Diss. ETH NO. 21806

Applying MALDI-ToF Mass Spectrometry to Study Non-covalent Interactions

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zürich)

presented by
FAN CHEN

Mphil, The Chinese University of Hong Kong, Hong Kong

born on 15.08.1986
citizen of
China

accepted on the recommendation of

Prof. Dr. Renato Zenobi, examiner
Prof. Dr. Kaspar Locher, co-examiner

2014

I. DISCIAIMER

I hereby declare that this dissertation represents the result of my own original work during my PhD studies in ETH, and it has not been submitted to any other institute/university for another degree or qualification. I also declare that no material written by another person without proper citation in the text.

Fan Chen **Date**

II. Publications

F.Chen, S. Gerber, K. Heuser, V. M. Korkhov, C. Lizak, S. Mireku, K. P. Locher, and R. Zenobi, High-mass MALDI Mass Spectrometry of Integral Membrane Proteins and their Complexes, Anal. Chem., 85, 2013, 3483-3488. (Editors' Highlight)

F. Chen, B. Gülbakan, R. Zenobi, Direct access to aptamer–protein complexes via MALDI-MS, Chem. Sci., 4, 2013, 4071-4078.

F. Chen, S. Nielsen, R. Zenobi, Understanding Chemical Reactivity for Homo- and Heterobifunctional Protein Cross-linking Agents, J. Mass. Spectrom., 48, 2013, 807-812.

F.Chen, S. Mädler, S. Weidmann and R. Zenobi, Detection of Non-covalent Interactions of Single Stranded DNA with Escherichia Coli Single-Stranded DNA-Binding Protein by MALDI-MS, J. Mass. Spectrom., 47, 2012, 560-566.

C. Lizak, S. Gerber, D. Zinne, G. Michaud, M. Schubert, **F. Chen**, M. Bucher, T. Darbre, R. Zenobi, J. Reymond, K. P. Locher, A Catalytically Essential Motif in the External Loop 5 of the Bacterial Oligosaccharyltransferase PglB, J. Biol. Chem. doi:10.1074/jbc.M113.524751

F. Chen, S. Gerber, V. M. Korkhov, S. Mireku, M. Bucher, K. Heuser, K. P. Locher, and R. Zenobi, Can NHS Esters Stabilize Integral Membrane Protein Complexes in Detergent Micelles? Submitted to Anal. Chem.

F. Chen, B. Gülbakan, S. Weidmann, A. J. Ibanez, S. R. Fagerer, R. Zenobi, Applying MALDI-ToF Mass Spectrometry to Study Non-covalent Biomolecular Complexes, Submitted to Mass Spectrom. Rev.

N. Ding, M. Fu, H. Xing, X. Chen, **F. Chen** and C. M. L. Wu, Ultrathin-shell BN Hollow Spheres as Sorbent for Dispersive Solid-phase Extraction, Submitted to Chem. Comm.

L. Tang, **F. Chen**, L. M. Yang and Q. Q. Wang, The Determination of Low-molecular-mass Thiols with 4-(Hydroxymercuric)benzoic Acid as A Tag using

HPLC Coupled Online with UV/HCOOH-Induced Cold Vapor Generation AFS, J. Chromatogr. B, 877, 2009, 3428-3433.

III. Conference Contributions

F. Chen, B. Gülbakan and R. Zenobi, EMBO Practical Course: Modern Biophysical Methods for Protein-Ligand Interactions, October 21-25, 2013, Oulu, Finland. (Poster Presentation)

F.Chen, B. Gülbakan and R. Zenobi, Fall Meeting of the Swiss Chemical Society 2013, September 6, 2012, EPFL, Switzerland. (Poster Presentation)

F.Chen, B. Gülbakan and R. Zenobi, 31th Informal Meeting on Mass Spectrometry, May 5 - 8, 2013, Palermo, Italy. (Oral Presentation)

B. Gülbakan, **F.Chen**, K. Barylyuk, and R. Zenobi, 61th Annual Conference on Mass Spectrometry and Allied Topics by the American Society of Mass Spectrometry, June 9-13, 2013, Minneapolis, MN, USA. (Poster Presentation)

F.Chen, and R. Zenobi, 2013 Gordon Research Conference on Gaseous Ions: Structures, Energetics & Reactions, Feb 24 – March 1, 2013, Galveston, TX, USA. (Poster Presentation)

F.Chen, S. Gerber, K. Heuser, V. M. Korkhov, C. Lizak, S. Mireku, K. P. Locher, and R. Zenobi, November 1-2, 2012, Swiss Group for Mass Spectrometry Meeting, Beatenberg, Swizterland. (Oral Presentation)

F.Chen, S. Gerber, K. Heuser, V. M. Korkhov, C. Lizak, K. P. Locher, and R. Zenobi, 19th International Mass Spectrometry Conference, September 15-21, 2012, Kyoto, Japan. (Award Oral Presentation)

F.Chen, K. P. Locher, and R. Zenobi, Fall Meeting of the Swiss Chemical Society 2012, September 13, 2012, ETH Zürich, Switzerland. (Oral Presentation)

F.Chen, K. P. Locher, and R. Zenobi, 30th Informal Meeting on Mass Spectrometry, April 29 – May 3, 2012, Palacky University, Olomouc, Czech Republic. (Oral Presentation)

F. Chen, S. Mädler, S. Weidmann and R. Zenobi, The 24th ASMS Sanibel Conference on Mass Spectrometry, January 19-22, 2012, St. Pete Beach, Florida, USA. (Poster Presentation)

F. Chen, S. Mädler, S. Weidmann and R. Zenobi, EMBO Practical Course: Protein-Protein and Protein-Nucleic Acid Cross-linking and Mass Spectrometry, October 23-29, 2011, Göttingen, Germany. (Poster Presentation)

F. Chen, S. Mädler, and R. Zenobi, Fall Meeting of the Swiss Chemical Society 2011, September 9, 2011, EPFL, Switzerland. (Poster Presentation)

IV. Acknowledgements

This thesis is the end of my journey in obtaining my Ph.D. This thesis has been through to completion with the support and encouragement from numerous people including my family, my friends, colleagues and my supervisors. At the beginning of my thesis, I would like to thank all those people who made this thesis possible and an unforgettable experience for me. It is a pleasant task to express my gratitude to all those who contributed in many ways to the success of this study.

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Renato Zenobi, for the continuous support of my Ph.D. study. This work would not have been possible without his guidance, support and encouragement. I truly appreciate the freedom to pursue my ideas in different fields and his support to present my work in numerous conferences and journals.

I am also extremely indebted to my co-examiner Prof. Dr. Kasper P. Locher for his motivation, enthusiasm and immense knowledge. I am very much thankful for allowing me as an external member in his lab during my Ph.D. His valuable advice, constructive criticism and his extensive discussion are significant to this thesis. I could not have imagined having a better co-examiner for my Ph.D study.

Words are inadequate in offering my tribute to Ms. Brigit Bräm, the most beautiful and kind lady I have ever met in Switzerland. Her understanding, encouragement, helps and smiles have provided wonderful basis for my Ph.D. research. My research could not go well without her organization, support and kindness.

I take this opportunity to sincerely acknowledge the Locher Lab. My gratitude goes to Dr. Sabina Gerber (PglK, PglB and "naughty" TonB), Dr. Jungin Kim (ABCG2), Dr. Christian Lizak (PglB and his knowledge on glycosylation), Dr. Jae-Sung Woo (ABCA1), Dr. Katrin Heuser (Cdr1p), Dr. Martina Niederer (MATE, our first MALDI spectrum on membrane proteins), Dr. Vladimir Korkhov (BtuCD,

BtuCDF and BtuCDX), Dr. Amer Alam (PgP), Dr. Camilo Perez (PglK), Dr. Ana Ramirez, Samantha Mireku (BtuCD and BtuCDFX-n) and Monika Bucher, for their sample preparations, biological knowledge sharing and extensive discussions. Special thanks go to Sabina and Jungin for their careful reading and comments for TonB and ABCG2 chapters.

I want to thank Prof. Dr. Andreas Plueckthun for his valuable advice and assistance to perform my work within his lab. Special thanks go to Dr. Alexander Batyuk, Matthias Hillenbrand and Marco Schütz for the interesting collaborations and discussions. My gratitude goes also to the Riek group, Dr. Silvia Campioni and Nadezda Nespovitaya, for the challenging studies together. Moreover, I am grateful to Dr. Xiaodan Li and Dr. Ching-ju Tsai for providing membrane protein samples and their knowledge, especially on lipids.

I was extremely lucky to work in our awesome high-mass MALDI team with all you guys. I would like to first thank Dr. Stefanie Mädler for introducing me to the practical aspects on chemical cross-linking and also MALDI-MS. I am deeply indebted to Dr. Simon Weidmann for his organizing, help and understanding. Special thanks to external high-mass MALDI team member, Dr. Basri Gülbakan, for guiding me into a new field: aptamer. I will for sure miss you guys no matter what I will pursue in my future. I would also like to thanks Dr. Claudia Bich for sharing her valuable knowledge on chemical cross-linking of protein and DNA.

It will never be easy to use one instrument on two extremely different projects. Thanks to our super lab chief, Dr. Stephan Fagerer, I am able to carry out most of my experiments smoothly. And my research also won't go that far without the understanding and help from the whole MALDI team: Dr. Alfredo Ibanez, Dr. Martin Pabst, Dr. Stephan Fagerer, Dr. Simon Weidmann, Jasmin Krismer and Robert Steinhoff. Special thanks go to Dr. Alfredo Ibanez, the brilliant scientist, thanks for your special skills to get free and useful instruments for our lab; Dr. Simon Weidmann, thank you for your numerous effort to solve the database problem.

I was extremely lucky to share the office with two senior scientists, Dr. Pablo Martinez-Lozano Sinues (also famous as DAGE after Christmas Eve, 2012) and

Dr. Vladimir Frankevich, and of course our external guest Dr. Konstantin Barylyuk (Dr. B). All the numerous and extremely fruitful discussions about science, career and the whole world made our office full of laughs. Special thanks go to Pablo, who always told me that "I am your Dage," for your help, understanding and support.

We do not only do science in Zenobi group. Thanks Guys. You made my stay much more colorful and enjoyable. I want to thank all the past and present members, especially Rui (for the best Chinese food and our amazing movie nights), Amy (to be my "best" drinking partner), Fanny (to be the best Zumba and fun partner), Jingjing (for her unique explanation of all the sophisticated phenomena), Carolin (for your endless help, understanding and my favorite dessert), Simone (for the best blueberry chocolate cake), Stefannie (for her encouragement through my whole Ph.D. study), Melissa (for the best beer-"alcohol free"), Jasmin, Agni, Liang (for introducing me to the group and the continuous friendship), Pablo (for being our amazing Madrid Christmas guide), Alfredo (for guiding me through the Ph.D. life; I do admit that I like your South-American style a lot), Simon (for organizing everything, from lab to conference trips), Kostya (for his broad knowledge of everything and great laughs), Stefan (for his "complain", extra spicy, fun and being a great steps counter), Stephan (for organizing all our Saturday drinks), Lothar (for his creative, curiosity and unbeatable food), Johannes (I won't bet against you anymore), Morphy (for being my short-term mentor and long-term friend), Jianfeng (for introducing your wife to us and your special "pig-feet"), Basri (for all his Turkish delights), Rob (for introducing the best bars in Zürich), Volker (for your best protection), Bruno (for being my short-term Greek teacher and maybe French as well), Robert (for your awesome pan cover), Thomas, Christian (for your unique explanation of basic Swiss drinking rules) and Jan (for taking care of my drinking partner). Thank you for the great time we shared at work, on holidays and during conferences. Big thanks for all the help from you guys and crazy events (party nights) together.

Besides the formal Zenobi members, my thanks also go in particular to the external Zenobi family members: Zaira (for your encouragement), Madina (for your afternoon tea time), Adrian (for being the best train driver), Eliane (for

your cuteness), Edu (for your clear fan), David (for your super surprise) and Twins-Paul & Augustus (for your definite “no”) and our party members: Twins-Bartosz & Tomasz (for all the crazy events and laughs) and Philip (for the party nights).

A million thanks to all the “Chinese Club” members: Liang, Rui, Hui, Amy, Jingjing and Jianfeng, and our VIP members: Fanny, Morphy and Pablo. Thanks for bringing “Chinese Culture” to Zürich. Your support and friendship is spread over this thesis; and without all you guys, life in Zürich would have been only half as great. Special thanks to Hui, Rui, Amy, Fanny and Jingjing for all the great trips we had together.

It is also a pleasure to mention all the friends I met in two EMBO practical courses. Special thanks go to Romina (for her organization and our margarita), Zloty (for supporting all my crazy ideas), Simone, Ilala (I promised that I will fill pig into the form), Mirka (for being my short-term roommate), Alexander (for the best photographer), Chris (for our bloody stories), Dario (to be an awesome friend in candy-crush) and Edward (for answering all my questions and forgiving my curiosity). Thank you for all the great time we had together during and after the courses. I would like to thank my three long-term roomies in Zürich: Rui, Emilie and Gizem. Thanks for all the nice dinners, wine and our warm atmosphere at home. Thank you, Falk. I won’t forget all the great time we had together. I would also like to take this opportunity to sincerely acknowledge Annegret, Sameul, Lukas, Max, and our case owner, Marcello, for the great experience during the leadership workshop.

Finally, and most importantly, I would like to pay high regards to my beloved family for their sincere encouragement and inspiration throughout my research work. Without their support and understanding, I would not have managed to come that far. Thank you so much for all your love, trust and support.

Besides all I earnestly thank each and every one from the bottom of my heart and I also pay my sincere apology to those if I have not offered my thankfulness.

V.Abstract

Non-covalent interactions are ubiquitous for the structural organization of biomacromolecules and play an important role in molecular recognition processes, such as the interactions between proteins, glycans, lipids, DNA and RNA. With the advent of soft ionization, namely electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), mass spectrometry (MS) has become an indispensable tool to investigate non-covalent interactions. Nowadays, native-ESI MS is heavily used in studies of non-covalent interactions and to understand the architecture of biomolecular complexes. Compared with native ESI-MS, MALDI-MS has contributed significantly less to the studies of biomolecular complexes. The primary reason for this is that MALDI causes dissociation of non-covalent interactions, either during sample preparation or the ionization process, and can sometimes also induce formation of nonspecific aggregates.

However, due to its high salt/detergent tolerance and the simplicity of data interpretation, MALDI-MS is also becoming increasingly useful in this realm. Our studies of membrane proteins highlight the potential of MALDI-MS. Currently, there are two major approaches to carry out studies on non-covalent complexes with MALDI-MS. In the first approach, different experimental and instrumental parameters are fine-tuned, such as the use of non-acidic matrices, or collecting first-shot spectra. With this approach, we were able to study non-covalent interactions in the complexes formed between single stranded DNA and a single stranded DNA binding protein, as well as those in several protein-aptamer complexes. Alternatively, interacting species can be stabilized by chemical crosslinking. We used this approach to study the stoichiometries of several membrane protein complexes, using both glutaraldehyde and chemically more specific cross-linkers, such as NHS esters. Both approaches have advantages to study non-covalently bound biomolecules. In summary, this work gives detailed insight into applications of MALDI-MS for investigating non-covalent interactions.

VI. Zusammenfassung

Nichtkovalente Wechselwirkungen sind allgegenwärtig in der strukturellen Organisation von Biomakromolekülen und spielen eine wichtige Rolle für molekulare Erkennungsprozesse, wie z.B. Wechselwirkungen zwischen Proteinen, Polysacchariden, Lipiden, sowie DNA- und RNA-Molekülen. Durch die Entwicklung weicher Ionisationsmethoden wie der Elektrospray-Ionisation (ESI) und der matrix-unterstützten Laser-Desorption/Ionisation (MALDI), konnte sich die Massenspektrometrie (MS) als eine unverzichtbare Methode für die Untersuchung nichtkovalenter Wechselwirkungen etablieren. Heutzutage wird insbesondere sogenannte „native“ ESI-MS eingesetzt, um den Aufbau von biomolekularen Komplexen zu verstehen. Im Vergleich zu nativer ESI-MS wurden deutlich weniger Studien mit MALDI-MS durchgeführt. Der Hauptgrund dafür liegt darin, dass der MALDI-Prozess - entweder bedingt durch die Probenpräparation oder den eigentlichen Ionisationsvorgang - die Dissoziation nichtkovalenter Wechselwirkungen herbeiführen kann. Darüber hinaus werden oft nichtspezifische Aggregate gebildet. Dennoch nimmt die Bedeutung von MALDI-MS auf diesem Gebiet aufgrund der hohen Toleranz gegenüber Salzen/Detergenzien sowie der einfachen Dateninterpretation zu. Unsere Untersuchungen von Membranproteinen veranschaulichen das Potential von MALDI-MS. Aktuell werden hauptsächlich die folgenden zwei Ansätze für die Analyse von nichtkovalenten Komplexen mit MALDI-MS verwendet: Einerseits können experimentelle und instrumentelle Parameter optimiert, nicht-azide Matrices verwendet oder so genannte „first-shot“-Spektren akkumuliert werden. Mit diesen Ansätzen wurden in dieser Arbeit Einzelstrang-DNA–Protein-Komplexe, sowie verschiedene Protein–Aptamer-Komplexe nachgewiesen. Andererseits können die Wechselwirkungen zwischen interagierenden Biomolekülen durch chemische Quervernetzung stabilisiert werden. In dieser Arbeit wurde diese Methode für Untersuchungen zur Stöchiometrie von mehreren Membranproteinkomplexen mit Hilfe von Glutaraldehyd und spezifischeren Quervernetzern, wie z.B. NHS-Estern, verwendet. Zusammenfassend gibt diese Dissertation detaillierte Einblicke in Anwendungen von MALDI-MS für die Untersuchung von nichtkovalenten Wechselwirkungen.

VII. Table of Contents

Chapter 1. Introduction and Background

This chapter provides a general overview of non-covalent interactions that are found in biomolecules, and current approaches to study these interactions. Thereafter, a discussion on the soft ionization mass spectrometry methods applied, especially MALDI-MS, is given. Following that, the motivation of the work described in this thesis is given.

a. Non-covalent Interactions In Biomolecules

After two centuries of intense study, the concept of covalent bonding is quite developed. The process of forming and breaking of covalent bonds is understood both on the experimental and theoretical levels. In contrast, the nature of non-covalent interactions is still elusive in comparison and unfortunately, calculations are frequently in conflict with the relevant experimental data. (Hobza & Müller-Dethlefs, 2010) The main reason behind this disparity is that, both in solution and in the solid phase, non-covalently bound molecules are surrounded by other molecules. Thus, these surrounding molecules significantly perturb the electronic system of the molecule and affect the non-covalent bonds. This perturbation depends on the strength and extent of the non-covalent interactions, e.g., ionic and H-bonding interactions.

Non-covalent interactions are responsible for the tertiary and quaternary structure of large biomolecules (also denominated biopolymers or biomacromolecules) such as DNA, RNA, proteins, carbohydrates, etc. The first three are major biomolecules essential for all known forms of life. Hydrogen bonds and vertical π-π interactions are two major examples of non-covalent interactions that define the form of those biomolecules, yet their relative strength is still unclear. Non-covalent interactions also play a key role in molecular recognition processes. For instance, they trigger for many cellular processes, including cell division, cell signaling, gene transcription and translation. In nature, these interactions come in different flavors, including protein-protein interactions, protein-small ligand (or peptide, metal, and carbohydrate) interactions, protein-DNA interactions, DNA-DNA interactions and DNA/RNA-small ligand interactions. A brief discussion of two types of interactions, protein-protein and protein-DNA interactions, is given below to highlight their importance.

Protein-protein interactions (PPIs) are ubiquitous in biology and are defined by the composition of the primary sequence (identical or non-identical chains), affinity and by the fact whether the association is permanent or transient. (Ngounou Wetie et al., 2013; Nooren, 2003) Based on the composition, PPIs can

occur between homo- or hetero-oligomers. Two different types of complexes can be distinguished on the basis of whether a complex is "obligate" or "non-obligate". In an obligate PPI, the protomers are not found as stable structures on their own in vivo. In many cases, such complexes are also functionally "obligate". In contrast, protomers in non-obligate interactions exist independently. Examples of such systems are antibody-antigen, receptor-ligand, and enzyme-inhibitor complexes. The components of such protein-protein complexes are often initially not co-localized and thus their components need to be independently stable. Moreover, PPIs can be differentiated by the lifetime of the complex. A permanent interaction is usually very stable; however, a transient interaction is a dynamic equilibrium and can easily associate or dissociate in vivo. A good example is the hetrotrimeric G protein complex, which dissociates into Gα and G$\beta\ \gamma$ subunits upon guanosine triphosphate (GTP) binding, but forms a stable trimer with guanosine diphosphate (GDP) bound.

Besides their important roles in PPIs, non-covalent interactions are at the center of protein-nucleic acid interactions. The nonspecific binding of proteins to DNA is largely governed by electrostatic forces and is driven by the displacement of condensed counterions. (Hippel, 2007) For example, single-stranded DNA binding proteins (SSBPs) form a central part of the phage T4 DNA replication complex and generally bind to ssDNA lattices with little, if any, base sequence specificity. (Chase & Williams, 1986) The reason behind the lack of specificity is the strong attractive force of electrostatic interactions and its involvement in ion condensation. (Manning, 1969; Oosawa, 1971) In the previous example, the positively charged protein side chains located in the binding site of the protein interact with the negative phosphate groups of the ssDNA backbone. However, specific proteins do also recognize and bind to DNA based on the structure. These protein-DNA interactions are strong and mediated by hydrogen bonding, ionic interactions (protein side chain-DNA backbone interactions), van der Waals, and hydrophobic interactions. The most intensely studied proteins of this kind are transcription factors, (Pabo & Sauer, 1992) which specifically bind to DNA sequences.

In summary, non-covalent interactions play a crucial role in most cellular processes. For this reason, the study of these biophysical interactions among biomacromolecules draws great attention. In the next section, we will list and discuss the most widely used analytical methods for studying non-covalent interactions among biomolecules.

b. Biophysical Methods to Investigate Non-covalent Interactions among Biomolecules

Currently, several established instrumental and theoretical methods are applied for studying of non-covalent interactions in macromolecule complexes. (Phizicky & Fields, 1995) Figure 1 highlights different aspects of a molecular assembly that can be investigated by the corresponding method. In addition these methods, analytical ultracentrifugation, isothermal titration calorimetry, various spectroscopic approaches (e.g., circular dichroism, light scattering, and fluorescence), differential scanning calorimetry, gel filtration chromatography, far western blot, surface plasma resonance, and affinity capillary electrophoresis are also used to study non-covalent interactions. Given the plethora of established techniques for measuring protein interactions (Fu, 2004; Schuck, 2007), we would like to survey and categorize generic and frequently applied techniques for identification and characterization of protein interactions here.

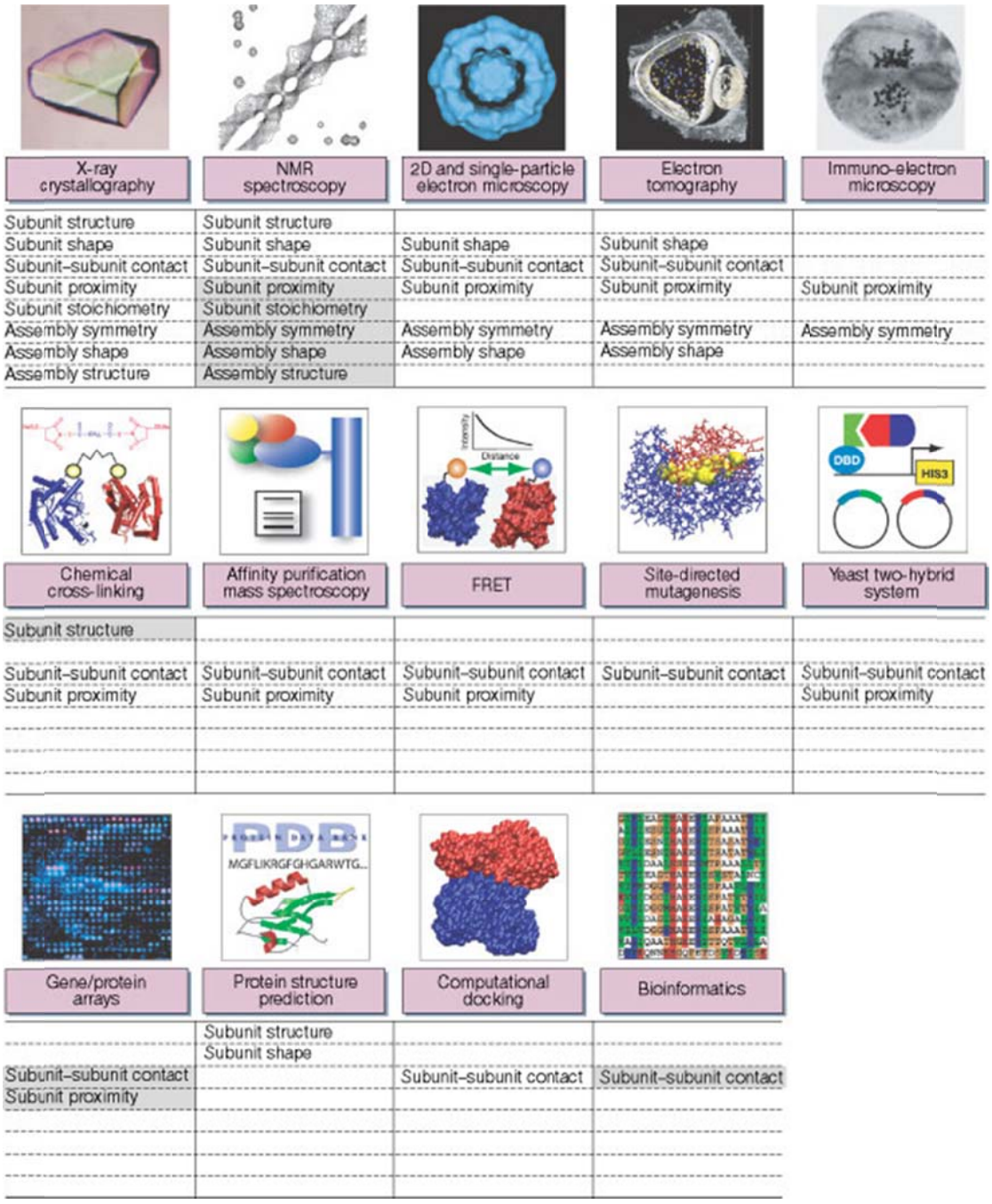


Figure 1. Illustration of experimental and theoretical methods that provide information on a macromolecular complex assembly structure. The aspects of an assembly that might be obtained by the corresponding method are listed in the annotations. Extreme difficulties in obtaining the corresponding information by a given method are indicated in grey boxes. Reproduced with permission from (Sali et al., 2003). Copyright © 2003 Nature Publishing Group.

As seen in Figure 1, x-ray crystallography is the major tool for determining static protein complexes at atomic resolution. Nuclear magnetic resonance (NMR)

generates unparalleled high-resolution structures of dynamic biomolecule complexes in solution. Both of these techniques require relatively large quantities of material (milligram scale) and a long processing time. Material to be studied also needs to be of high quality and must be extremely homogeneous, which can be difficult. Moreover, proteins might precipitate at the high concentrations necessary for NMR. Also, the molecular weight range that can be analyzed with NMR is limited, although larger magnetic fields allow the analysis of higher molecular weight complexes. For an analysis with X-ray crystallography, the protein must be present as a single crystal. Besides these two most widely accepted high-resolution methods, electron crystallography, electron tomography and immune-electron microscopy can also be used to understand subunit proximity, assembly symmetry and the assembly shape at intermediate resolution.

The other six methods presented in Figure 1, including chemical cross-linking, affinity purification, fluorescence resonance energy transfer (FRET), site directed mutagenesis, the yeast two-hybrid system, and protein arrays, are mostly used to determine subunit-subunit contacts. For instance, FRET can be applied to monitor protein interactions if one protein is fused to a fluorescence donor and its potential partner carries a fluorescence acceptor. The binding constant is the key aspect concerning subunit-subunit contacts. Besides FRET, isothermal titration calorimetry, circular dichroism, differential scanning calorimetry and surface plasma resonance are well-known methods to measure noncovalent binding constants. A very good comparison among these methods, including the accessible K_d range, time requirement, sample amount and limitation, has been presented in the literature. (Table 1) (Boeri Erba & Zenobi, 2011) This comparison is made between native mass spectrometry, mainly electrospray mass spectrometry, and other solution phase methods, and considered the accessible K_d range and required sample amount to determine binding constants.

Table 1. Comparison of various analytical methods to measure noncovalent binding constants. Adapted from (Boeri Erba & Zenobi, 2011).

Method	Accessible K_d range	Required Time	Required amount of sample (μ g)
		per single run	
Fluorescence-based methods (FRET, FCS, fluorescence polarization)	μ M - 10 pM	1 h	
CD	μ M - 10 pM	2-3 h	10-4000
ITC	10 μ M - 10 nM	2-3 h	500-5000
SPR	10 μ M - 10 pM	1h	50-400
Differential Scanning Calorimetry	nM -10 zM	1-2 h	50-500
ESI titration	mM - 0.1 μ M	1-2 h	10-15
Competition method	mM - fM	2h	10-15

Non-experimental methods used to provide information on macromolecular assembly structures are also listed in Figure 1. These include protein structure prediction, computational docking and a variety of bioinformatics techniques. Such approaches are mostly used when the protein structure cannot be obtained experimentally. Although protein structure prediction and computational docking are generally not sufficient to accurately predict whether two proteins actually interact with each other, they can, to some extent, correctly identify the interacting surfaces between two subunits. (Strynadka et al., 1996) Bioinformatic analysis of genomic sequences can indicate the presence and location of protein interaction interfaces.

c. Investigating Non-covalent Complexes via Mass Spectrometry

With the advent of soft ionization mass spectrometry tools namely electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), mass

spectrometry (MS) has become an indispensable tool to investigate non-covalent interactions. Compared with other approaches, MS requires only small amounts of sample, on the order of fmol, and can be readily carried out without extensive purification or labeling. (Table 1) By preserving the non-covalent interactions during the transition from solution to the gas phase, the mass spectra directly provide stoichiometric information about the composition of the complex. A number of reviews from different groups focus on the application and strategies to use native ESI and MALDI-MS in studies of non-covalent interactions. (Hernández & Robinson, 2007; Barrera et al., 2009; Barrera & Robinson, 2011; Marcoux & Robinson, 2013; Heck, 2008; Uetrecht et al., 2010b; Bich & Zenobi, 2009; Mädler et al., 2013). In this section, we will discuss current challenges of non-covalent interaction studies carried out with native ESI-MS and MALDI-MS..

c.1. Studying Non-covalent Complexes with native ESI-MS

McLafferty (McLafferty, 1981) has pointed out the three "S" advantages of mass spectrometry: specificity, sensitivity and speed; Loo (Loo, 2000) included another "S" advantage, stoichiometry, in the context of non-covalent interaction studies. Fairly soon after its introduction ESI-MS has demonstrated its utility for the detection and study of non-covalent complexes. ESI is achieved by applying a potential difference between the inlet of the mass spectrometer and a conductive capillary containing the analyte solution as shown in Figure 2. Thus, analytes are ionized by a combined process of desolvation and (de)protonation. Successive rounds of solvent evaporation and droplet fission occur until an analyte ion is formed. To account for the ion formation, two models have been proposed: ion evaporation model (Iribarne, 1976) and the charged residue model (Dole, 1968). In the ion evaporation model, analyte ions are expelled directly from droplets; whereas, in the charged residue model, analyte ions are finally formed after droplet fission and solvent evaporation processes. Current evidence suggests that folded protein ions are generated according to the charge residue model (Fernandez de la Mora, 2000). Recently, unified theories are also proposed (Hogan et al., 2009; Konermann et al., 2013).

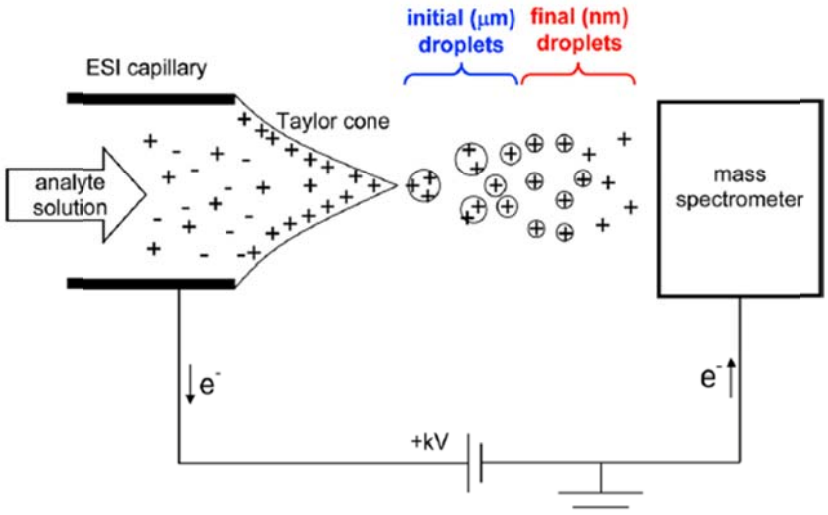


Figure 2. Scheme of an ESI source operated in positive ion mode. Reproduced with permission from (Konermann et al. 2013). Copyright © 2013 American Chemistry Society.

c.1.1. Native ESI-MS

In a typical ESI experiment, acidified organic solvent is usually employed to promote protonation of the analytes and to increase the sensitivity. Substantial sample volumes and high temperature are also required for the experiment. These conditions are generally not friendly to preserve of biomolecular complexes in solution. To overcome these difficulties, nanospray, which uses a smaller capillary diameter and lower flow rates, has been introduced. (Wilm & Mann, 1994 & 1996; Emmett & Caprioli, 1994; Gale & Smith, 1993; Jiang, 2006) The reduced flow rate has the added benefit of producing smaller initial droplet sizes, which further increases sensitivity and salt tolerance. It is also crucial to abolish organic solvents and high interface temperature. In native ESI-MS, aqueous solutions of non-covalent complexes are typically prepared at concentrations of 1-10 μ M in a volatile neutral buffer. The low concentration eliminates non-specific association during nESI. The buffer ensures a neutral pH to keep non-covalent complexes at near physiological conditions, and also to against electrochemical effects during the ionization process. A commonly used volatile buffer is ammonium acetate. (Heck, 2008) For native MS, ammonium acetate is applied at a concentration range from 5 mM to 1 M at neutral pH. To preserve the biomolecular complexes, mild instrument conditions are also required, i.e., low collision voltages and ion guide pressures. In conjunction with

nanoelectrospray, native MS become important for studying non-covalently bound biomolecules. Figure 3 illustrates the effect of native and non-native solution conditions and instrumental parameters, on detecting a protein complex of a small Heat shock protein (sHsp) *Ta*SHP16.9, which is an oligomeric species with six dimeric building blocks that forms a 12mer. (Hilton & Benesch, 2012) The HSP complex was well preserved under "near native conditions" (Figure 3 a); introducing organic solvent (Figure 3 b&c) or activating the complex (Figure 3 d) destroyed the complex.

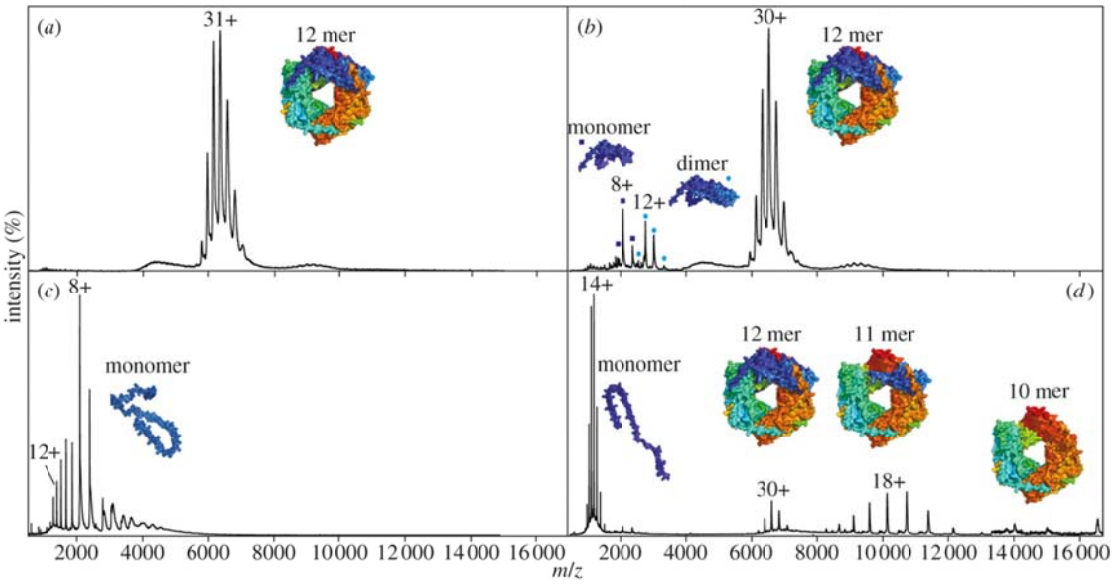


Figure 3. Mass spectra of the heat shock protein complex (HSP) subjected to different experimental conditions. a) Mass spectrum of HSP under "near native conditions"; b) example spectrum showing the effects of solution phase by introducing isopropanol (10%, v/v); c) denaturing conditions (acetonitrile: water: formic acid, 50:49.9:0.1, v:v:v); d) mass spectrum under activating conditions (high collision voltages and ion guide pressures). Reproduced with permission from (Hilton & Benesch, 2012) Copyright © 2012 The Royal Society.

The introduction of native ESI-MS to study biomolecular complexes brought new challenges to MS technology. The most important property for a mass analyzer in native ESI-MS is that it must be capable of handing ions with high mass-to-charge ratios (m/z). In early studies, triple-quadrupole mass spectrometers were used to study non-covalent complexes. They offer the advantage of allowing tandem-MS experiments but are typically limited to a maximum m/z of approximately 4000 Th. By operating the quadrupole at a lower frequency, it is

now possible to analyze 100 kDa proteins (Light-Wahl et al., 1994). However, the low resolving power at high m/z is a considerable disadvantage. In contrast, time-of-flight (ToF) mass analyzers can reach a high mass range, good mass resolution and sensitivity on a very fast timescale. The "hybrid" Q-ToF design has been recognized as the optimum instrument geometry for accessing the high m/z range, with the benefits of both the ToF for mass analysis and the selection capabilities of a quadrupole (Q). (Sobott et al., 2002) Commercial ESI-ToF and Q-ToF instruments are available from several companies (e.g., Waters, SCIEX, Bruker, Agilent) and special modifications are also available for native ESI-MS applications. Albeit being lower after these modifications, a mass resolution of several thousand at m/z of 10,000 can still be achieved.

Additionally, ion mobility (IM), a gas-phase separation technique, has been coupled to mass spectrometry as a new tool to analyze intact protein complexes. (Kaddis & Loo, 2007; Uetrecht et al., 2010a; Uetrecht et al., 2010b) It enables the direct determination of molecular size in terms of a rotationally averaged collisional cross section (CCS) in the gas phase. Some studies suggest that the experimental CCS of a protein is similar to the size estimated from atomic constraints (Scarff et al., 2008), which has also been observed in fragile protein complexes (Ruotolo et al., 2005).

c.1.2. Application of Native ESI-MS to Study Biomolecular Complexes

A series of exciting applications has been reported in the analysis of intact biomacromolecular complexes, including whole virus particles (Uetrecht et al., 2008; Uetrecht et al, 2010a; Uetrecht et al, 2010b), proteasomes (Loo et al., 2005; Sharon et al., 2006) and ribosomes (Gordiyenko et al., 2010; Hernández & Robinson, 2007; Rostom et al., 2000; Videler et al., 2005). A few examples are presented below. Loo et al. used native MS to characterize the structure of the 20S proteasome, a barrel-shaped catalytic core, from *Methanosarcina thermophile* and rabbit; they confirmed the stoichiometries of the 192 kDa subunit as an α7 ring and of the intact 690 kDa assembly as α7β7β7α7, respectively. (Loo et al., 2005) Heck and coworkers applied native ESI to study yeast RNA Pol II and III and refined the current structural model of RNA Pol III

by combining the known high-resolution structures of RNA Pol II and native MS data. (Mohammed et al., 2008) The studies of in vitro-assembled HBV capsid particles show a well-resolved charge state distributions of the T = 3 and T = 4 icosahedron symmetry particles with molecular weights of approximately 3 and 4 MDa. (Uetrecht et al., 2008)

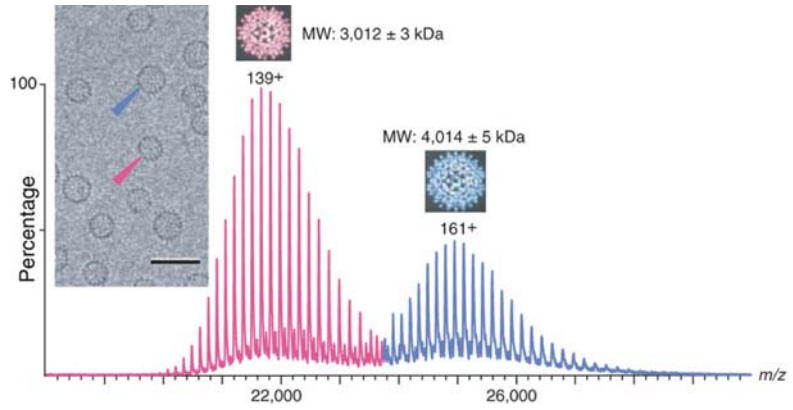


Figure 4. Native mass spectrum of intact HBV viral capsids reconstituted in vitro from truncated cp149 capsid monomers in 200 mM ammounium acetate, at pH 6.8. Capsid particles visualized by cryo-electron microscopy are shown in the inset on the left. Reproduced with permission from (Heck, 2008). Copyright © 2008 Nature Publishing Group.

Membrane protein complexes are among the most challenging protein systems for structural biology owing to their strict solubility requirements. Rather than removal of detergent prior to sample analysis, which is usually done by organic solvent extraction or reversed-phase or size-exclusion chromatography, Griffiths et al. first showed the possibility to detect an intact non-covalent homotrimer of rat microsomal glutathione transferase-1 by electrospray mass spectrometry. (Lengqvist et al., 2004) Robinson's group found that detergent micelles could be used to protect a membrane protein complex during the transfer from solution phase into the gas phase and that detergent could be removed by collisional activation. (Barrera et al., 2008; Barrera et al., 2009; Barrera & Robinson, 2011; Marcoux & Robinson, 2013) Besides determining the stoichiometry, this approach shows the potential to study the presence of small molecules and their influence on the structure and stability of the complex. (Zhou et al., 2011) Figure 5 shows the applicability of native ESI-MS to study membrane protein complexes.

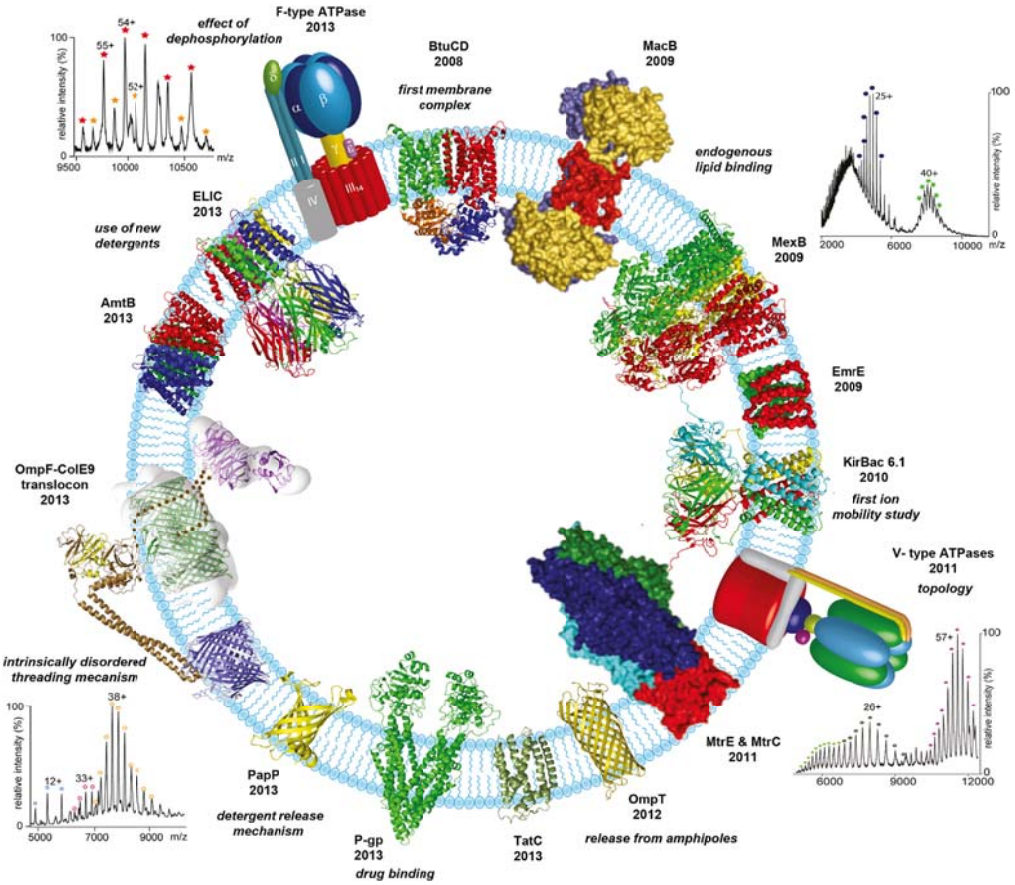


Figure 5. Membrane protein complexes studied by native ESI-MS in the past five years. Reprinted with permission from (Marcoux & Robinson, 2013). Copyright © 2013 Elsevier Inc.

Tandem mass spectrometry (MS/MS) experiments allow the determination of the arrangement of subunits in the complexes. Complex precursor ions can be fragmented via different dissociation pathways, including collision-induced dissociation (CID) (Hernández & Robinson, 2007), surface-induced dissociation (SID) (Zhou et al., 2012), and electron capture dissociation/ electron transfer dissociation (ETD) (Hofstadler & Griffey, 2001; Zhang et al., 2011). For example, Robinson and coworkers have used native ESI-IM-MS, combined with CID, to study a series of protein complexes, such as the membrane protein complexes discussed above. (Barrera et al., 2008; Barrera & Robinson, 2011; Hernández & Robinson, 2007; Marcoux & Robinson, 2013) The dissociation carried out with CID allows separate detection of the transmembrane or cytoplasmic subunits. Wysocki's group has found that, to a certain extent, SID yields a better-preserved

complex compared with CID. (Dongre et al., 1996; Zhou et al., 2013; Zhou et al., 2012)

The groundbreaking work described above highlights the potential of native MS for studying biomolecular complexes. However, the analysis of biomolecular complexes using ESI is not yet routine (van Duijn, 2010). Careful sample preparation is crucial for successful native-ESI-MS measurements. Because nano ESI-MS is sensitive to salts and is largely incompatible with detergents, these additives need to be removed before MS analysis. (Yin & Loo, 2009) Although attractive and simple techniques were developed to deplete unwanted salts and detergents, e.g., using molecular weight cut-off filters, they still pose several problems. Transient and weak interactions are easily lost during multi-step and/or harsh isolation procedures. The interactions are often sensitive to salt concentration, temperature and/or pH. At present, it appears that each biomolecular complex requires its own specific protocol, hampering high-throughput analysis and routine use.

c.2. Investigation of Biomolecular Complexes via MALDI-MS

As one of the two "soft" ionization techniques, MALDI allows the sensitive detection of large, non-volatile, and labile molecules by mass spectrometry. Compared with native ESI-MS, MALDI-MS has not yet contributed widely to the study of biomolecular complexes. The primary reason for this is that MALDI causes dissociation of the non-covalent interactions during sample preparation or in the ionization process, and can induce formation of nonspecific aggregates. However, under appropriate conditions, MALDI-MS can successfully be used and allows for the detection of non-covalent complexes. Bolbach (Bolbach, 2005) has provided an excellent review of this field. In this section, critical factors to preserve non-covalent interactions among biomolecular complexes in both the desorption/ionization processes (e.g., first shot phenomenon) and the target preparation, including the matrix selection, will be covered. The challenges to detect high molecular weight ions with MALDI-TOF are also addressed from the instrumental and methodological aspects. Various examples will be presented showing the potential of MALDI-MS application for the analysis of biomolecular complexes.

The term matrix-assisted laser desorption ionization (MALDI) was first coined in 1985 by Hillenkamp, Karas and their colleagues. (Karas et al., 1985) While they employed organic matrices, Tanaka and coworkers used ultrafine cobalt particles dispersed in glycerol to ionize biomolecules as large as the protein carboxypeptidase-A with a molecular weight 34,472 Da. (Tanaka et al., 1988) While Tanaka's soft laser desorption ionization approach was successful, MALDI offers more practical advantages and has been much more widely used. Currently, MALDI is routinely utilized for protein sequencing and proteomic research. Similar to ESI, MALDI is also a powerful tool for studying DNA/RNA, lipids, and glycoconjugates. However, MALDI generates ions from the solid phase, more precisely from a mixture of analyte and a large excess of an appropriate matrix, as shown in Figure 6. In MALDI, the analyte is first co-crystallized with a large molar excess of a matrix, which is usually a UV-absorbing, weak organic acid. Laser irradiation of this analyte-matrix mixture results in vaporization of the matrix, which carries the analyte with it. Therefore, the matrix does not only absorb the laser light, but also indirectly induces the analyte to vaporize. Moreover, the matrix also serves as a proton donor and acceptor, helping to ionize the analyte. (Hillenkamp et al., 1986)

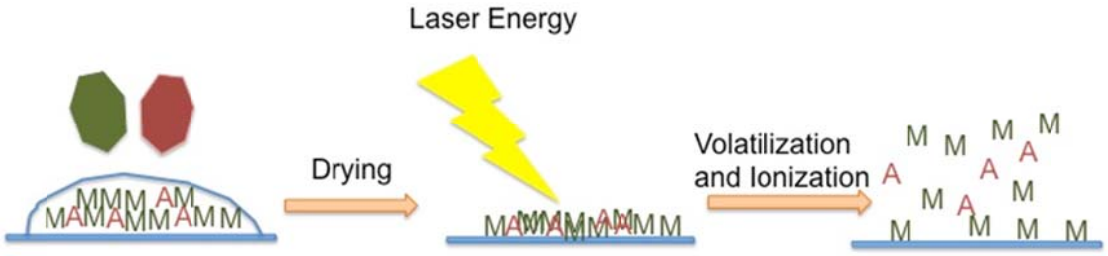


Figure 6. Principle of the MALDI process: an analyte (A) is co-crystallized in a large molar excess of matrix (M). Irradiation with a laser leads to volatilization and ionization of the matrix and the analyte.

