

Mass spectrometry. All ESI MS measurements were carried out with a QStar-XL (ABI/SCIEX, Toronto, Canada) hybrid quadrupole/time-of-flight MS equipped with a nanospray source. Mass profiling of activated Tf, activated Lz, the crude reaction product mixture, and the 1:1 conjugate isolated by cation exchange LC was carried out following extensive buffer-exchange of proteins and placing them in water/methanol/acetic acid (49:49:2) at a concentration of ca. 10 µM. Native ESI MS analyses of the conjugate, and its mixtures with TfR and NAG₃ were performed using 20
mM ammonium acetate as a solvent. To ensure the integrity of non-covalent complexes in the gas phase, the declustering potential in the ESI MS interface (DP1) was minimized, unless noted otherwise in the text. Ectodomain of transferrin receptor (TfR)\textsuperscript{112} used in binding assays was provided by Prof. Anne B. Mason (Univ. of Vermont College of Medicine, Burlington, VT), and N-acetylglucosamine trimer (NAG\textsubscript{3}) was purchased from Sigma-Aldrich (St. Louis, MO).

*Antimicrobial activity assay.* Antimicrobial activity of intact Lz, Lz dimers and Lz-Tf conjugate was measured using re-suspended dried cells of *Micrococcus lysodeikticus* (Worthington Biochemical Corp., Lakewood, NJ) as the substrate. The rate of cell wall lysis in 0.1 M sodium phosphate buffer pH 7.0 at 25°C was monitored by recording the transmission at 450 nm.\textsuperscript{113} The measurements were carried out in a 1 mL cuvette with a NanoDrop 2000c (Thermo Fisher Scientific, Rockford, IL) UV-Vis spectrophotometer.

2.3. Results

2.3.1. Production, purification and characterization of the Lz-Tf conjugate

The classical scheme of conjugating Tf to a protein payload involves derivatizing the Lys side chains and amino terminus of Tf with sulfo-SMCC and activating the protein payload at similar sites using Traut’s reagent, followed by reacting them with each other\textsuperscript{114}. This produces the same thio-ether linkage that was used in the production of TransMID,\textsuperscript{115} the only Tf-based biopharmaceutical product that ever reached Phase III clinical trials. It is expected that a 1:1 stoichiometry for the conjugate would minimally disturb the functionality of either molecule and was the desired product of our
synthesis. Placing a single maleimide group on Tf and a single free thiol group on Lz should lead to the formation of a 1:1 conjugate, with Lz dimers being the only by-product that can form via external disulfide bond formation (Figure 2.2). While the extent of Tf functionalization with sulfo-SMCC and Lz with Traut’s reagent can be varied over a wide range it is virtually impossible to limit the extent of activation of the two proteins to a single reactive group on each polypeptide chain (Figure 2.2). Moreover, with 58 Lys residues in Tf and 6 in Lz, an additional level of heterogeneity is introduced by the number of linker positions that can form a 1:1 conjugate. Our initial development embraced this possible heterogeneity, benefitting from a quicker development time. Identifying conjugation sites would be suited for final optimization of the protein drug conjugate.

We found that adequate yields of the conjugation reaction can be achieved only if multiple activation groups are placed on each protein. Placing multiple free thiol groups on Lz is likely to increase the incidence and extent of this protein’s polymerization via formation of external disulfide linkages. While homo-polymerization of the functionalized Tf was not expected to be as significant (at neutral pH maleimide groups are ca. 1000 less reactive towards free amines compared to free sulfhydryls), this process nonetheless was found to occur. Homo-polymerization was particularly apparent when concentrations of Tf in the reaction mixture were elevated compared to that of the activated payload (Lz), most likely due to the presence of a large number of free amine groups (Lys side chains) on the surface of Tf. Above and beyond the formation of Lzₙ (and, to a lesser extent, Tfₙ) homo-polymers, polyvalent
functionalization of Tf and Lz was also expected to contribute to the extent of heterogeneity of the conjugation products (Figure 2.2C). Balancing the extent of modification of each protein in order to optimize production of the 1:1 Lz-Tf conjugate was achieved in this work by controlling the following primary variables: reagent concentrations, temperature, incubation time and reaction pH. Mass spectrometry enabled us to determine the effect of altering these variables on the product, allowing us to screen these parameters in an iterative fashion that sought to optimize the yield of conjugation while minimizing undesirable side reactions.

Of the two modification reactions, thiolation by Traut’s reagent was the most problematic. Two significant issues that needed to be overcome were the reduction of intramolecular disulfide bonds and formation of N-substituted 2-iminothilane (NSI) products. The abundance of intramolecular disulfide bonds within Tf presented an unexpected obstacle when this protein was initially chosen for activation by Traut’s reagent. While the highest yield of thiolated Tf could be obtained at elevated pH and a high concentration of reagents, the native disulfide bonds of Tf were largely reduced, leading to a drastic and unacceptable change in its higher order structure (see Supporting Information Figure S1 for more detail). Disulfide reduction was still observed (albeit to a much lesser extent) in the thiolation of Lz by Traut’s reagent, but was practically eliminated by reducing reagent concentrations and performing the reaction on ice. Further investigation of this reaction led us to an interesting finding; In addition to thiolated proteins, other chemical modifications were observed with masses consistent with non-thiol by-products reported by Singh. These dead-end (non-
reactive) products form when the unstable thiol adduct breaks down into a non-reactive 5-membered ring and reduce the number of free thiol groups. Such instability in one of the activating groups can significantly hinder downstream conjugation even when all six Lys residues of Lz have been functionalized (Figure 2.3B). Formation of NSI by-products was also observed during and after the conjugation reaction, but fortunately could be minimized by lowering the reaction pH and temperature. We investigated the use of N-succinimidyl S-acetylthioacetate (SATA) as an alternative reagent to introduce the desired thiol group. SATA introduces a “protected sulfhydryl” which requires activation by a mild reducing reagent to expose the sulfhydryl prior to the conjugation reaction. Compared to Traut’s reagent, Lz functionalized with SATA had increased stability (did not form NSIs) and reduced heterogeneity (Figure 2.3B,D).

Optimizing the activating steps for each protein focused initially on minimizing deleterious or non-productive by-products such as the reduction of intramolecular disulfide bonds or the formation of NSIs as monitored by ESI MS. After screening pH values from 7.0 to 9.0, reaction temperatures from 0 to 37°C, and reaction times from 0.5 to 24 hours, the optimized thiolation reaction was performed at pH 8, 0°C, for 12 h. Notably, in the examined range of protein concentrations (50 µM to 500 µM), elevated protein concentration was found to result in a higher yield of modified Lz. Different ratios of protein primary amine to reagent from 2:1 to 1:4 were used to generate a series of differentially modified Lz. The optimal extent of Lz activation was determined in conjunction with the optimal extent of Tf activation selecting for values that produced the highest yield of the 1:1 Lz-Tf conjugate. The final conditions for Lz activation utilized
an equimolar ratio of primary amine groups relative to activating reagent and introduced 2 to 2.5 thiol groups per Lz. Similar reaction parameters were screened to optimize the activation of Tf with its thiol reactive group. Tf was optimally activated at pH 7.0, 25 °C using 50 μM of Tf and a ratio of Tf to activating reagent of 1:2 for 90 min. Under these conditions, the average number of functional groups incorporated into Tf was 1.5. In the final round of optimization, formation of the Lz-Tf conjugate was monitored as a function of pH, time, temperature as well as the extent of activation of each protein. The highest yield of a 1:1 Lz-Tf conjugate was obtained at pH 7 allowing optimally activated Tf and Lz (50 μM each) to react for 12 h at 4 °C. Nevertheless, even though ESI MS analysis of the optimized conjugation reaction (Figure 2.4A) clearly shows the presence of the 1:1 Lz-Tf conjugate (charge states assigned based on the calculated mass of 94.4 kDa are shown in Figure 2.4A with dotted lines), a large number of other species are also present in the mixture. Therefore, evaluation of various properties of Lz-Tf is impossible without its separation from other products and/or reagents. Since the incremental mass increase of Lz-Tf over intact Tf makes use of SEC impractical for purification of the conjugation products, alternative methods of separating the 1:1 Lz-Tf conjugate from other components of the reaction mixture were examined.

The significant difference in pl values for Tf (5.5-6.3) and Lz (11.0) made ion exchange chromatography (IXC) particularly attractive as a means of purifying the reaction products. Using a weak cation exchange stationary phase, a mobile phase buffered to pH 6.5, and a shallow salt gradient we have been able to achieve separation between the Tf and Lz peaks exceeding 15 minutes (Figure 2.4C), with Lz homo-
polymers having even longer elution times. The products of the Lz/Tf conjugation reaction elute within a wide (9-17 min) time period and are mostly unresolved, although a distinct peak is observed at 14 min elution time. Collection of a corresponding IXC fraction (13.5-14.5 min) followed by quick desalting and off-line ESI MS analysis yields a mass spectrum consistent with the 1:1 Lz-Tf conjugate as the major component of this fraction (Figure 2.4B).

Even though the collected IXC fraction represents a 1:1 Lz-Tf conjugate, there still might be a significant degree of micro-heterogeneity due to the presence of modified Lys side chains on the surfaces of both Tf and Lz. Indeed, native ESI MS analysis of this fraction spiked with intact Tf (Figure 2.5) clearly shows significantly broader peak shapes for multiply charged Lz-Tf ions compared to intact Tf ions. A more detailed analysis of the mass spectrum reveals very convoluted peak shapes for Lz-Tf ions (insets in Figure 2.5), where the broad mass distribution of ionic species is due to the presence of either unreacted maleimide groups on the surface of Tf and/or dead-end NSI groups on the surface of Lz. Despite these extensive modifications, no large scale conformational changes are apparent as a result of the conjugation reaction, as the charge state distribution of Lz-Tf ions is consistent with both components of the conjugate maintaining compact structures in solution (no ions were detected in the low m/z region, whose presence in ESI MS usually signals either partial or complete protein unfolding in solution117).

2.3.2. Influence of conjugation and chemical modifications on interaction with transferrin receptor (TfR)
Although examination of the Lz-Tf conjugate with native ESI MS suggests that neither protein undergoes unfolding as a result of the conjugation, the ability of both Tf and Lz to interact with their physiological partners and/or therapeutic targets may nonetheless be compromised as a result of unfavorable location of the cross-link, as well as multiple modification of Lys residues on the surface of either protein beyond the cross-link sites. For example, Lz cross-linked to Tf may interfere with the ability of the latter to bind to TfR, thereby rendering Lz-Tf incapable of crossing the BBB.

