

**Investigating Protein-Carbohydrate Interactions with
Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)**

by

Jingjing Zhang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry
University of Alberta

© Jingjing Zhang, 2014

Abstract

The application of hydrogen/deuterium exchange mass spectrometry (HDX-MS) to investigating protein-carbohydrate interactions is described. Proteins from three bacterial toxins, the B subunit homopentamers of Cholera toxin (CTB₅) and Shiga toxin type 1 (Stx1B₅) and a fragment of *Clostridium difficile* toxin A (TcdA-A2), and their interactions with native carbohydrate receptors, GM₁ pentasaccharide (GM₁-os), Pk trisaccharide and CD-grease, respectively, were first served as model systems for this study. The results suggested that HDX-MS can serve as a useful tool for localizing the ligand binding sites in carbohydrate-binding proteins. Following this, HDX-MS measurements were applied to explore the existence of distinct HMOs binding sites on toxins. Altogether, two toxins were studied, CTB₅ and TcdA-A2, and their interactions with HMOs, 2'-fucosyllactose (2'-FL) and lacto-N-tetraose (LNT), respectively. For CTB₅ and its interaction with 2'-FL, a novel binding site was localized for 2'-FL, different from the one for native receptor GM₁-os. For TcdA-A2 and its interaction with LNT, however, the localized binding site was the same as its native carbohydrate receptor CD-grease.

A HDX-MS based titration method Protein-Ligand Interactions in solution by Mass Spectrometry, Titration and hydrogen/deuterium Exchange (PLIMSTEX), was also applied to CTB₅ and its interactions GM₁-os, to test the reliability of using peptides as indicators to obtain the protein-carbohydrate binding affinities. The average apparent association constant measured for the addition of GM₁-os to CTB at pH 7.0 and 20 °C was found to be $(1.6 \pm 0.2) \times 10^6 \text{ M}^{-1}$. This is in reasonable

agreement with the reported value of $(3.2 \pm 0.2) \times 10^6 \text{ M}^{-1}$, which was measured using direct ESI-MS assay at pH 6.9 and room temperature.

Table of Contents

Chapter 1

Investigating Protein-Carbohydrate Interactions with Hydrogen/Deuterium

Exchange Mass Spectrometry (HDX-MS)	1
1.1 Introduction.....	1
1.2 Hydrogen/Deuterium Exchange (HDX).....	4
1.2.1 HDX mechanism.....	4
1.2.2 Factors affecting HDX rates	7
1.2.3 HDX kinetic model of amide H's in structured proteins	11
1.3 Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)	13
1.3.1 HDX-MS workflow	13
1.3.1.1 Protein labeling.....	14
1.3.1.2 HDX quench	15
1.3.1.3 Protein digestion	16
1.3.1.4 LC-MS analysis	17
1.3.1.5 HDX-MS data process and interpretation.....	18
1.3.2 HDX-MS instrumentation.....	20
1.3.2.1 PAL HTX- <i>xt</i> system	20
1.3.2.2 NanoACQUITY UPLC system	22
1.3.2.3 Synapt G2-S high definition mass spectrometer (HDMS)	24
1.3.2.3.1 Electrospray ionization (ESI).....	25
1.3.2.3.2 Quadrupole mass filter.....	28
1.3.2.3.3 Time of flight (TOF) analyzer.....	28

1.4 The Present Work	29
1.5 Literature Cited	31

Chapter 2

Localizing Carbohydrate Binding Sites in Proteins Using

Hydrogen/Deuterium Exchange Mass Spectrometry	39
2.1 Introduction	39
2.2 Experimental Methods	44
2.2.1 Materials	44
2.2.2 HDX-MS	45
2.2.3 Data analysis	47
2.3 Results and Discussion	48
2.3.1 CTB ₅ and its interaction with GM ₁ -os	48
2.3.2 Stx1B ₅ and its interaction with Pk-OH	51
2.3.3 TcdA-A2 and its interaction with CD-grease	55
2.3.4 Influence of protein-carbohydrate interactions on deuterium exchange rates	56
2.4 Conclusions	59
2.5 Literature Cited	60

Chapter 3

Localizing the Binding Site for Toxin-Human Milk Oligosaccharides (HMOs)

Interactions Using Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)	64
3.1 Introduction.....	64
3.2 Experimental Methods.....	66
3.2.1 Materials.....	66
3.2.2 HDX-MS.....	67
3.2.3 Data processing for HDX-MS.....	69
3.3 Results and Discussion.....	69
3.3.1 CTB ₅ and its interaction with 2'-FL.....	69
3.3.2 TcdA-A2 and its interaction with LNT.....	73
3.4 Conclusions.....	76
3.5 Literature Cited.....	77

Chapter 4

Measuring Protein-Carbohydrate Binding Affinities Using Protein-Ligand

Interactions in Solution by Mass Spectrometry, Titration and Hydrogen/Deuterium Exchange (PLIMSTEX)	80
4.1 Introduction.....	80
4.2 Experimental Methods.....	85
4.2.1 Materials.....	85
4.2.2 HDX-MS.....	86

4.2.3 Data processing.....	87
4.2.4 Fitting method.....	88
4.3 Results and Discussion	89
4.4 Conclusions.....	95
4.5 Literature Cited.....	95

Chapter 5

Conclusions and Future Work.....	100
5.1 Literature Cited.....	104

List of Tables

- Table 2.1** Summary of the putative intermolecular H-bonds identified in the crystal structures of the three model protein-carbohydrate complexes: (CTB₅ + 5GM₁-os), PDB 3CHB; (Stx1B₅ + 15Pk), PDB 1BOS; and (TcdA-A2 + 2CD-grease), PDB 2G7C. 49
- Table 4.1** K_a, D_P and D_{PL} values obtained from the nonlinear fitting based on Equation 4.5 for the addition of GM₁-os to CTB monomer. Errors were calculated as the standard deviation for triplicate measurements. 92

List of Figures

- Figure 1.1** Three types of H's within a protein: H's bonded to hetero atoms (Type 1); H's bonded to carbon atoms (Type 2); and backbone amide H's (Type 3). 5
- Figure 1.2** HDX rate constant dependence on solvent pH and temperature for backbone amide H's within freely exposed and unstructured polyalanine. Figure was obtained based on the results calculated using Equation 1.5, with k_A and k_B values of $10^{-1} \text{ M}^{-1}\text{s}^{-1}$ and $10^7 \text{ M}^{-1}\text{s}^{-1}$, respectively. 8
- Figure 1.3** Cartoon illustration of deuterium uptake level for backbone amide H's involved in H-bonding network (left) or buried inside the folded protein (right) after certain labeling time. 9
- Figure 1.4** Measured $\log P_f$ as a function of H-bond acceptor types for backbone amide H's (left) and the distance of the amides from the protein surface (right). Figure was adapted from references 25 and 26. 10
- Figure 1.5** Unique mass spectrometry signature of a peptide that follows EX1 or EX2 process. From top to bottom, the labeling time is increasing. 12
- Figure 1.6** Illustration of typical HDX-MS experimental workflow to determine deuterium uptake values for backbone amide H's at peptide level for investigation of protein-carbohydrate interactions. 14
- Figure 1.7** D_i value versus labeling time. With increasing of the labeling times, D_i value increases. 20

- Figure 1.8** Scheme of PAL HTX-xt system used in the study of this thesis for automatic sample preparation and injection. Figure was adapted from <http://www.leapwiki.com>. 21
- Figure 1.9** Illustration of HDX sample manager for HDX-MS experiments at its working modes: load mode (up) and injection mode (bottom). 23
- Figure 1.10** Schematic diagram of Synapt G2-S high definition mass spectrometer (HDMS). Figure was provided by Waters Corporation. 25
- Figure 1.11** Schematic representation of positive ion mode ESI and the processes of gas phase ion formation. 26
- Figure 1.12** Illustration of IEM pathway. Small analyte ion (red) emissions from the charged droplets (blue). 27
- Figure 2.1** Protein-carbohydrate complexes considered for HDX-MS analysis: (a) (CTB₅ + 5GM₁-os) complex (each subunit has one GM₁-os binding site), (b) (Stx1B₅ + 15Pk) complex (each subunit has three Pk binding sites) and (c) (TcdA-A2 + 2CD-grease) complex (each TcdA-A2 has two CD-grease binding sites). (d) The structures of the carbohydrate ligands, GM₁-os, Pk-OH and CD-grease. 42
- Figure 2.2** (a) Difference plot of peptide level D-uptake between free CTB₅ and (CTB₅ + 5GM₁-os) complex for a subset of peptides that cover 96.1% of CTB monomer sequence. The x-axis indicates the position (residue number) of the peptides considered. The y-axis shows the ΔD_i values for each peptide, summed over all of the labeling times. The errors

bars correspond to one standard deviation. (b) Cartoon illustration of ΔD_i values mapped onto the crystal structure of the (CTB₅ + 5GM₁-os) complex (PDB 3CHB). The coloured region corresponds to one representative CT B subunit. The regions highlighted in red (with corresponding residue numbers indicated) exhibited protection from deuterium exchange upon ligand binding, while the regions highlighted in cyan were unaffected by ligand binding. 50

Figure 2.3 (a) Difference plot of peptide level D-uptake between free Stx1B₅ and (Stx1B₅ + 15Pk-OH) complex for a subset of peptides that cover 95.6% of Stx1B monomer sequence. The x-axis indicates the position (residue number) of the peptides considered. The y-axis shows the ΔD_i values for each peptide, summed over all of the labeling times. The errors bars correspond to one standard deviation. (b) Cartoon illustration of ΔD_i values mapped onto the crystal structure of the (Stx1B₅ + 15Pk) complex (PDB 1BOS). The coloured region corresponds to one representative Stx1 B subunit. The regions highlighted in red (with corresponding residue numbers indicated) exhibited protection from deuterium exchange upon ligand binding, while the regions highlighted in cyan were unaffected by ligand binding. 53

Figure 2.4 (a) Difference plot of peptide level D-uptake between free TcdA-A2 and (TcdA-A2 + 2CD-grease) complex for a subset of peptides that

cover 92.4% of the TcdA-A2 sequence. The x-axis indicates the position (residue number) of the peptides considered. The y-axis shows the ΔD_i values for each peptide, summed over all of the labeling times. The errors bars correspond to one standard deviation. (b) Cartoon illustration of ΔD_i values mapped onto the crystal structure of the (TcdA-A2 + 2CD-grease) complex (PDB 2G7C). The regions highlighted in red and dark blue exhibited protection or de-protection, respectively, against deuterium exchange upon ligand binding; the regions shown in cyan were unaffected by ligand binding. 54

Figure 3.1 Difference plot of peptide level D-uptake between free CTB₅ and (CTB₅ + 5(2'-FL)) complex for a subset of peptides that were common to all repeated experiments covering 94.2% of CTB monomer sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments. 70

Figure 3.2 Difference plot of peptide level D-uptake between GM₁-os saturated CTB₅ and (CTB₅ + 5GM₁-os + 5(2'-FL)) complex for a subset of peptides that were common to all repeated experiments covering 94.2% of CTB monomer sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are

standard deviations of three independent experiments. 71

Figure 3.3 Cartoon illustration of ΔD_i values mapped onto the crystal structure of (CTB₅ + 5GM₁-os) complex (PDB 3CHB). Colored region is a representative CTB monomer. The red and purple regions of protein were protected from deuterium exchange upon GM₁-os and 2'-FL binding, respectively, containing specific residues shown as labeled, while the cyan regions were unaffected by the bound ligand. 72

Figure 3.4 Difference plot of peptide level D-uptake between free TcdA-A2 and (TcdA-A2 + LNT) complex for a subset of peptides that were common to all repeated experiments covering 85.5% of TcdA-A2 sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments. 74

Figure 3.5 Difference plot of peptide level D-uptake between CD-grease saturated TcdA-A2 and (TcdA-A2 + CD-grease + LNT) complex for a subset of peptides that were common to all repeated experiments covering 85.5% of TcdA-A2 sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments. 76

Figure 3.6 Cartoon illustration of ΔD_i values mapped onto the crystal structure of

(TcdA-A2 + 2CD-grease) complex (PDB 2G7C). The red and dark blue regions of protein were protected and de-protected from deuterium exchange upon LNT binding containing specific residues shown as labeled, while the cyan regions were unaffected by the bound ligand.

Figure 4.1 Illustration of detailed procedures for PLIMSTEX method at peptide level. 82

Figure 4.2 (a) (CTB₅ + 5GM₁-os) complex with yellow sticks represent the ligands. Colored region is a representative CTB monomer (includes one loop from adjacent subunit involved in binding). The red regions of protein were protected from deuterium exchange upon ligand binding, while the blue regions were unaffected by the bound ligand. (b) Chemical structures of GM₁-os. 85

Figure 4.3 Sequence coverage (97.1%) of peptic peptides from the CTB monomer. Each bar under the sequence indicates a common peptide reproducibly identified in all samples. A total of 13 peptides were selected for further analysis. Peptides bars labeled in red were the regions protected by the bound GM₁-os. Secondly structural information is also illustrated in the map. 90

Figure 4.4 With an increase in concentration of GM₁-os, the D_i values for the six GM₁-os protected peptides from CTB monomer, (a) peptide 4-15, (b) peptide 27-38, (c) peptide 49-56, (d) peptide 49-68, (e) peptide 49-76

and (f) peptide 86-98, decreased. Data was obtained with 20 μ M of CTB as initial concentration. 91

Figure 4.5 Nonlinear fitting based on Equation 4.5 (using Origin 9.1) for the six GM₁-os protected peptides from CTB monomer (a) peptide 4-15, (b) peptide 27-38, (c) peptide 49-56, (d) peptide 49-68, (e) peptide 49-76 and (f) peptide 86-98. Data was obtained with 20 μ M of CTB as initial concentration. 93

Figure 5.1 Cartoon representatives of nanodiscs. Purple belts are two copies of MSPs surrounding the phospholipid bilayer with incorporated receptors. 102

List of Abbreviations

ACN	Acetonitrile
Amide H's	Amide hydrogens
ASM	Auxiliary solvent manager
BSM	Binary solvent manager
CEM	Chain ejection model
CD-grease	α -Gal-(1 \rightarrow 3)- β -Gal-(1 \rightarrow 4)- β -GlcNAcO(CH ₂) ₈ CO ₂ CH ₃
CRM	Charge residue model
CTB ₅	Cholera toxin B subunit homopentamer
D	Deuterium
Da	Dalton
DC	Direct current
D _i	Deuterium uptake for a specific peptide i
ΔD_i	Relative deuterium uptake for a specific peptide i between free and ligand-bound protein
D-uptake	Deuterium uptake
D ₂ O	Deuterium oxide
ESI	Electrospray ionization
2'-FL	α -L-Fuc(1-2)- β -D-Gal(1-4)- β -D-Glc
GM ₁	Monosialotetrahexosylganglioside
GM ₁ -os	β -Gal-(1 \rightarrow 3)- β -GalNAc-(1 \rightarrow 4)[α -Neu5Ac-(2 \rightarrow 3)]- β -Gal- (1 \rightarrow 4)-Glc)

GndCl	Guanidine hydrochloride
H	Hydrogen
H-bond	Hydrogen bond
HDX	Hydrogen/deuterium exchange
HDX-MS	Hydrogen/deuterium exchange mass spectrometry
HMOs	Human milk oligosaccharides
HPLC	High performance liquid chromatography
IEM	Ion evaporation model
k_d	Diffusion-limit rate constant
k_A	Acid catalysis rate constant
K_a	Association constant
k_B	Base catalysis rate constant
K_D	Dissociation constant
K_{eq}	Equilibrium constant
k_{ex}	Apparent exchange rate constant for free protein
k_{HX}	Apparent exchange rate constant for ligand-bound protein
k_{int}	Intrinsic exchange rate constant
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
L	Ligand
LC	Liquid chromatography
LNT	β -D-Gal(1-3)- β -D-GlcNAc(1-3)- β -D-Gal(1-4)- β -D-Glc

MS	Mass spectrometry
MSP	Membrane scaffold protein
MW	Molecular weight
m/z	Mass-to-charge ratio
N	Nitrogen
ND	Nanodisc
NMR	Nuclear magnetic resonance
O	Oxygen
P _f	Protection factor
Pk-OH	α -Gal-(1→4)- β -Gal-(1→4)-Glc
PLIMSTEX	Protein-Ligand Interactions in solution by Mass Spectrometry, Titration and hydrogen/deuterium Exchange
RF	Radio frequency
S	Sulfur
Stx1B ₅	Shiga toxin B subunit homopentamer
<i>t</i>	Labeling time
TcdA-A2	A2 fragment of <i>Clostridium difficile</i> toxin A
TCEP	<i>Tris</i> (2-carboxyethyl)phosphine
TOF	Time of flight
UPLC	Ultra performance liquid chromatography
ZrO ₂	Zirconium oxide

Chapter 1

Investigating Protein-Carbohydrate Interactions with Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)

1.1 Introduction

Carbohydrates are one of the most important and abundant biological molecules in biological systems.^{1,2} They are generally found as glycolipids, glycopeptides and glycoproteins at the cell surface, and serve as receptors for other biomolecules like lectins, antibodies and carbohydrate-processing enzymes on the surface of other cells or in solution. Protein-carbohydrate interactions are implicated in a wide range of cellular processes, including cell-cell and cell-matrix interactions, signal transduction, inflammation, cancer metastasis, bacterial and viral infections and the immune response.³⁻⁶ Elucidating the structures of protein-carbohydrate complexes, as well as the kinetics and thermodynamics of the interactions, is an essential prerequisite for developing a complete understanding of many physiological and pathological cellular processes and to guide drug discovery and design efforts.⁷⁻⁹

High resolution techniques, such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, have been extensively applied to elucidate the three dimensional structures of biological macromolecules and their interactions with ligands.^{10,11} X-ray crystallography remains the gold standard for providing structural information of free protein and protein-ligand complex at atomic resolution.^{12,13} However, optimization of crystallization can be a lengthy process, and not all the proteins can be easily crystallized or co-crystallized with their ligands.¹⁴

Besides, it's not clear whether the protein conformation within the crystal is identical to that in bulk solution. NMR is also widely used to characterize the structures of biological molecules and their complexes in solution.^{11,15} However, NMR measurements are usually limited to relatively small proteins, with molecular weights (MWs) < 40 kDa.¹⁶ Additionally, NMR measurements generally require large amounts of sample (typically mg quantities) and are time consuming, which limit their application. Many protein-carbohydrate interactions are not amenable to these techniques due to limitations associated with protein size, solubility or ease of crystallization, as well as the cost and availability of pure oligosaccharide ligands.¹⁷⁻¹⁹ As a result, there is a need for alternative techniques capable of probing protein-carbohydrate interactions. Recently, the hydrogen/deuterium exchange mass spectrometry (HDX-MS) has emerged as a new tool to study protein-carbohydrate interactions.^{20,21} Using MS as the detector, this method provides rates of amide hydrogens (H's) on a protein backbone exchange with deuterium in the solvent. HDX rates can then be used to extract out structural information of the protein or protein-ligand complex.

The HDX method was conceived by Linderstrøm-Lang in the early 1950s, with the goal to finding H-bonded structures, i.e. α -helix and β -sheet, in proteins.²²⁻²⁴ However, results revealed that HDX rate of a protein also depends on several other factors rather than the H-bonded structures within that protein alone. Later on, using NMR as the detector for HDX rate, studies were carried out to study the mechanism of HDX phenomenon, and its relationship with protein structures.^{25,26} At this stage, proteins were studied as a whole and it was impossible to relate the HDX rate to

different structures within a protein and furthermore locate them. More recently, with the application of proteolytic fragmentation of the protein, the peptides produced can be used to extract structural information within a short region of the protein, and improved spatial resolution was achieved.²⁷⁻²⁹ The first application of MS as the detector for HDX rates was carried out in the early 1990s,³⁰ which took advantage of the gentle ionization method electrospray ionization (ESI). Since then, HDX-MS methods have improved rapidly due to incredible improvement in instrumentation and software. Some excellent recent reviews cover the methodologies and applications of HDX-MS to study protein structures and dynamics,³¹⁻³⁵ to investigate protein-ligand interactions,^{5,36,37} and also the illustration of key issues related to this technique.^{28,38-40}

Compared to techniques like X-ray crystallography and NMR, using HDX-MS to study protein-carbohydrate interactions possesses several advantages, includes improved sensitivity and also the ability to study large protein systems. Only micrograms of protein sample are normally required to obtain details of protein-ligand interactions. In principle, there is no size limit of the protein and protein-ligand complex. Besides, HDX-MS measurements are able to monitor the protein-carbohydrate interactions in their physiological buffers, which are more close to what happened *in vivo*.

However, like all other methods, HDX-MS has its limitations when being used to investigate protein-carbohydrate interactions. An important fundamental assumption in using HDX-MS measurements to investigate protein-ligand interactions is that the ligand will provide enough protection from deuterium exchange of the backbone amide H's within the ligand binding site. However, it's not

always the case.⁴¹ The potential difficulties of this assay will be discussed in more detail in Chapter 2. Although the relationship between HDX rates and the protein structures is not fully understood yet, the HDX-MS measurements have become a promising technique for characterizing protein-carbohydrate interactions.

Before describing the HDX-MS measurements in detail, it is necessary to first review the basic principles of HDX. An overview of the HDX mechanisms, followed by the workflow, is given below.

1.2 Hydrogen/Deuterium Exchange (HDX)

1.2.1 HDX mechanism

H's within a protein are exchanging with H's in the surrounding solvent unnoticeably all the time.⁴² However, when exposing a protein to a deuterated solvent (D₂O), H's in the protein will be replaced by deuteriums. Therefore, the mass of the protein will increase by one for each exchange event, which is detectable by MS. In principle, there are three types of H's within a protein (Figure 1.1). They are H's bonded to hetero atoms, i.e. N, O and S, on side chains of amino acids with fast exchange rates (less than ms) (Type 1); H's bonded to carbon atoms for which the exchange take much longer than several days (Type 2); and backbone amide H's, which have exchange rates in a range that can be easily measured (Type 3). Since each residue (except proline) has a backbone amide H, and those H's are involved in different structures, e.g. α -helix, β -sheet or loop, possess varied HDX rates, HDX-MS can be used to probe the structural features affecting the whole protein.

In principle, HDX of an unstructured and solvent exposed peptide can be catalyzed by acid and base.^{43,44} In the case of base catalysis, the deuteroxide ions and the peptide substrate (PH) will form an encounter complex, followed by proton transfer through H-bonding, which can be expressed as follows:

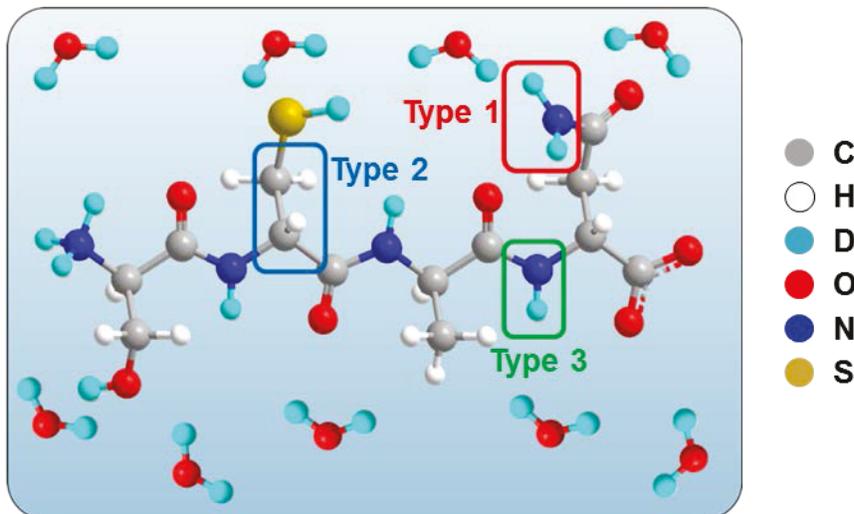
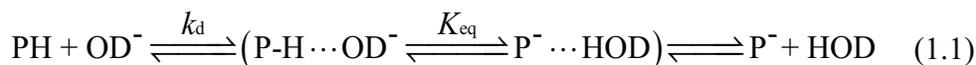


Figure 1.1 Three types of H's within a protein: H's bonded to hetero atoms (Type 1); H's bonded to carbon atoms (Type 2); and backbone amide H's (Type 3).



where k_d and is the diffusion-limit rate constant. The equilibrium reached within the brackets can be calculated from the equilibrium constant $K_{\text{eq}} = 10^{\Delta\text{pK}_a}$, where $\Delta\text{pK}_a = \text{pK}_{\text{HOD}} - \text{pK}_{\text{PH}}$ (pK_{HOD} and pK_{PH} are pK_a values for HOD and PH, respectively).

Therefore, the fraction of complex $\text{P}^- \cdots \text{HOD}$ can be expressed as:

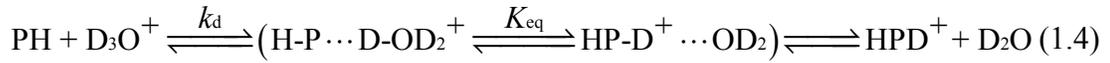
$$\frac{K_{\text{eq}}}{K_{\text{eq}} + 1} = \frac{10^{\Delta\text{pK}_a}}{10^{\Delta\text{pK}_a} + 1} \quad (1.2)$$

The overall rate constant for proton transfer, $k_{\text{H-transfer}}$, can be calculated using Equation 1.3:

$$k_{\text{H-transfer}} = k_d \times \frac{10^{\Delta pK_a}}{10^{\Delta pK_a} + 1} \quad (1.3)$$

Since the intermediate anion P^- can quickly convert to the deuterated product PD in the deuterated solvent D_2O , the rate-determining step of the HDX reaction is always the proton transfer step. Therefore, the complete base catalysis rate constant k_B is the same as $k_{\text{H-transfer}}$. And the 1st order HDX rate constant can be calculated as k_B times the concentration of the catalyst, or $k_B \times [\text{OD}^-]$. At 25 °C, for freely exposed and unstructured protein backbone amide H's, k_B is around $10^7 \text{ M}^{-1}\text{s}^{-1}$.⁴⁴

The acid catalysis pathway is similar as the base catalysis pathway, which is expressed as follows:



Similarly, the intermediate cation HPD^+ can quickly convert to the deuterated product PD. At 25 °C, for freely exposed and unstructured protein backbone amide H's, k_A is around $10^{-1} \text{ M}^{-1}\text{s}^{-1}$.⁴⁴

In D_2O , HDX reactions are catalyzed by both D^+ and OD^- . Therefore, the intrinsic exchange rate constant, k_{int} , of HDX reaction can be expressed by the following equation:

$$k_{\text{int}} = k_A [\text{D}^+] + k_B [\text{OD}^-] \quad (1.5)$$

1.2.2 Factors affecting HDX rates

A wide variety of factors can influence the HDX rate of a backbone amide H, such as side chain effects,⁴⁵ solvent pH,^{43,46} temperature,⁴⁷ H-bonding network,²⁶ as well as solvent accessibility.²⁵ As mentioned above, the pK_a value of an amide H will affect the proton transfer step, thus affecting its overall catalysis rate constant. At 25 °C, amide H's have a pK_a value around 18.5. However, for a specific amide H, this value will be affected by the side chain of the amino acid through inductive effect.⁴⁴ Besides, side chains of adjacent residues within a polypeptide will also have steric effect on the amide H. Both side chain effects have been quantitatively assessed using model dipeptides. It turns out that the inductive effect and the steric effect are additive.^{45,48} Generally, the side chain effects in polypeptides with different amino acid sequences can change the HDX rates by about tenfold.⁴⁹

In addition, according to Equation 1.5, the intrinsic rate constant k_{int} depends linearly on the concentration of D^+ and OD^- in solution. Therefore, the solvent pH is a major consideration for HDX rate. Figure 1.2 shows the relationship between the intrinsic exchange rate constant, k_{int} , and solvent pH for backbone amide H's in polyalanine at 25 °C (results were calculated using Equation 1.5). The slowest intrinsic rate constant occurred at pH around 2.6. HDX at pH on the lower and higher side is dominated by acid and base catalysis, respectively. And the rate constant is changed by around 1 order of magnitude for each unit of pH at both sides.

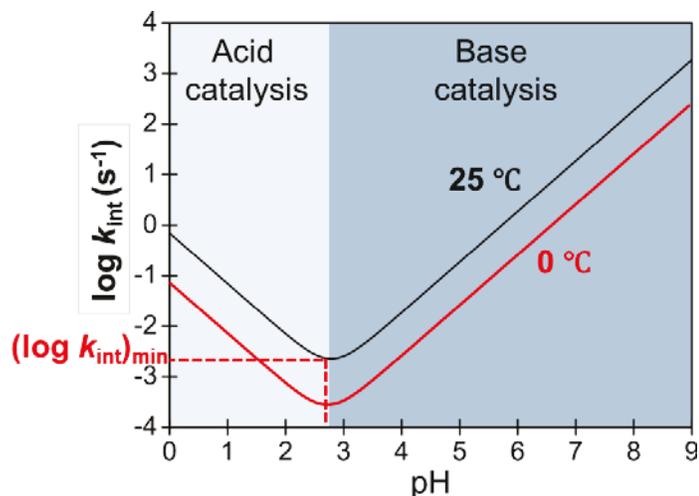


Figure 1.2 HDX rate constant dependence on solvent pH and temperature for backbone amide H's within freely exposed and unstructured polyalanine. Figure was obtained based on the results calculated using Equation 1.5, with k_A and k_B values of $10^{-1} \text{ M}^{-1}\text{s}^{-1}$ and $10^7 \text{ M}^{-1}\text{s}^{-1}$, respectively.

Temperature is also an important factor that affects the HDX rates for amide H's. The temperature dependence can be predicted based on the activation energies for k_A and k_B (14 and 17 kcal/mol, respectively⁴⁸). Generally, decreasing the temperature from 25 °C to 0 °C, the HDX rate is decreased by around 10 times.⁴⁴

The pH and temperature dependence of k_{int} is crucial for HDX measurements. Due to this dependence, protein labeling can be performed under neutral pH at around room temperature. Then the HDX reaction can be quenched by decreasing the pH and temperature to maintain the exchange profile before analysis.