The MALDI technique involves the processes of desorption and ionization of the analyte with the assistance of the matrix. Despite the importance of the matrix, very little is known about what makes a good matrix for a certain analyte. Some properties a matrix should have are solubility, absorptivity, reactivity, volatility and vacuum stability. (Hillenkamp & Peter-Katalinic, 2007) The analyte of interest should co-dissolve with the matrix; the matrix, rather than the analyte,

should efficiently absorb the energy from the laser; the matrix itself should not covalently modify analytes; it should have a sufficiently low vapor pressure to be stable in vacuum; and it should facilitate co-desorption of the analyte upon laser irradiation. For proteins/peptides, commonly used matrices include 2,5-dihydroxybenzoic acid (2,5-DHB), α-cyanohydroxycinnamic acid (HCCA or CHCA), and sinapinic acid (SA). For acidic oligosaccharides, 6-aza-2-thiothymine (ATT) and 2,4,6-trihydroxyacetophenone (THAP) were found to be particularly useful. In the field of nucleic acids, THAP, 3-hydroxypicolinic acid (3-HPA), picolinic acid (PA), ATT, and 2,5-DHB are standard choices. Lipids are low-molecular weight compounds, and MALDI analysis can be compromised due to matrix background. 2,5-DHB has been found to be the most effective matrix for various types of lipids and phospholipids by far. Among these, 2,5-DHB, HCCA, SA and PA are widely used in positive ion mode, and the rest are mostly used in negative ion mode.

Besides the analyte ions, non-specific cluster can be observed in MALDI mass spectra. These clusters may be generated during co-crystallization of the sample with the matrix or in the MALDI plume, depending on the laser energy. The thermochemical properties of the monomeric species and the chemical environment affect such cluster ion formation. Formation of cluster ions is influenced by analyte concentration and laser pulse energy. (Livadaris et al., 2000) To differentiate between non-specific cluster ions and specific non-covalent complexes, control experiments must be carried out. (Zenobi & Knochenmuss, 1998; Karas et al., 2000; Livadaris et al., 2000) The most powerful strategy is to modify one or several functionalities that are important for the formation of a specific non-covalent complex with an experiment where such a modification is absent and then cross-compare. Moreover, non-specific clusters exhibit an exponentially decaying intensity distribution; specific complexes are normally observed with more intense signals that stick out of the exponentially decaying distribution. Applying higher laser energy or higher analyte concentration will promote the formation of non-specific cluster ions.

c.2.1 MALDI Mechanism

Although MALDI has been introduced and further developed more than two decades already, the detailed mechanism of MALDI is still not fully understood. MALDI ionization is a complex phenomenon since multiple physical and chemical processes are involved in ions generation, which starts from the condensed phase. The time frame and mode of material ejection helps us to understand the events occurring after an ablation. (Figure 7) The laser pulse typically lasts a few nanoseconds, e.g., 5 ns for a typical Nd:YAG laser. The excited state decay, which extends into the time regime of secondary reactions, is many microseconds. Different physical properties of desorption and ablation by phase explosion and spallation have been investigated and reviewed by Dreisewerd 2003. (Dreisewerd et al, 2003)

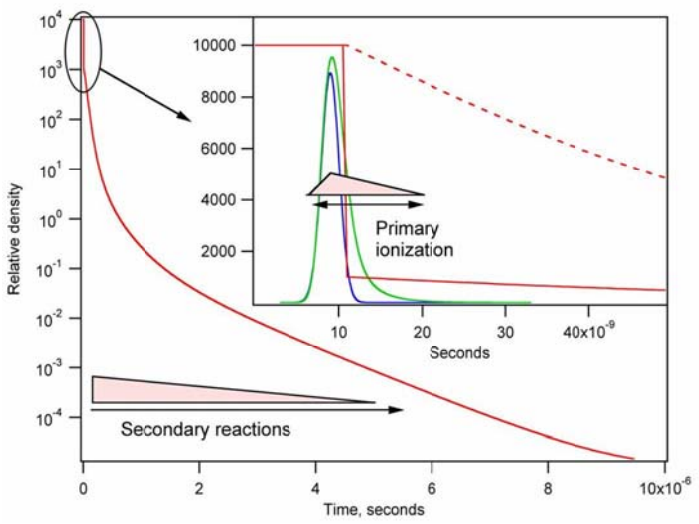


Figure 7. Origin of the two-step model: expansion of the MALDI plume and excited state decay. Density changes during the MALDI event: density is plotted versus time for a plume expanding as an adiabatic free jet. (In the inset showing the primary ionization process, the blue curve represents a typical laser pulse, green is for the population of excited states and red describes the relative density of the expanding plume.) (Reproduced with permission from (Knochenmuss, 2006; Knochenmuss & Zhigilei, 2010). Copyright © 2006 Royal Society of Chemistry.

A two-step framework, primary ion formation and secondary ionization reactions, was introduced to better explain MALDI ionization. (Zenobi & Knochenmuss, 1998) The primary ionization processes happen during and soon

after the laser pulse. In this period, energy and material density are sufficient for high-energy processes, i.e., first ion generation. Two major models are popular in primary ion generation process, namely photochemical ionization (PI) (Ehring et al., 1992) and cluster ionization (CI) (Karas et al., 2000; Karas & Krüger, 2003; Krüger et al., 2001). The secondary ionization continues for much longer than primary ionization. In this process, ion-molecule reactions in the desorption/ablation plume generate new ions. The two most important secondary reactions in the positive mode are proton transfer and cation transfer. (Zenobi & Knochenmuss, 1998; Knochenmuss & Zenobi, 2003)

c.2.1.1. Primary Ionization Processes

The first step of the two-step ionization process is the generation of primary ions; how the first ions arise remains controversial. Table 2 list possible models of primary processes. We will only briefly touch on major models here, namely photochemical ionization (PI) (Ehring et al., 1992) and cluster ionization (CI) (Karas et al., 2000; Karas & Krüger, 2003; Krüger et al., 2001). In the photochemical ionization model, analyte ions are considered to be produced from a protonation or deprotonation process involving an analyte molecule colliding with a matrix photo ion in the gas phase. The photochemical ionization model was first proposed by Ehring in 1992 to explain the MALDI process; (Ehring et al., 1992) energy pooling and multiphoton absorption have been suggested to lead to ionization of matrix molecules.

The cluster model was proposed and largely developed by the Karas group. (Karas et al., 2000; Karas & Krüger, 2003; Krüger et al., 2001). It began as the "lucky survivor" model; the ions were taken to be largely preformed in the solid matrix, and extensive neutralization was suggested to happen in the plume. The key processes in this and other cluster models are described in Figure 8. Clusters of various sizes, which statistically contain a net excess of positive charge (others of negative charge), are first ablated during laser irradiation. In a desolvatation phase, ions are liberated by evaporation of neutral matrix and analyte molecules. During this process, the charge will migrate to yield the thermodynamically most stable ions.

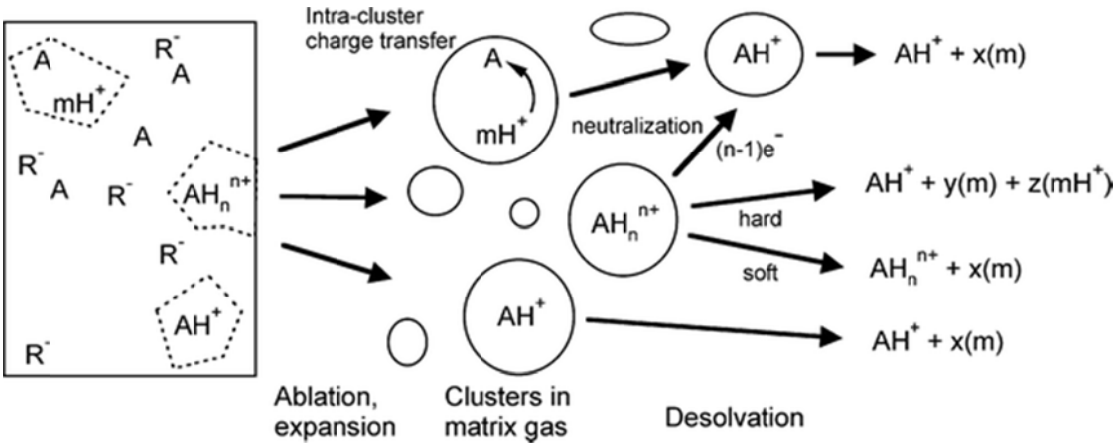


Figure 8. Illustration of the key processes in the cluster model of MALDI ionization. After undergoing desolvatation and neutralization, cluster ions become the "surviving" ions in the end. Reproduced with permission from (Knochenmuss, 2006). Copyright © 2006 Royal Society of Chemistry.

c.2.1.2. Secondary Processes

Depending on the mode of primary ion generation, primary ions could be matrix or analyte ions or both. Secondary processes start as soon as primary ions exist. As shown in Figure 7, the density drops slowly in the plume and many collisions occur during this period. Primary ions will react with neutral molecules to generate secondary ions, which will finally be recorded by the detector. Possible secondary reactions are summarized and listed in Table 3. The most important secondary reactions in positive ion mode are proton transfer and cation transfer. (Zenobi & Knochenmuss, 1998; Knochenmuss & Zenobi, 2003) Electron or anion transfer (McCarley et al., 1998) can also contribute to the secondary processes. Concerning proton transfer reactions, the gas phase proton affinity of the analytes is critical. If the proton affinity of the analyte is higher than that of the matrix, the proton-transfer will be exothermic and favorable. For many biomolecules, such as proteins or peptides, protonation reactions are dominant. For analysis of substances with low proton affinity, e.g., synthetic polymers, cationization by metal ions is often favored.

Table 2. Summary of primary processes in MALDI Ionization. (Adapted from Zenobi & Knochenmuss, 1998)

Name	Description	Remarks
Multiphoton Ionization	$M \xrightarrow{n(hv)} M^{+\cdot} + e^-$ $M \xrightarrow{hv} M^* \xrightarrow{m(hv)} M^{+\cdot} + e^-$	A 2-photon process would be required for ion formation.
Energy Pooling and Multicenter Models	$MM \xrightarrow{2hv} M^*M^* \xrightarrow{m(hv)} M + M^{+\cdot} + e^-$ $M^*M^* + A \rightarrow MM + A^{+\cdot} + e^-$	Explains how energetic processes can take place even in a diffusely excited solid. It is also used to explain matrix suppression effects.
Excited-State Proton Transfer	$M \xrightarrow{hv} M^*$ $M^* + A \rightarrow (M-H)^- + (AH)^+$ $M^* + M \rightarrow (M-H)^- + (MH)^+$	Commonly used matrices are not photo-active.
Disproportionation Reactions	$MM \xrightarrow{2hv} (MM)^* \rightarrow (M-H)^- + (MH)^+$ $MM \xrightarrow{2hv} (MM)^* \rightarrow M^- + M^+$	Although it brings down the instantaneous energy requirements to within the two-proton range, the strong experimental evidence is missing to support disproportionation reactions.
Desorption of Preformed Ions	$A^+X^- \xrightarrow{hv} A^+ + X^-$	It is difficult to be certain that ions finally observed in the mass spectrum are truly preformed, rather than the result of secondary gas-phase chemistry in the plume.
Thermal Ionization	$\alpha_i = C\exp((\phi - IP_i)/kT)$ $2M \xrightarrow{\Delta H} M^- + M^+ \;\; \alpha_i = C\exp((EA - IP_i)/kT)$	The temperature achieved cannot induce significant a thermal ionization/disproportionation reactions.
The “Lucky Survivor” or “Cluster”model	Figure 8	Ion-molecule ractions in the gas phase are no longer required. This model has difficulties to explain the predominant production of singly charged ions for biomolecules, and also sweet spot phenomenon and the matrix suppression effect.

Table 3. Summary of secondary processes in MALDI Ionization. (Adapted from Zenobi & Knochenmuss, 1998)

Name	Description	Remarks
Gas-Phase Proton Transfer		
a) Matrix-matrix reactions	$M + M^{+\cdot} \rightarrow MH^+ + (M-H)^{\cdot}$ $M + e^- \rightarrow (M-H)^- + H^{\cdot}$	These certainly occur in the plume.
b) Matrix-analyte reactions	$MH^+ + A \rightarrow M + AH^+$ $(M-H)^- + A \rightarrow (A-H)^- + M$	Protons transfer to a neutral gas-phase peptide or protein will almost always be thermodynamically favorable. (PA value)
Gas-Phase Cationization	$(M+Na)^+ + A \rightarrow M + (A+Na)^+$	This process is favorable for the compounds with low proton affinities, such as polymers.
Electron Transfer	$M^{+\cdot} + A \rightarrow M + A^{+\cdot}$ $M^{-\cdot} + A \rightarrow M + A^{-\cdot}$	This process can happen only if the analyte has a lower ionization potential than the matrix. (Not favorable for peptides and proteins)
Charge Ejection		Coulomb repulsion could induce the proton ejection into the gas phase.

c.2.2. High-Mass Detection in MALDI-MS

MALDI uses a pulsed laser for ionization and is therefore best coupled to a time-of-flight (TOF) mass analyzer. A TOF can record all the ions generated by each laser pulse very rapidly. Theoretically, the TOF instrument has an unlimited mass range. However, the mass range is limited by the performance of commonly used detectors, i.e., multichannel plate (MCP) detectors. MCPs suffer from very low sensitivity in the high mass range due to their low ion-to-electron conversion efficiency for ions with low velocity. Moreover, the long channel recovery time causes saturation effects in the presence of smaller ions. To overcome these drawbacks, other significant improvements to the detector design were needed. These improvements include cryodetectors, conversion dynodes, and inductive detectors. (Cole, 2011) The company Comet AG (Flamatt, Switzerland) developed a cryodetector instrument to increase the sensitivity for high-mass ion detection. (Wenzel et al., 2005) With this instrument, singly charged ions in the megadalton range, such as immunoglobulin M and von Willenbrand factor protein oligomers, could be detected (Figure 9).

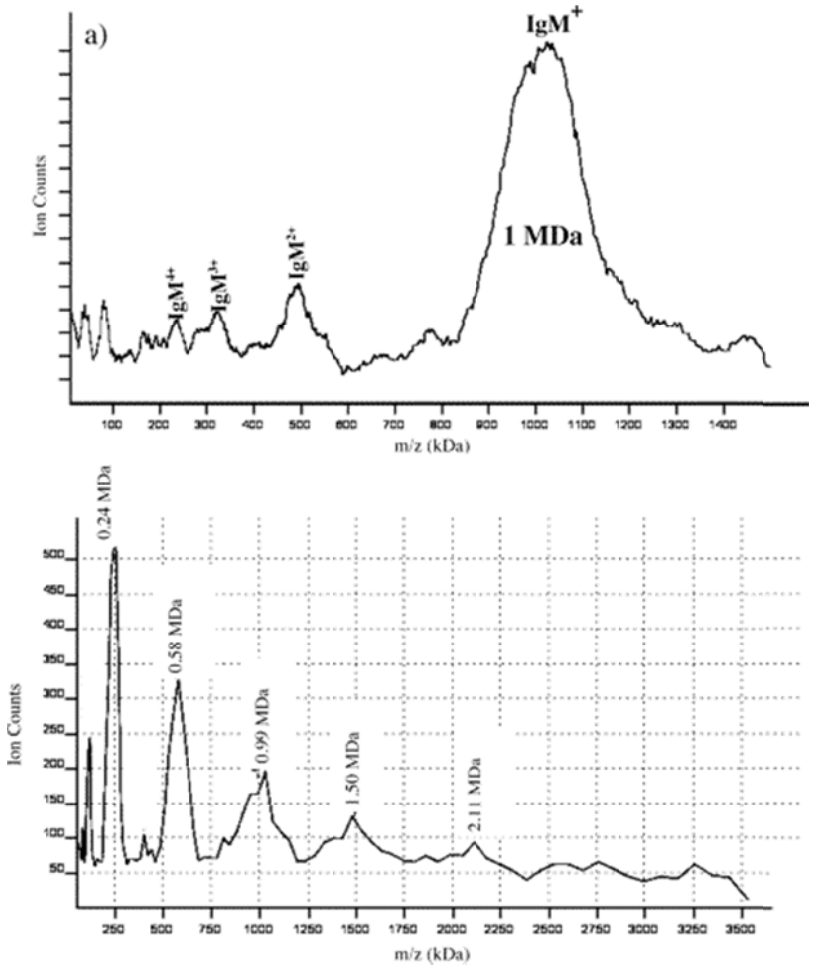


Figure 9. Mass spectra of two proteins recorded by high-mass MALDI-MS, a) immunoglobulin M at 1 MDa and b) von Willenbrand factor protein oligomers with a molecular weight up to 1.5 MDa. Reproduced with the permission from (Wenzel et al., 2005). Copyright © 2005 American Chemistry Society.

Unfortunately, these cryodetector instruments were very time- and cost-intensive to maintain, are not available any more, such that alternative detection principles for high-mass MADLI-MS were sought. To our knowledge, only commercially available high-mass detector is based on the ion conversion dynode (ICD) technology. Here, the high-mass ions (e.g., protein ions) are converted into secondary ions on a conversion dynode. These secondary ions are re-accelerated towards and finally detected with a discrete dynode secondary electron multiplier (SEM). (Hillenkamp et al., 2008; Nazabal et al., 2006) A high sensitivity for high-mass ions is achieved because the yield of sputtered secondary ions increases with increasing mass of the impacting ions. The

possibility of saturation caused by smaller ions is greatly reduced by adding capacitance buffers to the last dynodes of the SEM.

With the high-mass detector development, the *m/z* range has been extended to approximately 1.5 MDa and is usually limited by restrictions in the software used for recording the mass spectra. An accurate and reliable calibration is critical for MALDI-MS in studies of biomolecular complexes. In principle, the calibrant should cover a broad *m/z* range with good sensitivity. Several calibration standards, such as bovine serum albumin (Belgacem et al., 2002; Pimenova et al., 2008), immunoglobulin G (Pimenova et al., 2008) or Invitromass (a commercial high molecular weight mass calibration kit) (Kemptner et al., 2010; R. Müller & Allmaier, 2006) have been employed. Nonspecific aggregates of medium-size proteins are sometimes also used. (Ludwig et al., 2011) Our group recently designed a special system for calibration, especially for the high-mass region. By cloning several copies of maltodextrin-binding protein (MBP, 40 kDa) into the same expression vector and expression in *Escherichia coli*, different calibrants, MBP, MBP_2, MBP_3, MBP_4, and MBP_6 are obtained. (Weidmann et al., 2013a; 2013b) With this system, a mass range from 40 to approximately 400 kDa can be covered. Besides the broad *m/z* range covered, this new calibration system achieves a much better resolution compared to the other potential calibration candidates with similar molecular weight, such as bovine thyroglobulin, recombinant human fibronectin, and GroEL.

c.2.3. "Soft" MALDI to Study Non-Covalent Complexes

c.2.3.1. Principles of "Soft" MALDI

Analysis of non-covalent complexes in MALDI requires particular experimental conditions, including proper choice of matrix, solvent, pH, concentrations and laser energy to preserve the non-covalent interactions. Here, we use the term "soft" MALDI to describe the fine-tuned MALDI conditions in studies of non-covalent complexes. Compared with an acidic matrix, non-acidic matrices allow a better preservation of the non-covalent interactions. A number of non-acidic matrices, such as 2-amino-4-methyl-5-nitropyridine, para-nitroaniline and ATT, were introduced in the 1990s. (Jespersen et al., 1998; Lecchi et al, 1995)

Applying non-acidic matrices has advantages for analyzing intact peptide-oligonucleotide complexes or DNA duplexes. However, these matrices were not widely applied since they cannot easily generate signal from high molecular weight compounds with comparable spectral quality than acidic matrices, i.e., sinapinic acid.

In some cases, the observation of intact non-covalent complexes with MALDI is possible if data is only acquired from the first few laser shots at a new sample spot; this is called the "first shot phenomenon". (Rosinke et al, 1995; Cohen et al, 2000; Wortmann et al, 2007) Rosinke mentioned this phenomenon for the first time in 1995; (Rosinke et al., 1995) Cohen studied the first shot phenomenon systematically to improve the detection of non-covalent complexes by MALDI. (Cohen et al., 2000) To explain this observation, three models have been proposed by Cohen, which are summarized in Figure 10 (A, B and C). One possibility is that large molecules, including protein complexes, segregate at the crystal surface during crystal growth. A second possibility is that an intact protein complex survives only at the surface of the crystal and the protein complexes are present in dissociated form in the interior of the crystal after the crystallization process. Third, the laser pulse may have a strong influence on protein complexes below the crystal surface. Because of the pulsed irradiation, protein complexes around the crater may be dissociated. The first shot phenomenon was reported in quite a few publications (Figure 10 (D)); however, it also seems to be a sample-, preparation- and matrix-dependent phenomenon rather than a general fact. (Wortmann et al., 2007)

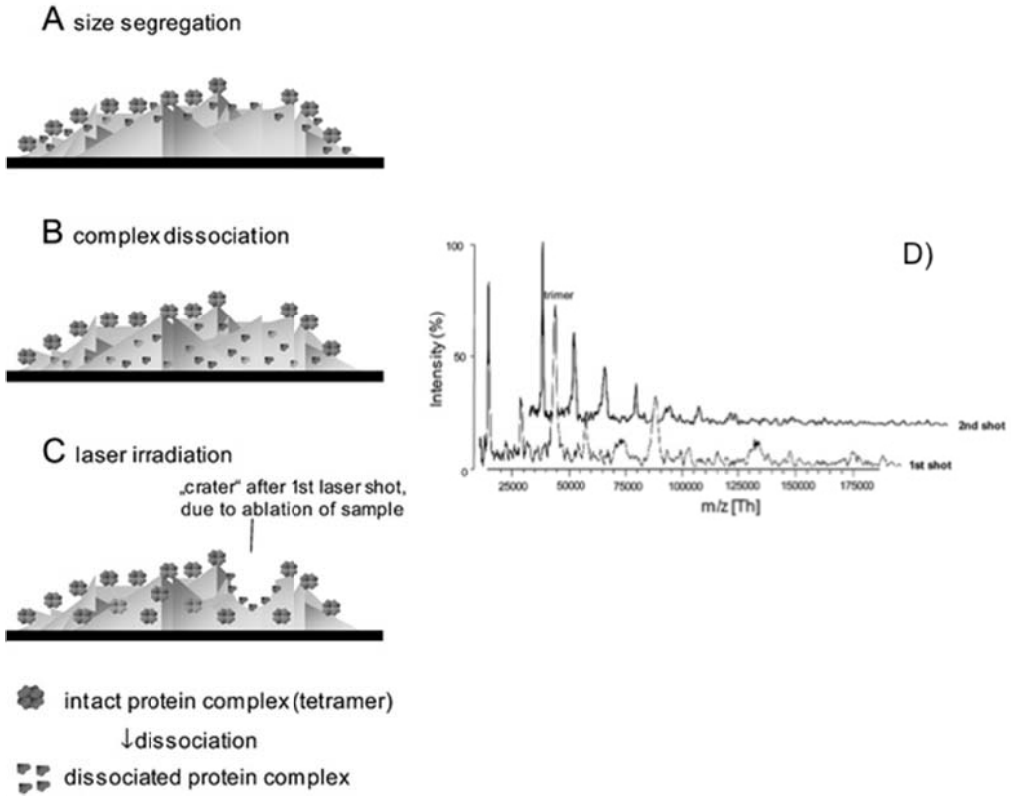


Figure 10. Illustration of three proposed models to explain the "first shot phenomenon" and chorismate mutase as the example for the first shot phenomenon. Reproduced with permission from (Wortmann et al., 2007). Copyright © 2007 Royal Society of Chemistry.

c.2.3.2. Application of "Soft" MALDI for Studying Non-Covalent Complexes

The first detection of a noncovalent protein complex by MALDI was reported by Hillenkamp's group, with nicotinic acid as the matrix dissolved in 10 % ethanol/water. (Karas et al., 1989; Hillenkamp et al., 1991) As shown in Figure 11, the predominant signal comes from the non-covalently bound tetramer of glucose isomerase. (Karas et al., 1989) A transmembrane Ompt trimeric porin was also reported (Rosinke et al., 1995) by the same group using ferulic acid as the matrix. 6-aza-2-thiothymine matrix was used by Glocker et al. without addition of any organic solvent, and an intact non-covalent protein complex, RNAseS, and also the non-covalent complexes between S-protein, S-peptide and specific dimers of coiled-coil leucine zipper polypeptide were observed. (Glocker et al.,1996) Moreover, 2,6-dihydroxyacetophenone (Zehl & Allmaier, 2004), 2,4,6-trihydroxyacetophenone (Thiede & Janta-Lipinski, 1998) or 4-nitroaniline (Jespersen et al., 1998) have been reported to preserve non-covalent complexes

in MALDI. Standard matrices buffered to physiological pH could also be used to detect intact complexes: sinapinic acid in acetonitrile (ACN)/0.2 M Bis-tris (v:v, 30:70, pH 7) was used as the matrix in the heterodimer human immune farnesyl protein transferase detection (Farmer & Caprioli, 1998); sinapinic acid dissolved in ACN/H_2O (v:v, 1:1), without addition of TFA, was used as the matrix to analyze the complex of β-lactoglobulin and polyclonal anti-β-lactoglobulin antibody (Schlosser et al., 2003); Cotter reported the detection of an enzyme•peptide complex by a saturated solution of sinapinic acid in ethanol/1M ammonium citrate (v:v 1:1) (Woods et al., 1995). All the examples show the possibility of detecting non-covalent complexes with MALDI-MS, once proper matrix is used.

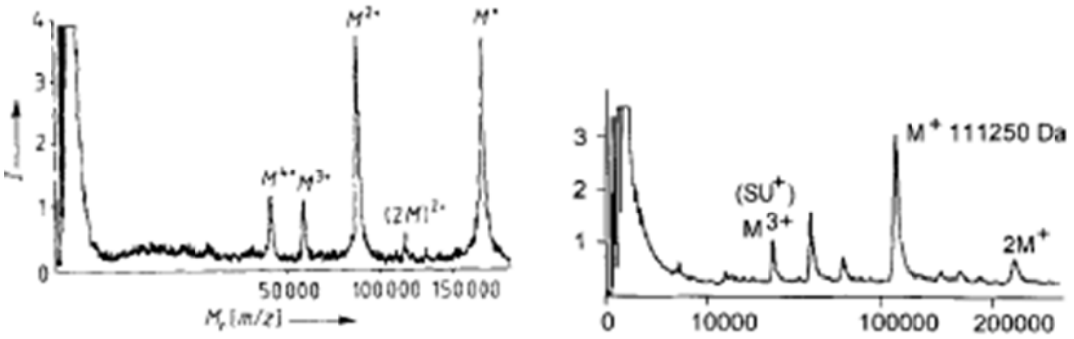


Figure 11. a) LDI (laser desorption ionization) of glucose isomerase. Reproduced with permission from (Karas et al., 1989). Copyright © 1989 John Wiley & Sons. (b)UV-MALDI mass spectra of porin with the matrix of ferulic acid in THF. Sum of first shots to on a given spot. Reproduced with permission from (Rosinke et al., 1995). Copyright © 1995 John Wiley & Sons.

Cohen reported that with 2,5-DHB as the matrix they could detect the non-covalent structure of homo-oligomeric complexes of streptavidin, alcohol dehydrogenase and beef liver catalase only from the first shot at a given position. (Cohen et al. 1997; Horneffer et al., 2006) By gathering the first shot signals, the intact allophycocyanine (APC, Figure 12), a noncovalent 107 kDa protein heterohexamer, was analyzed in 6-aza-2-thiothymine matrix. Moreover, the Omp F porin protein from *E. coli*, was recorded as a homotrimer by gathering first shot signals. (Rosinke et al., 1995)

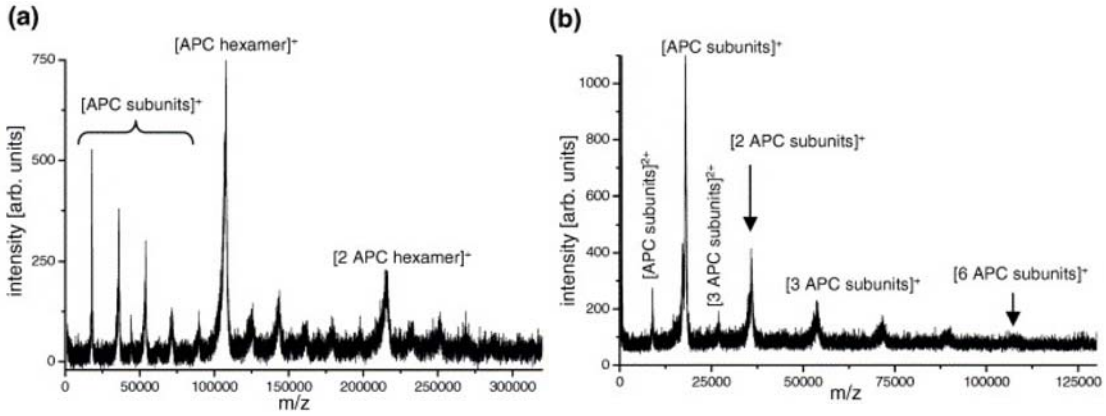


Figure 12. Mass spectra of APC with the matrix ATT: a) accumulation of only first shots on 20 different spots, b) accumulation of second and up to 30 following shots onto a given spot. The APC hexamer is only observed in the accumulation of first shots. Reproduced with permission from (Horneffer et al., 2006). Copyright © 2006 Elsevier B.V..

Comprehensive reviews have been provided by the groups of Hillenkamp (Hillenkamp 1998), Caprioli (Farmer & Caprioli, 1998) and Allmaier (Zehl & Allmaier, 2005). However, there are no clear guidelines for selecting the matrix to analyze non-covalent complexes, i.e., the success is not predictable. The use of non-acidic matrices sometimes suffers from low sensitivity for the analytes. Moreover, quite a few cases listed above are complexes between protein/peptide and oligonucleotides, where ionic interactions facilitate the MALDI measurement in the gas phase. It has also been noticed that the "first shot phenomenon" is not general. Thus, MALDI-MS is still not widely applied for the investigation of non-covalent interactions.

It has been noticed that chemical cross-linking, a technique that dates back to the 1920s, combined with mass spectrometry, especially MALDI-MS, provides a general approach to study non-covalent biomolecular complexes. (Farmer & Caprioli, 1998, 2005; Bich & Zenobi, 2009; Mädler et al., 2013) In the following section, we will address the nature and the applications of chemical cross-linking, combined with MALDI-MS, for the analysis of non-covalent complexes.

c.2.4. Chemical Cross-linking Combined with MALDI-MS

Cross-linking is a process that links two components through covalent bonds with the assistance of a cross-linking reagent. The cross-linked components

could be proteins, peptides, drugs, nucleic acids or even solid particles. Application of cross-linking in structural analysis, including the analysis of three-dimensional and quaternary structures, only started after its combination with mass spectrometry. (Kalkhof & Sinz, 2008; Leitner et al., 2010.; Müller & Sinz, 2012; Sinz, 2003; Sinz, 2006) There are quite a lot advantages of using chemical cross-linking combined with mass spectrometry, especially MALDI-MS, to study non-covalent biomolecular complexes. First, the chemistry of the cross-linking reagents is well studied; second, available cross-linking reagents target specific functional groups, including primary amines, sulfhydryls or carboxylic acids, and are available with different spacer arm length; last but not least, MALDI-MS is tolerant to the presence of excess cross-linking reagents. Due to its application in three-dimensional structure analysis, a brief description for each cross-linking strategy will be presented first. Afterwards, we will discuss currently widely used cross-linking reagents.

c.2.4.1. Cross-linking Strategies

In principle, two alternative strategies exist for chemical cross-linking experiments, which are commonly named the "bottom-up" approach and the "top-down" approach. (Figure 13) In the bottom-up approach, the cross-linked species will be enzymatically digested. The bottom-up approach allows a low-resolution three-dimensional structure determination. (Chen et al., 2010; Leitner et al., 2010; Sinz, 2003; Young et al., 2000) To confirm and optimize cross-linking protocols, SDS-PAGE and MALDI-TOF have been used after cross-linking. This process is also used in the determination of the quaternary structure of protein complexes. (Bich & Zenobi, 2009; Mädler et al., 2013) Rather than digestion of the cross-linked species, the top-down approach analyzes cross-linked products directly. The cross-linked products are screened and fragmented via different techniques, including sustained off-resonance irradiation collision-induced dissociation (SORI-CID), infrared multi-photon dissociation (IRMPD) or electron capture/transfer dissociation (ECD/ETD) (Kelleher et al., 1999; Kruppa, et al., 2003; Li et al., 2011; McLafferty, 1999; Novak et al., 2005; Novak & Giannakopulos, 2007). The top-down approach is always carried out with

electrospray ionization fourier-transform ion-cyclotron resonance (ESI-FTICR) mass spectrometry.

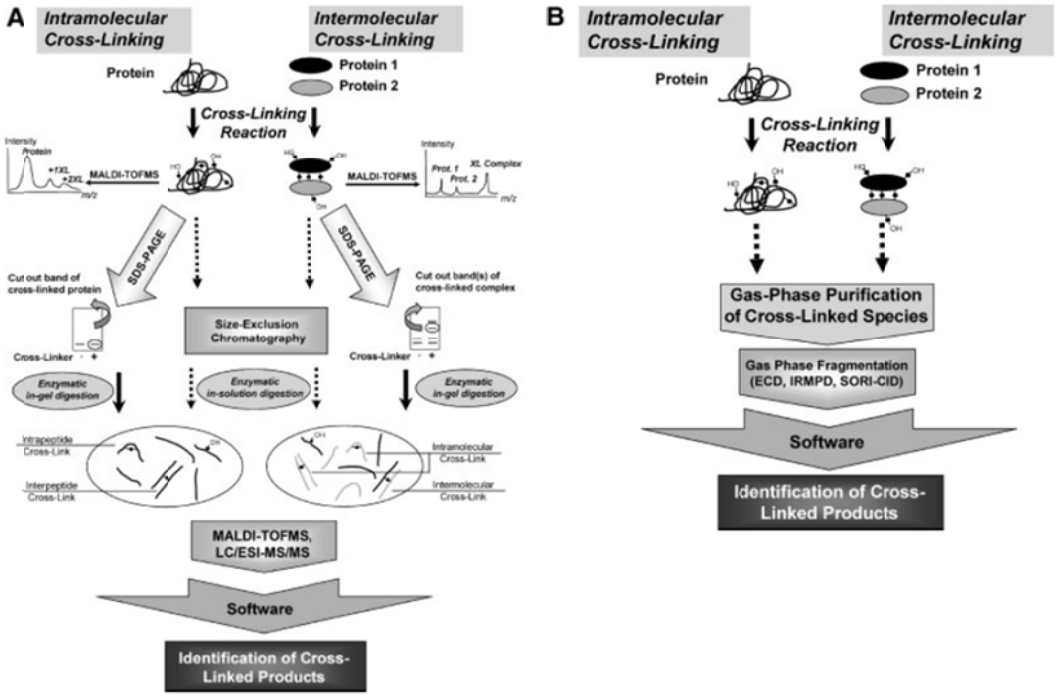


Figure 13. Two strategies for the low-resolution protein structure determination by chemical cross-linking and mass spectrometry: A) bottom-up approach, and B) top-down approach. Reproduced with permission from (Sinz, 2005). Copyright © 2005 Springer.

c.2.4.2. Cross-linking Reagents

Before discussing the chemistry of the cross-linkers, we will first present a general description of current cross-linker designs (Figure 14). Generally, there are four types: 1) homobifunctional cross-linking reagents basically contain identical functional groups on both reactive sites; 2) heterobifunctional cross-linking reagents contain two different reactive groups that target different functional groups; 3) zero-length cross-linkers, such as carboimides, with EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide) being the most popular representative; 4) trifunctional cross-linkers allow an additional link to a third protein to be fromed or incorporate a biotin moiety for affinity purification.

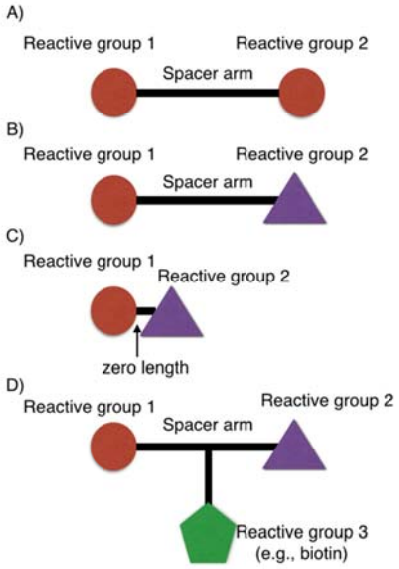


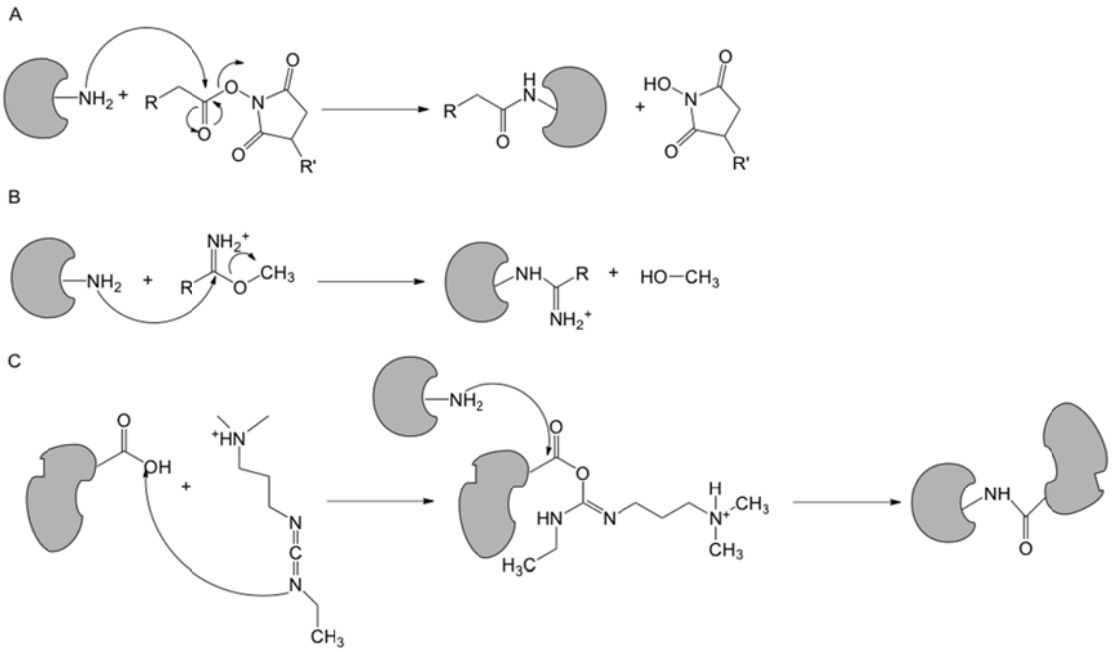
Figure 14. Different types of cross-linking reagents: A) homobifunctional cross-linker, B) heterobifunctional cross-linker, C) zero-length cross-linker and D) trifunctional cross-linker D. (Adapted from Sinz, 2006)

Chemical cross-linking introduces covalent bonds between different molecules (intermolecular) or within parts of the same molecule (intramolecular). Moreover, it can also form a mono-link (dead-end link). Hundreds of reagents that have been described in the literature (Hermanson, 2013; Wong, 1991) or are offered commercially are based on a small number of organic chemical reactions. The most widely used classes of cross-linking reagents are presented below, including amine-reactive cross-linkers, sulfhydryl-reactive cross-linkers, and photo-reactive cross-linkers.

c.2.4.2.1. Amine-Reactive Cross-linkers

The most frequently utilized cross-linkers target the primary amino group of lysine and the N-terminus. *N*-hydroxysuccinimide (NHS) esters, which were introduced more than 30 years ago, are almost exclusively used (Scheme 1(A)). Besides NHS esters, sulfo-NHS esters, their water-soluble analogs, are also used. Apart from acylation of the ε-amino group of lysines and the α-amino group of the N-terminus, side reactions with hydroxyl side chains have been also observed. (Mädler et al., 2009) The imidate functional group is one of the most

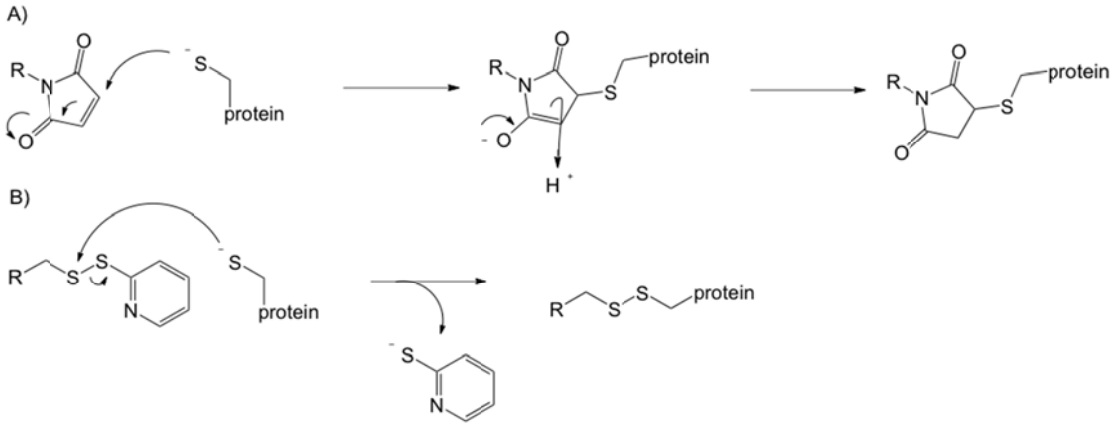
specific acylating groups to modify primary amines (Scheme 1(B)); imidoesters possess minimal cross-reactivity towards other nucleophilies. Imidoesters are highly water-soluble, but undergo degradation due to hydrolysis. As mentioned before, carbodiimides are so called “zero-length” cross-linking reagents in which the distance between the interaction partners does not change after cross-linking (Scheme 1(C)) and are used to mediate amide bond formation between spatially close groups (< 3 Å).



Scheme 1. Mechanism of reaction of primary amines in proteins with A) NHS esters, B) imidates, and C) “zero-length” cross-linker EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide).

c.2.4.2.2. Sulfhydryl-Reactive Cross-linkers

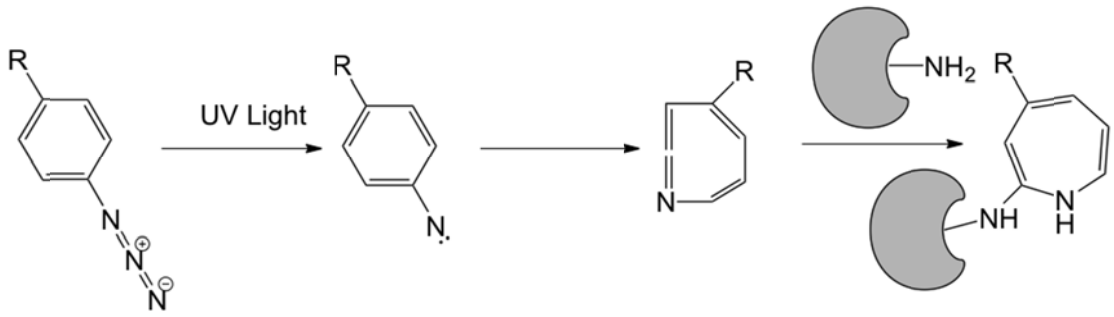
Cross-linking reagents targeting sulfhydryl-containing moieties are the second most commonly used reagents. Maleic acid imides (maleimides) are widely used as heterobifunctional cross-linking reagents (Scheme 2(A)); thiol-disulfide exchange reagents are applied to compounds containing a disulfide group. Maleimide can react with amines and be hydrolyzed at high pH values. (Brewer & Riehm, 1967) Disulfide exchange reactions occur over a broad range of conditions, from acidic to basic, and in the presence of buffer components. (Scheme 2(B))



Scheme 2. Mechanism of reaction of thiols in proteins with A) maleimides, B) pyridyl disulfides.

c.2.4.2.3. Photo-Reactive Cross-linkers

Photoreactive groups can only react with target molecules after exposure to UV light. The most popular type of photosensitive functionality is the aryl azide group. (Gilchrist & Rees, 1969) Short-lived nitrenes formed upon photolysis can insert non-specifically into chemical bonds of target molecules. (Scheme 3)



Scheme 3. Mechanism of aryl azides react with protein under UV light.

c.2.4.2.4. Glutaraldehyde

Among the many available protein cross-linking agents, glutaraldehyde has undoubtedly found the widest application in various fields, including histochemistry, microscopy, cytochemistry, the leather tanning industry, enzyme technology, chemical sterilization, and biomedical/pharmaceutical sciences. However, the structure of glutaraldehyde in aqueous solution has been the subject of more debate than for any of the other cross-linking reagents. Different glutaraldehyde structures have been found and have been summarized in the

literature (Figure 15). (Migneault et al., 2004) Glutaraldehyde can react with several functional groups of proteins, including amines, thiols, phenols, and imidazoles. The ability of different aldehydes to react with amino acids has been ranked in decreasing order of reactivity as follows: ε-amino, α-amino, guanidinyl, secondary amino, and hydroxyl groups. (Migneault et al., 2004) Despite its wide application, there is still no agreement about the chemical nature of the reaction of glutaraldehyde with proteins. It seems that no single mechanism is responsible for the reaction of glutaraldehyde with proteins.

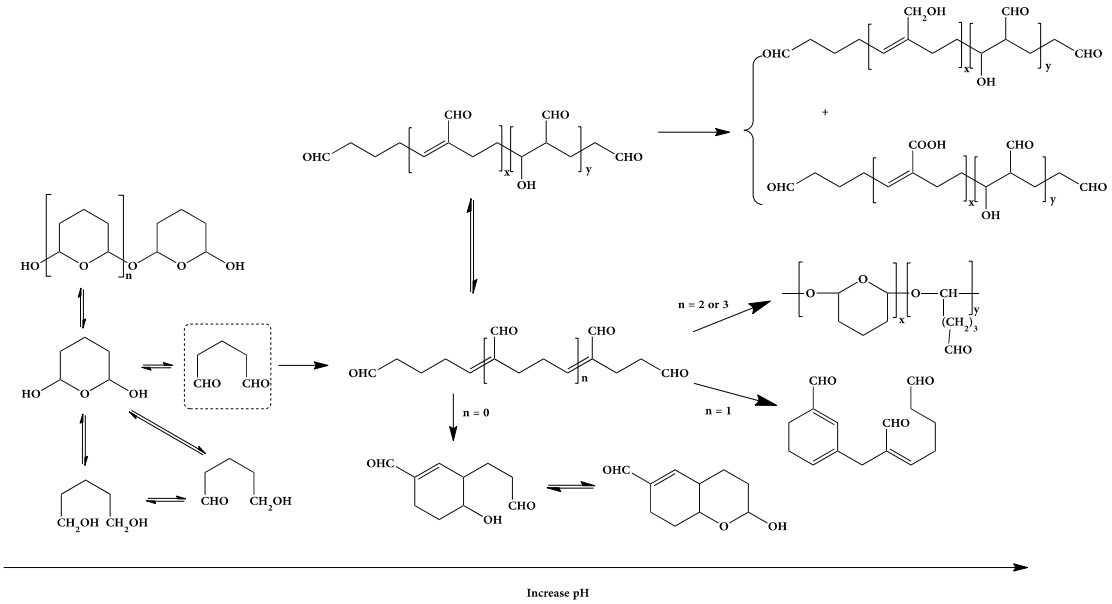


Figure 15. An overview of the possible forms of glutaraldehyde in aqueous solution. Adapted from (Migneault et al., 2004).

c.2.4.3. Application of Chemical Cross-linking and MALDI-MS in Biomolecular Complexes

Since the 1990s, chemical cross-linking has been combined with MALDI-MS analysis. Glutaraldehyde was used as the crosslinker to determine dimeric and terameric complexes, such as avidin and yeast alcohol dehydrogenase (Figure 17). (Farmer & Caprioli, 1991) However, to the best of our knowledge, glutaraldehyde was never widely used as the cross-linking reagent (Farmer & Caprioli, 1998). A possible explanation might be that the quality of the mass spectra deteriorates due the glutaraldehyde polymerization (Figure 15). Besides the application of glutaraldehyde for stabilizing complexes, numerous other

chemical cross-linkers, including nonselective photoreagents and site-specific linkers, have been applied for studying soluble complexes. Tomer reported the application of a heterobifunctional cross-linker, (sulfosuccinimido-2- (7-azido-4-methylcoumarin-3-acetamido)ethyl-1,3'dithiopropionate), to stabilize the noncovalent complex of human immunodeficiency virus glycoprotein 120 with its cellular receptor CD4 (Borchers & Tomer, 1999). In our previous studies, we demonstrated the combination of high-mass MALDI with chemical cross-linking, using mainly NHS esters, in different fields, such as epitope mapping, kinetic studies, sandwich assays for immunocomplexes (Nazabal et al, 2006; Bich et al., 2008; Pimenova et al, 2009), and monitoring of a ligand regulation mechanism (Bovet et al., 2008). Moreover, chemical cross-linking, combined with MS, can also be used for low-resolution structural studies. (Pimenova et al., 2008; Sinz, 2006) In addition to the increasing body of research in this field, the MALDI-MS instrument itself has been improved in the last decade, thus these technical improvement in combination with a careful selection of the experimental conditions can lead to the quality improvement of the spectrum.