Native ESI MS provides an easy way to evaluate protein binding to a variety of ligands, including both small molecules and biopolymers,\textsuperscript{118, 119} and in some instances allows the binding affinity to be estimated.\textsuperscript{120, 121} This approach has been used in the past to monitor TfR recognition by wild type Tf and its mutants under a variety of conditions,\textsuperscript{122, 123} and recently we were successful in using this approach to monitor interactions of a Tf-based fusion protein with TfR.\textsuperscript{124} However, native ESI MS has never been used to evaluate interaction of protein-protein conjugates with their physiological partners. An ESI mass spectrum of the Lz-Tf/TfR mixture acquired in this work under near native conditions (neutral pH, ionic strength 20 mM) clearly indicates that the receptor does recognize the conjugate, although the binding affinity is diminished compared to intact Tf (Figure 2.6). Indeed, no ionic signal of unbound Fe\textsubscript{2}Tf is detected in the mass spectrum of the Fe\textsubscript{2}Tf/TfR mixture, consistent with the receptor-binding affinity of Fe\textsubscript{2}Tf being in the sub-µM range (concentration of both proteins in the Fe\textsubscript{2}Tf/TfR mixture was in the low-µM range, 3 µM). At the same time, the presence of a weak, but detectable ionic signal of unbound Lz-Tf in the mass spectrum of the Lz-Tf/TfR
mixture acquired under identical conditions suggests that the TfR binding affinity of the conjugate is in the low-µM range. This affinity range is close to that of intact apo-Tf,\textsuperscript{123} even though the conjugate was saturated with iron following its isolation from the reaction mixture and its measured mass is in agreement with the diferric form. Nevertheless, even this lower receptor affinity should be sufficient for transient binding to TfR at the cell surface (endogenous Tf is only 30% saturated with iron), and may actually prove beneficial for dissociation from TfR upon crossing the BBB.

2.3.3. Influence of conjugation and chemical modification on enzymatic activity

In order to exhibit bacteriostatic properties, enzymatic activity of Lz must be preserved within the conjugate. Enzymatic activity of Lz and its variants is frequently probed using the short tri-saccharide NAG\textsubscript{3} as a surrogate substrate to demonstrate the substrate-binding competence of the protein\textsuperscript{125} (Figure 2.7A). Despite the proximity of the six Lys residues (targets of modification by Traut’s reagent) to the catalytic site of the enzyme, Lz-Tf retains the ability to bind NAG\textsubscript{3}, indicated by the presence of the protein-NAG\textsubscript{3} complexes in the mass spectrum of a Lz-Tf/NAG\textsubscript{3} mixture (Figure 2.7B) acquired under near-native conditions (neutral pH and 20 mM ionic strength), a behavior very similar to that exhibited by intact Lz.

Despite the promising substrate binding results obtained using NAG\textsubscript{3}, a photometric-based activity assay that measures the lysis of Gram-positive bacteria\textsuperscript{126} indicated a very significant loss of bacteriolytic activity by Lz-Tf (Figure 2.8). As can be seen in Table 2.1, the bacteriolytic activity of Lz-Tf is reduced by two orders of magnitude compared to the control (intact) Lz, which may be attributed to two factors. First, chemical modification
of Lys residues on the Lz surface changes the electrostatic properties of this protein (as reflected by its behavior in IXC, vide supra), which may introduce conformational changes in the vicinity of the catalytic site, thereby adversely affecting the ability of this enzyme to recognize large substrates. Second, the presence of a large protein (Tf) anchored to a Lys chain located in the vicinity of the catalytic site of Lz may introduce steric effects, which would make it more difficult for this protein to attack the bacterial cell wall even in the absence of any conformational changes.

Importantly, bacteriolytic activity of the Lz dimer is reduced four-fold, the actual activity of the dimer is 37.2 % compared to that of intact Lz, and is reported as 18.6 % in Table 2.1 after adjusting it by a factor of 2 due to the presence of two catalytic sites in a single Lz₂ molecule. The Lz dimer byproduct conveniently serves as an important control demonstrating that while chemical modification of primary amines on the enzyme surface by Traut’s reagent contributes to the reduction in catalytic activity, it is not the primary reason for Lz inactivation (this is also consistent with the observed ability of Lz-Tf to bind short substrate surrogates, vide supra). Therefore, the impaired ability of Lz-Tf to catalyze the hydrolysis of large glycans is largely due to the significant steric restraints introduced by a bulky “anchor” (Tf). Metaphorically, Tf can be viewed as an “elephant on a leash,” which allows the payload to be delivered to a desired location by overcoming physiological barriers, but also restricts its freedom, thereby making it less effective in exerting the desired therapeutic action.

One possibility to mitigate this negative effect is offered by longer linkers, which should increase the freedom of movement of Lz cross-linked to Tf, allowing it to attack
the bacterial cell walls more effectively. A possibility to create protein/protein conjugate with a longer linker is offered by amine-reactive SM(PEG)$_{12}$, which also introduces a thiol-reactive maleimide group on the protein surface (Figure 2.2C). Activation of Tf with this reagent and introduction of free thiol groups to Lz using, SATA (to avoid formation of dead-ended by-products introduced by the Traut’s reagent, see Figure 2.3 B and D) leads to formation of a conjugate that is recognized by TfR (see Supporting Information Figure S2) and has anti-bacterial activity over an order of magnitude higher than that of the conjugate with a short linker (Figure 2.8). Although a fraction of that increase could be attributed to using SATA as the enzyme-modifying reagent (note that the covalent dimer of SATA-modified Lz retains half the activity of the intact enzyme, see Table 2.1), the most significant gain is a result of using the longer (and more flexible) linker.

2.4. Discussion

The incidence of CNS infections (brain infections) is rising at an alarming rate, while the treatment options remain very limited.$^{104}$ Only a very small faction of existing small-molecule medicines can penetrate the BBB, and none of the currently approved protein therapeutics is capable of doing so. Coupled with the ever increasing resistance of pathogens to common antibiotics and dire side effects of the immune system’s inflammation response to infection (frequently leading to brain abscess), this presents the clinicians with a grave challenge. Clearly, there exists a significant and unmet need for novel (bio)pharmaceuticals that can control CNS infections in the most efficient way without eliciting an immune response from the host.
Chemical conjugation of a therapeutic payload (a small molecule medicine or a protein drug) to a transport protein, such as Tf, offers a convenient and inexpensive way to produce effective medicines that can be delivered to target tissues and cells. However, only one Tf-based therapeutic has ever reached Phase III clinical trials and was subsequently withdrawn.\textsuperscript{127, 128} This rather modest record of success can be explained by the tremendous complexity and heterogeneity of conjugate species, a feature that not only complicates the underlying biology, but also frequently prevents effective utilization of state-of-the-art analytical technologies at various stages of the drug development process. MS has been a critical component in the analytical armamentarium supporting protein drug development efforts,\textsuperscript{129, 130} but its applications in characterization of protein-protein conjugates, such as TransMID, have been very limited so far. Lower-end analytical techniques, such as size-exclusion chromatography (SEC) and circular dichroism (CD) spectroscopy, which have been traditionally applied to characterize protein-protein conjugates,\textsuperscript{131} do not provide sufficient resolution and may in fact be misleading when relied upon as a sole source of information to describe the conjugate’s molecular weight distribution or its conformation.

In this work we demonstrated that ESI MS can provide characterization of both the products and intermediates of protein-protein conjugation reactions at great detail. Characterization of the activated proteins with ESI MS provides an important feedback for optimization of the conjugation protocol, while MS analysis of the reaction products highlights the enormous degree of structural heterogeneity and underscores the need for chromatographic separation as a means of controlling the extent of heterogeneity.
MS characterization of the IXC-purified 1:1 conjugation product demonstrates that its heterogeneity is substantially reduced, but not completely eliminated, as the mass distribution reveals a significant number of additional chemical modifications to both protein components of the conjugate. Another level of structural heterogeneity is due to the large number of combinations that can be used to describe the spatial distribution of both the cross-link and unreacted modification sites on the surface of each protein component. Although structural heterogeneity at this level is not apparent when considering MS data alone, it can be easily visualized when methods of tandem mass spectrometry (MS/MS) are applied either in conjunction with enzymatic degradation of the conjugated protein, or alone (the so-called top-down MS/MS), an approach that has been used successfully in the recent past to characterize the distribution of conjugation sites in PEGylated proteins\(^\text{132}\) and protein-small molecule drug conjugates.\(^\text{133}\) Since it is likely that the location of the linker influences the activity of the conjugate, identifying specific conjugation sites utilizing mass spectrometry will be the focus of future work. Ideally, one would be able to correlate linker position with conjugate activity allowing for yet an additional level of optimization in the design of the protein conjugate.

Structural complexity and heterogeneity of the Lz-Tf conjugate highlighted by ESI MS brings to the fore the question of how conjugation and chemical modification may affect the ability of Tf to be recognized by its receptor, a crucial first step in receptor-mediated transcytosis, without which no delivery of the payload to CNS would be possible. ESI MS has been used in the past to monitor protein-receptor interactions involving protein
drugs, such as association of interferon β1a (Avonex™) with its cognate receptors, and to evaluate the modulation of its receptor-binding competence by a specific (well-defined) chemical modification. Heterogeneity of the 1:1 conjugation product of Lz and Tf does not allow a precise correlation between structural changes and receptor-binding properties to be established. Instead, TfR-binding competence is evaluated in this work for the entire ensemble of Lz-Tf species. Although native ESI MS analysis provides evidence for some loss of the receptor affinity, the conjugate is nonetheless clearly recognized by the receptor. The ability of Lz-Tf to associate with TfR suggests that transfer of this species from the bloodstream to the CNS via receptor-mediated transcytosis is possible (and likely), although the efficiency of this process can be estimated only by in vivo studies capable of measuring protein levels in various biological fluids (a direction actively pursued in our laboratory).

ESI MS also provides an exciting opportunity to evaluate the retention of enzymatic activity of Lz following its conjugation to Tf. Since the therapeutic targets of Lz (peptidoglycans from the cell walls of Gram-positive bacteria) are too large and heterogeneous for direct ESI MS analysis, most studies use a trisaccharide molecule NAG₃ as a surrogate substrate. Interestingly, native ESI MS measurements indicate robust binding of NAG₃ to Lz-Tf, while the actual biological activity test indicates a significant (by over two orders of magnitude) loss of its bacteriolytic activity compared to intact Lz. Auto-conjugation of Lz does not result in such a dramatic loss of bacteriolytic activity of this protein, suggesting that it is the presence of a large anchor
(Tf), rather than the chemical modification of the enzyme surface, that prevents Lz-Tf from effective execution of its desired therapeutic function.

An obvious solution to a steric hindrance problem is introduction of a longer and/or more flexible linker between the payload and the transport protein. Switching from SMCC (spacer arm length 8.3 Å) to SM(PEG)$_{12}$ (53.4 Å) results in a dramatic increase of autonomy for each protein component within the conjugate product, leading to a nearly 30-fold increase of its bacteriostatic activity. Such a dramatic recovery of the antibacterial activity of the conjugate is well beyond the much more modest enhancement provided by switching from the Traut’s reagent to SATA. In retrospect, utilization of longer and more flexible linkers may seem an obvious measure to enhance the potency of any conjugate by minimizing interaction of its two protein components; however, this consideration is not always brought to the fore when a specific type of conjugation chemistry is selected. In fact, it was the parallel analysis of ESI MS and biological activity data that illuminated this problem in the Lz-Tf conjugate and allowed for rational optimization of the conjugate product, improving its therapeutic potential.