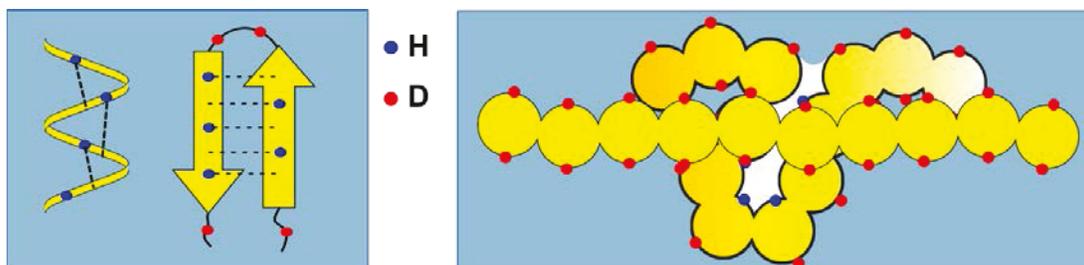


Figure 1.3 Cartoon illustration of deuterium uptake level for backbone amide H's involved in H-bonding network (left) or buried inside the folded protein (right) after certain labeling time.

Apart from the pH and temperature dependence, HDX rates of amide H's from a folded protein are affected by the H-bonding network and solvent accessibility as well (Figure 1.3).^{50,51} If the amide H's are located in structured regions like α -helix or β -sheet, and are involved in H-bonding network. They are less likely to form the encounter complex with the catalysts in solvent. Or if they are deeply buried inside the protein and are not accessible to the solvent, the catalysis pathways will also be affected due to the change in diffusion-limit rate constant (Equation 1.3). In both cases, the values of apparent exchange rate constant, k_{ex} , for the amide H's, are smaller than when they are in unstructured and exposed peptides. These effects are indicated by protection factor P_f (Equation 1.6).

$$P_f = \frac{k_{int}}{k_{ex}} \quad (1.6)$$

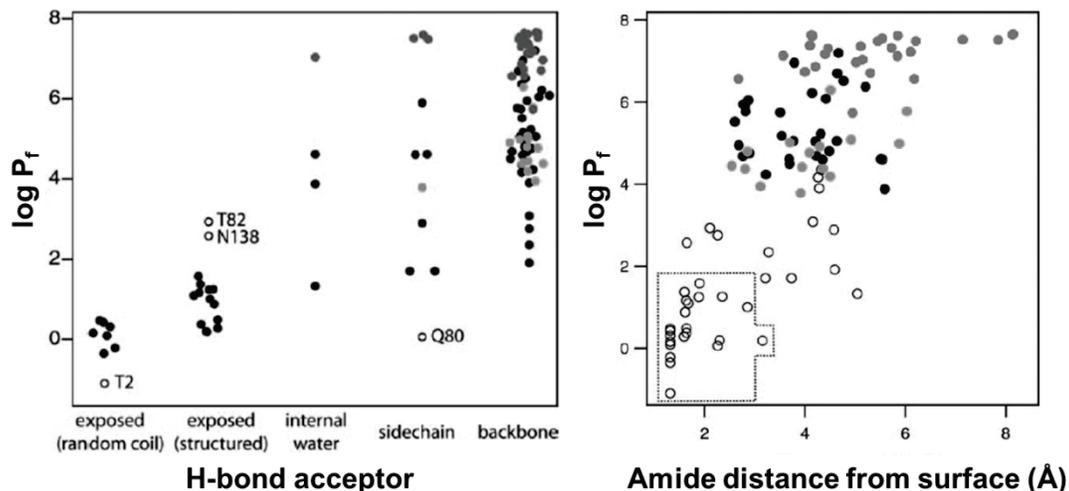
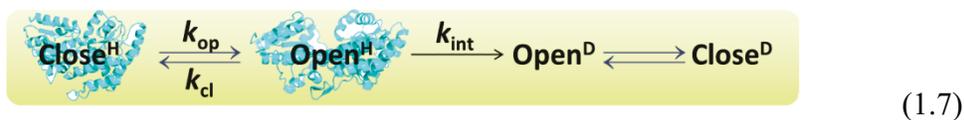


Figure 1.4 Measured $\log P_f$ as a function of H-bond acceptor types for backbone amide H's (left) and the distance of the amides from the protein surface (right). Figure was adapted from references 25 and 26.

Practically, k_{int} value can be predicted from modeling the unstructured peptides at any pH and temperature using Equation 1.5,⁵² and k_{ex} value can be measured through time resolved HDX experiments. Obtained P_f thus provides the opportunity to quantitatively assess the effect of protein secondary and tertiary structure on HDX rates of amide H's. According to previous studies,^{25,26} the fast exchanging amide H's with low P_f only exist within random structures at the protein surface (high solvent accessibility). However many others exchange far more slowly. The P_f can be as high as 10^8 , which makes HDX a sensitive probe to monitor the protein structure.

1.2.3 HDX kinetic model of amide H's in structured proteins

In general, HDX-MS is most widely used to study protein dynamics or conformational changes. Obtaining varied HDX rates for different states of a protein (e.g. free and ligand-bound protein) indicates the presence of a change to that protein. To have a better understanding of the factors affecting HDX kinetics of amide H's in a structured protein, a model had been developed by Linderstrøm-Lang and his coworkers, as illustrated as follows:⁴⁴



A protein exists in equilibrium between “open” and “close” forms determined by k_{op} and k_{cl} , which are the rate constants for protein “open” and “close” processes, respectively. In the open form, the intramolecular H-bonds protecting the amide H's from HDX may be broken and the amide H's can undergo exchange with the rate constant k_{int} . The closed form in this kinetic model is assumed to be un-exchangeable. Hence the apparent exchange rate k_{ex} will be affected by the equilibrium.

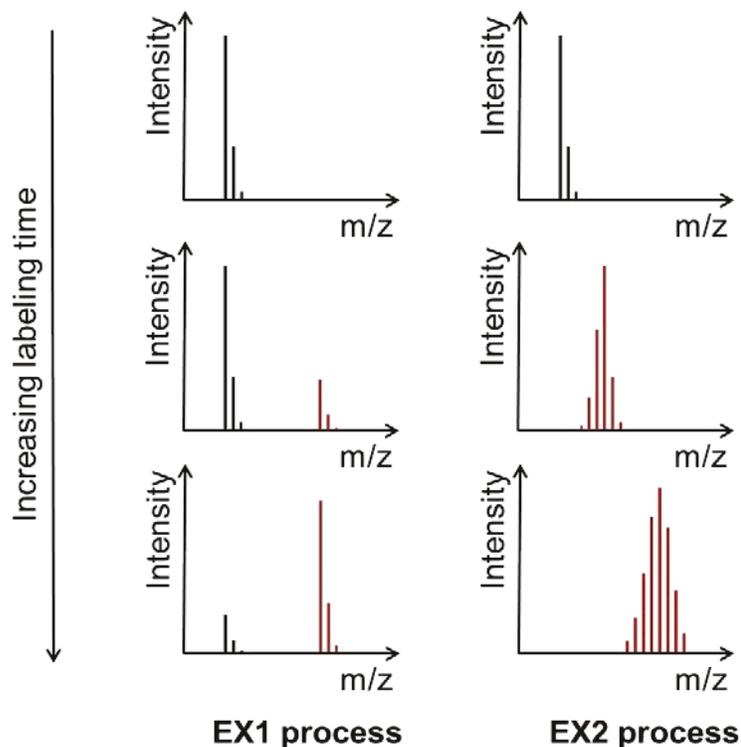


Figure 1.5 Unique mass spectrometry signature of a peptide that follows EX1 or EX2 process. From top to bottom, the labeling time is increasing.

When $k_{\text{int}} \gg k_{\text{cl}} > k_{\text{op}}$ (for a stable structured protein, $k_{\text{cl}} > k_{\text{op}}$), the open process is the rate limiting step, and k_{ex} is controlled by k_{op} , or $k_{\text{ex}} = k_{\text{op}}$, which is known as EX1 process. When $k_{\text{int}} \ll k_{\text{cl}}$, using steady-state approximation, k_{ex} can be calculated as:

$$k_{\text{ex}} = \frac{k_{\text{op}}}{k_{\text{cl}}} \times k_{\text{int}} \quad (1.8)$$

This is known as EX2 process. EX2 process is typically encountered rather than EX1 for proteins under native conditions.^{53,54} It is important to separate the two kinetic processes when interpreting the HDX results. This can be done by detecting their

unique signatures in HDX-MS experiments (Figure 1.5). Once the kinetic process is figured out, k_{op} or k_{op}/k_{cl} can be derived from the observed HDX rates.

The kinetic model illustrated in Equation 1.7 also forms the basis when using HDX measurements to investigate protein-ligand interactions. A detailed discussion will be provided in Chapter 2.

1.3 Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)

1.3.1 HDX-MS workflow

The general HDX-MS procedures to study protein-carbohydrate interactions are known as bottom-up continuous labeling workflow. The detailed steps are shown as Figure 1.6.^{55,56} Both free protein (a) and protein-carbohydrate complex (b) samples are first incubated in D₂O for a series of labeling times to establish their HDX profiles. The reactions are then quenched by decreasing the pH and temperature to maintain the HDX profiles prior to further analysis. After that, the proteins are digested by a protease (normally pepsin), followed by liquid chromatography-MS (LC-MS) analysis. All these steps require precise control of pH, temperature and reaction incubation time. Isotope distributions of peptides are obtained from the LC-MS analysis. The average molecular mass for each peptide is calculated as the centroid of the entire envelope of corresponding isotopic peaks. The mass difference between labeling samples ($t > 0$) and a control sample ($t = 0$) defines the deuterium uptake (D-uptake, D_i) for a specific peptide. Relative D-uptake ($\Delta D_i = D_p - D_{PL}$) values are also calculated to compare the D-uptake difference between free protein

($D_i = D_P$) and ligand-bound protein ($D_i = D_{PL}$) to detect localized structural changes and the ligand binding site on the protein.

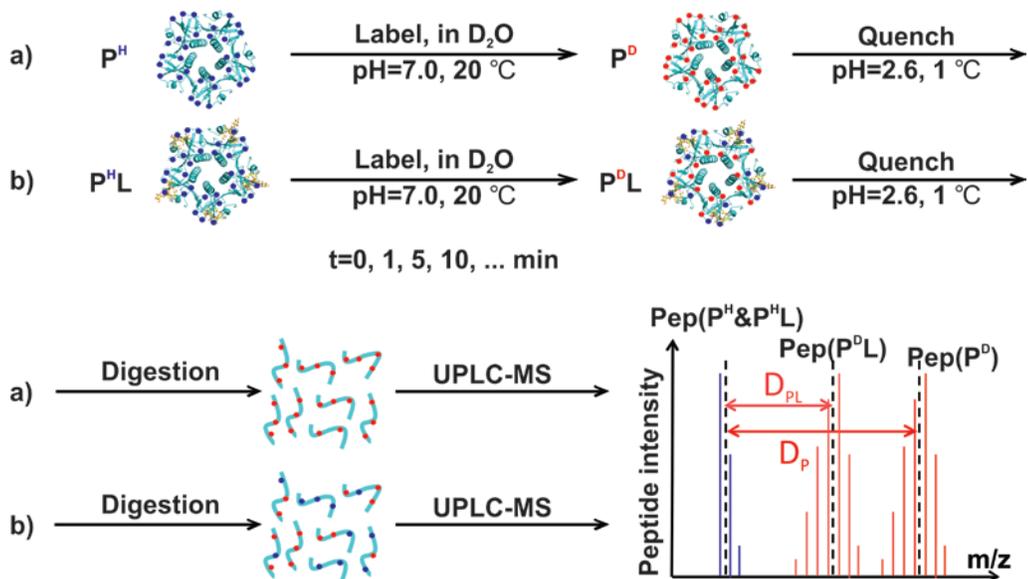


Figure 1.6 Illustration of typical HDX-MS experimental workflow to determine deuterium uptake values for backbone amide H's at peptide level for investigation of protein-carbohydrate interactions.

1.3.1.1 Protein labeling

The removal of salts in protein sample solution can be achieved by using LC prior to MS analysis. This salts removal step thus makes it possible to perform protein labeling in a deuterated buffer which is suitable for the protein to maintain its native state. This is a notable advantage over conventional ESI-MS studies in which protein signal will be suppressed by the salts in the sample solution, thus only limited choices of buffers (e.g. ammonium acetate) are available. The most commonly used labeling

buffer is phosphate buffer because its pK_a values ($pK_{a,1} = 2.15$, $pK_{a,2} = 6.82$) are of perfect match for HDX-MS studies, which can provide excellent buffering capacity at both labeling and quench steps.

In practice, the method to introduce the deuterated buffer into protein sample is by dilution. Dilution of 15-fold or greater will produce a final deuterium concentration greater than 95%, which will force the labeling reaction towards one direction rather than both directions for deuteration and de-deuteration. However, since the original protein sample is diluted through this procedure, care must be taken of dilution-fold to make sure the final concentration of the sample is compatible with the sample amount needed for the mass spectrometer.

1.3.1.2 HDX quench

Following protein-labeling after incubation period, the samples are quenched to maintain the D-uptake profiles before further analysis. This is usually achieved by acidifying the sample pH to around 2.6 and decreasing the temperature to ~ 0 °C. Under this condition, the HDX reaction rates are being slowed by $\sim 10^5$ times (Figure 1.2). In practice, the most commonly used quench buffer is the phosphate buffer.^{57,58} For proteins resistant to pepsin digestion, additives can be added to quench buffer to denature the target protein and improve the digestion.⁵⁹⁻⁶¹ Guanidine hydrochloride (GndCl) is usually used to denature the protein, while *tris*(2-carboxyethyl)phosphine (TCEP) is used to reduce the disulfide bonds within the protein. There are also studies using purely acid to quench the HDX reaction in order to minimize the dilution of the sample.^{62,63} Although after quench, the HDX rate is far slower than the rate under

labeling condition, some back exchange during the protein digestion and LC-MS analysis is inevitable. Therefore, precise control of the pH and temperature at the quench step is crucial to obtain reproducible results.

1.3.1.3 Protein digestion

In order to localize the structural and dynamic changes within a protein, digestion is applied to fragment the protein under quench condition. Pepsin is an ideal protease choice due to its maximum enzymatic activity at acidic condition (pH ~ 2). Besides, since it is a nonspecific enzyme,⁶⁴ a considerable number of peptic fragments with overlapping sequence can be produced. For the overlapping peptides share the same start or end residues, the D-uptake of non-overlapping segment can be considered the same as the difference between the D-uptake for the two overlapping peptides. In this way the spatial resolution of HDX-MS measurements is improved.^{65,66} Despite its non-specificity, pepsin is still a reproducible enzyme. When a protein is digested at identical conditions, the same peptides can be obtained.⁶⁷

To perform pepsin digestion, pepsin can simply be added to the protein solution. At low temperature, 0 °C, a large amount of pepsin is needed with enzyme/substrate ratio of 1:1 or even larger to achieve desired enzymatic activity. Multiple enzymes are used in some studies to generate shorter peptides and improve the spatial resolution of HDX-MS measurements.^{65,68,69} Nowadays, immobilized enzyme column is also available, which improves the digestion efficiency.^{70,71}

1.3.1.4 LC-MS analysis

After protein digestion, the obtained peptides will be analyzed by a LC-MS system. The LC separation can help to increase the number of peptides detected, by allowing peptides with similar molecular mass be eluted at varied retention time. Since there is a desalting step during the LC step, salts and additives used for labeling and quench can be effectively removed. Therefore, samples can be prepared as needed without concern regarding signal suppression in MS. However, the LC separation is performed with protic solvent. As a result, the back exchange is a major concern.⁷² To maintain the quench condition, the mobile phase solutions for LC are prepared with pH ~2.6, and the separation columns as well as tubes needed, are stored in a chamber with temperature ~0 °C. The half-life for deuterium loss under this condition is 20-500 min,⁴⁸ hence a short gradient with high flow rates is needed to minimized the analysis time. Instead of conventional HPLC, UPLC has the advantage of improved separations, shorter separation times, therefore it is widely used for HDX-MS measurements.^{73,74}

Following LC separation, MS acquisition of peptides is performed simultaneously. Due to the non-specificity, the pepsin cleavage sites can't be predicted, and the resulting peptides can't be reliably identified by the molecular mass alone. As a consequence, a further fragmentation of obtained peptide for identification is required. For this study, the peptide identification is achieved with MS^E acquisition mode with a Synapt G2-S HDMS mass spectrometer (Waters, UK). MS^E acquisition method rapidly alternates between two functions with low and high energy, respectively.⁷⁵ With low-energy, exact mass of precursor peptide ions are

obtained, while with elevated-energy, spectra of fragment ions are acquired. The precursor and fragment spectra for each peptide are aligned according to the retention time in LC separations, and peptide identification is achieved by combining those results. Consequently, MS^E records data without discrimination or pre-selection. To reduce the MS measurement error from random fluctuations of the environment and improve the reproducibility, lock-mass correction using [Glu]-Fibrinopeptide is applied.

1.3.1.5 HDX-MS data process and interpretation

To obtain the D-uptake value for peptides of interest, HDX-MS data process starts with the identification of the proteolytical peptides produced from pepsin digestion of the target protein. In the present work, this was achieved by using program ProteinLynx Global SERVER (PLGS, Waters), which contains the sequence for the target protein. By applying non-specific digestion, a database containing the MS ion spectra for all possible peptides is generated by PLGS, as well as the MS/MS ion spectra for peptide product ions. By comparing the spectra for both peptide molecular ion and its product ions between HDX-MS results and the database, the peptides produced in the experiment can be identified, including peptides with overlapping sequence.

The identified peptides lists for all the control samples (i.e. $t = 0$ s) then are imported into DynamX (Waters), which calculates the MW for peptides in the list and determines their D-uptake values at different labeling times. To minimize redundancy, DynamX is capable of generating a smaller peptide list consisting of

peptides that were detected in all replicate measurements of the control samples and that provided maximum sequence coverage. The average MW for each peptide was calculated using the centroid of the entire envelope of the corresponding isotopic peaks, as expressed in Equation 1.9:

$$\text{MW} = \frac{\sum_i I_i \times m_i}{\sum_i I_i} \quad (1.9)$$

where I_i is intensity of the isotopic peak for peptide isotopic form i with the MW m_i . For each peptide more than one charge states may be considered, and the peptide MW is averaged of all the considered charge states. Once the MWs of a peptide i in both labeled and control (no exchange) samples are obtained, the absolute D-uptake value (D_i , units of Da) for the peptide i can be determined as the difference between the two MWs.

To visualize the data, the D_i values are often converted into a graph as shown in Figure 1.7. With increasing of the labeling time, the D_i will also increase until it reaches a plateau in which the peptide is fully deuterated. Because the apparent rate constant for HDX in structured proteins can span 10^8 magnitudes, choosing a series of appropriate labeling times is important when applying HDX to protein studies. Associating the D-uptake values of the peptides in the list to the regions of the intact protein can provide structural information for the target protein. If a stimulus, such as ligand binding, is applied to the target protein, finding different D_i values for the same peptides at the same labeling time is strong evidence that a change (in structure or dynamic or both) has occurred to that part of the protein.

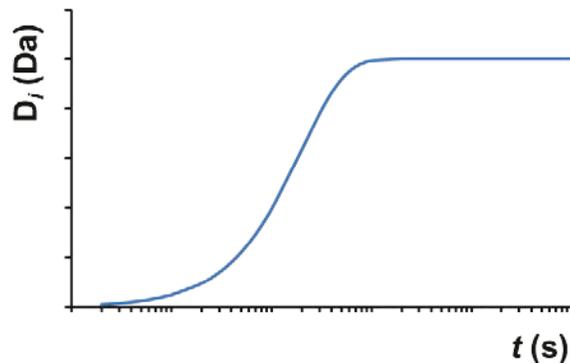


Figure 1.7 D_i value versus labeling time. With increasing of the labeling times, D_i value increases.

1.3.2 HDX-MS instrumentation

In the present study, Synapt G2-S high definition mass spectrometer (HDMS) (Waters, UK) equipped with a nanoACQUITY UPLC system with HDX technology, or HDX sample manager (Waters, UK) and a robot, PAL HTX- xt system (LEAP Technologies, Carrboro, NC, USA) for sample automation, were used.

1.3.2.1 PAL HTX- xt system

To improve the reproducibility and reduce the artificial error during sample preparation of HDX-MS measurements (for precise control of pH, temperature and time), PAL HTX- xt system was used for automatic sample preparation and injection (Figure 1.8). This system is able to accomplish the jobs of sample labeling, quench, as well as sample injection into the LC-MS system for further analysis.⁷⁶⁻⁷⁸

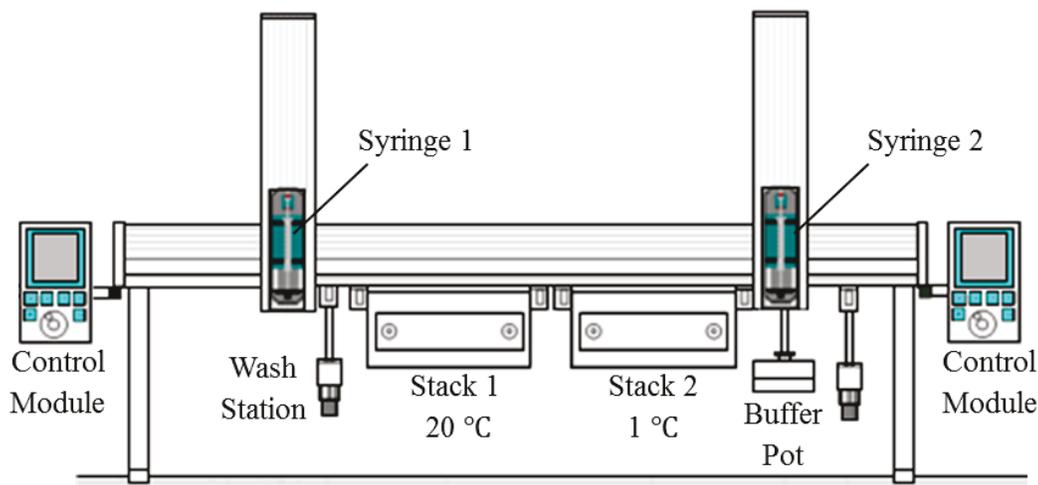


Figure 1.8 Scheme of PAL HTX-*xt* system used in the study of this thesis for automatic sample preparation and injection. Figure was adapted from <http://www.leapwiki.com>.

As illustrated in Figure 1.8, PAL system includes two syringes, one dedicated to sample preparation (draw protein stock sample and place in reaction vial) with a volume of 10 μL (Syringe 2), and the other one syringe help with both sample preparation (labeling and quench) and injection with a volume of 250 μL (Syringe 1). The system also contains two sample stacks (each consists of two trays) with their temperatures be controlled precisely and independently. During the HDX-MS measurement, stack 1 is controlled at 20 $^{\circ}\text{C}$ for protein labeling, whereas stack 2 is controlled at 1 $^{\circ}\text{C}$ for protein stock sample storage as well as HDX reaction quench. Buffers for equilibrium, labeling and quench are stored in the buffer pot. Two different solutions, 3% or 5% ACN in water, are used at the wash station for syringe

wash. Altogether, there are three different modes of operation depending on labeling time: fast (10-19 s), medium (20-105 s) and slow (> 105 s), which is setup by software LEAP Shell 3.0.1. With the automated sample preparation and injection, the precision and reproducibility is improved.⁷⁸

1.3.2.2 NanoACQUITY UPLC system

In the HDX-MS experiments, separation of the peptides produced is a crucial step, and LC has been used to achieve the peptide separation. The LC step consists of a digestion stage as sample is passing through an enzymatic column. Digested fragments are collected onto a trap column. Due to use of salts during the labeling and quench steps, a few minutes (1-3 min) of desalting prior to elution is desired, therefore a trap column (Figure 1.9). After the protein digestion and trapping steps, the peptides produced are eluted to a reversed-phase analytical column for separation and then to the MS. For present work, a NanoACQUITY UPLC system, or HDX sample manager, is used to achieve the peptide separation, which is optimized for high-resolution separations at nanoflow rates (0.2-100 $\mu\text{L}/\text{min}$).

The NanoACQUITY UPLC system mainly consists of three components: auxiliary solvent manager (ASM), binary solvent manager (BSM) and HDX sample manager. In the HDX-MS measurements, a two-pump trapping step is applied, using both ASM and BSM, with a pepsin digestion column, a trap column and an analytical column kept inside the HDX sample manager. The two modes for two-pump trapping are illustrated in Figure 1.9.

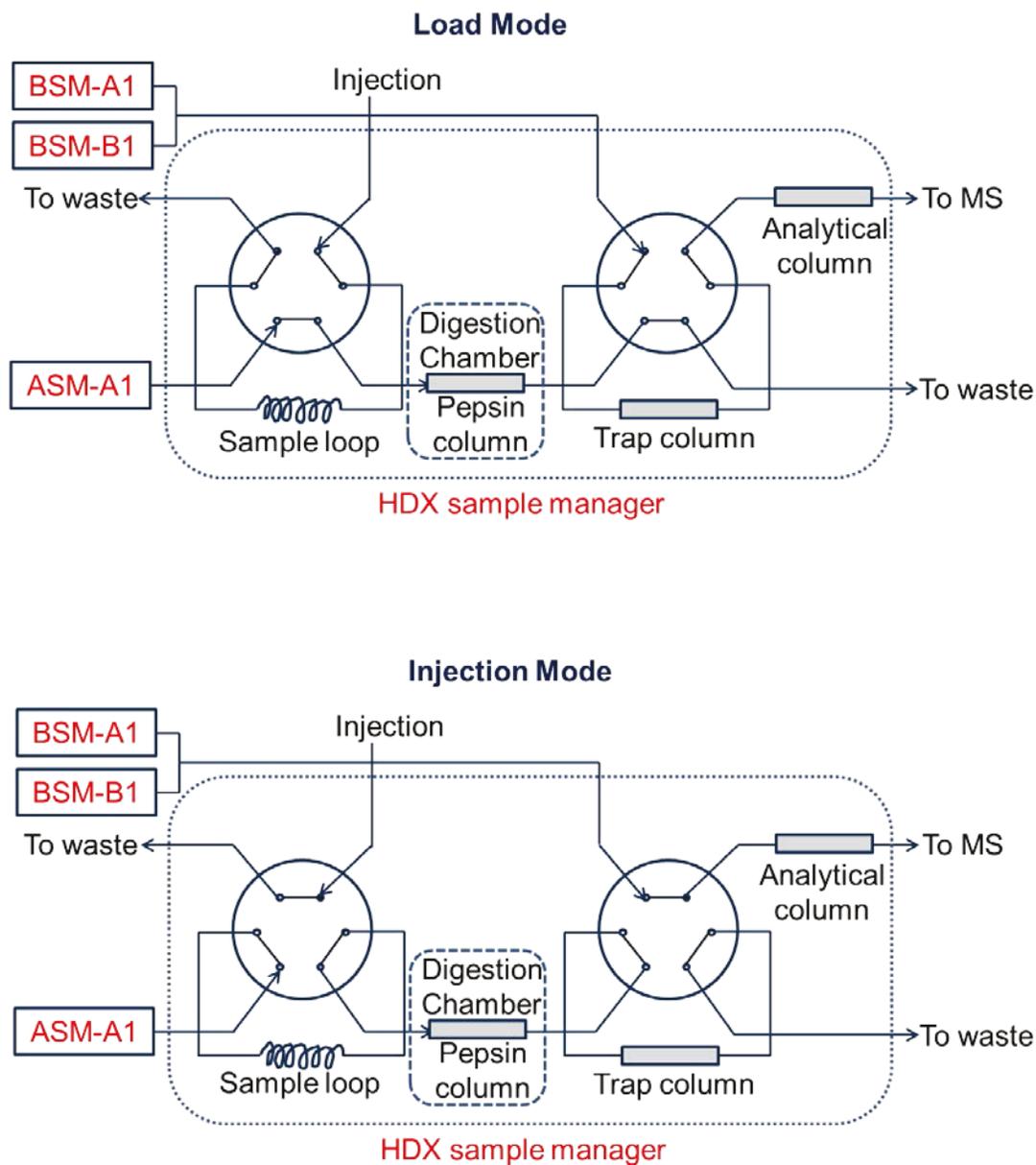


Figure 1.9 Illustration of HDX sample manager for HDX-MS experiments at its working modes: load mode (up) and injection mode (bottom).

Before sample injection, the HDX sample manager is in the load mode. The quenched sample is injected and fills up the sample loop. Once the sample is injected,

the system switches to the injection mode immediately. A dedicated trapping pump at the ASM (pump A1) then pushes the sample in the loop passing through the pepsin column and the trap column. The unwanted solutes like salts flush through the trap column and elute to waste while the peptides produced by the pepsin column are retained on the trap column. The trapping usually takes a few minutes (1-3 min) to allow for protein digestion and also sample desalting. At the end of trapping, the system switches back to the load mode. Gradient elution proceeds as the BSM pumps solvents (pump A1 and B1 for water and acetonitrile, respectively) through the trap and analytical columns, allowing the peptic peptides to be washed off the trap column and separated by the analytical column via flow, and being delivered to the mass spectrometer. During the HDX-MS measurement, the temperature of HDX sample manager is kept at 1 °C to maintain the quench condition, with the exception of the separated digestion chamber which is being kept at 20 °C to ensure efficient digestion.

1.3.2.3 Synapt G2-S high definition mass spectrometer (HDMS)

In this study, the detector of the peptide HDX profiles was a Synapt G2-S high definition mass spectrometer (HDMS) (Waters, UK) equipped with an ESI source (Figure 1.10). Briefly, gaseous ions produced by ESI source are introduced into the mass spectrometer and entering into the stepwave transfer optic, which can minimize neutral contamination and enhance the signal-to-noise ratio. The resulting ions are then transmitted through a quadrupole mass filter to the ion mobility section of the instrument (Triwave, not used in this work). The ions are then detected by an

orthogonal acceleration (oa)-TOF mass analyzer (QuanTOF™) equipped with a high field pusher and a dual-stage reflectron. Detailed working mechanisms of ESI method, quadrupole mass filter and TOF mass analyzer are given below.