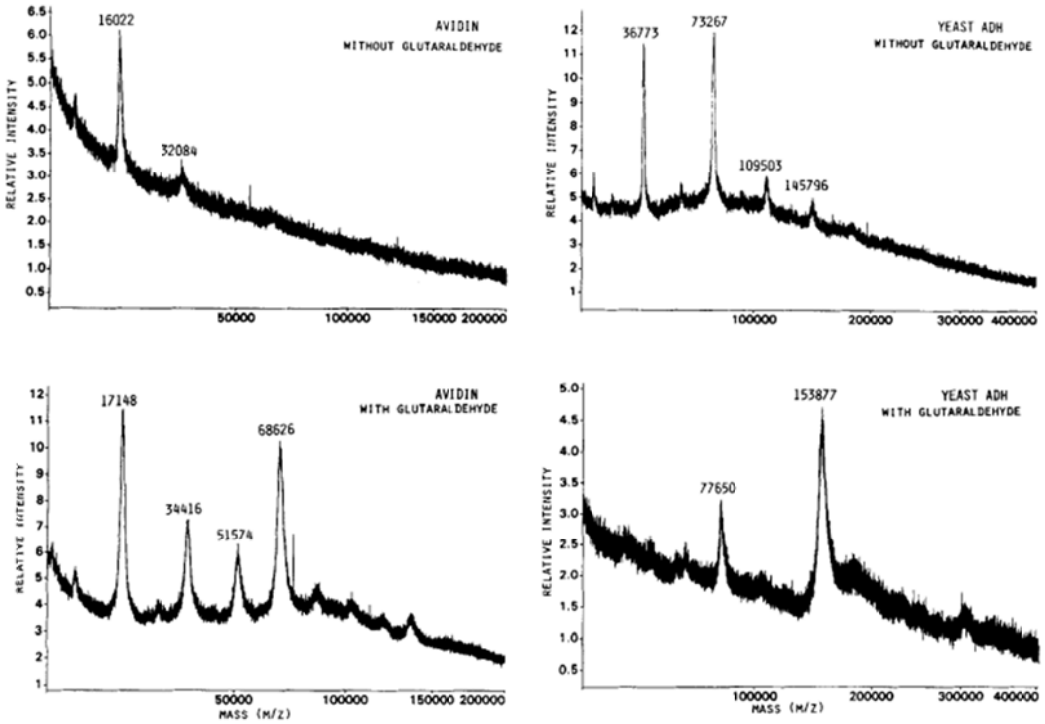


Figure 16. MALDI mass spectra of avidin and yeast alcohol dehydrogenase, before and after reaction with glutaraldehyde. Reproduced with permission from (Farmer & Caprioli, 1991). Copyright © 1991 John Wiley & Sons.

Chemical cross-linking, combined with MALDI-MS, could also be used in the studies of protein-oligonucleotide complexes, which require different cross-linkers. (Geyer et al., 2004; Golden et al., 2008; Kühn-Hölsken et al., 2010; Robinette et al., 2006) Cross-linking could be carried out based on photochemical reactions, with the natural UV reactivity of the nucleobases and of modified oligonucleotides. A typical UV-crosslinked protein-DNA sample, SSB with a DNA 19-mer, is shown in Figure 17. (Steen et al., 2001) In this example, MALDI-MS was applied after in-gel digestion. Our group reported the observation of a specific RAR•RXR•DNA (retinoic acid receptor, *9-cis* retinoic acid receptor and DR5) complex after applying NHS esters. However, we believe that NHS esters only stabilized the RAR•RXR part of the complex. (Bich et al., 2010) The observation of the intact RAR•RXR•DNA complex could be the result of strong electrostatic interactions in the gas phase. (Figure 18) Compared with studies of protein-protein complexes, chemical cross-linking combined with MALDI-MS faces more challenges in studies of protein-oligonucleoitide complexes: 1) difficulties to covalently bind protein-oligonucleotide complexes, 2) different matrix requirements for proteins and oligonucleotides, 3) a thorough purification step is always required to remove salts, and 4) the high probability that non-specific cluster ions from proteins and oligonucleotides are generated.

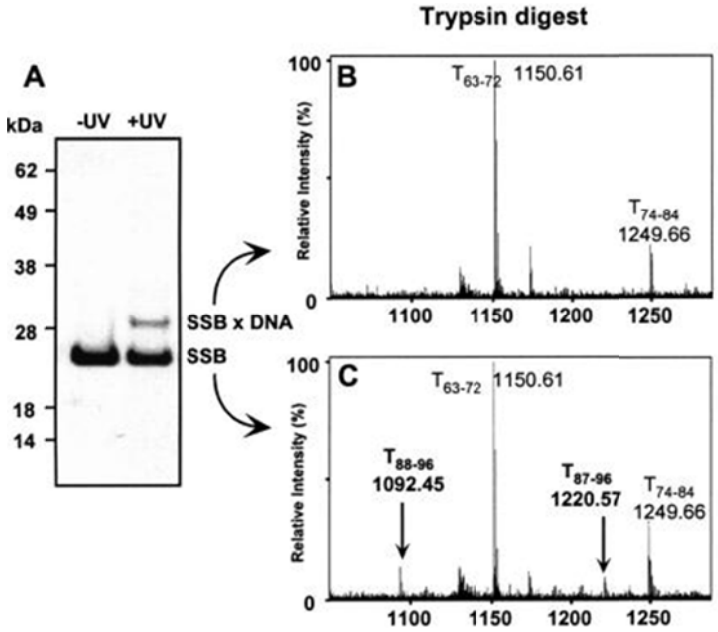


Figure 17. SDS-PAGE and differential peptide mass mapping of SSB•DNA complex. Reproduced with permission from (Steen et al., 2001). Copyright © The Protein Society.

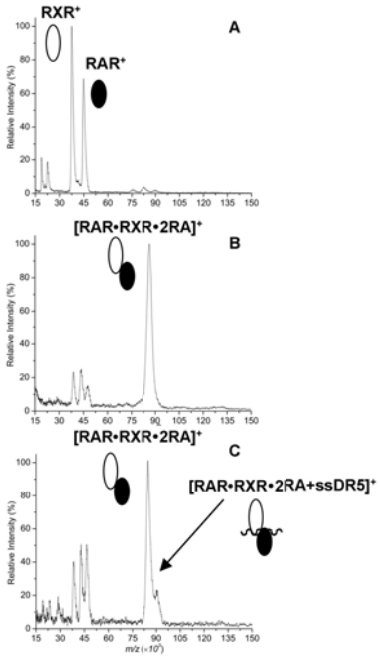


Figure 18. High-mass MALDI mass spectra of the RAR•RXR heterodimer (in the presence of RA. (a) Control experiment b) after chemical cross-linking and c) after incubation with a DNA strand, DR5. Reproduced with permission from (Bich et al., 2010). Copyright © Springer.

d. Motivation and Scope of this Thesis

As discussed above, MALDI-MS, with the assistance of chemical cross-linking, has been applied to study non-covalent interactions in biomolecules. However, its application is still quite limited. Up to now, published studies have focused on soluble protein complexes and some protein-oligonucleotide complexes. Considering the importance of non-covalent interactions in biomolecules and the advantages of MALDI-MS, the goal of this thesis was to develop this approach further to study challenging biomolecules, including protein-DNA/aptamer complexes and membrane protein complexes.

In Chapter 2, the Escherichia coli single-stranded DNA binding protein (SSB), which selectively binds single-stranded (ss) DNA and participates in the processes of DNA replication, recombination and repair, was chosen as the target for studying protein-oligonulcletide interactions. SSB forms a stable homotetramer in solution, but only the monomeric species can be detected with standard MALDI-MS. With chemical cross-linking, the quaternary structure of SSB is conserved and the tetramer was observed. We found that ssDNA also functions as a stabilizer to conserve the quaternary structure of SSB, as evidenced by the detection of a SSB•ssDNA complex even in the absence of chemical cross-linking. We noticed that the stability of the SSB•ssDNA complex with MALDI strongly depends on the length and strand of oligonucleotides and the stoichiometry of the SSB•ssDNA complex, which can be attributed to electrostatic interactions that are enhanced in the gas phase.

In Chapter 3, protein•aptamer complexes were targeted. With the application of a non-acidic matrix, the complexes of thrombin and two different thrombin binding aptamers (TBAs) could be observed directly. The complexes of PDGF-AB/BB and the specific PDGF binding aptamer (Apt-35) with a stoichiometry of 1:2 were observed. Detection of the complex between lysozyme and its corresponding aptamer further confirmed the capability of MALDI-MS to study such systems. The stronger thrombin•TBA29 complex showed a larger signal at the m/z of the intact complex than the weaker thrombin•TBA15 complex; the complex signal of Apt-35 and PDGF-BB was stronger in MALDI compared with

that of PDGF-AB and PDGF-AA. These observations indicate that the non-covalent interaction strength in solution is reflected in the MALDI mass spectra.

In Chapter 4, chemical cross-linking, from the aspects of the protein structures and the cross-linking reagents involved, was studied. By using two structurally well-known proteins, glyceraldehyde 3-phosohate dehydrogenase and ribonuclease S, the chemical cross-linking reactivity was compared using a series of homo- and hetero-bifunctional cross-linkers. This study suggested that the protein structure itself, especially the distances between target amino acid residues, was a determining factor for the cross-linking efficiency. Moreover, the reactive groups of the chemical cross-linker also play an important role; a higher cross-linking reaction efficiency was found for maleimides compared to 2-pyrimidyldithiols. The reaction between maleimides and sulfhydryl groups is more favorable than that between N-hydroxysuccinimide esters and amine groups, although cysteine residues are less abundant in proteins compared to lysine residues.

Analyzing purified membrane proteins and membrane protein complexes by mass spectrometry has been notoriously challenging, and requires highly specialized buffer conditions, sample preparation methods, and apparatus. In Chapter 5, I show that a standard MALDI protocol, if used in combination with a high-mass detector, allows straightforward mass spectrometric measurements of integral membrane proteins and their complexes, directly following purification in detergent solution. Molecular weights can be determined precisely (mass error ≤ 0.1%) such that high-mass MALDI-MS was able to identify the site for N-linked glycosylation of the eukaryotic multidrug ABC transporter Cdr1p without special purification steps, which is impossible by any other current approach. After chemical cross-linking with glutaraldehyde in the presence of detergent micelles, the subunit stoichiometries of a series of integral membrane protein complexes, including the homomeric PglK and the heteromeric BtuCD, as well as BtuCDF, were unambiguously resolved. This thus adds a valuable tool for biophysical characterization of integral membrane proteins.

Based on the approach developed in Chapter 5, the stoichiometry of integral membrane protein complexes was determined by MAIDI-MS, following chemical cross-linking via glutaraldehyde. However, the structure of glutaraldehyde in solution is not clear due to the formation of oligomeric species. Moreover, glutaraldehyde can react non-specifically with different functional groups of proteins. This further limit its application in structure determination of protein complexes. In Chapter 6, we investigated the capability of a series of chemically much more specific cross-linkers, *N*-hydroxysuccinimide (NHS) esters, to stabilize membrane protein complexes, e.g., PglK and $BtuC_2D_2$. We found that NHS esters could stabilize membrane protein complexes in situ. Stabilization occurred in both detergent systems applied, DDM and C12E8. It was also found that the stabilization efficiency strongly depends on the membrane protein structure, including the primary structure (the number of primary amino groups) and the tertiary structure (distances between two primary amino groups). The number of primary amino groups assures that cross-linkers can attach to the membrane protein; while the distances between two primary amino groups dictate whether the cross-linker with a specific spacer arm length can reach these two targets.

With the established and optimized approach to study membrane protein complexes based on high-mass MALDI-MS, a study of TonB complex was carried out and is presented in Chapter 7. Two partner proteins, ExbB and ExbD, are required for TonB to function, and it has been suggested that TonB, ExbB and ExbD form a complex in the cytoplasmic membrane; understanding the structure of the TonB complex is essential for studying energy transduction. However, the stoichiometry of TonB complex is still missing. After stabilization with glutaraldehyde, the oligomeric structure of ExbB was determined to be predominantly $ExbB_6$. For the ExbB-ExbD complex, two major species, $ExbB_4ExbD_2$ and $ExbB_5ExbD2$, were observed; for the ExbB-ExbD-TonB complex, the major species was either $ExbB_5ExbD_2$ or $ExbB_4ExbD_2$TonB. At the current stage, the origin of the observed lower oligomeric state is hard to define. It could be the result of insufficient reaction or the natural form in the detergent micelle.

In Chapter 8, the approach was further used to understand the oligomeric structure of ABCG2, a breast cancer resistance protein, which was discovered in multi-drug-resistant cancer cells. Since ABCG2 is considered to be a half-transporter, it is believed to at least form a homodimer, or possibly oligomers, in order to function. Although the minimal functional unit of ABCG2 is widely accepted to be the homodimer, higher order oligomers, a tetramer of dimers, has been observed via the cryonegative stain electron microscopy. The dimeric structure of ABCG2 and the existence of disulfide bond(s) between the two subunits was confirmed. Moreover, the ABCG2 prepared in liposome was detected for the first time.

Finally, the thesis ends with Chapter 9, which contains conclusions and an outlook.

Chapter 2. MALDI-MS Detection of Non-covalent Interactions of Single Stranded DNA with Escherichia Coli Single-Stranded DNA-Binding Protein

This chapter is adapted from:

Chen, F., Mädler, S., Weidmann, S., & Zenobi, R. (2012). MALDI-MS detection of noncovalent interactions of single stranded DNA with Escherichia coli single-stranded DNA-binding protein. *J. Mass Spectrom.*, *47*(5), 560–566.

Abstract

The Escherichia coli single-stranded DNA binding protein (SSB) selectively binds single-stranded (ss) DNA and participates in the process of DNA replication, recombination and repair. Different binding modes have previously been observed in SSB•ssDNA complexes, due to the four potential binding sites of SSB. Here, chemical cross-linking, combined with high-mass matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS), is used to determine the stoichiometry of the SSB•ssDNA complex. SSB forms a stable homotetramer in solution, but only the monomeric species (*m/z* 19,100) can be detected with standard MALDI-MS. With chemical cross-linking, the quaternary structure of SSB is conserved and the tetramer (*m/z* 79,500) was observed. We found that ssDNA also functions as a stabilizer to conserve the quaternary structure of SSB, as evidenced by the detection of a SSB•ssDNA complex at *m/z* 94,200 even in the absence of chemical cross-linking. The stability of the SSB•ssDNA complex with MALDI strongly depends on the length and strand of oligonucleotides and the stoichiometry of the SSB•ssDNA complex, which could be attributed to electrostatic interactions that are enhanced in the gas phase. The key factor affecting the stoichiometry of the SSB•ssDNA complex is how ssDNA binds to SSB, rather than the protein-to-DNA ratio. This further suggests that detection of the complex by MALDI is a result of specific binding, and not due to non-specific complexes formed in the MALDI plume.

2.1 Introduction

Specific, non-covalent interactions between proteins and nucleic acids play an important role in numerous biochemical processes, including DNA replication, recombination and repair. Since noncovalent interactions can be maintained, mass spectrometry (MS), using electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI), in theory provides a powerful approach for studying macromolecular complex. Compared with the traditional methods for studying protein-nucleic acid interactions, e.g., electrophoretic mobility shift assays and nitrocellulose filter binding assay, MS requires only small amounts of the sample, on the order of μmol, and can be readily carried out without extensive purification or labeling. By preserving noncovalent interactions during the transition from solution to the gas phase, the mass of the complex directly provides stoichiometric information. ESI-MS, as a major approach, has been applied to study a number of protein-DNA complexes to probe the stoichiometry and the binding affinity of DNA to proteins. Griffey and co-workers showed the binding of a 20-mer phosphorothioate oligonucleotide to bovine serum albumin by ESI-MS. (Greig et al., 1995) The intact 30S subunit of the *E. coli* ribosome containing 21 protein subunits and a 16S RNA molecule has also been detected by using nano-ESI. (Rostom, 2000)

In contrast, noncovalent interactions are more easily disrupted with MALDI, either during sample preparation or ion formation. Moreover, non-specific clusters formed during the ionization step also complicate the unambiguous detection of noncovalent complexes in MALDI. Nevertheless, noncovalent complexes often require the presence of salts, detergent or other stabilization agents, which suppress ESI efficiency. MALDI is more tolerant to the presence of salts and detergents than ESI, rendering MALDI more useful in some cases. In addition, MALDI-MS generates predominately singly charged ions, which greatly simplifies data interpretation for complex mixtures. In order to detect ions with high *m/z* values, a special high-mass detector is required. Chemical cross-linking provides a possibility to chemically stabilize noncovalent interactions before MALDI-MS analyzing. (Sinz, 2006) In our previous studies, MALDI-MS combined

with chemical cross-linking and high-mass detection has been successfully used to study protein complexes, such as immunocomplexes or protein multimers. (Nazabal et al., 2006; Pimenova et al., 2009) Besides the detection of protein complexes stabilized by cross-linking with N-hydroxysuccinimide (NHS) esters, we also observed a heterodimer of the retinoic acid nuclear receptor (RAR) and the retinoid X acid nuclear receptor (RXR) bound to a single DNA strand (DR5) with high-mass MALDI-MS. (Bich et al., 2010) This suggested the possibility to detect protein-DNA complexes with MALDI, although it is still unclear how noncovalent interactions in protein-DNA complexes are preserved in MALDI after introducing an NHS ester, which is highly reactive towards primary amino groups (no reactive group on DNA).

To further explore the capability of high-mass MADLI-MS in protein-DNA complexes determination, the single-stranded DNA-binding protein (SSB) from *Escherichia coli* (*E. coli*), as one of the first discovered and the most thoroughly investigated SSBs, (Lohman & Ferrari, 1994) was studied here. SSBs are essential components in the process of DNA replication, recombination and repair. (Chase & Williams, 1986; Shereda et al., 2008) They associate with single-stranded DNA (ssDNA) with high affinity and in a sequence-independent manner, to remove and/or prevent the formation of secondary structures such as hairpins and cruciforms. *E. coli* SSB forms a stable tetramer in solution, with each subunit comprising an unstructured C-terminus and an oligonucleotide/oligosaccharide-binding (OB) domain on the N-terminus as shown in Figure 1. OB domains bind the phosphodiester backbone and nucleotide bases of ssDNA through electrostatic and base-stacking interactions. Two major binding modes for the ssDNA binding, named $(SSB)_{35}$ and $(SSB)_{65}$, have been observed in electron microscopy (Chrysogelos & Griffith, 1982) and tryptophan-fluorescence quenching (Bujalowski & Lohman, 1986a; Lohman & Overman, 1985). The concentration of monovalent salts (Lohman & Overman, 1985), as well as the protein-to-DNA ratio (Wei, Bujalowski, & Lohman, 1992), are found to be key factors that affect the relative stability of the two binding modes. In the $(SSB)_{35}$ binding mode, on average only two SSB subunits interact with a ssDNA 35-mer; whereas in the $(SSB)_{65}$ binding mode, 65 nucleotides wrap around all four

subunits of the tetramer. High-resolution structural (Figure 1) information for the chymotryptic fragment of *E. coli* SSB in complex with two molecules of $(dC)_{35}$ is also available. [14] Thermodynamic studies [15] indicate that electrostatic interactions play a major role in SSB•ssDNA binding, since the acetylation of lysine residues (Lys 43, Lys 62, Lys73 and Lys 87) and the terminal amine greatly reduced ssDNA binding. The structure also shows that these lysine residues, as well as the terminal amine, are within contact distance of the ssDNA backbone.

10	20	30	40
ASRGVN**K**VIL	VGNLGQDPEV	RYMPNGGAVA	NITLATSESW
RD**K**ATGEM**K**E	QTEWHRVVLF	G**K**LAEVASEY	LR**K**GSQVYIE
GQLRTR**K**WTD	QSGQDRYTTE	VVVNVGGTMQ	MLGGRQGGGA
PAGGNIGGGQ	PQGGWGQPQQ	PQGGNQFSGG	AQSRPQQSAP
AAPSNEPPMD	FDDDIPF		

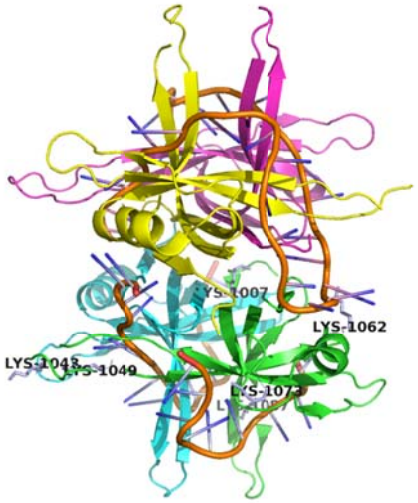


Figure 1. SSB sequence. The DNA binding domain is underlined and lysines are highlighted in bold. Ribbon diagram of the SSB tetramer bound to two 35-mer ssDNA (PDB ID code 1EYG). Lysine in chain A (green) has been labeled.

Here, high-mass MALDI combined with chemical cross-linking was applied to explore the noncovalent interactions in the SSB•ssDNA complex. We will present our results of the detection not only of ionic interactions between ssDNA and SSB, but also of specific noncovalent interactions in the SSB tetramer that were stabilized by cross-linking or ssDNA-wrapping. The results suggest that high-

mass MALDI-MS provides straightforward determination of the SSB•ssDNA complex stoichiometry. Moreover, we investigated the stability of the SSB•ssDNA complex in MALDI is affected when changing the length of the oligonucleotides and the effect of protein-to-DNA ratio on the SSB•ssDNA complex stoichiometry.

2.2 Experimental

2.2.1. Materials

SSB from *E. coli* (2-178, MW=18843.8 Da) was purchased from Connectorate AG (Dietikon, Switzerland) and buffer exchanged against a phosphate buffer (10 mM Na_2HPO_4/NaH_2PO_4, pH=7.2) by using Vivaspin 500 columns (5,000 MWCO PES, Sartorius, Göttingen, Germany). Concanavalin A (Canavalia ensiformis) and aldolase (rabbit muscle) were obtained from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). The protein concentration was determined by UV absorption at 280 nm (NanoDrop 1000, Thermo Fisher Scientific, Wilmington, USA). All the oligonucleotides used below were purchased from Eurogentec (Seraing, Belgium) and no further purification was applied. The oligonucleotide concentrations were determined by UV absorption at 260 nm. The theoretical molecular weight of each poly-d(T) sequence and also the complex formed with the SSB dimer, as well as the SSB tetramer, are listed in Table 1.

Table 1. Theoretical average molecular weights of the poly-d(T)$_n$ studied and theoretical average molecular weights of the SSB•ssDNA complex.

poly d(T)$_n$	Exact MW (Da)	Theoretical average MW of the SSB•ssDNA complexes (Da)			
		SSB_{dimer}•(1)ssDNA	SSB_{dimer}•(2)ssDNA	$SSB_{tetramer}$•(1)ssDNA	$SSB_{tetramer}$•(2)ssDNA
(dT)$_{10}$	2980	40668	43648	78355	81335
(dT)$_{15}$	4501	42189	46690	79876	84377
(dT)$_{25}$	7543	45231	52774	82198	90461
(dT)$_{35}$	10585	48273	58858	85960	96545
(dT)$_{45}$	13627	51315	64942	89002	102629
(dT)$_{55}$	16669	54357	71026	92044	108713
(dT)$_{65}$	19711	57399	77110	95086	114797

The cross-linker disuccinimidy suberate (DSS) was obtained from Pierce Protein Research Products (Thermo Fisher Scientific, Rockford, USA). The matrix sinapinic acid (SA) was purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). Trifluoroacetic acid (TFA) was obtained from Acros Organics (Thermo Fisher Scientific, Geel, Belgium). All commercial solvents and reagents were obtained in the highest available purity and used without further purification.

2.2.2. Chemical Cross-linking

To chemically cross-link SSB, 10 µL of a 2 µM protein solution (monomer concentration) was incubated with 1 µL of DSS dissolved in DMF (100-fold molar excess) for one hour at room temperature. Another 10 µL of SSB solution was used as a control after adding 1 µL of DMF. To investigate protein-DNA cross-linking, 10 µL of the 2 µM protein solution were mixed with 1 µL of 5 µM poly-$d(T)_n$ for 15 minutes at room temperature. The mixture was then incubated with a DSS solution in a 10/1 (v/v, protein/DSS) ratio for one hour at room temperature. A control experiment without adding cross-linker was also carried out. Different concentrations of poly-$d(T)_n$ were mixed with the same protein solution as used before in a volume ratio of 1/10 (v/v, DNA/protein) for different protein-to-DNA ratio required. No further purification steps were carried out.

2.2.3. Mass Spectrometric Detection

A commercial MALDI-TOF/TOF™ mass spectrometer (model 4800 plus, AB Sciex, Darmstadt, Germany) equipped with a high-mass detector (HM2, CovalX AG, Zurich, Switzerland) was used. All measurements were performed in the linear positive ion mode with standard settings. Ionization was achieved with a Nd:YAG laser (355 nm) with the energy just above the threshold for ion formation. Each mass spectrum was the average of 200 laser shots acquired at random sample positions. Sinapinic acid dissolved at 5 mg/mL and 20 mg/mL in water/acetonitrile/TFA (49.95/49.95/0.1, v/v/v) was used as a matrix. The 5 mg/mL sinapinic acid solution was spotted as a matrix-only under layer onto a

stainless steel plate and allowed to dry at ambient conditions. A 0.5 μL sample solution without matrix was then added and a second layer of sinapinic acid solution at 20 mg/mL was spotted on top, after the sample solution had dried. It is difficult to achieve a very high mass accuracy over such a broad mass range, essentially due to the absence of suitable calibrants in the high mass range. Though an internal calibration could improve the reproducibility for mass determination, it will cause mutual ion suppression, especially for the sample such as protein-DNA complexes studied here. A mixture of cytochrome C and bovine serum albumin was used for external calibration. The resulting error mass determination is estimated to be ± 1%. Before data processing, each MALDI mass spectrum was background subtracted and smoothed using a Savitzky-Golay algorithm by Igor Pro 6.2, WaveMetrics, Oregon, USA. Peak integrals were also calculated using this software.

2.3. Results and Discussion

2.3.1. Noncovalent Interactions of ssDNA with SSB

The SSB was investigated first with high mass MALDI-MS combined with cross-linking. A spectrum of SSB before cross-linking is given in Figure 2(a). Only a signal corresponding to the monomer of SSB (*m/z* 19,100, theoretical average MW=18843.8 Da) was detected. The mass error, around + 1%, could come from the calibration error, but also from adduct ions that are not resolved. A peak of the SSB homodimer (*m/z* 38,200) was observed at low intensity, and attributed to the non-specific dimer, which could be the result of clustering occurring in the MALDI plume. (Livadaris & Blais, 2000) After cross-linking, a major peak at *m/z* 79,500 was observed and assigned to the tetramer of SSB (Figure 2(b)). The doubly charged homotetramer was also detected at *m/z* 39,800. The observation of a tetramer instead of the monomer after cross-linking suggests that DSS "freezes" the quaternary structure of SSB in solution and that disruption of the quaternary structural does not take place during transfer into the gas phase.

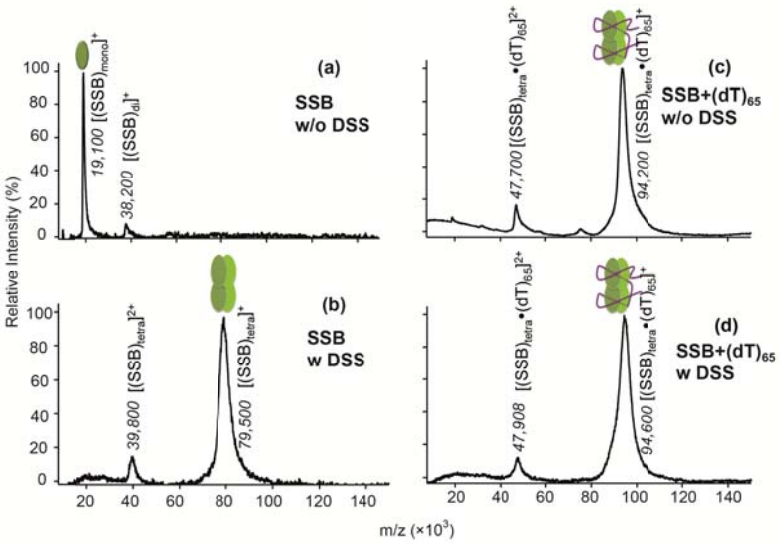


Figure 2. High-mass MALDI mass spectra of SSB (a) without cross-linking and (b) after cross-linking; SSB interacting with $(dT)_{65}$ at a protein-to-DNA molar ratio of 4:1 (c) without cross-linking and (d) after cross-linking.

In the cross-linking step with DSS, the accessibility of nucleophilic residues, mostly the primary amine groups of lysine or the N-terminus, and to a lesser extent hydroxyl groups of serine, threonine, or tyrosine (Kalkhof & Sinz, 2008; Mädler et al., 2009), on the exterior surface of the proteins is critical. The successful application of homobifunctional NHS esters to stabilize the tetramer of SSB suggests that there are sufficient primary amine-containing groups in each subunits to allow cross-linking. The observed mass of the SSB homotetramer is increased by 4125 Da, or around 5%, compared with the theoretical molecular weight of the SSB tetramer without cross-linking (theoretical average MW=75375.2 Da). To stabilize the tetramer, a minimum of three cross-links between the tetramer is required, corresponding to a molecular weight increase of 3×138 Da, which equals 414 Da. The molecular weight increase for each covalently bound DSS (without hydrolysis) is 253 Da. The mass increase after cross-linking suggests that there are at least 21 covalently bound DSS molecules introduced. The accessible number of the primary amine groups is in agreement with the estimated number of DSS molecules introduced based on the molecular weight increase, as a maximum of 25 covalently bound DSS could bind to the six lysine residues and also the N-terminal in each SSB subunit (4 × 6 + 4 – 3 = 25, Figure 1).

In previous studies, it had been found that ssDNA could bind with either two subunits or four subunits of SSB in solution, dubbed $(SSB)_{35}$ or $(SSB)_{65}$, respectively. In the $(SSB)_{35}$ binding mode, the oligonucleotides only interact with two protomers of the tetramer; in the $(SSB)_{65}$ binding mode, the oligonucleotides bind with the whole tetramer. However, noncovalent interactions rarely survive the MALDI process due to the sample preparation conditions and the high laser energy necessary for desorption/ionization of high-mass assemblies. Noncovalently bound complexes of peptides with ssDNA or dsDNA have been detected with MALDI by applying nonacidic matrixes, e.g., 6-azathiothymine in a solution with a pH of 7.0. (Lin & Cotter, 1998; Luo et al., 2004) Here, the matrix is sinapinic acid dissolved in 0.1% TFA, which is a good choice for proteins with molecular weights over 10,000 Da. However, sinapinic acid could disrupt the tertiary structure as only monomeric species of SSB were detected without cross-linking (Figure 2(a)). The noncovalent interactions between SSB and ssDNA were examined with MALDI by applying a ssDNA 65-mer, using a protein-to-DNA molar ratio of 4:1. A major peak, corresponding to the complex of the tetramer of SSB and one $(dT)_{65}$ strand was detected, with a minor peak of doubly charged complex even without applying DSS (Figure 2(c)). A similar result was obtained after chemical cross-linking, as shown in Figure 2(d). The observation of the SSB•ssDNA complex even without cross-linking suggests that $(dT)_{65}$ stabilizes the noncovalent interactions within the SSB tetramer during the sample preparation as well as in the ionization processes. As negative controls, we also incubated $(dT)_{65}$ with concanavalin A (theoretical average MW=25,539 Da, theoretical average $MW_{tetramer}$ =102,156 Da) and aldolase (theoretical average MW=39,420 Da, theoretical average $MW_{tetramer}$ =157,680 Da) using a range of protein-to-DNA ratios. No specific protein-DNA binding is expected to take place in any of these cases. Experimentally, no protein-DNA complexes were ever observed (data not shown). This suggests that the observation of the SSB•ssDNA complex in the absence of cross-linking is based on a specific interaction. The molecular weight of the SSB•ssDNA complex after chemical cross-linking is only slightly higher than that of the complex without chemical cross-linking, by 0.4%. In comparison to the 5% increase observed after cross-

linking of SSB alone, the data suggest limited availability of primary amine groups after ssDNA binding. Our interpretation is that primary amines are involved in the process of ssDNA binding before cross-linking, which is in accordance with results from studies carried out by chemical modification (Chen et al., 1998) and X-ray diffraction (Raghunathan et al., 2000). Moreover, steric hindrance after binding of ssDNA could also lead to a limited availability of lysine residues for reaction with DSS. The observation of the SSB•ssDNA complex suggests the noncovalent protein-DNA complex could also survive in the commonly used sinapinic acid matrix environment. Additionally, the phenomenon can probably be attributed to electrostatic interactions that are strengthened in the gas phase when compared to the solution phase, because of solvent shielding in solution (Loo, 2000; Schug & Lindner, 2005). If only two subunits were connected by the ssDNA as in the $(SSB)_{35}$ binding mode, the noncovalent interactions in the SSB tetramer would hardly be fully protected and the peak corresponding to two subunits of SSB with ssDNA would show a higher intensity. This is indeed observed for $(dT)_{35}$ binding to SSB, which was described below. The high stability of the SSB•ssDNA complex with MALDI might result from $(dT)_{65}$ wrapping around the SSB tetramer.

The $(SSB)_{65}$ mode is favored at high monovalent salt concentrations (>0.2 M NaCl) (Bujalowski & Lohman, 1986b), but also in solutions containing >50 mM Mg^{2+} (Wei et al., 1992). In previous studies, under low-salt conditions, ssDNA tightly binds to the first two subunits, but much more weakly to the second two subunits; increasing the salt concentration results in ssDNA binding to all four subunits. However, the $(SSB)_{65}$ mode was observed here even at a low salt concentration prepared in the solution. This may be caused by the MALDI sample preparation procedure: during crystallization, the salt concentration increases due to solvent evaporation. This salt concentration increase may affect the binding between ssDNA and SSB, which might lead to the observation of the $(SSB)_{65}$ binding mode. There is, however, an alternative explanation. Previous binding constant measurements between SSB and ssDNA were done in the solution phase; electrostatic interactions are enhanced in the gas phase, which

could strengthen the binding between ssDNA and SSB, followed by observation of the $(SSB)_{65}$ binding mode.

2.3.2. Oligonucleotide Length Dependence of SSB•ssDNA Stability

Several ssDNA strands with different lengths were applied, with a protein-to-DNA molar ratio of 4:1, to explore the stability of SSB•ssDNA complex with MALDI. In the mixture of SSB and $(dT)_{10}$, the major peak is assigned to the SSB monomer, without any evidence for association with the DNA as listed in Table 2. Following cross-linking, the high-mass MALDI spectra displayed a major peak assigned to the tetramer of SSB with at least two strands of $(dT)_{10}$ at *m/z* 86,100, because the *m/z* was increased by around 6,700 Da/e compared to the *m/z* of SSB alone after chemical cross-linking. The doubly charged complex was found at *m/z* 43,000. The fact that the complex of SSB and $(dT)_{10}$ was observed only after applying DSS indicates that $(dT)_{10}$ cannot stabilize the quaternary structure of SSB. The $(dT)_{10}$ is probably simply too short to wrap around SSB, let alone the SSB dimer. The observation of the SSB•2$(dT)_{10}$ complex after DSS was introduced suggests that there are still available lysine side chains to form covalent bonds with DSS after $(dT)_{10}$ binding to SSB. Although the observation of the complex requires the assistance of DSS, it is still not enough evidence to suggest that DSS introduces a covalent link between DNA and protein. Our interpretation is that $(dT)_{10}$ might bind to more than one subunit of SSB and the stabilization of the SSB tetramer by DSS could preserve the interactions between $(dT)_{10}$ and SSB multimers. Similar results were observed when $(dT)_{15}$ or $(dT)_{25}$ was used (data not shown).

Table 2. Major species detected of SSB interacting with $(dT)_{10}$ and $(dT)_{35}$ without cross-linking and after cross-linking at a protein-to-DNA molar ratio of 4:1.

poly-$d(T)_n$	Without cross-linking		After Cross-linking	
	$[M]^+$	$[2M]^+$	$[M]^+$	$[M]^{2+}$
$(dT)_{10}$	19,100	38,200	86,100	43,000
	$[(SSB)_{mono}]^+$	$[(SSB)_{di}]^+$	$[(SSB)_{tetra}•2\ (dT)_{10}]^+$	$[(SSB)_{tetra}•2\ (dT)_{10}]^{2+}$
	$[M]^+$	$[M]^{2+}$	$[M]^+$	$[M]^{2+}$
$(dT)_{35}$	95,100	47,900	97,100	49,100
	$[(SSB)_{tetra}•2\ (dT)_{35}]^+$	$[(SSB)_{tetra}•2\ (dT)_{35}]^{2+}$	$[(SSB)_{tetra}•2\ (dT)_{35}]^+$	$[(SSB)_{tetra}•2\ (dT)_{35}]^{2+}$

When increasing the length of oligonucleotides to 35 bases, a peak assigned to the homotetramer of SSB binding two strands of $(dT)_{35}$ at *m/z* 95,100 and another peak at *m/z* 47,900 were detected directly without cross-linking (Table 2.). The observation of the complex without cross-linking suggests that the interaction of $(dT)_{35}$ with SSB can preserve the quaternary structure of the protein to a significant extent. The different stabilities of the SSB•ssDNA complexes in MALDI suggest ssDNA length is important in stabilizing the tertiary structure. Electrostatic interactions are probably responsible for preserving the complex throughout the MALDI process; the strength of electrostatic interactions increases with the length of ssDNA. The integral of the peak at *m/z* 47,900 relative to that of the peak at *m/z* 95,100 was around 28% when $(dT)_{35}$ was applied; while the integral of the peak at *m/z* 49,100 relative to that of the peak at *m/z* 97,100 was only around 9% in $(dT)_{65}$. Besides doubly charged ions, the peak at *m/z* 47,900 could also be attributed to two subunits of SSB binding one strand of $(dT)_{35}$, which would indicate that each $(dT)_{35}$ connects two subunits of SSB. The *m/z* increase of the SSB•ssDNA complex indicates that the cross-linker again has sufficient access to covalently bind to SSB. However, the accessible number of amine groups for DSS reaction after incubating together with $(dT)_{35}$ is much less than the SSB protein alone, as the mass only increase by 2,000 Da. The abundance of SSB tetramer in the presence of two strands of $(dT)_{35}$ vs. one

strand of $(dT)_{65}$ also suggests that the length increase of oligonucleotide will enhance the binding strength between ssDNA and SSB.

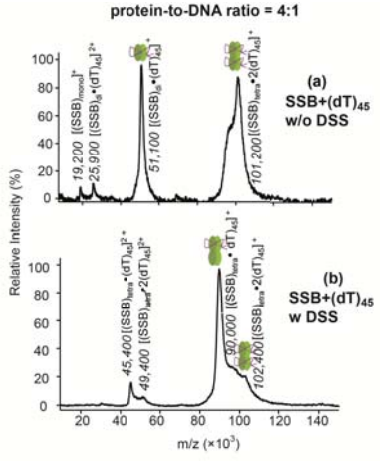


Figure 3. Comparison of MALDI mass spectra of SSB interacting with $(dT)_{45}$ at a protein-to-DNA ratio of 2:1 without cross-linking (a) and after cross-linking (b).

Binding of the 45-mer and 55-mer oligonucleotides to SSB was also explored. The binding pathway of $(dT)_{55}$ to SSB was found to be the same as that of $(dT)_{65}$, i.e. one strand wrapping around the entire SSB tetramer (data not shown). In contrast, for $(dT)_{45}$, both species with one and two strands of oligonucleotides binding to the tetramer of SSB were detected simultaneously after cross-linking, and single strand binding was dominant (Figure 3). Without cross-linking, the two major peaks were assigned to the complexes containing either one oligonucleotide strand and two subunits of SSB (at *m/z* 51,100) or two oligonucleotide strands and the tetramer (at *m/z* 101,200) (Figure 3). The doubly charged of these two binding species were also detected. This indicates that the 45-mer is still not long enough to completely wrap around the tetramer compared with the 55- or 65- mers. The difference of the binding mode before and after cross-linking observed for the $(dT)_{45}$ suggests that the binding of the second oligonucleotide to the SSB tetramer is even weaker than in the case of the 35-mer. Steric hindrance induced by the presence of longer oligonucleotides could perhaps lower the binding with the second oligonucleotide strand. Previous stopped-flow kinetic studies of the $(SSB)_{65}$ binding mode using the oligonucleotide, $(dT)_{35}$, indicate that binding occurs in at least two steps and that

there is an approximately 10-fold difference between the rate constants for binding of the first and second strand of $(dT)_{35}$. (Kozlov & Lohman, 2002) The corresponding dissociation rate constants are expected to be below 0.3 $M^{-1}s^{-1}$, which is around 8 orders of magnitude lower than the association rate constants. When introducing cross-linker, DSS competes with the second DNA to occupy the potential binding sites, which could also weaken the interaction of a second 45-mer with SSB especially in the gas phase.

2.3.3. Molar Ratio Dependence of the SSB•ssDNA Stoichiometry

As presented above, by comparing the results generated before and after cross-linking, we could determine the stoichiometry and subunit assembly of the SSB•ssDNA complex directly. Different protein-to-DNA molar ratios were used to further investigate the binding stoichiometry of the protein-DNA complex generated by MALDI. Each subunit of *E. coli* SSB comprises an unstructured C-terminal domain and an oligonucleotide/oligosaccharide-binding (OB) domain near the N-terminus. To probe the dependence of observed complex stoichiometry on the relative concentration of the binding partners, we decreased the protein-to-DNA molar ratio to from 4:1 to 2:1 and 1:1. The stoichiometry of the SSB and $(dT)_{65}$ complex was explored first when decreasing the protein-to-DNA molar ratio to 2:1.

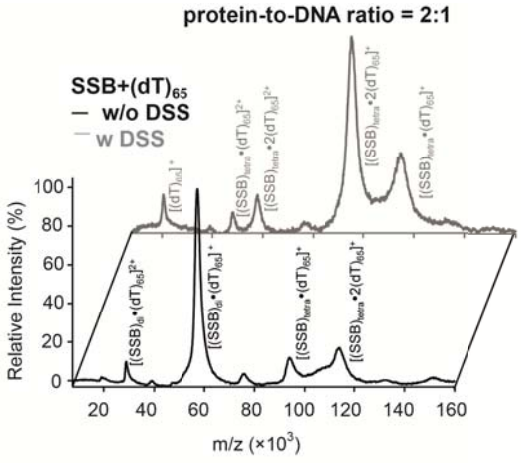


Figure 4. Comparison of MALDI mass spectra of SSB interacting with $(dT)_{65}$ at a protein-to-DNA ratio of 2:1 without cross-linking and after cross-linking (shown in grey).

As shown in Figure 4, the *m/z* value of the dominant peak corresponded to a complex containing two subunits of SSB and one strand of $(dT)_{65}$ when the ratio was 2:1. After adding cross-linker, the peaks at *m/z* 95,600 and 113,400 appear with greatly enhanced intensities. Binding of two strands of $(dT)_{65}$ to the SSB tetramer was not observed when SSB was incubated with $(dT)_{65}$ at a protein-to-DNA ratio of 4:1 (Figure 2(c,d)). This suggests that the protein-to-DNA ratio affects the stoichiometry of the complex, although the same binding mode prevails. Binding of two $(dT)_{65}$ strands to the SSB tetramer weakens the hydrogen bonding in each of the two subunits of SSB. This might explain why the same stoichiometry is followed when $(dT)_{35}$ and a protein-to-DNA ratio of 4:1 is used, although the stability of SSB and two strands of $(dT)_{65}$ is weaker. When decreasing the ratio to 1:1, the preferential stoichiometry included the assembly of two strands of $(dT)_{65}$ with SSB (Figure 5 (b)). However, the stoichiometry was not changed in $(dT)_{35}$ even when the protein-to-DNA ratio dropped to 1:1 (Figure 5 (a)).

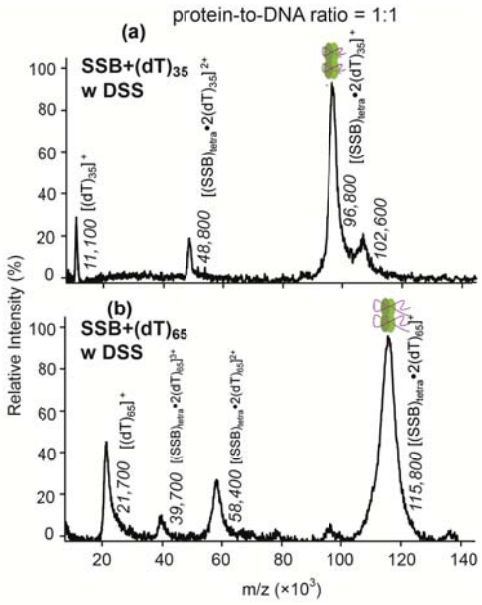


Figure 5. High-mass MALDI mass spectra of SSB interacting with (a) $(dT)_{35}$ and (b) $(dT)_{65}$ at a protein-to-DNA molar ratio of 1:1 after cross-linking. (The mass difference between the same species in Table 2 is likely due to fluctuation in the mass determination in different experiments.)

This suggests that the key factor affecting the stoichiometry of the complex is how ssDNA binds to SSB, rather than the molar ratio. The crystal structure shows that $(dT)_{35}$ can connect the two subunits of SSB, but not penetrate the interface between the two subunits, which is consistent with the stoichiometry found here, as only two strands of $(dT)_{35}$ bind to the SSB tetramer. It further confirms that the stoichiometry observed by high-mass MALDI-MS is due to specific binding rather than non-specific clustering occurring in the MALDI plume. However, it is not possible to guess which lysine binds to ssDNA directly only from the molecular weight increase of the whole complex. Moreover, ssDNA affects the chemical cross-linking, by directly binding to some primary amine groups, and also via steric hindrance leading to a limited availability of lysine residues. Even when the protein-to-DNA ratio was increased to 6:1, the protein-DNA complex, and also the protein monomer was observed like the detection of $(dT)_{35}$ or $(dT)_{65}$ when the ratio down to 1:1 (data not shown).

2.4. Conclusions

By exploring the stoichiometry of various SSB•ssDNA complexes, protein-DNA complexes were observed directly with high-mass MALDI-MS for the first time. The preservation of noncovalently associated SSB dimer or tetramer in MALDI with various lengths of oligonucleotides is compared to covalently, by chemical cross-linking. Combined with the results after chemical cross-linking, the stability of the complex with MALDI was shown, the stoichiometry of how subunits assemble could be determined, and also the accessible number of amine groups could be estimated. The effect of the ssDNA length on the complex stabilization is probably due to electrostatic interactions. In addition, results generated from different protein-to-DNA molar ratios further confirm the stoichiometry results generated from MALDI-MS support the interpretation that specific noncovalently bound complexes are formed, rather than non-specific complexes formed in MALDI plume. Thus, these studies open up new ways for studying protein-DNA interactions.

Chapter 3. Direct Access to Aptamer-Protein Complexes via MALDI-MS

This chapter is adapted from:

Chen, F., Gulbakan, B., & Zenobi, R. (2013) Direct access to aptamer-protein complexes via MALDI-MS. *Chem. Sci.,* 4(10), 4071-4078.

Abstract

We report on the direct detection of protein•aptamer complexes by matrix-assisted laser desorption ionization (MALDI) mass spectrometry (MS). By using optimized conditions, we were able to observe the complexes of thrombin and two different thrombin binding aptamers (TBAs) directly. We also detected the complex of PDGF-AB/BB with the specific PDGF binding aptamer (Apt-35) in a 1:2 stoichiometry. Detection of the complex between lysozyme and its corresponding aptamer further confirmed the capability of MALDI-MS for studying such systems. All these analyses could be performed with very low sample concentrations (1 pmol) and volumes (1-10 μL). Well-designed control experiments confirmed that the complex observation is due to specific non-covalent interactions, rather than non-specific clusters formed in the MALDI plume. The stronger thrombin•TBA29 complex showed a larger signal at the m/z of the intact complex than the weaker thrombin•TBA15 complex; the complex signal of Apt-35 and PDGF-BB was stronger in MALDI compared with that of PDGF-AB and PDGF-AA. These observations indicate that the non-covalent interaction strength in solution is reflected in the MALDI mass spectra.

3.1 Introduction

Protein-nucleic acid interactions are vital for all living organisms, and these interactions control many important biological processes such as the transport and translation of RNA, packaging of DNA, genetic recombination, replication, and DNA repair. (Forterre, 2002) As a notable example, histones interact with DNA to help the formation of chromatin assembly. (Jenuwein & Allis, 2001; Hathaway et al., 2012) Several transcription factors interact with particular DNA loci to mediate protein expression. (Farnham, 2009) In efforts of developing new affinity binders, a new class of oligonucleotide probes called "aptamers" has emerged. (Ellington & Szostak, 1992; Tuerk & Gold, 1990) Aptamers are generated through an iterative process called SELEX (Systematic Evolution of Ligands by EXpotential enrichment) in vitro. They bind to a broad range of targets including metal ions, small organic molecules, peptides, proteins, cancer cells and viruses with high specificity and selectivity. (Bunka & Stockley, 2006) Therefore aptamers rival antibodies in many analytical applications, clinical diagnosis and therapeutic applications. (Jayasena, 1999; Lee et al., 2006) Owing to their unique binding properties, aptamers have been the subject of immense of research in the past 10 years. (Fang & Tan, 2010; Vinkenborg et al., 2011)

To understand the interactions, comprehensive and detailed studies concerning molecular recognition properties, mapping of aptamer and target binding regions, kinetic studies of the induced conformation changes including aptamer and targets, binding stoichiometry and measurement of equilibrium constants are required. Several different analytical methods such as nuclear magnetic resonance (NMR) (Nonin-Lecomte et al., 2001), surface plasmon resonance (SPR) (Balamurugan et al., 2008), X-ray crystallography (Russo Krauss et al., 2011), quartz crystal microbalance (QCM) (Yao et al., 2009), isothermal titration calorimetry (ITC) (P.-H. Lin et al., 2008), circular dichroism (CD) (Johnson et al., 2013), fluorescence correlation spectroscopy (FCS) (Schürer et al., 2001), fluorescence anisotropy (Zhang et al., 2011) were employed to characterize aptamer-protein interactions. However, most of these techniques require a fluorescent label or surface immobilization, which could affect the binding

between an aptamer and its target. In the case of X-ray crystallography, very homogenous and high purity crystals are needed for the analysis. However, the crystal growth process is lengthy, tedious, and often unsuccessful. Therefore only a handful of aptamer-ligand structures have so far been elucidated using x-ray crystallography. (Ruigrok et al., 2012a) Nuclear magnetic resonance provides detailed information, but it requires a lot of material, a relatively long measurement time, and is limited to relatively small structures. Chemical shifts are intrinsically small, which renders structural assignments difficult. (Latham et al., 2009) Therefore, development of a label-free, sensitive, and reliable method for studying aptamer-protein interactions would be of great value.