2.5. Conclusions

A suite of ESI MS-based methods has been applied to characterize the structural and conformational integrity of a model bacteriostatic agent (Lz) conjugated to a transport protein (Tf), as well as its interaction with a physiological partner (TfR) critically important for delivery of this product to the CNS. Interaction of Lz-Tf with therapeutic targets was evaluated initially using ESI MS to monitor binding to a small surrogate substrate (NAG$_3$) followed by measuring its bacteriolytic activity, and comparing its level
to that of the intact Lz and Lz dimer. Analysis of these data led to the conclusion that steric hindrance imposed by a large protein anchored closely to the Lz surface reduced its biological activity. Increasing the autonomy of Lz by lengthening the linker lead to a dramatic increase in the bacteriolytic activity of the conjugate. ESI MS has already become an indispensable tool facilitating all stages of the protein drug development process,\textsuperscript{129} and this work demonstrates the enormous potential of this technique as a means to facilitate development of a range of therapeutically effective protein-drug conjugates. While mass spectrometry is beginning to enjoy wider acceptance in the biopharmaceutical community beyond the trivial tasks of primary structure elucidation,\textsuperscript{134, 136} its applications for the analysis of protein-drug conjugates have been limited primarily to measuring stoichiometry of the conjugation.\textsuperscript{137-139} This field has experienced an explosive growth in the past several years due to extensive efforts invested in developing antibody-drug conjugates (ADC),\textsuperscript{140, 141} and ESI MS clearly has a tremendous potential in this arena by providing invaluable information beyond mass measurement that can be used to optimize protein drug conjugate structures during early stages of development, and further catalyzing the drug design efforts.

\textbf{SUPPORTING INFORMATION}

ESI mass spectra of intact Tf and Tf modified with the Traut's reagent acquired under denaturing conditions, and an ESI mass spectrum of the mixture of Lz-Tf conjugate (produced with SM(PEG)\textsubscript{12} as a linker) and TfR acquired under near-native conditions. This material is available free of charge via the Internet at \url{http://pubs.acs.org}.
ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institutes of Health R01 GM061666. The authors are grateful to Prof. Anne B. Mason (Univ. of Vermont College of Medicine, Burlington, VT) for providing TfR sample and to Dr. Stephen J. Eyles (UMass-Amherst) for useful discussions.
## TABLES

**Table 2.1.** Bacteriolytic activity of Lz-Tf and related proteins

<table>
<thead>
<tr>
<th>Samples</th>
<th>Rate (mA/min)</th>
<th>Specific activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blank (negative control)</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Tf (20 µM)</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Lz control (2.5 µM)</td>
<td>124</td>
<td>100</td>
</tr>
<tr>
<td>Lz-Tf conjugate (6.7 µM)</td>
<td>1.5</td>
<td>0.45</td>
</tr>
<tr>
<td>Lz-Tf conjugate (30.6 µM)</td>
<td>6.4</td>
<td>0.42</td>
</tr>
<tr>
<td>Lz-Tf longer linker (6µM)</td>
<td>35</td>
<td>11.8</td>
</tr>
<tr>
<td>Lz dimer 2IT(7.5 µM)</td>
<td>138</td>
<td>18.6</td>
</tr>
<tr>
<td>Lz dimer SATA (2.5 µM)</td>
<td>125</td>
<td>50.6</td>
</tr>
</tbody>
</table>
FIGURES

**Figure S1.** ESI mass spectra of intact Tf and Tf modified with the Traut’s reagent acquired under denaturing conditions. The three traces correspond to an unmodified Tf (blue), 1:10 (red) and 1:20 (purple) concentration ratio of Tf to Traut’s reagent in the reaction. The higher charge state peaks of activated Tf compare to control Tf indicate further unfolding due to the reduction of intramolecular disulfide bonds.
Figure S2. ESI mass spectrum of the mixture of Lz-Tf conjugate (produced with SM(PEG)12 as a linker) and TfR acquired under near-native conditions. BSA that was present as an impurity in the TfR sample also served as a negative control to ensure specificity of binding for the Lz-Tf conjugate.
Figure 2.2. Schematic diagrams illustrating conjugation of Tf to Lz using the Traut’s reagent and SMCC (A) and SATA and SM(PEG)12 (B), and possible side reactions due to excessive activation of the two proteins with the Traut’s reagent and SMCC (C).
Figure 2.3. ESI mass spectra of activated Tf, charge state +32 (A) and Lz, charge state +10 (B) showing a range of reactive groups attached to the surface of each protein. The three traces shown in panel A correspond to a 1:2, 1:4 and 1:20 concentration ratio of Tf to sulfo-SMCC in the reaction. The two traces shown in panel B correspond to Lz modified with Traut’s reagent for 12 hours on ice (the gray trace corresponds to the protein-to-reagent molar ratio 1:2, the product kept at room temperature for 24 hours prior to MS; and the black trace corresponds to the protein-to-reagent molar ratio 1:7, MS analysis of the product was carried out immediately upon reaction completion). The multiple peaks shown in panel B are due to the presence of both chemically active (thiols) and de-activated (rings) groups on the surface of Lz. All mass spectra were acquired under denaturing condition (10μM total protein concentration in water/methanol/acetic acid, 49:49:2 by volume). Panels (C) and (D) show Tf and Lz activated with SM(PEG)12 and SATA, respectively. The series of peaks indicated with an asterisk represent salt adducts.
Figure 2.4. ESI mass spectra of crude (A) and IXC-purified (B) conjugation products of Lz and Tf. Panel C shows IXC chromatogram of the crude mixture (the fraction whose mass spectrum is shown in panel B is highlighted in the chromatogram).
Figure 2.5. ESI mass spectra of the purified 1:1 Lz-Tf conjugate (short linker) spiked with intact Tf (A) and a 1:1 conjugate produced with a longer linker (B) acquired under near-native conditions (3 μM of each protein in 20 mM ammonium acetate, pH 7.1). Insets zoom in on a selected charge state for each of the conjugates.
Figure 2.6. ESI mass spectra of Tf/TfR (A) and Lz-Tf/TfR (B) mixtures acquired under near-native conditions (3 µM of each protein in 20mM ammonium acetate, pH 7.1).
Figure 2.7. ESI mass spectra of NAG$_3$/Lz (A), NAG$_3$/Lz-Tf (B), NAG$_3$/Lz$_2$ (C) and NAG$_3$/Lz-Tf longer linker (D) mixtures acquired under near-native conditions (5µM of proteins and 10µM NAG$_3$ in 20mM ammonium acetate, pH 7.1).
Figure 2.8. Antibacterial activity data for Lz-Tf conjugates compared to that of intact Lz and Tf (A) and Lz dimers (B).
CHAPTER 3

BIODISTRIBUTION STUDIES OF TRANSFERRIN-BASED DRUG IN ANIMAL MODELS BY INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY (ICP-MS)

Son N. Nguyen, Hanwei Zhao, Cedric E. Bobst, Gokhan Elci and Igor A. Kaltashov

Department of Chemistry, University of Massachusetts Amherst, Amherst, MA

Figure 3.1. Chapter 3 TOC graphic

*address correspondence to:
Igor A. Kaltashov
Department of Chemistry
University of Massachusetts-Amherst
710 North Pleasant Street, LGRT 104
Amherst, MA 01003
Tel: (413) 545-1460
Fax: (413) 545-4490
Email: Kaltashov@chem.umass.edu
3.1. Introduction

Human serum transferrin (Tf) is a promising drug carrier with the ability to cross the blood-brain barrier that may enable specific targeting to cancer cells and non-invasive delivery of drugs to the central nervous system (CNS). In our previous work, lysozyme-transferrin conjugates (Lz-Tf) were developed as a model CNS-targeting drug. Lz-Tf was successfully synthesized and well characterized including in-vitro studies, however, in-vivo quantitation proved challenging due to the very high background of endogenous Tf (at least three orders of magnitude higher than injected Tf). Initially, a proteomics-inspired technique developed in our laboratory (based on $^{18}$O-labeling of cysteine residues in the protein) was used to study the biodistribution of exogenous Tf in animal models (rats). Despite the high sensitivity (0.2 ppm) of this method that enables the quantitation of exogenous Tf in blood up to 24 hours following the injection, it was not sufficient for studies of physiological barriers crossing (e.g., exogenous Tf could not be detected in cerebrospinal fluid). Additionally, due to the high background level of endogenous Tf, a Ni column was used to purify and preconcentrate the His-tag containing recombinant form of Tf from the biological samples post-administration. However, since recombinant therapeutic proteins have their His-tags removed prior to administration, reliance on the His-tag may limit the utility of this method. Thus, a very high sensitivity universal method without reliance on a His-tag for quantitation of exogenous Tf and Tf-based drugs in biological fluids and tissues is needed. Moreover, multiplexing capable method that allows highly selective measurements of a reference Tf alongside its variants and/or Tf-drug conjugates to be carried out in a single animal
following injection is highly desired. This will not only significantly reduce the number of animals needed for these studies but also results in a dramatic increase in the accuracy of measurements.

In order to achieve this goal, our lab has recently developed a new sensitive and selective method that uses metal tracers and inductively coupled plasma (ICP) MS with the possibility of multiplexing for quantitation of exogenous Tf and Tf-based drugs in biological tissues and fluids. ICP MS is a very powerful technique for elemental analysis,\textsuperscript{88-90} especially for measuring metal- and metalloid-containing compounds.\textsuperscript{91, 92} In the last several years, ICP-MS has been applied to many different areas, particularly bioanalysis.\textsuperscript{93} Although Tf is a metalloprotein, its cognate metal ion (iron) is very ubiquitous in all biological tissues and fluids and cannot be used as a tracer of Tf. In order to solve this problem, Indium (In), a non-endogenous metal that bind tightly to Tf in form of a stable In\textsubscript{2}Tf complex were used as the metal tag.

Furthermore, metal tagged Tf and Tf-based drugs also enable the biodistribution study of proteins at tissue level by imaging LA-ICP-MS. Although imaging LA-ICP-MS for the analysis of biological samples has been widely used in the last few years,\textsuperscript{94-98} this is the first time this imaging method was used for monitoring protein drugs with the help of a metal tracer.