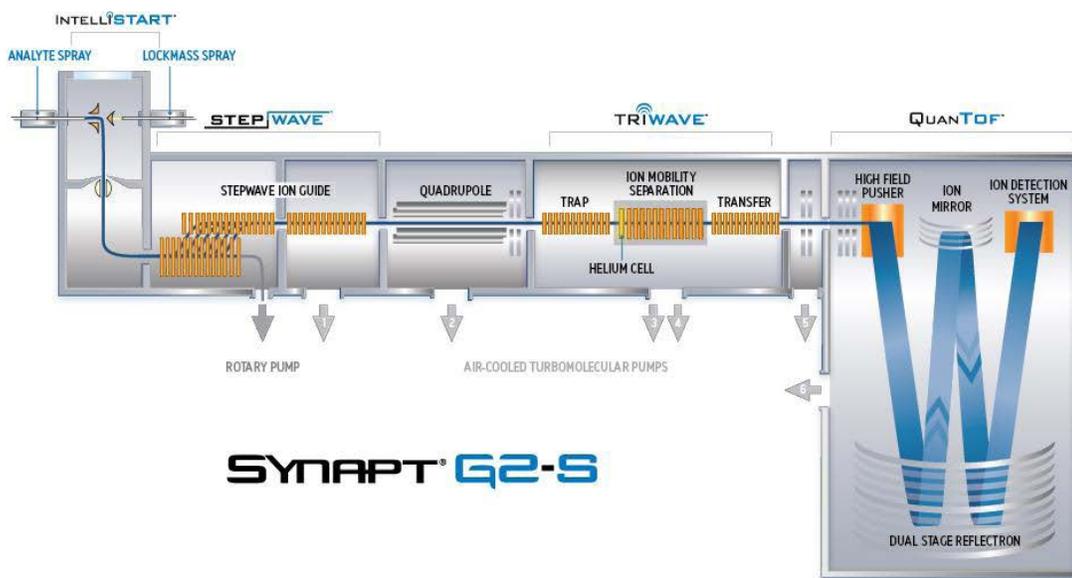


Figure 1.10 Schematic diagram of Synapt G2-S high definition mass spectrometer (HDMS). Figure was provided by Waters Corporation.

1.3.2.3.1 Electrospray ionization (ESI)

As a soft ionization technique, ESI usually generates intact, multiply charged ions at atmospheric pressure (Figure 1.11). The mechanism of the ESI process, as described by Kebarle and coworkers, involves three major steps:⁷⁹

- a) Production of charged droplets at the ESI tip,
- b) Shrinkage of the charged droplets due to solvent evaporation and droplet disintegrations,

c) Formation of fine, highly charged droplets from which gas-phase ions are produced.

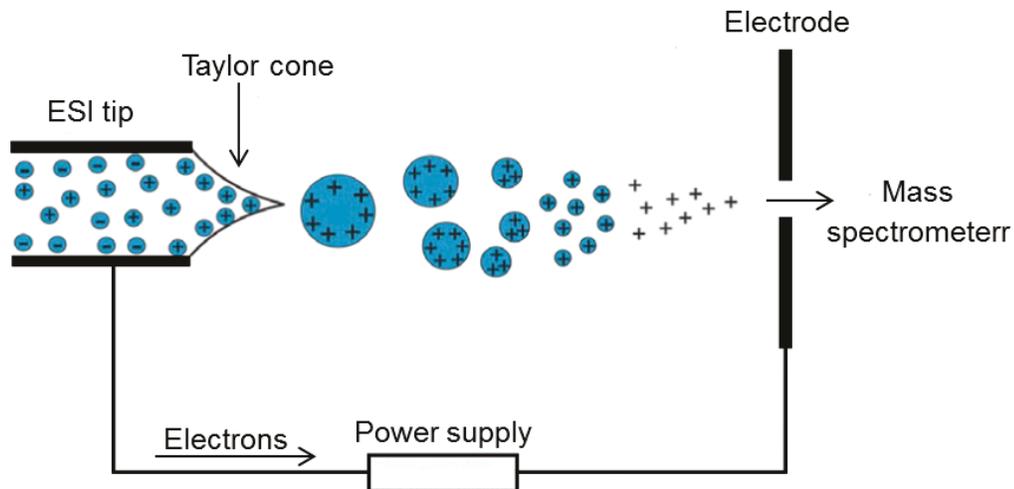


Figure 1.11 Schematic representation of positive ion mode ESI and the processes of gas phase ion formation.

Figure 1.11 shows a schematic diagram describing the ESI process in positive ion mode. A high positive voltage (~ 3 kV) is applied to the capillary, inducing charge separation in solution. The positive ions drift towards the liquid surface, followed by the formation of a liquid cone known to as Taylor cone.⁸⁰ Taylor cone is a stable liquid cone exists with competing forces between downfield forces generated by electric field and resistance by surface tension of the liquid. The Taylor cone will become unstable at a sufficiently high electric field, leading to the emission of a thin liquid filament whose surface is enriched in positive ions.⁸⁰ Afterwards, the thin liquid filament breaks up into small positively charged droplets. Due to rapid solvent

evaporation, the droplets shrink, leading to an increase in charge density on the droplets surface. When the Columbic repulsion of the surface charges overcomes the surface tension of the droplet, the droplets become unstable and undergoing Coulomb fission, forming small, highly charged offspring droplets. Repeated evaporation/fission events eventually yield the final generation of ESI droplets with radii of a few nanometers. There are three proposed mechanisms for the production of gas-phase ions from these droplets, i.e. the ion evaporation model (IEM),⁸¹ the charged residue model (CRM)⁸² and the chain ejection model (CEM),⁸³ which apply for low molecular weight analytes, large globular species and disordered polymers, respectively.^{84,85} In this study, the analytes were peptides with average length around ten amino acids, therefore it should follow the IEM mechanism (Figure 1.12), which was proposed by Iribarne and Thomson.⁸⁶ It predicts that direct ion emission from the droplets will occur after the radii of the droplets shrink to radii less than 10 nm.^{81,86}

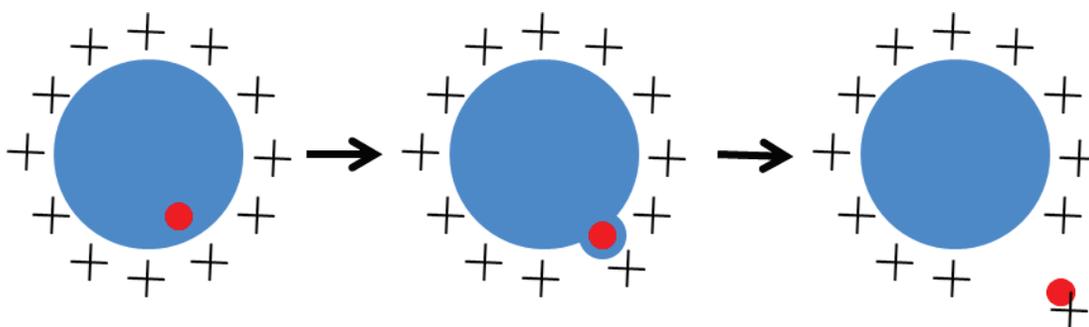


Figure 1.12 Illustration of IEM pathway. Small analyte ion (red) emissions from the charged droplets (blue).

1.3.2.3.2 Quadrupole mass filter

The quadrupole is consisted of four cylindrical metal rods which are accurately positioned in a radial array, with the diametrically opposed rods paired with each other (Figure 1.10). A direct current (DC) potential and a radiofrequency (RF) potential, are applied to each pair of rods (the same absolute potential with different sign to each rod) to create a hyperbolic field.⁸⁷ By applying specific voltage and frequency, ions which possess a small range of mass-to-charge (m/z) values can be selected and transmitted through the quadrupole, whereas other ions with m/z values out of the small range will hit the rods and are discharged. The width of the bandpass region is controlled by the DC and RF potentials applied to the rods.⁸⁸ By operating in RF only mode, the quadrupole can also act as a broad bandpass filter, which transmits and guides ions over a wide m/z range to other components of the apparatus.

1.3.2.3.3 Time of flight (TOF) analyzer

TOF analyzers measure the flight time of the ions to move through a flight tube between the source and detector, and determine their m/z values using Equation 1.12:^{89,90}

$$\sqrt{\frac{m}{z}} = t \left(\frac{\sqrt{2eV_s}}{L} \right) \quad (1.12)$$

where m is the mass of the ion, z is the charge state of the ion, t is the flight time, e is the elementary charge, V_s is the acceleration potential, and L is the length of the flight

tube. In general, time, V_s and L are kept constant during analysis. According to this equation, the lower the m/z of the ion, the faster it will reach the detector.

There are two types of TOF analyzers: linear TOF analyzer and reflectron TOF analyzer. Due to initial energy distribution, when analyzed by a linear TOF analyzer, the ions of the same m/z value may reach the detector at different times, resulting in peak broadening and poor resolution. However, with the reflectron TOF analyzer, this initial energy distribution is compensated by using an ion mirror (Figure 1.10) which consists of successive sets of electric plates of increasing potential. The ion mirror can deflect the ions and reverse their flight direction. The fast ions penetrate deeper into the field and take a longer time to return than slow ions. Therefore, fast and slow ions with the same m/z are focused in time at the detector. The net effect is improved mass resolution, typically in the range of 10,000-20,000 with minimal losses in sensitivity.

1.4 The Present Work

HDX-MS has been applied to study a variety of non-covalent protein complexes, including protein-protein and multiprotein complexes,^{91,92} antibody-antigen,^{92,94} protein-peptide,⁹⁵ and protein-small molecule interactions.^{96,97} In most of the cases, it was used to monitor the ligand-induced changes in protein conformation and dynamics.^{60,62,98-101} To date, however, there have been few HDX-MS studies reported for protein-carbohydrate interactions.^{20,21,102} Therefore, the present work is focused on the application of HDX-MS assay to study protein-carbohydrate interactions.

The work described in chapter 2 focuses on the application of hydrogen/deuterium exchange mass spectrometry (HDX-MS) to localize ligand binding sites in carbohydrate-binding proteins. Proteins from three bacterial toxins, the B subunit homopentamers of Cholera toxin (CTB₅) and Shiga toxin type 1 (Stx1B₅), and a fragment of Clostridium difficile toxin A (TcdA-A2), and their interactions with native carbohydrate receptors, GM₁ pentasaccharide (GM₁-os, β -Gal-(1 \rightarrow 3)- β -GalNAc-(1 \rightarrow 4)[α -Neu5Ac-(2 \rightarrow 3)]- β -Gal-(1 \rightarrow 4)-Glc), Pk trisaccharide (Pk-OH, α -Gal-(1 \rightarrow 4)- β -Gal-(1 \rightarrow 4)-Glc) and CD-grease (α -Gal-(1 \rightarrow 3)- β -Gal-(1 \rightarrow 4)- β -GlcNAcO(CH₂)₈CO₂CH₃), respectively, served as model systems to test the reliability of using HDX-MS for localizing carbohydrate binding sites. In addition, we would also like to establish an HDX-MS kinetic model for protein-ligand interactions and identify potential pitfalls of using HDX-MS to investigate protein-carbohydrate interactions.

Chapter 3 describes using HDX-MS to explore the existence of distinct HMOs binding sites on bacterial toxins on the basis of the model system studies. Altogether, two toxins were studied, including CTB₅ and TcdA-A2, and their interactions with HMOs, 2'-fucosyllactose (2'-FL, α -L-Fuc(1 \rightarrow 2)- β -D-Gal(1 \rightarrow 4)- β -D-Glc) and lacto-N-tetraose (LNT, β -D-Gal(1 \rightarrow 3)- β -D-GlcNAc(1 \rightarrow 3)- β -D-Gal(1 \rightarrow 4)- β -D-Glc), respectively. Both direct binding of the HMO and competitive binding of the HMO with the native ligand of the toxin was carried out to help locate the HMO binding site.

In Chapter 4, a HDX-MS based method known as Protein-Ligand Interactions in solution by Mass Spectrometry, Titration and hydrogen/deuterium Exchange

(PLIMSTEX) carried out at peptide level was applied to CTB₅ and its interactions with native carbohydrate receptor GM₁-os, to test the reliability of using peptides as indicators to obtain the protein-carbohydrate binding affinities. A mathematical fitting was carried out to derive the apparent association constant from the HDX-MS data.

1.5 Literature Cited

- (1) Dwek, R. A., *Chem. Rev.*, **1996**, *96*, 683-720.
- (2) Ghazarian, H., Idoni, B., Oppenheimer, S. B., *Acta Histochem.*, **2011**, *113*, 236-247.
- (3) Vuignier, K., Schappler, J., Veuthey, J. L., Carrupt, P. A., Martel, S., *Anal. Bioanal. Chem.*, **2010**, *398*, 53-66.
- (4) Renaud, J. P., Delsuc, M. A., *Curr. Opin. Pharmacol.*, **2009**, *9*, 622-628.
- (5) Pacholarz, K. J., Garlish, R. A., Taylor, R. J., Barran, P. E., *Chem. Soc. Rev.*, **2012**, *41*, 4335-4355.
- (6) Zeng, X., Andrade, C. A. S., Oliveira, M. D. L., Sun, X. L., *Anal. Bioanal. Chem.*, **2012**, *402*, 3161-3176.
- (7) Lee, Y. C., Lee, R. T., *Accounts Chem. Res.*, **1995**, *28*, 321-327.
- (8) Holgersson, J., Gustafsson, A., Breimer, M. E., *Immunol. Cell Biol.*, **2005**, *83*, 694-708.
- (9) Kamiya, Y., Yagi-Utsumi, M., Yagi, H., Kato, K., *Current Pharm. Design*, **2011**, *17*, 1672-1684.
- (10) Schotte, F., Lim, M. H., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Wulff, M., Anfinrud, P. A., *Science*, **2003**, *300*, 1944-1947.

- (11) Fernandez-Alonso, M. D., Diaz, D., Berbis, M. A., Marcelo, F., Canada, J., Jimenez-Barbero, J., *Curr. Protein Pept. Sci.*, **2012**, *13*, 816-830.
- (12) Quioco, F. A., Spurlino, J. C., Rodseth, L. E., *Structure*, **1997**, *5*, 997-1015.
- (13) Duan, X. Q., Quioco, F. A., *Biochemistry*, **2002**, *41*, 706-712.
- (14) Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., Deng, S. J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L. P., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., Shewchuk, L. M., *Acta Crystallogr. Sect. D-Biol. Crystallogr.*, **2007**, *63*, 72-79.
- (15) Mittermaier, A., Kay, L. E., *Science*, **2006**, *312*, 224-228.
- (16) Sprangers, R., Kay, L. E., *Nature*, **2007**, *445*, 618-622.
- (17) Kohn, J. E., Afonine, P. V., Ruscio, J. Z., Adams, P. D., Head-Gordon, T., *PLoS Comput. Biol.*, **2010**, *6*, e1000911.
- (18) Weis, W. I., Drickamer, K., Hendrickson, W. A., *Nature*, **1992**, *360*, 127-134.
- (19) Ehring, H., *Anal. Biochem.*, **1999**, *267*, 252-259.
- (20) King, D., Lumpkin, M., Bergmann, C., Orlando, R., *Rapid Commun. Mass Spectrom.*, **2002**, *16*, 1569-1574.
- (21) Seyfried, N. T., Atwood, J. A., Yongye, A., Almond, A., Day, A. J., Orlando, R., Woods, R. J., *Rapid Commun. Mass Spectrom.*, **2007**, *21*, 121-131.
- (22) Baldwin, R. L., *Proteins*, **2011**, *79*, 2021-2026.
- (23) Hvidt, A., Linderstromlang, K., *Biochim. Biophys. Acta*, **1954**, *14*, 574-575.
- (24) Englander, S. W., Mayne, L., Bai, Y., Sosnick, T. R., *Protein Sci.*, **1997**, *6*, 1101-1109.

- (25) Skinner, J. J., Lim, W. K., Bedard, S., Black, B. E., Englander, S. W., *Protein Sci.*, **2012**, *21*, 987-995.
- (26) Skinner, J. J., Lim, W. K., Bedard, S., Black, B. E., Englander, S. W., *Protein Sci.*, **2012**, *21*, 996-1005.
- (27) Rosa, J. J., Richards, F. M., *J. Mol. Biol.*, **1979**, *133*, 399-416.
- (28) Brock, A., *Protein Expr. Purif.*, **2012**, *84*, 19-37.
- (29) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.
- (30) Katta, V., Chait, B. T., *Rapid Commun. Mass Spectrom.*, **1991**, *5*, 214-217.
- (31) Konermann, L., Tong, X., Pan, Y., *J. Mass Spectrom.*, **2008**, *43*, 1021-1036.
- (32) Konermann, L., Pan, J. X., Liu, Y. H., *Chem. Soc. Rev.*, **2011**, *40*, 1224-1234.
- (33) Wei, H., Mo, J. J., Tao, L., Russell, R. J., Tymiak, A. A., Chen, G. D., Iacob, R. E., Engen, J. R., *Drug Discov. Today*, **2014**, *19*, 95-102.
- (34) Jaswal, S. S., *BBA-Proteins Proteomics*, **2013**, *1834*, 1188-1201.
- (35) Wang, L., Smith, D. L. In *Current Protocols in Protein Science*, Coligan, J.E., John Wiley and Sons.: New York, **2002**, Vol. Chapter 17, p Unit 17.6.
- (36) Percy, A. J., Rey, M., Burns, K. M., Schriemer, D. C., *Anal. Chim. Acta*, **2012**, *721*, 7-21.
- (37) Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., Griffin, P. R., *Expert Rev. Proteomic*, **2011**, *8*, 43-59.
- (38) Iacob, R. E., Engen, J. R., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1003-1010.
- (39) Engen, J. R., *Anal. Chem.*, **2009**, *81*, 7870-7875.
- (40) Engen, J. R., Smith, D. L., *Anal. Chem.*, **2001**, *73*, 256A-265A.
- (41) Sowole, M. A., Konermann, L., *Anal. Chem.*, **2014**, *86*, 8.

- (42) Englander, S. W., Downer, N. W., Teitelba, H., *Annu. Rev. Biochem.*, **1972**, *41*, 903-924.
- (43) Eigen, M., *Angew. Chem.-Int. Edit.*, **1964**, *3*, 1-72.
- (44) Englander, S. W., Kallenbach, N. R., *Q. Rev. Biophys.*, **1983**, *16*, 521-655.
- (45) Molday, R. S., Englander, S. W., Kallenrg, *Biochemistry*, **1972**, *11*, 150-158.
- (46) Hvidt, A., *Comptes Rendus Des Travaux Du Laboratoire Carlsberg*, **1963**, *34*, 299-317.
- (47) Englander, S. W., Sosnick, T. R., Englander, J. J., Mayne, L., *Curr. Opin. Struct. Biol.*, **1996**, *6*, 18-23.
- (48) Bai, Y. W., Milne, J. S., Mayne, L., Englander, S. W., *Proteins*, **1993**, *17*, 75-86.
- (49) Smith, D. L., Deng, Y. Z., Zhang, Z. Q., *J. Mass Spectrom.*, **1997**, *32*, 135-146.
- (50) Truhlar, S. M. E., Croy, C. H., Torpey, J. W., Koeppe, J. R., Komives, E. A., *J. Am. Soc. Mass Spectrom.*, **2006**, *17*, 1490-1497.
- (51) Jane Clarke, L. S. I., *Curr. opin. struct. biol.*, **1998**, *8*, 112-118.
- (52) Lobanov, M. Y., Suvorina, M. Y., Dovidchenko, N. V., Sokolovskiy, I. V., Surin, A. K., Galzitskaya, O. V., *Bioinformatics*, **2013**, *29*, 1375-1381.
- (53) Ferraro, D. M., Robertson, A. D., *Biochemistry*, **2004**, *43*, 587-594.
- (54) Wales, T. E., Engen, J. R., *J. Mol. Biol.*, **2006**, *357*, 1592-1604.
- (55) Morgan, C. R., Engen, J. R. In *Current Protocols in Protein Science*, Coligan, J.E., John Wiley and Sons.: New York, **2009**, Vol. Chapter 17, p Unit 17.617.
- (56) Tsutsui, Y., Wintrode, P. L., *Curr. Med. Chem.*, **2007**, *14*, 2344-2358.
- (57) Karageorgos, I., Wales, T. E., Janero, D. R., Zvonok, N., Vemuri, V. K., Engen, J. R., Makriyannis, A., *Biochemistry*, **2013**, *52*, 5016-5026.

- (58) Lee, C. T., Graf, C., Mayer, F. J., Richter, S. M., Mayer, M. P., *Embo J.*, **2012**, *31*, 1518-1528.
- (59) Pan, L. Y., Salas-Solano, O., Valliere-Douglass, J. F., *Anal. Chem.*, **2014**, *86*, 2657-2664.
- (60) Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., Chen, G. D., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 498-504.
- (61) Zhang, H. M., McLoughlin, S. M., Frausto, S. D., Tang, H. L., Emmett, M. R., Marshall, A. G., *Anal. Chem.*, **2010**, *82*, 1450-1454.
- (62) Pan, Y., Brown, L., Konermann, L., *J. Am. Chem. Soc.*, **2011**, *133*, 20237-20244.
- (63) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (64) Palashoff, M. H., Chemistry Master's Theses, Northeastern University, **2008**.
- (65) Kan, Z. Y., Walters, B. T., Mayne, L., Englander, S. W., *P Natl. Acad. Sci. U.S.A.*, **2013**, *110*, 16438-16443.
- (66) Fajer, P. G., Bou-Assaf, G. M., Marshall, A. G., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1202-1208.
- (67) Ahn, J., Cao, M. J., Yu, Y. Q., Engen, J. R., *BBA-Proteins Proteomics*, **2013**, *1834*, 1222-1229.
- (68) Rey, M., Yang, M. L., Burns, K. M., Yu, Y. P., Lees-Miller, S. P., Schriemer, D. C., *Molecular & Cellular Proteomics*, **2013**, *12*, 464-472.
- (69) Cravello, L., Lascoux, D., Forest, E., *Rapid Commun. Mass Spectrom.*, **2003**, *17*, 2387-2393.
- (70) Wang, L. T., Pan, H., Smith, D. L., *Mol. Cell. Proteomics*, **2002**, *1*, 132-138.

- (71) Hamuro, Y., Coales, S. J., Molnar, K. S., Tuske, S. J., Morrow, J. A., *Rapid Commun. Mass Spectrom.*, **2008**, *22*, 1041-1046.
- (72) Sheff, J. G., Rey, M., Schriemer, D. C., *J. Am. Soc. Mass Spectrom.*, **2013**, *24*, 1006-1015.
- (73) Wu, Y., Engen, J. R., Hobbins, W. B., *J. Am. Soc. Mass Spectrom.*, **2006**, *17*, 163-167.
- (74) Wales, T. E., Fadgen, K. E., Gerhardt, G. C., Engen, J. R., *Anal. Chem.*, **2008**, *80*, 6815-6820.
- (75) Plumb, R. S., Johnson, K. A., Rainville, P., Smith, B. W., Wilson, I. D., Castro-Perez, J. M., Nicholson, J. K., *Rapid Commun. Mass Spectrom.*, **2006**, *20*, 1989-1994.
- (76) Chalmers, M. J., Busby, S. A., Pascal, B. D., He, Y. J., Hendrickson, C. L., Marshall, A. G., Griffin, P. R., *Anal. Chem.*, **2006**, *78*, 1005-1014.
- (77) Chalmers, M. J., Busby, S. A., Pascal, B. D., Southern, M. R., Griffin, P. R., *J. Biomol. Technol.*, **2007**, *18*, 194-204.
- (78) Burkitt, W., O'Connor, G., *Rapid Commun. Mass Spectrom.*, **2008**, *22*, 3893-3901.
- (79) Kebarle, P., Tang, L., *Anal. Chem.*, **1993**, *65*, A972-A986.
- (80) Wu, X. Y., Oleschuk, R. D., Cann, N. M., *Analyst*, **2012**, *137*, 4150-4161.
- (81) Thomson, B. A., Iribarne, J. V., *J. Chem. Phys.*, **1979**, *71*, 4451-4463.
- (82) Dole, M., Mack, L. L., Hines, R. L., *J. Chem. Phys.*, **1968**, *49*, 2240-2249.
- (83) Ahadi, E., Konermann, L., *J. Phys. Chem. B*, **2012**, *116*, 104-112.
- (84) Konermann, L., Ahadi, E., Rodriguez, A. D., Vahidi, S., *Anal. Chem.*, **2013**, *85*, 2-9.

- (85) Konermann, L., Vahidi, S., Sowole, M. A., *Anal. Chem.*, **2014**, *86*, 213-232.
- (86) Iribarne, J. V., Thomson, B. A., *J. Chem. Phys.*, **1976**, *64*, 2287-2294.
- (87) De Hoffmann, E., Stroobant, V., *Mass Spectrometry: Principles and Applications*, 3 ed., John Wiley & Sons.: New York, **2007**.
- (88) Miller, P. E., Denton, M. B., *J. Chem. Educ.*, **1986**, *63*, 617-622.
- (89) Wollnik, H., *Mass Spectrom. Rev.*, **1993**, *12*, 89-114.
- (90) El-Aneed, A., Cohen, A., Banoub, J., *Appl. Spectrosc. Rev.*, **2009**, *44*, 210-230.
- (91) Huang, R. Y. C., Wen, J., Blankenship, R. E., Gross, M. L., *Biochemistry*, **2012**, *51*, 187-193.
- (92) Chetty, P. S., Mayne, L., Kan, Z. Y., Lund-Katz, S., Englander, S. W., Phillips, M. C., *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, *109*, 11687-11692.
- (93) Ahn, J., Engen, J. R., *Chim. Oggi-Chem. Today*, **2013**, *31*, 25-28.
- (94) Clementi, N., Mancini, N., Criscuolo, E., Cappelletti, F., Clementi, M., Burioni, R. In *Monoclonal Antibodies: Methods and Protocols*, 2 ed., Ossipow, V., Fischer, N., Humana Press Inc: Totowa, **2014**, Vol. 1131, p 427-446.
- (95) Jorgensen, T. J. D., Gardsvoll, H., Dano, K., Roepstorff, P., Ploug, M., *Biochemistry*, **2004**, *43*, 15044-15057.
- (96) Dai, S. Y., Burris, T. P., Dodge, J. A., Montrose-Rafizadeh, C., Wang, Y., Pascal, B. D., Chalmers, M. J., Griffin, P. R., *Biochemistry*, **2009**, *48*, 9668-9676.
- (97) Hamuro, Y., Coales, S. J., Morrow, J. A., Molnar, K. S., Tuske, S. J., Southern, M. R., Griffin, P. R., *Protein Sci.*, **2006**, *15*, 1883-1892.
- (98) Pan, Y., Piyadasa, H., O'Neil, J. D., Konermann, L., *J. Mol. Biol.*, **2012**, *416*, 400-413.

- (99) Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Garcia-Ordonez, R. D., Pascal, B. D., Burris, T. P., Dodge, J. A., Griffin, P. R., *Structure*, **2010**, *18*, 1332-1341.
- (100) Sperry, J. B., Huang, R. Y. C., Zhu, M. M., Rempel, D. L., Gross, M. L., *Int. J. Mass Spectrom.*, **2011**, *302*, 85-92.
- (101) Zhang, X., Chien, E. Y. T., Chalmers, M. J., Pascal, B. D., Gatchalian, J., Stevens, R. C., Griffin, P. R., *Anal. Chem.*, **2010**, *82*, 1100-1108.
- (102) King, D., Bergmann, C., Orlando, R., Benen, J. A. E., Kester, H. C. M., Visser, J., *Biochemistry*, **2002**, *41*, 10225-10233.

Chapter 2

Localizing Carbohydrate Binding Sites in Proteins Using Hydrogen/Deuterium Exchange Mass Spectrometry

2.1 Introduction

Protein-carbohydrate interactions are implicated in a wide range of cellular processes, including cell-cell and cell-matrix interactions, signal transduction, inflammation, cancer metastasis, bacterial and viral infections and the immune response.¹ Elucidating the structures of protein-carbohydrate complexes, as well as the kinetics and thermodynamics of the interactions, is vital to a complete understanding of many physiological and pathological cellular processes and can guide drug discovery and design efforts.²⁻⁴ High resolution techniques, such as X-ray crystallography and nuclear magnetic resonance spectroscopy, have been extensively used to elucidate the three dimensional structures of protein-carbohydrate complexes.⁵⁻⁷ However, many such interactions are not amenable to these techniques due to limitations associated with protein size, solubility or ease of crystallization, as well as the cost and availability of pure oligosaccharide ligand.⁸ Consequently, there is a need for alternative structural techniques capable of probing protein-carbohydrate interactions.