With the advent of soft ionization methods namely electrospray ionization (ESI) and matrix-assisted laser desorption/ ionization (MALDI), mass spectrometry (MS) has emerged as very valuable tools to study biomolecules and their complexes. (Bich & Zenobi, 2009; Boeri Erba & Zenobi, 2011; Heuvel, 2004; Loo, 2000) These methods enable gentle transfer of biomolecules to the gas phase, and noncovalent interactions can often be preserved using ESI and MALDI processes with proper sample preparation. In stark contrast to other analytical methods, MS offer several advantages for studying biomolecule complexes: only very low sample quantities are required, there is no need for chemical labeling, analyses can be performed in a very short time, and relatively simple spectra are obtained.

ESI-MS has generally been the major approach for studying non-covalent complexes by MS, as it allows direct transfer of solution-phase species to the gas phase. ESI has been applied to investigate a number of protein-DNA complexes to probe the stoichiometry and the binding affinity of DNA to proteins (Greig et al., 1995; Rostom, 2000). Under carefully chosen experimental and instrumental conditions, ESI can be applied to study aptamer-ligand interactions, yielding information on both the stoichiometry of the complex and the binding affinity. (Keller et al., 2005; Ruigrok et al., 2012b) Compared to ESI, noncovalent interactions are more easily disrupted during MALDI analysis, either by the sample preparation or during the ion formation, as samples need to be co-crystallized with a chemical matrix and then irradiated with a strong laser pulse. Nevertheless, noncovalent complexes often require the presence of salts,

detergent additives, or other stabilization agents. MALDI is more tolerant to the presence of salts and detergents than ESI, rendering MALDI the method of choice if such additives are required, for example, in the present case, to keep the aptamer–protein binding intact. Moreover, salts required in the synthesis of aptamers would strongly suppress ESI signals. Finally, MALDI-MS generates predominately singly charged ions, which greatly simplifies data interpretation for complex mixtures.

In previous studies, noncovalently bound complexes of peptides with ssDNA or dsDNA have been detected with MALDI by working with nonacidic matrixes, e.g., 6-azathiothymine, in a solution with a pH of 7.0. (Lin et al., 1998; Luo et al., 2004) In more recent work, we observed noncovalently bound complexes of single-stranded DNA-binding protein (SSB) with single-stranded DNA with high-mass MALDI-MS, even when using an acidic matrix, sinapinic acid. (Chen et al., 2012) In other studies, aptamers were immobilized on a solid surface and used to selectively enrich the target compound(s) to be detected by MALDI-MS. (Cole et al., 2007; Dick & McGown, 2004; Gülbakan et al., 2010; Ocsoy et al., 2013) To the best of our knowledge, direct MALDI mass spectrometric detection of aptamer-protein complexes has never been achieved before. Based on observations made in our laboratory on similar systems, we hypothesized that MALDI-MS could be employed to directly visualize aptamer-protein complexes. We first carried out a series optimization experiments, to identify the most suitable MALDI matrix and sample preparation method. The thrombin binding aptamer (TBA) was then studied. We observed the thrombin-TBA and complex in MALDI-MS directly, using the nonacidic matrix 6-azathiothymine (ATT). We also extended the MALDI-MS study of aptamer-protein interactions to PDGF-AA/-AB/-BB with a PDGF aptamer (Apt-35) and lysozyme with a lysozyme binding aptamer (LBA). The observation of a complex was found to strongly correlate with its binding constant in solution, the laser energy applied, and choice of the proper matrix. This suggests that specific aptamer-protein interactions are preserved in MALDI-MS.

3.2. Experimental

3.2.1. Materials

Human alpha-thrombin was purchased from CellSystems, Biotechnologie Vertrieb GmbH (Troisdorf, Germany), recombinant human PDGF-AA, PDGF-AB, PDGF-BB were purchased from R&D Systems Inc. (Minneapolis, MN) and recombinant human Lysozyme was obtained from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). The protein concentration was determined by UV absorption at 280 nm (NanoDrop 1000, Thermo Fisher Scientific, Wilmington, USA). All the oligonucleotides/aptamers used below were purchased from Eurogentec (Seraing, Belgium) and no further purification was applied (Table 1).

Table 1. Theoretical average molecular weights of the ssDNA (used as a control, see Supporting Information), the ssDNA-aptamers, and the proteins employed in this study.

Name	Sequence	Exact MW (Da)
ssDNA	TCT-GAC-CTT-TGA-CCT-ACT-GAC-CTT-TGA-CCT-CT	9,652.3
LBA	ATC-TAC-GAA-TTC-ATC-AGG-GCT-AAA-GAG-TGC-AGA-GTT-ACT-TAG	12,985
TBA15	GGT-TGG-TGT-GGT-TGG	4,726
TBA29	AGT-CCG-TGG-TAG-GGC-AGG-TTG-GGG-TGA-CT	9,086
Apt-35	CAG-GCT-ACG-GCA-CGT-AGA-GCA- TCA-CCA-TGA-TCC-TG	10,726
Thrombin		36,144
PDGF-AA		28,605
PDGF-AB		26,594
PDGF-BB		24,582
Lysozyme		14,700

The matrices 3-hydroxypicolinic acid (3-HPA) and 2,4,6-trihydroxyacetophenone (THAP) were purchased from ProteaBio Europe (Langlade, France). The other matrixes, sinapinic acid (SA) and 6-azathiothymine (ATT), and ammonium citrate were purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). Trifluoroacetic acid (TFA) was obtained from Acros Organics (Thermo Fisher Scientific, Geel, Belgium). All commercial solvents and reagents were obtained in the highest available purity and used without further purification.

3.2.2. Aptamer-Protein Interactions

Different concentrations of aptamer were incubated with the protein solution in a volume ratio of 1/10 (v/v, DNA/protein) for 1 hour at room temperature. No further purification steps were carried out. The mixture was further diluted before analysis. 3-hydroxypicolinic acid, 2,4,6-trihydroxyacetophenone, and 6-azathiothymine were dissolved at 20 mg/mL in water/acetonitrile/0.1 M ammonium citrate (40/40/10, v/v/v) and sinapinic acid dissolved at 20 mg/mL in water/acetonitrile/TFA (49.95/49.95/0.1, v/v/v). Dried-droplet and modified sandwich sample preparation methods were used respectively. The difference between our modified sandwich preparation and the sandwich preparation reported in 1996 (Li et al., 1996) is we applied the traditional matrix (non-volatile solvent) for matrix-only layer instead of fast-evaporated matrix.

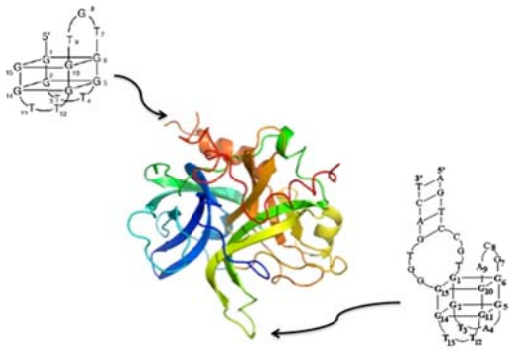
3.2.3. Mass Spectrometric Detection

A commercial MALDI-TOF/TOF™ mass spectrometer (model 4800 plus, AB Sciex, Darmstadt, Germany) equipped with a high-mass detector (HM2, CovalX AG, Zurich, Switzerland) was used. All measurements were performed in the linear positive ion mode with standard settings. Ionization was achieved with a Nd:YAG laser (355 nm) with the proper energy. Each mass spectrum was the average of 1000 laser shots acquired at random sample positions. Before data processing, each MALDI mass spectrum was background subtracted and smoothed using a Savitzky-Golay algorithm by Igor Pro 6.2, WaveMetrics, Oregon, USA.

3.3. Results and Discussion

3.3.1 Detection of Thrombin-Aptamer complexes

Alpha-thrombin is a trypsin-like serine protease that plays a pivotal role in haemostasis considering it is the only enzyme capable of catalyzing the conversion of soluble fibrinogen in insoluble fibrin strands and is the most potent platelet activator. (Huntington, 2005) Thrombin has been the first protein targeted for single-stranded DNA aptamer selection. (Macaya, 1993; Russo Krauss et al., 2011; Schultze et al, 1994) It presents two exosites, 1 and 2, one for the binding of fibrinogen and the other for heparin, which are also the binding sites of 15-mer (TBA15) and the 29-mer (TBA29) as shown in Scheme 1, respectively.



Scheme 1. TBA 15 and TBA 29 binding to thrombin (PDB 1PPB).

The dissociation constant K_d of TBA 29 is around 0.7 nM, much lower than that of TBA15 (up to 450 nM). (Kim et al., 2008) In test experiments with several MALDI matrices (see Supporting Information, SI1), we found ATT as the matrix and the modified sandwich method to provide quite a homogeneous crystallization pattern as well as a good sensitivity for detecting the aptamer. We applied these conditions to study the thrombin•TBA29 complex first. As shown in Figure 1(a), the thrombin•TBA29 complex at m/z 45,000 was detected as the dominant species with a small amount of thrombin at m/z 36,300. The nonspecific heterodimer of thrombin and the thrombin•TBA29 complex and the homodimer of two thrombin•TBA29 complexes were detected at m/z 81,700 and 89,600, respectively. The species at m/z 40,800 could be a complex of thrombin with an impurity in the aptamer solution or the doubly charged nonspecific heterodimer of thrombin and the thrombin•TBA29 complex. The minor species at m/z 53,700 is probably a non-specific adduct of TBA29 with the thrombin•TBA29 complex. Sinapinic acid (SA), which is often used as the matrix of choice for proteins with molecular weights over 10,000 Da was also tried to study aptamer-protein complex. However, in this case, the dominant species was the thrombin itself (Figure 1(b)); in contrast, with ATT, the thrombin•TBA29 complex was well preserved and readily detected (Figure 1(a)). Although electrostatic interactions are enhanced in the gas phase, which could strengthen the binding between thrombin and TBA-29, the fact that the thrombin•TBA29 complex is detected with ATT, but not with SA, suggests that the acidic matrix SA suppresses the non-covalent interactions in the thrombin•TBA29 complex in solution. Our earlier observations of non-covalently bound complexes between single-stranded DNA-

binding protein (SSB) and single-stranded DNA, which survived in the gas phase even when an acidic matrix was used, is believed to be the result of the ssDNA wrapping itself around the SSB tetramer. (Chen et al., 2012) Apparently, maintenance of the thrombin•TBA29 complex depends on the presence of an intact binding pocket of thrombin.

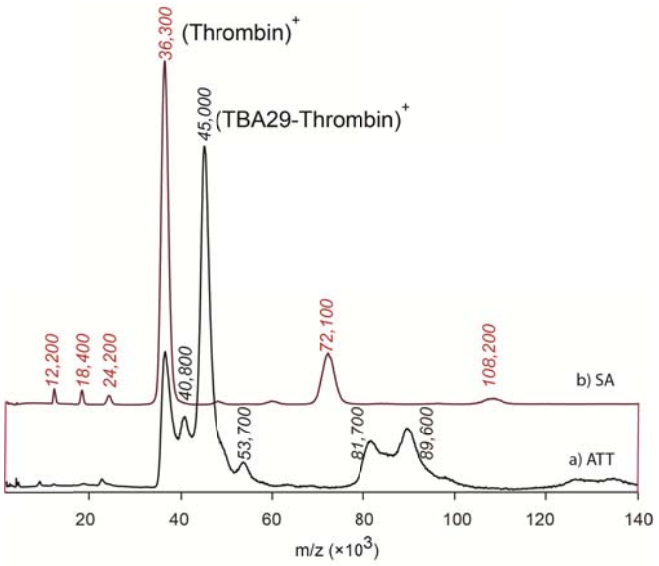


Figure 1. Comparison of high-mass MALDI mass spectra of thrombin interacting with TBA29 at a thrombin-to-TBA29 molar ratio of 1:1 with different matrixes (a) ATT and (b) SA.

The effect of solution pH on aptamer-ligand binding has been addressed in many different studies. Electrochemistry and quartz crystal microbalance measurements on thrombin binding aptamer revealed that optimum binding occurs at pH 7.4; lowering the pH results in poor binding. (Kang et al., 2008) In a similar report (Neves et al., 2010) optimum binding pH of a cocaine aptamer was found to be 7.4, below which binding was significantly compromised. The deteriorating effect of pH on aptamer-ligand binding (Huang & Liu, 2012) was used to develop a biosensor based on graphene oxide. In this study, the pH was changed from 7.5 to 3.5; at lower pH, the aptamer/GO binding was enhanced while aptamer/target binding is weakened. Cowen et al. (Cowan, 2000) also showed that 80% of the binding energy of the neomycin B aptamer was contributed by hydrogen bonding. Since aptamers are strongly negatively charged at physiological pH values, electrostatic interactions and hydrogen bonding play a

dominant role in aptamer binding. All these previous literature findings on similar systems point out the fact that altering the pH significantly changes the strength of the electrostatic interactions and hydrogen bonding, and may partially unfold the aptamer to a thermodynamically less favorable configuration, which leads to poor binding. Therefore the low pH value (3.0) of the sinapinic acid matrix most likely affects the conformation of thrombin aptamer, which could distort the binding pocket that accommodates thrombin. To confirm our reasoning we have slightly acidified the neutral ATT matrix by incorporating 0.1% TFA in the sample. As can easily be seen in SI1, Figure 3, aptamer-protein binding is completely lost after 0.1 % TFA addition, which clearly supports our hypothesis.

To ensure the peaks observed correspond to specific protein-aptamer complexes rather than nonspecific aggregates or clusters, we studied the effects of laser energy and the sample concentration on the complex formation. Non-specific complexes observation in the gas phase could be formed by two processes: either a direct desorption from the solid phase (preformed clusters) or clustering reactions in the plume. (Karas et al., 2000; Livadaris & Blais, 2000) In both processes, increasing the laser energy should favor the formation of larger clusters. Conversely, specific non-covalent complex would be destroyed at higher laser energy and thus decreasing the laser energy should favor observation of the complex. In Figure 2(1), we present the mass spectra recorded at different laser energies, from 0.18 to 0.28 J/cm^2. The intensity of the thrombin•TBA29 complex decreased while the intensity of thrombin itself increased with increasing laser energy. This directly suggests that the observation of the complex results from a specific interaction. Similarly, a higher concentration will favor the formation of larger clusters in MALDI. However, the specific complex would not be strongly affected by the concentration of the analytes. Lowering the concentration of the analyte should eliminate non-specific complexes. We did not find a decrease of the thrombin•TBA29 complex intensity after decreasing the concentration from 10 μM to 2 μM, which confirms that the observation of the thrombin•TBA29 complex is a result of specific non-covalent interactions.

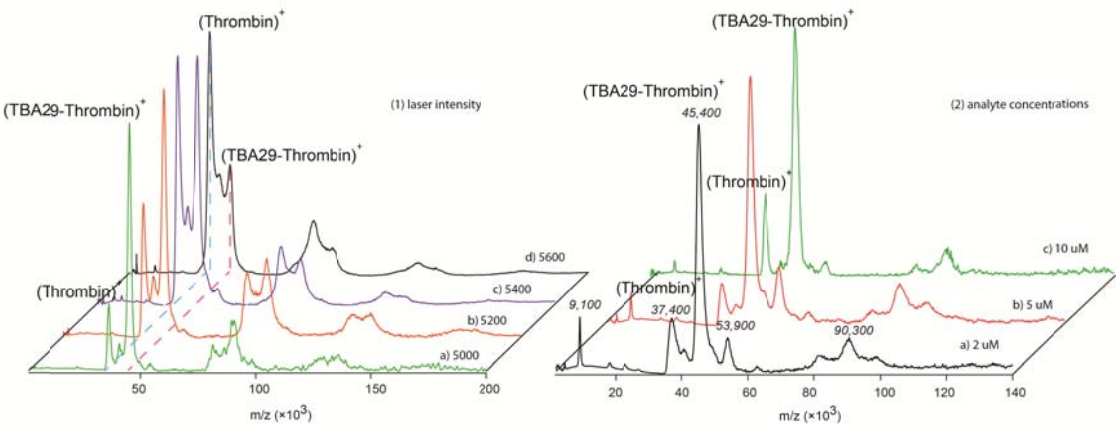


Figure 2. High-mass MALDI mass spectra of thrombin interacting with TBA29 at a thrombin-to-TBA29 molar ratio of 1:1 with different laser energies (1): (a) 5000 a.u. (corresponding to ≈ 0.18 J/cm²), (b) 5200 a.u., (c) 5400 a.u. and (d) 5600 a.u.; and different analyte concentrations (2): (a) 25 μM, (b) 10 μM, (c) 5 μM and (d) 2 μM.

Although the dissociation constants K_d of TBA 29 and TBA15 are strongly dependent on the measurement method used, TBA15 is well known to bind thrombin much more weakly compared to TBA29. (Kim et al., 2008) We used the exact same conditions, i.e., the same protein to aptamer ratio, laser energy and concentration, to analyze binding between thrombin and these two aptamers. As discussed already, at the proper laser energy and molar ratio, the thrombin•TBA29 complex was the dominant species as shown in Figure 1(a); however, the thrombin•TBA15 complex was detected with a much lower intensity, with thrombin being the dominant species (Figure 3). The relative intensity of the thrombin-TBA signal indicates that the noncovalent interaction strength in solution is reflected in the MALDI mass spectra.

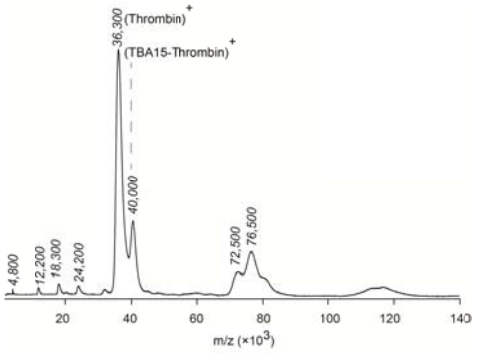


Figure 3. High-mass MALDI mass spectra of thrombin interacting with TBA15 at a thrombin-to-TBA molar ratio of 1:1.

3.3.2 Detection of Aptamer binding to PDGF isomers

To demonstrate the wide applicability of our MALDI based method in studying aptamer-protein complexes, we have chosen a platelet-derived growth factor binding aptamer. Platelet-derived growth factor (PDGF) is a growth factor protein found in human platelets. (Heldin, 1992) It has growth-promoting activity toward fibroblasts, smooth muscle cells, and glial cells. Three disulfide-linked dimers that composed of two homologous chains, A and B, form three known PDGF isoforms, including PDGF-AA, -BB, and –AB. Green et al. reported the aptamer selected from a randomized DNA library could bind to PDGF-AB and PDGF-BB with relatively higher affinity ($K_d \approx 0.1$ nM) and to PDGF-AA with lower affinity ($K_d \approx 80$ nM) (Fredriksson et al., 2002; Green et al., 1996); differently, Fang et al. found the binding difference between PDGF-AB and PDGF-BB could be distinguished via fluorescence quenching assay (Fang et al., 2003). Here, we applied MALDI-MS to study these three isomeric PDGF molecules, PDGF-AA, -AB and –BB, binding with the specific aptamer (Apt-35). In Figure 4(a), we observed Apt-35 at m/z 10,700 and the non-specific homodimer of Apt-35 at m/z 21,400. PDGF-AA, PDGF-AB and PDGF-BB were recorded at m/z of 28,100, 26,600 and 25,200 in Figure 4(b), (c) and (d), respectively. The PDGF-AA, -AB or -BB and Apt-35 were incubated for one hour at 1:1 molar ratio. The dominant species observed after the incubation between PDGF-AA and Apt-35 is Apt-35 at m/z 10,700 (Figure 4(e)). We recorded the complex of PDGF-AB and Apt-35 with the stoichiometry of 2:1 at m/z 47,700 though the dominant species is still Apt-35 (Figure 4(f)). However, the dominant species observed in Figure 4(g) were assigned to the complex of PDGF-BB with Apt-35 (PDGF-BB+2Apt-35) at m/z 46,200. The observation of the complexes involving PDGF-AB and –BB suggests that the aptamer binds to PDGF in a 2:1 fashion, which coincides with a previous report (Fredriksson et al., 2002). The fact that the complex in the MALDI mass spectra is the dominant species only between Apt-35 and PDGF-BB, but neither for PDGF-AB nor PDGF-AA, suggests that PDGF-BB binds to Apt-35 most strongly among the three isomers, and the binding between PDGF-AB and Apt-35 is stronger compared with that of PDGF-AA and Apt-35. This gives further evidence that MALDI-MS can distinguish the three

isomers of PDGF based on not only the different molecular weight, but also the binding strength to Apt-35.

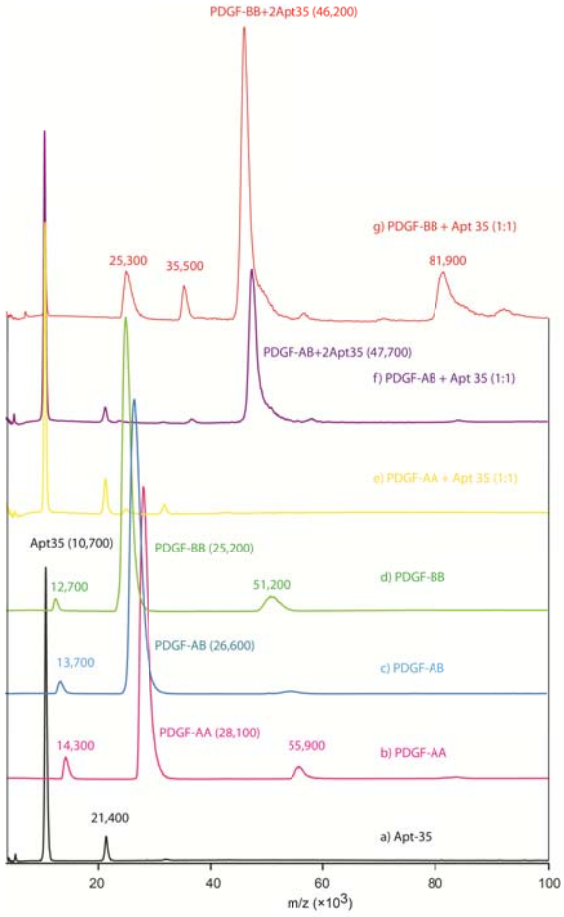


Figure 4. High-mass MALDI mass spectra (a) Apt-35, (b) PDGF-AA, (c) PDGF-AB, (d) PDGF-BB, (e) PDGF-AA reaction with Apt-35, (f) PDGF-AB reaction with Apt-35 and (e) PDGF-BB reaction with Apt-35 at a molar ratio of 1:1.

3.3.3 Detection of Lysozyme-Aptamer complexes

As another example we were also able to detect the complex between lysozyme (Lys) and its aptamer (LBA) by MALDI-MS. (Bamrungsap et al., 2011) Lysozyme is an acid hydrolase, which destructs bacterial cells walls, which can be used as biomarker to diagnose or monitor the treatment efficiency of blood diseases. (McKenzie & White, 1991) It has also been linked to cancer and host resistance of macrophages. (Cox & Ellington, 2001) Being a strongly basic protein, lysozyme is positively charged under physiological pH and is strongly adsorbed onto the channel wall due to the hydrophobic and electrostatic interactions. The binding

affinity of LBA to Lys is quite high, with a dissociation constant K_d of 30 nM. (Kirby et al., 2004) Different molar ratios of Lys to LBA, from 5:1 to 1:5, were studied in Figure 5. Lys (theoretical average MW= 14,700) and LBA have similar molecular weights. As shown in Figure 5(a), the dominant species at m/z 13,000 can be assigned to LBA. The species at m/ 25,800 and 27,700 were assigned to the non-specific homodimer of LBA and the specific heterodimer of Lys•LBA, respectively. By decreasing the molar ratio of LBA to Lys to 2:1 (Figure 5(b)), the Lys•LBA complex at m/z 27,700 became the dominant species. We also observed the signal corresponding to Lys at m/z 14,900. Decreasing the amount of LBA further, the signal intensities for Lys and Lys•LBA became more similar, as seen in Figure 5(c). The signal corresponding to Lys became dominant species when the molar ratio of LBA to Lys was decreased to 1:2 although the Lys•LBA complex could still be observed (Figure 5(d)). However, the Lys•LBA complex could not be detected anymore after further lowering the molar ratio of LBA to Lys to 1:5 (Figure 5(e)).

The observation of Lys•LBA once more confirmed the capability of MALDI-MS to study non-covalent interactions in protein-aptamer complexes. MALDI-MS has shown the potential to investigate noncovalent interactions of protein/peptide-DNA complexes, (Jensen et al., 1993; Juhasz & Biemann, 1994; Tang et al., 1995) with the assistance of cross-linking or in the presence of strong electrostatic interactions between highly acidic DNA and basic peptides. The dependence of thrombin•TBA15/29, PDGF-AB/BB•2Apt-35 and Lys•LBA complex detection in MALDI on the solution-phase binding constant, rather than the isoelectric point of the protein or the acidity of aptamer, suggests that in the case of these aptamer - protein complexes, it is the noncovalent interaction involving the binding pocket that dictates the outcome of the MALDI-MS experiments.

MALDI could be used to rank the affinities of aptamers to their targets but quantification is still quite difficult based on the peak intensity, although introducing an internal standard could help quantification but may suppress the signal and introduce non-specific aggregation. Ionization response in either ESI or MALDI is not a linear function and a so called “response factor” is always needed for calculation. (Gabelica et al., 2003) For aptamer-small molecule complexes whereby the MW of the ligand is relatively very small compared to aptamer,

response factor may be assumed as unity but for larger systems such as thrombin-TBA or lysozyme-LBA where the MW of the interacting species are significantly different, response factors can no longer be assumed as unity. More-

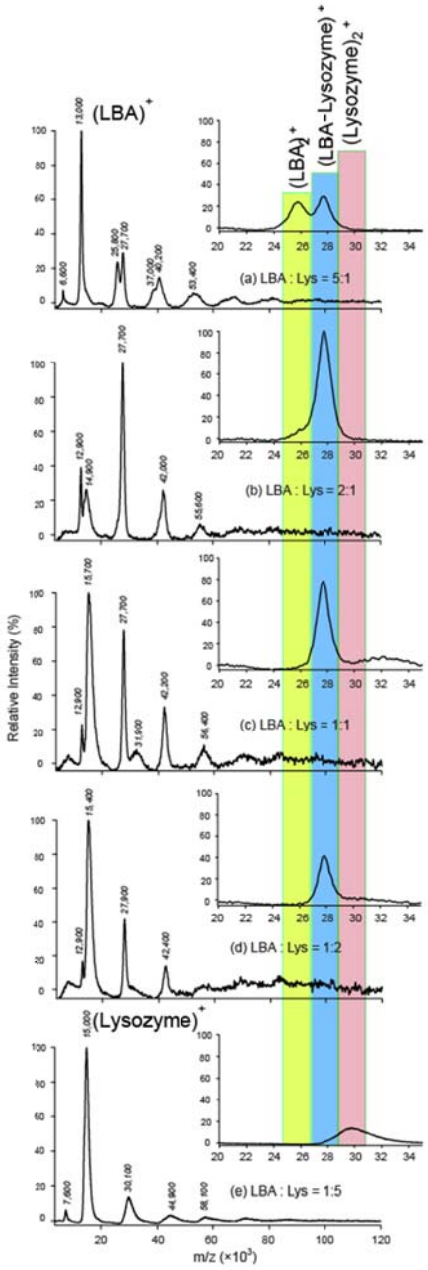


Figure 5. High-mass MALDI mass spectra of lysozyme interacting with LBA at different LBA-to-Lysozyme ratios (a) 5:1, (b) 2:1, (c) 1:1, (d) 1:2 and (e) 1:5.

-over no universal and reliable method yet exists to accurately calculate response factors in MALDI. We have recently addressed the importance of mass discrimination effects in MALDI and how to minimize the effect of response

factors. Even for homomeric protein complexes response factors deviate significantly from unity. (Weidmann et al., 2013b) Therefore an accurate quantitative calculation of binding constants with MALDI-MS is not yet possible and only an affinity ranking can be made.

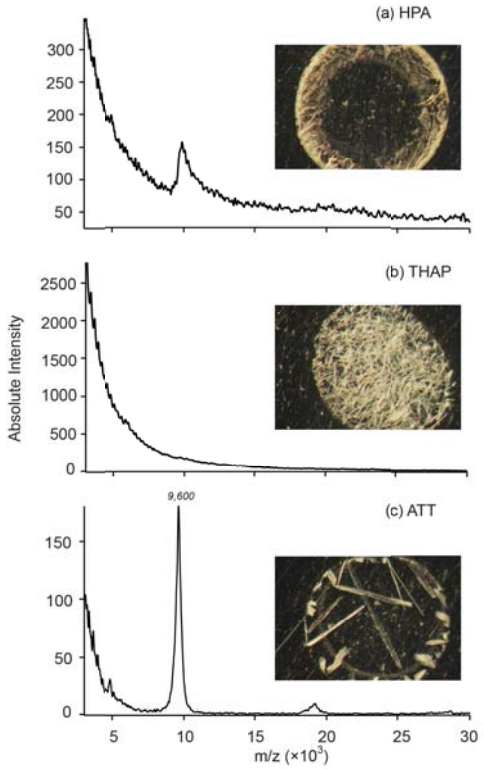
3.4. Conclusions

We report for the first time that high-mass MALDI-MS with a proper matrix, ATT, is able to preserve the non-covalent interactions in protein-aptamer complexes without any chemical cross-linking or other stabilization. This was observed for three systems, thrombin•TBA15/29, PDGF-AA/AB/BB•2Apt-35 and Lysozyme•LBA. Systematic variation of experimental parameters, including the matrix, laser energy and the sample concentration confirmed the observation of the complex to be due to specific non-covalent interactions. The peak intensity of the complexes compared with these of the protein or aptamer subunits in the mass spectra agrees with the solution-phase binding constants, indicating that the non-covalent interaction strength in solution is reflected in the MALDI mass spectra. Thus, these studies open up new mass spectrometry-based avenues for studying protein-aptamer interactions, First, our method provides a fast, specific, sensitive, and single-step assay of aptamer-protein interactions, distinguishing molecular variants of proteins. It provides a complementary structure characterization of the complex without the need for chemical cross-linking. Therapeutic aptamers have become more important in recent years. A typical SELEX process yields several different “aptamer candidates”, i.e., a rapid ranking of their binding affinity is beneficial for post-SELEX applications.

3.5. Supporting Information

Because aptamers are single-stranded DNAs, a single stranded control DNA (ssDNA, TCT-GAC-CTT-TGA-CCT-ACT-GAC-CTT-TGA-CCT-CT, theoretical average MW= 9652.3 Da) was investigated first with high mass MALDI-MS with different matrixes in the positive ion detection mode. Supporting Figure 1 shows mass spectra of 500-fmol ssDNA via dried droplet sample preparation with three different matrixes, including HPA, THAP and ATT. (Sudha & Zenobi, 2002) Only a weak signal corresponding to the ssDNA was detected when HPA was used

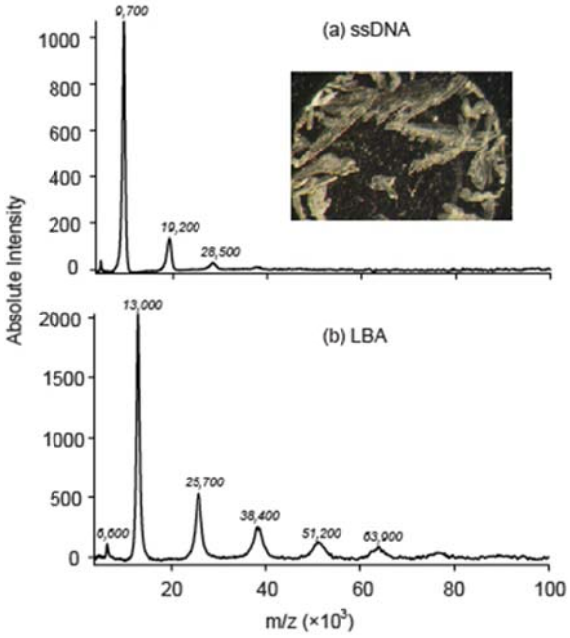
(Supporting Figure 1(a)). In the microscope image, a clear aggregation of high amounts of matrix or analyte crystals was observed in a ring around the edge of the droplet. Although very homogeneous crystals were noticed in the image with THAP as the matrix, we hardly observe any signal corresponding to ssDNA (Supporting Figure 1(b)). ATT with ammonium citrate as the comatrix provides a much better signal of ssDNA as observed in Supporting Figure 1(c). It is well known the presence of alkali metal ions (Na^+ and K^+), which will adduct to the negatively charged phosphate backbone of oligonucleotides, would result in peak broadening and loss of sensitivity. Analyzing ssDNA directly, without any further purification, might limit the sensitivity here.



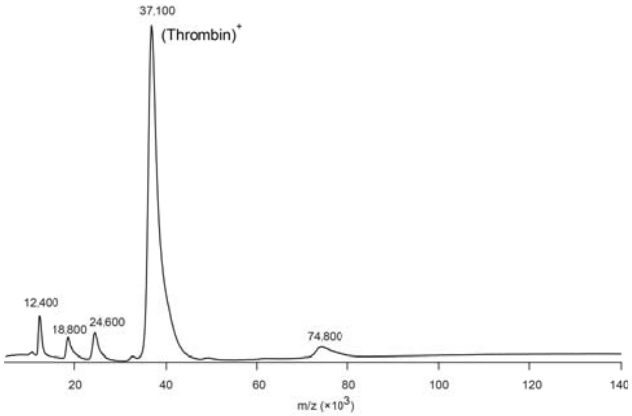
Supporting Figure 1 High-mass MALDI mass spectra of ssDNA with different matrixes (a) HPA, (b) THAP and (c) ATT by using dried droplet preparation.

To enhance the detection of the ssDNA in high-mass MALDI-MS we also optimized the sample preparation. Supporting Figure 2(a) showed the mass spectra of the same amount of the ssDNA as used in Supporting Figure 1 by using modified sandwich method. The signal corresponding to the ssDNA monomer was improved dramatically. Besides the dominant monomeric ssDNA species, the non-specific aggregates, including the ssDNA dimer and the ssDNA trimer that could be the result of clustering occurring in the MALDI plume, and also the doubly

charged species were also detected. A much more homogeneous crystallization pattern, which was observed when modified sandwich sample preparation applied, could explain the signal enhancement. In Supporting Figure 2(b), we recorded the lysozyme binding aptamer (LBA, ATC-TAC-GAA-TTC-ATC-AGG-GCT-AAA-GAG-TGC-AGA-GTT-ACT-TAG, theoretical average MW= 12985.5 Da). A major peak at *m/z* 13,000 was observed and assigned to the monomer of LBA. The peaks of the LBA homodimer (m/z 25,700), the homotrimer (m/z 38,400), the homotetramer (m/z 51,200), and the homopentamer (m/z 63,900) were observed at low intensity, and attributed to the non-specific complexes. The doubly charged LBA was also detected at *m/z* 6,600.



Supporting Figure 2. High-mass MALDI mass spectra of (a) ssDNA and (b) LBA by using the modified sandwich method.



Supporting Figure 3. High-mass MALDI mass spectra of thrombin interacting with TBA29 at a thrombin-to-TBA29 molar ratio of 1:1 with ATT (0.1% TFA) as matrix.

Chapter 4. Understanding Chemical Reactivity for Homo- and Heterobifunctional Protein Cross-linking Agents

This chapter is adapted from:

Chen, F., Nielsen, S., & Zenobi, R. (2013). Understanding chemical reactivity for homo- and heterobifunctional protein cross-linking agents. *J. Mass Spectrom.*, *48*(7), 807–812.

Abstract

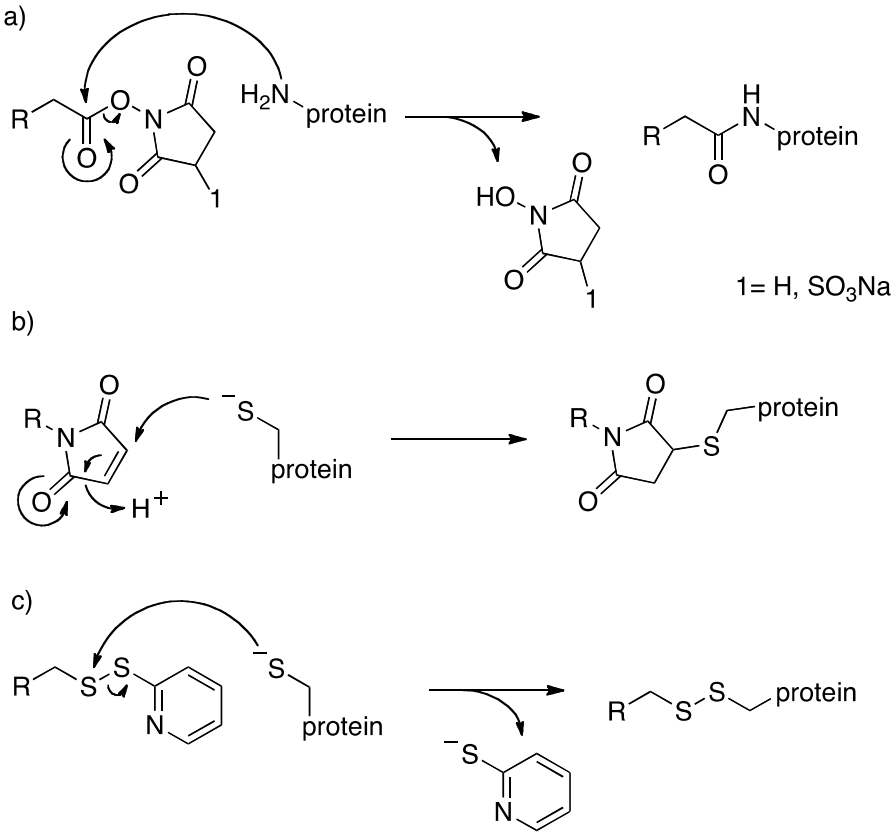
Chemical cross-linking, combined with mass spectrometry, has been applied to map three-dimensional protein structures and protein-protein interactions. Proper choice of the cross-linking agent, including its reactive groups and spacer arm length, is of great importance. However, studies to understand the details of reactivity of the chemical cross-linkers with proteins are quite limited. In this study, we investigated chemical cross-linking from the aspects of the protein stuctures and the cross-linking reagents involved, by using two structurally well-known proteins, glyceraldehyde 3-phosohate dehydrogenase and ribonuclease S. Chemical cross-linking reactivity was compared using a series of homo- and hetero-bifunctional cross-linkers, including bis(sulfosuccinimidyl) suberate, dissuccinimidyl suberate, bis(succinimidyl) penta (ethylene glycol), bis(succinimidyl) nona (ethylene glycol), *m*-maleimidobenzoyl-*N*-hydroxysulfosuccinimide ester, 2-pyridyldithiol-tetraoxaoctatriacontane-*N*-hydrosuccinimide, and succinimidyl-[(*N*-maleimidopropionamido)-tetracosaethyleneglycol]ester. The protein structure itself, especially the distances between target amino acid residues, was found to be a determining factor for the cross-linking efficiency. Moreover, the reactive groups of the chemical cross-linker also play an important role; a higher cross-linking reaction efficincy was found for maleimides compared to 2-pyrimidyldithiols. The reaction between maleimides and sulfhydryl groups is more favorable than that between *N*-hydroxysuccinimide esters and amine groups, although cysteine residues are less abundant in proteins compared to lysine residues.

4.1. Introduction

Analysis of the three-dimensional structure of proteins is of great importance for both protein function and drug design. Mass spectrometry in combination with chemical cross-linking can give rapid information about distance relationships of different side chains, as an alternative approach to high-resolution methods for structural analysis such as X-ray crystallography and NMR spectroscopy. (Leitner et al., 2010; Sinz, 2003; 2006) Besides its application in protein structure determination, chemical cross-linking combined with MALDI-MS and especially high-mass detection, has been demonstrated to be a valuable tool for the detection of non-covalent interactions in biomolecules. (Chen et al., 2012; Nazabal et al., 2006; Pimenova et al., 2009) The primary aim of the cross-linking reaction is to generate a covalent bond between the interaction partners. Cross-linking can occur between two residues within a single strand or between two separate strands of a protein or peptide. For chemical cross-linking, a specific reagent containing at least two functionalities capable of reacting with the protein is required. These cross-linking agents are commercially available with many different reactive groups and many different spacer lengths. (Brunner, 1992; Farmer & Caprioli, 1991; 1998; Melcher, 2004; Singh et al., 2010) The most commonly used cross-linkers are homobifunctional, i.e., they have two identical reactive groups separated by a spacer. (Melcher, 2004) Heterobifunctional cross-linkers, where the spacer separates two different functionalities, are less frequently used, as they pose additional difficulties to data analysis. (Leitner et al., 2010)

Most reagents that are commercially available or reported in the literature are based on a small number of organic reactions. The commonly used chemical cross-linkers include amine-reactive cross-linkers, sulfhydryl-reactive cross-linkers, and photo-reactive cross-linkers. The most frequently utilized cross-linkers target the primary amino groups of lysine and the N-terminus. *N*-hydroxysuccinimide (NHS) esters, which were introduced more than 30 years ago (Bragg & Hou, 1975; Lomant & Fairbanks, 1976) are almost exclusively used (Scheme 1(a)). Several advantages of cross-linking to lysine residues exit, such as high prevalence of lysine residues in proteins (about 6%) and relatively high

reaction specificity of the NHS esters. Under carefully controlled reaction conditions, side reactions of NHS esters with other amino acid residues do not occur at relevant levels. (Mädler et al., 2009) Similar to NHS esters reacting with amine groups, maleimides undergo specific cross-linking reactions with cysteine residues. (Gorin et al., 1966; Gregory, 1955; Smyth et al., 1964) (Scheme 1(b)). Beside maleimides, disulfide reagents, such as the commonly used Ellman's reagent, 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB), also react with sulfhydryl groups *via* disulfide interchange (Scheme 1(c)). (Kimura et al., 1982) Nevertheless, lower abundance of cysteine residues in proteins (lower than 2%) renders them a less attractive for cross-linking reactions.



Scheme 1: Mechanism of the reactions between the cross-linkers and the proteins; a) NHS-ester with lysine, b) *N*-maleimide with cysteine and c) 2- pyridyldithiol with cysteine.

Despite the importance of cysteine and lysine for protein structure, there are only a limited number of studies on heterobifunctional cross-linkers, especially amine-to-sulfhydryl cross-linkers. In this context, the desired products we are interested in are cross-links, rather than loop-links or mono-links, considering only cross-

links can stabilize the protein quaternary structures. Here we investigated the factors that determine the cross-linking efficiency for homo- and heterobifunctional reagents, including protein structure, the reactive group of the cross-linkers used, and their spacer arms. Two proteins with well-known structures, namely glyceraldehyde 3-phosphate dehydrogenase (GAPDH) with a homotetrameric structure and ribonuclease S (RNaseS) with a heterodimeric complex containing an S-protein and an S-peptide, were studied here to address these questions. We present results on the chemical cross-linking of GAPDH and RNAses using a number of different cross-linkers, including homobifunctional cross-linkers bis(sulfosuccinimidyl) suberate (BS^3), dissuccinimidyl suberate (DSS), bis(succinimidyl) penta (ethylene glycol) ($BS(PEG)_5$) and bis(succinimidyl) nonaethyleneglycol ($BS(PEG)_9$) as well as heterobifunctional cross-linkers *m*-maleimidobenzoyl-*N*-hydroxysulfosuccinimide ester (Sulfo-MBS), 2-pyridyldithiol-tetraoxaoctatriacontane-*N*-hydrosuccinimide (PEG_{12}-SPDP) and succinimidyl-[(N-maleimidopropionamido)-tetracosaethyleneglycol] -ester ($SM(PEG)_{24}$). The cross-linking efficiency was evaluated based on the peak intensity of the protein complexes after the reaction. The observations for different chemical cross-linkers suggest that chemical cross-linking efficiency strongly depends on the structure of the protein, specifically on the distance distribution of the protein's target amino acid residues compared to the cross-linker arm length. The nature of the reactive groups of the chemical cross-linkers was also found to have an influence.

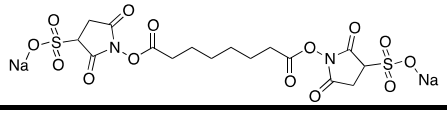
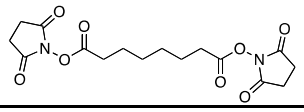
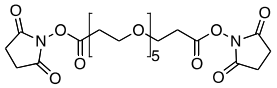
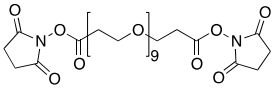
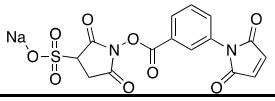
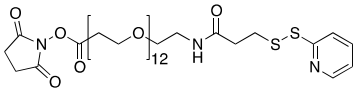
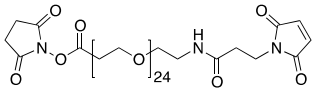
4.2. Experimental

4.2.1. Materials

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH, rabbit muscle), Ribonuclease S (RNaseS, bovine pancreas) and sinapinic acid (SA) were purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). All cross-linkers, bis(sulfosuccinimidyl) suberate (BS^3), dissuccinimidyl suberate (DSS), bis(succinimidyl) penta (ethylene glycol) ($BS(PEG)_5$), bis(succinimidyl) nonaethyleneglycol ($BS(PEG)_9$), *m*-maleimidobenzoyl-*N*-hydroxysulfosuccinimide ester (Sulfo-MBS), 2-pyridyldithiol-tetraoxaoctatriacontane-*N*-hydrosuccinimide (PEG_{12}-SPDP) and succinimidyl-[(N-maleimidopropionamido)-

tetracosaethyleneglycol]ester ($SM(PEG)_{24}$), which are listed in table 1, were obtained from Pierce Protein Research Products (Thermo Fisher Scientific, Rockford, USA). Tris(2-carboxyethyl)phosphine (TCEP) hydrochloride was purchased at ABCR Chemicals (Karlsruhe, Germany). Trifluoroacetic acid (TFA) was obtained from Acros Organics (Thermo Fisher Scientific, Geel, Belgium). All commercial reagents and solvents were obtained in the highest purity available and used without further purification.

Table 1: Name, Structures and distances of spacer arms for used cross-linkers.

Cross-linkers	Structure	Distance of spacer (Å)
BS^3 Bis(sulfosuccinimidyl) suberate		11.4
DSS Dissuccinimidyl suberate		11.4
$BS(PEG)_5$ Bis(succinimidyl) penta (ethylene glycol)		21.7
$BS(PEG)_9$ Bis(succinimidyl) nona (ethylene glycol)		35.8
Sulfo-MBS *m*-Maleimidobenzoyl-*N*-hydroxysulfosuccinimide ester		7.3
PEG_{12}-SPDP 2-Pyridyldithiol-tetraoxaoctatriacontane-*N*-hydrosuccinimide		54.1
$SM(PEG)_{24}$ Succinimidyl-[(*N*-maleimidopropionamido)-tetracosaethyleneglycol]ester		95.2

4.2.2. Chemical Cross-linking

GAPDH and RNaseS were dissolved in phosphate buffer (10 mM Na_2HPO_4/NaH_2PO_4, pH=7.2) to final concentrations of 20 µM. Cross-linker solutions were prepared at a concentration of 20 mM. BS^3 and sulfo-MBS were dissolved in water, while DSS, $BS(PEG)_5$, $BS(PEG)_9$, PEG_{12}-SPDP and $SM(PEG)_{24}$ were dissolved in acetonitrile. To chemically cross-link the proteins, the cross-linker solutions (100-fold molar excess) were mixed with the protein solution in a

1:10 v/v ratio. For RNaseS, reduction was carried out by adding a tris(2-carboxyethyl)phosphine (TCEP) solution (500 µM) to the protein solution in a 1:10 v/v ratio.

4.2.3. Mass Spectrometry

A commercial MALDI-TOF/TOF mass spectrometer (model 4800 plus, AB Sciex, Darmstadt, Germany) equipped with a high-mass detector (HM2, CovalX AG, Zurich, Switzerland) was used. Measurements were performed in positive and negative ion mode with standard settings. Ionization was done with a Nd:YAG laser (355 nm) with the pulse energy just above the threshold for ion formation. Each mass spectrum was the average of 1000 laser shots fired at random sample positions. Sinapinic acid (20 mg/mL in water/acetonitrile/TFA, 49.95/49.95/0.1, v/v/v) was used as matrix. The samples were directly mixed with the matrix in a 1/1 (v/v) ratio. For each sample, 1 µL of the mixture was spotted on a stainless steel plate and allowed to dry under ambient conditions. All mass spectra were baseline corrected and smoothed by utilizing a Savitzky-Golay algorithm by Igor Pro 6.2, WaveMetrics, Oregon, USA. Distance calculations between specific amino acid residues, based on the protein structure from the Protein Data Bank, were carried out using UCSF Chimera (version 1.6.2, University of California, San Francisco, USA).

4.3. Results and Discussion

4.3.1. Chemical Cross-linking of GAPDH

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was investigated with high-mass MALDI-MS combined with different cross-linking reactions. Besides catalyzing the sixth step of glycolysis, GAPDH has also been proposed to play an important role in the initiation of apoptosis. (Tarze et al., 2007) The quaternary structure of GAPDH is a homotetramer as shown in Figure 1. Each monomer contains 26 lysine residues and 3 cysteine residues, as well as an S-nitrosocysteine.

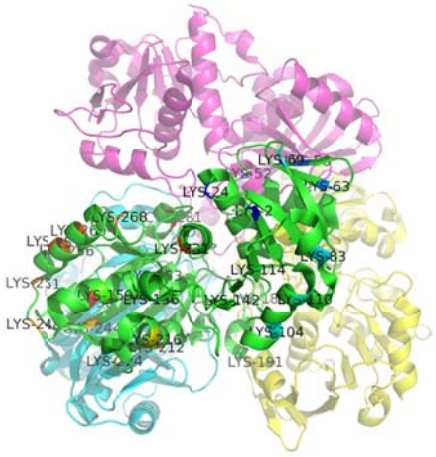


Figure 1. Ribbon diagram of the GAPDH (PDB ID code 1J0X). Lysine and cysteine residues in chain A (green) have been labeled.

The reaction of lysine with NHS esters takes between 10 and 20 minutes to run to completion, (Lindsay, 1972) while the reaction of cysteine with 2-pyridyldithiol needs about 30 minutes. (Carlsson et al., 1978) The reaction of maleimide with cysteine is considered to be quite fast, 1000 times faster than its reaction with amines. (Wong, 1991) Hydrolysis competes with the desired cross-linking reactions. However, hydrolysis of NHS esters has been found to take four to five hours at pH 7.5, (Lomant & Fairbanks, 1976) and also for maleimides it is known to occur at a far lower rate than the reaction with cysteine. (Heitz et al., 1968) An incubation time of two hours was therefore chosen to ensure that the reaction had reached equilibrium.