In this work, we apply this new sensitive and selective ICP-MS based method for quantitative profiling of exogenous Tf and Tf-drug conjugate in biological fluids and their corresponding distribution patterns in tissues.
3.2. Experimental

3.2.1. Preparation of indium (In) loaded Lz-Tf conjugate

Lz-Tf conjugate was synthesized as described previously.\textsuperscript{48} Lz from chicken egg white (Sigma-Aldrich, St. Louis, MO) was activated using N-succinimidyl-S-acetyltioacetate (SATA; Pierce Biotechnology, Rockford, IL), while human Tf (Sigma-Aldrich, St. Louis, MO) was decorated with thiol-reactive maleimide groups (succinimidyl-([N-maleimidopropionamido]-dodeca-ethyleneglycol) ester (SM(PEG)\textsubscript{12}; Pierce Biotechnology, Rockford, IL). The 1:1 Lz-Tf conjugate was isolated by cation exchange chromatography on a 4.6 x 100 mm PolyCATA\textsuperscript{\texttrademark} column (5 μM, 1000 Å, PolyLC Inc., Columbia, MD) using an Agilent 1100 (Agilent Technologies, Palo Alto, CA) HPLC system. Indium (In) was then loaded to Lz-Tf following the removal of iron as described previously.\textsuperscript{343}

3.2.2. Animal Models

Experiments were carried at the animal facilities of the University of Massachusetts, Amherst on 10-week-old, male Wistar rats purchased from Charles River Laboratories (Wilmington, MA, USA). Rats were housed in controlled light (12 h light/dark cycle), temperature and humidity conditions with free access to food and water. All animal experiments were complied with the National Institute of Health Guidelines for Animal Care and were approved by the University of Massachusetts, Amherst Institutional Animal Care and Use Committee.
Animals were divided into three groups: G1: control rats (i.v. injected with 100 mM sterilized PBS solution); G2: Tf rats (i.v. injected with 5mg/kg dose of In\textsubscript{2}Tf); G3: Tf-Lz rats (i.v. injected with 5mg/kg dose of In\textsubscript{2}Tf-Lz). All In\textsubscript{2}Tf and In\textsubscript{2}Tf-Lz samples were buffer exchanged into 100mM sterilized PBS buffer solution prior to injections. Rats were euthanized by CO\textsubscript{2} overdose 24 h after injection. Blood samples were collected in heparin coated tube to prevent coagulation. Organs of interest (liver, spleen, lungs, kidneys, heart, intestine and brain) were harvested, rinsed in PBS, dried, weighed and stored in a -80\textdegree C freezer. The experimental sequence is presented in Figure 3.1.

3.2.3. ICP-MS

Small pieces cut organs (50mg) were oxidized in 0.5 mL of acid mixture (67% HNO\textsubscript{3} and 30% H\textsubscript{2}O\textsubscript{2} at a 3:1 volume ratio) and shaken overnight at 50\textdegree C until they dissolved fully. After the addition of 1ppb Rh as internal standard and 9.5 mL of deionized water, the sample was centrifuged at 3200 \times g for 15 min to remove cell debris, leaving the supernatant for Indium content analysis using a NexION 300X ICP-MS (Perkin Elmer, Waltham, MA, USA). Prior to the analysis, daily tuning was done to make sure the instrument was in optimum conditions. All signals were obtained using the collision cell with kinetic energy discrimination mode (KED). Operating conditions for the experiment are illustrated in Table 3.1. A calibration curve of various signal ratio of In/Rh was used to measure the In content, using 2.5% HNO\textsubscript{3} and 0.37% H\textsubscript{2}O\textsubscript{2} as the blank. All data were reported as In concentration in ppb (ng/g) and Tf concentration in ppm (µg/g) as well as the injected dose percentages. Error bars indicate 1SD from three different rats in each group.
3.2.4. Imaging LA-ICP-MS

Organs of rats were embedded in Tissue-Tek® O.C.T. compound (Sakura Finetek USA, Torrance, CA, USA) and cryocut at -20°C into slices of 20 µm thickness with a LEICA CM1850 cryostat (Leica Microsystems GmbH, Wetzlar, Germany). The slices were placed on regular glass slides and air dried. Experiments were conducted with a CETAC LSX-213 G2 laser ablation system (Nebraska, USA) coupled to the Perkin Elmer NexION 300X ICP-MS. LA unit was synchronized with the ICP-MS in external triggering mode. Organs sections were inserted into the cell and ablated line by line. The operation conditions for laser ablation system are listed in Table 3.1. Elemental images were reconstructed using ImageJ, an open source image generator software (http://rsbweb.nih.gov/ij/), following the extraction and compilation of individual element list of pixels from raw data files, each of which was acquired from a single ablated line contains values for all three metal measured.

3.3. Results and discussion

3.3.1. In Vivo Distribution of Tf and Lz-Tf

The ICP-MS based method for transferrin-based therapeutics quantitation using $^{115}$In as a metal tracer recently developed in our lab was used in this work for in-vivo tracking of the injection of In$_2$Tf and In$_2$Lz-Tf by quantitating their localization to key organs. Exogenous Tf and Tf-Lz concentration from spleen, liver, kidney, lung, heart, intestine and brain as well as body fluids i.e. blood, urine and CSF collected 24 hours post-injection was determined by measuring $^{115}$In concentration. Despite the complexity of
the sample matrix as well as the presence of large number of impurities, largely from inexpensive H₂O₂ and HNO₃ used for sample digestion prior to analysis, two stable isotopes of In (¹¹³In and ¹¹⁵In) could be readily resolved by ICP-MS. The distribution of the In₂Tf and In₂Lz-Tf in organs and body fluids 24 hours after intravenous injection is shown in Figure 3.1. Panel A shows absolute In concentration from 3 groups directly measured from organs, while panel B indicates the In contents as a fraction of the injected dose. Data from control group indicates that the sources of error were mainly from chemical impurities and instrumental noises other than that of method and human error. At the organ level, spleen (15.2±3.2µg/g), liver (14.1±3.5µg/g), kidney (14.4±3.6 µg/g) and intestine (4.6±1.6 µg/g), as expected, are the highest uptake organs given the fact that they all play a very important role in clearance of blood-circulating drugs. Spleen is the body’s largest filter of the blood in which old erythrocytes are removed and iron is recycled.¹⁴⁴ On the other hand, liver is the largest internal organ in which chemicals and metabolized drugs are detoxified and filtered blood coming from the digestive tract is passed to the rest of the body.¹⁴⁵ Contrary to that, a relatively low level of In accumulation is observed in organs that do not play a role in drugs clearance such as the heart (2.1±0.4µg/g) and lungs (3.4±0.4µg/g). Additionally, there was a noticeable abundance of In₂Tf in the blood circulation (4.7±1.9µg/g). However, when the size of organs taken into consideration, ¹¹⁵In was largely accumulated in liver and blood (~30% injected dose) and to a lesser extent, kidney (~9%), intestine (~9%), spleen (~3%), lung (~1%) and heart (~1%) (figure 3.2 B). The fairly low In₂Tf contents in the brain (220±11ng/g) and CSF (92±28ng/g) may be attributed to the fact that human Tf was
used in all experiments instead of rat Tf. Importantly, with the high sensitivity of the method, In can still be detected in the CSF of dosed rats. However, a better quantitation limit (LOQ) is still needed in order to confidently quantitate human Tf in rat CSF and to facilitate in-vivo studies of physiological barrier crossing and determine the physiological fate of an injected Lz-Tf conjugate. The reported ICP-MS instrumental detection limit of In is at sub-ppt levels, which equates to sub-ppb levels of Tf. Theoretically this represents a substantial improvement over the currently used procedure based on $^{18}$O labeling that relies upon a His-tag containing recombinant form of Tf. Our method LOQ for In, however, currently reaches the sub-ppb level but efforts to further improve this are underway.

3.3.2. Distribution of Lz-Tf at tissue level by imaging LA-ICP-MS.

As a powerful imaging technique, LA-ICP-MS has recently been used to produce distribution images of elements in thin human or rodent tissue sections. In order to assess distribution of In$_2$Lz-Tf in organs of interest we attempted to apply this technique for mapping the In$_2$Lz-Tf at the tissue level. In this work, organs were cryosected into thin layers of 50 µM thickness and mounted on glass slides. LA-ICP-MS was operated in line scanning ablation mode (line by line) with all the optimized experimental parameters provided in Table 3.1. The quadrupole mass filter was set to simultaneously measured signals from $^{115}$In, $^{57}$Fe and $^{63}$Cu (iron and copper are two ubiquitous endogenous metals).

The metal distribution images in the spleen and brain section are shown in Figure 3.3 and Figure 3.4. Panel A in both figures show the micro-photographs of the ablated
slices illustrating the location of histological features in the samples. Panels B, C and D show the distribution images of $^{57}$Fe, $^{63}$Cu and $^{115}$In respectively. Although the resolution achieved (100µm x 7µm pixel) was adequate for the differentiation of the distribution of elements in specific histological areas, it could still be improved more by reducing spot size and scan speed. As expected, there were some similarities between the distribution of Fe and In whereas Cu shows a more homogenous distributed.

In spleen, a clear accumulation of Fe and In in the red pulp could be observed (Figure 3.4 B and D). An explanation for this observation may be the fact that red pulp area is where all the blood filtration units, the sinusoids are located. Blood cells as well as bloodstream clutter particles, i.e. nuclear remnants, platelets, and denatured hemoglobin are mechanically filtered here as they enter the spleen. It is worth noting that red pulp also functions as the red blood cell reservoir. The high blood circulation in this region explains the high abundance of Fe and In as they are components of plasma proteins. On the other hand, the white pulp which closely resembles the structure of a lymph node and where antibodies are synthesized is organized with T- and B-cell compartments. Here, antibody-coated bacteria and antibody-coated blood cells are removed via blood and lymph node circulation. That explains the low abundance of Fe and In in these areas.

Similar to what was observed from homogenized brain measurements, the low abundance of In showed in LA-ICP-MS images from dosed rats’ brain (figure 3.4) was most likely due to the use of human Tf in all experiments instead of rat Tf. Despite this caveat, the signal from In was still distinguished from the tissue background in
reconstructed image which shows similar histological features to the photograph of brain section. Based on the data acquired from the measurements of homogenized thin brain cross-sections, in content was calculated to have an average of $5 \times 10^{-15} \text{g}$ per ablated pixel ($700 \mu m^2$) assuming it is homogenously distributed. This example not only demonstrates the outstanding capabilities of LA-ICP-MS for mapping metal coded protein drugs at cellular level (less than $100 \mu m$) but also provides us with a powerful tool to monitor the successful delivery of cancer-targeting Tf conjugates to animal cancer models.