Recently, hydrogen/deuterium exchange mass spectrometry (HDX-MS) has emerged as a promising method for characterizing the interactions between proteins and their ligands.⁹⁻¹³ When exposed to a deuterated solvent (usually D₂O), the acidic hydrogen (H) atoms of the protein undergo exchange with the deuterium (D) atoms of

the solvent. The rate of the exchange reaction for a specific functional group depends on a number of factors - pK_a (which reflects the chemical nature of the functional group and the formation of intra- or intermolecular H-bonds, as well as solvent accessibility), pH and temperature.¹⁰ Exchange rates of acidic H's associated with heteroatoms of amino acid side chains are too fast to be reliably measured. In contrast, the exchange rates for amide H's are much slower (ms – yr).^{14,15} The “global” amide exchange rates of proteins and protein-ligand complexes can be established through time-resolved electrospray ionization (ESI)-MS analysis.¹⁶ To establish exchange rates associated with specific amino acids (or peptides), the exchange reaction is quenched (by acidifying the solution and reducing the temperature) and the protein is digested with a protease, usually pepsin, and the resulting peptides analyzed by MS.^{9,13} The extent of deuterium uptake (D-uptake) of each peptic peptide reflects the average deuterium exchange rate of the amide H's (with the exception of the one at the N-terminus which experiences fast back exchange rate) within that peptide. When a ligand binds to the protein some of the amide H's may become (more) protected against exchange (due to the formation of new or stronger inter- or intramolecular H-bonds or a reduction in solvent accessibility), resulting in decreased rates of exchange for the peptides containing these groups.^{10,17} Consequently, comparison of the D-uptake for peptides produced from the ligand-bound and unbound forms of the protein can, in principle, reveal the residues involved in ligand binding.¹⁸

HDX-MS has been applied to a variety of non-covalent protein complexes, including protein-protein and multiprotein complexes,^{16,19} antibody-antigen,²⁰

protein-peptide,²¹ and protein-small molecule complex.^{22,23} Additionally, it has been used to study ligand-induced changes in protein conformation and dynamics.^{18,24-26} To date, however, there have been few HDX-MS studies reported for protein-carbohydrate interactions.²⁷⁻²⁹ One possible reason for this is the low affinities that are typical of protein-carbohydrate interactions (association constants (K_a) of $\sim 10^3$ M⁻¹).³ In order to obtain detectable differences in peptide D-uptake, high ligand occupancy (approaching saturation of the binding site) is needed. A large and sometimes prohibitive amount of carbohydrate ligand is required to achieve this condition in the case of low affinity interactions. Another consideration is the relatively small size of many carbohydrate ligands, typically mono-, di- or trisaccharide. Binding of small ligands, which form few intermolecular interactions, generally affords protection to only a few residues in the protein. Moreover, binding is usually dominated by intermolecular interactions involving amino acid side chains, rather than the peptide backbone, which may offer limited protection to the amide H's. Finally, carbohydrate binding can be accompanied by changes in protein structure and dynamics. These changes may lead to a decrease or increase in D-uptake of residues remote from the carbohydrate binding site.^{29,30} Consequently, the differences in peptide D-uptake may reflect the formation of intermolecular interactions or ligand binding-induced changes in protein conformation or dynamics, or a combination of these effects.^{9,10} It has been proposed that docking simulations or site-directed mutagenesis, in combination with HDX-MS may help to localize the ligand binding site.^{18,30} Alternative strategies to distinguish direct ligand protection from ligand-induced changes in protein dynamics or conformation involve comparing

the HDX-MS profiles measured for the protein binding to a homologous series of ligands³¹ or through the use of a second ligand, which binds at an alternative site, to suppress allosteric effects.³²

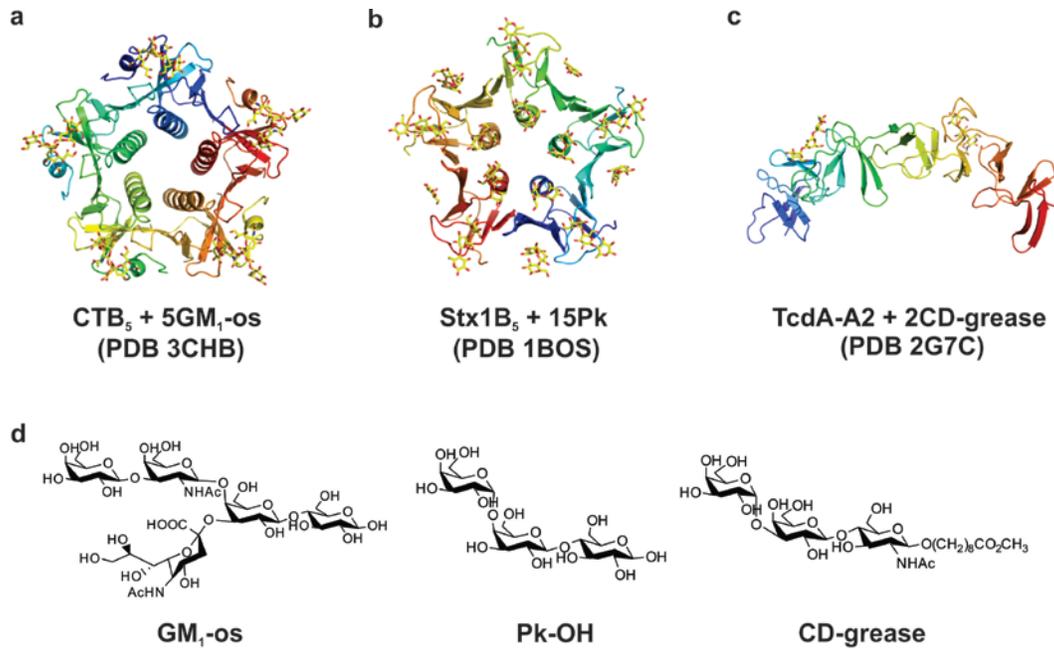


Figure 2.1 Protein-carbohydrate complexes considered for HDX-MS analysis: (a) (CTB₅ + 5GM₁-os) complex (each subunit has one GM₁-os binding site), (b) (Stx1B₅ + 15Pk) complex (each subunit has three Pk binding sites) and (c) (TcdA-A2 + 2CD-grease) complex (each TcdA-A2 has two CD-grease binding sites). (d) The structures of the carbohydrate ligands, GM₁-os, Pk-OH and CD-grease.

Here, HDX-MS was applied to proteins from three bacterial toxins, the B subunit homopentamers of Cholera toxin (CTB₅) and Shiga toxin 1 (Stx1B₅) and a fragment of *Clostridium difficile* toxin A (TcdA-A2), and their interactions with native carbohydrate receptors, GM₁ pentasaccharide (GM₁-os, β -Gal-(1 \rightarrow 3)- β -GalNAc-(1 \rightarrow 4)[α -Neu5Ac-(2 \rightarrow 3)]- β -Gal-(1 \rightarrow 4)-Glc), Pk trisaccharide (Pk-OH, α -Gal-(1 \rightarrow 4)- β -Gal-(1 \rightarrow 4)-Glc), which is the oligosaccharide of globotriose Gb₃, and CD-grease (α -Gal-(1 \rightarrow 3)- β -Gal-(1 \rightarrow 4)- β -GlcNAcO(CH₂)₈CO₂CH₃), respectively, to test the reliability of the method for localizing carbohydrate binding sites. Crystal structures have been reported for all three protein-carbohydrate complexes (Figure 2.1). It is known that CTB₅ has five GM₁-os binding sites and the step-wise association constants range from 10⁶ to 10⁷ M⁻¹.³³ According to the crystal structure reported for the (CTB₅ + 5GM₁-os) complex (PDB 3CHB), each GM₁-os interacts primarily with a single subunit, with the binding site formed by three loops (loop 1-loop 3) from the same subunit and a fourth loop (loop 4) containing residues from the adjacent subunit (Figure 2.1a).³⁴ Although it is known that the stepwise binding of GM₁-os to CTB₅ exhibits positive cooperativity,³³ no obvious protein conformational change is detectable (based on X-ray crystallography) upon ligand binding.³⁴ Stx1B₅ is structurally similar to CTB₅. According to the reported crystal structure (PDB 1BOS), each subunit of Stx1B₅ can bind up to three Pk trisaccharide ligands (Figure 2.1b). The three binding sites (referred to as *site 1*, *site 2* and *site 3*) are located on the same face of the homopentamer.³⁵ *Site 1* is composed of residues within a single subunit, while *site 2* and *site 3* also contain residues from adjacent subunits. According to available binding data, the three binding sites are independent and non-

equivalent, with Pk binding preferentially to *site 2*, although with very low affinity, $\sim 10^3 \text{ M}^{-1}$.^{35,36} Based on crystallographic data available for free (PDB 2XSC) and Pk-bound Stx1B₅, no significant conformational change occurs upon Pk binding.³⁷ The TcdA-A2 fragment is from the C-terminal repetitive domain of TcdA, which contains nine short repeats separated by two long repeats.³⁸ Within the TcdA-A2 fragment, a long repeat and the following short repeat create a shallow carbohydrate binding site, and two carbohydrate binding sites are present in the fragment (PDB 2G7C).³⁸ Previous studies showed that TcdA-A2 displays a low affinity for CD-grease, with an apparent association constant of $\sim 500 \text{ M}^{-1}$ at 25 °C, and the two binding sites exhibiting similar affinities.³⁹

2.2 Experimental Methods

2.2.1 Materials

CTB₅ (MW 58,020 Da) was purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). A stock solution (60 μM) of CTB₅ was prepared by dissolving in ultrafiltered water (Milli-Q; Millipore, Billerica, MA, USA) and stored at 4 °C until needed. Stx1B₅ (MW 38,450 Da) was a gift from Prof. G. Armstrong (Univ. of Calgary) as a stock solution prepared in 0.05 M *Tris* buffer (pH 7.5). TcdA-A2 (MW 29,575 Da) was expressed and purified as previously described [38]. The TcdA-A2 stock solution was at a concentration of 57.5 μM in 60 mM imidazole (pH 7.0), 150 mM NaCl and 50 g L⁻¹ glycerol (>99.5% purity). Both Stx1B₅ and TcdA-A2 stock solutions were stored at -80 °C until needed. Prior to analysis, the protein solutions were diluted with Milli-Q water to the desired concentrations. GM₁-os (β -Gal-(1 \rightarrow 3)-

β -GalNAc-(1 \rightarrow 4)[α -Neu5Ac-(2 \rightarrow 3)]- β -Gal-(1 \rightarrow 4)-Glc, MW 998.9 Da), and Pk-OH (α -Gal-(1 \rightarrow 4)- β -Gal-(1 \rightarrow 4)-Glc, MW 504.4 Da) were purchased from Elicityl SA (Crolles, France). CD-grease (α -Gal-(1 \rightarrow 3)- β -Gal-(1 \rightarrow 4)- β -GlcNAcO(CH₂)₈CO₂CH₃, MW 715.7 Da) was a gift from Prof. D. Bundle (Univ. of Alberta). Stock solutions of each of the carbohydrates were prepared by dissolving a known amount of solid in Milli-Q water to give a final concentration of 1 mM (GM₁-os) or 0.4 M (Pk-OH and CD-grease); the stock solutions were stored at -20 °C until needed.

2.2.2 HDX-MS

The HDX-MS experiments were carried out using a Synapt G2-S HDMS mass spectrometer equipped with a nanoACQUITY UPLC system with HDX technology (Waters, UK), and a PAL HTX-*xt* automatic sample preparation and injection system. Two sample stacks in the PAL system provided accurate temperature control for the labeling reactions (20 °C), and for the reactions quench (1 °C), respectively. Protein solutions (4 μ M, 6 μ M and 12 μ M for CTB₅, Stx1B₅ and TcdA-A2, respectively) alone or in the presence of excess ligand (0.2 mM GM₁-os, 0.2 M Pk-OH and 0.2 M CD-grease) were diluted 15-fold with either equilibrium buffer (10 mM potassium phosphate in H₂O at pH 7.0) for control experiments, or labeling buffer (10 mM potassium phosphate in D₂O at pD 7.0) for labeling experiments. For the labeling experiments, diluted samples were incubated at 20 °C for time intervals of 1, 5 and 10 min. After that, samples (both for control and labeling experiments) were quenched with quench buffer (4 M guanidine hydrochloride and 0.5 M *tris*(2-carboxyethyl)phosphine (TCEP) in H₂O at pH 2.6 for

CTB₅ and Stx1B₅, and 4 M guanidine hydrochloride in 100 mM potassium phosphate in H₂O at pH 2.6 for TcdA-A2) using a 1:1 dilution ratio at 1 °C. Quenched samples were incubated for 30 s prior to injection into a 50 µL injection loop of a nanoACQUITY UPLC system with HDX technology. Online digestion was performed using an immobilized pepsin column (Life Technologies, Burlington, Canada) with 0.1% formic acid in H₂O at a flow rate of 200 µL min⁻¹ at 20 °C. Peptic peptides were trapped online using an ACQUITY UPLC BEH C18 1.7 µm VanGuard Pre-column at 1 °C and desalted for 2 min. Peptide separation was carried using an ACQUITY UPLC C18 1.7 µm 1.0×100mm column with a 12 min gradient elution at a flow rate of 40 µL min⁻¹. The content of solvent A (Solvent A, 0.1% formic acid and 5% acetonitrile in H₂O; solvent B, 0.1% formic acid in acetonitrile) in the mobile phase was decreased over a 7 min period from 95% to 63% and held constant for 1 min before a further reduction from 63% to 16% over a 0.5 min period. After 0.5 min at 16% for 0.5 min, solvent A was increased back to 95% over a 0.5 min period. The eluent was introduced to the Synapt G2-S HDMS using the ESI source. Mass spectra were acquired in MS^E mode from m/z 50 to 2000 with a scan rate of 0.4 s scan⁻¹ and lock-mass correction (using [Glu]-Fibrinopeptide). The capillary and cone voltages were kept constant at 3 kV and 40 V, respectively. At a given labeling time (*t*), the free and ligand-bound protein samples were analyzed back-to-back, with a blank sample (water) in between, to avoid the effects of sample carry-over.

2.2.3 Data analysis

ProteinLynx Global Server 2.5.2 software (PLGS, Waters) was used to identify peptic peptides produced for each protein, in the absence and presence of the ligand, prior to HDX (*i.e.*, at $t = 0$ min). To minimize redundancy, the DynamX 2.0 software (Waters) was used to generate a smaller peptide list consisting of peptides that were detected in all replicate measurements and that provided maximum sequence coverage. The average MW for each peptide was calculated by DynamX using the centroid of the entire envelope of the corresponding isotopic peaks; for each peptide one or more than one charge states were considered. The absolute D-uptake value (D_i , units of Da) for peptide i was determined as the difference of the MW measured for this peptide in labeled and control (no exchange) samples, for all labeling times. The relative D-uptake (ΔD_i , units of Da) was calculated as the difference in the D_i values for peptide i measured in the absence and presence of ligand. Because the extent of back exchange (during protein digestion and LC separation^{14,40}) was expected to be the same for peptides produced in the absence and presence of ligand (under identical experimental conditions), no correction for back exchange was carried out. To compare the HDX rates for peptides produced in the absence and presence of ligand the ΔD_i values were summed over all labeling times.¹³ Errors were calculated as the standard deviation for triplicate measurements and the ΔD_i values were considered significant if the values were greater than three times the standard deviation.

2.3 Results and Discussion

Time-resolved HDX-MS measurements were performed on the free and ligand-bound forms of CTB₅, Stx1B₅ and TcdA-A2. A summary of the results obtained for each protein-carbohydrate interaction is given below.

2.3.1 CTB₅ and its interaction with GM₁-os

Pepsin digestion of CTB₅ under denaturing conditions produced approximately eighty different peptides per analysis, of which forty were identified in all measurements. In order to decrease redundancy (while maintaining a high sequence coverage), ΔD_i values of only fourteen of these reproducible peptides, covering 96.1% of CTB monomer sequence, were considered (Figure 2.2a). It can be seen that, in the presence of GM₁-os, seven of the peptides (4-15, 27-38, 41-55, 49-66, 57-72, 83-94 and 80-103) exhibited a significant decrease in D_i values (leading to positive ΔD_i values ranging from 1.1 to 4.6 Da). In Figure 2.2b the ΔD_i values are mapped onto the structure of the (CTB₅ + 5GM₁-os) complex for one subunit, with the peptides exhibiting positive ΔD_i values highlighted in red. Notably, the peptides with positive ΔD_i values are from the four loops that make up the binding site for GM₁-os.^{41,42}

Table 2.1 Summary of the putative intermolecular H-bonds identified in the crystal structures of the three model protein-carbohydrate complexes: (CTB₅ + 5GM_{1-os}), PDB 3CHB; (Stx1B₅ + 15Pk), PDB 1BOS; and (TcdA-A2 + 2CD-grease), PDB 2G7C.

Protein-carbohydrate complex	Amino acid residues that participate in H-bonds with the carbohydrate ligands^a
(CTB₅ + 5GM_{1-os})	E11(O), H13(N, ND1), G33 (N), ^{b,c} E51(OE1), Q56(O), Q61(NE2), N90(OD1, ND2) and K91(NZ)
(Stx1B₅ + 15Pk)	Site 1: D17(OD2), T21(OG1), E28(OE2), G60(N, O) Site 2: D16(OD2), ^c N32(N, OD1), R33(NH2, NE), N55(N, OD1) and F63(N, O) Site 3: D18(OD1, OD2), ^c W34(N) and N35(N)
(TcdA-A2 + 2CD-grease)	Site 1: D92(OD2), Q99(NE2), R102(N), N119(O), S121(OG) and K122(NZ) Site 2: D183(OD2), Q190(NE2), R193(N), N210(O), S212(OG) and K213(NZ)

a. Specific atoms of the amino acid residue that participate in the intermolecular H-bonds. b. Water mediated H-bond between the specified residue and ligand. c. H-bond involves residue from an adjacent subunit.

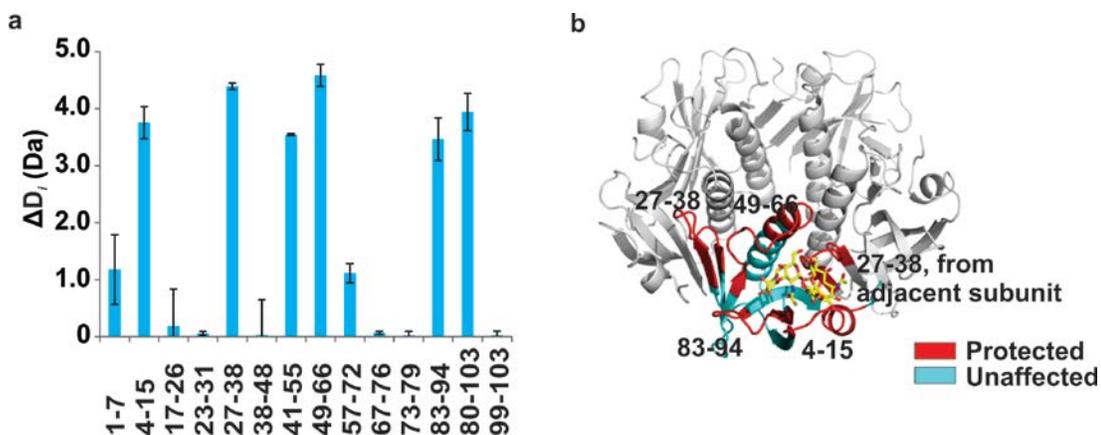


Figure 2.2 (a) Difference plot of peptide level D-uptake between free CTB₅ and (CTB₅ + 5GM₁-os) complex for a subset of peptides that cover 96.1% of CTB monomer sequence. The x-axis indicates the position (residue number) of the peptides considered. The y-axis shows the ΔD_i values for each peptide, summed over all of the labeling times. The errors bars correspond to one standard deviation. (b) Cartoon illustration of ΔD_i values mapped onto the crystal structure of the (CTB₅ + 5GM₁-os) complex (PDB 3CHB). The colored region corresponds to one representative CT B subunit. The regions highlighted in red (with corresponding residue numbers indicated) exhibited protection from deuterium exchange upon ligand binding, while the regions highlighted in cyan were unaffected by ligand binding.

According to the crystal structure, GM₁-os forms direct intermolecular H-bonds with seven residues located within loops 1-3, namely residues E11, H13, E51, Q56, Q61, N90 and K91 (Table 2.1). With the exception of H13, all of these residues

interact with the ligand through their side chains or backbone carbonyl oxygens, rather than amide H's. Importantly, the present results indicate that the protein-ligand intermolecular interactions involving amino acid side chains can slow down the exchange rate of the associated backbone amide H's of the peptides containing that residue. The protection may originate from reduced solvent accessibility or changes in the local dynamics of the protein. According to the reported crystal structure, there exists a solvent-mediated H-bond between GM₁-os and the backbone amide H of residue G33 from the adjacent subunit. It is interesting to note that peptide 27-38 exhibited a positive ΔD_i (Figure 2.2b), suggesting that solvent mediated H-bonds involving amide H's can also influence the rate of exchange.⁴³

2.3.2 Stx1B₅ and its interaction with Pk-OH

Pepsin digestion of Stx1B₅ produced approximately seventy different peptides, of which thirty-four were identified in every analysis. Since the Stx1B subunit is a small protein, consisting of only 69 residues, there was significant overlap amongst the proteolytic peptides and the D-uptake of only seven of these, which provided 95.6% sequence coverage, was considered. The corresponding ΔD_i values measured for these seven peptides are shown in Figure 2.3a. Upon binding of Pk-OH, peptides 12-20 (combining results from peptide 1-20 and 12-20), 30-40, and 52-66 exhibited modest protection, with ΔD_i values ranging from 0.6 Da to 0.7 Da. For peptides 40-48, 65-69, the ΔD_i values are smaller (0.2 and 0.3 Da, respectively), suggesting little change in protection upon ligand binding. In Figure 2.3b, the peptides with significant ΔD_i values are highlighted in red for one B subunit in the

(Stx1B₅ + 15Pk) complex. As noted above, each Stx1B subunit has three Pk binding sites located on one the face of the homopentamer [33]. The residues that make up the binding sites are present in either a single subunit (*site 1*) or two adjacent subunits (*sites 2* and *3*). Consistent with the reported crystal structure, the HDX-MS results reveal that the peptides exhibiting the greatest protection against exchange upon ligand binding are located within ligand binding sites, and contain residues involved in direct H-bonds (through side chains or amide groups) with the ligand (Table 2.1). Because three of the peptides exhibiting protection (namely 12-20, 30-40 and 52-69) include amino acid residues that belong to all three sites (*site 1*: D17 and G60, *site 2*: D16, N32, R33, N55 and F63 and *site 3*: D18, W34 and N35), it is not possible to delineate the contribution of ligand binding to each of these sites. However, it is notable that for peptide 20-31, which contains residues T21 and E28 and form direct H-bonds with Pk in *site 1*, no protection was observed upon ligand binding. Examination of the crystal structure reveals that these two residues are located in a β -sheet and interact with ligand through their side chains. It is possible that the absence of protection reflects the fact that these amide H's are located in structured regions and, therefore, already experience a certain degree of protection against exchange. Consequently, the influence of ligand binding on the rates of exchange may not be pronounced, such that no significant difference in D-uptake is observed.

However, it is also possible that the absence of protection for this peptide reflects the fact that no ligand binding occurs at this site. It has been suggested, based on solution NMR data,⁴⁴ that the Pk trisaccharide ligand only binds to *site 2*. Since residues T21 and E28 belong to *site 1* (Table 2.1), the absence of protection can be

explained by an absence of binding or, at least, very low ligand occupancy of this binding site at the ligand concentration used.

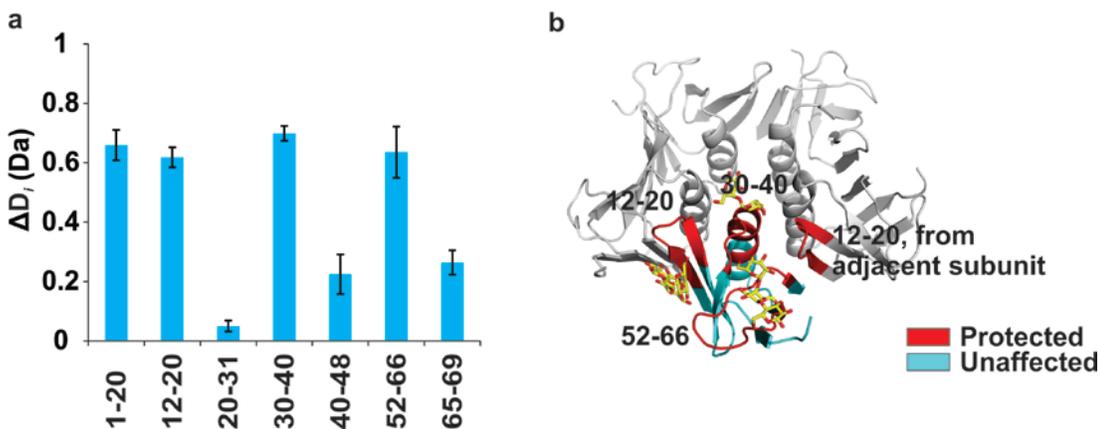


Figure 2.3 (a) Difference plot of peptide level D-uptake between free Stx1B₅ and (Stx1B₅ + 15Pk-OH) complex for a subset of peptides that cover 95.6% of Stx1B monomer sequence. The x-axis indicates the position (residue number) of the peptides considered. The y-axis shows the ΔD_i values for each peptide, summed over all of the labeling times. The errors bars correspond to one standard deviation. (b) Cartoon illustration of ΔD_i values mapped onto the crystal structure of the (Stx1B₅ + 15Pk) complex (PDB 1BOS). The colored region corresponds to one representative Stx1 B subunit. The regions highlighted in red (with corresponding residue numbers indicated) exhibited protection from deuterium exchange upon ligand binding, while the regions highlighted in cyan were unaffected by ligand binding.

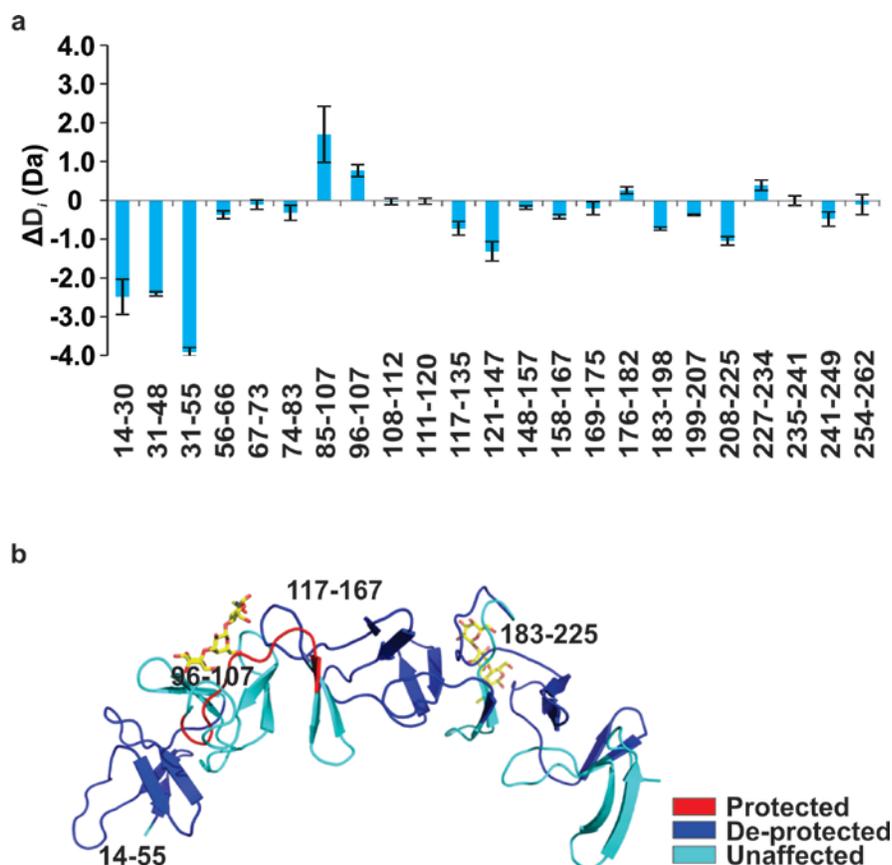


Figure 2.4 (a) Difference plot of peptide level D-uptake between free TcdA-A2 and (TcdA-A2 + 2CD-grease) complex for a subset of peptides that cover 92.4% of the TcdA-A2 sequence. The x-axis indicates the position (residue number) of the peptides considered. The y-axis shows the ΔD_i values for each peptide, summed over all of the labeling times. The errors bars correspond to one standard deviation. (b) Cartoon illustration of ΔD_i values mapped onto the crystal structure of the (TcdA-A2 + 2CD-grease) complex (PDB 2G7C). The regions highlighted in red and dark blue exhibited protection or de-protection, respectively, against deuterium exchange upon ligand binding; the regions shown in cyan were unaffected by ligand binding.

2.3.3 TcdA-A2 and its interaction with CD-grease

Pepsin digestion of TcdA-A2 produced approximately a hundred different peptides, fifty-four of which were identified in every analysis carried out. The D-uptake for twenty-three of these peptides, covering 92.4% of the TcdA-A2 sequence, was considered. The ΔD_i values measured for these peptides are shown in Figure 2.4a. Unlike the two other protein-carbohydrate complexes considered in this work, ligand binding led to both regions of increased and decreased protection. Notably, none of the peptides associated with *site 2* exhibited protection in the presence of CD-grease. Also, peptides 111-120, 117-135 and 121-147, which contain residues (N119, S121 and K122) that interaction with the ligand in *site 1* exhibited no protection. However, the two overlapping peptides 85-107 and 96-107 did exhibit positive ΔD_i values. These two peptides contain three residues (D92, Q99 and R102) that are predicted, based on the crystal structure, to interact directly with CD-grease in *site 1*. This result is consistent with ligand binding to *site 1* under the experimental conditions investigated. A number of the peptic peptides (14-30, 31-48, 31-55, 117-135, 121-147, 148-157, 158-167, 183-198, 199-207 and 208-225) exhibited negative ΔD_i values (as large as -3.9 Da), indicating a significant increase in D-uptake upon CD-grease binding. In Figure 2.4b, the ΔD_i values are mapped onto the crystal structure of the (TcdA-A2 + 2CD-grease) complex and highlighted in red (protected) and dark blue (de-protected). It can be seen that de-protected regions are mainly located close to the N-terminus, which is the artificial truncation point of the fragment.³⁹ These data suggest that changes in the structure or dynamics of the N-terminal region of the TcdA-A2 fragment occur upon binding of CD-grease.

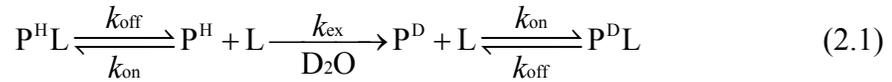
Overall, these results indicate that the binding of CD-grease to the two crystallographically-identified binding sites in TcdA-A2 does not correlate with a simple pattern of protection in all of the peptides where intermolecular H-bonds are observed in the crystal structure. This is, perhaps, not surprising given that only a single direct H-bond is formed between the bound CD-grease and backbone amide groups in each of the two binding sites. Instead, indirect effects on exchange due to changes in protein conformation and dynamics seem to play a larger role in the observed pattern of deuterium exchange.

2.3.4 Influence of protein-carbohydrate interactions on deuterium exchange rates

Several conclusions regarding the relationship between the nature of protein-carbohydrate interactions and the measured changes in the exchange rates can be drawn from the results of this study. First, intermolecular H-bonds between the ligand and backbone amides generally lead to a decrease in deuterium uptake. Protection is also conferred by water-mediated H-bonds involving backbone amides. For example, for the CTB₅ interaction with GM₁-os, positive ΔD_i values were observed for peptides containing residues H13 (direct H-bond with ligand) and G33 (water mediated H-bond). In contrast, putative intermolecular H-bonds involving backbone carbonyl oxygens or side chains did not always lead to protection. For example, residues D16, D17, D18, T21 and E28 of Stx1B₅ participate in direct H-bonds with Pk through their side chains (Table 2.1). However, based on HDX-MS results, only residues D16, D17 and D18 in peptide 12-20 were protected, while T21 and E28, which are contained in

peptide 20-31, were unaffected by ligand binding. It should be emphasized that this analysis is predicated on the assumption that all of the carbohydrate binding sites identified by X-ray crystallography are occupied under the solution conditions investigated.