A spectrum of GAPDH before cross-linking is given in Figure 2(a). The major signal corresponding to the GAPDH monomer is found at m/z 35,900 (theoretical m/z = 35,780, mass error = 0.3 %). Besides the monomeric species, signals of the dimer (m/z 70,300), the trimer (m/z 106,700) and the tetramer (m/z 142,100) were observed at low intensities. These are attributed to non-specific multimers, which are the result of clustering occurring in the MALDI plume. (Livadaris & Blais, 2000)After cross-linking with BS3 (Figure 2(b)), the signal corresponding to the dimer is increased, but the signal corresponding to the tetramer is still not dominant. Similarly, the signal corresponding to the dimer is increased after reacting with DSS (Figure 2(c)). DSS is well known to have an essentially identical

crosslinking reactivity towards primary amines as BS3; however, as opposed to BS3, DSS does not contain sulfonate substituents at either end of the molecule. By using $BS(PEG)_5$, which has an identical reactive group as DSS, the species corresponding to the dimer becomes dominant and the intensity of the tetramer species is also increased (Figure 2(d)). Contrarily, dominant tetramer species are observed after applying $BS(PEG)_9$ (Figure 2(e)). The significant difference of the tetramer intensity for DSS, $BS(PEG)_5$ and $BS(PEG)_9$ cannot be from the functional groups in the chemical cross-linkers nor the number of lysine residues in proteins. This difference must be caused by the polyethyleneglycol spacers.

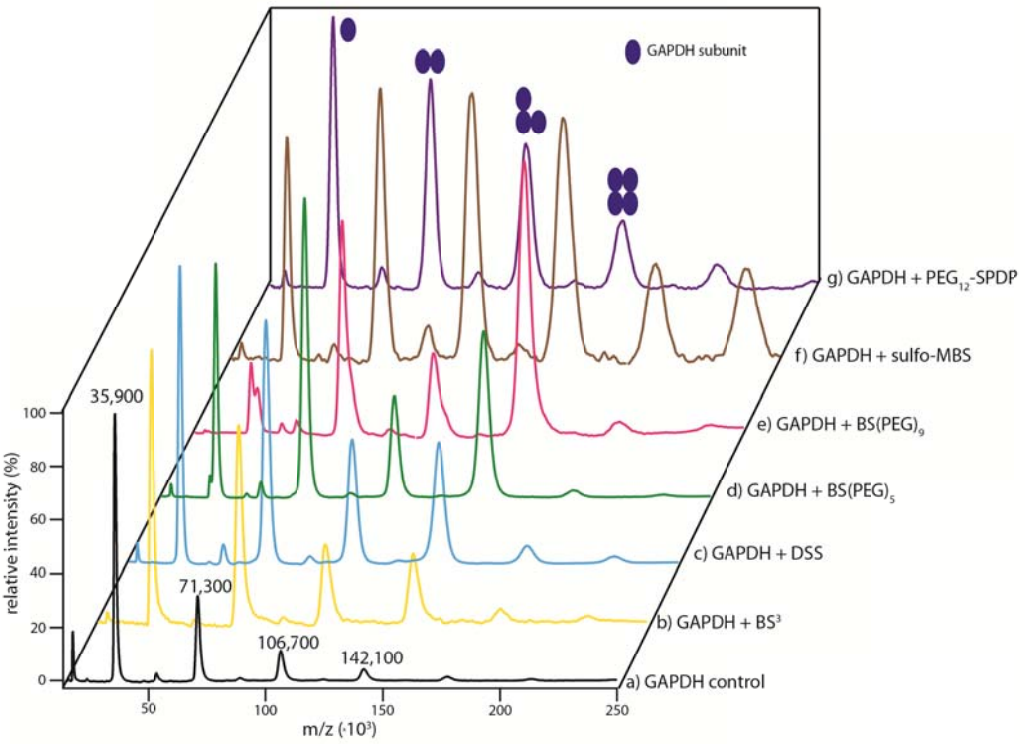


Figure 2. High-mass MALDI mass spectra of GAPDH control (a) and GAPDH with BS^3(b), DSS (c), $BS(PEG)_5$ (d), $BS(PEG)_9$ (c) and GAPDH with sulfo-MBS (e) and PEG_{12}-SPDP (f).

Besides the reaction with homobifunctional NHS esters, heterobifunctional amine-to-sulfhydryl cross-linkers were also explored. After introducing sulfo-MBS (Figure 2(f)), the signals representing dimer, trimer and tetramer show intensities similar to that of the monomer signal. This implies that sulfo-MBS is definitely capable of stabilizing the quaternary structure of GAPDH prior to

MALDI-MS analysis. Signals corresponding to pentamer and hexamer species are also observed. The similar intensities of the different species, including monomer, dimer, trimer and to some extent tetramer, promoted the production of the non-specific pentamer and hexamer species, probably in the plume. PEG_{12}-SPDP, another amine-to-sulfhydryl cross-linker with 2-pyrimidyldithiol groups to target sulhydryls, proved to have cross-linking capabilities comparable to those of BS3, but not as good as sulfo-MBS (Figure 2(f)). The lower stabilizing ability of PEG_{12}-SPDP, compared with sulfo-MBS, might be due to the 2-pyrimidyldithiol group being less reactive with sulfhydryls than maleimide. An alternative explanation is the spacer arm of PEG_{12}-SPDP, which is much longer than that of sulfo-MBS. This affects the capability to stabilize the quaternary structure.

4.3.3. Chemical Cross-linking of RNaseS

Ribonuclease S (RNaseS) was studied in the same fashion as GAPDH. Bovine pancreatic RNaseS is enzymatically cleaved with subtilizing into an S-peptide (residues 1-20, MW 2,166 Da) and an S-protein (residues 21-124, MW 11,541 Da). For its size, RNaseS contains a larger than usual numbers of lysine and cysteine residues; it has 10 lysine residues (8%) and eight cysteine residues (6%). All cysteine residues are paired up in disulfide bonds, giving the protein the required structure to bind single-stranded RNA. Only two of the ten lysine residues are found in the S-peptide, while none of the cysteine residues are located there. (Wyckoff et al., 1967)

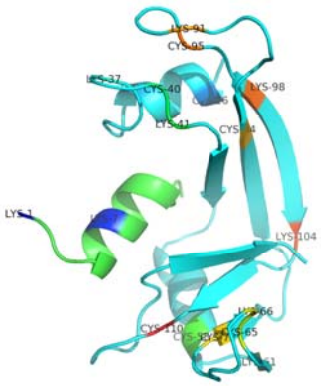


Figure 3. Ribbon diagram of the RNaseS (PDB ID code 1CJR). Lysine and cysteine residues have been labeled.

The major signal corresponding to the singly charged S-protein is observed at *m/*z 11,600 (mass error 0.5 %) (Figure 4(a)). After reaction with BS^3, two species corresponding to the S-protein and the S-protein•S-peptide complex are detected. The S-protein•S-peptide complex is found to be the dominant species (Figure 4 (b)). A similar observation is made when DSS is applied (Figure 4 (c)). However, the intensity of the S-protein•S-peptide complex is decreased dramatically when $BS(PEG)_5$ is introduced (Figure 4(d)). Only a low intensity signal of the complex is observed after cross-linking with $BS(PEG)_5$ or $BS(PEG)_9$ (Figure 4(e)). Again, the significant difference of the S-protein•S-peptide complex intensity for DSS, $BS(PEG)_5$ and $BS(PEG)_9$ must come from the polyethyleneglycol spacers. To conduct sulfhydryl cross-linking, RNaseS needed to be reduced in order to generate free cysteine residues. This was done by addition of tris(2-carboxyethyl)phosphine (TCEP), a frequently used reducing agent. Comparison of the native and reduced states of RNAseS in cross-linking experiments is shown in Supplementary Figure 1. The S-protein•S-peptide complex was observed as the dominant species after treating reduced RNaseS with sulfo-MBS (Figure 4(f)). Interestingly, $SM(PEG)_{24}$, which contains the same functional groups as sulfo-MBS, but has a much longer spacer arm, did not efficiently stabilize the S-protein•S-peptide complex, but only produced monolinks on the S-protein (Figure 4(g)). In both homo- and hetero-bifunctional cross-linkers, limited stabilization efficiency was observed with the longer polyethyleneglycol spacers.

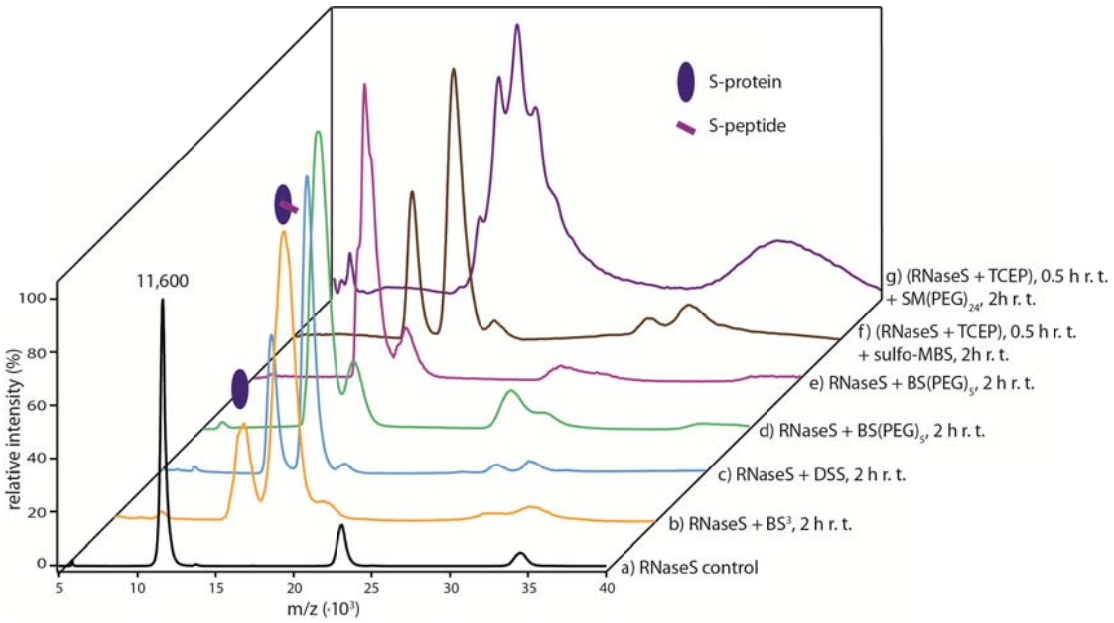


Figure 4. High-mass MALDI mass spectra of RNAseS control (a) and its reaction with BS³(b), DSS (c), BS(PEG)5 (d) and BS(PEG)9 (c). Reduced RNaseS with sulfo-MBS (e) and SM(PEG)$_{24}$ (f).

4.3.3. Factors affecting the Chemical Cross-linking Efficiency

Although the number of cysteine residues in both proteins is lower than that of lysine residues, a high stabilization capability was observed with sulfo-MBS as the cross-linking agent. Interestingly, different stabilization capabilities were observed for GAPDH and RNaseS when the spacer arm length of the cross-linkers with the same functional groups was changed.

To determine the key factors for the stabilization efficiency, we systematically compared the cross-linking efficiency, which was calculated based on Equation 1 (GAPDH) and Equation 2 (RNaseS), where S represents the integral of the signal in both cases. The value from a control measurement without cross-linking is subtracted to eliminate the influence of nonspecific interactions. The cross-linking efficiencies obtained are listed in Table 2.

$$\left(\frac{S_{tetramer}^{\ (cross-linking)}}{S_{monomer}+S_{dimer}+S_{trimer}+S_{tetramer}} - \frac{S_{tetramer}^{\ (control)}}{S_{monomer}+S_{dimer}+S_{trimer}+S_{tetramer}}\right) \cdot 100\%$$

Equation 1: Cross-linking efficiency calculation for GAPDH

$$\left(\frac{S_{complex}^{\ (cross-linking)}}{S_{protein}+S_{complex}} - \frac{S_{complex}^{\ (control)}}{S_{protein}+S_{complex}}\right) \cdot 100\%$$

Equation 2: Cross-linking efficiency calculation for RNAseS.

Table 2. Calculated chemical cross-linking efficiencies based on Equations 1 and 2. (n.d. = not determined)

Protein	BS^3	DSS	$BS(PEG)_5$	$BS(PEG)_9$	Sulfo-MBS	PEG_{12}-SPDP	$SM(PEG)_{24}$
GAPDH	12%	15%	27%	46%	28%	9%	n.d.
RNaseS	79%	72%	23%	24%	73% (Reduced)	n.d.	0%

BS3 and DSS both contain an NHS ester at each end of an eight-carbon spacer arm; BS3 and DSS thus should have essentially identical crosslinking activity toward primary amines. In GAPDH, a cross-linking efficiency of 15 % is observed with DSS, which is slightly higher than BS3 (12 %). A control experiment was carried out by using the same amount of acetonitrile in the reaction of GAPDH with BS3. The cross-linking efficiency remained the same, 12 %. This suggests that the small amount of acetonitrile introduced with the cross-linker does not affect the cross-linking efficiency. BS3 and DSS thus have comparable crosslinking efficiency, which was also observed for RNaseS (79 and 72 %).

The most relevant difference between DSS, $BS(PEG)_5$, and $BS(PEG)_9$ is the spacer length. DSS has a carbon chain with a length of 11.4 Å, while $BS(PEG)_5$ and $BS(PEG)_9$ have PEG chain with lengths of 21.7 Å and 35.8 Å, respectively. For

GAPDH, $BS(PEG)_9$ shows a cross-linking efficiency of 46 %, while DSS only has a cross-linking efficiency of 15 %. In contrast, for RNaseS, a 72 % cross-linking efficiency was found for DSS, but only 24 % for $BS(PEG)_9$. To understand this opposite trend with increasing length of the spacer arms, we calculated the distances between lysine residues in the different subunits in GAPDH, shown in Figure 5 (a). Although the flexibity of the lysine side chain is expected to support the reach of DSS, only 4 % of the lys-lys distances between different subunits of GAPDH were found below 15 Å. On the contrary, about 10 % and 16 % of measured intersubunit lysine-lysine distances are in the range of 20 to 30 Å and 30 to 40 Å, respectively. DSS can simply not reach as many of them; $BS(PEG)_5$ and $BS(PEG)_9$ are much better suited to bridge the lysine-lysine distances in GAPDH. This is our interpretation for the better cross-linking efficiencies for $BS(PEG)_5$ and $BS(PEG)_9$ compared to DSS in stabilizating the GAPDH tetramer. This observation indicates that the efficiency of chemical cross-linking requires a good match between the spacer arm of the cross-linker and the distance distribution of the target residues in the protein structure. In RNaseS, around 20 % of the lysine-lysine distances were found in the region from 6 to 20 Å, whereas only about 6 % are in the 31 to 40 Å range. Instead of an extended PEG linker, a contracted or bent PEG linker is required to reach the second target residue. The lower entropy of a compressed or bent PEG linker will increase the Gibbs free energy of the transition state, i.e., decrease the reaction rate. This is why BS^3/DSS manages to stabilize the S-protein•S-peptide complex better than $BS(PEG)_9$. However, more contacts are found in the region of $BS(PEG)_5$ than DSS in RNaseS. Different from GAPDH, the S-peptide and S-protein in RNaseS have different primary sequences; only two of the lysine residues are located on the S-peptide. The number of inter-subunit contacts between the S-peptide and the S-protein are smaller than the number of intra-subunit contacts. Interestingly, more intra-subunits contacts are found in the range reachable by $BS(PEG)_5$ compared to $BS(PEG)_9$ or DSS (Figure 5 (c)). It is well known that intra-subunit contacts can form loop-links, which will negatively affect the formation of intermolecular cross-links. This explains the better cross-linking efficiency of DSS compared with $BS(PEG)_5$ in stabilizating the complex. These observations support the notion that efficiency of chemical cross-linking requires a good match between the spacer arm of the cross-linker and the

distance distribution of the target residues in the protein structure, including inter- and intra-subunits.

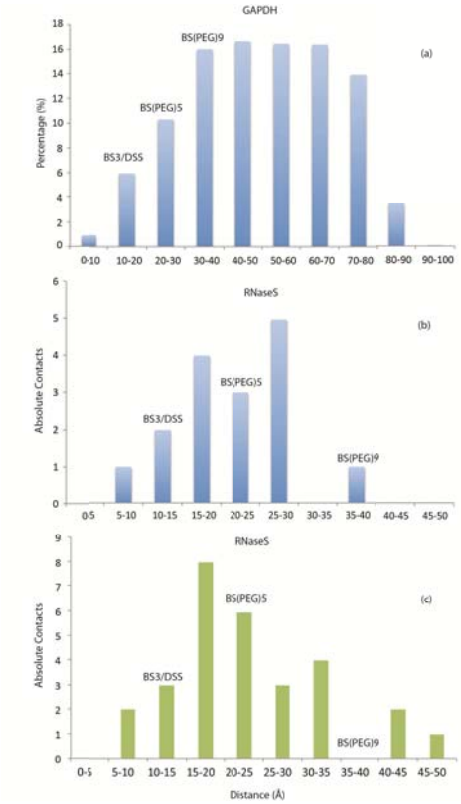


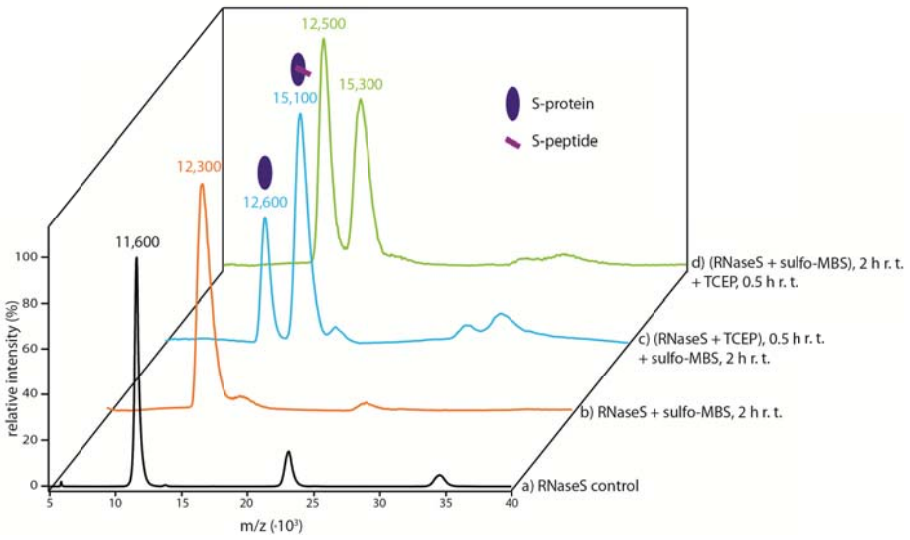
Figure 5. Inter-subunit lysine-lysine distances in the GAPDH tetramers (a) and in the RNaseS-protein•S-peptide complex (b), and intra-subunit lysine-lysine distances in RNaseS (c).

Besides the effect of protein structure, a difference in the reactivity of the functional groups of the utilized cross-linkers exists, too. BS^3 and sulfo-MBS, which have comparable spacer arm lengths, give quite different cross-linking efficiencies in GAPDH. Sulfo-MBS has a cross-linking efficiency around 28 %, while BS^3 shows an efficiency of only 12 %. Moreover, PEG_{12}-SPDP shows 9 % cross-linking, even though it should have a better chance of cross-linking if only the lysine-cysteine distances were considered. This suggests a preference of maleimide targeting cysteine over both NHS esters targeting lysine and 2-pyrimidyldithiols targeting cysteine.

4.4. Conclusions

We conducted chemical cross-linking on two well-known proteins, GAPDH and RNaseS, with a series of homo- and hetero-bifunctional cross-linkers, including BS^3, DSS, $BS(PEG)_5$, $BS(PEG)_9$, sulfo-MBS, PEG_{12}-SPDP and $SM(PEG)_{24}$. The same chemical cross-linker was found to behave quite differently for the two proteins complexes. We compared the cross-linking efficiency among different chemical cross-linkers. The results suggest that the distances between the target amino acid residues, including the inter- and intra-subunits, play a key role in successfully stabilization of a protein complex with a given cross-linker. Additionally, the functional groups on the chemical cross-linkers also affects the reactivity. The *N*-maleimide shows a stronger reaction capability with cysteine than the NHS ester with lysine or 2-pyrimidyldithiol with cysteine. This study provides a better understanding of chemical cross-linking in proteins, especially cross-linking efficiency.

4.5. Supporting Information



Supporting Figure. High-mass MALDI mass spectra of RNAseS control (a) its reaction with sulfo-MBS(b), reduced RNaseS with sulfo-MBS(c) and RNaseS with sulfo-MBS following reduction(d). This result suggests the importance of free cysteines when conducting cross-linking to sulfhydryls.

Chapter 5. High-mass MALDI Mass Spectrometry of Integral Membrane Proteins and their Complexes

This chapter is adapted from:

Chen, F., Gerber, S., Heuser, K., Korkhov, V. M., Lizak, C., Mireku, S., Locher, K. P. & Zenobi, R. (2013). High-mass matrix-assisted laser desorption ionization-mass spectrometry of integral membrane proteins and their complexes. *Anal. Chem.*, *85*(7), 3483–3488.

Abstract

Analyzing purified membrane proteins and membrane protein complexes by mass spectrometry has been notoriously challenging, and required highly specialized buffer conditions, sample preparation methods, and apparatus. Here we show that a standard matrix-assisted laser desorption/ionization (MALDI) protocol, if used in combination with a high-mass detector, allows straightforward mass spectrometric measurements of integral membrane proteins and their complexes, directly following purification in detergent solution. Molecular weights can be determined precisely (mass error ≤ 0.1%) such that high-mass MALDI-MS was able to identify the site for N-linked glycosylation of the eukaryotic multidrug ABC transporter Cdr1p without special purification steps, which is impossible by any other current approach. After chemical cross-linking with glutaraldehyde in the presence of detergent micelles, the subunit stoichiometries of a series of integral membrane protein complexes, including the homomeric PglK and the heteromeric BtuCD, as well as BtuCDF, were unambiguously resolved. This thus adds a valuable tool for biophysical characterization of integral membrane proteins.

5.1. Introduction

Mass spectrometry (MS) has become an indispensable bioanalytical tool in structural studies of soluble proteins. However, investigating membrane proteins, which require detergents for solubilization, by MS is generally a struggle. Only a handful of examples have been described in the literature where integral membrane proteins or membrane protein complexes have been successfully measured by MS. Highly specialized buffer conditions, sample preparation and purification methods, and apparatus are required to accomplish this, i.e., it is far from routine. Recently, Robinson's group detected some integral membrane complexes, e.g., the intact $BtuC_2D_2$ complex, in electrospray ionization (ESI) MS by keeping the nonionic detergent, n-dodecyl-β-D-maltoside (DDM), just above the critical micelle concentration (CMC) (Barrera et al, 2008; 2009); Braun's group revealed the oligomeric structure of ExbB and ExbB-ExbD by laser-induced liquid bead ion desorption (LILBID) MS, a highly specialized technique (Pramanik et al., 2011). Besides, matrix-assisted laser desorption/ionization (MALDI) MS was applied to analyze membrane proteins especially a series condition optimizations including the sample preparation, matrices and detergents for standard membrane proteins, as rhodopsin or bacteriorhodopsin, porins, and also cholesterolesterase (Grey et al., 2009; Keller & Li, 2006; Marty et al., 2012; Rosinke et al., 1995; Schey et al., 1992; Trimpin & Deinzer, 2007; Zhang et al., 2001); by using the "ultrathin layer" sample preparation method, Chait's group determined the molecular weight (MW) of the KcsA potassium channel by MALDI MS (Cadene & Chait, 2000). However, in most of these MALDI studies, only model membrane proteins are examined and relatively low MWs were achieved. The biggest challenge is probably to ionize membrane proteins efficiently in the presence of detergents, and to be able to access high MW membrane proteins. Indeed, a routine MS method to detect integral membrane proteins, to accurately determine their MW, to find post-translational modifications, and define the subunit stoichiometry of complexes would be very helpful to assist high-resolution structural determination and functional analysis, but has so far been lacking.

MALDI is known to be tolerant against the presence of salts, buffers, detergents, and impurities, but has hardly been used to study high-MW membrane proteins and their complexes. We reasoned that with a commercially available high-mass detector, it should be possible to detect membrane proteins, many of which have MWs up to 300 kDa. The detector used here relies on secondary ion production upon impact of large primary ions, and has been shown to be much more sensitive for high-mass ions, e.g., soluble proteins with MWs up to 1 MDa and above, than the widely used microchannel plate detectors, which employ ion-to-electron conversion (Nazabal et al., 2006). Here we report how we succeeded in studying high MW membrane proteins and their complexes by using the traditional dried-droplet sample preparation method and sinapinic acid (SA), a favored matrix for large biomolecules rather than non-acidic matrices, which work well for studying non-covalent assemblies, but perform poorly at very high m/z. This success, based on the high-mass MALDI-MS and good knowledge of handling membrane proteins, is demonstrated to a standard protocol for membrane protein analysis. We further examined the potential of high-mass MALDI-MS for analyzing integral membrane proteins, for pinpointing the glycosylation site of a membrane protein and for defining the subunit stoichiometry of membrane protein complexes, using site-directed mutagenesis, and chemical cross-linking, respectively.

5.2. Experimental Sections

5.2.1. Purification of PglB

PglB from *C. lari* was expressed and purified as described previously (Lizak et al., 2011). The purified protein was desalted into 10 mM MES-NaOH, pH6.5; 100 mM NaCl; 0.5 mM EDTA; 3% glycerol (v/v); 0.016% N-dodecyl-β-D-maltopyranoside (w/v) (DDM, Anatrace) using a HiPrep 26/10 desalting column (GE Healthcare). The proteins variants prepared for the experiments were mutants that differed in their masses: The theoretical average molecular weight of the PglB variant shown in Figure. 1 was 84185 Da, whereas that shown in Figure. 4 was 84277 Da (all the theoretical molecular weights in this paper are average molecular weights).

5.2.2. Purification of PglK

The PglK gene from *C.jejuni* was expressed in *E.coli* BL21(DE3) with a N-terminal decahistidine affinity tag. It was solubilized from cell membranes using n-dodecyl-

・ -D-maltopyranoside (DDM, Anatrace) as a detergent and purified using Ni-NTA affinity chromatography. The buffer was then exchanged to 10mM Tris-HCl pH 8.0 or 10mM Bicine-NaOH pH 8.2, 500mM NaCl, 0.5mM EDTA-NaOH pH 8.0, 10% (w/v) glycerol and 0.016% (w/v) n-dodecyl-・ -D-maltopyranoside using a HiPrep 26/10 desalting column (GE Healthcare).

5.2.3. Purification of BtuCD and BtuCDF

BtuCD was purified as described previously (Locher et al., 2002). In brief, *E. coli* cells expressing BtuCD were disrupted by sonication and BtuCD was solubilized in 1% LDAO. The protein was purified using Ni-NTA following standard procedures, in a final buffer of following composition: 50 mM Na-phosphate, pH 7, 500 mM NaCl, 0.5 mM EDTA, 0.1% LDAO. To purify the BtuCDF complex, the BtuCD in LDAO was mixed with purified BtuF (Korkhov et al., Locher, 2012) in a 1:3 molar ratio and bound to Ni-NTA agarose. Upon detergent exchange (LDAO to $C_{12}E_8$), elution and concentration of the Ni-NTA eluate, the full BtuCDF complex was isolated using preparative size-exclusion chromatography (Superdex 200 10/300 GL), in a buffer containing 10 mM Na-phosphate, pH 7.6, 100 mM NaCl, 0.02% DDM.

5.2.4. Purification of Cdr1p

Wild-type Cdr1p and the mutants N535Q, N724Q, N1303Q and N*all*Q were expressed in *S.cerevisiae* strain FGY217 (Kota et al., 2007) under the control of the GAL promoter. Each construct carried a C-terminal histidine tag. The solubilized proteins were purified using Ni-NTA affinity chromatography and the buffer exchanged to 25mM MES pH 6.2, 100mM NaCl, 10% (v/v) glycerol, 0.5mM EDTA, 0.016% (w/v) DDM (Anagrade) and 0.0032% (w/v) cholesterol hemisuccinate. For deglycosylation, samples were incubated for 2h at 4°C with Endoglycosidase F at a molar ratio of 1:5 (EndoF1 : Cdr1p).

5.2.5. Materials

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH, rabbit muscle), sinapinic acid (SA) matrix and glutaraldehyde (50 wt.% solution in H_2O) were purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). Trifluoroacetic acid (TFA) was obtained from Acros Organics (Thermo Fisher Scientific, Geel, Belgium). All commercial solvents and reagents were obtained in the highest available purity and used without further purification.

5.2.6. Chemical Cross-linking

The protein complexes were mixed with a 10% glutaraldehyde (GA) solution in a 10/1 (v/v) ratio for a certain time at room temperature in buffer solutions containing detergent micelles. A lower concentration of glutaraldehyde, i.e., 1%, was used in Supplementary Figure 1. The mixture was further diluted with the original protein buffer solution or water prior to mass spectrometric analysis.

5.2.7. Mass Spectrometry

A commercial MALDI-TOF/TOF™ mass spectrometer (model 4800 plus, AB Sciex, Darmstadt, Germany) equipped with a high-mass detector (HM2, CovalX AG, Zurich, Switzerland) was used. All measurements were performed in both linear positive and negative ion modes with standard settings. Ionization was achieved with a Nd:YAG laser (355 nm), with the energy just above the threshold for ion formation. Each mass spectrum was the average of 1000 laser shots acquired at random sample positions without searching for hot spot. Sinapinic acid (20 mg/mL in water/acetonitrile/TFA, 49.95/49.95/0.1, v/v/v) was used as the matrix. The samples were directly mixed with the matrix in a 1/2 (v/v) ratio. 1 μL of the mixture was spotted on a stainless steel plate and allowed to dry under ambient conditions. The dried droplet method gives comparable results to other sample preparation methods, i.e., the ultra-thin layer method. A mixture of cytochrome C and bovine serum albumin was used for external calibration. All mass spectra were baseline corrected and smoothed using a Savitzky-Golay algorithm by Igor Pro 6.2, WaveMetrics, Oregon, USA.

5.3. Results and Discussion

5.3.1. Visualization of PglB via High-mass MALDI-MS

We first applied our approach to PglB of *Campylobacter lari*, a bacterial oligosaccharyltransferase (OST) (Lizak et al., 2011). A signal corresponding to the PglB (m/z = 84,300, mass error = 0.1 %) was readily observed. The mass accuracy obtained by high-mass MALDI-MS is hard or impossible to achieve with SDS-PAGE. (Figure 1) The detection of PglB by high-mass MALDI-MS without any buffer exchange (Barrera et al., 2008; 2009) or sophisticated sample handling steps (Bechara et al., 2012; Cadene & Chait, 2000; Keller & Li, 2006; Marty et al., 2012; Trimpin & Deinzer, 2007; N. Zhang et al., 2001) suggests that high-mass MALDI-MS is robust and straightforward, and has excellent potential to become routine for the mass spectrometric investigation of integral membrane proteins. Moreover, the spectrum recorded by the high mass detector is clearly superior to previous published spectra of membrane proteins recorded with standard MALDI detectors in a comparable mass range. The observation of bare PglB without DDM aggregates in the MALDI mass spectrum further highlights the value of this methodology in assisting structure determination of integral membrane proteins.

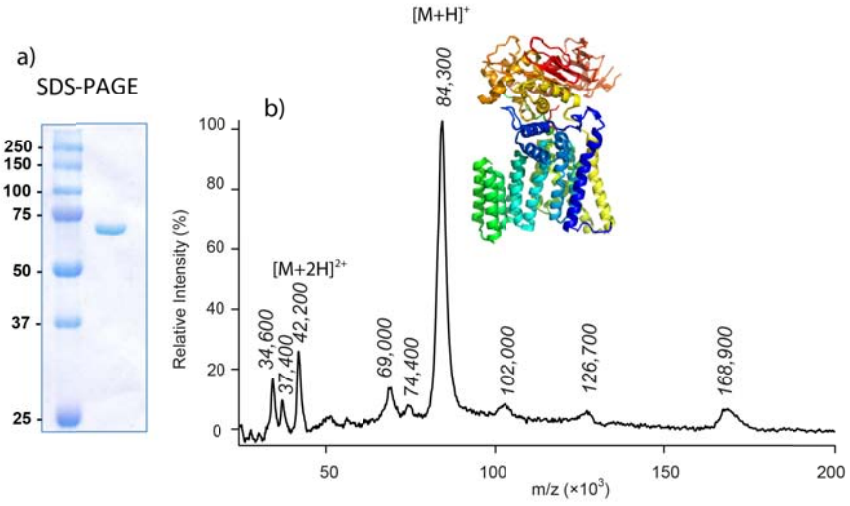


Figure 1. SDS-PAGE and high-mass MALDI mass spectrum of PglB (theoretical mass 84,185 Da). In SDS-PAGE, PglB results in only a band with a molecular weight lower than 75 kDa, which has a mass error around 25 %. MALDI-MS presents the major species of singly charged PglB, with a small intensity of doubly charged PglB at m/z = 42,200. Peaks of the PglB homodimer (m/z = 168,900) and homotrimer (m/z = 254,700), are non-specific

multimers that result from clustering occurring in the MALDI plume (Livadaris & Blais, 2000). The doubly charged PglB trimer was observed at m/z = 126,700. The peaks at low molecular weight, e.g., at m/z = 34,600 and 37,400 (and also 15,200 and 18,700, data not shown) and the peaks close to the PglB, such as at m/z of 69,100 (74, 400) are likely from impurities or N-terminal degradation, which could not be detected by SDS-PAGE. The peak at m/z = 102,000 is probably a non-specific complex between PglB and unassigned species at m/z = 18,700.

5.3.2. Determining the Glycosylation Site of Cdr1p

Motivated by the high mass accuracy achieved, we next studied post-translational modifications of integral membrane proteins that are hard to detect by any other method. Here, *Candida albicans* drug resistance protein 1 (Cdr1p), an ABC transporter containing three potential sites for N-linked glycosylation as shown in Figure 2, was investigated. N-linked glycosylation plays an important role in many cellular processes (Lis & Sharon, 1993; Varki, 1993) and affects the function and stability of proteins (Beretta et al., 2010; Draheim et al., 2010; Wakabayashi-Nakao et al., 2009); nevertheless, different locations of the same N-glycan (macro-heterogeneity) or distinct glycan compositions (micro-heterogeneity) may alter or influence protein function. To date, it is still unclear which, if any, of the potential sites in Cdr1p are glycosylated (Spiro, 2002).

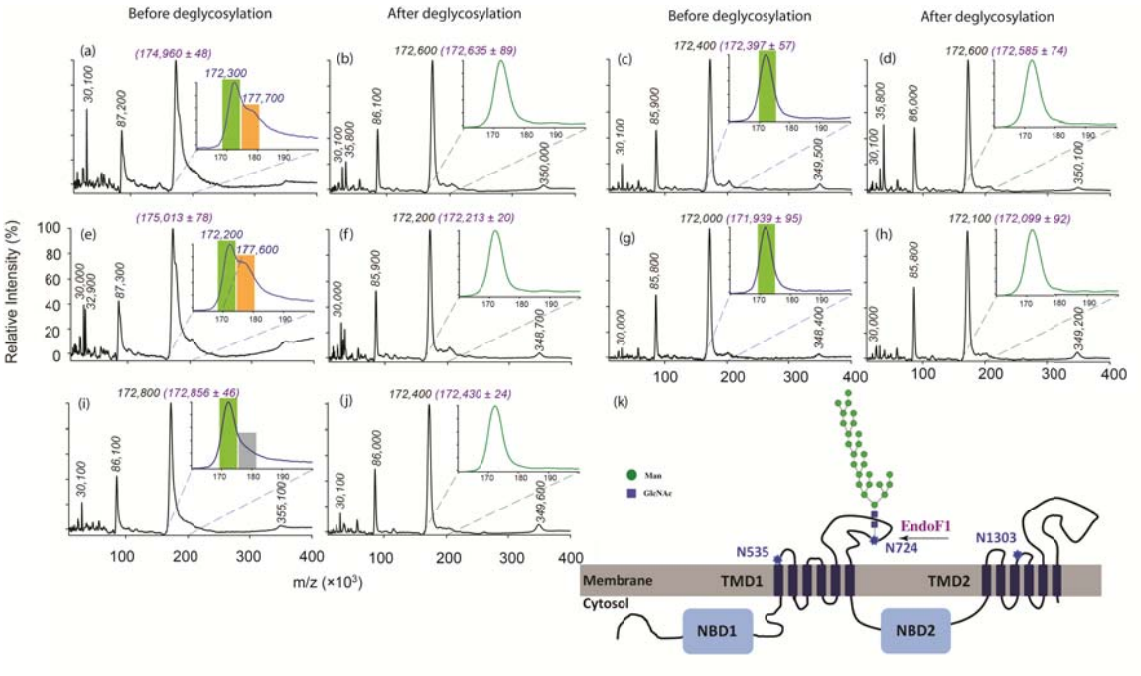


Figure 2. High-mass MALDI mass spectra of wild-type Cdr1p (a) and (b), NallQ (c) and (d), N535Q (e) and (f), N724Q (g) and (h), and N1303Q (i) and (j) before and after deglycosylation by Endo F1. In the lower right, the topology of Cdr1p is shown (k). Error bars correspond to the standard deviation determined from six replicate measurements. Before deglycosylation, peaks at *m/z* = 172,300 and 177,700 were observed simultaneously in wild-type Cdr1p (a); after deglycosylation, only a sharp peak at *m/z* = 172,600 was detected (b). Similar results were also observed for the N535Q mutation, where two peaks at *m/z* = 172,200 and 177,600 before deglycosylation (e) and single peak at *m/z* = 172,200 after glycosylation (f) were observed. Only a single peak was observed for the NallQ mutant ((c) and (d)) and the N724Q mutation ((g) and (h)) before and after deglycosylation. The peak at m/z = 172,800 for the N1303Q mutant before deglycosylation (i) was much broader than that after deglycosylation (j); the second peak is not fully resolved.

We unambiguously determined the site of glycosylation of Cdr1p by comparing the mass of Cdr1p, the wild-type and four different mutants including the full replacement of all three potentially glycosylated asparagines (NallQ) and three individual amino acid replacements (N535Q, N724Q and N1303Q), before and after treatment with Endoglycosidase F1 (EndoF1). Glycosylation of Cdr1p was clearly observed in wild-type and N535Q by the appearance of shoulder to higher mass (Δm ≈ 5.7kDa, consistent with two N-acteylglucosamine and 25-29 mannose moieties). This indicates that N535 in the 1st extracellular loop does not carry any glycans. However, hyperglycosylation is variable in yeast, i.e., the number of saccharide units attached to the acceptor asparagines is not homogenerous.

Therefore, there is not only one peak representing glycosylated Cdr1p, but a broad shoulder peak. The N1303Q mutant only shows a less well-resolved shoulder (Figure 2 (i)), but this is still enough evidence against N1303 carrying another oligosaccharide. Conversely, glycosylation of Cdr1p neither occurred in the NallQ nor in the N724Q mutant. Hence, we believe that N724 in the 3rd extracellular loop carries the oligosaccharide. Functional analysis showed that the N1303Q and the NallQ mutations both decreased the resistance to the antifungal drug fluconazole and accumulated in internal compartments of the cell rather than being trafficked to the plasma membrane (not shown). This indicates that the N1303Q mutation has a negative effect on folding or processing of Cdr1p. There is, however, an alternative explanation: it might also affect N724 in the 3rd extracellular loop from becoming fully glycosylated. The high mannose structure attached to Cdr1p (Shimma et al., 1997), but in particular the heterogeneity in glycosylation, would likely be detrimental for crystallization. The clear determination N724 in the 3rd extracellular loop as the N-linked glycosylation site therefore helps to circumvent problems due to heterogeneous glycosylation in future crystallization studies. The direct analysis of the mutations with high mass MALD-MS without special purification steps (Whitelegge et al., 1998) underscores the capability of high-mass MALDI-MS to identify the glycosylation site in integral membrane proteins.

5.3.3. Defining the Subunit Stoichiometry of Integral Membrane Protein Complexes

Next, we targeted even larger MW membrane protein complexes. A well-known difficulty is that noncovalent interactions are easily disrupted in the MALDI process. We therefore combined chemical cross-linking with high-mass MALDI-MS to study non-covalently bound membrane protein complexes, rather than choosing non-acidic matrices to attempt to maintain non-covalent interactions (Schey et al., 1992). We have previously shown that this approach works well for soluble protein complexes (Nazabal et al., 2006; Pimenova et al., 2009). Glutaraldehyde, one of the most effective protein cross-linking reagents reacting with primary amine groups, (Migneault et al., 2004)was used for chemical stabilization of the noncovalent complexes in the presence of detergent micelles.

Results for three integral membrane protein complexes are shown here, including the glycolipid flippase PglK (Alaimo et al., 2006), the heteromeric vitamin B_{12} importer BtuCD (Locher et al., 2002), and its cognate binding protein BtuF (Hollenstein et al., 2007; Hvorup et al., 2007; Lewinson et al., 2010). PglK is supposed to be homodimer as a structural homolog of Sav1866 (Dawson & Locher, 2006) but without any experimental data for the subunit stoichiometry up to now; the crystal structures of BtuCD and BtuCDF have been recently resolved for the functional heterotetramer ($BtuC_2D_2$) (Locher et al., 2002)and the heteropentameric ($BtuC_2D_2F$) (Hvorup et al., 2007).

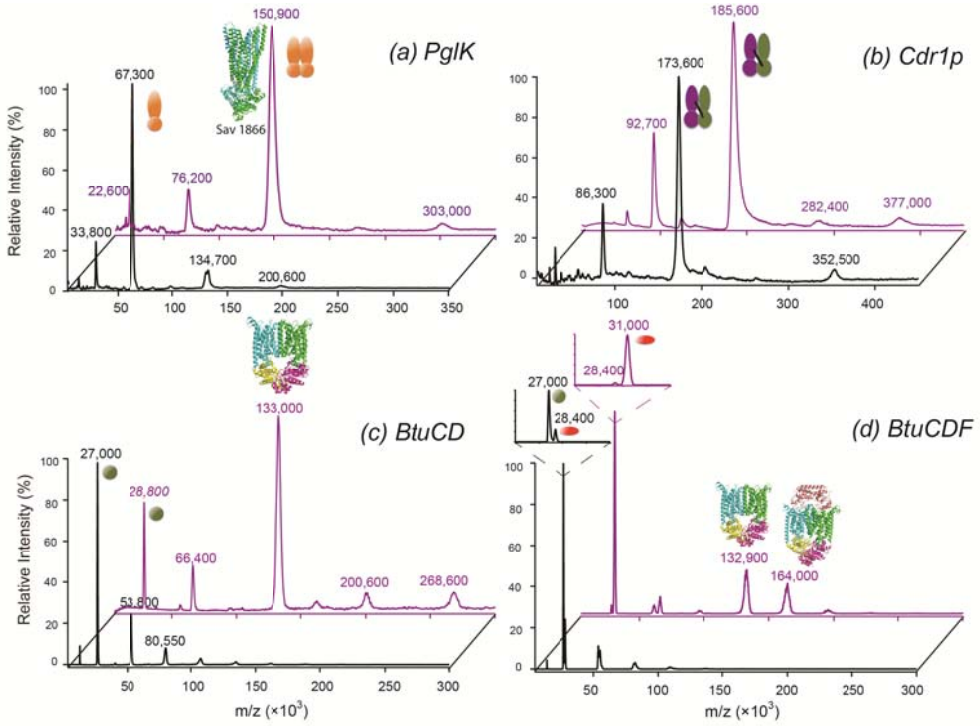


Figure 3. High-mass MALDI mass spectra of (a) PglK and (b) wild-type Cdr1p after deglycosylation in positive mode, (c) BtuCD and (d) BtuCDF in negative mode before (black) and after (purple) reaction with 1 % glutaraldehyde for 1 hour. The ribbon diagram in Fig 3a is that of Sav 1866 (Dawson & Locher, 2006), a structural homologue of PglK.

In Figure 3 (a), a peak corresponding to the monomer of PglK (*m/z* 67,300, theoretical MW=67,327 Da, mass error = 0.04 %) is observed; after cross-linking, a major peak at *m/z* 150,900 was observed and assigned to the PglK homodimer. Stabilization of the non-covalent interactions in PglK before MALDI by reaction with glutaraldehyde easily allows us to first demonstrate the quaternary structure of PglK experimentally to be homodimeric. A control experiment carried

out by incubating Cdr1p, which does not form a complex, with glutaraldehyde (Figure 3 (b)) resulted only in a 6.5 % increase in MW. The intensity of the Cdr1p dimer, which did not increase dramatically after reaction with glutaraldehyde as that observed in PglK, suggests that the Cdr1p dimer come from clustering occurring in the MALDI plume. This shows that the cross-linking chemistry acts on specific complexes. We further recorded BtuCD and BtuCDF in negative ion mode (Figure 3 (c) and (d)), which was recently shown to often generate cleaner spectra than positive ion mode (Mädler et al., 2012). Before cross-linking, only the BtuD (*m/z* 27,000, theoretical MW=27,063 Da, mass error = 0.2 %) subunit, but not the BtuC subunit, was observed, due to the low ionization efficiency for the transmembrane domain. After cross-linking, the BtuCD complex was observed at *m/z* 133,000 (theoretical MW=129,354 Da, Figure 3(c)). Clearly, glutaraldehyde allows stabilizing the integral membrane complexes, not only between nucleotide-binding domains (or transmembrane domains) but also between nucleotide-binding and transmembrane domains. Notably, two species, assigned to $BtuC_2D_2$ (m/z 132,900) and $BtuC_2D_2F$ (m/z 164,000) respectively, were detected after treating the BtuCDF complex with glutaraldehyde (Figure 3(d)). The crystal structure of BtuCDF (Hvorup et al., 2007) suggests the only lysine residue and N-terminus of BtuC to be located in the cytoplasm, close to BtuD but not BtuF. This should directly affect the cross-linking efficiency between the BtuF subunit and the BtuC subunits, which explains the simultaneous observation of both the BtuCD and BtuCDF complexes after adding glutaraldehyde.

We further explored the correlation between the number of amine groups in each subunit, i.e., the lysines and the N-terminus, and the MW increase after reaction with glutaraldehyde (1%, for one hour). If the mass increase introduced by the reaction with glutaraldehyde were accurately known, it should become possible to calculate the MW of a complex without cross-linker, which is critical for determining the stoichiometry of more complicated heteromeric complexes. A relationship exhibiting excellent linearity between the number of amine groups and the MW increase was indeed found (Figure 4), due to polarity of lysine residues located on the protein surface and their reactivity towards

glutaraldehyde. This allows correcting for the glutaraldehyde "decoration", and yields MWs of the bare complexes (0.5% mass accuracy of PglK).

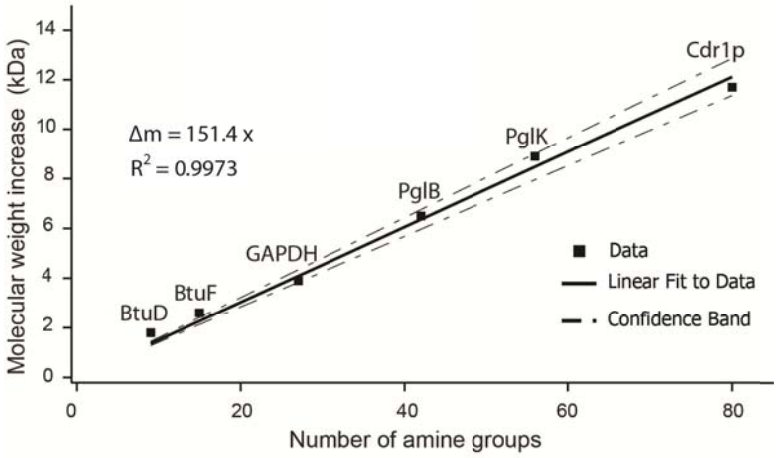


Figure 4. Plot of the molecular weight increase of a series of the protein subunits after one-hour incubation with 1% glutaraldehyde vs. the number of amine groups per subunit (a confidence level of 95% is indicated). The slope is higher than expected if all amine groups were to react with exactly one glutaraldehyde, which indicates that glutaraldehyde oligomerizes in the cross-linking step. As observed for Cdr1p and PglK (Supplementary Figure 1), the oligomerization of glutaraldehyde will not generate any artificial multimers, but only results in a molecular weight increase.

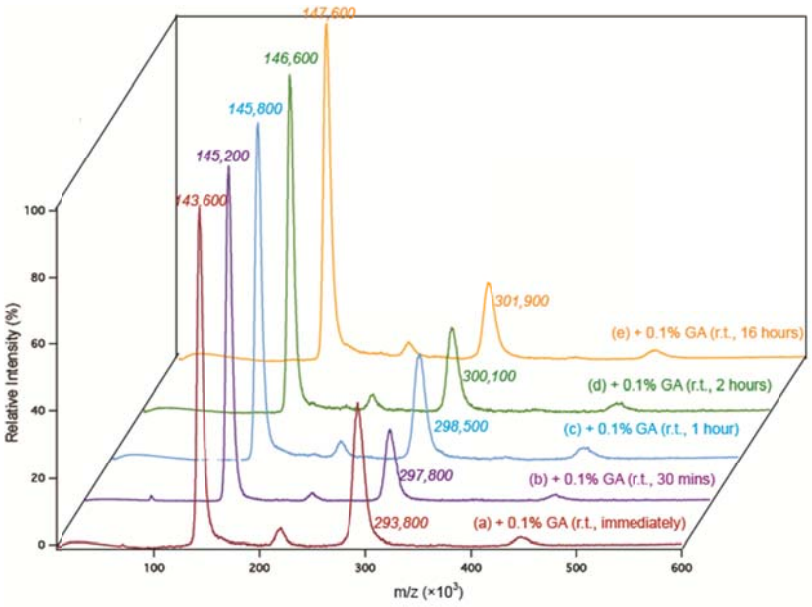
5.4. Conclusions

In summary, we have shown that with standard high-mass MALDI-MS, integral membrane proteins can be detected in a straightforward fashion, and their MW determined with high mass accuracy. This method has excellent potential to become a standard protocol to membrane protein analysis. This high precision allowed the determination of the N-linked glycosylation site in Cdr1p, which is crucial to circumvent problems due to heterogeneous glycosylation in crystallization studies. By introducing glutaraldehyde for chemical cross-linking, we are able to unambiguously define the quaternary structure of integral membrane protein complexes, such as the homodimeric structure of PglK and heteromeric of BtuCD/BtuCDF. We expect MALDI-MS to become a routine method to study integral membrane proteins, the stoichiometry of their complexes, and post-translational modifications, in a straightforward and elegant fashion. Given

the significant number of membrane proteins that are drug targets, our method adds a valuable tool for their study.