3.4. Conclusions

Although attempts to study the biodistribution of exogenous Tf and Tf-based therapeutics in animal models (rats) using a proteomics-inspired $^{18}$O-labeling technique made the detection of Tf in blood possible, it was not sufficient for studies of physiological barriers crossing (e.g., exogenous Tf could not be detected in cerebrospinal fluid). A new sensitive and selective ICP-MS based method using non-cognate metal tracers was applied for the quantitation of exogenous Tf and Tf-based drugs in biological tissues and fluids. Our in-vivo studies tracked the injection of In$_2$Tf and In$_2$Tf-Lz by quantitating its localization to key organs. We found that Lz-Tf had accumulated largely in spleen ($15.2 \pm 3.2 \mu g/g$), liver ($14.1 \pm 3.5 \mu g/g$), kidney ($14.4 \pm 3.7 \mu g/g$), intestine ($4.6 \pm 1.6 \mu g/g$) and blood ($4.7 \pm 1.9 \mu g/g$), and to a lesser extent in the lungs ($3.4 \pm 0.4 \mu g/g$), heart ($2.1 \pm 0.4 \mu g/g$) and brain ($220 \pm 11 \text{ng/g}$). Although the used of human Tf in all experiments instead of rat Tf may have resulted in low Lz-Tf contents in CSF ($92 \pm 28 \text{ng/g}$), it can still be detected thanks to the high sensitivity of the method. However, better quantitation
limit (LOQ) is still needed. LA ICP-MS in cryosections of organs of interest was also assessed. We have demonstrated the imaging capabilities of LA-ICP-MS for studying biodistribution of metal coded protein drugs at cellular level. This is also a promising tool for the tracking of cancer-targeting Tf conjugates delivery to tumor cells.
**Table 3.1. Operating parameters of LA-ICP-MS**

<table>
<thead>
<tr>
<th>ICP-MS parameters</th>
<th>Laser Ablation parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nebulizer Gas Flow</td>
<td>Spot Size</td>
</tr>
<tr>
<td>0.9-1 L/min</td>
<td>100 µm</td>
</tr>
<tr>
<td>Auxiliary Gas Flow</td>
<td>Scan rate</td>
</tr>
<tr>
<td>1.4 L/min</td>
<td>20 µm/sec</td>
</tr>
<tr>
<td>Plasma Gas Flow</td>
<td>Laser energy</td>
</tr>
<tr>
<td>16.5 L/min</td>
<td>3.34 J</td>
</tr>
<tr>
<td>ICP RF power</td>
<td>Frequency</td>
</tr>
<tr>
<td>1600 W</td>
<td>10 Hz</td>
</tr>
<tr>
<td>Analog Stage Voltage</td>
<td>Shutter delay</td>
</tr>
<tr>
<td>-1600 V</td>
<td>10 sec</td>
</tr>
<tr>
<td>Pulse Stage Voltage</td>
<td>Gas flow rate</td>
</tr>
<tr>
<td>950 V</td>
<td>600 mL/min</td>
</tr>
<tr>
<td>Deflector Voltage</td>
<td>Space between lines</td>
</tr>
<tr>
<td>-12 V</td>
<td>0 µm (no space)</td>
</tr>
<tr>
<td>Dwell time</td>
<td></td>
</tr>
<tr>
<td>50 ms</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.2. Schematic of the ICP-MS based method for tracking Tf-based therapeutics distribution in organs and fluids.
Figure 3.3. Organ distribution of 115In over a 24 hours period after intravenous injection (A). Indium contents of some key organs are normalized to % injected dose (B). Each result represents the mean ± 1 SD from three rats.
Figure 3.4. LA-ICP-MS images for the simultaneous monitoring of $^{57}$Fe, $^{63}$Cu, and $^{115}$In on spleen sections from rats treated with a single dose of In$_2$Lz-Tf. Optical image of parallel sections for comparison purposes.
Figure 3.5. LA-ICP-MS images for the simultaneous monitoring of $^{57}$Fe, $^{63}$Cu, and $^{115}$In on brain sections from rats treated with a single dose of In$_2$Lz-Tf. Optical image of parallel sections for comparison purposes.
CHAPTER 4

STUDY OF IN-VIVO AGGREGATION OF TRANSFERRIN-BASED DRUG BY SIZE EXCLUSION CHROMATOGRAPHY AND INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY

Son N. Nguyen and Igor A. Kaltashov

Department of Chemistry, University of Massachusetts Amherst, Amherst, MA

4.1. Introduction

Previously, lysozyme-transferrin conjugates (Tf-Lz) were developed as a model CNS-targeting drug. Tf-Lz was successfully synthesized and well characterized including stability testing.\textsuperscript{48} In-vitro studies have indicated that although the Tf-Lz conjugate was produced as a monomer, it has the propensity to aggregate. Protein aggregation is a process that impairs protein activity both directly and indirectly (e.g., by reducing solubility).\textsuperscript{148} This phenomenon is of serious concern for protein drugs not just because it economically affects the production process but it also raises safety issues.\textsuperscript{149} However, all studies to date have been only done in-vitro. It is obviously a serious concern that Lz-Tf conjugates might aggregate even more when injected into live animal. Thus, it is important to study the aggregation of Tf-based drugs in-vivo in order

\textsuperscript{48} address correspondence to:
Igor A. Kaltashov
Department of Chemistry
University of Massachusetts-Amherst
710 North Pleasant Street, LGRT 104
Amherst, MA 01003
Tel: (413) 545-1460
Fax: (413) 545-4490
Email: Kaltashov@chem.umass.edu
to assess this propensity. However worth noting that, with the existing tools, it is extremely challenging to track the aggregation process *in-vitro*, let alone *in-vivo* study.

The most commonly used methods for the detection of soluble protein aggregates include size exclusion chromatography (SEC), Capillary Electrophoresis (CE), analytical ultracentrifugation (AUC), light scattering techniques (MALLS, DLS), field flow fractionation (FFF) and electrospray ionization mass spectrometry (ESI-MS). Among those techniques, SEC is undoubtedly the most widely used for characterization of aggregate size and content thanks to its ease of use, low cost, and high throughput. However, none of the aforementioned methods have the ability to detect protein aggregation *in-vivo* due to the complexity and high abundance of background proteins.

In order to achieve this goal, our lab has developed a new method that uses metal tracers and inductively coupled plasma (ICP) MS with the help of SEC for *in-vivo* quantitation of Tf-based drugs aggregation post injection. This method relies on SEC to separate different forms of In₂Tf-Lz conjugates (oligomers (aggregates) or monomers) and uses ICP-MS to quantitate the concentration of each form via the In concentration. ICP MS is a very powerful technique for elemental analysis, especially for measuring metal- and metalloid-containing compounds. In the last several years, ICP-MS has been applied to many different areas, particularly bioanalysis. Additionally, ICP-MS-based methods are not affected by the abundance background proteins of biological samples (i.e. body fluids) thanks to their high degree of selectivity toward metals, thus, no purification and preconcentration steps are required. Our previous studies have proved that Indium (In), a non-endogenous metal can be used as the Tf-based drugs
metal tracer since it binds tightly to Tf forming a stable In$_2$Tf complex. It is clear that this ICP MS based method using metal tracers enables protein quantitation, however, it does not provide any information about protein aggregation. Fortunately, it can be used as a very powerful protein detector for separation techniques that are capable of characterizing protein aggregation such as SEC and CE. In this work, we use the combination of SEC and ICP-MS for in-vivo quantitation of Tf-based drugs aggregation post injection.

4.2. Materials and methods

4.2.1. Preparation of indium (In) loaded Lz-Tf conjugate

Tf-Lz conjugate was synthesized using the protocol that was described previously. Indium (In) was then loaded to Tf-Lz following the removal of iron as also described previously.

4.2.2. Animal Models

Experiments were carried at the animal facilities of the University of Massachusetts, Amherst on 10-week-old, male Wistar rats purchased from Charles River Laboratories (Wilmington, MA, USA). Rats were housed in controlled light (12 h light/dark cycle), temperature and humidity conditions with free access to food and water. All animal experiments complied with the National Institute of Health Guidelines for Animal Care and were approved by the University of Massachusetts, Amherst Institutional Animal Care and Use Committee.
Animals were divided into two groups: G1: control rats (i.v. injected with 100mM sterilized PBS solution); G2: Tf-Lz rats (i.v. injected with 5mg/kg dose of In2Tf-Lz). All In2Tf-Lz samples were buffer exchanged into 100mM sterilized PBS buffer solution prior to injections. Rats were euthanized by CO2 overdose 24 h after injection. Blood samples were obtained in heparin coated tube to prevent coagulation and divided into aliquots of 200µL each, stored in the -80°C freezer. Rat plasma was prepared by centrifuging blood from control rats immediately after collection for a minimum of 10 minutes at 1000-2000 RCF (generally 1300 RCF) at room temperature. Supernatant (plasma) was then carefully aspirated, aliquoted into vials (200µL) and stored at -80°C. The experimental sequence is presented in Figure 4.1. Storage conditions experiments are shown in the bold black boxes; in-vitro, ex-vivo and in-vivo studies showed in bold red, blue and green boxes respectively.

4.2.3. Instrumentation.

An Agilent 1100 series liquid chromatography (Agilent Technologies, Santa Clara, CA) equipped with a TSK Gel 3000 SWXL column (TOSOH Bioscience, King of Prussia, PA) was used for SEC analysis. The separation was performed with 150 mM ammonium acetate (pH 7.0) at a flow rate of 0.8 mL/min and was monitored with 280 nm UV detection. The injected sample volume was 20 µL for storage condition tests (5µM Lz-Tf sample) and 10µL for all other samples.

All ESI MS measurements were carried out with a QStar-XL (ABI/SCIEX, Toronto, Canada) hybrid quadrupole/time-of-flight MS equipped with a nanospray source. The
position of the ESI emitter and spray voltage were carefully adjusted to optimize the spray efficiency, which directly affected the resolution of the mass spectra.

Three SEC fractions collected F1 (5.5 min - 10.6 min), F2 (10.8 min – 13 min) and F3 (13.2 min – 17 min) were oxidized in 0.2 mL of acid mixture (67% HNO₃ and 30% H₂O₂ at a 3:1 volume ratio) and shaken for 1 hour at 50°C. After the addition of 1 ppb Rh as internal standard, deionized water was added to bring the final volume to 5.0 mL. Indium content was analysis using a NexION 300X ICP-MS (Perkin Elmer, Waltham, MA, USA). Prior to the analysis, daily tuning was done to make sure the instrument was in optimum conditions. All signals were obtained using the collision cell with kinetic energy discrimination mode (KED). Operating conditions for the experiment are illustrated in Table 4.1. A calibration curve of various signal ratio of In/Rh was used to measure the In content, using 2.5% HNO₃ and 0.37% H₂O₂ as the blank. All ICP-MS data were reported as a fraction of total In. Error bars indicate 1 standard deviation from triplicate measurements. Post-acquisition analysis of both SEC and MS data was carried out using Origin 8.0 (OriginLab Corporation, Northampton, MA) software.

4.3. Results and discussion

4.3.1. Aggregation study of In₂Tf-Lz under typical storage conditions

In order to assess the stability (aggregation and degradation) of the Tf-Lz conjugate as a result of storage in typical conditions (e.g. in 4°C fridge or -20°C freezer), SEC, ESI-MS and ICP-MS methods were applied. Freshly prepared In₂Tf-Lz conjugate was divided into 50 µL aliquots of 5 µM protein concentration and stored either at 4°C or lyophilized
and stored in a -20°C freezer. After one month, the stored samples were reconstituted in 150 mM NH₄Ac solution, pH 7.0 and analyzed by SEC, native ESI-MS and ICP-MS. Figure 4.1 shows the experimental protocols for the test of Tf-Lz conjugate stability. Figure 4.2 shows data acquired from SEC and native ESI-MS; Panel A indicates SEC chromatograms of the freshly prepared samples along with 1 month stored samples either at 4°C or -20°C. The early elution fraction (F1) corresponds to a high molecular weight fraction being oligomers of Tf-Lz i.e. aggregation. The second fraction (F2) shows the Tf-Lz monomer peak and lastly, F3 shows the small molecular weight fraction which is where Tf-Lz degradation products eluted; Panel B shows native ESI-MS data of the corresponding unfractionated samples. Degradation products of Tf-Lz as well as unfolded proteins and those with compromised (partially reduced or scrambled) disulfide networks will appear at low m/z region (1000-3000) while oligomers of Tf-Lz shows signals at high m/z region (above 5000). All samples had the same concentration in the same 150 mM NH₄Ac buffer, pH 7.0 for all SEC measurements and were diluted to a final concentration of 1 µM in 150 mM NH₄Ac, pH 7.0 for native Nanospray-MS analysis.