The present results also suggest that the extent of protection (*i.e.*, the magnitude of the ΔD_i values) correlates with the affinity of the protein-carbohydrate interaction, with larger ΔD_i values observed for higher affinity interactions. This observation can be rationalized by considering a simple kinetic model to describe the exchange process,⁹ Equation 2.1:



where P^D and P^H are the deuterated and non-deuterated forms of the free protein, respectively, and $P^D L$ and $P^H L$ are the deuterated and non-deuterated forms of the ligand-bound protein, respectively, L is the ligand and k_{off} , k_{on} and k_{ex} are the dissociation, association and deuterium exchange rate constants, respectively. An important underlying assumption of this model is that exchange only occurs in absence of ligand binding. In other words, the amide H's in the protein-ligand complex are assumed to be unexchangeable. On the basis of Equation 2.1, the rate of exchange (v_{HDX}), which is equal to the rate of consumption of “unexchanged” species (*i.e.*, P^H and $P^H L$), can be written as, Equation 2.2:

$$v_{HDX} = -d[\text{unexchanged}]/dt = k_{ex} [P^H] \quad (2.2)$$

and the rate of consumption of P^H can be expressed by Equation 2.3:

$$-d[P^H]/dt = k_{on} [P^H][L] - k_{off} [P^H L] + k_{ex} [P^H] \quad (2.3)$$

Assuming that the on-off kinetics for the protein-ligand complex are much faster than the rate of exchange, such that the steady-state approximation can be applied to $[P^H]$, $[P^HL]$ can be written in terms of $[P^H]$ (Equation 2.4):

$$[P^HL] = [P^H] (k_{on} [L] + k_{ex})/k_{off} \quad (2.4)$$

It follows that the concentration of unexchanged species can be described by Equation 2.5:

$$[\text{unexchanged}] = [P^H] + [P^HL] = [P^H] (k_{off} + k_{on}[L] + k_{ex})/k_{off} \quad (2.5)$$

Combining Equation 2.2 and 2.5 gives Equation 2.6:

$$v_{HDX} = -d[\text{unexchanged}]/dt = k_{ex}k_{off} [\text{unexchanged}]/(k_{off} + k_{on}[L] + k_{ex}) \quad (2.6)$$

and

$$k_{HDX} = k_{ex}k_{off}/(k_{off} + k_{on} [L] + k_{ex}) = k_{ex}K_D/(K_D + [L] + k_{ex}/k_{on}) \quad (2.7a)$$

where k_{HDX} is the *apparent* exchange rate constant, and K_D is the dissociation equilibrium constant. In order to saturate the binding sites, the ligand is present in large excess. In such cases, $K_D \ll [L] \approx [L]_o$ (initial concentration of ligand), Equation 2.7a can be written as Equation 2.7b:

$$k_{HDX} \approx k_{ex}K_D/([L]_o + k_{ex}/k_{on}) \quad (2.7b)$$

Furthermore, when $k_{ex}/k_{on} \ll [L]$, the rate of exchange can be further simplified, Equation 2.7c:

$$k_{HDX} \approx k_{ex}K_D/[L]_o \quad (2.7c)$$

This simple analysis predicts that the rate of exchange will decrease with increasing affinity (*i.e.*, decreasing K_D). Consequently, in the absence of other effects, larger ΔD_i values are anticipated for higher affinity interactions, which is consistent with the experimental observations.

2.4 Conclusions

In summary, the application of HDX-MS to localize ligand binding sites for three model carbohydrate-binding proteins is described. Comparison of the differences in D-uptake for peptic peptides produced in the absence and presence of native carbohydrate ligand revealed regions of the protein that are protected against deuterium exchange upon ligand binding. For all three proteins, the peptides exhibiting protection contain residues known to make up the carbohydrate binding sites, as identified by X-ray crystallography. For the interaction between CTB₅ and GM₁-os, peptides associated with each of the four loops in the CTB monomer that make up the GM₁ binding site were found to exhibit protection against exchange upon ligand binding. For the interaction between Stx1B₅ and Pk-OH, peptides containing residues associated with each of the three ligand binding sites were also found to exhibit protection in the presence of ligand. However, because the peptides exhibiting protection include amino acid residues that belong to all three sites, it was not possible to establish unambiguously whether all three sites were occupied under the experimental conditions investigated. For the interaction between TcdA-A2 and CD-grease, one (*site 1*) of the two known binding sites was identified based on the observation of protection against exchange in the presence of ligand. However, ligand binding was also found to induce changes in either the structure or the dynamics of the protein that resulted in significant de-protection of the peptic peptides associated with the N-terminus of the protein. Taken together, the results of this study suggest that HDX-MS can serve as a useful tool for localizing the ligand binding sites in carbohydrate-binding proteins. However, a detailed interpretation of the changes in

deuterium exchange can be challenging due to the presence of indirect effects on the structure and dynamics of the protein.

2.5 Literature Cited

- (1) Vuignier, K., Schappler, J., Veuthey, J. L., Carrupt, P. A., Martel, S., *Anal. Bioanal. Chem.*, **2010**, *398*, 53-66.
- (2) Lee, Y. C., Lee, R. T., *Accounts Chem. Res.*, **1995**, *28*, 321-327.
- (3) Holgersson, J., Gustafsson, A., Breimer, M. E., *Immunolo. Cell Biol.*, **2005**, *83*, 694-708.
- (4) Kamiya, Y., Yagi-Utsumi, M., Yagi, H., Kato, K., *Curr. Pharm. Design*, **2011**, *17*, 1672-1684.
- (5) Duan, X. Q., Hall, J. A., Nikaido, H., Quioco, F. A., *J. Mol. Biol.*, **2001**, *306*, 1115-1126.
- (6) Duan, X. Q., Quioco, F. A., *Biochemistry*, **2002**, *41*, 706-712.
- (7) Fernandez-Alonso, M. D., Diaz, D., Berbis, M. A., Marcelo, F., Canada, J., Jimenez-Barbero, J., *Curr. Protein Pept. Sc.*, **2012**, *13*, 816-830.
- (8) Kohn, J. E., Afonine, P. V., Ruscio, J. Z., Adams, P. D., Head-Gordon, T., *PLoS Comput. Biol.*, **2010**, *6*, e1000911.
- (9) Percy, A. J., Rey, M., Burns, K. M., Schriemer, D. C., *Anal. Chim. Acta*, **2012**, *721*, 7-21.
- (10) Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., Griffin, P. R., *Expert Rev. Proteomic*, **2011**, *8*, 43-59.
- (11) Iacob, R. E., Engen, J. R., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1003-1010.

- (12) Garcia, R. A., Pantazatos, D., Villarreal, F. J., *Assay Drug Dev. Technol.*, **2004**, 2, 81-91.
- (13) Houde, D., Berkowitz, S. A., Engen, J. R., *J. Pharm. Sci.*, **2011**, 100, 2071-2086.
- (14) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, 2, 522-531.
- (15) Pacholarz, K. J., Garlish, R. A., Taylor, R. J., Barran, P. E., *Chem. Soc. Rev.*, **2012**, 41, 4335-4355.
- (16) Huang, R. Y. C., Wen, J., Blankenship, R. E., Gross, M. L., *Biochemistry*, **2012**, 51, 187-193.
- (17) Wildes, D., Marqusee, S., *Protein Sci.*, **2005**, 14, 81-88.
- (18) Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Garcia-Ordonez, R. D., Pascal, B. D., Burris, T. P., Dodge, J. A., Griffin, P. R., *Structure*, **2010**, 18, 1332-1341.
- (19) Chetty, P. S., Mayne, L., Kan, Z. Y., Lund-Katz, S., Englander, S. W., Phillips, M. C., *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, 109, 11687-11692.
- (20) Ahn, J., Engen, J. R., *Chim. Oggi-Chem. Today*, **2013**, 31, 25-28.
- (21) Jorgensen, T. J. D., Gardsvoll, H., Dano, K., Roepstorff, P., Ploug, M., *Biochemistry*, **2004**, 43, 15044-15057.
- (22) Dai, S. Y., Burris, T. P., Dodge, J. A., Montrose-Rafizadeh, C., Wang, Y., Pascal, B. D., Chalmers, M. J., Griffin, P. R., *Biochemistry*, **2009**, 48, 9668-9676.
- (23) Hamuro, Y., Coales, S. J., Morrow, J. A., Molnar, K. S., Tuske, S. J., Southern, M. R., Griffin, P. R., *Protein Sci.*, **2006**, 15, 1883-1892.
- (24) Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., Chen, G. D., *J. Am. Soc. Mass Spectrom.*, **2012**, 23, 498-504.

- (25) Pan, Y., Brown, L., Konermann, L., *J. Am. Chem. Soc.*, **2011**, *133*, 20237-20244.
- (26) Pan, Y., Piyadasa, H., O'Neil, J. D., Konermann, L., *J. Mol. Biol.*, **2012**, *416*, 400-413.
- (27) King, D., Bergmann, C., Orlando, R., Benen, J. A. E., Kester, H. C. M., Visser, J., *Biochemistry*, **2002**, *41*, 10225-10233.
- (28) King, D., Lumpkin, M., Bergmann, C., Orlando, R., *Rapid Commun. Mass Spectrom.*, **2002**, *16*, 1569-1574.
- (29) Seyfried, N. T., Atwood, J. A., Yongye, A., Almond, A., Day, A. J., Orlando, R., Woods, R. J., *Rapid Commun. Mass Spectrom.*, **2007**, *21*, 121-131.
- (30) Huzil, J. T., Chik, J. K., Slysz, G. W., Freedman, H., Tuszynski, J., Taylor, R. E., Sackett, D. L., Schriemer, D. C., *J. Mol. Biol.*, **2008**, *378*, 1016-1030.
- (31) Lee, C. T., Graf, C., Mayer, F. J., Richter, S. M., Mayer, M. P., *Embo J.*, **2012**, *31*, 1518-1528.
- (32) Bennett, M. J., Barakat, K., Huzil, J. T., Tuszynski, J., Schriemer, D. C., *Chem. Biol.*, **2010**, *17*, 725-734.
- (33) Lin, H., Kitova, E. N., Klassen, J. S., *J. Am. Soc. Mass Spectrom.*, **2014**, *25*, 104-110.
- (34) Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., Hol, W. G. J., *Protein Sci.*, **1997**, *6*, 1516-1528.
- (35) Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L., Read, R. J., *Biochemistry*, **1998**, *37*, 1777-1788.
- (36) Johannes, L., Romer, W., *Nat. Rev. Microbiol.*, **2010**, *8*, 105-116.

- (37) Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L., Read, R. J., *Nature*, **1992**, 355, 748-750.
- (38) Greco, A., Ho, J. G. S., Lin, S. J., Palcic, M. M., Rupnik, M., Ng, K. K. S., *Nat. Struct. Mol. Biol.*, **2006**, 13, 460-461.
- (39) Dingle, T., Wee, S., Mulvey, G. L., Greco, A., Kitova, E. N., Sun, J. X., Lin, S. J., Klassen, J. S., Palcic, M. M., Ng, K. K. S., Armstrong, G. D., *Glycobiology*, **2008**, 18, 698-706.
- (40) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, 50, 8722-8732.
- (41) Merritt, E. A., Sarfaty, S., Vandenakker, F., Lhoir, C., Martial, J. A., Hol, W. G. J., *Protein Sci.*, **1994**, 3, 166-175.
- (42) Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K., Ho, W. G. J., *J. Mol. Biol.*, **1998**, 282, 1043-1059.
- (43) Skinner, J. J., Lim, W. K., Bedard, S., Black, B. E., Englander, S. W., *Protein Sci.*, **2012**, 21, 996-1005.
- (44) Shimizu, H., Field, R. A., Homans, S. W., Donohue-Rolfe, A., *Biochemistry*, **1998**, 37, 11078-11082.

Chapter 3

Localizing the Binding Site for Toxin-Human Milk Oligosaccharides (HMOs) Interactions Using Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)

3.1 Introduction

Human milk contains a wide variety of components, such as proteins, carbohydrates, lipids, mineral, vitamins etc.¹⁻³ It is not only superior food to provide infants essential nutrition for growth and development, but also a protection for newborns against a number of infection diseases until their own immune systems are functioning properly.⁴⁻⁷ Human milk oligosaccharides (HMOs) are the third largest component of human milk composed of thousands of components, among which some are related to the immunological protective effect of human milk.⁷⁻⁹ The protective effect of HMOs is thought to primarily result from inhibition of pathogens binding to host cells.^{10,11} HMOs are similar to some glycan receptors on mucosal cell surface in structure. As a result, the HMOs can serve as decoy ligands and inhibit the binding of microbial lectins to host cell receptors, and consequently preventing infection of the host by these organisms.¹² Although there are lots of detailed studies into the protective effect of HMOs against infections, investigations into the nature of the pathogen-HMOs interactions such as the structures of the pathogen-HMOs complexes, and the corresponding thermodynamic and kinetic parameters of the interactions are few.¹³⁻¹⁶

Recently, our laboratory applied the direct electrospray ionization mass spectrometry (ESI-MS) assay to identify and quantify the binding of a HMO library, consisting of twenty of the most abundant HMOs, to two recombinant fragments (A2 and B1) of *Clostridium difficile* toxins A (TcdA) and B (TcdB), respectively.¹³ According to the direct ESI-MS results, fragments of both toxins A and B recognize a number of neutral and acidic HMOs, despite the weak interactions with apparent association constants ($K_{a,app}$) in the 10^3 M^{-1} range. Following these results, the same assay was also performed on some other bacterial exotoxins like cholera toxin (CT) and Shiga toxin 1 (Stx1). Competitive binding measurements of HMOs and the native ligands of the toxins were carried out. The results showed that even when the toxin was saturated by its native ligand, it can still bind to some HMOs like 2'-fucosyllactose, suggesting that toxins might possess distinct binding sites for native receptors and HMOs. Therefore, efforts are needed to explore such possibilities.

Here, Hydrogen/deuterium exchange mass spectrometry (HDX-MS) was applied to explore the existence of distinct HMOs binding sites on toxins. Altogether, two toxins were studied, including the B subunit homopentamers of Cholera toxin (CTB₅) and a recombinant fragment of *Clostridium difficile* toxins A (TcdA-A2) and their interactions with HMOs, 2'-fucosyllactose (2'-FL, α -L-Fuc(1→2)- β -D-Gal(1→4)- β -D-Glc) and lacto-N-tetraose (LNT, β -D-Gal(1→3)- β -D-GlcNAc(1→3)- β -D-Gal(1→4)- β -D-Glc), respectively. Cholera toxin belongs to the family of AB₅ toxins, which consists of an enzymatic A component and a cell-binding B component that is a pentamer of identical subunits that assemble into a doughnut-shape.^{17,18} The native carbohydrate receptor for CTB₅ is GM₁ pentasaccharide (GM₁-os, β -Gal-

(1→3)-β-GalNAc-(1→4)[α-Neu5Ac-(2→3)]-β-Gal-(1→4)-Glc).¹⁹ According to the crystal structure reported for the (CTB₅ + 5GM₁-os) complex, the ligand binding site is formed by three loops from the same subunit and a fourth loop containing residues from the adjacent subunit.^{18,20} *Clostridium difficile* toxins A is a single unit protein with multi-domain structure. A2 is a fragment from its C-terminal repetitive domain, which can bind to carbohydrate receptors on the cell surface.^{21,22} Currently, the only known native receptor for TcdA is the trisaccharide α-Gal-(1→3)-β-Gal(1→4)-β-GlcNAc, which is found on the surface of rabbit erythrocytes, hamster brush border membranes and bovine thyroglobulin.²³ Within the A2 fragment, a long repeat and the hairpin turn of the following short repeat create a shallow carbohydrate binding site, and there are two carbohydrate binding sites present in the fragment.²¹

3.2 Experimental Methods

3.2.1 Materials

CTB₅ (MW 58,020 Da) was purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). A stock solution (60 μM) of CTB₅ was prepared by dissolving in ultrafiltered water (Milli-Q; Millipore, Billerica, MA, USA) and stored at 4 °C until needed. TcdA-A2 was a gift from Prof. K. Ng (Univ. of Calgary) as a stock solution prepared in 60 mM imidazole (pH 7.0), 150 mM NaCl and 50 g/L glycerol (>99.5% purity, Sigma) with a concentration of 57.5 μM. The stock solutions were stored at -80 °C until needed. Prior to analysis, the protein solutions were diluted with Milli-Q water to a desired concentration. 2'-fucosyllactose (2'-FL, α-L-Fuc(1→2)-β-D-Gal(1→4)-β-D-Glc, MW 488.43 Da) and lacto-N-tetraose (LNT, β-D-Gal(1→3)-β-

D-GlcNAc(1→3)-β-D-Gal(1→4)-β-D-Glc, MW 708.62 Da) were purchased from Elicityl SA (Crolles, France). The stock solutions of the HMOs were prepared by dissolving in Milli-Q water to yield a final concentration of 0.4 M and were stored at -20 °C until needed.

3.2.2 HDX-MS

The HDX-MS measurements procedures was similar to that was described in Chapter 2. Briefly, the HDX-MS experiments were carried out using a Synapt G2-S HDMS mass spectrometer equipped with a nanoACQUITY UPLC system with HDX technology (Waters, UK) and PAL HTX-xt automatic sample preparation and injection system. The temperature of labeling reactions and the reactions quench was accurately controlled at 20 °C and 1 °C, respectively, by the two sample stacks from PAL system. Free protein sample (4 μM and 12 μM for CTB₅ and TcdA-A2, respectively) or protein together with excess ligand (0.2 M HMO with 4 μM CTB₅ or 12 μM TcdAQ-A2) was diluted 15-fold with either equilibrium buffer (10 mM potassium phosphate in H₂O at pH 7.0) for control experiments, or labeling buffer (10 mM potassium phosphate in D₂O at pD 7.0 (pD = pH + 0.4)) for labeling experiments. Once labeling reaction was initiated, samples were incubated at 20 °C for time intervals of 0 (control experiments), 1, 5 and 10 min (labeling experiments). The protein samples in the absence or presence of ligand were analyzed back-to-back for a given labeling time (*t*). In order to minimize the carry-over effect, a water sample was injected as blank after each pair of protein samples following the same procedures. Quench buffer (4 M guanidine hydrochloride and 0.5M *tris*(2-

carboxyethyl)phosphine (TCEP) in H₂O at pH 2.6 for CTB₅, and 4 M guanidine hydrochloride in 100 mM potassium phosphate in H₂O at pH 2.6 for TcdA-A2) was used to quench the labeling reaction through a 1:1 dilution at 1 °C. Prior to injection into a injection loop (50 µL) on the nano-ACQUITY UPLC system, quenched samples were held for 30 s under quench condition. Immobilized pepsin column (Life Technologies, Burlington, Canada) was used to perform the online digestion at 20 °C. The mobile phase flow through the pepsin column was made up with 0.1% formic acid in H₂O at a flow rate 200 µL min⁻¹. Peptides produced were trapped online by an ACQUITY UPLC BEH C18 1.7 µm VanGuard Precolumn (Waters, UK) at 1 °C. Protein digestion and peptides trapping were taken place at the same time for 2 min. An ACQUITY UPLC C18 1.7 µm 1.0×100mm column was employed for peptides separation in a 12 min gradient elution with a flow rate of 40 µL min⁻¹. During the elution, solvent A in mobile phase (Solvent A, 0.1% formic acid and 5% acetonitrile in H₂O; solvent B, 0.1% formic acid in acetonitrile). was first decreased over a 7 min period from 95% to 63%, following by reduction from 63% to 16% over 0.5 min after being held constant for 1 min, and was increased back to 95% over 0.5 min after being held constant for another 0.5 min. The separated peptides were introduced into the Synapt G2-S HDMS (Waters, UK) with ESI source together with lock-mass correction (using [Glu]-Fibrinopeptide and the peak with m/z 785.8426). Mass spectra were acquired in MS^E mode with m/z range from 50 to 2000. The scan rate was 0.4 sec/scan, with the capillary and cone voltages kept constant at 3 kV and 40 V, respectively.

3.2.3 Data processing for HDX-MS

ProteinLynx Global Server 2.5.2 software (PLGS, Waters, UK) was applied to identify peptides from control samples (i.e., $t = 0$) with a databank containing sequence for protein of interest. Only the subset of peptides identified in all control and labeling samples from all replicates were considered further. According to the common peptides identified by PLGS, DynamX 2.0 software (Waters, UK) was used to generate a peptide list, and assign corresponding peaks for peptides in both control and labeling samples. The average molecular mass for a specific peptide i was determined by DynamX as the centroid of the entire envelope of corresponding isotopic peaks. The absolute D-uptake value (D_i , unit of Da) for peptide i was calculated as the molecular mass difference for this peptide between labeling and control samples at each labeling time. Because of the unavoidable occurrence of back exchange after sample quench,^{24,25} relative D-uptake (ΔD_i , unit of Da) values for each peptide were calculated to compare between free and ligand-bound protein. No back exchange correction was applied. The summation of ΔD_i values across all the labeling time points was also assessed for all peptides of interest. Differences were considered significant if the values were greater than three times the calculated standard deviation.

3.3 Results and Discussion

3.3.1 CTB₅ and its interaction with 2'-FL

Pepsin digestion of denatured CTB produced around 80 peptides for each sample, among which 40 were identified in all replicates of all samples. In order to

decrease the redundancy and ensure maximum sequence coverage, a peptide list only containing 15 peptides, covering 94.2% of CTB monomer sequence, was generated for further analysis (Figure 3.1).

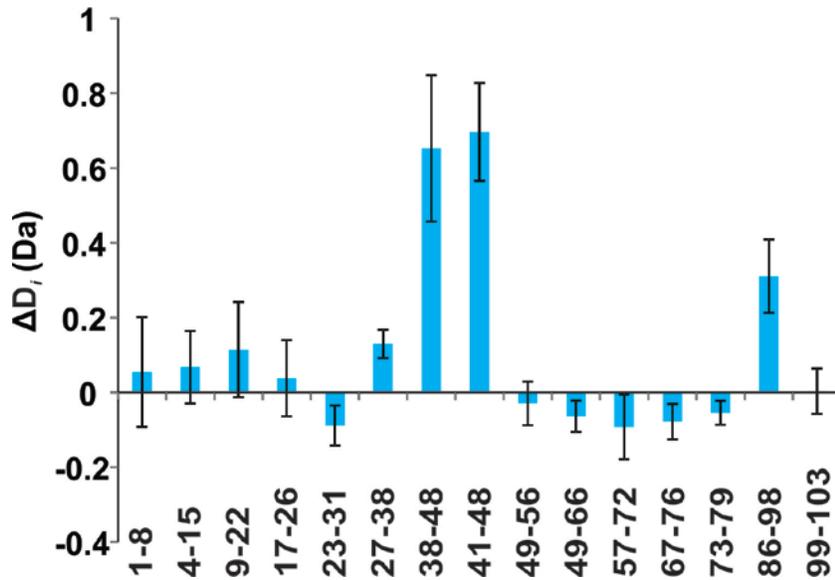


Figure 3.1 Difference plot of peptide level D-uptake between free CTB₅ and (CTB₅ + 5(2'-FL)) complex for a subset of peptides that were common to all repeated experiments covering 94.2% of CTB monomer sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments.

In the presence of 2'-FL, two peptides 38-48 and 41-48 from each CTB monomer showed a significant decrease in D_i values with ΔD_i value equal to 0.4 and 0.5 Da, respectively. Notably, the region protected by 2'-FL were different from those

protected by GM₁-os which consist of four individual parts (Chapter 2).^{18,20} This result thus suggesting the existence of a distinct 2'-FL binding site.

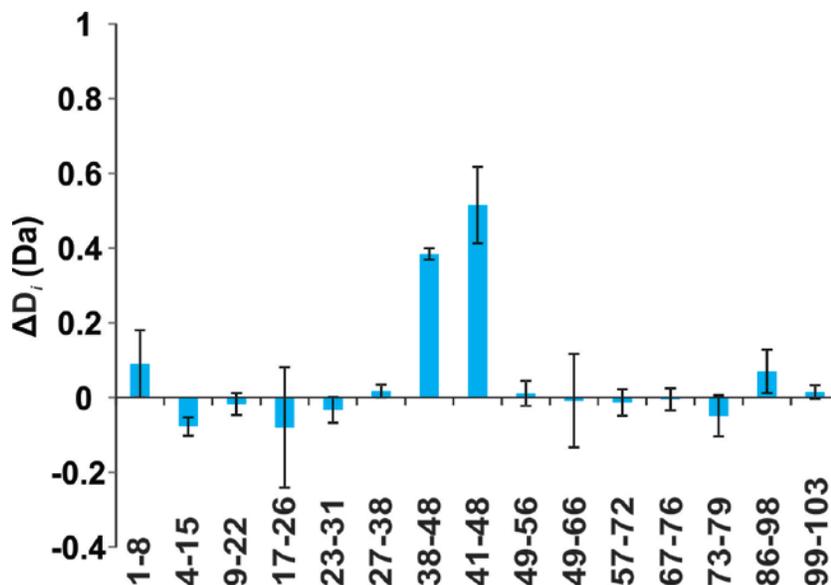


Figure 3.2 Difference plot of peptide level D-uptake between GM₁-os saturated CTB₅ and (CTB₅ + 5GM₁-os + 5(2'-FL)) complex for a subset of peptides that were common to all repeated experiments covering 94.2% of CTB monomer sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments.

For further verification, a HDX-MS binding measurement using GM₁ saturated CTB₅ to interact with 2'-FL was carried out. The same list containing exactly the same peptides was generated to compare the D-uptake difference between

GM₁-os saturated CTB₅ in the absence and presence of 2'-FL (Figure 3.2). It can be seen that peptides 38-48 and 41-48 clearly showed protection in the presence of 2'-FL for GM₁-os saturated CTB₅. This confirmed 2'-FL binds to a distinct site from GM₁-os.

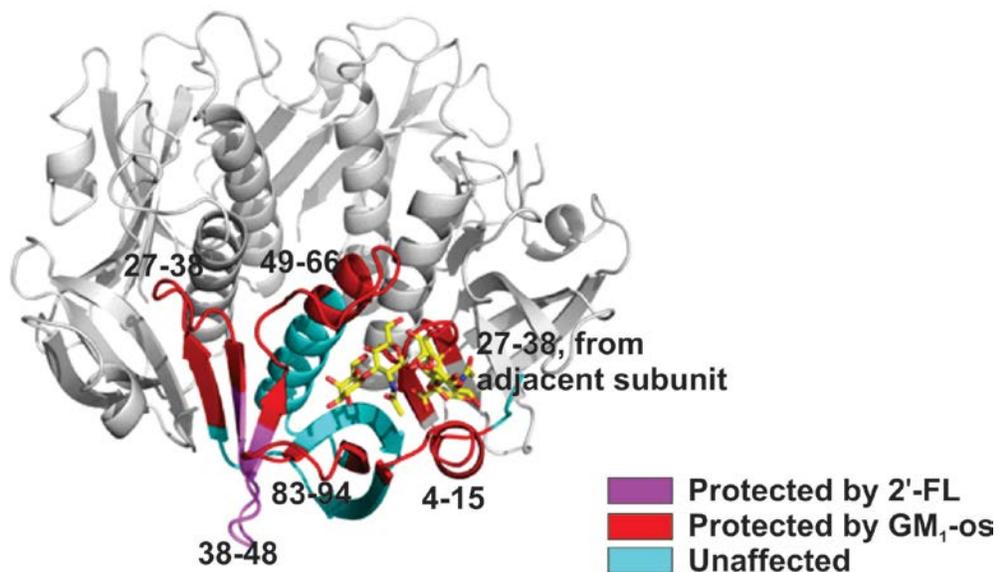


Figure 3.3 Cartoon illustration of ΔD_i values mapped onto the crystal structure of (CTB₅ + 5GM₁-os) complex (PDB 3CHB). Colored region is a representative CTB monomer. The red and purple regions of protein were protected from deuterium exchange upon GM₁-os and 2'-FL binding, respectively, containing specific residues shown as labeled, while the cyan regions were unaffected by the bound ligand.

To locate the protected peptides, the ΔD_i values were mapped onto the crystal structure of (CTB₅ + 5GM₁-os) complex (PDB 3CHB) for one subunit (Figure 3.3). The residues protected by the bound 2'-FL were labeled in red. The protected region

contains a loop oriented on the opposite side of GM₁-os binding site, which is most likely to be the 2'-FL binding site. For specific residues involved in CTB-2'-FL interaction, more studies are needed such as docking simulation.^{26,27}

3.3.2 TcdA-A2 and its interaction with LNT

Digestion of TcdA-A2 produced approximately 100 different peptides, of 52 were identified in every analysis carried out. The deuterium uptake for 20 of these peptides, covering 85.5% of TcdA-A2 sequence, was investigated.

The ΔD_i values measured for these peptides are shown in Figure 3.4. The two protected peptides are 85-107 and 92-107, with ΔD_i values of 0.6 and 0.4 Da, respectively. The protected region was similar as CD-grease binding (Chapter 2),²¹ suggesting the two ligands, CD-grease and LNT, share the same binding site on TcdA-A2. Besides, only one peptide 15-30 exhibited a negative ΔD_i value of -0.8 Da, indicating it was de-protected from deuterium exchange, which was most likely to be originated from the allosteric effects of the toxin upon ligand binding. However, this allosteric effects caused by LNT binding was much smaller than CD-grease binding.

A competitive HDX-MS binding measurement was also carried out to compare the D-uptake difference between CD-grease saturated TcdA-A2 in the absence and presence of LNT (Figure 3.5). According to it, no significant difference in D-uptake was observed across the whole TcdA-A2 fragment. This result furthermore verified that CD-grease and LNT compete to each other to bind to the same site on TcdA-A2, thus no extra effect was found on the toxin upon ligand binding.

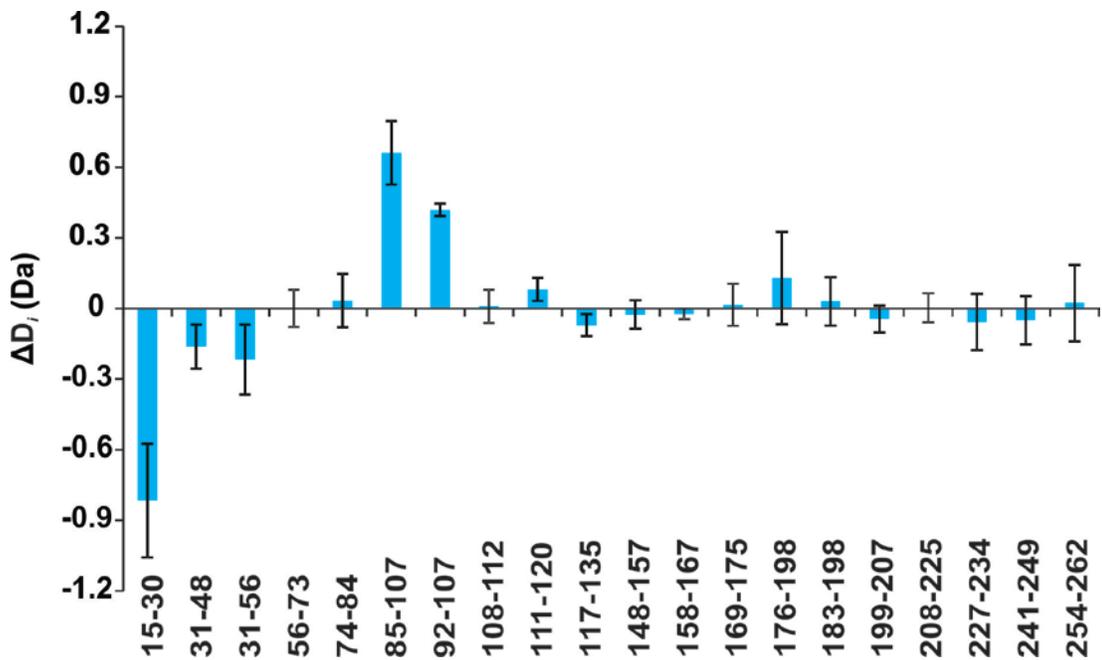


Figure 3.4 Difference plot of peptide level D-uptake between free TcdA-A2 and (TcdA-A2 + LNT) complex for a subset of peptides that were common to all repeated experiments covering 85.5% of TcdA-A2 sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments.