5.5. Supporting Information

The reaction time of PglK with glutaraldehyde affects the observed molecular weight of the PglK dimer, which increases from 143,600 to 147,600 with longer reaction time. The observation of dimeric species immediately after incubation of PglK with glutaraldehyde highlights the high reaction efficiency of glutaraldehyde. No higher order oligomeric species were observed at all, even after 16 hours reaction. This suggests that glutaraldehyde only crosslinks specific complexes, and will not generate spurious non-specific complexes.



Supporting Figure. High-mass MALDI mass spectra of PglK reaction with 0.1 % GA in positive mode (a) 0 minute, (b) 30 minutes, (c) 60 minutes, (d) 2 hours and (e) 16 hours.

Chapter 6. Can NHS Esters Stabilize Integral Membrane Protein Complexes in Detergent Micelles?

Abstract

In an earlier study, we presented a straightforward approach based on high-mass matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) to study membrane proteins. Based on this approach, the stoichiometry of integral membrane protein complexes was determined by MAIDI-MS, following chemical cross-linking via glutaraldehyde. However, the structure(s) of glutaraldehyde in solution is not clear due to its polymerization. Moreover, glutaraldehyde can unspecificity react with different targets in proteins. These further limit its application in structure determination for protein complexes. In this study, we investigated the capability of a series of chemically much more specific cross-linkers, *N*-hydroxysuccinimide (NHS) esters, to stabilize membrane protein complexes, e.g., PglK and $BtuC_2D_2$. We found that NHS esters could stabilize membrane protein complexes, even in two different detergent systems, DDM and C12E8. The stabilization efficiency strongly depends on the membrane protein structure, including the primary structure (the number of primary amino groups) and the tertiary structure (distances between two primary amino groups). The number of primary amino groups assures cross-linkers could attach to the membrane protein; and distances between two primary amine groups decide whether the cross-linker with a specific arm length could reach these two targets. This study highlights the possibility to map membrane protein structure via chemical cross-linking, combined with MS.

6.1. Introduction

Mass spectrometry (MS), using electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI), in theory provides a powerful approach for studying macromolecular complexes. However, applying MS in the study of membrane proteins' studies is generally a challenge, because detergents are required for solubilization membrane proteins. Recently, Marcoux and Robinson provide a good review on the progress in studying membrane proteins and their complexes by native ESI-MS in last five years. (Marcoux & Robinson, 2013) The first mass spectrum of a membrane protein complex in detergent micelles recorded by native ESI-MS was the heteromeric vitamin B_{12} importer $BtuC_2D_2$. (Barrera et al., 2008) Laser-induced liquid bead ion desorption (LILBID) MS, a highly specialized technique, was applied to study the oligomeric structure of ExbB and ExbB-ExbD. (Pramanik et al., 2011) We recently presented a straightforward method to study integral membrane proteins based on MALDI-MS, with a high-mass detection capabilities. (Chen et al., 2013a) It allows the direct study of membrane proteins, including determination of their molecular weights, glycosylation sites and the subunit stoichiometries of membrane protein complexes, without further purification or optimization.

Noncovalent interactions are easily disrupted in MALDI, either during sample preparation or ion formation. Applying chemical cross-linkers in advance can effectively stabilize noncovalent interactions chemically before analyzing. (Farmer & Caprioli, 1998; Pimenova et al., 2009) Although glutaraldehyde was confirmed to react with membrane protein complexes (Chen et al., 2013a), the unclear structure(s) of glutaraldehyde in aqueous solution, due to its polymerization, and the unspecific reaction with several functional groups of proteins strongly (Migneault et al., 2004) limit its application in structure determination. For instance, different polymeric forms of glutaraldehyde bring difficulties to map the distances between different side chains, which is the core of three-dimensional structure analysis based on chemical cross-linking combined with MS.

N-hydroxysuccinimide (NHS) esters are among most widely applied chemical cross-linkers. They have been used to analyze the three-dimensional structure of proteins, and have several advantages in the structural analysis of soluble proteins: the high popularity of lysine residues in proteins (about 6%) and relatively high reaction specificity. Under carefully controlled reaction conditions, side reactions of NHS esters with other amino acid residues do not occur at relevant levels. (Mädler et al., 2009) Cross-linking protocols, mass spectrometric analysis of cross-linked samples, and also the further data analysis have been well established, as described in some recent reviews (Kalkhof & Sinz, 2008; Müller & Sinz, 2012; Sinz, 2003; 2006). Applying NHS esters for the structural analysis of membrane proteins has, unfortunately, been neglected so far. Before applying tandem mass spectrometry (top-down approach) or in-solution digestion (bottom-up approach) for structure determination, it is critical to establish that NHS esters can react with membrane proteins (or their complexes) effectively in the presence of detergent micelles.

To answer the question whether NHS esters can effectively react with membrane proteins in detergent micelles, we applied a series of NHS esters two membrane protein complexes, specifically two ATP binding cassette (ABC) transporters, (Locher, 2004; 2009) PglK and $BtuC_2D_2$. We will approach this question, NHS-esters cross-linking on integral membrane protein complexes, from two main directions: the properties of the cross-linkers and the structure of the membrane proteins. Applying four chemical cross-linkers, including bis(sulfosuccinimidyl) suberate (BS^3), disuccinimidyl suberate (DSS), bis(succinimidyl) penta(ethylene glycol) ($BS(PEG)_5$) and bis(succinimidyl) nonaethyleneglycol ($BS(PEG)_9$), (SI3, Table 1) we found that NHS esters are capable of stabilizing membrane protein complexes. The stabilization efficiency strongly depends on the protein structure, including the primary and tertiary structure. Succeed in cross-linking PglK in DDM, as well as C12E8, highlights the capability of applying NHS esters to cross-link membrane protein complexes in different detergent micelles.

6.2. Materials and Methods

6.2.1. Materials

PglK and $BtuC_2D_2$ were purified as described previously. (Chen et al., 2013a; Locher et al., 2002) After buffer exchange, the PglK was in a buffer of 10mM Bicine-NaOH, at pH 8.2, containing 500mM NaCl, 0.5mM EDTA-NaOH, 10% (w/v) glycerol, and 0.016% (w/v) n-dodecyl-· -D-maltopyranoside (DDM, Anatrace). Dodecyl octaethylene glycol ether (C12E8) was used as another detergent to solubilize PglK. The BtuCD protein was in a final buffer with the following composition: 10 mM Na-phosphate, pH 7.6, 100 mM NaCl, 0.02% DDM. Sinapinic acid (SA) was purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland). All chemical cross-linkers were purchased from Pierce Protein Research Products (Thermo Fisher Scientific, Rockford, USA). Trifluoroacetic acid (TFA) was obtained from Acros Organics (Thermo Fisher Scientific, Geel, Belgium). All commercial reagents and solvents were obtained in the highest purity available and used without further purification.

6.2.2. Chemical Cross-linking Protocol

Cross-linker solutions were prepared at a concentration 1000 times higher than that of protein complexes. BS^3 was dissolved in water and the other cross-linkers used here were dissolved in acetonitrile. To chemically cross-link the proteins, the cross-linker solutions were mixed with the protein solution in a 1:10 volume ratio at room temperature for 2 hours. The mixture was further diluted with the original protein buffer solution or water prior to mass spectrometric analysis.

6.2.3. Mass Spectrometry

A commercial MALDI-TOF/TOF mass spectrometer (model 4800 plus, AB Sciex, Darmstadt, Germany) equipped with a high-mass detector (HM2, CovalX AG, Zurich, Switzerland) was used. All measurements were performed in linear positive ion mode with standard settings. Ionization was achieved with a Nd:YAG laser (355 nm) with the pulse energy set just above the threshold for ion formation. Each mass spectrum was the average of 1000 laser shots acquired at random sample positions. Sinapinic acid (20 mg/mL in water/acetonitrile/TFA,

49.95/49.95/0.1, v/v/v) was used as the matrix. The samples were directly mixed with the matrix solution in a 1/2 (v/v) ratio. 1 μ L of the mixture was spotted on to a stainless steel plate and allowed to dry under ambient conditions. All mass spectra were baseline corrected and smoothed using a Savitzky-Golay algorithm available within Igor Pro (Version 6.2, WaveMetrics, Oregon, USA). Distance calculations between specific amino acid residues, based on the protein structure from the Protein Data Bank, were carried out using UCSF Chimera (version 1.6.2, University of California, San Francisco, USA).

6.3. Results and Discussion

NHS esters, which are used widely to map the three-dimensional structure of proteins, i.e., soluble proteins, were applied here to study their capabilities to cross-link membrane protein (complexes) in detergent micelles. The structure of PglK (Alaimo et al., 2006), a Campylobacter jejuni-encoded ABC transporter, is still not well understood; it is supposed to be a structural homolog of Sav1866, a well-defined homodimeric multidrug ABC transporter consisting of 12 transmembrane helices (Dawson & Locher, 2006a). In our earlier study, the oligomeric state of PglK was characterized experimentally for the first time via high-mass MALDI-MS with glutaraldehyde. (Chen et al., 2013a) The monomeric species of PglK at m/z 67,500 (theoretical mass 67,327 Da, mass error = 0.2 %) was detected as the dominant peak in the absence of any chemical cross-linking (Figure 1 (a)). After cross-linking with BS^3 (Figure 1 (b)), the signal corresponding to the dimer becomes the dominant species with a m/z 144,500. The observed molecular weight here is bigger than the theoretical mass the dimer of PglK (134654 Da), which should come from the BS^3 “decoration” as also observed previously (Chen et al., 2013a). The observation of a homodimer instead of the monomer after cross-linking suggests that PglK forms a non-covalent dimer in solution and these non-covalent interactions can be preserved in MALDI after cross-linking with BS^3. Moreover, Dominant dimer species are also observed after adding $BS(PEG)_5$ (Figure 1 (c)) and $BS(PEG)_9$ (Figure 1 (d)), respectively.

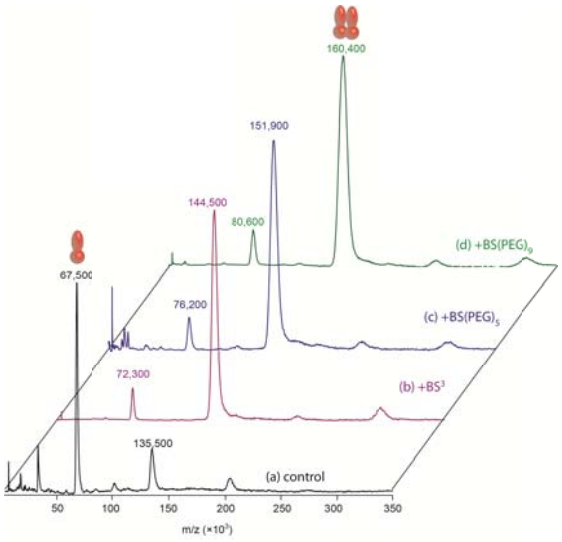


Figure 1. High-mass MALDI mass spectra of PglK in DDM a) without any cross-linker and with b) BS^3, c) $BS(PEG)_5$ and d) $BS(PEG)_9$, respectively.

To further study chemical cross-linking in membrane protein complex, we carried out the cross-linking of PglK in another detergent micelles, C12E8. DDM (dodecyl-β-D-maltoside) and C12E8 (octaethyleneglycol monododecyl ether), which are usually used to solubilize and stabilize membrane proteins, both are nonionic detergents. The observation of the monomeric species of PglK and a low intensity of PglK dimer in C12E8 suggests that high-mass MALDI-MS is a membrane protein friendly approach (Figure 2 (a)). Again, we observed the dimer of PglK as the dominant species after chemical crosslinking with, BS^3 (Figure 2 (b)), $BS(PEG)_5$, (Figure 2 (d)) and $BS(PEG)_9$ (Figure 2 (e)). Moreover, applying DSS to PglK in C12E8 also provided the dominant species with a m/z 146,400. (Figure 2 (c)) The success of cross-linking PglK in different detergent micelles, either DDM or C12E8, suggested that NHS esters could stabilize membrane protein complexes in detergent micelles.

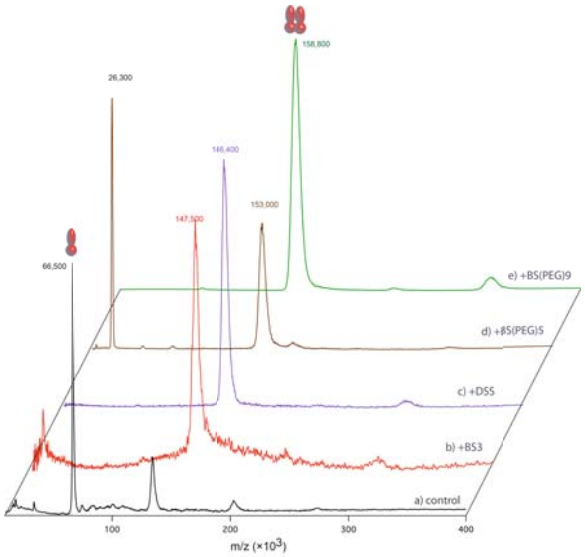


Figure 2. High-mass MALDI mass spectra of PglK in C12E8 a) without any cross-linker and with b) BS^3, c) DSS, d) $BS(PEG)_5$ and e) $BS(PEG)_9$, respectively.

Following, we studied NHS esters cross-linking on $BtuC_2D_2$, a vitamin B_{12} importer from *Escherichia Coli* (Locher et al., 2002). In the absence of any cross-linker, we observed the BtuD subunit at m/z 26,800 as the dominant species, and the BtuC subunit at m/z 37,200 with a much low intensity. (Figure 3 (a)) Cross-linking of $BtuC_2D_2$ with BS^3 or DSS did not yield any peak corresponding to the complex of $BtuC_2D_2$ (Figure 3(b) and (c)). The molecular weight corresponding to the BtuD subunit increased in both cross-linking reagents, suggesting that the cross-linker did react with amine group in the BtuD subunit. After adding $BS(PEG)_5$ (Figure 3 (d)), we observed a higher intensity of the BtuD dimer peak at m/z 57,000 and of the BtuC monomer at m/z 38,600. Moreover, complex species, including the species at m/z 67,100 (BtuCD), 95,100 ($BtuCD_2$), 133,700 ($BtuC_2D_2$), can be seen in Figure 3 (d). Similarly, after introducing $BS(PEG)_9$ (Figure 3 (e)) the $BtuC_2D_2$ complex of BtuCD was stabilized as seen from observing the species at m/z 68,800 (BtuCD), 97,700 ($BtuCD_2$), and 136,800 ($BtuC_2D_2$).

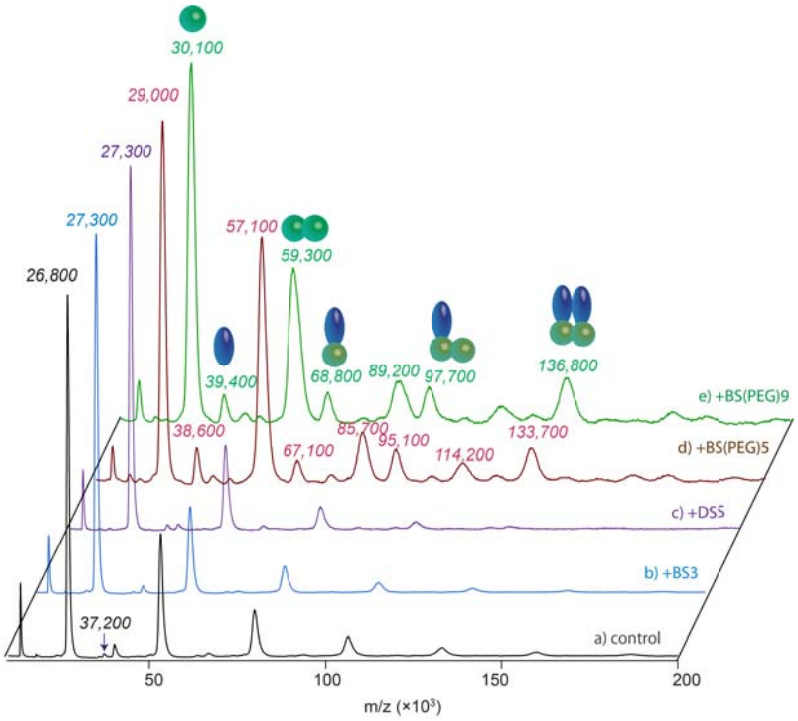


Figure 3. High-mass MALDI mass spectra of $BtuC_2D_2$ a) without any cross-linker control and b) BS^3, c) DSS, d) $BS(PEG)_5$ and e) $BS(PEG)_9$, respectively. Blue stands for BtuC subunit and green stands for BtuD subunit.

In PglK, we observed a comparable cross-linking efficiency between all NHS esters and glutaraldehyde. SI3 Table 2 presents the cross-linking efficiency of all NHS esters applied in PglK, which has been prepared in two different detergent micelles. In both detergents, all the cross-linkers share the similar cross-linking efficiency, 63 % in DDM and 71 % in C12E8. In a previous report, a comparable cross-linking efficiency was observed in NCoA-1•STAT6Y with the K_D around 30 nM (Madler et al., 2010). It indicates that PglK has the K_D in the nanomolar range. Besides the contribution of the binding affinity of the PglK dimer, we also believe the high abundance of lysine residues in PglK (55 lysine residues, 9.4 %), which is higher than average (around 6 %), to play an important role in the stabilization process via chemical cross-linking. The high abundance of lysine residues in PglK helps NHS-ester cross-linkers to efficiently stabilize the PglK dimer.

Different as the high cross-linking efficiency of all NHS esters applied in PglK, BS^3 or DSS could not stabilize the $BtuC_2D_2$ complex. Although $BS(PEG)_5$ or $BS(PEG)_9$ shows its capability to stabilize the $BtuC_2D_2$ complex, the cross-linking efficiency is much lower compared with that of glutaraldehyde as the $BtuC_2D_2$ complex was

the dominant species after reaction with glutaraldehyde (Chen et al., 2013a). The cross-linking efficiency in stabilize the $BtuC_2D_2$ complex between DSS and $BS(PEG)_5$ or $BS(PEG)_9$ should come from the most relevant difference among the NHS esters, i.e., the spacer arm length. DSS has a carbon chain with a length of 11.4 Å; while $BS(PEG)_5$ has a 21.7 Å spacer arm (35.8 Å for $BS(PEG)_9$). In our previous work, we found that the cross-linking efficiency in soluble protein complexes is strongly influenced by the distances of target amino acid residues. (Chen et al., 2013c) Therefore, we calculated the distances of aliphatic primary amine groups, including N-terminus and lysine residues, between the different subunits of $BtuC_2D_2$ based on its crystal structure (PDB 1L7V, Figure 4 (a)). We plotted the distances of between primary amine groups in the B subunit (BtuC in green) and those groups in the other three units in Figure 4(b). For the only two primary amine groups, one lysine residue and its N-terminus, in B subunit, the distances are mostly located in the range of 25 Å to 70 Å (Figure 4(b)); the distances below 20 Å (12.3 Å and 18.2 Å) are only found against two lysine residues in the D subunit (BtuD). The same for the C subunit (BtuD in pink) is presented in Figure 4 (c). For the target amine groups in D subunit (BtuD subunit), the distances under 20 Å to the amine groups in other subunits are mostly found in the other BtuD subunit, which explains the observation of the dimer of BtuD owns a higher intensity compared with other complex forms, such as BtuCD, $BtuCD_2$ and $BtuC_2D_2$. The amine-amine distances thus explain why $BS(PEG)_5$ or $BS(PEG)_9$ can stabilize the $BtuC_2D_2$ complex, but not DSS. Moreover, we believe that the limited number of lysine residues in the BtuC subunit is the reason for the low cross-linking efficiency, even for $BS(PEG)_5$ or $BS(PEG)_9$. The correlation between the chemical cross-linking efficiency and the spacer arm length of chemical cross-linkers suggested that the NHS esters could be used to map membrane protein structure in detergent micelles.

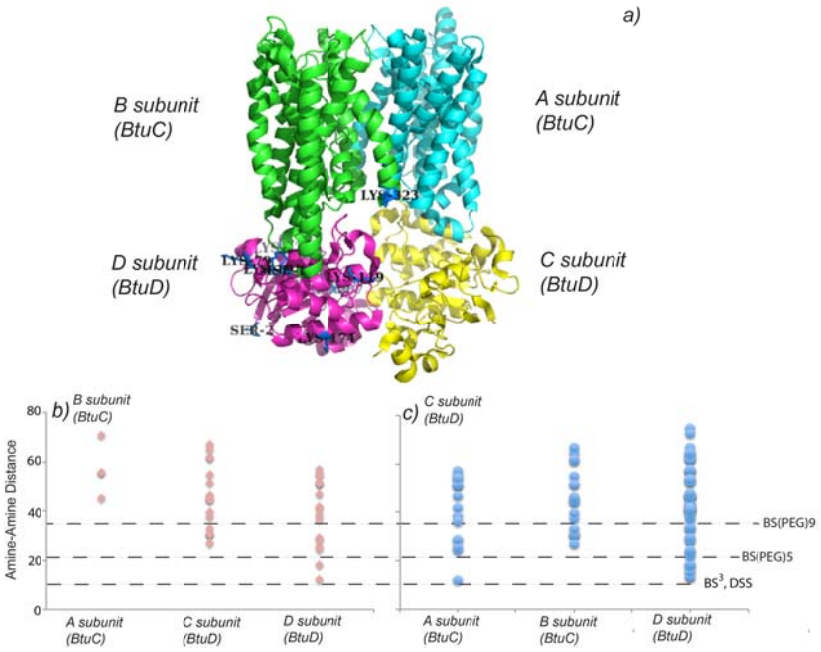


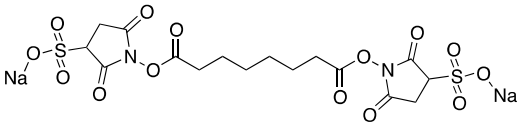
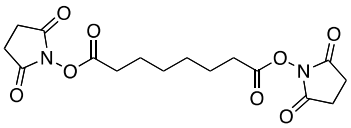
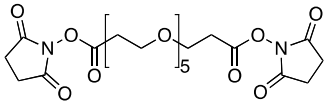
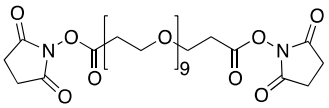
Figure 4. Ribbon diagram of BtuCD (PDB ID code 1L7V) (a); distribution of inter-subunit amine-amine distances in the B subunit of BtuCD (b) and in the D subunit of BtuCD (c).

6.4. Conclusions

In this study, we conducted specific chemical cross-linking, i.e., NHS-esters, on two membrane protein complexes, PglK and $BtuC_2D_2$. The cross-linking experiment was carried out with a series of chemical cross-linkers with different spacer arms. We observed that NHS esters could be used to stabilize or cross-link membrane protein complexes in different detergents, such as DDM and C12E8 applied here. The stabilizing capabilities differences among the cross-linkers applied should due to the protein structure, including the number of lysine residues and the amine-amine distances between different subunits. The available number of lysine residues should be the reason for the relatively low stabilization efficiency in $BtuC_2D_2$, but high stabilization efficiency in PglK. The successful cross-linking membrane protein complexes in different detergent micelles via NHS esters highlight the possibility to map membrane protein structure by chemical cross-linking, combined with mass spectrometry.

6.5. Supporting Information

Supporting Table 1: Name, structures, and distances of spacer arms for used cross-linkers.

Cross-linkers	Structure	Distance of spacer (Å)
BS^3 Bis(sulfosuccinimidyl) suberate		11.4
DSS Dissuccinimidyl suberate		11.4
$BS(PEG)_5$ Bis(succinimidyl) penta (ethylene glycol)		21.7
$BS(PEG)_9$ Bis(succinimidyl) nona (ethylene glycol)		35.8

To estimate the cross-linking efficiency, we used the equation presented in the Table 2, where S represents signal integrals. The value from a control measurement without cross-linking is subtracted to eliminate the influence of nonspecific interactions. We observed a similar stabilization efficiency for four NHS esters, BS^3, DSS, $BS(PEG)_5$ and $BS(PEG)_9$, although they have different spacer arm length (from 11.4 Å to 35.8 Å).

Table 2: Calculated chemical cross-linking efficiencies

	$\left(\frac{S_{dimer(cross-linked)}}{S_{dimer}+S_{monomer}}-\frac{S_{dimer(control)}}{S_{dimer}+S_{monomer}}\right)\times 100\%$			
Cross-linkers	BS3	DSS	$BS(PEG)_5$	$BS(PEG)_9$
		DDM		
	63 %	N.D. (not determined)	63 %	63 %
		C12E8		
	71 %	71 %	71 %	71 %

Chapter 7. Stoichiometry Study of TonB Complexes via High-mass MALDI-MS with Chemical Cross-linking

Abstract

Energy-coupled transport of iron-siderophores and vitamin B_{12} across the outer membrane of Escherichia coli and other Gram-negative bacteria requires the TonB protein. Two partner proteins, ExbB and ExbD, are required for TonB to function. TonB/ExbB/ExbD is suggested to form a complex in the cytoplasmic membrane. However, the stoichiometry of the TonB complex remains unknown. Here, we studied ExbB oligomers, the ExbB-ExbD complex and the ExbB-ExbD-TonB complex, which was extracted from the cytoplasmic membrane with the detergent n-Dodecyl-β-maltoside (DDM), by high-mass matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS). Combined with chemical cross-linking using glutaraldehyde, $ExbB_6$ was the predominant oligomeric state of ExbB. Two major species, $ExbB_4ExbD_2$ and $ExbB_5ExbD_2$, were observed in the ExbB-ExbD complex, respectively. The major species, either $ExbB_5ExbD_2$ or $ExbB_4ExbD_2$TonB, was found in the ExbB-ExbD-TonB complex. However, the identification of the biological relevant stoichiometry for the functional unit turned out to be challenging due to partial cross-linking of complexes or the presence of complex subsets and the very similar molecular weights of monomers ExbB and TonB.

7.1. Introduction

Iron is an essential element for growth in nearly all living organisms. (Wandersman & Delepelaire, 2004) Due to its oxidization in the presence of oxygen and water, iron exists either in the form of insoluble ferric hydroxides or is sequestered into host proteins. Specifically, ferric siderophores pass through the outer membrane (OM) in Gram-negative bacteria by energy-coupled transport, which is catalyzed by transport proteins consisting of a β-barrel with a pore and the plug. (Figure 1, (Krewulak & Vogel, 2011)) To release the substrate from the high-affinity binding sites into the channel, the energy required for transport is derived from the proton motive force of the cytoplasmic membrane, which is transferred through the TonB complex to the outer membrane transporters. (Braun, 1995; Braun & Braun, 2002)

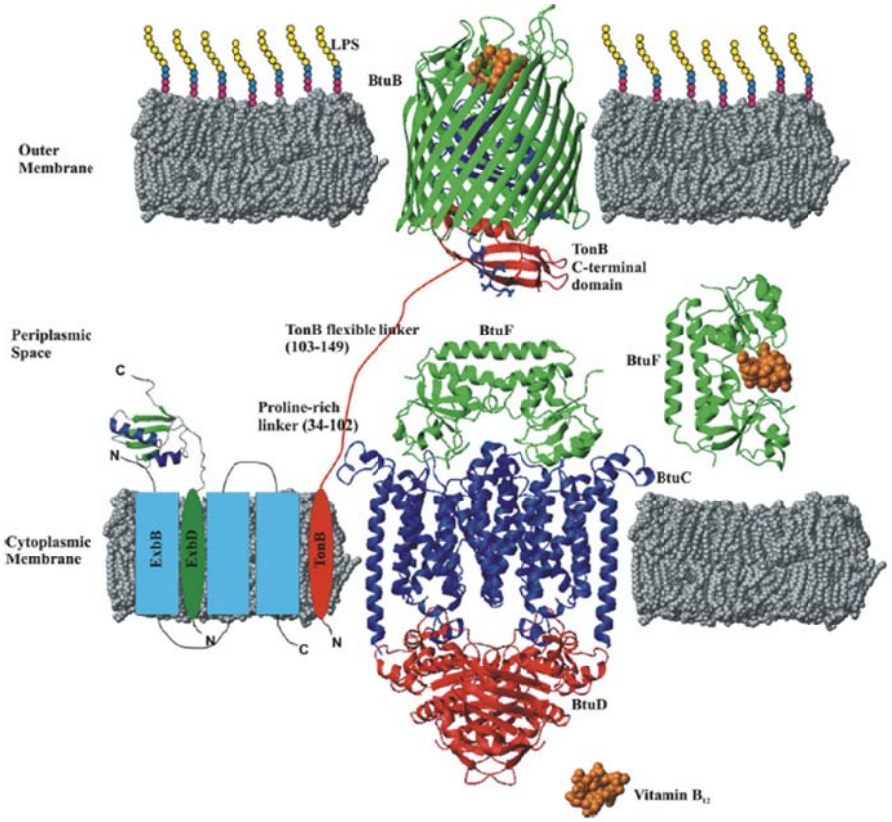


Figure 1. Schematic illustration of TonB-dependent transport and regulation in Gram-negative bacteria. All of these pathways require a TonB-dependent outer membrane receptor, a periplasmic binding protein, a cytoplasmic membrane ATP-binding cassette

(ABC) transporter and the TonB complex. Reproduced with permission from (Krewulak & Vogel, 2011). Copyright © Canadian Science Publishing.

ExbB is thought to transverse the cytoplasmic membrane three times with the N-terminus in the periplasm and the C-terminus in the cytoplasm. (Kampfenkel & Braun, 1993) In contrast, ExbD (Kampfenkel & Braun, 1992) and TonB (Postle & Skare, 1988) have a single transmembrane helix with a cytoplasmic N-terminus and a periplasmic C-terminus. Whereas the C-terminus of TonB is unstructured and spans the cytoplasm to allow interaction with outer membrane proteins. (Braun, 2003; Postel & Kadner, 2003; Wiener, 2005) Although abundant evidence suggests that the TonB complex consists of the cytoplasmic membrane proteins TonB, ExbB and ExbD, solid data of the nature oft he complex formed by the three proteins in the cytoplasmic membrane is still missing. Understanding stoichiometry and arrangement of TonB, ExbB and ExbD within the complex will help to understand how the interaction results in the transport of the substrate.

A careful determination of the protein concentrations (TonB/ExbB/ExbD) in cells, with the assistance of the specific incorporation of a quantifiable radiolabel into the subunit and inmunoblots, resulted in a 7:2:1 ratio for ExbB:ExbD:TonB (Higgs et al., 2002); however, it is unclear whether this reflects the stoichiometry of the proteins in the complex. In the complex, ExbB is supposed to stabilize TonB and ExbD, because TonB and ExbD are proteolytically degraded in the absence of ExbB (Karlsson et al., 1993); four different techniques, including blue native gel electrophoresis, size exclusion chromatography, transmission electron microscopy and small-angle X-ray scattering indicate that ExbB forms a stable homo-oligomer with 4 to 6 monomers. (Pramanik et al., 2010) Homodimers of ExbB, ExbD and TonB, as well as heterodimers of After cross-linking with formaldehyde, TonB with ExbD and ExbB with ExbD were observed in the cell. (Ollis et al., 2009)

Mass spectrometry (MS) has been widely accepted and used as a bioanalytical tool in structural studies. Laser-induced liquid bead ion desorption mass spectrometry (LILBID-MS) has been used to study membrane protein complexes. (Pramanik et al., 2011) With a fine tuned laser energy, non-covalent bondscan be preserved, to some extent, in LILBID-MS. Recently, we developed a standard

matrix-assisted laser desorption/ionization (MALDI) protocol, in combination with a high-mass detector, to study integral membrane proteins. (Chen et al., 2013a) If combined with chemical cross-linking, high-mass MALDI-MS can also be used to determine the subunit stoichiometry of the integral membrane protein complex. (Chen et al., 2013a)

Here, we present a study of the oligomeric state of the TonB complex, by high-mass MALDI-MS with chemical cross-linking via glutaraldehyde. We carried out the studies on the intact complex, i.e., ExbB-ExbD-TonB, and also on subsets of the complex, ExbB and ExbB-ExbD. Our results indicate that ExbB, which forms a stable hexamer ($ExbB_6$), is responsible to stabilize the whole complex. Major species that were observed include $ExbB_4ExbD_2$ and $ExbB_5ExbD_2$ in the ExbB-ExbD complex. $ExbB_5ExbD_2$ or $ExbB_4ExbD_2$TonB were observed in the ExbB-ExbD-TonB complex.

7.2 Experimental

7.2.1. Protein Expression, Solubilization and Purification

ExbB oligomers, the ExbB-ExbD complex and the ExbB-ExbD-TonB complex were prepared and purified as described previously. (Gerber, 2008) (Experiments were done by Sabina Gerber and Monika Bucher, Institute of Molecular Biology and Biophysics, ETH Zürich.)

7.2.2. Chemical Cross-linking

The protein complexes were incubated with a 10% glutaraldehyde (GA) solution in a 10/1 (v/v) ratio for 1 h at room temperature. The mixture was further diluted with the original protein buffer solution or water prior to mass spectrometric analysis.

7.2.3. Mass Spectrometry

The measurements were carried out as described previously. (Chen et al., 2013a)

7.3. Results and Discussion

7.3.1. Oligomeric structure of the ExbB complex

Previous studies suggested that ExbB is the scaffold for the entire complex. It probably forms a stable homo-oligomer consisting of four to six monomers based on results from blue native gel electrophoresis, size exclusion chromatography, transmission electron microscopy and small-angle X-ray scattering. (Pramanik et al., 2010) At moderate laser intensities, LILBID-MS displayed a mass of 154.8 kDa, which has been assigned to be a hexamer of ExbB with a theoretical molecular weight of 163.8 kDa (mass error = -5%). (Pramanik et al., 2011) However, the experimental molecular weight is expected to be larger than the theoretical molecular weight, because detergent molecules may still attach to the complex in moderate laser intensity. A possible albeit unlikely reason might simply be poor calibration of the instrument. In our previous study (Chen et al., 2013a), high-mass MALDI-MS has been applied to detect membrane proteins with high mass accuracy. Moreover, a correlation extracted from the number of amine groups and the MW increase after reaction with glutaraldehyde allowed us to correct for "decoration" by glutaraldehyde monolinks, and to derive MWs of the bare complexes. Here, we first studied the oligomeric state of ExbB by using high-mass MALDI-MS for the first time. (Chen et al., 2013a)

In Figure 2 (a), a peak corresponding to the monomer of ExbB (*m/z* 27,300, theoretical MW=27,352 Da, mass error = 0. 2 %) was observed. The doubly charged ExbB and the non-specific homodimer were also detected. After cross-linking by glutaraldehyde (Figure 2 (b)), a major peak at *m/z* 174,500 was observed and assigned to the ExbB hexamer. The observed mass is higher than the theoretical molecular mass of the ExbB hexamer, which is 164,112 Da. The correlation established in our previous study would suggest an estimated molecular weight of 172,288 Da due to the glutaraldehyde "decoration" (Table 1). (Chen et al., 2013a) This estimated molecular mass is reasonably close to the experimental molecular mass, with a mass error of 1 %. Besides $ExbB_6$, the monomeric ExbB at *m/z* 29,200 was also detected. The observation of monomeric ExbB probably comes from insufficient chemical cross-linking. Species such as $ExbB_2$ at *m/z* 58,400, $ExbB_3$ (partially also from doubly charged

$ExbB_6$) at *m/z* 88,100 and $ExbB_4$ at *m/z* 116,700, could stem from partially cross-linked proteins.

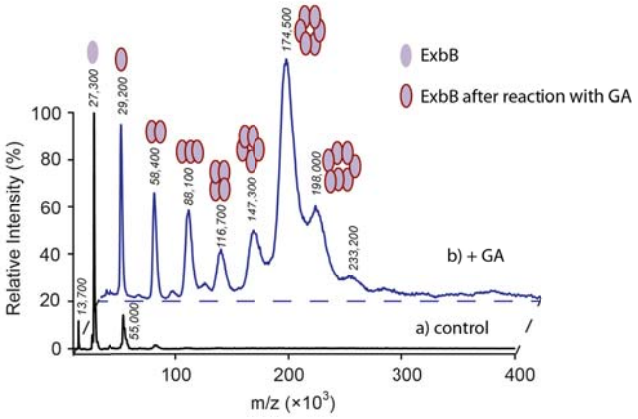


Figure 2. High-mass MALDI mass spectra of ExbB from *E. coli* control (a) and ExbB following cross-linking with GA (b).

The theoretical molecular weights, the estimated molecular weights with GA "decoration", and the experimental molecular weights of different species are listed in Table 1. The relative intensity of $ExbB_5$ suggests that $ExbB_5$ was generated before MALDI ionization considering that non-specific clusters should exhibit an exponentially decaying intensity distribution. However, it is unclear whether $ExbB_5$ existed already in the solution before cross-linking or not. Interestingly, besides pentameric ExbB, heptameric ExbB was also observed. This could be the result of clustering occurring in the MALDI plume because of the high intensity of ExbB and $ExbB_6$. Another possible explanation is that non-specific species are generated by chemical cross-linking, although non-specific species are seldom formed via cross-linking under well controlled condition. Hence, based on the relative intensities, the dominant species, $ExbB_6$, presents the major oligomeric state of ExbB.

Table 1. Theoretical average molecular weights, estimated molecular weights with glutaraldehyde "decoration" and experimental molecular weight of different oligomeric states of ExbB complexes from *E. coli*. Estimated molecular weights were calculated based on the correlation established between the number of free amine groups and the glutaraldehyde "decoration". (Chen et al., 2013a)

	ExbB (*E. Coli*)							
	$ExbB_1$	$ExbB_2$	$ExbB_3$	$ExbB_4$	$ExbB_5$	**$ExbB_6$**	$ExbB_7$	$ExbB_8$
	Estimated MW increase = 151.4 * No (amine groups)							
Theoretical MW	**27,352**	**54,701**	**82,056**	**109,408**	**136,760**	**164,112**	**191,464**	**218,816**
Estimated MW with glutaraldehyde "decoration"	*28,714*	*57,429*	*86,144*	*114,858*	*143,573*	*172,288*	*201,002*	*229,717*
Experimental MW	27,300	55,000						
Observed MW (+ GA)	**29,200**	58,400	88,100	116,700	147,300	**174,500**	198,000	233,200

7.3.2. Oligomeric state of the ExbB-ExbD complex

The study of the stoichiometry of ExbB-ExbD complexes carried out by LILBID-MS showed complexes with a size of 171.2 kDa ($ExbB_6ExbD_1$, theoretical mass 180.2 kDa) and a size of 145.4 kDa ($ExbB_5ExbD_1$, theoretical mass 152.9 kDa). (Pramanik et al., 2011) The conclusion was drawn that $ExbB_6$ tightly binds to one ExbD subunit, and that this was the major species observed. Here, we propose another assignment of the peaks at around 171.2 kDa: $ExbB_5ExbD_2$ (theoretical mass 169.3 kDa) and around 145.4 kDa: $ExbB_4ExbD_2$ (theoretical mass 142 kDa). This assignment is based on a significantly lower mass error.

High-mass MALDI-MS, combined with chemical cross-linking, i.e., glutaraldehyde, was applied in exactly the same fashion to study the stoichiometry of the ExbB-ExbD complex from three species, *E. coli, S. typhimurium* and *V. cholerae*. In an earlier study carried out by Gerber (Gerber, 2008), the complexes from the three microorganisms *E. coli, S. typhimurium* and *V. cholerae* were the most promising homologues for structural studies.

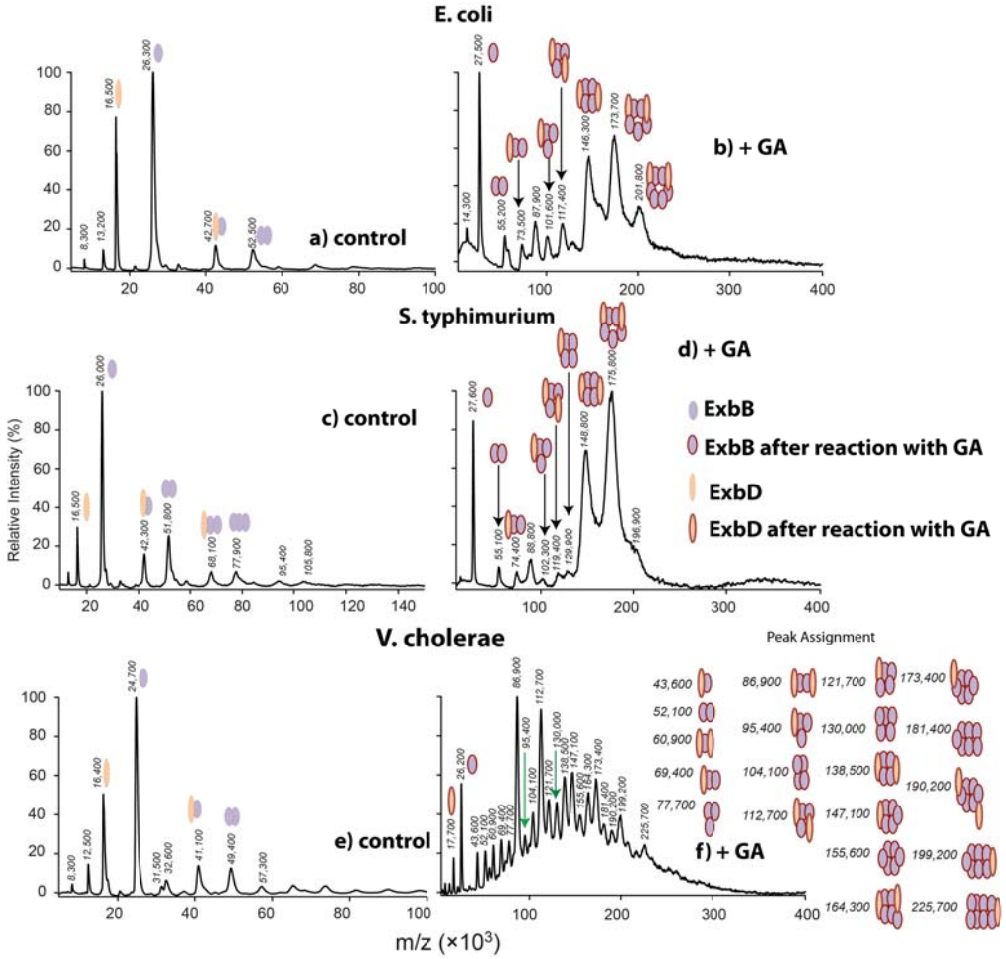


Figure 3 High-mass MALDI mass spectra of ExbB-ExbD control (a) from *E. coli*, (c) from *S. typhimurium* and (e) from *V. cholera*, and ExbB-ExbD following cross-linking with GA (b) from *E. coli*, (d) from *S. typhimurium* and (f) from *V. cholera*. The graphic representation of various hetero complex and their estimated MWs after reaction with GA are given in Table 2.

In Figure 3 (a), peaks corresponding to the monomer of ExbB (*m/z* 26,200, theoretical MW=26,287 Da, mass error = 0.33 %) and the monomer of ExbD (*m/z* 16,500, theoretical MW=16,610 Da, mass error = 0.66 %) can be seen, for *E. coli*. The higher intensity of the peak assigned to ExbB, compared with that of ExbD, suggests that ExbB is the major component in the ExbB-ExbD complex rather than ExbD. After chemical cross-linking via glutaraldehyde (Figure 3(b)), two major peaks at *m/z* 173,700 and *m/z* 146,300 were detected besides ExbB at *m/z* 27,500 (theoretical MW=26,287 Da, estimated MW with glutaraldehyde "decoration" =27,650 Da). We assigned the stoichiometries $ExbB_5ExbD_2$ (theoretical MW=164,655 Da, estimated MW with glutaraldehyde

decoration=174,799 Da) and $ExbB_4ExbD_2$ (theoretical MW=138,368 Da, estimated MW with glutaraldehyde "decoration" =147,149 Da) to the two species observed at *m/z* 173,700 and *m/z* 146,300 (as shown in Table 2). Moreover, we assigned some minor species at *m/z* 117,400 ($ExbB_3ExbD_2$, estimated MW with glutaraldehyde "decoration" =119,500 Da), 101,600 ($ExbB_3ExbD_1$, estimated MW with glutaraldehyde "decoration" =101,224 Da), 201,800 ($ExbB_6ExbD_2$, estimated MW with glutaraldehyde "decoration" =202,448 Da). The experimental MWs are slightly smaller than the estimated MWs with glutaraldehyde "decoration", probably due to a higher concentration of the TonB complexes used here compared to the concentration used in the earlier study. (Chen et al., 2013a) The observation of $ExbB_6ExbD_2$ at *m/z* 201,800 could be due to closer ions formed in MALDI. The observation of two dominant species, $ExbB_5ExbD_2$ and $ExbB_4ExbD_2$, in the *E. coli* ExbB-ExbD complex could come from: 1) insufficient chemical cross-linking and 2) a possible equilibrium between $ExbB_5ExbD_2$ and $ExbB_4ExbD_2$ in solution. Interestingly, the number of ExbB subunits in the ExbB-ExbD complex was less than that in the ExbB complex. This suggests that the construction of the ExbB-ExbD complex is not simply based on $ExbB_6$ even though the ExbB subunit was proposed to stabilize the whole complex.

Table 2. Theoretical average molecular weights, estimated molecular weights with glutaraldehyde "decoration" and experimental molecular weight of different oligomeric states of ExbB-ExbD subcomplexes from *E. coli*.

ExbB*ExbD (*E. Coli*)								
Estimated MW increase = 151.4 * No (amine groups)								
	$ExbB_1$	$ExbB_2$	$ExbB_3$	$ExbB_4$	$ExbB_5$	$ExbB_6$	$ExbB_7$	$ExbB_8$
Theoretical MW	**26,287**	**52,574**	**78,861**	**105,148**	**131,435**	**157,722**	**184,009**	**210,296**
Estimated MW with glutaraldehyde "decoration"	*27,650*	*55,300*	*82,949*	*110,598*	*138,248*	*165,898*	*193,547*	*221,197*
	$ExbB_4ExbD_1$	$ExbB_5ExbD_1$	$ExbB_6ExbD_1$	$ExbB_7ExbD_1$	$ExbB_4ExbD_2$	$ExbB_5ExbD_2$	$ExbB_6ExbD_2$	$ExbB_7ExbD_2$
Theoretical MW	**121,758**	**148,045**	**174,332**	**217,229**	**138,368**	**164,655**	**190,942**	**217,229**
Estimated MW with glutaraldehyde "decoration"	*128,874*	*156,523*	*184,173*	*211,823*	*147,149*	*174,799*	*202,448*	*230,098*
	$ExbB_3ExbB_3$	$ExbB_4ExbD_3$	$ExbB_5ExbD_3$	$ExbB_6ExbD_3$				
Theoretical MW	**128,691**	**154,978**	**181,265**	**207,552**				
Estimated MW with glutaraldehyde "decoration"	*137,775*	*165,427*	*193,074*	*220,724*				
Experimental Data								
Observed MW (+ GA)	**27,500**	55,200	73,500	87,900	101,600	117,400	**146,300**	**173,700**
Assignment	**$ExbB_1$** ***(27,650)***	$ExbB_2$ *(55,300)*	$ExbB_2$ $ExbD_1$ *(73,574)*	($ExbB_5$ $ExbD_2$)$^{2+}$ *(87,400)*	$ExbB_3$ $ExbD_1$ *(101,224)*	$ExbB_3$ $ExbD_2$ *(119,500)*	**$ExbB_4$ $ExbD_2$** ***(147,799)***	**$ExbB_5ExbD_2$** ***(174,799)***
Observed MW (+ GA)	201,800							
Assignment	$ExbB_6$ $ExbD_2$ *(202,448)*							

The high-mass MALDI mass spectrum of the ExbB-ExbD complex from *S. typhimurium* is shown in in Figure 3 (c). Peaks corresponding to the monomer of ExbB (*m/z* 26,000, theoretical MW=26,179 Da, mass error = 0.68 %) and the monomer of ExbD (*m/z* 16,500, theoretical MW=16,442 Da, mass error = 0.35 %) were observed. The relative intensity of the peak assigned to ExbB is much higher that of ExbD, suggesting that ExbB, rather than ExbD, is also the major component in the ExbB-ExbD complex rather than ExbD in *S. typhimurium*. Cross-linking via glutaraldehyde as recorded in Figure 3(d) generated two major peaks at *m/z* 175,800 and *m/z* 148,800, and also ExbB at *m/z* 27,500 (theoretical MW=26,179 Da, estimated MW with glutaraldehyde "decoration" =27,538 Da). The stoichiometries $ExbB_5ExbD_2$ (theoretical MW=163,799 Da, estimated MW with glutaraldehyde "decoration" =173,896 Da) and $ExbB_4ExbD_2$ (theoretical MW=137,600 Da, estimated MW with glutaraldehyde "decoration" =146,358 Da), were assigned to the two peaks observed species at *m/z* 175,800 and *m/z* 148,800. (Table 3) Different to the ExbB-ExbD complex from *E. coli* (Figure 3(b)), the relative intensities of minor species are much lower here. We also did not observe any higher oligomeric states beyond $ExbB_5ExbD_2$. The observation of $ExbB_5ExbD_2$ and $ExbB_4ExbD_2$, both in *E. coli* and *S. typhimurium*, suggests that the ExbB-ExbD complex could originate from an equilibrium between $ExbB_5ExbD_2$ and $ExbB_4ExbD_2$. The assignments we observed in high-mass MALDI-MS happens to coincides the stoichiometry we proposed for the ExbB-ExbD observed in LILBID (Pramanik et al., 2011) based on a significantly lower mass error, which is different than the one proposed by Braun (Pramanik et al., 2011).

Table 3. Theoretical average molecular weights, estimated molecular weights with glutaraldehyde "decoration" and experimental molecular weights of different oligomeric state of ExbB-ExbD subcomplexes from *S. typhimurium*.