The SEC chromatogram of Tf-Lz sample stored at -20°C shows that there was not much change in the F1 region compared to that of freshly prepared Tf-Lz. Sample stored at 4°C, however, showed increasing of UV absorption in this fraction indicating aggregation has occurred after 1 month in storage in the fridge and is ~ less than 5% compared to the monomer peak shown in F2 region. Thus it appears that, almost no hydrolysis products were observed in any of the samples indicating good stability of
both proteins and chemical linkers. Native ESI-MS data of the corresponding samples are in good agreement with the SEC data. No noticeable Tf-Lz degradation products were found although low signal of Tf-Lz oligomers were observed in all samples. Furthermore, ICP-MS measurements of In content of 3 collected fractions from SEC runs indicate a low abundance of Tf-Lz aggregation species (Figure 4.4). Interestingly, there were noticeable In signals from F3 fraction of both freshly prepared Tf-Lz conjugate and the 1 month-old sample indicating hydrolysis products. However, a closer look at the results, indicated that those In signals were most likely coming from the tailing effect of monomer peak in fraction F2 due to the high Tf-Lz concentration used for these studies. The good agreement among the three orthogonal methods namely SEC, native ESI-MS and ICP-MS indicates that Tf-Lz conjugate is stable under all typical storage conditions.

4.3.2. In-vitro, Ex-vivo and In Vivo aggregation study of In\_2Tf-Lz

To further our understanding of the Tf-Lz aggregation process, in-vitro, ex-vivo and in-vivo stability studies were performed using the method developed. For in-vitro and ex-vivo studies, lyophilized In\_2Tf-Lz sample was first reconstituted into 50 µL of 150 mM NH\_4Ac, pH 7.0 to make roughly 5 µM solution of which 5 µL was either mixed with 45 µL of 150 mM NH\_4Ac, pH 7.0 (in-vitro) or 45µL rat plasma (ex-vivo), followed by their 24h incubation at 37°C. Each experiment was done in triplicate. The incubated samples along with rat blood samples were then injected (10 µL) onto a SEC column and three fractions were collected for each sample. Figure 4.3 illustrated SEC chromatograms of In\_2Tf-Lz kept under different conditions including freshly prepared (black trace, control sample), 24h incubation in NH\_4Ac at 37°C (red trace, in-vitro), 24h incubation in rat
plasma at 37°C (blue trace, *ex-vivo*) and blood of injected rats (green trace, *in-vivo*). All traces were normalized to their most abundant peaks. SEC data showed that the aggregation and hydrolysis of Tf-Lz conjugate after incubation in NH₄Ac for 24h at 37°C is not significant compared to that of fresh Tf-Lz. Indeed, ICP-MS measurement indicates similar result (Figure 4.4).

As expected, SEC chromatograms of rat plasma and whole blood do not give us any useful information due to the high abundance of background proteins. However, three fractions were still collected based on their elution times as defined earlier (F1 (5.5min-10.6min), F2 (10.8min-13min) and F3 (13.2min-17min)). Figure 4.4 shows In content of SEC fractions measured by ICP-MS. Data were reported as fraction (percentage) of total In concentration measured from all three collected SEC fractions. It is worth noting that, ICP-MS quantitation of total In concentration in blood circulation indicated that Tf-Lz conjugate concentration is 4.7±1.9µg/mL which is 10 times lower than in-vitro studies. Although background proteins are present at high abundance and blood samples were diluted roughly 500 times due to the collecting process, In contents in all fractions were confidently quantitated with high accuracy and precision. Importantly, the ICP-MS results of *in-vivo* studies indicated that there were up to 15% In signal in F1 fraction. Beside protein aggregation, there might be a possibility of conjugate bound antibody and/or conjugate bound soluble Tf receptor (sTfR) in this fraction. This is still an encouraging result since blood and plasma contain a number of proteases and other enzymes as well as scavenger proteins that may affect and modify the structure of Tf-Lz. Even 24 hours post-injection over 85% of Tf-Lz conjugate was still in its monomeric
form. Importantly, this ICP-MS based method can be coupled as a detector online to many separation and/or fractionation techniques such as capillary electrophoresis (CZE), SEC or field flow fraction (FIFFF) to take full advantage of this powerful method of characterizing the aggregation size, forms and make distinction of protein aggregation from conjugate bound antibody and/or conjugate bound soluble Tf receptor (sTfR) as well as achieving high throughput measurements. Others have attempted to study the aggregation of biopharmaceuticals when mixed with human plasma using methods like light microscopy, Fluorescence Single Particle Tracking (fSPT)$^{154}$, light-scattering (DLS) and surface plasmon resonance (SPR)$^{155}$. However, this is the first time in-vivo quantitation of a small soluble protein drug aggregation and degradation post-injection was studied.
4.4. Conclusions

Protein aggregation is a phenomenon that is a serious concern for protein drugs regarding safety issues\textsuperscript{149} Traditional methods do not have the ability to detect protein aggregation due to the complexity and high abundance of background proteins, even with the help of purification and preconcentration.

This study has demonstrated that an ICP-MS based method using metal tracer has the ability to probe the protein stability post-injection and yields useful data not accessible by other methods. We found that, whilst the Tf-Lz conjugate appears to be stable under typical storage conditions as well as in the \textit{in-vitro} study i.e. it does not show significant aggregation and degradation products, \textit{ex-vivo} and \textit{in-vivo} studies indicate a noticeable aggregation formed. These results emphasize the importance of analytical methods for \textit{in-vivo} quantitation of aggregation in biological fluids since aggregation profile may drastically change post-injection.
**TABLES**

**Table 4.1.** Operating parameters of ICP-MS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nebulizer Gas Flow</td>
<td>0.9-1 L/min</td>
</tr>
<tr>
<td>Auxiliary Gas Flow</td>
<td>1.4 L/min</td>
</tr>
<tr>
<td>Plasma Gas Flow</td>
<td>16.5 L/min</td>
</tr>
<tr>
<td>ICP RF power</td>
<td>1600 W</td>
</tr>
<tr>
<td>Analog Stage Voltage</td>
<td>-1600 V</td>
</tr>
<tr>
<td>Pulse Stage Voltage</td>
<td>950 V</td>
</tr>
<tr>
<td>Deflector Voltage</td>
<td>-12 V</td>
</tr>
<tr>
<td>Dwell time</td>
<td>50 ms</td>
</tr>
</tbody>
</table>
Figure 4.1. Experimental scheme of the Tf-Lz conjugate stability test
Figure 4.2. SEC chromatograms and native ESI-MS spectra of Tf-Lz under storage conditions; (A) SEC chromatograms of the freshly prepared samples along with 1 month stored samples either at 4°C or -20°C. The early elution fraction (F1) corresponds to high molecular weight fraction being oligomers of Tf-Lz i.e. aggregation. The second fraction (F2) shows Tf-Lz monomer peak and lastly, F3 shows the small molecular weight fraction which is where Tf-Lz degradation products eluted out; (B) native ESI-MS spectra of corresponding samples. Degradation products of Tf-Lz, if any, will appear at low m/z region (1000-3000) while oligomers of Tf-Lz shows signals at high m/z region (above 5000). All samples had the same concentration in the same buffer of 150mM NH4Ac, pH 7.0 for all SEC measurements and were diluted to 1µM solution with 150mM NH4Ac, pH 7.0 for native Nanospray-MS analysis.
Figure 4.3. SEC chromatograms of In2Tf-Lz in different conditions including freshly prepared (black trace, control sample), 24h incubation in NH₄Ac at 37°C (red trace, in-vitro), 24h incubation in rat plasma at 37°C (blue trace, ex-vivo) and bloods of
Figure 4.4. In contents of SEC collected fractions measured by ICP-MS. Data were reported as fraction (percentage) of total In concentration.
CHAPTER 5
A NEW $^{18}$O LABELING METHOD FOR MODIFICATION SITES QUANTITATION OF
PROTEIN-DRUGS CONJUGATES

Son N. Nguyen, Alex J. Johnson and Igor A. Kaltashov

Department of Chemistry, University of Massachusetts Amherst, Amherst, MA

5.1. Introduction

Protein-drug conjugates have emerged as a very promising approach for cancer treatment. They combine the best features of both components, i.e. specificity of carrier proteins and the potency of cytotoxic drugs. Usually, carrier proteins, such as monoclonal antibodies (mAbs) or transferrin (Tf), are linked to cell-killing drugs via various chemical linkers which are crucial for the successful delivery. It is well known that conjugation chemistry, particularly lysine-based methods, introduces heterogeneity and modifies the surface charges of the carrier proteins. In spite of recent advances, characterization of intrinsic heterogeneity that may affect the pharmacokinetics, tissue

---

address correspondence to:
Igor A. Kaltashov
Department of Chemistry
University of Massachusetts-Amherst
710 North Pleasant Street, LGRT 104
Amherst, MA 01003
Tel: (413) 545-1460
Fax: (413) 545-4490
Email: Kaltashov@chem.umass.edu
distribution\textsuperscript{156} and the physicochemical stability of the conjugates\textsuperscript{157} still remains challenging.

The site of conjugation is reported to have a significant impact on the stability (i.e. aggregation, poor solubility and low stability in circulation) and bioactivity of the conjugates.\textsuperscript{158} Ideally, the correlation of linker position to conjugate activity will allow the optimization in design of the protein conjugate.\textsuperscript{159} It is worth noting that the heterogeneity of conjugation sites is usually the most difficult to characterize.\textsuperscript{160} This type of structural heterogeneity can be characterized using methods of tandem mass spectrometry (MS/MS) with the help of enzymatic degradation of the conjugated protein (peptide mapping) or top-down MS/MS, a successful approach that have been applied to characterize the distribution of conjugation sites in PEGylated proteins\textsuperscript{132} and protein-small molecule drug conjugates.\textsuperscript{133}

Peptide mapping using chromatographic methods combined with MS analysis is a universal tool for the detection and characterization of a wide variety of protein modifications including PEGylation, oxidation, glycation, disulfide scrambling, and crosslinking.\textsuperscript{161} More importantly, it can provide structural information about the locations of conjugation sites. It is well known that trypsin cannot cleave at modified lysines, hence, larger and more hydrophobic peptides are generated. Moreover, modified peptides usually elute later than unmodified ones in reverse phase HPLC due to the changes in surface charge and increase in hydrophobicity. Therefore, enzymatic digests’ chromatograms of a protein drug conjugate are considerably different from those of the unmodified protein thus helping identify specific modification sites of drug.
Using this approach, modification sites of lysine-modified maytansinoid-monoclonal antibody immunoconjugate, huN901-DM1, were mapped using both trypsin and Asp-N as proteases. It is reported that all observed sites are partially modified and mostly found in locations that are more exposed to solvent as well as those that have structural flexibility.