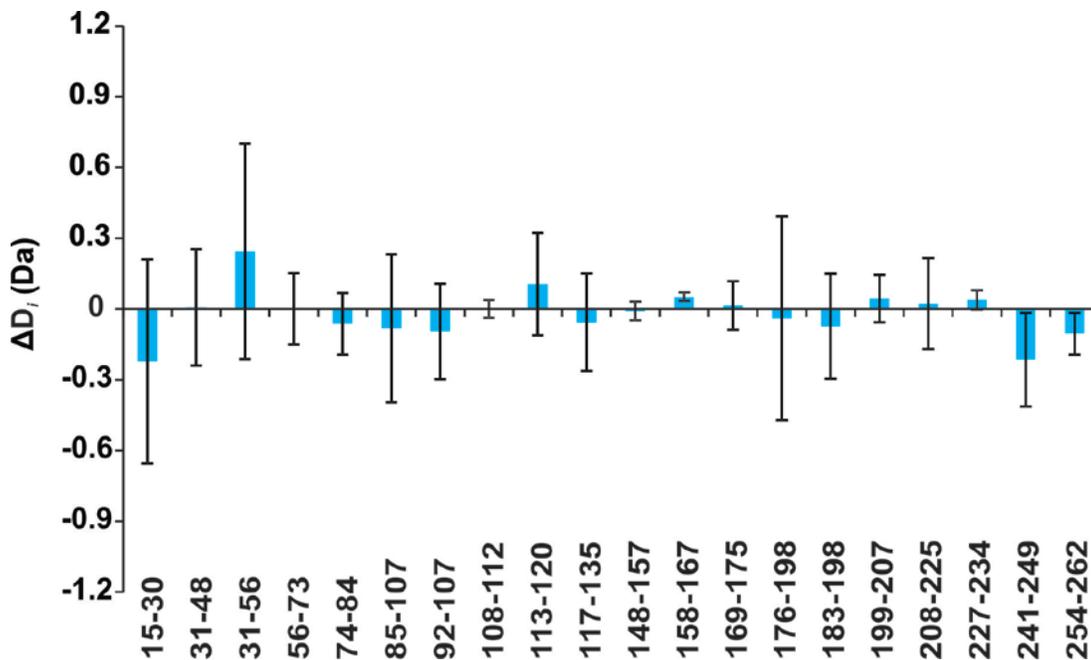


Figure 3.5 Difference plot of peptide level D-uptake between CD-grease saturated TcdA-A2 and (TcdA-A2 + CD-grease + LNT) complex for a subset of peptides that were common to all repeated experiments covering 85.5% of TcdA-A2 sequence. The x-axis shows the position of individual peptides. The y axis is the summation of ΔD_i values for each peptide across the three labeling time points. Error bars are standard deviations of three independent experiments.

In Figure 3.6 the ΔD_i values are mapped onto the crystal structure of the (TcdA-A2 + 2CD-grease) complex (PDB 2G7C). The peptides protected and de-protected by LNT were labeled in red and dark blue. The region protected by LNT was consistent with the CD-grease binding site. However, the existence of the second ligand binding site for CD-grease was not observed for LNT. Considering the HDX-

MS results for TcdA-A2 and CD-grease binding (Chapter 2), we are not sure whether TcdA-A2 has two binding sites for LNT or not. To investigate the interactions between TcdA-A2 and LNT at atomic level, however, more studies are required.



Figure 3.6 Cartoon illustration of ΔD_i values mapped onto the crystal structure of (TcdA-A2 + 2CD-grease) complex (PDB 2G7C). The red and dark blue regions of protein were protected and de-protected from deuterium exchange upon LNT binding containing specific residues shown as labeled, while the cyan regions were unaffected by the bound ligand.

3.4 Conclusions

HDX-MS was applied to explore the existence of distinct HMOs binding sites on bacterial toxins. For CTB₅, a novel binding site was localized for 2'-FL, different from the one for native receptor GM₁. When it comes to TcdA-A2, however, the localized LNT binding site was the same as its native carbohydrate receptor CD-grease. Nevertheless, to examine these interactions at atomic level, more studies are required.

3.5 Literature Cited

- (1) German, J. B., Freeman, S. L., Lebrilla, C. B., Mills, D. A. In *Personalized Nutrition for the Diverse Needs of Infants and Children*, Bier, D. M., German, J. B., Lonnerdal, B., Karger: Basel, **2008**, Vol. 62, p 205-222.
- (2) Field, C. J., *J. Nutr.*, **2005**, *135*, 1-4.
- (3) Lara-Villoslada, F., Olivares, M., Sierra, S., Rodriguez, J. M., Boza, J., Xaus, J., *Br. J. Nutr.*, **2007**, *98*, S96-S100.
- (4) Hoppu, U., Kalliomaki, M., Laiho, K., Isolauri, E., *Allergy*, **2001**, *56*, 23-26.
- (5) Morrow, A. L., Ruiz-Palacios, G. M., Jiang, X., Newburg, D. S., *J. Nutr.*, **2005**, *135*, 1304-1307.
- (6) Newburg, D. S., *J. Nutr.*, **2005**, *135*, 1308-1312.
- (7) Newburg, D. S., Ruiz-Palacios, G. M., Morrow, A. L. In *Annu. Rev. Nutr.*, Annual Reviews: Palo Alto, **2005**, Vol. 25, p 37-58.
- (8) Morrow, A. L., Ruiz-Palacios, G. M., Altaye, M., Jiang, X., Guerrero, M. L., Meinzen-Derr, J. K., Farkas, T., Chaturvedi, P., Pickering, L. K., Newburg, D. S., *J. Pediatr.*, **2004**, *145*, 297-303.
- (9) Kunz, C., Rudloff, S., *Acta Paediatr.*, **1993**, *82*, 903-912.
- (10) Kunz, C., Rudloff, S., Baier, W., Klein, N., Strobel, S., *Annu. Rev. Nutr.*, **2000**, *20*, 699-722.
- (11) Ruiz-Palacios, G. M., Cervantes, L. E., Ramos, P., Chavez-Munguia, B., Newburg, D. S., *J. Biol. Chem.*, **2003**, *278*, 14112-14120.
- (12) Bode, L., *Glycobiology*, **2012**, *22*, 1147-1162.

- (13) El-Hawiet, A., Kitova, E. N., Kitov, P. I., Eugenio, L., Ng, K. K. S., Mulvey, G. L., Dingle, T. C., Szpacenko, A., Armstrong, G. D., Klassen, J. S., *Glycobiology*, **2011**, *21*, 1217-1227.
- (14) Kobata, A., *Proc. Jpn. Acad. Ser. B-Phys. Biol. Sci.*, **2010**, *86*, 731-747.
- (15) Martin-Sosa, S., Martin, M. J., Hueso, P., *J. Nutr.*, **2002**, *132*, 3067-3072.
- (16) Bode, L., *J. Nutr.*, **2006**, *136*, 2127-2130.
- (17) Merritt, E. A., Sarfaty, S., Vandenaeker, F., Lhoir, C., Martial, J. A., Hol, W. G. J., *Protein Sci.*, **1994**, *3*, 166-175.
- (18) Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., Hol, W. G. J., *Protein Sci.*, **1997**, *6*, 1516-1528.
- (19) Masco, D., Vandewalle, M., Spiegel, S., *J. Neurosci.*, **1991**, *11*, 2443-2452.
- (20) Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K., Ho, W. G. J., *J. Mol. Biol.*, **1998**, *282*, 1043-1059.
- (21) Greco, A., Ho, J. G. S., Lin, S. J., Palcic, M. M., Rupnik, M., Ng, K. K. S., *Nat. Struct. Mol. Biol.*, **2006**, *13*, 460-461.
- (22) Dingle, T., Wee, S., Mulvey, G. L., Greco, A., Kitova, E. N., Sun, J. X., Lin, S. J., Klassen, J. S., Palcic, M. M., Ng, K. K. S., Armstrong, G. D., *Glycobiology*, **2008**, *18*, 698-706.
- (23) Krivan, H. C., Clark, G. F., Smith, D. F., Wilkins, T. D., *Infect. Immun.*, **1986**, *53*, 573-581.
- (24) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.
- (25) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (26) Trott, O., Olson, A. J., *J. Comput. Chem.*, **2010**, *31*, 455-461.

(27) Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Garcia-Ordonez, R. D., Pascal, B. D., Burris, T. P., Dodge, J. A., Griffin, P. R., *Structure*, **2010**, *18*, 1332-1341.

Chapter 4

Measuring Protein-Carbohydrate Binding Affinities Using Protein-Ligand Interactions in Solution by Mass Spectrometry, Titration and Hydrogen/Deuterium Exchange (PLIMSTEX)

4.1 Introduction

Protein-carbohydrate interactions play a crucial role in biophysics and drug design.^{1,2} Therefore, investigations into the protein structures and thermodynamic properties upon carbohydrate binding are of major interest for understanding the fundamental biological processes, which can serve as the guidance of disease diagnosis and new therapeutics development.^{3,4} Up to now, there are several established methods available to characterize and quantify protein-carbohydrate interactions, such as surface plasmon resonance (SPR) spectroscopy,⁵⁻⁷ isothermal titration calorimetry (ITC),⁸⁻¹⁰ enzyme-linked immunosorbent assay (ELISA),¹¹ and nuclear magnetic resonance (NMR) spectroscopy.^{3,12} These methods are all capable to provide association constant (K_a) of the interactions. Besides, SPR can also provide kinetic parameters like ligand on and off rate constants, and ITC is able to give thermodynamic parameters like Gibbs free energy, enthalpy and entropy of binding.¹³ Although SPR and ELISA afford high sensitivity and only need a small amount of sample, they require the immobilization one of the binding partners on a surface, which may affect the nature of binding interactions.¹¹ Also, most ELISAs rely on enzyme-mediated amplification of signal to achieve high sensitivity, which can limit their applicability. For ITC and NMR, however, they suffer from low sensitivity

which requires a large amount (~mg) of pure protein and ligand for each analysis and the measurements are time-consuming.^{10,14}

The hydrogen/deuterium exchange mass spectrometry (HDX-MS) measurements are emerging as a promising method to investigate protein-ligand interactions.¹⁵⁻¹⁹ Most HDX-MS studies were aiming at obtaining differences in HDX kinetics between free and ligand-bound protein, so as to extract out changes in protein conformation or dynamic upon ligand binding, or localize the ligand binding site, similar as the studies in Chapter 2 and 3 of this thesis.²⁰⁻²⁴ However, only a few examples employed HDX-MS for quantitative analysis of protein-ligand interactions. One example is a method known as Protein-Ligand Interactions in solution by Mass Spectrometry, Titration and hydrogen/deuterium Exchange (PLIMSTEX), which requires a change of D-uptake occurs during a titration (Figure 4.1).²⁵ First, protein (with a fixed initial concentration) is incubated with different concentrations of ligand in aqueous buffer to reach the equilibrium. Similar to the continuous labeling method mentioned in Chapter 1, the HDX reactions are initiated by adding deuterated buffer. After a certain labeling time, the exchange is quenched by decreasing both pH and temperature, followed by enzymatic digestion and LC-MS analysis. Due to ligand protection, free and ligand-bound protein are expected to show different D-uptake values within the ligand binding site, as D_P and D_{PL} . If the sample is a mixture of both free and ligand-bound protein, the peak from the two species will overlap with each other, and the D-uptake value, D_i , will fall between D_P and D_{PL} . The magnitude of D_i , however, depends on the ratio of free and ligand-bound protein in solution under labeling conditions. Therefore, by changing the initial ligand concentration with fixed

initial protein concentration, a ligand titration experiment can be carried out using HDX-MS measurement. And the D_i obtained can relate the initial ligand concentration to the fraction of protein-ligand complex under labeling condition.

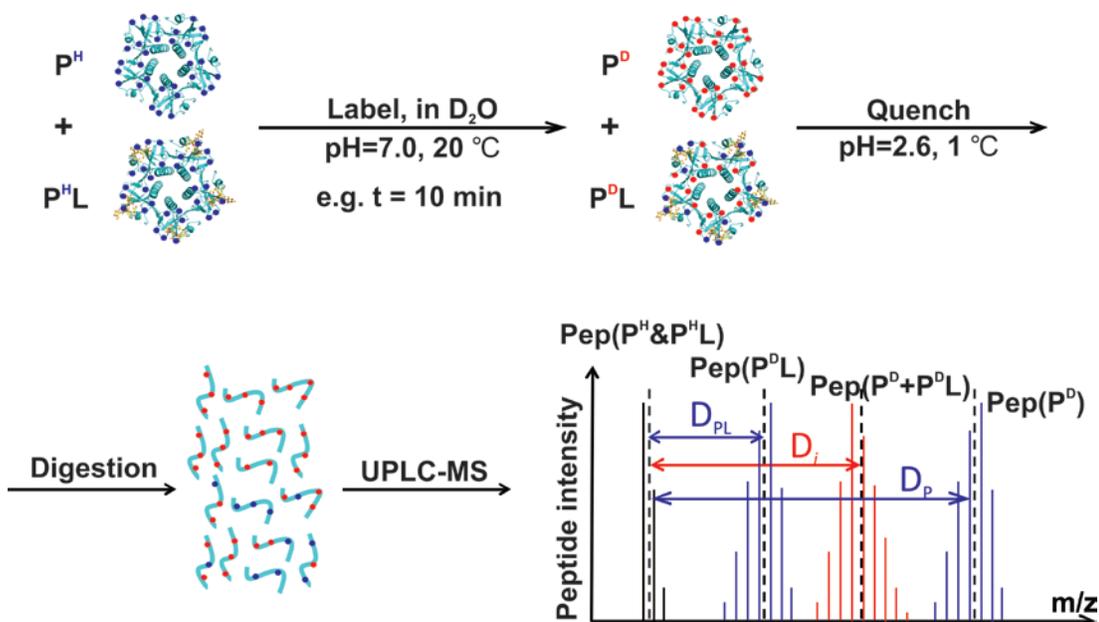


Figure 4.1 Illustration of detailed procedures for PLIMSTEX method at peptide level.

PLIMSTEX was first developed by Gross and coworkers for ligand titration of a global protein without enzymatic fragmentation step.^{26,27} Using global PLIMSTEX method, they determined the association constants, K_a , and binding stoichiometry for interactions of proteins with various ligands, including metal ions, small organic molecules, peptides, proteins and DNA.²⁸⁻³¹ Altogether, there are several advantages of using global PLIMSTEX to quantify protein-ligand interactions. First, with LC desalting step, this method is amenable to different

protein-ligand systems in their physiological buffers, and provides minimal perturbation of the binding equilibrium. In addition, global PLIMSTEX measures the mass of a protein after HDX rather than the abundance, thus overcome the problems such as non-uniform response factors,³² nonspecific binding, and in-source dissociation when using direct ESI-MS measurements to obtain the binding affinities.³³⁻³⁵ Nevertheless, PLIMSTEX requires a measurable change of protein molecular mass after HDX during the titration. For a large protein, it's difficult to determine the accurate molecular mass, leading to a large error when calculating the relative D-uptake between the free and ligand-bound protein. Consequently, a large error will show up for the binding affinity derived from the D-uptake values. Moreover, it is impossible for global PLIMSTEX method to obtain site specific binding affinities for a protein with multiple ligand binding sites, which is the case for a number of protein-carbohydrate interactions.³⁶⁻³⁸ Gross and coworkers also tried PLIMSTEX carried out at peptide level. With careful selection of peptide indicators, they obtained the binding affinity for a protein-DNA interaction successfully.²⁵ However, up to now, this is the only peptide-level PLIMSTEX study. To test if the method is general and can be applied to obtain the affinities for protein-carbohydrate interactions, peptide level PLIMSTEX experiments were carried out in the study of this chapter. And the B subunit homopentamers of Cholera toxin (CTB₅) and their interactions with native carbohydrate receptors, GM₁ pentasaccharide (GM₁-os, β -Gal-(1 \rightarrow 3)- β -GalNAc-(1 \rightarrow 4)[α -Neu5Ac-(2 \rightarrow 3)]- β -Gal-(1 \rightarrow 4)-Glc), was chosen as the model system.

CTB₅ displays a high affinity (intrinsic association constant $K_{a,int}$ between 10^6 and 10^7 M⁻¹) for GM₁-os and can bind simultaneously to five GM₁-os. According to the direct ESI-MS assay, the stepwise binding of GM₁ to CTB₅ exhibits positive cooperativity, and binding was sensitive to the number of ligand-bound nearest neighbor subunits.³⁹ Previous studies presented the binding of GM₁ to CTB₅ as a function of the sialic acid “thumb” (H-bonds between sialic acid and residues E11, H13, G33 and Q56) and the terminal galactose and galNAc “finger” (H-bonds between galactose and residues E51, Q56, Q61, N90 and K91).⁴⁰ While the thumb is sufficient for recognition, the finger is mainly to stabilize the CTB-GM₁ interaction. Base on the reported crystal structure for the (CTB₅ + 5GM₁-os) complex (PDB 3CHB), the GM₁-os binding site is consisted of three separated loops (close to each other in three-dimensional structure) from the same subunit and a fourth loop containing residues from the adjacent subunit.^{41,42} According to the results of Chapter 2, each of the four loops is covered by at least one peptide produced in the pepsin digestion step of the HDX-MS experiment. Therefore, there are at least four peptides can serve as indicators to obtain the binding affinity for CTB₅ and GM₁-os interactions.

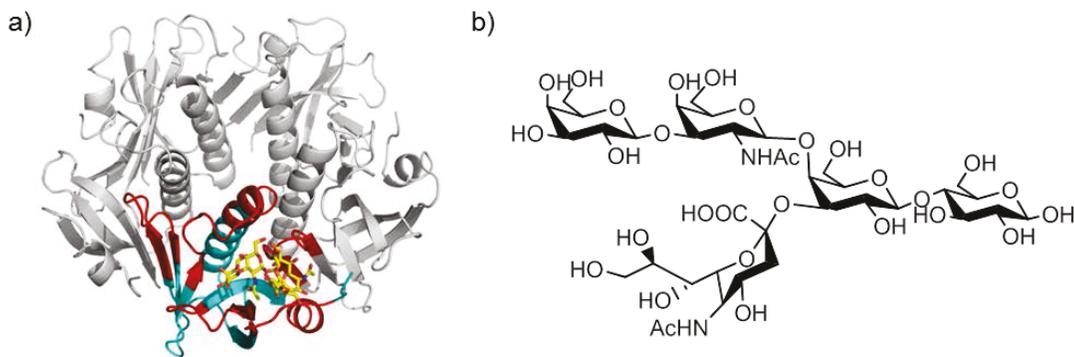


Figure 4.2 (a) (CTB₅ + 5GM₁-os) complex with yellow sticks represent the ligands. Colored region is a representative CTB monomer (includes one loop from adjacent subunit involved in binding). The red regions of protein were protected from deuterium exchange upon ligand binding, while the blue regions were unaffected by the bound ligand. (b) Chemical structures of GM₁-os.

4.2 Experimental Methods

4.2.1 Materials

CTB₅ (MW 58,020 Da) was purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). A stock solution (60 μM) of CTB₅ was prepared by dissolving in ultrafiltered water (Milli-Q; Millipore, Billerica, MA, USA) and stored at 4 °C until needed. Prior to analysis, the protein solutions were diluted with Milli-Q water to a desired concentration. GM₁-os (β-Gal-(1→3)-β-GalNAc-(1→4)[α-Neu5Ac-(2→3)]-β-Gal-(1→4)-Glc, MW 998.9 Da) were purchased from Elicityl SA (Crolles, France). The stock solutions of GM₁-os were prepared by dissolving in Milli-Q water to yield a final concentration of 1 mM and were stored at -20 °C until needed.

4.2.2 HDX-MS

The PLIMSTEX experiments at peptide level was carried out using a Synapt G2-S HDMS mass spectrometer equipped with a nanoACQUITY UPLC system with HDX technology (Waters, UK) and a PAL HTX-xt system for automatic sample preparation and injection. As described similarly in Chapter 2, two sample stacks provided accurate temperature control for both labeling reactions (20 °C) and the reactions quench (1 °C). Protein samples (4 μ M CTB₅) alone or in the presence different concentrations of ligand (5, 16, 26, 40, 100 or 400 μ M GM₁-os) were diluted 15-fold with either equilibrium buffer (10 mM potassium phosphate in H₂O at pH 7.0) for control experiments ($t = 0$), or labeling buffer (10 mM potassium phosphate in D₂O at pD 7.0 (pD = pH + 0.4)) for labeling experiments. All the labeling samples were incubated for 10 min at 20 °C. After that, samples (both for control and labeling experiments) were quenched by quench buffer (4 M guanidine hydrochloride and 0.5 M *tris*(2-carboxyethyl)phosphine (TCEP) in H₂O at pH 2.6) with a 1:1 dilution at 1 °C. After being held for 30 s, quenched samples were injected into a 50 μ L injection loop of a nanoACQUITY UPLC system with HDX technology. Online digestion of intact protein was performed using an immobilized pepsin column (Life Technologies, Burlington, Canada) with 0.1% formic acid in H₂O with a flow rate of 200 μ L min⁻¹ at 20 °C. Peptide trapping was carried out online using an ACQUITY UPLC BEH C18 1.7 μ m VanGuard Precolumn (Waters) at 1 °C. The digestion, trapping and desalting process took 2 min. Peptide separation was achieved using an ACQUITY UPLC C18 1.7 μ m 1.0 \times 100mm column (Waters) with a 12 min gradient elution at a flow rate of 40 μ L min⁻¹. The content of solvent A in mobile phase was

decreased over a 7 min period from 95% to 63% and held constant for 1 min before reduction from 63% to 16% over 0.5 min. After holding constant at 16% for 0.5 min, solvent A was increased back to 95% over 0.5 min. (Solvent A, 0.1% formic acid and 5% acetonitrile in H₂O; solvent B, 0.1% formic acid in acetonitrile). The eluent was directly introduced to a Synapt G2-S HDMS using ESI source. Mass spectra were acquired in MS^E mode (m/z 50 to 2000) with a scan rate of 0.4 s/scan and lock-mass correction (using [Glu]-Fibrinopeptide). The capillary and cone voltages were kept constant at 3 kV and 40 V, respectively.

4.2.3 Data processing

ProteinLynx Global Server 2.5.2 software (PLGS, Waters) together with a databank containing Cholera toxin B monomer sequence was used to identify detected peptides from control samples (i.e., $t = 0$). The subset of peptides observed in all replicated control and labeling samples were considered further. DynamX 2.0 software (Waters, UK) was used to generate a peptide list, according to the common peptides of all samples identified by PLGS, and to assign corresponding peaks (with possibility of having more than one charge state) for peptides in control and labeling samples. The average molecular mass (\bar{M}) for each peptide was calculated by DynamX as the centroid of the entire envelope of corresponding isotopic peaks (all observed charge states were considered). To cancel out the back exchange during protein digestion and LC separation of peptides,^{43,44} relative D-uptake (D_i , unit of Da) value for a peptide at certain initial ligand concentration ($[L]_{0,i}$) was determined from Equation 4.1:

$$D_i = \overline{M}_{\text{label}, i} - \overline{M}_{\text{control}, i} \quad (4.1)$$

in which D_i is the difference of the average molecular mass obtained for certain peptide between labeled and control samples.

4.2.4 Fitting method

In the titration experiment, by fixing the initial protein concentration ($[P]_0$) and varying the initial ligand concentration ($[L]_{0,i}$), a number of corresponding D_i values for each peptide were obtained. At certain $[L]_{0,i}$, provided that the ratio of the ligand-bound protein ($[PL]_{\text{eq}}$) to the total protein ($[P]_{\text{T}} = [P]_{\text{eq}} + [PL]_{\text{eq}} = [P]_0$) under labeling condition can be calculated by Equation 4.2:²⁷

$$\frac{[PL]_{\text{eq}}}{[P]_0} = \frac{D_i - D_P}{D_{\text{PL}} - D_P} \quad (4.2)$$

where D_P and D_{PL} corresponding the D_i values of a certain peptide from free and ligand-bound CTB monomer. Therefore, Equation 4.2 related the experimentally obtained value D_i to $[PL]_{\text{eq}}$.

In this study, the positive cooperativity of the stepwise ligand binding was neglected and the five subunits of CTB₅ were assumed to be identical and bound to GM₁-os independently. Therefore, it can be treated as 1:1 binding between CTB monomer and GM₁-os. The general expression of K_a , for a 1:1 attachment of L to P (Equation 4.3) is given by Equation 4.4:³³



$$K_a = \frac{\frac{[PL]_{eq}}{[P]_0}}{\left(1 - \frac{[PL]_{eq}}{[P]_0}\right)([L]_{0,i} - [PL]_{eq})} \quad (4.4)$$

For the PLIMSTEX experiment, by using D_i to indicate $[PL]_{eq}$, the following expression can be obtained after rearrangement:

$$\frac{D_i - D_P}{D_{PL} - D_P} = \frac{1 + K_a[P]_0 + K_a[L]_{0,i} - \sqrt{(1 + K_a[P]_0 - K_a[L]_{0,i})^2 + 4K_a[L]_{0,i}}}{2K_a[P]_0} \quad (4.5)$$

A nonlinear fitting was then carried out based on Equation 4.5 using Origin 9.1. $[L]_{0,i}$ and D_i are dependent and independent variables for the fitting, K_a , D_P and D_{PL} are unknown parameters, and $[P]_0$ is a constant. K_a , D_P and D_{PL} can all be obtained through the fitting.

4.3 Results and Discussion

PLIMSTEX measurements were carried out in aqueous solutions of 4 μM CTB₅ (or 20 μM CTB monomer) and GM₁-os at concentrations ranging from 0 to 400 μM (0, 5, 16, 26, 40, 100, 200, 300, 400 μM , respectively) before diluting by equilibrium or labeling buffer. Same as the results in Chapter 2, pepsin digestion of CTB monomer produced around eighty peptides for each sample, among which around forty peptides were identified in all samples (both free and ligand-bound CTB monomer). In order to decrease the redundancy of each residue and ensure maximum sequence coverage, a peptide list containing 13 peptides were generated by DynamX for further analysis, covering 97.1% of CTB₅ sequence (Figure 4.3). The average

standard deviation of the D_i values for the 13 peptides was about 0.11 Da, based on three replicated data sets acquired on different days.

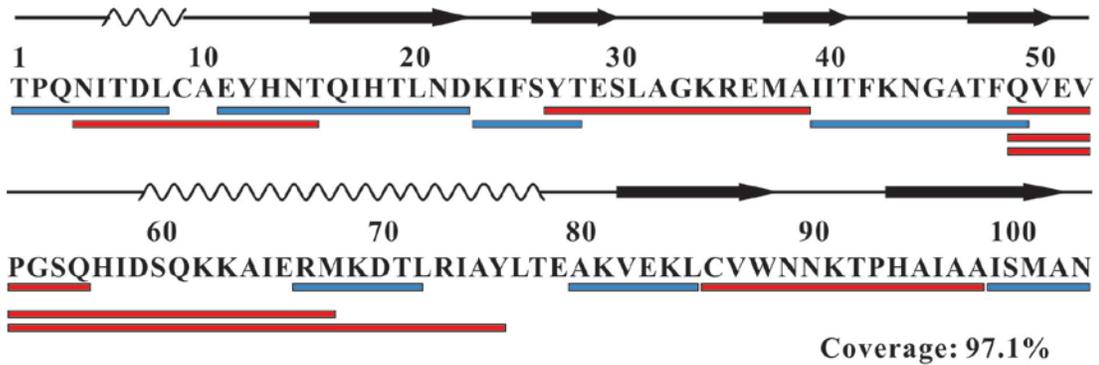


Figure 4.3 Sequence coverage (97.1%) of peptic peptides from the CTB monomer.

Each bar under the sequence indicates a common peptide reproducibly identified in all samples. A total of 13 peptides were selected for further analysis. Peptides bars labeled in red were the regions protected by the bound GM_1 -os. Secondly structural information is also illustrated in the map.

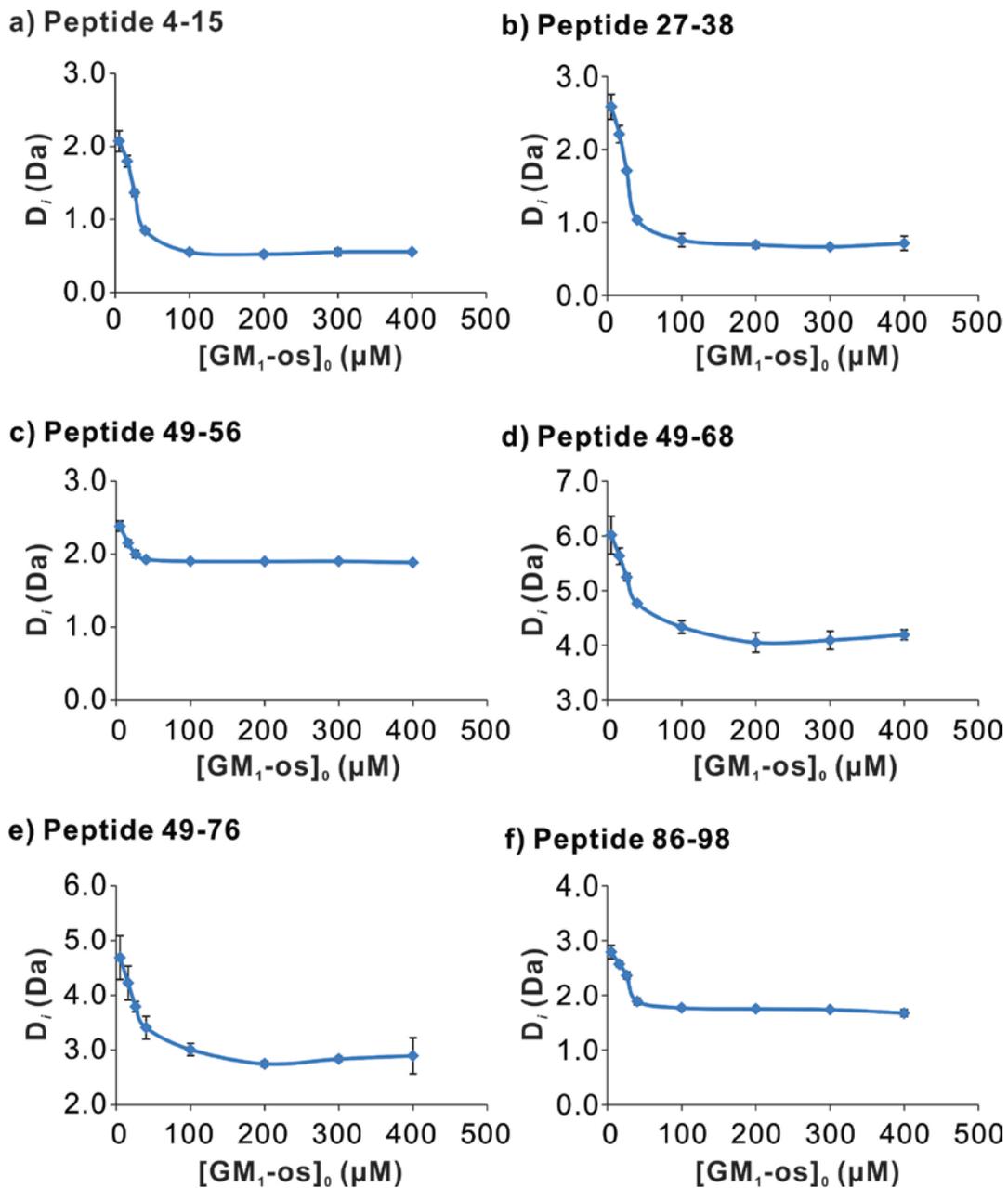


Figure 4.4 With an increase in concentration of GM_{1-os}, the D_i values for the six GM_{1-os} protected peptides from CTB monomer, (a) peptide 4-15, (b) peptide 27-38, (c) peptide 49-56, (d) peptide 49-68, (e) peptide 49-76 and (f) peptide 86-98, decreased. Data was obtained with 20 µM of CTB as initial concentration.