ExbB*ExbD (*S. typhimurium*)								
Estimated MW increase = 151.4 * No (amine groups)								
	$ExbB_1$	$ExbB_2$	$ExbB_3$	$ExbB_4$	$ExbB_5$	$ExbB_6$	$ExbB_7$	$ExbB_8$
Theoretical MW	**26,179**	**52,358**	**78,537**	**104,716**	**130,895**	**157,074**	**183,253**	**209,432**
Estimated MW with glutaraldehyde "decoration"	*27,538*	*55,076*	*82,614*	*110,152*	*137,690*	*165,228*	*192,766*	*220,304*
	$ExbB_4ExbD_1$	$ExbB_5ExbD_1$	$ExbB_6ExbD_1$	$ExbB_7ExbD_1$	$ExbB_4ExbD_2$	$ExbB_5ExbD_2$	$ExbB_6ExbD_2$	$ExbB_7ExbD_2$
Theoretical MW	**121,158**	**147,337**	**173,516**	**199,695**	**137,600**	**163,779**	**189,958**	**216,137**
Estimated MW with glutaraldehyde "decoration"	*128,254*	*155,792*	*183,330*	*210,869*	*146,358*	*173,896*	*201,434*	*228,972*
	$ExbB_3ExbB_3$	$ExbB_4ExbD_3$	$ExbB_5ExbD_3$	$ExbB_6ExbD_3$				
Theoretical MW	**127,773**	**153,952**	**180,131**	**206,310**				
Estimated MW with glutaraldehyde "decoration"	*137,116*	*164,654*	*192,192*	*219,730*				
Experimental Data								
Observed MW (+ GA)	**27,600**	55,100	74,400	88,800	102,300	119,400	129,900	**148,800**
Assignment	**$ExbB_1$** ***(27,538)***	$ExbB_2$ *(55,076)*	$ExbB_2$ $ExbD_1$ *(73,179)*	$(ExbB_5$ $ExbD_2)^{2+}$ *(86,948)*	$ExbB_3$ $ExbD_1$ *(100,717)*	$ExbB_3$ $ExbD_2$ *(118,820)*	$ExbB_4$ $ExbD_1$ *(128,254)*	**$ExbB_4ExbD_2$** ***(146,358)***
Observed MW (+ GA)	**175,800**							
Assignment	**$ExbB_5$** **$ExbD_2$** ***(173,896)***							

A high-mass MALDI mass spectrum of the ExbB-ExbD complex from *V. cholerae* is shown in Figure 3 (e). Peaks corresponding to the monomer of ExbB (*m/z* 24,700, theoretical MW=24,720 Da, mass error = 0.08 %) and the monomer of ExbD (*m/z* 16,400, theoretical MW=16,304 Da, mass error = 0.58 %) were observed. Cross-linking with glutaraldehyde for the ExbB-ExbD complex from *V. cholerae* (Figure 3(f)) did not generate a very clear dominant species. A series of peaks from the heterodimer ExbBExbD to $ExbB_8ExbD$ were observed and are labeled in Figure 3(f). (See also Table 4) Three dominant species were ExbB at *m/z* 26,200 (theoretical MW=24,720 Da, estimated MW with glutaraldehyde "decoration" =26,083 Da), $ExbB_2ExbD_2$ at *m/z* 86,900 (estimated MW with glutaraldehyde "decoration" =88,103 Da) and $ExbB_3ExbD_2$ at *m/z* 112,700 (estimated MW with glutaraldehyde "decoration" =114,186 Da). The detection of a series of complexes could be the result of poor chemical cross-linking of the ExbB-ExbD complex from *V. cholerae*.

7.3.3. Oligomeric structure of the ExbB-ExbD-TonB complex

Crystallization trials have been done on subsets of the TonB complex, such as ExbB and ExbB-ExbD, however, crystallographic investigations on the whole assembly turned out to be very challenging because TonB has a long C-terminus, which is believed to be highly unstructured when unbound to outer membrane receptors. (Gerber, 2008) In stoichiometric studies, MS does not require a well-structured protein, as opposed for X-ray or NMR. Here, we targeted the entire complex, ExbB-ExbD-TonB, from *E. coli*.

In Figure 4 (a), peaks corresponding to the monomer of ExbB (*m/z* 26,200, theoretical MW=26,287 Da, mass error = 0.33 %) and the monomer of ExbD (*m/z* 16,500, theoretical MW=16,610 Da, mass error = 0.66 %) are shown. However, it is hard to detect a separate peak corresponding to TonB, because the theoretical MW of TonB is 26,094 Da, which is very close to that of ExbB. To overcome this problem, a tag was added to the TonB subunit to increase its molecular weight. TonB-Tag was observed at *m/z* 28,800 (theoretical MW=28,891 Da, mass error = 0.48 %) in Figure 4(c). The relative intensities of the signals of ExbB, ExbD and TonB-Tag in Figure 4(c) suggest that ExbB is the

major component in the ExbB-ExbD complex. However, it is inconclusive to compare the number of TonB and ExbD subunits based on the intensities of their signals.

After introducing glutaraldehyde, a dominant peak at *m/z* 175,000 was observed (Figure 4(b)). Moreover, two other peaks at *m/z* 147,200 ($ExbB_4ExbD_2$, estimated MW with glutaraldehyde "decoration" =147,100 Da), and *m/z* 202,400 were detected besides ExbB at *m/z* 28,000 (theoretical MW=26,287 Da, estimated MW with glutaraldehyde "decoration" =27,650 Da). The dominant species at *m/z* 175,000 could be assigned to $ExbB_5ExbD_2$ (theoretical MW=164,655 Da, estimated MW with glutaraldehyde "decoration" =174,799 Da) or $ExbB_4ExbD_2TonB$ (theoretical MW=164,462 Da, estimated MW with glutaraldehyde "decoration" =176,100 Da). The species at *m/z* 202,400 could be assigned to either $ExbB_6ExbD_2$ (theoretical MW=190,942 Da, estimated MW with glutaraldehyde "decoration" =202,448 Da) or $ExbB_5ExbD_2TonB$ (theoretical MW=190,749 Da, estimated MW with glutaraldehyde "decoration" =203,762 Da). After treating the TonB-Tag complexes with glutaraldehyde (Figure 4(d)), we observed dominant peaks at *m/z* 176,000 and 27,700, and also minor peaks at *m/z* 149,000, 157,600, and 202,900. The dominant species at *m/z* 176,000 could be assigned to $ExbB_5ExbD_2$ (theoretical MW=174,655 Da, estimated MW with glutaraldehyde "decoration" =174,799 Da) or $ExbB_4ExbD_2TonB$ (theoretical MW=167,260 Da, estimated MW with glutaraldehyde "decoration" =178,918 Da). The species at *m/z* 202,900 could be assigned to $ExbB_6ExbD_2$ (theoretical MW=190,942 Da, estimated MW with glutaraldehyde "decoration" =202,448 Da) or $ExbB_5ExbD_2TonB$ (theoretical MW=193,547 Da, estimated MW with glutaraldehyde "decoration" =206,567 Da). The species at *m/z* 149,000 could be assigned to $ExbB_4ExbD_2$ (theoretical MW=138,368 Da, estimated MW with glutaraldehyde "decoration" =147,149 Da) and the species at *m/z* 157,600 could be assigned to $ExbB_5ExbD_1$ (theoretical MW=148,045 Da, estimated MW with glutaraldehyde "decoration" =156,523 Da). Unfortunately, at this stage, we could not derive an unambiguous assignment of the stoichiometry of the entire TonB complex, because the mass errors for two different assignments are both reasonably low.

Table 4. Theoretical average molecular weights, estimated molecular weights with glutaraldehyde "decoration" and experimental molecular weights of different oligomeric states of ExbB-ExbD subcomplexes from *V. cholerae*.

ExbB*ExbD (*Vibrio Cholerae*)								
Estimated MW increase = 151.4 * No (amine groups)								
	$ExbB_1$	$ExbB_2$	$ExbB_3$	$ExbB_4$	$ExbB_5$	$ExbB_6$	$ExbB_7$	$ExbB_8$
Theoretical MW	**24,720**	**49,400**	**74,200**	**123,600**	**123,600**	**148,300**	**173,000**	**197,800**
Estimated MW with glutaraldehyde "decoration"	*26,083*	*52,165*	*78,248*	*104,330*	*130,413*	*156,496*	*182,578*	*208,661*
	$ExbB_1ExbD_1$	$ExbB_2ExbD_1$	$ExbB_3ExbD_1$	$ExbB_4ExbD_1$	$ExbB_5ExbD_1$	$ExbB_6ExbD_1$	$ExbB_7ExbD_1$	$ExbB_8ExbD_1$
Theoretical MW	**41,024**	**65,734**	**90,464**	**115,184**	**139,904**	**164,624**	**189,345**	**214,064**
Estimated MW with glutaraldehyde "decoration"	*44,052*	*70,134*	*96,217*	*122,299*	*148,382*	*174,465*	*200,547*	*226,630*
	$ExbB_1ExbD_2$	$ExbB_2ExbD_2$	$ExbB_3ExbD_2$	$ExbB_4ExbD_2$	$ExbB_5ExbD_2$	$ExbB_6ExbD_2$	$ExbB_7ExbD_2$	$ExbB_8ExbD_2$
Theoretical MW	**57,328**	**82,048**	**106,768**	**131,488**	**156,208**	**180,928**	**205,648**	**230,368**
Estimated MW with glutaraldehyde "decoration"	*62,021*	*88,103*	*114,186*	*140,268*	*166,351*	*192,434*	*218,516*	*244,599*
	$ExbB_1ExbD_3$	$ExbB_2ExbD_3$	$ExbB_3ExbD_3$	$ExbB_4ExbD_3$	$ExbB_5ExbD_3$	$ExbB_6ExbD_3$	$ExbB_7ExbD_3$	$ExbB_8ExbD_3$
Theoretical MW	**73,632**	**98,352**	**123,072**	**147,792**	**172,512**	**197,232**	**221,952**	**246,672**
Estimated MW with glutaraldehyde "decoration"	*79,990*	*106,072*	*132,155*	*158,237*	*184,320*	*210,403*	*236,485*	*262,568*
Experimental Data								
Observed MW (+ GA)	17,700	26,200	43,600	52,100	60,900	69,400	77,700	**86,900**
Assignment	$ExbD_1$ *(17,969)*	$ExbB_1$ *(26,083)*	$ExbB_1$ $ExbD_1$ *(44,051)*	$ExbB_2$ *(52,165)*	$ExbB_1$ $ExbD_2$ *(62,020)*	$ExbB_2$ $ExbD_1$ *(70,134)*	$ExbB_3$ *(78,248)*	**$ExbB_2ExbD_2$** ***(88,103)***
Observed MW (+ GA)	95,400	104,100	**112,700**	121,700	130,000	138,500	147,100	155,600
Assignment	$ExbB_3$ $ExbD_1$ *(96,217)*	$ExbB_4$ *(104,330)*	**$ExbB_3$ $ExbD_2$** ***(114,186)***	$ExbB_4$ $ExbD_1$ *(122,299)*	$ExbB_5$ *(130,413)*	$ExbB_4$ $ExbD_2$ *(140,268)*	$ExbB_5$ $ExbD_1$ *(148,382)*	$ExbB_6$ *(156,496)*
Observed MW (+ GA)	164,300	173,400	199,200	225,700				
Assignment	$ExbB_5$ $ExbD_2$ *(166,351)*	$ExbB_6$ $ExbD_1$ *(174,465)*	$ExbB_7$ $ExbD_1$ *(200,547)*	$ExbB_8$ $ExbD_1$ *(226,630)*				

Although we observed TonB-Tag in Figure 4(c) directly, we cannot assure the dominant species observed in Figures 4(b) and (d) contain the TonB subunit. Interestingly, the dominant species observed in the TonB complex has a comparable size with the ExbB-ExbD complex; we also observed this size similarity between the ExbB and ExbB-ExbD complexes. This observation suggests that the TonB complex is not simply constructed by adding ExbD or TonB subunit into the ExbB hexameric structure.

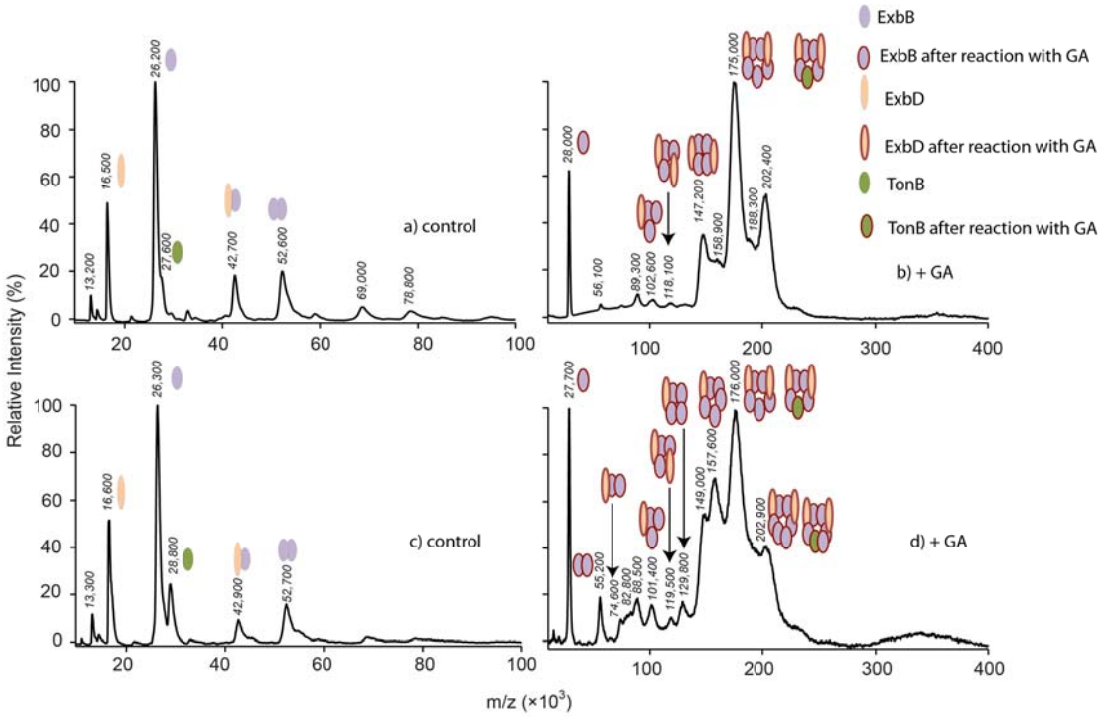


Figure 4. High-mass MALDI mass spectra of ExbB-ExbD-TonB from *E. coli* without cross-linking (a) and ExbB-ExbD-TonB following cross-linking with GA (b), ExbB-ExbD-TonB-Tag from *E. coli* without cross-linking (c) and ExbB-ExbD-TonB-Tag following cross-linking with GA (d).

Table 5. Theoretical average molecular weights, estimated molecular weights with glutaraldehyde "decoration" and experimental molecular weights of different oligomeric states of ExbB-ExbD-TonB complexes from *E. Coli*.

ExbB*ExbD*TonB (*E. Coli*)								
Estimated MW increase = 151.4 * No (amine groups)								
	Exb B_1	$ExbB_2$	$ExbB_3$	$ExbB_4$	$ExbB_5$	$ExbB_6$	$ExbB_7$	$ExbB_8$
Theoretical MW	**26,287**	**52,574**	**78,861**	**105,148**	**131,435**	**157,722**	**184,009**	**210,296**
Estimated MW with glutaraldehyde "decoration"	*27,650*	*55,300*	*82,949*	*110,598*	*138,248*	*165,898*	*193,547*	*221,197*
	$ExbB_4ExbD_1$	$ExbB_5ExbD_1$	$ExbB_6ExbD_1$	$ExbB_7ExbD_1$	$ExbB_4ExbD_2$	$ExbB_5ExbD_2$	$ExbB_6ExbD_2$	$ExbB_7ExbD_2$
Theoretical MW	**121,758**	**148,045**	**174,332**	**200,619**	**138,368**	**164,655**	**190,942**	**217,229**
Estimated MW with glutaraldehyde "decoration"	*128,874*	*156,523*	*184,173*	*211,823*	*147,149*	*174,799*	*202,448*	*230,098*
	$ExbB_4ExbD_1TonB_1$	$ExbB_5ExbD_1TonB_1$	$ExbB_6ExbD_1TonB_1$	$ExbB_7ExbD_1TonB_1$	$ExbB_4ExbD_2TonB_1$	$ExbB_5ExbD_2TonB_1$	$ExbB_4ExbD_2TonB_1$	$ExbB_4ExbD_2TonB_1$
Theoretical MW	**147,852**	**174,139**	**200,426**	**226,713**	**164,462**	**190,749**	**217,036**	**243,323**
Estimated MW with glutaraldehyde "decoration"	*157,837*	*185,486*	*184,173*	*211,823*	*176,112*	*203,762*	*231,411*	*259,061*
	$ExbB_4ExbD_1TonB_2$	$ExbB_5ExbD_1TonB_2$	$ExbB_6ExbD_1TonB_2$	$ExbB_7ExbD_1TonB_2$	$ExbB_4ExbD_2TonB_2$	$ExbB_4ExbD_2TonB_2$	$ExbB_4ExbD_2TonB_2$	$ExbB_4ExbD_2TonB_2$
Theoretical MW	**173,946**	**200,233**	**226,520**	**252,807**	**190,556**	**216,843**	**243,130**	**269,417**
Estimated MW with glutaraldehyde "decoration"	*186,800*	*214,449*	*242,099*	*269,749*	*215,763*	*243,412*	*271,062*	*298,712*
Experimental Data								
Observed MW (+ GA)	**28,000**		147,200	**175,000**	**202,400**			
Assignment		**$ExbB_1$ *(27,650)***	$ExbB_4$ $ExbD_2$ *(147,100)*	**$ExbB_5$ $ExbD_2$ *(174,800)* or $ExbB_4ExbD_2TonB_1$ *(176,100)***	**$ExbB_6$ $ExbD_2$ *(202,400)* or $ExbB_5ExbD_2TonB_1$ *(203,800)***			

Table 6. Theoretical average molecular weights, estimated molecular weights with glutaraldehyde "decoration" and experimental molecular weights of different oligomeric states of ExbB-ExbD-TonB-Tag complex from *E. Coli*.

ExbB*ExbD*TonB-Tag (*E. Coli*)								
Estimated MW increase = 151.4 * No (amine groups)								
	$ExbB_1$	$ExbB_2$	$ExbB_3$	$ExbB_4$	$ExbB_5$	$ExbB_6$	$ExbB_7$	$ExbB_8$
Theoretical MW	**26,287**	**52,574**	**78,861**	**105,148**	**131,435**	**157,722**	**184,009**	**210,296**
Estimated MW with glutaraldehyde "decoration"	*27,650*	*55,300*	*82,949*	*110,598*	*138,248*	*165,898*	*193,547*	*221,197*
	$ExbB_4ExbD_1$	$ExbB_5ExbD_1$	$ExbB_6ExbD_1$	$ExbB_7ExbD_1$	$ExbB_4ExbD_2$	$ExbB_5ExbD_2$	$ExbB_6ExbD_2$	$ExbB_7ExbD_2$
Theoretical MW	**121,758**	**148,045**	**174,332**	**200,619**	**138,368**	**164,655**	**190,942**	**217,229**
Estimated MW with glutaraldehyde "decoration"	*128,874*	*156,523*	*184,173*	*211,823*	*147,149*	*174,799*	*202,448*	*230,098*
	$ExbB_4ExbD_1TonB_1$	$ExbB_5ExbD_1TonB_1$	$ExbB_6ExbD_1TonB_1$	$ExbB_7ExbD_1TonB_1$	$ExbB_4ExbD_2TonB_1$	$ExbB_5ExbD_2TonB_1$	$ExbB_4ExbD_2TonB_1$	$ExbB_4ExbD_2TonB_1$
Theoretical MW	**150,650**	**176,937**	**203,224**	**229,511**	**167,260**	**193,547**	**219,834**	**246,121**
Estimated MW with glutaraldehyde "decoration"	*160,642*	*188,292*	*215,942*	*243,591*	*178,918*	*206,567*	*234,217*	*261,867*
	$ExbB_4ExbD_1TonB_2$	$ExbB_5ExbD_1TonB_2$	$ExbB_6ExbD_1TonB_2$	$ExbB_7ExbD_1TonB_2$	$ExbB_4ExbD_2TonB_2$	$ExbB_4ExbD_2TonB_2$	$ExbB_4ExbD_2TonB_2$	$ExbB_4ExbD_2TonB_2$
Theoretical MW	**179,542**	**205,829**	**232,116**	**258,403**	**196,152**	**222,439**	**248,726**	**275,013**
Estimated MW ($-NH_2$)	*192,411*	*220,061*	*247,710*	*275,360*	*210,686*	*238,336*	*265,986*	*293,635*
Experimental Data								
Observed MW (+ GA)	**27,700**	101,400	**129,800**	**157,600**	**176,000**	**202,900**		
Assignment	**$ExbB_1$** ***(27,650)***	$ExbB_4$ $ExbD_2$ *(101,224)*	**$ExbB_4$ $ExbD_1$** ***(128,874)***	**$ExbB_5$ $ExbD_1$** ***(156,523)***	**$ExbB_5$ $ExbD_2$** ***(174,799)*** **or** **$ExbB_4ExbD_2TonB_1$** ***(178,918)***	**$ExbB_6$ $ExbD_2$** ***(202,448)*** **or** **$ExbB_5ExbD_2TonB_1$** ***(206,567)***		

7.4. Conclusions

In this work, we studied ExbB oligomers, the ExbB-ExbD complex and the ExbB-ExbD-TonB complex, by high-mass MALDI-MS with chemical cross-linking. We determined the oligomeric state of ExbB to be predominantly $ExbB_6$. The architecture of the ExbB-ExbD complex is $ExbB_5ExbD_2$, perhaps potentially in equilibrium with $ExbB_4ExbD_2$. In the entire TonB complex comprising all three subunits, we proposed the dominant oligomeric state to be $ExbB_5ExbD_2$ or $ExbB_4ExbD_2TonB$, which gave a reasonably small mass error. Two reasons were proposed to explain the diverse oligomeric state observed in the MALDI mass spectra: 1) chemical cross-linking might be not sufficient to cross-link all species, 2) the TonB complex may not be very stable.

Chapter 8. Study of ABCG2 via Chemical Cross-linking with High-mass MALDI-MS

Abstract

In our earlier studies, we have presented a straightforward approach to study membrane proteins with high-mass MALDI-MS. Combined with chemical cross-linking, this methodology could be used to determine the stoichiometry of integral membrane protein complexes. ABCG2 (breast cancer resistance protein), one of the three human ATP binding cassette transporters, is functionally capable of exporting a diverse range of substrates from cells. Although the minimal functional unit of ABCG2 is widely accepted to be a homodimer, higher order oligomers, i.e., a tetramer of dimers, have been observed in a study carried out by cryonegative stain electron microscopy. In this chapter, we investigate the oligomeric state of ABCG2 via high-mass MALDI-MS, combined with chemical cross-linking. Two identical subunits, which are connected via intermolecular disulfide bond(s), form ABCG2. For the first time, we observed ABCG2, which was prepared imbedded in liposomes, to be a homodimer, based on MALDI-MS data.

8.1. Introduction

Despite decades of research aimed at early diagnosis and efficient treatment of cancer, the efficiency of cancer chemotherapy is limited by the capability of cells to develop resistance to chemotherapeutic agents. This ability, which is called multi-drug resistance (MDR), is often mediated by over-expression of ATP-binding cassette (ABC) transporters. (Kerr et al., 2010; Xu, 2004) Among the 48 human ABC transporters, there are three transporters most often associated with MDR: ABCB1 (P-glycoprotein), ABCC1 (multidrug resistance protein-1) and ABCG2 (breast cancer resistance protein, BCRP). Among these three ABC transporters, ABCB1 and ABCC1 are believed to be monomeric proteins. (Higgins, 1992; McDevitt et al., 2006) The situation regarding ABCG2 is more complicated. The cDNA encodes a 655 amino acid protein, which comprises a single N-terminal nucleotide binding domain (NBD) and a single c-terminal transmembrane domain (TMD) (Figure 1). (Polgar et al., 2008; Rosenberg et al., 2010) It has been widely accepted that a functional ABC transporter requires two MSDs and two NBDs to form the central substrate translocation pathway. (Kage et al., 2002a; Litman et al., 2002; Ni et al., 2010; Ozvegy et al., 2001) Therefore, ABCG2 is considered to be a half-transporter. Consequently, ABCG2 is believed to be at least a homodimer, or possibly oligomer complex in order to be biologically active.

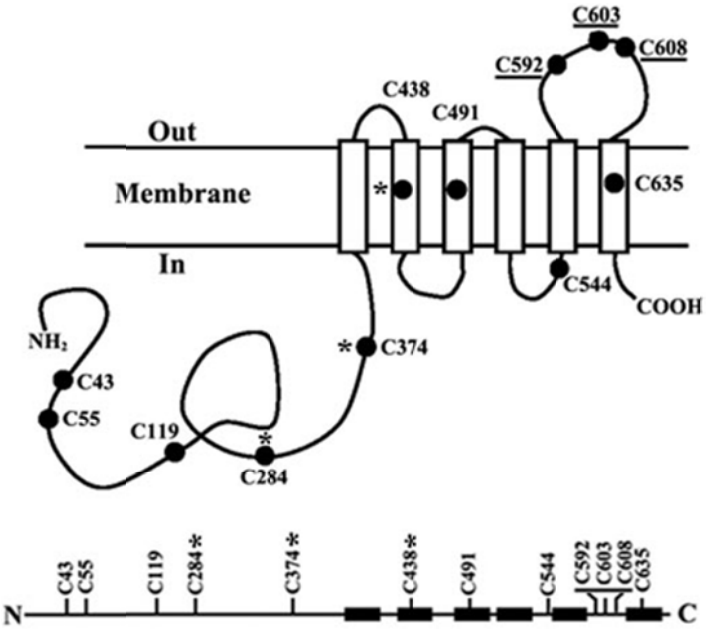


Figure 1. Schematic illustration of the protein structure of human ABCG2. Produced with permission from (Liu et al., 2008). Copyright © American Society for Pharmacology and Experimental Therapeutics.

An earlier study found that ABCG2 migrated as a 70 kDa band in SDS-PAGE under reducing condition, but a 140 kDa complex without reducing agents. (Henriksen et al., 2002b) This suggested that ABCG2 could form a disulfide-bridged homodimer. Immunoblot analysis performed under non-reducing conditions and after cross-linking demonstrated a molecular weight shift from 72 kDa to several bands of 180 kDa and higher molecular weight. (Litman et al., 2002) Chimeric fusion proteins containing two ABCG2 subunits joined either with or without a flexible linker peptide were expressed in the plasma membrane and maintained drug transport activity. (Bhatia et al., 2005) Although the minimal functional unit of ABCG2 is widely accepted to be a homodimer, the presence of higher order oligomers is also possible. The SDS-PAGE of non-reduced ABCG2 or chemically cross-linked AGCG2 presents several diffuse bands above the molecular weight of the expected homodimer. (Kage et al., 2002b) With sucrose density gradient sedimentation and non-denaturing gel electrophoresis, the formation of homotetramer was observed for detergent solubilized ABCG2. (Xu, 2004) Also, electron microscopy (EM) analysis of ABCG2 in detergent solutions identified an octameric state. (McDevitt et al., 2006)

Recently, studies carried out by cryo-EM analysis of 2D crystals indicated that ABCG2 could form a tetrameric complex in one unit cell of the projection maps. (Rosenberg et al., 2010) Furthermore, using fluorescence resonance energy transfer (FRET) analysis of CFP/YFP tagged ABCG2 in whole cells, the oligomeric ABCG2 was recorded without the exact size of the complex.(Ni et al., 2010) At present, the role of oligomerization in its function is not clear. It has been suggested that the function of ABCG2 could be regulated by the dynamic association/dissociation of ABCG2 monomers in high oligomeric state(s) via protein and protein interactions. (Henriksen et al., 2005)

As the extraction and separation methods used in the above studies might lead to the detection of oligomers that do not represent any physiologically relevant form, a clear answer about the oligomeric state of ABCG2 is still missing. Mass spectrometry (MS) is increasingly used as a bioanalytical tool for structural studies. We have presented a standard matrix-assisted laser desorption/ionization (MALDI) protocol, in combination with a high-mass detector that allows straightforward measurements of integral membrane proteins and their complexes. (Chen et al., 2013a; Kerr et al., 2010) Besides giving high mass accuracy, high-mass MALDI-MS can be used to determine the subunit stoichiometry of integral membrane protein complexes. (Chen et al., 2013a; Higgins, 1992)

Here, we studied the oligomeric state of ABCG2, and its reduced form, by high-mass MALDI-MS with chemical cross-linking via glutaraldehyde. The results suggest that ABCG2 is a cysteine-linked homodimer and non-covalent interactions are found in the ABCG2 homodimer. Chemical cross-linking of ABCG2 did not yield any higher oligomeric state of ABCG2. We further investigated ABCG2, which was prepared in liposome, via high-mass MALDI-MS.

8.2. Experimental

8.2.1. Protein Expression, Solubilization and Purification

Experiments were done by Jungin Kim, Institute of Molecular Biology and Biophysics, ETH Zürich.

8.2.2. Chemical Cross-linking and DTT Reducing

DTT (20 mM) was introduced into the protein in a volume ratio of 1/10 (v/v). The protein samples, with and without DTT, were mixed with a 10% glutaraldehyde (GA) solution in a 10/1 (v/v) ratio for 10 mins at room temperature. The solution was further diluted with the original protein buffer solution or water before mass spectrometric analysis.

8.2.3. Experiment Materials and Mass Spectrometry

Experimental materials are prepared as described previously and mass spectrometry method was the same as previously carried out. (Chen et al., 2013a; Polgar et al., 2008)

8.3. Results and Discussion

8.3.1. Cystine-linked dimerization of ABCG2

In SDS-PAGE (Figure 2), ABCG2, without TCEP, results in several bands, including a band with the molecular weight lower than 130 kDa and some unclear bands with the molecular weight even above 250 kDa. We applied MALDI to ABCG2 directly (Figure 3 (a)). A signal corresponding to the dimer of ABCG2 (m/z = 141,600) was readily observed. The detection of the ABCG2 dimer by high-mass MALDI-MS without any cross-linking suggests that the two subunits of ABCG2 have been covalently bound. After introducing DTT into ABCG2, which reduces the disulfide bonds, the ABCG2 monomer at m/z of 70,200 (theoretical MW= 70120.7 Da, mass error 0.1 %) as the dominant species was detected (Figure 3(b)). The observation of the ABCG2 monomer instead of the dimer after applying the reducing agent DTT, indicates that intermolecular disulfide bridge(s) stabilize the ABCG2 dimer.

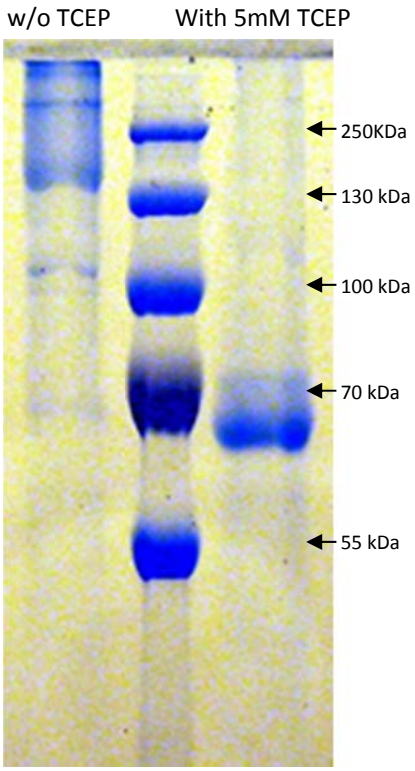


Figure 2. SDS-PAGE of ABCG2 and its reduced form.

Next, we tried to understand whether the high oligomeric state of ABCG2, i.e., tetramer of dimer, which was previously observed by cryonegative stain electron microscopy really exists. In our earlier studies, we have shown that the combination of chemical cross-linking and high-mass MALDI-MS allowed the detection of membrane protein complexes. (Chen et al., 2013a; Kage et al., 2002a; Litman et al., 2002; Ozvegy et al., 2001) Glutaraldehyde, one of the most effective protein cross-linking reagents, was found to stabilize membrane protein complexes. (Chen et al., 2013a; Kage et al., 2002b)

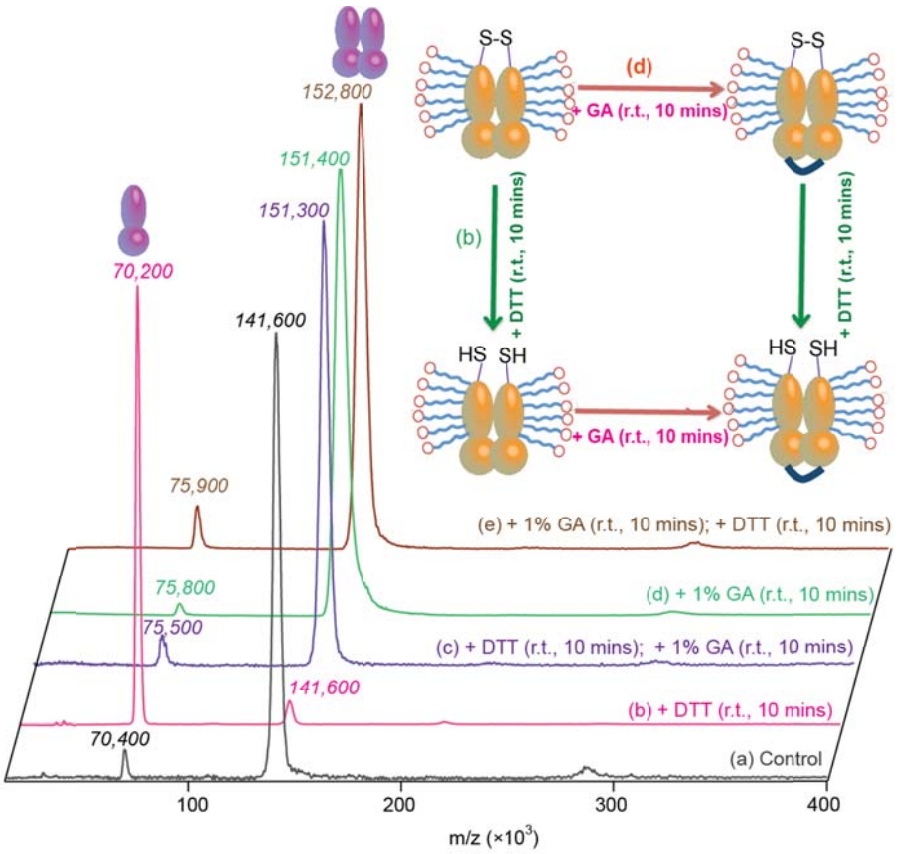


Figure 3. High mass MALDI mass spectra of ABCG2 alone (a), ABCG2 with DTT (b), reduced ABCG2 with GA (c), ABCG2 with GA (d) and cross-linked ABCG2 with DTT (e). The experiment procedure is illustrated in the inset on the right.

In Figure 3(d), a major peak corresponding to the dimer of ABCG2 (*m/z* 151,400) is observed after cross-linking. The incubation of ABCG2 with glutaraldehyde resulted only in a 6.9 % increase of MW, but not the observation of any higher oligomeric states. A molecular weight increase, but not the formation of higher oligomeric state, was previously also observed before for Cdr1p before, which is a monomeric membrane protein. (Chen et al., 2013a; Litman et al., 2002) Because cross-linking chemistry acts only on specific complexes, we believe that ABCG2 is a cysteine-linked homodimer.

We carried out chemical cross-linking on the reduced form of ABCG2. In Figure 3(c), we observed the ABCG2 dimer (*m/z* 151,300) as the dominant species. The observation of ABCG2 dimer suggests that the reduction reaction, to some extent, does not denature ABCG2. Moreover, non-covalent interactions in ABCG2 could also contribute to the formation of ABCG2 dimer, which further allows the

reduced ABCG2 to be stabilized by chemical cross-linking. After chemical cross-linking ABCG2 with glutaraldehyde for 10 mins, we added DTT to break the disulfide bridge(s). The result is shown in Figure 3(e). We observed a major peak corresponding to the dimer of ABCG2 (*m/z* 152,800). The 1 % increase of the MW compared with that in Figure 3(d) is due to the extra 10 mins incubation time with glutaraldehyde.

As shown in Figure 1, three cysteine residues in the extracellular loop are close to each other, i.e., Cys-592, Cys-603 and Cys-608. (Bhatia et al., 2005; Henriksen et al., 2005) They could either form intermolecular disulfide bridge(s), or intramolecular disulfide bridge(s). With the current data, we cannot identify the specific cysteine residues that contribute to the intermolecular disulfide bridges. Site-directed mutagenesis, combined with high-mass MALDI-MS and chemical cross-linking, could be an approach to pinpoint the cysteine residues, that are responsible for intermolecular disulfide bridge formation.

8.3.2. Detection of ABCG2 in Liposomes

Liposomes have been shown to be an excellent tool for investigating the function of membrane proteins. Proteoliposomes consist of a self closed phospholipid bilayer, where the purified membrane protein is incorporated. (Kage et al., 2002b; Rigaud & Lévy, 2003) However, the ability to investigate membrane proteins in liposomes has long been limited. One of the limiting factors is the lack of analytical method to identify the quality of the membrane protein after its incorporation into liposomes. Here, we applied high-mass MALDI-MS to analyze ABCG2 in liposome for the first time. In Figure 4, we directly observed the ABCG2 dimer at *m/z* 145,800 as the dominant species. The peak at *m/z* 72,700 (theoretical MW= 70120.7 Da, mass error 3.7 %) correspond to the doubly charged ABCG2 dimer. The high concentration of lipids in liposome could be the reason for the relatively high mass error. Moreover, the high concentration of analytes (around 10 μM) and the high laser energy promote the formation of the doubly charged ABCG2 dimer in MALDI. These two factors also led to the observation of cluster ions, such as the species at *m/z* 295,200. The detection of ABCG2 in liposome by high-mass MALDI-MS without any further purification

step(s) suggests that high-mass MALDI-MS has excellent potential for studying integral membrane proteins.

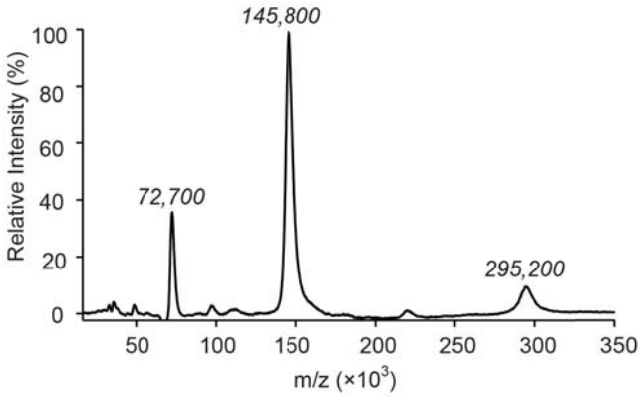


Figure 4. High-mass MALDI-MS spectrum of ABCG2 in liposomes.

8.4. Conclusions

In summary, we identified the oligomeric structure of ABCG2 with standard high-mass MALDI-MS. By introducing glutaraldehyde for chemical cross-linking or DTT for reduced the disulfide bridge, we confirmed that ABCG2 is a cysteine-linked homodimer. The chemical cross-linking studies suggested that there is no higher order oligomeric state of ABCG2 existed in detergent micelles. We detected ABCG2, which was prepared in liposomes, for the first time with high-mass MALDI-MS. Given the importance of liposomes in function studies of membrane proteins, high-mass MALDI-MS could be a valuable tool for these studies.

Chapter 9. Conclusions and Outlook

This thesis began with a overview of non-covalent interactions in biomolecules, which play an important role in molecular recognition processes, and presented current approaches to study these interactions. Mass spectrometry (MS) shows great potential in the study of non-covalent interactions, in particular since the introduction of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). Nowadays, native-ESI MS is heavily used in studying non-covalent interactions and to understand the architecture of biomolecular complexes. However MALDI-MS is also becoming increasingly useful. There are two major approaches to study non-covalent interactions with MALDI-MS. In the first approach, different experimental and instrumental parameters can be fine-tuned such as applying non-acidic matrices and collecting first-shot spectra. In the second approach, interacting species are stabilized by chemical crosslinking. The main objective of this thesis was to investigate the application of MALDI-MS in some challenging systems.

We started with the *E. Coli* single-stranded DNA binding protein (SSB), which selectively binds single-stranded (ss) DNA. Different binding modes have previously been observed in SSB•ssDNA complexes, due to the four potential binding sites of SSB. SSB forms a stable homotetramer in solution, but only the monomeric species can be detected with standard MALDI-MS. With chemical cross-linking, the quaternary structure of SSB is conserved and the tetramer was observed. Interestingly, ssDNA was found to function as a stabilizer to conserve the quaternary structure of SSB during MALDI process, as evidenced by the detection of a SSB•ssDNA complex at *m/z* 94,200 even in the absence of chemical cross-linking. This observation could be attributed to electrostatic interactions that are enhanced in the gas phase. The key factor affecting the stoichiometry of the SSB•ssDNA complex is how ssDNA binds to SSB, rather than the protein-to-DNA ratio. We found that the stability of the SSB•ssDNA complex in MALDI strongly depends on how ssDNA binds to SSB, including the binding oligonucleotides and the stoichiometry of the SSB•ssDNA complex. This study suggests that MALDI could be used to study protein-DNA complex, even in the acidic matrix condition, i.e., sinapinic acid.

In the following chapter, we addressed the application of MALDI-MS application to investigate a special protein-oligonucleotide system, i.e., protein-aptamer complexes. Aptamers are generated through an iterative process called SELEX in vitro. They bind to a broad range of targets including metal ions, small organic molecules, peptides, proteins, cancer cells and viruses with high specificity and selectivity. Owing to their unique binding properties, aptamers have been the subject of immense of research in the past 10 years. We report on the direct detection of protein•aptamer complexes by MALDI-MS. We report for the first time that high-mass MALDI-MS with a proper matrix, ATT, is able to preserve the non-covalent interactions in protein-aptamer complexes without any chemical cross-linking or other stabilization. This was observed for three systems, thrombin•TBA15/29, PDGF-AA/AB/BB•2Apt-35 and Lysozyme•LBA. Systematic variation of experimental parameters, including the matrix, laser energy and the sample concentration confirmed the observation of the complex to be due to specific non-covalent interactions. The peak intensity of the complexes compared with these of the protein or aptamer subunits in the mass spectra agrees with the solution-phase binding constants, indicating that the non-covalent interaction strength in solution is reflected in the MALDI mass spectra. Thus, these studies open up new mass spectrometry-based avenues for studying protein-aptamer interactions, First, our method provides a fast, specific, sensitive, and single-step assay of aptamer-protein interactions, distinguishing molecular variants of proteins. It provides a complementary structure characterization of the complex without the need for chemical cross-linking.

Starting with Chapter 4, we applied high-mass MALDI-MS to study protein-protein interactions. Chemical cross-linking, combined with mass spectrometry, has previously been applied to map three-dimensional protein structures and protein-protein interactions. Proper choice of the cross-linking agent, including its reactive groups and spacer arm length, is of great importance. Considering the limited knowledge about the reactivity of chemical cross-linkers with proteins, investigated chemical cross-linking from the aspects of the protein stuctures and the cross-linking reagents involved, by using two structurally well-known proteins, glyceraldehyde 3-phosohate dehydrogenase and ribonuclease S. We

conducted chemical cross-linking with a series of homo- and hetero-bifunctional cross-linkers, including BS^3, DSS, $BS(PEG)_5$, $BS(PEG)_9$, sulfo-MBS, PEG_{12}-SPDP and $SM(PEG)_{24}$. The same chemical cross-linker was found to behave quite differently for the two proteins complexes. By comparing the cross-linking efficiency among different chemical cross-linkers, we found that the distances between the target amino acid residues, including the inter- and intra-subunits, play a key role in successfully stabilization of a protein complex with a given cross-linker. Additionally, the functional groups on the chemical cross-linkers also affects the reactivity. *N*-maleimide showed a stronger reaction capability with cysteine than NHS esters with lysine or 2-pyrimidyldithiol with cysteine. This study provides a better understanding of chemical cross-linking in proteins, especially cross-linking efficiency.

In the following, we targeted purified membrane proteins and membrane protein complexes, which have been notoriously challenging to investigate, and required highly specialized buffer conditions, sample preparation methods, and apparatus. In chapter 5, we showed that with standard high-mass MALDI-MS, integral membrane proteins can be detected in a straightforward fashion, and their MW determined with high mass accuracy. The high mass accuracy, i.e., mass error $\leq 0.1\%$, allows the identification of the N-linked glycosylation site for the Cdr1p, an eukaryotic multidrug ABC transporter, without special purification steps. This information is crucial to circumvent problems due to heterogeneous glycosylation in crystallization studies. Combined with chemical cross-linking, such as glutaraldehyde, in the presence of detergent micelles, we could determine the subunit stoichiometries of a series of integral membrane protein complexes, including the homomeric PglK and the heteromeric BtuCD, as well as BtuCDF. With these achievements, we expect MALDI-MS to become a routine method to study integral membrane proteins, the stoichiometry of their complexes, and post-translational modifications, in a straightforward and elegant fashion.

The following three chapters focus on membrane protein complexes. The first question we addressed was whether specific chemical cross-linkers can stabilize membrane protein complexes in detergent micelles. Within this study, we conducted specific chemical cross-linking agents, i.e., NHS-esters, on two

membrane protein complexes, PglK and $BtuC_2D_2$. The cross-linking experiments were carried out with a series of chemical cross-linkers with different spacer arms. We observed that NHS esters could be used to cross-link membrane protein complexes in different detergents, such as DDM and C12E8 used here. The differences of stabilizing capabilities among the cross-linkers applied depend on the protein structure, including the number of lysine residues, and on the amine-amine distances between different subunits. The limited number of lysine residues is the reason for the relatively low stabilization efficiency in $BtuC_2D_2$, but the high stabilization efficiency in PglK. The successful cross-linking of membrane protein complexes in different detergent micelles via NHS esters highlights the possibility to map membrane protein structures by chemical cross-linking combined with mass spectrometry.

In the following, we targeted ExbB oligomers, the ExbB-ExbD complex, and the ExbB-ExbD-TonB complex. We determined the oligomeric state of ExbB to be predominantly $ExbB_6$. The architecture of the ExbB-ExbD complex is $ExbB_5ExbD_2$, perhaps in equilibrium with $ExbB_4ExbD_2$. In the entire TonB complex, we proposed the dominant oligomeric state to be $ExbB_5ExbD_2$ or $ExbB_4ExbD_2$TonB, which gave a reasonably small mass error. Two reasons were proposed to explain the diverse oligomeric state observed in the MALDI mass spectra: 1) chemical cross-linking might be not sufficient to cross-link all species, 2) the TonB complex may not be very stable.

We further identified the oligomeric structure of ABCG2 with standard high-mass MALDI-MS. By introducing glutaraldehyde for chemical cross-linking and DTT (TCEP) to reduce the disulfide bridge, we confirmed that ABCG2 is a cysteine-linked homodimer. No higher order oligomeric state of ABCG2 was observed in MALDI-MS, combined with chemical cross-linking, in detergent micelles. We detected ABCG2, which was prepared in liposomes, for the first time with high-mass MALDI-MS. Given the importance of liposomes in studying membrane protein function, high-mass MALDI-MS could be a valuable tool for these studies.

To overcome the possible dissociation of non-covalent complexes during matrix incorporation and/or the ionization process in MALDI, selecting a proper matrix

or chemical cross-linking agent, in combination with high-mass MALDI-MS, has been shown to be important for studying non-covalent biomolecular complexes. However, there are still some open questions left in this area.

1) Although MALDI-MS has been used to study membrane proteins in detergent micelles, as well as in liposome (a first trial of ABCG2), systematic studies of detergent effect on membrane proteins' signal intensities are still missing. Also, it would be interesting to study membrane proteins in nanodiscs, which are formed from phospholipid bilayers and two molecules of an amphipathic alpha-helical protein. This knowledge could further assist the quantification of membrane proteins by MALDI-MS. Because chemical cross-linkers can be used to stabilize membrane protein complexes, it might be possible to carry out the characterization of membrane protein complex association/dissociation under different conditions.

2) Chemical cross-linking, especially photo cross-linking, can be applied to stabilize protein-DNA complexes. However, the lower cross-linking efficiency of photo cross-linkers, compared with NHS esters, is a major problem. Currently, amino-modified oligonucleotides have been routinely employed in solid support and label attachment chemistries. Thus, it is possible to study complexes of proteins and oligonucleotides, which have been amino-modified, via cross-linking using NHS esters and MALDI-MS.

3) Chemical cross-linking, combined with MALDI-MS, has been proven to provide information on relative binding constants. It will be useful to design a platform to rapidly screen potential binding candidates for a target protein or DNA.

4) Although we have applied MALDI-MS in several non-covalent interactions systems, such as protein-protein complexes, protein-DNA complexes and membrane protein complexes, the study of non-covalent interactions between proteins and oligosaccharides are far behind. For instance, many proteins of widely differing functionality and structure are capable of

binding heparin. It would be interesting and challenging to study the structure of protein-heparin complexes by MALDI-MS.

References

Alaimo, C., Catrein, I., Morf, L., Marolda, C. L., Callewaert, N., Valvano, M. A., et al. (2006). Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. The EMBO Journal, 25(5), 967–976.

Balamurugan, S., Obubuafo, A., McCarley, R. L., Soper, S. A., & Spivak, D. A. (2008). Effect of linker structure on surface density of aptamer monolayers and their corresponding protein binding efficiency. Analytical Chemistry, 80(24), 9630–9634.

Bamrungsap, S., Shukoor, M. I., Chen, T., Sefah, K., & Tan, W. (2011). Detection of lysozyme magnetic relaxation switches based on aptamer-functionalized superparamagnetic nanoparticles. Analytical Chemistry, 83(20), 7795–7799.

Barrera, N. P., Di Bartolo, N., Booth, P. J., & Robinson, C. V. (2008). Micelles Protect Membrane Complexes from Solution to Vacuum. Science, 321(5886), 243–246.

Barrera, N. P., Isaacson, S. C., Zhou, M., Bavro, V. N., Welch, A., Schaedler, T. A., et al. (2009). Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions. Nature Methods, 6(8), 585–587.

Barrera, N. P., & Robinson, C. V. (2011). Advances in the Mass Spectrometry of Membrane Proteins: From Individual Proteins to Intact Complexes. *Annual Review Biochemistry*, *80*(1), 247–271.