Despite large number of studies published on this topic, most set out only to identify the modification sites of protein drug conjugates, very few attempted to quantitate them. However, quantitation of modification sites is commonly used in chemical probing methods for characterization of protein structure, function, and ligand binding. In these studies, modified lysine sites were quantitated either using extracted ion chromatogram (XIC) estimation or iTRAQ (isobaric tags for relative and absolute quantification) method. As an example, a chromogenic chemical probe targeting lysines, termed “NN”, was used to study the surface accessibilities/reactivities of lysozyme. Peak areas of all modified and unmodified peptides identified in the digests from the XIC were used to calculate the reactivity of each lysine residue (epsilon-amine) and the N-terminus. Although the XIC-based method is a simple method, it suffers from a fundamental error i.e. modified peptides have different ionization efficiencies compared to unmodified ones. On the other hand, iTRAQ is a very powerful method with multiplexing capability for relative and absolute quantitation of peptides and proteins. However, iTRAQ quantitation requires commercial isobaric tags and is only possible with tandem MS analysis, thus making it an expensive approach.
Herein, we describe a simple and cost-effective $^{18}$O labeling-based method for quantitation of lysine modification sites of protein drug conjugates. The principle of this method is based on the fact that trypsin cannot cleave at modified lysines, thus the modification percentage at those sites can be calculated by comparing the amount of modified proteins tryptic peptides containing lysine to those of unmodified proteins. Enzyme catalyzed $^{18}$O labeling is a simple technique incorporating a maximum of two $^{18}$O atoms at the C-terminal carboxyl group of proteolytic peptides in the presence of $H_2^{18}$O. This results in a mass shift of 4 Da between $^{18}$O/$^{16}$O labeled peptides, thus, enabling peptides quantitation. The $^{18}$O based method has proved to be a useful tool not only for protein quantitation in serum and tissues but also for quantitating changes of protein phosphorylation (post-translational modification). Despite the simplicity and low costs, incomplete labeling, mostly due to the different enzyme substrate specificity, oxygen back exchange, pH and peptide physicochemical properties (e.g. size, charge, hydrophobicity...) make this method technically challenging. However, the problems can be overcome either by using correction algorithms or by minimizing back exchange i.e. decreasing the pH, using heat deactivated trypsin or using immobilized trypsin.

5.2. Materials and methods

5.2.1. Chemicals and Reagents.

$H_2^{18}$O (97% purity) from Cambridge Isotope Laboratories (Andover, MA), iodoacetic acid (IAA) and mass spectrometry-grade trypsin were purchased from Sigma-Aldrich
Chemical Co. (St. Louis, MO, USA). All other chemicals and solvents used in this work were of analytical grade or higher.

5.2.2. Preparation of SATA-Lysozyme conjugate

Lysozyme (Lz) from chicken egg white (Sigma-Aldrich, St. Louis, MO) was activated by N-succinimidyl-S-acetyltioacetate (SATA; Pierce Biotechnology, Rockford, IL) as described in Nguyen et.al.,48. The reaction was carried out by incubating 200 µL consisted of 50 µM Lz, 200 µM of SATA (freshly prepared in DMSO) in 50 mM phosphate buffer with 100 mM NaCl pH 8.0 for 2h at 37°C. The control reaction was done in parallel under the same conditions with DMSO added instead of SATA. Both control Lz and SATA-Lz conjugate were purified by size exclusion chromatography on a TSK Gel 3000 SWXL column (TOSOH Bioscience, King of Prussia, PA) using an Agilent 1100 (Agilent Technologies, Palo Alto, CA) HPLC system. The separation was performed in 150 mM ammonium acetate (pH 7.0) at a flow rate of 0.8 mL/min and was monitored with 280 nm UV detection. Protein concentrations were measured using a NanoDrop 2000c (Thermo Fisher Scientific, Rockford, IL) UV-Vis spectrophotometer.

5.2.3. Protein Digestion and Post-digestion Labeling

Control Lz samples and SATA-Lz conjugate (10µg) were reduced in 10 mM DTT, 15% acetonitrile (ACN) in 100 mM ammonium bicarbonate, pH 8.0 for 1h at 60°C followed by alkylation with 20 mM iodoacetic acid (IAA) for 45 min in the dark. Digestion was carried out overnight at 37°C with sequence grade trypsin (1:30 protease/protein). The digested samples were dried using the N₂ purging method and then re-suspended in 50 µL of either H₂¹⁶O or H₂¹⁸O containing 10% DMSO, 0.3 µg trypsin and incubated at 37°C for 48
hours. Samples were removed, vortexed, spun down and formic acid was added immediately to make a final concentration of 1% (v/v) to quench trypsin activity. LC MS/MS measurements were performed immediately and any remaining samples were stored at -20°C.

5.2.4. Instrumentation.

Ion exchange chromatography of SATA-Lz conjugate was performed on a 4.6 x 100 mm PolyCATA™ weak cation exchange column (5 μM, 300 Å, PolyLC Inc., Columbia, MD). The $^{16}\text{O}/^{18}\text{O}$ labeled mixtures of digested samples were subjected to LC/MS/MS analysis on a PepMap100 75μm i.d. x 15cm capillary column using an LC Packings Ultimate (Dionex/Thermo Fisher Scientific, Sunnyvale, CA, USA) nano-HPLC system coupled with a Qstar-XL (AB SCIEX, Toronto, Canada) hybrid quadrupole/TOF MS. Steps gradient. Post-acquisition analysis of both SEC and MS data was carried out using Origin 8.0 (OriginLab Corporation, Northampton, MA) software.

5.3. Results and discussion

5.3.1. $^{18}\text{O}$ method for quantitation of modification sites

This $^{18}\text{O}$ based method is first demonstrated for SATA-Lysozyme (Lz) conjugate. Lz was modified with SATA using a 4:1 SATA:Lz molar ratio. There are 7 possible modification sites in Lz including six lysine residues (K1, K13, K33, K96, K97 and K116) and the N-terminus. ESI-MS of the SATA-Lz conjugate (+10 charge state) showed in Figure 5.1 indicates that the most abundant product is a 1:1 conjugate, and up to three SATA had been conjugated to Lz. The conjugate and control Lz were then subjected to trypsin digestion followed by trypsin catalytic $^{18}\text{O}$ labeling process as described in the
The LC MS/MS measurements were carried out immediately after mixing $^{16}$O labeled samples being unmodified or modified Lz with $^{18}$O labeled control Lz as internal standard. Figure 5.2 shows all the tryptic peptides that are monitored in order to quantitate lysine modification sites. The peptides coded $CT_x$ are control peptides which neither include a lysine residue nor immediately follow a lysine residue in the protein sequence. Whereas, $K_x$ peptides ($x$ is the location of lysine residue in Lz sequence) are peptides directly related to the modification at location “$x$”. The modification percentages at those lysine residues are proportional to the depletion of corresponding peptides ($K_x$). Figure 5.3 indicates the $^{18}$O labeled internal standard strategy used to quantitate modification sites. Panel (A) shows the peptide MS1 spectra of either $^{16}$O (gray trace) or $^{18}$O (black trace) labeled control Lz before and after 1:1 mixing (control mix). Since both samples are unmodified Lz, we expect to have the ratio of $^{16}$O species ($I_0$) to $^{18}$O species ($I_2$ and $I_4$) approximately 1.0. However, when mixing 1 to 1 ratio of $^{16}$O labeled modified Lz and $^{18}$O labeled control Lz, the $^{16}$O/$^{18}$O ratio will be less than 1 if there was a modification at adjacent lysine site (Panel B). The $^{16}$O/$^{18}$O ratio was calculated using the following equation adopted from Yao et. al.,$^{177}$

$$\text{ratio}(\frac{^{16}O}{^{18}O}) = I_0 - \frac{M_4}{M_0}I_0 - \frac{M_2}{M_0}I_2 - \left(\frac{M_2}{M_0}\right)I_0$$

(1)

$M_0$, $M_2$, and $M_4$ are the relative intensities for the first, third and fifth peak in theoretical isotope pattern of peptides which can be calculated based on peptide
sequences using Analyst QS calculator tool; whereas \( I_0, I_2, \) and \( I_4 \) are the observed intensities from the \(^{16}\text{O}/^{18}\text{O}\) peptides mixture.

Additionally, calibration plots were used in order to accurately quantitate modification percentages. Standard samples were prepared by mixing \(^{16}\text{O}\) and \(^{18}\text{O}\) labeled digested control Lz at various ratios being 1:1, 0.75:1, 0.5:1, 0.25:1, 0:1 \((^{16}\text{O}:^{18}\text{O})\). Figure 5.4 shows calibration plots of monitored peptides. Although all of them have good linearity; different peptides show different slopes and intercepts with the y axis. The factors that could contribute to this issue include incomplete labeling, different enzyme substrate specificity, oxygen back exchange, pH and peptide physicochemical properties such as size, charge and hydrophobicity. After closely inspecting isotopic distributions of labeled peptides, we realized that the observed isotopic patterns are mainly due to the incomplete incorporation of \(^{18}\text{O}\) to a subset of peptides, not because of oxygen back-exchange. Subsequently, labeling conditions have been optimized in order to overcome this insufficient labeling issue. However we can only manage to minimize the process but not eliminate it. Others have tried to fix this issue. For example, Bezstarosti et. al., has reported a two-step \(^{18}\text{O}\) labeling using immobilized trypsin that helps to completely label \(^{18}\text{O}\) to peptides.\(^{175}\) Furthermore, it was found that the hydrophobicity of long peptides, such as K96 and K97, makes them unstable in aqueous solution (buffer A; 0.1% Formic acid) and renders peptides prone to either aggregation or interaction with and retaintion by the silica capillary tubing. Therefore, mixing the peptides in 10% DMSO or ACN and immediate measurements after mixing are required for more consistent results.
Selection of Reference Peptide. It is worth noting that lysines are among the most abundant amino acids, e.g. 58 lysines in transferrin and ~90 lysines in a typical monoclonal antibody; thus it is impractical to construct calibration plots and analyze all of them. In fact, only one good reference peptide is needed for this method. The criteria for selection of the reference peptide in this $^{18}$O based modification sites quantitation are stability, labeling completeness and signal intensity. In order to find the best reference, a group of ten peptides was investigated by comparing their $^{16}$O to $^{18}$O ratios calculated from triplicate equally mixed control samples. Finally, CT3 (WWCNDGDGR) was selected as a reference since it can be easily and stably detected at a very low concentration. More importantly, as shown in Figure 5.5, normalized $^{16}$O/$^{18}$O ratios of all peptides against the ratio of CT3, nearly all results were well within ±15% range of the calibrated value.