Altogether, there were six peptides showed that a decrease in D_i value in the presence of GM₁-os (Figure 4.4), whereas the D-uptake of the other seven peptides were not affect upon ligand binding. Based on previous X-ray crystallography data, the six affected peptides are located within the ligand binding site.^{42,45} Therefore they are protected from deuterium exchange by the bound ligand. And with an increase in ligand initial concentration, the fraction of protein-ligand complex also increased, leading to a decrease of D_i value. For example, in the presence of 5 μ M of GM₁-os, the D_i value for peptide 27-38 after 10 min labeling was 2.6 Da. When the GM₁-os initial concentration increased to 400 μ M, D_i decreases dramatically to only 0.7 Da.

Table 4.1 K_a , D_P and D_{PL} values obtained from the nonlinear fitting based on Equation 4.5 for the addition of GM₁-os to CTB monomer. Errors were calculated as the standard deviation for triplicate measurements.

Peptide	$K_{a,ave}$ ($10^6 M^{-1}$)	D_P (Da)	D_{PL} (Da)
4-15	1.61 \pm 0.16	2.5 \pm 0.2	0.4 \pm 0.0
27-38	1.72 \pm 0.20	3.2 \pm 0.2	0.6 \pm 0.1
49-56	13.66 \pm 12.03	2.5 \pm 0.1	1.9 \pm 0.0
49-68	1.24 \pm 0.47	6.5 \pm 0.4	4.0 \pm 0.2
49-76	1.77 \pm 0.85	5.1 \pm 0.5	2.8 \pm 0.3
86-98	1.42 \pm 0.32	3.1 \pm 0.1	1.6 \pm 0.1

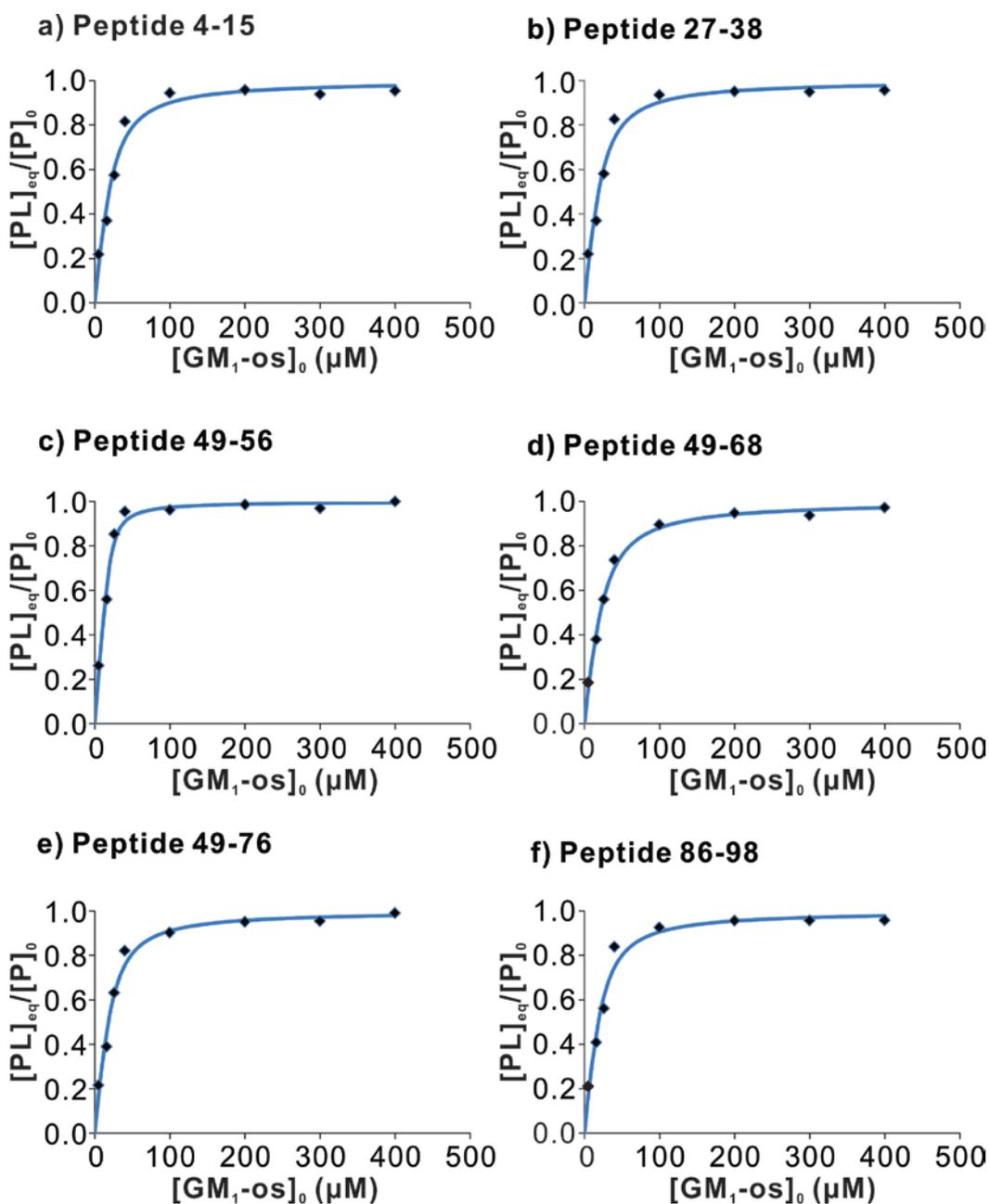


Figure 4.5 Nonlinear fitting based on Equation 4.5 (using Origin 9.1) for the six GM_1-os protected peptides from CTB monomer (a) peptide 4-15, (b) peptide 27-38, (c) peptide 49-56, (d) peptide 49-68, (e) peptide 49-76 and (f) peptide 86-98. Data was obtained with 20 μM of CTB as initial concentration.

The nonlinear fitting was then carried out for the six ligand-protected peptides based on Equation 4.5 (Figure 4.5). With known initial concentrations of CTB monomer and GM₁-os, and the experimentally determined [PL]_{eq} (Equation 4.2), the apparent binding affinity (K_a), as well as the two unknown parameters, D_P and D_{PL} , was obtained through fitting (Table 4.1).

According to Table 4.1, with the exception of peptide 49-56, the fitting for all the other five peptides gave an affinity of GM₁-os binding to CTB between $(1\sim 2) \times 10^6 \text{ M}^{-1}$ in average. The average ($D_{PL} - D_P$) values obtained from fitting for these peptides ranged from 0.6 to 2.6 Da, with peptide 49-56 exhibited the smallest ($D_{PL} - D_P$) values of 0.6 Da. This explains the large error of obtained affinity value associated with peptide 49-56. The typical error in D-uptake (or D_i) for HDX-MS measurements is about 0.15 Da (0.11 Da for this study).^{20,46} When using Equation 4.2, 0.11 Da error in D_i leads to 18% (0.11 Da divided by 0.6 Da) uncertainty in [PL]_{eq} values. Based on this, in order to use peptides as indicators to obtain accurate affinity data using PLIMSTEX measurements, a large difference between D_P and D_{PL} is required. The consistence of K_a obtained from the five peptides suggested that PLIMSTEX measurements using peptides as indicators are capable to obtain the protein-carbohydrate binding affinities. The average affinity obtained from the five peptides was of $(1.6 \pm 0.2) \times 10^6 \text{ M}^{-1}$. This number is in reasonable agreement with the reported value of $(3.2 \pm 0.2) \times 10^6 \text{ M}^{-1}$, which was measured using direct ESI-MS assay at pH 6.9 and room temperature when assuming the five binding sites of CTB₅ are equivalent and independent.³⁹

4.4 Conclusions

In summary, the application of peptide level PLIMSTEX measurement to obtain the binding affinity for protein-carbohydrate interaction is described. Using the experimentally acquired peptide D-uptake value to indicate the fraction of protein-ligand complex at equilibrium, together with known protein and ligand initial concentration, a nonlinear fitting was carried out to derive the unknown parameter, apparent association constant, for the interaction. Although the five ligand binding sites on CTB₅ are found to exhibit positive cooperativity, they were assumed equivalent and independent in this study. In this way, the average apparent association constant measured for the addition of GM₁-os to CTB at pH 7.0 and 20°C was found to be $(1.6 \pm 0.2) \times 10^6 \text{ M}^{-1}$. This is in reasonable agreement with the reported value of $(3.2 \pm 0.2) \times 10^6 \text{ M}^{-1}$, which was measured using direct ESI-MS assay at pH 6.9 and room temperature with the same assumption. Taken together, the results of this study suggest that peptide level PLIMSTEX can serve as effective method to obtain the binding affinities for protein-carbohydrate interactions. However, to obtain reliable affinity results, the peptide indicator is better to exhibit a large difference in D-uptake in the absence and presence of the ligand.

4.5 Literature Cited

- (1) Ghazarian, H., Idoni, B., Oppenheimer, S. B., *Acta Histochem.*, **2011**, *113*, 236-247.
- (2) Dwek, R. A., *Chem. Rev.*, **1996**, *96*, 683-720.

- (3) Roldos, V., Canada, F. J., Jimenez-Barbero, J., *Chem. Bio. Chem.*, **2011**, *12*, 990-1005.
- (4) Avci, F. Y., Kasper, D. L. In *Annual Review of Immunology*, Paul, W. E., Littman, D. R., Yokoyama, W. M., Annual Reviews: Palo Alto, **2010**, Vol. 28, p 107-130.
- (5) Daghestani, H. N., Day, B. W., *Sensors*, **2010**, *10*, 9630-9646.
- (6) Karlsson, R., *J. Mol. Recognit.*, **2004**, *17*, 151-161.
- (7) Homola, J., *Anal. Bioanal. Chem.*, **2003**, *377*, 528-539.
- (8) Bundle, D. R., Eichler, E., Gidney, M. A. J., Meldal, M., Ragauskas, A., Sigurskjold, B. W., Sinnott, B., Watson, D. C., Yaguchi, M., Young, N. M., *Biochemistry*, **1994**, *33*, 5172-5182.
- (9) Bundle, D. R., Sigurskjold, B. W., *Methods Enzymol.*, **1994**, *247*, 288-305.
- (10) Saboury, A. A., *J. Iran Chem. Soc.*, **2006**, *3*, 1-21.
- (11) Larsen, K., Thygesen, M. B., Guillaumie, F., Willats, W. G. T., Jensen, K. J., *Carbohydr. Res.*, **2006**, *341*, 1209-1234.
- (12) Fernandez-Alonso, M. D., Diaz, D., Berbis, M. A., Marcelo, F., Canada, J., Jimenez-Barbero, J., *Curr. Protein Pept. Sci.*, **2012**, *13*, 816-830.
- (13) Lundquist, J. J., Toone, E. J., *Chem. Rev.*, **2002**, *102*, 555-578.
- (14) Wishart, D., *Curr. Pharm. Biotechnol.*, **2005**, *6*, 105-120.
- (15) Hernychova, L., Man, P., Verma, C., Nicholson, J., Sharma, C. A., Ruckova, E., Teo, J. Y., Ball, K., Vojtesek, B., Hupp, T. R., *Proteomics*, **2013**, *13*, 2512-2525.
- (16) Percy, A. J., Rey, M., Burns, K. M., Schriemer, D. C., *Anal. Chim. Acta*, **2012**, *721*, 7-21.

- (17) Pacholarz, K. J., Garlish, R. A., Taylor, R. J., Barran, P. E., *Chem. Soc. Rev.*, **2012**, *41*, 4335-4355.
- (18) Ling, J. M. L., Silva, L., Schriemer, D. C., Schryvers, A. B., *Methods Mol. Biol.*, **2012**, *799*, 237-252.
- (19) Iacob, R. E., Engen, J. R., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1003-1010.
- (20) Houde, D., Berkowitz, S. A., Engen, J. R., *J. Pharm. Sci.*, **2011**, *100*, 2071-2086.
- (21) Jaswal, S. S., *BBA-Proteins Proteomics*, **2013**, *1834*, 1188-1201.
- (22) Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., Griffin, P. R., *Expert Rev. Proteomic*, **2011**, *8*, 43-59.
- (23) Engen, J. R., *Anal. Chem.*, **2009**, *81*, 7870-7875.
- (24) King, D., Lumpkin, M., Bergmann, C., Orlando, R., *Rapid Commun. Mass Spectrom.*, **2002**, *16*, 1569-1574.
- (25) Sperry, J. B., Huang, R. Y. C., Zhu, M. M., Rempel, D. L., Gross, M. L., *Int. J. Mass Spectrom.*, **2011**, *302*, 85-92.
- (26) Zhu, M. M., Rempel, D. L., Du, Z. H., Gross, M. L., *J. Am. Chem. Soc.*, **2003**, *125*, 5252-5253.
- (27) Zhu, M. M., Rempel, D. L., Gross, M. L., *J. Am. Soc. Mass Spectrom.*, **2004**, *15*, 388-397.
- (28) Zhu, M. M., Chitta, R., Gross, M. L., *Int. J. Mass Spectrom.*, **2005**, *240*, 213-220.
- (29) Sperry, J. B., Shi, X. G., Rempel, D. L., Nishimura, Y., Akashi, S., Gross, M. L., *Biochemistry*, **2008**, *47*, 1797-1807.

- (30) Tu, T. T., Dragusanu, M., Petre, B. A., Rempel, D. L., Przybylski, M., Gross, M. L., *J. Am. Soc. Mass Spectrom.*, **2010**, *21*, 1660-1667.
- (31) Huang, R. Y. C., Rempel, D. L., Gross, M. L., *Biochemistry*, **2011**, *50*, 5426-5435.
- (32) Lin, H., Kitova, E. N., Klassen, J. S., *Anal. Chem.*, **2013**, *85*, 8919-8922.
- (33) Kitova, E. N., El-Hawiet, A., Schnier, P. D., Klassen, J. S., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 431-441.
- (34) El-Hawiet, A., Liu, L., Kitova, E. N., Shoemaker, G., Klassen, J. S., *Biophys. J.*, **2012**, *102*, 1A-1A.
- (35) El-Hawiet, A., Kitova, E. N., Klassen, J. S., *Biochemistry*, **2012**, *51*, 4244-4253.
- (36) Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L., Read, R. J., *Biochemistry*, **1998**, *37*, 1777-1788.
- (37) Johannes, L., Romer, W., *Nat. Rev. Microbiol.*, **2010**, *8*, 105-116.
- (38) Sacchettini, J. C., Baum, L. G., Brewer, C. F., *Biochemistry*, **2001**, *40*, 3009-3015.
- (39) Lin, H., Kitova, E. N., Klassen, J. S., *J. Am. Soc. Mass Spectrom.*, **2014**, *25*, 104-110.
- (40) Kuziemko, G. M., Stroh, M., Stevens, R. C., *Biochemistry*, **1996**, *35*, 6375-6384.
- (41) Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K., Ho, W. G. J., *J. Mol. Biol.*, **1998**, *282*, 1043-1059.
- (42) Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., Hol, W. G. J., *Protein Sci.*, **1997**, *6*, 1516-1528.
- (43) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.

- (44) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (45) Merritt, E. A., Sarfaty, S., Vandenakker, F., Lhoir, C., Martial, J. A., Hol, W. G. J., *Protein Sci.*, **1994**, *3*, 166-175.
- (46) Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., Chen, G. D., *J. Am. Soc. Mass, Spectrom.*, **2012**, *23*, 498-504.

Chapter 5

Conclusions and Future Work

This work describes the application of HDX-MS methods to study the non-covalent protein-carbohydrate interactions. The first two research projects focus on the localization of carbohydrate binding sites on proteins. The third research project highlighted the potential of the HDX-MS based method, PLIMSTEX, for quantifying protein-carbohydrate binding affinities.

In chapter 2, the application of HDX-MS to localize carbohydrate binding sites for three different proteins is described as model systems. Comparisons of the differences in deuterium uptake measured for peptic peptides, produced from the bacterial toxins in the absence and presence of ligand, revealed regions of the protein that are protected against deuterium exchange upon ligand binding. For all three toxins, the peptides found to exhibit protection upon ligand binding are associated with, at least some of the carbohydrate binding sites identified by X-ray crystallography. For CTB₅ and GM₁-os interactions, peptides associated with the four loops of CTB monomer that made up the ligand binding site were found to be protected by the bound ligand. For Stx1B₅ and Pk-OH interactions, protection of peptides within the three known Pk-OH binding sites was also observed. However, occupation of all the three ligand binding sites could not be established under our experimental conditions. For TcdA-A2 and CD-grease interactions, one of the two known binding sites close to the N-terminus was identified. However, large regions of TcdA-A2 showed de-protection upon CD-grease binding, indicating the presence of

ligand induced protein conformational change for this flexible fragment. Taken together, the results of this study suggest that HDX-MS can serve as an effective tool for localizing ligand binding sites in carbohydrate-binding proteins.

In chapter 3, HDX-MS was applied to explore the existence of distinct HMOs binding sites on bacterial toxins. For CTB₅, a novel binding site was localized for 2'-FL, different from the one for native receptor GM₁. For TcdA-A2, however, the LNT binding site was the same as its native carbohydrate receptors. These results are consistent with previous direct ESI-MS data, suggesting that HMOs may have distinct binding sites on toxins depending on the nature of the toxin.

In Chapter 4, the application of peptide level PLIMSTEX measurement to obtain the binding affinity for protein-carbohydrate interaction is described. Using the experimentally acquired peptide D-uptake value to indicate the fraction of protein-ligand complex at equilibrium, together with known protein and ligand initial concentration, a nonlinear fitting was carried out to derive the unknown parameter, apparent association constant, for the interaction. Although the five ligand binding sites on CTB₅ are found to exhibit positive cooperativity, they were assumed equivalent and independent in this study. In this way, the average apparent association constant measured for the addition of GM₁-os to CTB at pH 7.0 and 20°C was found to be $(1.6 \pm 0.2) \times 10^6 \text{ M}^{-1}$. This is in reasonable agreement with the reported value of $(3.2 \pm 0.2) \times 10^6 \text{ M}^{-1}$, which was measured using direct ESI-MS assay at pH 6.9 and room temperature with the same assumption. Taken together, the results of this study suggest that peptide level PLIMSTEX can serve as effective method to obtain the binding affinities for protein-carbohydrate interactions.

However, to obtain more reliable affinity results, the peptide indicators are better to exhibit a large difference in D-uptake in the absence and presence of the ligand.

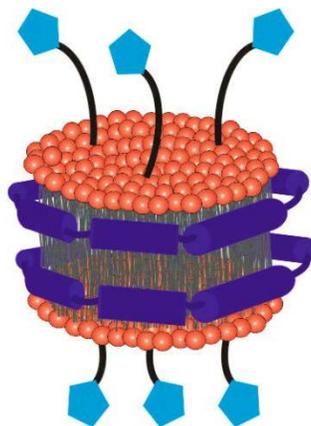


Figure 5.1 Cartoon representatives of nanodiscs. Purple belts are two copies of MSPs surrounding the phospholipid bilayer with incorporated receptors.

There are several possible extensions of the current studies. As mentioned in chapter 2, soluble carbohydrate portion of the receptors were used to localize the ligand binding sites. However, the native receptors for the three toxins in *vivo* exist as glycolipids with their lipid tails buried inside the cell membrane. Undoubtedly, direct investigation of the interactions between proteins and glycolipids are better reflections of what happens in *vivo*. The question then becomes how to dissolve the glycolipids in aqueous solutions. Fortunately, the development of nanodiscs (NDs) work in our lab can help with this.¹ NDs consist of a discoidal phospholipid bilayer which is encirculated by two copies of an amphipathic membrane scaffold protein (MSP) (Figure 5.1).² Because of the amphipathic characteristic of MSPs, the hydrophobic acyl chain is shielded from the aqueous solution, thus the NDs are

water-soluble.³ When the glycolipids are incorporated into the ND, they become soluble in the aqueous solution as well, and more importantly, remain active in a native-like bilayer environment. The uniform size of NDs are achieved by optimizing the lipid : MSP : receptor stoichiometry in the self-assembly process.⁴

However, the phospholipids from the digested NDs are not compatible with the LC-MS system for HDX analysis. This is because the phospholipids tend to be retained on the reverse phase column during LC separation and dramatically reduce the column life time and retention time reproducibility.⁵ Therefore, a phospholipid removal step is needed prior to chromatographic separation of peptides. Using zirconium oxide (ZrO_2) coated silica beads to remove the phospholipids, Engen and his coworkers have successfully carried out HDX-MS measurement using NDs for membrane protein studies.⁵⁻⁷ In the presence of phospholipids, ZrO_2 serves as a Lewis acid and interacts strongly with the phosphate group, thus removing the phospholipids. However, during the sample incubation with ZrO_2 coated silica beads, peptide adsorption to the silica is inevitably, leading to a peptide loss. Efforts of looking for better ways of dealing with NDs and phospholipid removal are still needed.

An alternative to ND is Saposin A disc (SapA disc or picodisc), which is a soluble phospholipid bilayer encirculated by two copies of SapA protein.⁸ Since much less phospholipids are incorporated into a SapA disc (~15 lipids/disc) than a ND (~200 lipids/disc), it's easier to work with Sap A disc and achieve the same purpose, that is to dissolve the glycolipid in aqueous solution.

Another possible extension of the current work is to investigate the relationship between protein-carbohydrate interactions and observed D-uptake difference between free and ligand-bound protein. According to the model system studies, in most cases ligand binding can lead to a decrease in D-uptake. However, increased or unchanged D-uptake was also observed upon ligand binding. What's the origin behind these phenomena in general? As mentioned in Chapter 2, the intermolecular H-bonds formed between the ligand and the backbone amide H's are more likely to affect the HDX rate of a protein than the H-bonds formed between the ligand and the protein side chains or carbonyl oxygens. Besides, ligand induced protein conformation or dynamic change will also change the HDX rate of the protein. Recently, a paper has been published by Konermann and his coworkers, looking into the ligand effect on protein HDX rates.⁹ They found that ligand binding can decrease, increase, or unaffected the HDX rates of backbone amide H's throughout the protein due to protein free energy change upon ligand binding. All these factors explain part of the story of the effect of ligand binding on protein HDX kinetics. How to differentiate between these factors? Are there more possible affecting factors? To answer these questions, however, more efforts are needed.

5.1 Literature Cited

- (1) Zhang, Y. X., Liu, L., Daneshfar, R., Kitova, E. N., Li, C. S., Jia, F., Cairo, C. W., Klassen, J. S., *Anal. Chem.*, **2012**, *84*, 7618-7621.
- (2) Bayburt, T. H., Sligar, S. G., *FEBS Lett.*, **2010**, *584*, 1721-1727.
- (3) Shaw, A. W., McLean, M. A., Sligar, S. G., *FEBS Lett.*, **2004**, *556*, 260-264.

- (4) Boldog, T., Li, M. S., Hazelbauer, G. L. In *Two-Component Signaling Systems*, Simon, M. I., Crane, B. R., Crane, A., Elsevier Academic Press Inc: San Diego, **2007**, Vol. 423, p 317-335.
- (5) Hebling, C. M., Morgan, C. R., Stafford, D. W., Jorgenson, J. W., Rand, K. D., Engen, J. R., *Anal. Chem.*, **2010**, *82*, 5415-5419.
- (6) Nasr, M. L., Shi, X. M., Bowman, A. L., Johnson, M., Zvonok, N., Janero, D. R., Vemuri, V. K., Wales, T. E., Engen, J. R., Makriyannis, A., *Protein Sci.*, **2013**, *22*, 774-787.
- (7) Parker, C. H., Morgan, C. R., Rand, K. D., Engen, J. R., Jorgenson, J. W., Stafford, D. W., *Biochemistry*, **2014**, *53*, 1511-1520.
- (8) Popovic, K., Holyoake, J., Pomes, R., Prive, G. G., *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, *109*, 2908-2912.
- (9) Sowole, M. A., Konermann, L., *Anal. Chem.*, **2014**, *86*, 8-15.

Literature Cited

Chapter 1

- (1) Dwek, R. A., *Chem. Rev.*, **1996**, *96*, 683-720.
- (2) Ghazarian, H., Idoni, B., Oppenheimer, S. B., *Acta Histochem.*, **2011**, *113*, 236-247.
- (3) Vuignier, K., Schappler, J., Veuthey, J. L., Carrupt, P. A., Martel, S., *Anal. Bioanal. Chem.*, **2010**, *398*, 53-66.
- (4) Renaud, J. P., Delsuc, M. A., *Curr. Opin. Pharmacol.*, **2009**, *9*, 622-628.
- (5) Pacholarz, K. J., Garlish, R. A., Taylor, R. J., Barran, P. E., *Chem. Soc. Rev.*, **2012**, *41*, 4335-4355.
- (6) Zeng, X., Andrade, C. A. S., Oliveira, M. D. L., Sun, X. L., *Anal. Bioanal. Chem.*, **2012**, *402*, 3161-3176.
- (7) Lee, Y. C., Lee, R. T., *Accounts Chem. Res.*, **1995**, *28*, 321-327.
- (8) Holgersson, J., Gustafsson, A., Breimer, M. E., *Immunol. Cell Biol.*, **2005**, *83*, 694-708.
- (9) Kamiya, Y., Yagi-Utsumi, M., Yagi, H., Kato, K., *Current Pharm. Design*, **2011**, *17*, 1672-1684.
- (10) Schotte, F., Lim, M. H., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Wulff, M., Anfirud, P. A., *Science*, **2003**, *300*, 1944-1947.
- (11) Fernandez-Alonso, M. D., Diaz, D., Berbis, M. A., Marcelo, F., Canada, J., Jimenez-Barbero, J., *Curr. Protein Pept. Sci.*, **2012**, *13*, 816-830.
- (12) Quijcho, F. A., Spurlino, J. C., Rodseth, L. E., *Structure*, **1997**, *5*, 997-1015.
- (13) Duan, X. Q., Quijcho, F. A., *Biochemistry*, **2002**, *41*, 706-712.

- (14) Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, H. L., Deng, S. J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L. P., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., Shewchuk, L. M., *Acta Crystallogr. Sect. D-Biol. Crystallogr.*, **2007**, *63*, 72-79.
- (15) Mittermaier, A., Kay, L. E., *Science*, **2006**, *312*, 224-228.
- (16) Sprangers, R., Kay, L. E., *Nature*, **2007**, *445*, 618-622.
- (17) Kohn, J. E., Afonine, P. V., Ruscio, J. Z., Adams, P. D., Head-Gordon, T., *PLoS Comput. Biol.*, **2010**, *6*, e1000911.
- (18) Weis, W. I., Drickamer, K., Hendrickson, W. A., *Nature*, **1992**, *360*, 127-134.
- (19) Ehring, H., *Anal. Biochem.*, **1999**, *267*, 252-259.
- (20) King, D., Lumpkin, M., Bergmann, C., Orlando, R., *Rapid Commun. Mass Spectrom.*, **2002**, *16*, 1569-1574.
- (21) Seyfried, N. T., Atwood, J. A., Yongye, A., Almond, A., Day, A. J., Orlando, R., Woods, R. J., *Rapid Commun. Mass Spectrom.*, **2007**, *21*, 121-131.
- (22) Baldwin, R. L., *Proteins*, **2011**, *79*, 2021-2026.
- (23) Hvidt, A., Linderstromlang, K., *Biochim. Biophys. Acta*, **1954**, *14*, 574-575.
- (24) Englander, S. W., Mayne, L., Bai, Y., Sosnick, T. R., *Protein Sci.*, **1997**, *6*, 1101-1109.
- (25) Skinner, J. J., Lim, W. K., Bedard, S., Black, B. E., Englander, S. W., *Protein Sci.*, **2012**, *21*, 987-995.
- (26) Skinner, J. J., Lim, W. K., Bedard, S., Black, B. E., Englander, S. W., *Protein Sci.*, **2012**, *21*, 996-1005.
- (27) Rosa, J. J., Richards, F. M., *J. Mol. Biol.*, **1979**, *133*, 399-416.