A linear calibration curve of reference peptide CT3 was plotted from 0% to 100% modification percentage using the measured $^{16}$O/$^{18}$O ratios versus the corresponding mixing ratios being 1:1, 0.75:1, 0.5:1, 0.25:1 and 0:1 ($^{16}$O:$^{18}$O) (Figure 5.6). Triplicate measurements were conducted for each data point, and the standard deviation was presented as the error bar. This calibration plot was then used to quantitate the modification percentage at different lysine sites. Figure 5.7 shows modification percentages at all six lysine sites along with measured control peptide percentages (gray bars) for method validation purposes. Blue bars correspond to SATA modified to Lz with the molar ratio of 4 to 1 (SATA:Lz), while Red bars correspond to 8:1 SATA:Lz molar ratio reaction. Triplicate measurements were conducted for each sample, and the standard
deviation is presented as the error bar. Importantly, the percentages of all control peptides were found in the range of the method error. Although two different peptides were used to quantitate modification at K33, they showed similar results with the error within ±15%. Moreover, the large error showed at K96 was due to instability observed for the peptide used to quantitate that site. Of all the modification sites quantitated, we found that the following order for the extent of modification: K97 ≈ K33 > K116 ≈ K13 > K1(ε-NH₂) ≈ K96. It is worth noting that, ¹⁸O method can only be used to quantitate modification of lysine side chain (ε-NH₂), it cannot measure the modification at the N-terminus. In order to correlate the relative reactivities of lysine SATA modification with structural parameters of lysozyme, quantitation results were compared to predicted surface accessibility (SA) and pKa values of lysozyme (4SEQ) (Table 5.1). It was found that reactivities correlate more with the SA values than pKa. The high modification percentage of K97 corresponds to its highest accessibility (53%) whereas K96 shows low reactivity and lowest SA value (26%).

5.3.2. Comparison of ¹⁸O method and XIC estimation method for modification sites quantitation

To further evaluate the validity of this method, the results were compared to those of the extracted ion chromatograms (XIC) estimation method. The reactivities of lysines were calculated using following equation:

\[
\text{Reactivity (\%)} = \frac{\sum \text{Area of all modified peptides containing lysine } X}{\sum \text{Area of all peptides containing lysine } X} \times 100
\]  \hspace{1cm} (2)
While the XIC method is simple and very easy to use, it suffers from fundamental errors. First of all, SATA replaces positively charge group of primary amine by its neutral molecule, thus, modified peptides tend to have less charge than unmodified one. Figure 5.8 shows typical problems of XIC method when used for quantitation of modification sites. Panel (A) shows +2 charge state of modified and unmodified K97 peptides with the modification percentage calculated to be ~84%. However, reactivity calculated for the same peptides at +3 charge state was only ~11% (Figure 5.8 B). This problem would be less significant if the linker had a charge group similar to primary amine it replaced. Second, very long peptides shown in Figure 5.8 D tend to have much lower intensity than shorter ones. Therefore, if equation (2) is used, the calculated value is not accurate. It is noticed that this problem only occurs with long peptides so a combination of two or three enzymes would make peptides shorter, thus, the problem could be solved. Despite all those problems, XIC method is still a quick, useful tool for semi-quantitating modification sites of protein drug conjugate.

Figure 5.9 shows results from two methods. Panel A shows data from 4:1 SATA:Lz modification while Panel B shows data from 8:1 SATA:Lz modification. Blue bars show results from $^{18}$O method while red bars indicate XIC results. The results from two methods appear to agree with each other except for the K33 where XIC method over estimates due to the calculation based solely on the long peptides. Additionally, XIC method shows results for both lysine-1 side chain and N-terminus modification whereas $^{18}$O only show reactivity for lysine side chain, thus higher percentage at K1 is expected in XIC method. Although N- terminus is partially shielded by Thr-40$^{178}$ and has low
accessibility, moderate reactivity at this site can be explained by its nucleophilicity (pKa\textasc/documents/9.0) and/or a possible conformational change upon modification of Lys-1. These results have proved that $^{18}$O labeling method is a universal tool and can be used for quantitating lysine based modification, despite the fact it lacks information about any N-terminus modification. This method can serve as a complementary method to XIC estimation for more accurate quantitating reactivities of lysine side chains.

5.4. Conclusions

Although heterogeneity of conjugation sites is usually the most difficult to characterize, we have demonstrated $^{18}$O labeling method is a simple and cost-effective way of quantitating lysine modification sites of protein drug conjugates. It quantitates the depletion of lysine-adjacent peptides based on the fact that trypsin cannot cleave at modified lysines. Among factors affecting the accuracy of this method, incomplete labeling is the major one that is very hard to eliminate. We found that reactivities correlate more with the SA values than pKa. The high modification percentage of K97 corresponds to its highest accessibility (53\%) whereas K96 shows low reactivity and lowest SA value (26\%).

A comparison between $^{18}$O and XIC method has also been carried out. While XIC method is simple and very easy to use, it suffers from a fundamental error namely differences in ionization efficiency of modified and unmodified peptides. The results from the two methods appear to agree with each other. Although the N-terminus was partially shielded by Thr-40$^{178}$ leading to low accessibility, moderate reactivity at this site can be explained by its nucleophilicity (pKa\textasc/documents/9.0) and/or a possible conformational
change upon modification of Lys-1. The results have proved that $^{18}$O labeling method can serve as an alternative method to XIC estimation or iTRAQ for accurate quantitation of modification sites of protein drug conjugates. Moreover, this method can also be applied for the quantitation of lysine site chain modification by chemical probes normally used to characterize protein structure, function, and ligand binding.
TABLES

**Table 5.1. Predicted Lysine Surface Accessibilities and pKa Values**

<table>
<thead>
<tr>
<th>Residue</th>
<th>Surface accessibilities, %</th>
<th>pKa</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminus</td>
<td>4.1</td>
<td>9.0</td>
</tr>
<tr>
<td>Lys-1</td>
<td>37.7</td>
<td>11.25</td>
</tr>
<tr>
<td>Lys-13</td>
<td>36.8</td>
<td>11.60</td>
</tr>
<tr>
<td>Lys-33</td>
<td>35.6</td>
<td>10.15</td>
</tr>
<tr>
<td>Lys-96</td>
<td>26.0</td>
<td>10.15</td>
</tr>
<tr>
<td>Lys-97</td>
<td>53.3</td>
<td>10.45</td>
</tr>
<tr>
<td>Lys-116</td>
<td>52.1</td>
<td>10.17</td>
</tr>
</tbody>
</table>

Surface accessibilities and pKa values for pdb structure 4QEQ were calculated using online software GETAREA and PROPKA 3.0.
Figure 5.1. ESI-MS of the SATA-Lz conjugate (+10 charge state) SATA:Lz (4:1) molar ratio.

Figure 5.2. Tryptic peptides that are monitored in order to quantitate lysine modification sites.
Figure 5.3. $^{18}$O labeled internal standard strategy used to quantitate modification sites. (A) peptide MS1 spectra of either $^{16}$O (gray trace) or $^{18}$O (black trace) labeled control Lz before and after 1:1 mixing (control mix). (B) $^{16}$O labeled modified Lz (gray trace) and $^{18}$O labeled control Lz (black trace) before and after 1:1 mixing (modified mix).
Figure 5.4. Calibration plots of monitored peptides.
Figure 5.5. Normalized $^{16}\text{O}/^{18}\text{O}$ ratios of all peptides against the ratio of CT3

Figure 5.6. A calibration curve of reference peptide CT3 was plotted using the measured $^{16}\text{O}/^{18}\text{O}$ ratios versus the corresponding mixing ratios.
Figure 5.7. Modification percentages at all six lysine sites along with measured control peptide percentages (gray bars) for method validation purposes. Blue bars correspond to SATA modified to Lz with the molar ratio of 4 to 1 (SATA:Lz), while Red bars correspond to 8:1 SATA:Lz molar ratio modification. Triplicate measurements were conducted for each sample, and the standard deviation was presented as the error bar.
Figure 5.8. Typical problems of XIC method when used for quantitating modification sites. (A) +2 charge state of modified and unmodified K97 peptides with the modification percentage calculated to be ~84%. (B) Reactivity calculated for the same peptides at +3 charge state was only ~11%. 
Figure 5.9. Results from $^{18}$O method and XIC method. (A) shows data from 4:1 SATA:Lz modification; (B) data from 8:1 SATA:Lz modification. Blue bars show results from $^{18}$O method while red bars indicate XIC results.
CHAPTER 6
SUMMARY AND FUTURE OUTLOOK

Mass spectrometry has already become an indispensable tool for facilitating all stages of modern medicine development and their characterization. It can be applied for a variety of tasks ranging from analyzing covalence structure, conformations and protein-receptor interaction to quantitative measuring biodistribution of protein therapeutics. In this work, we have demonstrated the potential of ESI MS in this field by providing invaluable information used for optimizing protein drug conjugate structures during early stages of development, and further catalyzing the drug design efforts. Importantly, this method can readily be applied to other protein drug conjugates development such as antibody drug conjugates (ADCs).

Additionally, a new method based on metal tracers and (ICP) MS was developed in our lab has been applied for ultrasensitive quantitation of exogenous Tf and Tf-based drugs in biological tissues and fluids. With this method, injected Tf and Tf-based drugs were successfully detected in the cerebrospinal fluid (CSF) in which proteomic $^{18}$O based method developed in out lab fails. Furthermore, laser ablation (LA) ICP-MS in cryosections of organs of interest was also assessed. LA-ICP-MS have been demonstrated its imaging capabilities for studying biodistribution of metal coded protein drugs at cellular level. This is also a promising tool that can be applied for the tracking of cancer-targeting Tf conjugates delivery to tumor cells.
Another application of ICP-MS based method was also assessed. When combine to separation methods such as SEC and CE, ICP-MS with the help of a metal tracer becomes a powerful detector for structural study of protein drugs. This method has been successfully applied to probe into protein stability post-injection and yields useful data not accessible by other methods. It is the first time a small soluble protein aggregation of injected protein drug was studied in live animals. It is speculated that the application of ICP-MS based method using metal tracer in combination with separation (SEC, CE) or fractionation (FlFFF) techniques will prove to be a useful tool for probing protein aggregation post-injection.

Finally, a simple and cost-effective $^{18}$O labeling-based method has been developed for quantitating lysine modification sites of protein drug conjugates and has been successfully applied for SATA-Lz conjugate. The results have proved that $^{18}$O labeling method can serve as an alternative method to XIC estimation or iTRAQ for accurate quantitation of modification sites of protein drug conjugates. Moreover, this method can also be applied for quantitating chemical probes modified lysine site chain in order to characterize protein structure, function, and ligand binding.
BIBLIOGRAPHY


78. Mendler, C. T.; Friedrich, L.; Schlapschy, M.; Schweiger, M.; Wester, H.-J.; Skerra, A. In High contrast tumor imaging with radio-labelled antibody Fab fragments tailored for optimized pharmacokinetics via PASylation, mAbs, 2014; Taylor & Francis: pp 00-00.


PEGylated dendrimers show differential patterns of deposition and activity in tumour-burdened lymph nodes after intravenous and subcutaneous administration in rats. *Molecular Pharmaceutics 2014.*


160. Abzalimov, R. R.; Frimpong, A.; Kaltashov, I. A. Structural characterization of protein–polymer conjugates. I. Assessing heterogeneity of a small PEGylated protein and mapping conjugation sites using ion exchange chromatography and