- (28) Brock, A., *Protein Expr. Purif.*, **2012**, *84*, 19-37.
- (29) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.
- (30) Katta, V., Chait, B. T., *Rapid Commun. Mass Spectrom.*, **1991**, *5*, 214-217.
- (31) Konermann, L., Tong, X., Pan, Y., *J. Mass Spectrom.*, **2008**, *43*, 1021-1036.
- (32) Konermann, L., Pan, J. X., Liu, Y. H., *Chem. Soc. Rev.*, **2011**, *40*, 1224-1234.
- (33) Wei, H., Mo, J. J., Tao, L., Russell, R. J., Tymiak, A. A., Chen, G. D., Jacob, R. E., Engen, J. R., *Drug Discov. Today*, **2014**, *19*, 95-102.
- (34) Jaswal, S. S., *BBA-Proteins Proteomics*, **2013**, *1834*, 1188-1201.
- (35) Wang, L., Smith, D. L. In *Current Protocols in Protein Science*, Coligan, J.E., John Wiley and Sons.: New York, **2002**, Vol. Chapter 17, p Unit 17.6.
- (36) Percy, A. J., Rey, M., Burns, K. M., Schriemer, D. C., *Anal. Chim. Acta*, **2012**, *721*, 7-21.
- (37) Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., Griffin, P. R., *Expert Rev. Proteomic*, **2011**, *8*, 43-59.
- (38) Jacob, R. E., Engen, J. R., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1003-1010.
- (39) Engen, J. R., *Anal. Chem.*, **2009**, *81*, 7870-7875.
- (40) Engen, J. R., Smith, D. L., *Anal. Chem.*, **2001**, *73*, 256A-265A.
- (41) Sowole, M. A., Konermann, L., *Anal. Chem.*, **2014**, *86*, 8.
- (42) Englander, S. W., Downer, N. W., Teitelba.H, *Annu. Rev. Biochem.*, **1972**, *41*, 903-924.
- (43) Eigen, M., *Angew. Chem.-Int. Edit.*, **1964**, *3*, 1-72.
- (44) Englander, S. W., Kallenbach, N. R., *Q. Rev. Biophys.*, **1983**, *16*, 521-655.
- (45) Molday, R. S., Englander, S. W., Kallenrg, *Biochemistry*, **1972**, *11*, 150-158.

- (46) Hvidt, A., *Comptes Rendus Des Travaux Du Laboratoire Carlsberg*, **1963**, *34*, 299-317.
- (47) Englander, S. W., Sosnick, T. R., Englander, J. J., Mayne, L., *Curr. Opin. Struct. Biol.*, **1996**, *6*, 18-23.
- (48) Bai, Y. W., Milne, J. S., Mayne, L., Englander, S. W., *Proteins*, **1993**, *17*, 75-86.
- (49) Smith, D. L., Deng, Y. Z., Zhang, Z. Q., *J. Mass Spectrom.*, **1997**, *32*, 135-146.
- (50) Truhlar, S. M. E., Croy, C. H., Torpey, J. W., Koeppe, J. R., Komives, E. A., *J. Am. Soc. Mass Spectrom.*, **2006**, *17*, 1490-1497.
- (51) Jane Clarke, L. S. I., *Curr. opin. struct. biol.*, **1998**, *8*, 112-118.
- (52) Lobanov, M. Y., Suvorina, M. Y., Dovidchenko, N. V., Sokolovskiy, I. V., Surin, A. K., Galzitskaya, O. V., *Bioinformatics*, **2013**, *29*, 1375-1381.
- (53) Ferraro, D. M., Robertson, A. D., *Biochemistry*, **2004**, *43*, 587-594.
- (54) Wales, T. E., Engen, J. R., *J. Mol. Biol.*, **2006**, *357*, 1592-1604.
- (55) Morgan, C. R., Engen, J. R. In *Current Protocols in Protein Science*, Coligan, J.E., John Wiley and Sons.: New York, **2009**, Vol. Chapter 17, p Unit 17.617.
- (56) Tsutsui, Y., Wintrode, P. L., *Curr. Med. Chem.*, **2007**, *14*, 2344-2358.
- (57) Karageorgos, I., Wales, T. E., Janero, D. R., Zvonok, N., Vemuri, V. K., Engen, J. R., Makriyannis, A., *Biochemistry*, **2013**, *52*, 5016-5026.
- (58) Lee, C. T., Graf, C., Mayer, F. J., Richter, S. M., Mayer, M. P., *Embo J.*, **2012**, *31*, 1518-1528.
- (59) Pan, L. Y., Salas-Solano, O., Valliere-Douglass, J. F., *Anal. Chem.*, **2014**, *86*, 2657-2664.

- (60) Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., Chen, G. D., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 498-504.
- (61) Zhang, H. M., McLoughlin, S. M., Frausto, S. D., Tang, H. L., Emmett, M. R., Marshall, A. G., *Anal. Chem.*, **2010**, *82*, 1450-1454.
- (62) Pan, Y., Brown, L., Konermann, L., *J. Am. Chem. Soc.*, **2011**, *133*, 20237-20244.
- (63) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (64) Palashoff, M. H., Chemistry Master's Theses, Northeastern University, **2008**.
- (65) Kan, Z. Y., Walters, B. T., Mayne, L., Englander, S. W., *P Natl. Acad. Sci. U.S.A.*, **2013**, *110*, 16438-16443.
- (66) Fajer, P. G., Bou-Assaf, G. M., Marshall, A. G., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1202-1208.
- (67) Ahn, J., Cao, M. J., Yu, Y. Q., Engen, J. R., *BBA-Proteins Proteomics*, **2013**, *1834*, 1222-1229.
- (68) Rey, M., Yang, M. L., Burns, K. M., Yu, Y. P., Lees-Miller, S. P., Schriemer, D. C., *Molecular & Cellular Proteomics*, **2013**, *12*, 464-472.
- (69) Cravello, L., Lascoux, D., Forest, E., *Rapid Commun. Mass Spectrom.*, **2003**, *17*, 2387-2393.
- (70) Wang, L. T., Pan, H., Smith, D. L., *Mol. Cell. Proteomics*, **2002**, *1*, 132-138.
- (71) Hamuro, Y., Coales, S. J., Molnar, K. S., Tuske, S. J., Morrow, J. A., *Rapid Commun. Mass Spectrom.*, **2008**, *22*, 1041-1046.
- (72) Sheff, J. G., Rey, M., Schriemer, D. C., *J. Am. Soc. Mass Spectrom.*, **2013**, *24*, 1006-1015.

- (73) Wu, Y., Engen, J. R., Hobbins, W. B., *J. Am. Soc. Mass Spectrom.*, **2006**, *17*, 163-167.
- (74) Wales, T. E., Fadgen, K. E., Gerhardt, G. C., Engen, J. R., *Anal. Chem.*, **2008**, *80*, 6815-6820.
- (75) Plumb, R. S., Johnson, K. A., Rainville, P., Smith, B. W., Wilson, I. D., Castro-Perez, J. M., Nicholson, J. K., *Rapid Commun. Mass Spectrom.*, **2006**, *20*, 1989-1994.
- (76) Chalmers, M. J., Busby, S. A., Pascal, B. D., He, Y. J., Hendrickson, C. L., Marshall, A. G., Griffin, P. R., *Anal. Chem.*, **2006**, *78*, 1005-1014.
- (77) Chalmers, M. J., Busby, S. A., Pascal, B. D., Southern, M. R., Griffin, P. R., *J. Biomol. Technol.*, **2007**, *18*, 194-204.
- (78) Burkitt, W., O'Connor, G., *Rapid Commun. Mass Spectrom.*, **2008**, *22*, 3893-3901.
- (79) Kebarle, P., Tang, L., *Anal. Chem.*, **1993**, *65*, A972-A986.
- (80) Wu, X. Y., Oleschuk, R. D., Cann, N. M., *Analyst*, **2012**, *137*, 4150-4161.
- (81) Thomson, B. A., Iribarne, J. V., *J. Chem. Phys.*, **1979**, *71*, 4451-4463.
- (82) Dole, M., Mack, L. L., Hines, R. L., *J. Chem. Phys.*, **1968**, *49*, 2240-2249.
- (83) Ahadi, E., Konermann, L., *J. Phys. Chem. B*, **2012**, *116*, 104-112.
- (84) Konermann, L., Ahadi, E., Rodriguez, A. D., Vahidi, S., *Anal. Chem.*, **2013**, *85*, 2-9.
- (85) Konermann, L., Vahidi, S., Sowole, M. A., *Anal. Chem.*, **2014**, *86*, 213-232.
- (86) Iribarne, J. V., Thomson, B. A., *J. Chem. Phys.*, **1976**, *64*, 2287-2294.
- (87) De Hoffmann, E., Stroobant, V., *Mass Spectrometry: Principles and Applications*, 3 ed., John Wiley & Sons.: New York, **2007**.
- (88) Miller, P. E., Denton, M. B., *J. Chem. Educ.*, **1986**, *63*, 617-622.
- (89) Wollnik, H., *Mass Spectrom. Rev.*, **1993**, *12*, 89-114.

- (90) El-Aneed, A., Cohen, A., Banoub, J., *Appl. Spectrosc. Rev.*, **2009**, *44*, 210-230.
- (91) Huang, R. Y. C., Wen, J., Blankenship, R. E., Gross, M. L., *Biochemistry*, **2012**, *51*, 187-193.
- (92) Chetty, P. S., Mayne, L., Kan, Z. Y., Lund-Katz, S., Englander, S. W., Phillips, M. C., *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, *109*, 11687-11692.
- (93) Ahn, J., Engen, J. R., *Chim. Oggi-Chem. Today*, **2013**, *31*, 25-28.
- (94) Clementi, N., Mancini, N., Criscuolo, E., Cappelletti, F., Clementi, M., Burioni, R. In *Monoclonal Antibodies: Methods and Protocols*, 2 ed., Ossipow, V., Fischer, N., Humana Press Inc: Totowa, **2014**, Vol. 1131, p 427-446.
- (95) Jorgensen, T. J. D., Gardsvoll, H., Dano, K., Roepstorff, P., Ploug, M., *Biochemistry*, **2004**, *43*, 15044-15057.
- (96) Dai, S. Y., Burris, T. P., Dodge, J. A., Montrose-Rafizadeh, C., Wang, Y., Pascal, B. D., Chalmers, M. J., Griffin, P. R., *Biochemistry*, **2009**, *48*, 9668-9676.
- (97) Hamuro, Y., Coales, S. J., Morrow, J. A., Molnar, K. S., Tuske, S. J., Southern, M. R., Griffin, P. R., *Protein Sci.*, **2006**, *15*, 1883-1892.
- (98) Pan, Y., Piyadasa, H., O'Neil, J. D., Konermann, L., *J. Mol. Biol.*, **2012**, *416*, 400-413.
- (99) Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Garcia-Ordonez, R. D., Pascal, B. D., Burris, T. P., Dodge, J. A., Griffin, P. R., *Structure*, **2010**, *18*, 1332-1341.
- (100) Sperry, J. B., Huang, R. Y. C., Zhu, M. M., Rempel, D. L., Gross, M. L., *Int. J. Mass Spectrom.*, **2011**, *302*, 85-92.
- (101) Zhang, X., Chien, E. Y. T., Chalmers, M. J., Pascal, B. D., Gatchalian, J., Stevens, R. C., Griffin, P. R., *Anal. Chem.*, **2010**, *82*, 1100-1108.

(102) King, D., Bergmann, C., Orlando, R., Benen, J. A. E., Kester, H. C. M., Visser, J., *Biochemistry*, **2002**, *41*, 10225-10233.

Chapter 2

(1) Vuignier, K., Schappler, J., Veuthey, J. L., Carrupt, P. A., Martel, S., *Anal. Bioanal. Chem.*, **2010**, *398*, 53-66.

(2) Lee, Y. C., Lee, R. T., *Accounts Chem. Res.*, **1995**, *28*, 321-327.

(3) Holgersson, J., Gustafsson, A., Breimer, M. E., *Immunolo. Cell Biol.*, **2005**, *83*, 694-708.

(4) Kamiya, Y., Yagi-Utsumi, M., Yagi, H., Kato, K., *Curr. Pharm. Design*, **2011**, *17*, 1672-1684.

(5) Duan, X. Q., Hall, J. A., Nikaido, H., Quijcho, F. A., *J. Mol. Biol.*, **2001**, *306*, 1115-1126.

(6) Duan, X. Q., Quijcho, F. A., *Biochemistry*, **2002**, *41*, 706-712.

(7) Fernandez-Alonso, M. D., Diaz, D., Berbis, M. A., Marcelo, F., Canada, J., Jimenez-Barbero, J., *Curr. Protein Pept. Sc.*, **2012**, *13*, 816-830.

(8) Kohn, J. E., Afonine, P. V., Ruscio, J. Z., Adams, P. D., Head-Gordon, T., *PLoS Comput. Biol.*, **2010**, *6*, e1000911.

(9) Percy, A. J., Rey, M., Burns, K. M., Schriemer, D. C., *Anal. Chim. Acta*, **2012**, *721*, 7-21.

(10) Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., Griffin, P. R., *Expert Rev. Proteomic*, **2011**, *8*, 43-59.

(11) Jacob, R. E., Engen, J. R., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 1003-1010.

- (12) Garcia, R. A., Pantazatos, D., Villarreal, F. J., *Assay Drug Dev. Technol.*, **2004**, *2*, 81-91.
- (13) Houde, D., Berkowitz, S. A., Engen, J. R., *J. Pharm. Sci.*, **2011**, *100*, 2071-2086.
- (14) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.
- (15) Pacholarz, K. J., Garlish, R. A., Taylor, R. J., Barran, P. E., *Chem. Soc. Rev.*, **2012**, *41*, 4335-4355.
- (16) Huang, R. Y. C., Wen, J., Blankenship, R. E., Gross, M. L., *Biochemistry*, **2012**, *51*, 187-193.
- (17) Wildes, D., Marqusee, S., *Protein Sci.*, **2005**, *14*, 81-88.
- (18) Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Garcia-Ordonez, R. D., Pascal, B. D., Burris, T. P., Dodge, J. A., Griffin, P. R., *Structure*, **2010**, *18*, 1332-1341.
- (19) Chetty, P. S., Mayne, L., Kan, Z. Y., Lund-Katz, S., Englander, S. W., Phillips, M. C., *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, *109*, 11687-11692.
- (20) Ahn, J., Engen, J. R., *Chim. Oggi-Chem. Today*, **2013**, *31*, 25-28.
- (21) Jorgensen, T. J. D., Gardsvoll, H., Dano, K., Roepstorff, P., Ploug, M., *Biochemistry*, **2004**, *43*, 15044-15057.
- (22) Dai, S. Y., Burris, T. P., Dodge, J. A., Montrose-Rafizadeh, C., Wang, Y., Pascal, B. D., Chalmers, M. J., Griffin, P. R., *Biochemistry*, **2009**, *48*, 9668-9676.
- (23) Hamuro, Y., Coales, S. J., Morrow, J. A., Molnar, K. S., Tuske, S. J., Southern, M. R., Griffin, P. R., *Protein Sci.*, **2006**, *15*, 1883-1892.
- (24) Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., Chen, G. D., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 498-504.
- (25) Pan, Y., Brown, L., Konermann, L., *J. Am. Chem. Soc.*, **2011**, *133*, 20237-20244.

- (26) Pan, Y., Piyadasa, H., O'Neil, J. D., Konermann, L., *J. Mol. Biol.*, **2012**, *416*, 400-413.
- (27) King, D., Bergmann, C., Orlando, R., Benen, J. A. E., Kester, H. C. M., Visser, J., *Biochemistry*, **2002**, *41*, 10225-10233.
- (28) King, D., Lumpkin, M., Bergmann, C., Orlando, R., *Rapid Commun. Mass Spectrom.*, **2002**, *16*, 1569-1574.
- (29) Seyfried, N. T., Atwood, J. A., Yongye, A., Almond, A., Day, A. J., Orlando, R., Woods, R. J., *Rapid Commun. Mass Spectrom.*, **2007**, *21*, 121-131.
- (30) Huzil, J. T., Chik, J. K., Slysz, G. W., Freedman, H., Tuszynski, J., Taylor, R. E., Sackett, D. L., Schriemer, D. C., *J. Mol. Biol.*, **2008**, *378*, 1016-1030.
- (31) Lee, C. T., Graf, C., Mayer, F. J., Richter, S. M., Mayer, M. P., *Embo J.*, **2012**, *31*, 1518-1528.
- (32) Bennett, M. J., Barakat, K., Huzil, J. T., Tuszynski, J., Schriemer, D. C., *Chem. Biol.*, **2010**, *17*, 725-734.
- (33) Lin, H., Kitova, E. N., Klassen, J. S., *J. Am. Soc. Mass Spectrom.*, **2014**, *25*, 104-110.
- (34) Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., Hol, W. G. J., *Protein Sci.*, **1997**, *6*, 1516-1528.
- (35) Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L., Read, R. J., *Biochemistry*, **1998**, *37*, 1777-1788.
- (36) Johannes, L., Romer, W., *Nat. Rev. Microbiol.*, **2010**, *8*, 105-116.
- (37) Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L., Read, R. J., *Nature*, **1992**, *355*, 748-750.

- (38) Greco, A., Ho, J. G. S., Lin, S. J., Palcic, M. M., Rupnik, M., Ng, K. K. S., *Nat. Struct. Mol. Biol.*, **2006**, *13*, 460-461.
- (39) Dingle, T., Wee, S., Mulvey, G. L., Greco, A., Kitova, E. N., Sun, J. X., Lin, S. J., Klassen, J. S., Palcic, M. M., Ng, K. K. S., Armstrong, G. D., *Glycobiology*, **2008**, *18*, 698-706.
- (40) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (41) Merritt, E. A., Sarfaty, S., Vandenakker, F., Lhoir, C., Martial, J. A., Hol, W. G. J., *Protein Sci.*, **1994**, *3*, 166-175.
- (42) Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K., Ho, W. G. J., *J. Mol. Biol.*, **1998**, *282*, 1043-1059.
- (43) Skinner, J. J., Lim, W. K., Bedard, S., Black, B. E., Englander, S. W., *Protein Sci.*, **2012**, *21*, 996-1005.
- (44) Shimizu, H., Field, R. A., Homans, S. W., Donohue-Rolfe, A., *Biochemistry*, **1998**, *37*, 11078-11082.

Chapter 3

- (1) German, J. B., Freeman, S. L., Lebrilla, C. B., Mills, D. A. In *Personalized Nutrition for the Diverse Needs of Infants and Children*, Bier, D. M., German, J. B., Lonnerdal, B., Karger: Basel, **2008**, Vol. 62, p 205-222.
- (2) Field, C. J., *J. Nutr.*, **2005**, *135*, 1-4.
- (3) Lara-Villoslada, F., Olivares, M., Sierra, S., Rodriguez, J. M., Boza, J., Xaus, J., *Br. J. Nutr.*, **2007**, *98*, S96-S100.
- (4) Hoppu, U., Kalliomaki, M., Laiho, K., Isolauri, E., *Allergy*, **2001**, *56*, 23-26.

- (5) Morrow, A. L., Ruiz-Palacios, G. M., Jiang, X., Newburg, D. S., *J. Nutr.*, **2005**, *135*, 1304-1307.
- (6) Newburg, D. S., *J. Nutr.*, **2005**, *135*, 1308-1312.
- (7) Newburg, D. S., Ruiz-Palacios, G. M., Morrow, A. L. In *Annu. Rev. Nutr.*, Annual Reviews: Palo Alto, **2005**, Vol. 25, p 37-58.
- (8) Morrow, A. L., Ruiz-Palacios, G. M., Altaye, M., Jiang, X., Guerrero, M. L., Meinzen-Derr, J. K., Farkas, T., Chaturvedi, P., Pickering, L. K., Newburg, D. S., *J. Pediatr.*, **2004**, *145*, 297-303.
- (9) Kunz, C., Rudloff, S., *Acta Paediatr.*, **1993**, *82*, 903-912.
- (10) Kunz, C., Rudloff, S., Baier, W., Klein, N., Strobel, S., *Annu. Rev. Nutr.*, **2000**, *20*, 699-722.
- (11) Ruiz-Palacios, G. M., Cervantes, L. E., Ramos, P., Chavez-Munguia, B., Newburg, D. S., *J. Biol. Chem.*, **2003**, *278*, 14112-14120.
- (12) Bode, L., *Glycobiology*, **2012**, *22*, 1147-1162.
- (13) El-Hawiet, A., Kitova, E. N., Kitov, P. I., Eugenio, L., Ng, K. K. S., Mulvey, G. L., Dingle, T. C., Szpacenko, A., Armstrong, G. D., Klassen, J. S., *Glycobiology*, **2011**, *21*, 1217-1227.
- (14) Kobata, A., *Proc. Jpn. Acad. Ser. B-Phys. Biol. Sci.*, **2010**, *86*, 731-747.
- (15) Martin-Sosa, S., Martin, M. J., Hueso, P., *J. Nutr.*, **2002**, *132*, 3067-3072.
- (16) Bode, L., *J. Nutr.*, **2006**, *136*, 2127-2130.
- (17) Merritt, E. A., Sarfaty, S., Vandenakker, F., Lhoir, C., Martial, J. A., Hol, W. G. J., *Protein Sci.*, **1994**, *3*, 166-175.

- (18) Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., Hol, W. G. J., *Protein Sci.*, **1997**, *6*, 1516-1528.
- (19) Masco, D., Vandewalle, M., Spiegel, S., *J. Neurosci.*, **1991**, *11*, 2443-2452.
- (20) Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K., Ho, W. G. J., *J. Mol. Biol.*, **1998**, *282*, 1043-1059.
- (21) Greco, A., Ho, J. G. S., Lin, S. J., Palcic, M. M., Rupnik, M., Ng, K. K. S., *Nat. Struct. Mol. Biol.*, **2006**, *13*, 460-461.
- (22) Dingle, T., Wee, S., Mulvey, G. L., Greco, A., Kitova, E. N., Sun, J. X., Lin, S. J., Klassen, J. S., Palcic, M. M., Ng, K. K. S., Armstrong, G. D., *Glycobiology*, **2008**, *18*, 698-706.
- (23) Krivan, H. C., Clark, G. F., Smith, D. F., Wilkins, T. D., *Infect. Immun.*, **1986**, *53*, 573-581.
- (24) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.
- (25) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (26) Trott, O., Olson, A. J., *J. Comput. Chem.*, **2010**, *31*, 455-461.
- (27) Zhang, J., Chalmers, M. J., Stayrook, K. R., Burris, L. L., Garcia-Ordonez, R. D., Pascal, B. D., Burris, T. P., Dodge, J. A., Griffin, P. R., *Structure*, **2010**, *18*, 1332-1341.

Chapter 4

- (1) Ghazarian, H., Idoni, B., Oppenheimer, S. B., *Acta Histochem.*, **2011**, *113*, 236-247.
- (2) Dwek, R. A., *Chem. Rev.*, **1996**, *96*, 683-720.
- (3) Roldos, V., Canada, F. J., Jimenez-Barbero, J., *Chem. Bio. Chem.*, **2011**, *12*, 990-1005.

- (4) Avci, F. Y., Kasper, D. L. In *Annual Review of Immunology*, Paul, W. E., Littman, D. R., Yokoyama, W. M., Annual Reviews: Palo Alto, **2010**, Vol. 28, p 107-130.
- (5) Daghestani, H. N., Day, B. W., *Sensors*, **2010**, *10*, 9630-9646.
- (6) Karlsson, R., *J. Mol. Recognit.*, **2004**, *17*, 151-161.
- (7) Homola, J., *Anal. Bioanal. Chem.*, **2003**, *377*, 528-539.
- (8) Bundle, D. R., Eichler, E., Gidney, M. A. J., Meldal, M., Ragauskas, A., Sigurskjold, B. W., Sinnott, B., Watson, D. C., Yaguchi, M., Young, N. M., *Biochemistry*, **1994**, *33*, 5172-5182.
- (9) Bundle, D. R., Sigurskjold, B. W., *Methods Enzymol.*, **1994**, *247*, 288-305.
- (10) Saboury, A. A., *J. Iran Chem. Soc.*, **2006**, *3*, 1-21.
- (11) Larsen, K., Thygesen, M. B., Guillaumie, F., Willats, W. G. T., Jensen, K. J., *Carbohydr. Res.*, **2006**, *341*, 1209-1234.
- (12) Fernandez-Alonso, M. D., Diaz, D., Berbis, M. A., Marcelo, F., Canada, J., Jimenez-Barbero, J., *Curr. Protein Pept. Sci.*, **2012**, *13*, 816-830.
- (13) Lundquist, J. J., Toone, E. J., *Chem. Rev.*, **2002**, *102*, 555-578.
- (14) Wishart, D., *Curr. Pharm. Biotechnol.*, **2005**, *6*, 105-120.
- (15) Hernychova, L., Man, P., Verma, C., Nicholson, J., Sharma, C. A., Ruckova, E., Teo, J. Y., Ball, K., Vojtesek, B., Hupp, T. R., *Proteomics*, **2013**, *13*, 2512-2525.
- (16) Percy, A. J., Rey, M., Burns, K. M., Schriemer, D. C., *Anal. Chim. Acta*, **2012**, *721*, 7-21.
- (17) Pacholarz, K. J., Garlish, R. A., Taylor, R. J., Barran, P. E., *Chem. Soc. Rev.*, **2012**, *41*, 4335-4355.

- (18) Ling, J. M. L., Silva, L., Schriemer, D. C., Schryvers, A. B., *Methods Mol. Biol.*, **2012**, 799, 237-252.
- (19) Iacob, R. E., Engen, J. R., *J. Am. Soc. Mass Spectrom.*, **2012**, 23, 1003-1010.
- (20) Houde, D., Berkowitz, S. A., Engen, J. R., *J. Pharm. Sci.*, **2011**, 100, 2071-2086.
- (21) Jaswal, S. S., *BBA-Proteins Proteomics*, **2013**, 1834, 1188-1201.
- (22) Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., Griffin, P. R., *Expert Rev. Proteomic*, **2011**, 8, 43-59.
- (23) Engen, J. R., *Anal. Chem.*, **2009**, 81, 7870-7875.
- (24) King, D., Lumpkin, M., Bergmann, C., Orlando, R., *Rapid Commun. Mass Spectrom.*, **2002**, 16, 1569-1574.
- (25) Sperry, J. B., Huang, R. Y. C., Zhu, M. M., Rempel, D. L., Gross, M. L., *Int. J. Mass Spectrom.*, **2011**, 302, 85-92.
- (26) Zhu, M. M., Rempel, D. L., Du, Z. H., Gross, M. L., *J. Am. Chem. Soc.*, **2003**, 125, 5252-5253.
- (27) Zhu, M. M., Rempel, D. L., Gross, M. L., *J. Am. Soc. Mass Spectrom.*, **2004**, 15, 388-397.
- (28) Zhu, M. M., Chitta, R., Gross, M. L., *Int. J. Mass Spectrom.*, **2005**, 240, 213-220.
- (29) Sperry, J. B., Shi, X. G., Rempel, D. L., Nishimura, Y., Akashi, S., Gross, M. L., *Biochemistry*, **2008**, 47, 1797-1807.
- (30) Tu, T. T., Dragusanu, M., Petre, B. A., Rempel, D. L., Przybylski, M., Gross, M. L., *J. Am. Soc. Mass Spectrom.*, **2010**, 21, 1660-1667.
- (31) Huang, R. Y. C., Rempel, D. L., Gross, M. L., *Biochemistry*, **2011**, 50, 5426-5435.
- (32) Lin, H., Kitova, E. N., Klassen, J. S., *Anal. Chem.*, **2013**, 85, 8919-8922.

- (33) Kitova, E. N., El-Hawiet, A., Schnier, P. D., Klassen, J. S., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 431-441.
- (34) El-Hawiet, A., Liu, L., Kitova, E. N., Shoemaker, G., Klassen, J. S., *Biophys. J.*, **2012**, *102*, 1A-1A.
- (35) El-Hawiet, A., Kitova, E. N., Klassen, J. S., *Biochemistry*, **2012**, *51*, 4244-4253.
- (36) Ling, H., Boodhoo, A., Hazes, B., Cummings, M. D., Armstrong, G. D., Brunton, J. L., Read, R. J., *Biochemistry*, **1998**, *37*, 1777-1788.
- (37) Johannes, L., Romer, W., *Nat. Rev. Microbiol.*, **2010**, *8*, 105-116.
- (38) Sacchettini, J. C., Baum, L. G., Brewer, C. F., *Biochemistry*, **2001**, *40*, 3009-3015.
- (39) Lin, H., Kitova, E. N., Klassen, J. S., *J. Am. Soc. Mass Spectrom.*, **2014**, *25*, 104-110.
- (40) Kuziemko, G. M., Stroh, M., Stevens, R. C., *Biochemistry*, **1996**, *35*, 6375-6384.
- (41) Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K., Ho, W. G. J., *J. Mol. Biol.*, **1998**, *282*, 1043-1059.
- (42) Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., Hol, W. G. J., *Protein Sci.*, **1997**, *6*, 1516-1528.
- (43) Zhang, Z. Q., Smith, D. L., *Protein Sci.*, **1993**, *2*, 522-531.
- (44) Keppel, T. R., Howard, B. A., Weis, D. D., *Biochemistry*, **2011**, *50*, 8722-8732.
- (45) Merritt, E. A., Sarfaty, S., Vandenakker, F., Lhoir, C., Martial, J. A., Hol, W. G. J., *Protein Sci.*, **1994**, *3*, 166-175.
- (46) Wei, H., Ahn, J., Yu, Y. Q., Tymiak, A., Engen, J. R., Chen, G. D., *J. Am. Soc. Mass Spectrom.*, **2012**, *23*, 498-504.

Chapter 5

- (1) Zhang, Y. X., Liu, L., Daneshfar, R., Kitova, E. N., Li, C. S., Jia, F., Cairo, C. W., Klassen, J. S., *Anal. Chem.*, **2012**, *84*, 7618-7621.
- (2) Bayburt, T. H., Sligar, S. G., *FEBS Lett.*, **2010**, *584*, 1721-1727.
- (3) Shaw, A. W., McLean, M. A., Sligar, S. G., *FEBS Lett.*, **2004**, *556*, 260-264.
- (4) Boldog, T., Li, M. S., Hazelbauer, G. L. In *Two-Component Signaling Systems*, Simon, M. I., Crane, B. R., Crane, A., Elsevier Academic Press Inc: San Diego, **2007**, Vol. 423, p 317-335.
- (5) Hebling, C. M., Morgan, C. R., Stafford, D. W., Jorgenson, J. W., Rand, K. D., Engen, J. R., *Anal. Chem.*, **2010**, *82*, 5415-5419.
- (6) Nasr, M. L., Shi, X. M., Bowman, A. L., Johnson, M., Zvonok, N., Janero, D. R., Vemuri, V. K., Wales, T. E., Engen, J. R., Makriyannis, A., *Protein Sci.*, **2013**, *22*, 774-787.
- (7) Parker, C. H., Morgan, C. R., Rand, K. D., Engen, J. R., Jorgenson, J. W., Stafford, D. W., *Biochemistry*, **2014**, *53*, 1511-1520.
- (8) Popovic, K., Holyoake, J., Pomes, R., Prive, G. G., *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, *109*, 2908-2912.
- (9) Sowole, M. A., Konermann, L., *Anal. Chem.*, **2014**, *86*, 8-15.