Bechara, C., Bolbach, G., Bazzaco, P., Sharma, K. S., Durand, G., Popot, J.-L., Zito, F., & Sagan, S., (2012). Maldi-tof mass spectrometry analysis of amphipol-trapped membrane proteins. *Analytical Chemistry, 84*(14), 6128–6135.

Belgacem, O., Buchacher, A., Pock, K., Josic, D., Sutton, C., Rizzi, A., & Allmaier, G. (2002). Molecular mass determination of plasma-derived glycoproteins by ultraviolet matrix-assisted laser desorption/ionization time-of-flight mass spectrometry with internal calibration. *Journal of Mass Spectrometry, 37*(11), 1118–1130.

Beretta, G. L., Benedetti, V., Cossa, G., Assaraf, Y. G. A., Bram, E., Gatti, L., Corna, E., Carenini, N., Colangelo, D., Howell, S. B., Zunino, F., & Perego, P. (2010). Increased levels and defective glycosylation of MRPs in ovarian carcinoma cells resistant to oxaliplatin. Biochemical Pharmacology, 79(8), 1108–1117.

Bhatia, A., Schäfer, H.-J., & Hrycyna, C. A. (2005). Oligomerization of the human ABC transporter ABCG2: evaluation of the native protein and chimeric dimers. Biochemistry, 44(32), 10893–10904.

Bich, C., Bovet, C., Rochel, N., Peluso-Iltis, C., Panagiotidis, A., Nazabal, A., Moras, D. & Zenobi, R. (2010). Detection of nucleic acid-nuclear hormone receptor complexes with mass spectrometry. Journal of the American Society for Mass Spectrometry, 21(4), 635–645.

Bich, C., & Zenobi, R. (2009). Mass spectrometry of large complexes. Current Opinion in Structural Biology, 19(5), 632–639.

Bolbach, G. (2005). Matrix-assisted laser desorption/ionization analysis of non-covalent complexes: fundamentals and applications. *Current Pharmaceutical Design*, *11*(20), 2535–2557.

Borchers, C., & Tomer, K B. (1999). Characterization of the noncovalent complex of human immunodeficiency virus glycoprotein 120 with its cellular receptor CD4 by matrix-assisted laser desorption/ionization mass spectrometry. *Biochemistry*, *38*(36), 11734–11740.

Bovet, C., Ruff, M., Eiler, S., Granger, F., Wenzel, R., Nazabal, A., Moras, D. & Zenobi, R., (2008). Monitoring ligand modulation of protein-protein interactions by mass spectrometry: estrogen receptor alpha-SRC1. *Analytical Chemistry*, *80*(20), 7833–7839.

Bragg, P. D., & Hou, C. (1975). Subunit composition, function, and spatial arrangement in the Ca2+- and Mg2+-activated adenosine triphosphatases of Escherichia coli and Salmonella typhimurium. *Archives of Biochemistry and Biophysics, 167*(1), 311–321.

Braun, V. (1995). Energy-coupled transport and signal transduction through the Gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. *FEMS Microbiology Reviews, 16*(4), 295–307.

Braun, V. (2003). Iron uptake by Escherichia coli. *Frontiers in Bioscience, 8*, s1409–21.

Braun, V., & Braun, M. (2002). Iron transport and signaling in Escherichia coli. FEBS Letters, 529(1), 78–85.

Brewer, C. F., & Riehm, J. P. (1967). Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins. *Analytical Biochemistry*, *18*(2), 248–255.

Brunner, J. (1992). New photolabeling and crosslinking methods. Annual Review of Biochemistry, 62, 483–514.

Bujalowski, W., & Lohman, T. M. (1986a). Escherichia coli single-strand binding protein forms multiple, distinct complexes with single-stranded DNA. Biochemistry, 25(24), 7799–7802.

Bujalowski, W., & Lohman, T. M. (1986b). Escherichia coli single-strand binding protein forms multiple, distinct complexes with single-stranded DNA. *Biochemistry, 25*(24), 7799–7802.

Bunka, D. H. J., & Stockley, P. G. (2006). Aptamers come of age - at last. *Nature Reviews Microbiology, 4*(8), 588–596.

Cadene, M., & Chait, B. T. (2000). A Robust, Detergent-Friendly Method for Mass Spectrometric Analysis of Integral Membrane Proteins. *Analytical Chemistry, 72*(22), 5655–5658.

Carlsson, J., Drevin, H., & Axén, R. (1978). Protein thiolation and reversible protein-protein conjugation. N-Succinimidyl 3-(2-pyridyldithio)propionate, a new heterobifunctional reagent. *The Biochemical Journal, 173*(3), 723–737.

Chase, J. W., & Williams, K. R. (1986). Single-Stranded DNA Binding Proteins Required for DNA Replication. *Annual Review of Biochemistry, 55*(1), 103–136.

Chen, F., Gerber, S., Heuser, K., Korkhov, V. M., Lizak, C., Mireku, S., Locher, K. P., & Zenobi, R. (2013a). High-Mass Matrix-Assisted Laser Desorption Ionization-Mass Spectrometry of Integral Membrane Proteins and Their Complexes. *Analytical Chemistry, 85*(7), 3483–3488.

Chen, F., Gulbakan, B., & Zenobi, R. (2013b). Direct access to aptamer–protein complexes via MALDI-MS. *Chemical Science*, *4*(10), 4071-4078.

Chen, F., Mädler, S., Weidmann, S., & Zenobi, R. (2012). MALDI-MS detection of noncovalent interactions of single stranded DNA with Escherichia coli single-stranded DNA-binding protein. *Journal of Mass Spectrometry, 47*(5), 560–566.

Chen, F., Nielsen, S., & Zenobi, R. (2013c). Understanding chemical reactivity for homo- and heterobifunctional protein cross-linking agents. *Journal of Mass Spectrometry, 48*(7), 807–812.

Chen, J., Smith, D. L., & Griep, M. A. (1998). The role of the 6 lysines and the terminal amine of Escherichia coli single-strand binding protein in its binding of single-stranded DNA. *Protein Science, 7*(8), 1781–1788.

Chen, Z. A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., et al. (2010). Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry. *The EMBO Journal*, *29*(4), 717–726.

Chrysogelos, S., & Griffith, J. (1982). Escherichia coli single-strand binding protein organizes single-stranded DNA in nucleosome-like units. *Proceedings of the National Academy of Sciences of the United States of America, 79*(19), 5803–5807.

Cohen, L. R. H., Strupat, K. & Hillenkamp, F. (1997). Analysis of quaternary protein ensembles by matrix assisted laser desorption/ionization mass spectrometry. *Journal of American Society for Mass Spectrometry*, *8*(10), 1046–1052.

Cohen, S. L., Padovan, J. C., & Chait, B. T. (2000). Mass Spectrometric Analysis of Mercury Incorporation into Proteins for X-ray Diffraction Phase Determination. *Analytical Chemistry, 72(3), 574–579.*

Cole, J. R., Dick, L. W., Morgan, E. J., & McGown, L. B. (2007). Affinity capture and detection of immunoglobulin E in human serum using an aptamer-modified surface in matrix-assisted laser desorption/ionization mass spectrometry. *Analytical Chemistry, 79*(1), 273–279. doi:10.1021/ac061256b

Cole, R. B. (2011). *Electrospray and MALDI Mass Spectrometry*. Fundamentals, Instrumentation, Practicalities, and Biological Applications. John Wiley & Sons.

Cowan, J. A. (2000). Recognition of a cognate RNA aptamer by neomycin B: quantitative evaluation of hydrogen bonding and electrostatic interactions. *Nucleic Acids Research, 28*(15), 2935–2942.

Dawson, R. J. P., & Locher, K. P. (2006). Structure of a bacterial multidrug ABC transporter. *Nature, 443*(7108), 180–185.

Dick, L. W., & McGown, L. B. (2004). Aptamer-Enhanced Laser Desorption/Ionization for Affinity Mass Spectrometry. *Analytical Chemistry, 76*(11), 3037–3041.

Dole, M. (1968). Molecular Beams of Macroions. *Journal of Chemical Physics*, *49*(5), 2240.

Dongre, A. R., Somogyi, A. & Wysocki, V. H. (1996) Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *Journal of Mass Spectrometry*, *31*, 339-350.

Draheim, V., Reichel, A., Weitschies, W., & Moenning, U. (2010). N-glycosylation of ABC transporters is associated with functional activity in sandwich-cultured rat hepatocytes. *European Journal of Pharmaceutical Sciences, 41*(2), 201–209.

Ehring, H., Karas, M. & Hillenkamp. F. (1992) Role of photoionization and photochemistry in ionization processes of organic molecules and relevance for matrix‐assisted laser desorption lonization mass *spectrometry. Organic Mass. Spectrometry,* 27, 472-480.

Emmett, M. R., & Caprioli, R. M. (1994). Micro-electrospray mass spectrometry: Ultra-high-sensitivity analysis of peptides and proteins. *Journal of American Society for Mass Spectrometry*, *5*(7), 605–613.

Erba, E. B., & Zenobi, R. (2011). Mass spectrometric studies of dissociation constants of noncovalent complexes. Annual Reports on the Progress of Chemistry, Section C: Physical Chemistry, 107(0), 199–228.

Fang, X., & Tan, W. (2010). Aptamers generated from cell-SELEX for molecular medicine: a chemical biology approach. *Accounts of Chemical Research, 43*(1), 48–57.

Fang, X., Sen, A., Vicens, M., & Tan, W. (2003). Synthetic DNA aptamers to detect protein molecular variants in a high-throughput fluorescence quenching assay. *Chembiochem, 4*(9), 829–834.

Farmer, T. B., & Caprioli, R. M. (1991). Assessing the multimeric states of proteins: studies using laser desorption mass spectrometry. *Biological Mass Spectrometry, 20*(12), 796–800.

Farmer, T. B., & Caprioli, R. M. (1998). Determination of protein-protein interactions by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry*, 33(8), 697–704.

Farnham, P. J. (2009). Insights from genomic profiling of transcription factors. *Nature Reviews Genetics, 10*(9), 605–616.

Fernandez de la Mora, J. (2000). Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytical Chimica Acta*, *406*(1), 93–104.

Forterre, P. (2002) The origin of DNA genomes and DNA replication proteins. Current Opinion in Microbiology, 5(5), 525-532.

Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gústafsdóttir, S. M., Östman, A., & Landegren, U. (2002). Protein detection using proximity-dependent DNA ligation assays. *Nature Biotechnology, 20*(5), 473–477.

Fu, H. (2004). *Protein-Protein Interactions*. Methods and Applications. Springer.

Gabelica, V., Galic, N., Rosu, F., Houssier, C., & De Pauw, E. (2003). Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. *Journal of Mass Spectrometry, 38*(5), 491–501.

Gale, D. C., & Smith, R. D. (1993). Small volume and low flow-rate electrospray lonization mass spectrometry of aqueous samples. *Rapid Commununication in Mass Spectrometry*, *7*(11), 1017–1021.

Gerber, S. M. (2008). Structural and functional studies of TonB complexes and ABC transporters. *Diss ETH* No. 17918.

Geyer, H., Geyer, R., & Pingoud, V. (2004). A novel strategy for the identification of protein-DNA contacts by photocrosslinking and mass spectrometry. *Nucleic Acids Research*, *32*(16), e132.

Gilchrist, T. L., & Rees, C. W. (1969). *Carbènes, nitrènes and arynes*. Nelson.

Glocker, M. O., Bauer, S. H., Kast, J., Volz, J., & Przybylski, M. (1996). Characterization of specific noncovalent protein complexes by UV matrix-assisted laser desorption ionization mass spectrometry. *Journal of Mass Spectrometry, 31*(11), 1221–1227.

Golden, M. C., Resing, K. A., Koch, T. H., Collins, B. D., & Willis, M. C. (2008). Mass spectral characterization of a protein-nucleic acid photocrosslink. *Protein Science*, *8*(12), 2806–2812.

Gordiyenko, Y., Videler, H., Zhou, M., McKay, A. R., Fucini, P., Biegel, E., Müller, V., & Robinson, C. V. (2010). Mass spectrometry defines the stoichiometry of ribosomal stalk complexes across the phylogenetic tree. *Molecular & Cellular Proteomics, 9*(8), 1774–1783.

Gorin, G., Martic, P. A., & Doughty, G. (1966). Kinetics of the reaction of N-ethylmaleimide with cysteine and some congeners. *Archives of Biochemistry and Biophysics, 115*(3), 593–597.

Green, L. S., Jellinek, D., Jenison, R., Ostman, A., Heldin, C. H., & Janjic, N. (1996). Inhibitory DNA ligands to platelet-derived growth factor B-chain. *Biochemistry, 35*(45), 14413–14424.

Gregogy, J. D. (1955). The Stability of N-Ethylmaleimide and Its Reaction with Sulfhydryl Groups.

Journal of the American Chemical Society, 77(14), 3922–3923.

Greig, M. J., Gaus, H., Cummins, L. L., Sasmor, H., & Griffey, R. H. (1995). Measurement of Macromolecular Binding Using Electrospray Mass Spectrometry. Determination of Dissociation Constants for Oligonucleotide: Serum Albumin Complexes. *Journal of the American Chemical Society*, 117(43), 10765–10766.

Grey, A. C., Chaurand, P., Caprioli, R. M., & Schey, K. L. (2009). MALDI Imaging Mass Spectrometry of Integral Membrane Proteins from Ocular Lens and Retinal Tissue. *Journal of Proteome Research, 8*(7), 3278–3283.

Gulbakan, B., Yasun, E., Shukoor, M. I., Zhu, Z., You, M., Tan, X., Sanchez, H., Powell, D. H., Dai, H., & Tan, H., (2010). A Dual Platform for Selective Analyte Enrichment and Ionization in Mass Spectrometry Using Aptamer-Conjugated Graphene Oxide. *Journal of the American Chemical Society, 132*(49), 17408–17410.

Hathaway, N. A., Bell, O., Hodges, C., Miller, E. L., Neel, D. S., & Crabtree, G. R. (2012). Dynamics and memory of heterochromatin in living cells. *Cell, 149*(7), 1447–1460.

Heck, A. J. R. (2008). Native mass spectrometry: a bridge between interactomics and structural biology. *Natural Methods, 5*(11), 927–933.

Heitz, J. R., Anderson, C. D., & Anderson, B. M. (1968). Inactivation of yeast alcohol dehydrogenase by N-alkylmaleimides. *Archives of Biochemistry and Biophysics, 127*(1), 627–636.

Heldin, C. H. (1992). Structural and functional studies on platelet-derived growth factor. *The EMBO Journal, 11*(12), 4251-4259.

Henriksen, U., Fog, J. U., Litman, T., & Gether, U. (2005). Identification of intra- and intermolecular disulfide bridges in the multidrug resistance transporter ABCG2. *The Journal of Biological Chemistry, 280*(44), 36926–36934.

Hermanson, G. T. (2013). *Bioconjugate Techniques.* Academic Press.

Hernández, H., & Robinson, C. V. (2007). Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry *Natural Protocol, 2*(3), 715–726.

Heuvel, R. H., & Heck A. J. (2004). Native protein mass spectrometry: from intact oligomers to functional machineries. *Current Opinion in Chemical Biology, 8*(5), 519-526.

Hillenkamp, F. (1998). New Methods for the Study of Biomolecular Complexes. *NATO ASI Series C: Mathematical and Physical Science*, *510*, 181-191

Hillenkamp, F., Karas, M., Beavis, R. C. & Chait, B. T. (1991) Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Analytical Chemistry, 63,* 1193-1203*.*

Hillenkamp, F., Karas, M., Holtkamp, D. & Klusener, P. (1986) Energy Deposition in Ultraviolet Laser Desorption Mass Spectrometry of Biomolecules. *International Journal of Mass Spectrometry, 69*(3), 265–276.

Hillenkamp, F., Nazabal, A., Roehling, U., & Wenzel, R., (2008, January 4). A detector device for high mass ion detection, a method for analyzing ions. WO2009086642A1. US Patent Office.

Hillenkamp, F. & Peter-Katalinic, J. (2007) MALDI MS - a practical guide to instrumentation, methods, and application. Wiley-VCH, Weinheim.

Hilton, G. R., & Benesch, J. L. P. (2012). Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of the Royal*

Society Interface, 9(70), 801–816.

Higgins, C. F. (1992). Abc Transporters - From Microorganisms to Man. *Annual Review of Cell Biology, 8*, 67–113.

Higgs, P. I., Larsen, R. A., & Postle, K. (2002) Quantification of known components of the Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA. *Molecular Microbiology, 44* (1), 271-281.

Hippel, von, P. H. (2007). From "simple" DNA-protein interactions to the macromolecular machines of gene .expression. *Annual Review of Biophysics and Biomolecular Structure 36*, 79–105.

Hobza, P., & Müller-Dethlefs, K. (2010). *Non-covalent Interactions*. Theory and Experiment (p. 225). Royal Society of Chemistry.

Hogan, C. J. J., Carroll, J. A., Rohrs, H. W., Biswas, P., & Gross, M. L. (2009). Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Analytical Chemistry*, *81*(1), 369–377.

Hofstadler, S. A., & Griffey, R. H. (2001). Analysis of Noncovalent Complexes of DNA and RNA by Mass Spectrometry. *Chemical Reviews*, *101*(2), 377–390.

Hollenstein, K., Frei, D. C., & Locher, K. P. (2007). Structure of an ABC transporter in complex with its binding protein. *Nature, 446*(7132), 213–216.

Horneffer, V., Strupat, K., & Hillenkamp, F. (2006). Localization of noncovalent complexes in MALDI-preparations by CLSM. *Journal of American Society for Mass Spectrometry, 17*(11), 1599–1604.

Huang, P.-J. J., & Liu, J. (2012). DNA-Length-Dependent Fluorescence Signaling on Graphene Oxide Surface. Small, 8(7), 977–983.

Huntington, J. A. (2005). Molecular recognition mechanisms of thrombin. Journal of Thrombosis and Haemostasis, 3(8), 1861–1872.

Hvorup, R. N., Goetz, B. A., Niederer, M., Hollenstein, K., Perozo, E., & Locher, K. P. (2007). Asymmetry in the Structure of the ABC Transporter-Binding Protein Complex BtuCD-BtuF. S*cience, 317*(5843), 1387–1390.

Iribarne, J. V. (1976). On the evaporation of small ions from charged droplets. *Journal of Chemical Physics, 64*(6), 2287–2294.

Ellington, A. D., & Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature, 346*, 818-822.

Jayasena, S. D. (1999). Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clinical Chemistry, 45*(9), 1628–1650.

Jensen, O. N., Barofsky, D. F., Young, M. C., Hippel, von, P. H., Swenson, S., & Seifried, S. E. (1993). Direct observation of UV-crosslinked protein-nucleic acid complexes by matrix-assisted laser desorption ionization mass spectrometry. Rapid Communications in Mass Spectrometry, 7(6), 496–501.

Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. Science, 293(5532), 1074–1080.

Jespersen, S., Niessen, W. M., Tjaden, U. R., & van der Greef, J. (1998). Basic matrices in the analysis of non-covalent complexes by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry, 33*(11), 1088–1093.

Jiang, H. (2006). *Nanoelectrospray Mass Spectrometry for Molecular Non-covalent Interactions and Proteomics*. ProQuest.

Johnson, J., Okyere, R., Joseph, A., Musier-Forsyth, K., & Kankia, B. (2012). Quadruplex formation as a molecular switch to turn on intrinsically fluorescent nucleotide analogs, *Nucleic Acids Research, 41*(1), 220-228.

Juhasz, P., & Biemann, K. (1994). Mass spectrometric molecular-weight determination of highly acidic compounds of biological significance via their complexes with basic polypeptides. *Proceedings of the National Academy of Sciences of the United States of America, 91*(10), 4333–4337.

Kaddis, C. S., & Loo, J. A. (2007) Native protein MS and ion mobility: large flying proteins with ESI. *Analytical Chemistry, 79*(5), 1778-1784

Kage, K., Tsukahara, S., Sugiyama, T., Asada, S., Ishikawa, E., Tsuruo, T., & Sugimoto, Y. (2002a). Dominant-negative inhibition of breast cancer resistance protein as drug efflux pump through the inhibition of S-S dependent homodimerization. *International Journal of Cancer, 97*(5), 626–630.

Kalkhof, S., & Sinz, A. (2008). Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. *Analytical and Bioanalytical Chemistry, 392*(1-2), 305–312.

Kampfenkel, K., & Braun, V. (1992). Membrane topology of the Escherichia coli ExbD protein. *Journal of Bacteriology, 174*(16), 5485–5487.

Kampfenkel, K., & Braun, V. (1993). Topology of the ExbB protein in the cytoplasmic membrane of Escherichia coli. *The Journal of Biological Chemistry, 268*(8), 6050–6057.

Kang, Y., Feng, K.-J., Chen, J.-W., Jiang, J.-H., Shen, G.-L., & Yu, R.-Q. (2008). Electrochemical detection of thrombin by sandwich approach using antibody and aptamer. *Bioelectrochemistry, 73*(1), 76–81.

Karlsson, M., Hannavy, K., & Higgins, C. F. (1993). ExbB acts as a chaperone-like protein to stabilize TonB in the cytoplasm. *Molecular Microbiology, 8*(2), 389–396.

Karas, M., Bachmann, D. & Hillenkamp, F. (1985). Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Analytical Chemistry, 57 (14),* 2935–2939.

Karas, M., Bahr, U., Ingendoh, A. & Hillenkamp, F. (1989) Laser Desorption/Ionization Mass Spectrometry of Proteins of Mass 100 000 to 250 000 Dalton. *Angewandte Chemie International Edition, 28*(6), 760-761.

Karas, M., Gluckmann, M. & Schafer, J (2000) Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *Journal of Mass Spectrometry*, *35*(1), 1-12.

Karas M, & Kruger, R. (2003) Ion formation in MALDI: the cluster ionization mechanism, *Chemical Reviews*, *103(2)*, 427-439.

Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K., & McLafferty, F. W. (1999). Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *Journal of American Chemistry Society, 121*(4), 806–812.

Keller, B. O., & Li, L. (2006). Three-layer matrix/sample preparation method for MALDI MS analysis of low nanomolar protein samples. *Journal of the American Society for Mass Spectrometry, 17*(6), 780–785.

Keller, K. M., Breeden, M. M., Zhang, J., Ellington, A. D., & Brodbelt, J. S. (2005). Electrospray ionization of nucleic acid aptamer/small molecule complexes for screening aptamer selectivity. *Journal of Mass Spectrometry, 40*(10), 1327–1337.

Kemptner, J., Marchetti-Deschmann, M., Müller, R., Ivens, A., Turecek, P., Schwarz, H. P., & Allmaier, G. (2010). A comparison of nano-electrospray gas-phase electrophoretic mobility macromolecular analysis and matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometry for the characterization of the recombinant coagulation glycoprotein von Willebrand factor. *Rapid Communication in Mass Spectrometry, 24*(6), 761–767.

Kerr, I. D., Jones, P. M., & George, A. M. (2010). Multidrug efflux pumps: The structures of prokaryotic ATP-binding cassette transporter efflux pumps and implications for our understanding of eukaryotic P-glycoproteins and homologues. *FEBS Journal, 277*(3), 550-563.

Kim, Y., Cao, Z., & Tan, W. (2008). Molecular assembly for high-performance bivalent nucleic acid inhibitor. *Proceedings of the National Academy of Sciences, 105*(15), 5664–5669.

Kimura, T., Matsueda, R., Nakagawa, Y., & Kaiser, E. T. (1982). New reagents for the introduction of the thiomethyl group at sulfhydryl residues of proteins with concomitant spectrophotometric titration of the sulfhydryls: methyl 3-nitro-2-pyridyl disulfide and methyl 2-pyridyl disulfide. *Analytical Biochemistry, 122*(2), 274–282.

Kirby, R., Cho, E. J., Gehrke, B., Bayer, T., Park, Y. S., Neikirk, D. P., McDevitt, J. T., & Ellington, A. D. (2004). Aptamer-Based Sensor Arrays for the Detection and Quantitation of Proteins. *Analytical Chemistry, 76*(14), 4066–4075.

Knochenmuss, R. (2006). Ion formation mechanisms in UV-MALDI. *Analyst*, *131*(9), 966-986.

Knochenmuss, R. & Zenobi, R. (2003) MALDI ionization: the role of in-plume processes. *Chemical Reviews*, 103, 441-452.

Knochenmuss, R & Zhigilei, L. V. (2010). Molecular dynamics simulations of MALDI: Laser fluence and pulse with dependence of plume characteristics and consequences for matrix and analyte ionization. *Journal of Mass Spectrometry, 45*, 333–346.

Konermann, L., Ahadi, E., Rodriguez, A. D., & Vahidi, S. (2013). Unraveling the mechanism of electrospray ionization. *Analytical Chemistry*, *85*(1), 2–9.

Korkhov, V. M., Mireku, S. A., Hvorup, R. N., & Locher, K. P. (2012). Asymmetric states of vitamin B_{12} transporter BtuCD are not discriminated by its cognate substrate binding protein BtuF. *FEBS Letters, 586*(7), 972–976.

Kota, J., Gilstring, C. F., & Ljungdahl, P. O. (2007). Membrane chaperone Shr3 assists in folding amino acid permeases preventing precocious ERAD. *The Journal of Cell Biology, 176*(5), 617–628.

Kozlov, A. G., & Lohman, T. M. (2002). Stopped-flow studies of the kinetics of single-stranded DNA binding and wrapping around the Escherichia coli SSB tetramer. *Biochemistry, 41*(19), 6032–6044.

Krewulak, K. D., & Vogel, H. J. (2011). TonB or not TonB: is that the question? *Biochemistry and Cell Biology, 89*(2), 87–97.

Kruppa, G. H., Schoeniger, J., & Young, M. (2003). A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry. *Rapid Communication in Mass Spectrometry, 17*(2), 155–162.

Krüger, R., Pfenninger, A., Fournier, I., Glückmann, M., & Karas, M. (2001). Analyte Incorporation

and Ionization in Matrix-Assisted Laser Desorption/Ionization Visualized by pH Indicator Molecular Probes. *Analytical Chemistry*, *73*(24), 5812–5821.

Kühn-Hölsken, E., Lenz, C., Dickmanns, A., Hsiao, H.-H., Richter, F. M., Kastner, B., Ficner, R. & Urlaub, H. (2010). Mapping the binding site of snurportin 1 on native U1 snRNP by cross-linking and mass spectrometry. *Nucleic Acids Research*, *38*(16), 5581–5593.

Latham, M. P., Zimmermann, G. R. & Pardi, A. (2009) NMR Chemical Exchange as a Probe for Ligand-Binding Kinetics in a Theophylline-Binding RNA Aptamer. *Journal of the American Chemical Society*, 132, 5052-5053.

Lecchi, P., Le, H. M. & Pannell, L. K. (1995). 6-Aza-2-thiothymine: a matrix for MALDI spectra of oligonucleotides. *Nucleic Acids Research*, *23*(7), 1276-1277.

Lee, J. F., Stovall, G. M., & Ellington, A. D. (2006). Aptamer therapeutics advance. *Current Opinion in Chemical Biology, 10*(3), 282–289.

Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., & Aebersold, R. (2010). Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Molecular & Cellular Proteomics, 9*(8), 1634–1649.

Lengqvist, J., Svensson, R., Evergren, E., Morgenstern, R., & Griffiths, W. J. (2004). Observation of an intact noncovalent homotrimer of detergent-solubilized rat microsomal glutathione transferase-1 by electrospray mass spectrometry. *Journal of Biological Chemistry, 279*(14), 13311–13316.

Lewinson, O., Lee, A. T., Locher, K. P., & Rees, D. C. (2010). A distinct mechanism for the ABC transporter BtuCD-BtuF revealed by the dynamics of complex formation. *Nature Structural & Molecular Biology, 17*(3), 332–338.

Li, H., Lin, T.-Y., Van Orden, S. L., Zhao, Y., Barrow, M. P., Pizarro, A. M., Qi, Y., Sadler, P. J. & O'Connor, P. B. (2011). Use of top-down and bottom-up Fourier transform ion cyclotron resonance mass spectrometry for mapping calmodulin sites modified by platinum anticancer drugs. *Analytical Chemistry*, *83*(24), 9507–9515.

Li, L., Golding, R. E., & Whittal, R. M. (1996). Analysis of single mammalian cell lysates by mass spectrometry. *Journal of the American Chemical Society, 118*(46), 11662–11663.

Light-Wahl, K. J., Schwartz, B. L., & Smith, R. D. (1994). Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. *Journal of the American Chemical Society, 116*(12), 5271–5278.

Lin, P.-H., Yen, S.-L., Lin, M.-S., Chang, Y., Louis, S. R., Higuchi, A., & Chen, W.-Y. (2008). Microcalorimetrics studies of the thermodynamics and binding mechanism between L-tyrosinamide and aptamer. *The Journal of Physical Chemistry. B, 112*(21), 6665–6673.

Lin, S., Cotter, R., & Woods, A. S. (1998). Detection of non-covalent interaction of single and double stranded DNA with peptides by MALDI-TOF. *Proteins: Structure, Function, and Bioinformatics, 33*(S2) ,12-21.

Lindsay, D. G. (1972). Intramolecular cross-linked insulin. *FEBS Letters, 21*(1), 105–108.

Lis, H., & Sharon, N. (1993). Protein glycosylation. Structural and functional aspects. *European Journal of Biochemistry, 218*(1), 1–27.

Litman, T., Jensen, U., Hansen, A., Covitz, K.-M., Zhan, Z., Fetsch, P., Abati, A., Hansen, P. R., Horn, T., Skovasgaard T., & Bates, S. E. (2002). Use of peptide antibodies to probe for the mitoxantrone resistance-associated protein MXR/BCRP/ABCP/ABCG2. Biochimica Et Biophysica Acta, 1565(1), 6–16.

Livadaris, V., & Blais, J. (2000). Formation of non-specific protein cluster ions in matrix-assisted laser desorption/ionization: abundances and dynamical aspects. *European Journal of Mass Spectrometry, 6*, 409-413.

Lizak, C., Gerber, S., Numao, S., Aebi, M., & Locher, K. P. (2011). X-ray structure of a bacterial oligosaccharyltransferase. *Nature, 474*(7351), 350–355.

Locher, K. P. (2004). Structure and mechanism of ABC transporters. *Current Opinion in Structural Biology, 14*(4), 426–431.

Locher, K. P. (2009). Review. Structure and mechanism of ATP-binding cassette transporters. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1514), 239–245.

Locher, K. P., Lee, A. T., & Rees, D. C. (2002). The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science, 296*(5570), 1091–1098.

Lohman, T. (1985). Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *Journal of Biological Chemistry. 260*(6), 3594-3603.

Lohman, T. M., & Ferrari, M. E. (1994). Escherichia coli single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annual Review of Biochemistry, 63*, 527–570.

Lomant, A. J. A., & Fairbanks, G. G. (1976). Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis-(succinimidyl propionate). *Journal of Molecular Biology, 104*(1), 243–261.

Loo, J. (2000). Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *International Journal of Mass Spectrometry. 200*(1-3), 175–186.

Loo, J. A., Berhane, B., Kaddis, C. S., Wooding, K. M., Xie, Y., Kaufman, S. L., & Chernushevich, I. V. (2005). Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *Journal of American Society for Mass Spectrometry*, *16*(7), 998–1008.

Ludwig, K., Habbach, S., Krieglstein, J., Klumpp, S., & König, S. (2011). MALDI-TOF high mass calibration up to 200 kDa using human recombinant 16 kDa protein histidine phosphatase aggregates. *PloS One*, *6*(8), e23612.

Luo, S.-Z., Li, Y.-M., Qiang, W., Zhao, Y.-F., Abe, H., Nemoto, T., Qin, X. R., & Nakanishi, H. (2004). Detection of specific noncovalent interaction of peptide with DNA by MALDI-TOF. *Journal of the American Society for Mass Spectrometry, 15*(1), 28–31.

Manning, G. S. (1969). Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions. III. An Analysis Based on the Mayer Ionic Solution Theory. *Journal of Chemical Physics, 51*(8), 3249.

Macaya, R. F. (1993). Thrombin-Binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proceedings of the National Academy of Sciences, 90*(8), 3745–3749.

Marcoux, J., & Robinson, C. V. (2013). Twenty years of gas phase structural biology. *Structure, 21*(9), 1541–1550.

Marty, M. T., Das, A., & Sligar, S. G. (2012). Ultra-thin layer MALDI mass spectrometry of membrane proteins in nanodiscs. *Analytical and Bioanalytical Chemistry, 402*(2), 721–729.

Mädler, S., Barylyuk, K., Boeri Erba, E., Nieckarz, R. J., & Zenobi, R. (2012). Compelling advantages of negative ion mode detection in high-mass MALDI-MS for homomeric protein complexes.

Journal of the American Society for Mass Spectrometry, 23(2), 213–224.

Mädler, S., Bich, C., Touboul, D., & Zenobi, R. (2009). Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *Journal of Mass Spectrometry, 44*(5), 694–706.

Mädler, S., Erba, E. B., & Zenobi, R. (2013). MALDI-ToF mass spectrometry for studying noncovalent complexes of biomolecules, *Topics in Current Chemistry*, *331*, 1–36.

Mädler, S., Seitz, M., Robinson J., & Zenobi, R. (2010). Does Chemical Cross-Linking with NHS Esters Reflect the Chemical Equilibrium of Protein–Protein Noncovalent Interactions in Solution? J*ournal of the American Society for Mass Spectrometry, 21*(10), 1775–1783.

McDevitt, C. A., Collins, R. F., Conway, M., Modok, S., Storm, J., Kerr, I. D., Ford, R. C., & Callaghan, R. (2006). Purification and 3D structural analysis of oligomeric human multidrug transporter ABCG2. *Structure, 14*(11), 1623–1632.

McKenzie, H. A., & White, F. H. (1991). Lysozyme and alpha-lactalbumin: structure, function, and interrelationships. *Advances in Protein Chemistry, 41*, 173–315.

Melcher, K. (2004). New chemical Crosslinking methods for the identification of transient protein-protein interactions with multiprotein complexes. *Current Protein & Peptide Science, 5*(4), 287–296.

McLafferty, F. W. (1981). Tandem mass spectrometry *Science*, *214*(4518), 280–287.

McLafferty, F. W. (1999). Biomolecule Mass Spectrometry. *Science*, *284*(5418), 1289–1290.

Migneault, I., Dartiguenave, C., Bertrand, M. J., & Waldron, K. C. (2004). Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *BioTechniques, 37*(5), 790–6, 798–802.

Mohammed, S., Lorenzen, K., Kerkhoven, R., van Breukelen, B., Vannini, A., Cramer, P., & Heck, A. J. R. (2008). Multiplexed proteomics mapping of yeast RNA polymerase II and III allows near-complete sequence coverage and reveals several novel phosphorylation sites. *Analytical Chemistry*, *80*(10), 3584–3592.

Müller, M. Q., & Sinz, A. (2012). Chemical cross-linking and high-resolution mass spectrometry to study protein-drug interactions. *Methods in Molecular Biology, 803*, 205–218.

Müller, R., & Allmaier, G. (2006). Molecular weight determination of ultra-high mass compounds on a standard matrix-assisted laser desorption/ionization time-of-flight mass spectrometer: PAMAM dendrimer generation 10 and immunoglobulin M. *Rapid Communications in Mass Spectrom*etry, *20*(24), 3803–3806.

Nazabal, A., Wenzel, R. J., & Zenobi, R. (2006). Immunoassays with Direct Mass Spectrometric Detection. Analytical Chemistry, 78(11), 3562–3570.

Neves, M. A. D., Reinstein, O., Saad, M., & Johnson, P. E. (2010). Defining the secondary structural requirements of a cocaine-binding aptamer by a thermodynamic and mutation study. *Biophysical Chemistry, 153*(1), 9–16.

Ngounou Wetie, A. G., Sokolowska, I., Woods, A. G., Roy, U., Loo, J. A., & Darie, C. C. (2013). Investigation of stable and transient protein-protein interactions: Past, present, and future. *Proteomics*, *13*(3-4), 538–557.

Ni, Z., Mark, M. E., Cai, X., & Mao, Q. (2010). Fluorescence resonance energy transfer (FRET) analysis demonstrates dimer/oligomer formation of the human breast cancer resistance protein (BCRP/ABCG2) in intact cells. *International Journal of Biochemistry and Molecular Biology, 1*(1), 1–11.

Nonin-Lecomte, S., Lin, C. H., & Patel, D. J. (2001) Additional Hydrogen Bonds and Base-Pair Kinetics in the Symmetrical AMP-DNA Aptamer Complex. *Biophysical Journal, 81*(6), 3422-3421.

Nooren, I. M. A. (2003). Diversity of protein-protein interactions. *The EMBO Journal, 22*(14), 3486–3492.

Novak, P., Haskins, W. E., Ayson, M. J., Jacobsen, R. B., Schoeniger, J. S., Leavell, M. D., Young, M. M & Kruppa, G. H. (2005). Unambiguous assignment of intramolecular chemical cross-links in modified mammalian membrane proteins by Fourier transform-tandem mass spectrometry. *Analytical Chemistry*, *77*(16), 5101–5106.

Novak, P., & Giannakopulos, A. E. (2007). Chemical cross-linking and mass spectrometry as structure determination tools. *European Journal of Mass Spectrometry, 13*(2), 105–113.

Ocsoy, I., Gulbakan, B., Shukoor, M. I., Xiong, X., Chen, T., Powell, D. H., & Tan, W. (2013). Aptamer-conjugated multifunctional nanoflowers as a platform for targeting, capture, and detection in laser desorption ionization mass spectrometry. *ACS Nano, 7*(1), 417–427.

Ollis, A. A., Manning, M., Held, K. G., & Postle, K. (2009). Cytoplasmic membrane protonmotive force energizes periplasmic interactions between ExbD and TonB. *Molecular Microbiology, 73*(3), 466–481.

Oosawa, F. (1971). Polyelectrolytes. New York: Marcel Dekker.

Ozvegy, C., Litman, T., Szakács, G., Nagy, Z., Bates, S., Varadi, A., & Sarkadi, B. (2001). Functional Characterization of the Human Multidrug Transporter, ABCG2, Expressed in Insect Cells. *Biochemical and Biophysical Research Communications, 285*(1), 111-117.

Pabo, C. O., & Sauer. R. T., (1992) Transcription factors: structural families and principles of DNA recognition. *Annual Review of Biochemistry*, *61*, 1053-1095.

Phizicky, E. M., & Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiology Reviews*, *59*(1), 94–123.

Pimenova, T., Nazabal, A., Roschitzki, B., Seebacher, J., Rinner, O., & Zenobi, R. (2008). Epitope mapping on bovine prion protein using chemical cross-linking and mass spectrometry. *Journal of Mass Spectrometry*, *43*(2), 185–195.

Pimenova, T., Pereira, C. P., Schaer, D. J., & Zenobi, R. (2009). Characterization of high molecular weight multimeric states of human haptoglobin and hemoglobin-based oxygen carriers by high-mass MALDI MS. *Journal of Separation Science, 32*(8), 1224–1230.

Polgar, O., Robey, R. W., & Bates, S. E. (2008). ABCG2: structure, function and role in drug response. *Expert Opinion on Drug Metabolism & Toxicology, 4*(1), 1–15.

Postle, K., Kadner, R. (2003) Touch and go: tying TonB to transport. *Journal of Molecular Microbiolology, 49*(4), 869-882.

Postle, K., & Skare, J. T. (1988). Escherichia coli TonB protein is exported from the cytoplasm without proteolytic cleavage of its amino terminus. *Journal of Biological Chemistry, 263*(22), 11000-11007.

Pramanik, A., Hauf, W., Hoffmann, J., Cernescu, M., Brutschy, B., & Braun, V. (2011). Oligomeric structure of ExbB and ExbB-ExbD isolated from Escherichia coli as revealed by LILBID mass spectrometry. *Biochemistry, 50*(41), 8950–8956.

Pramanik, A., Zhang, F., Schwarz, H., Schreiber, F., & Braun, V. (2010). ExbB protein in the cytoplasmic membrane of Escherichia coli forms a stable oligomer. *Biochemistry, 49*(40),

8721–8728.

Raghunathan, S., Kozlov, A. G., Lohman, T. M., & Waksman, G. (2000). Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nature Structural Biology, 7*(8), 648–652.

Rigaud, J.-L., & Lévy, D. (2003). Reconstitution of Membrane Proteins into Liposomes. *Methods in Enzymology, 372*, 65–86.

Robinette, D., Neamati, N., Tomer, Kenneth B, & Borchers, C. H. (2006). Photoaffinity labeling combined with mass spectrometric approaches as a tool for structural proteomics. *Expert Review of Proteomics*, *3*(4), 399–408.

Rosenberg, M. F., Bikadi, Z., Chan, J., Liu, X., Ni, Z., Cai, X., Ford R. C., & Mao, Q. (2010). The human breast cancer resistance protein (BCRP/ABCG2) shows conformational changes with mitoxantrone. *Structure, 18*(4), 482–493.

Rosinke, B., Strupat, K., Hillenkamp, F., Rosenbusch, J., Dencher, N., Krüger, U., & Galla, H. (1995). Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) of membrane proteins and non-covalent complexes. *Journal of Mass Spectrometry, 30*(10), 1462-1468.

Rostom, A. A. (2000). Detection and selective dissociation of intact ribosomes in a mass spectrometer. *Proceedings of the National Academy of Sciences, 97*(10), 5185–5190.

Ruigrok, V. J. B., Levisson, M., Hekelaar, J., Smidt, H., Dijkstra, B. W., & van der Oost, J. (2012a). Characterization of Aptamer-Protein Complexes by X-ray Crystallography and Alternative Approaches. *International Journal of Molecular Sciences, 13*(12), 10537–10552.

Ruigrok, V. J. B., van Duijn, E., Barendregt, A., Dyer, K., Tainer, J. A., Stoltenburg, R., Strehlitz, B., Levisson, M., Smidt, H., & van der Oost, J., (2012b). Kinetic and stoichiometric characterisation of streptavidin-binding aptamers. *Chembiochem, 13*(6), 829–836.

Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H., & Robinson, C. V. (2005). Evidence for macromolecular protein rings in the absence of bulk water. *Science*, *310*(5754), 1658–1661.

Russo Krauss, I., Merlino, A., Giancola, C., Randazzo, A., Mazzarella, L., & Sica, F. (2011). Thrombin-aptamer recognition: a revealed ambiguity. *Nucleic Acids Research, 39*(17), 7858–7867.

Sali, A., Glaeser, R., Earnest, T., & Baumeister, W. (2003). From words to literature in structural proteomics. *Nature*, *422*(6928), 216–225.

Schey, K. L., Papac, D. I., Knapp, D. R., & Crouch, R. K. (1992). Matrix-assisted laser desorption mass spectrometry of rhodopsin and bacteriorhodopsin. *Biophysical Journal, 63*(5), 1240–1243.

Schuck, P. W. (2007). *Protein Interactions*. Biophysical Approaches for the Study of Complex Reversible Systems. Springer.

Schlosser, G., Pocsfalvi, G., Malorni, A., Puerta, A., de Frutos, M., & Vékey, K. (2003). Detection of immune complexes by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Communication in Mass Spectrom*etry, *17*(24), 2741–2747.

Schug, K. A., & Lindner, W. (2005). Noncovalent binding between guanidinium and anionic groups: focus on biological- and synthetic-based arginine/guanidinium interactions with phosph[on]ate and sulf[on]ate residues. *Chemical Reviews, 105*(1), 67–114.

Schultze, P., Macaya, R. F., & Feigon, J. (1994). Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *Journal of Molecular Biology, 235*(5), 1532–1547.

Schürer, H., Buchynskyy, A., Korn, K., Famulok, M., Welzei, P., & Hahn, U. (2001). Fluorescence correlation spectroscopy as a new method for the investigation of aptamer/target interactions. *Biological Chemistry, 382*(3), 479–481.

Sharon, M., Taverner, T., Ambroggio, X. I., Deshaies, R. J., & Robinson, C. V. (2006). Structural organization of the 19S proteasome lid: insights from MS of intact complexes. *PLoS Biology*, *4*(8), e267.

Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M., & Keck, J. L. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology, 43*(5), 289–318.

Shimma, Y., Nishikawa, A., bin Kassim, B., Eto, A., & Jigami, Y. (1997). A defect in GTP synthesis affects mannose outer chain elongation in Saccharomyces cerevisiae. *Molecular & General Genetics : MGG, 256*(5), 469–480.

Singh, P., Panchaud, A., & Goodlett, D. R. (2010). Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique. *Analytical Chemistry, 82*(7), 2636–2642.

Sinz, A. (2003). Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *Journal of Mass Spectrometry, 38*(12), 1225–1237.

Sinz, Andrea. (2005). Chemical cross-linking and FTICR mass spectrometry for protein structure characterization. *Analytical and Bioanalytical Chemistry*, *381*(1), 44–47.

Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions, *Mass Spectrometry Reviews, 25*, 663-682.

Smyth, D. G., Blumenfeld, O. O., & Konigsberg, W. (1964). Reactions of N-ethylmaleimide with peptides and amino acids. *The Biochemical Journal, 91*(3), 589–595.

Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A., & Robinson, C. V. (2002). A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Analytical Chemistry*, *74*(6), 1402–1407.

Spiro, R. G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology, 12*(4), 43R-56R.

Steen, H., Petersen, J., Mann, M., & Jensen, O. N. (2001). Mass spectrometric analysis of a UV-cross-linked protein-DNA complex: tryptophans 54 and 88 of E. coli SSB cross-link to DNA. *Protein Science, 10*(10), 1989–2001.

Strynadka, N. C. J., Eisenstein, M., Katchalski-Katzir, E., Shoichet, B. K., Kuntz, I. D., Abagyan, R., Totrov, M., et al. (1996). Molecular docking programs successfully predict the binding of a β-lactamase inhibitory protein to TEM-1 β-lactamase. *Nature Structural & Molecular Biology*, *3*(3), 233–239.

Sudha, R., & Zenobi, R. (2002) The detection and stability of DNA duplexes probed by MALDI mass spectrometry. Helvetica Chimica Acta, 58(10), 3163-3143.

Tang, X., Callahan, J. H., Zhou, P., & Vertes, A. (1995). Noncovalent protein-oligonucleotide interactions monitored by matrix-assisted laser desorption/ionization mass spectrometry. *Analytical Chemistry, 67*(24), 4542–4548.

Tarze, A., Deniaud, A., Le Bras, M., Maillier, E., Molle, D., Larochette, N., Zamzami, N., Jan, G., Kroemer, G. & Brenner, C. (2007). GAPDH, a novel regulator of the pro-apoptotic mitochondrial membrane permeabilization. *Oncogene, 26*(18), 2606–2620.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. & Matsuo, T. (1988). Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*, 2 (20): 151–153.

Thiede, B., & Janta-Lipinski, von, M. (1998). Noncovalent RNA-peptide complexes detected by matrix-assisted laser desorption/ionization mass spectrometry *Rapid communications in mass spectrometry*, *12*(23), 1889–1894.

Trimpin, S., & Deinzer, M. L. (2007). Solvent-free MALDI-MS for the analysis of a membrane protein via the mini ball mill approach: case study of bacteriorhodopsin. *Analytical Chemistry, 79*(1), 71–78.

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science, 249*(4968), 505–510.

Uetrecht, C, Versluis, C., Watts, N. R., Roos, W. H., Wuite, G. J. L., Wingfield, P. T., Steven, A. C., & Heck, A. J. R., (2008). High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids. *Proceedings of the National Academy of Sciences, 105*(27), 9216–9220.

Uetrecht, Charlotte, Barbu, I. M., Shoemaker, G. K., van Duijn, E., & Heck, A. J. R. (2010a). Interrogating viral capsid assembly with ion mobility–mass spectrometry. *Natural Chemistry*, *3*(2), 126–132.

Uetrecht, C., Rose. J., van Duijn, E., Lorenzen, K., & Heck, A. J. R. (2010b). Ion mobility mass spectrometry of proteins and protein assemblies. *Chemical Society Reviews*, *39*(5), 1633–1655.

van Duijn, E. (2010) Current limitations in native mass spectrometry based structural biology. *Journal of American Society for Mass Spectrometry, 21*(6), 971-978.

Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. Glycobiology, 3(2), 97–130.

Vinkenborg, J. L., Karnowski, N., & Famulok, M. (2011). Aptamers for allosteric regulation. Nature Chemical Biology, 7(8), 519–527.

Videler, H., Ilag, L. L., McKay, A. R. C., Hanson, C. L., & Robinson, C. V. (2005). Mass spectrometry of intact ribosomes. *FEBS Lett. 579*(4), 943–947.

Wakabayashi-Nakao, K., Tamura, A., Furukawa, T., Nakagawa, H., & Ishikawa, T. (2009). Quality control of human ABCG2 protein in the endoplasmic reticulum: ubiquitination and proteasomal degradation. *Advanced Drug Delivery Reviews, 61*(1), 66–72.

Wandersman, C., & Delepelaire, P. (2004) Bacterial iron sources: from siderophores to hemophores. *Annual Review of Microbiology, 58*(1), 611-647.

Wei, T. F., Bujalowski, W., & Lohman, T. M. (1992). Cooperative binding of polyamines induces the Escherichia coli single-strand binding protein-DNA binding mode transitions. *Biochemistry, 31*(26), 6166–6174.

Weidmann, S., Barylyuk, K., Nespovitaya, N., Mädler, S., & Zenobi, R. (2013a). A new, modular mass calibrant for high-mass MALDI-MS. *Analytical Chemistry*, *85*(6), 3425–3432.

Weidmann, S., Mikutis, G., Barylyuk, K., & Zenobi, R. (2013b). Mass Discrimination in High-Mass MALDI-MS. Journal of the American Society for Mass Spectrometry. 24(9), 1396–1404.

Wenzel, R. J., Matter, U., Schultheis, L., & Zenobi, R. (2005). Analysis of megadalton ions using cryodetection MALDI time-of-flight mass spectrometry. *Analytical Chemistry*, *77*(14), 4329–

4337.

Whitelegge, J. P., Gundersen, C. B., & Faull, K. F. (1998). Electrospray-ionization mass spectrometry of intact intrinsic membrane proteins. Protein Science, 7(6), 1423–1430.

Wiener, M. C. (2005). TonB-dependent outer membrane transport: going for Baroque? *Current Opinion in Structural Biology, 15*(4), 394–400.

Wong, S. S. (1991). Chemistry of Protein Conjugation and Cross-Linking. CRC Press.

Wyckoff, H. W., Hardman, K. D., Allewell, N. M., Inagami, T., Tsernoglou, D., Johnson, L. N., & Richards, F. M. (1967). The structure of ribonuclease-S at 6 A resolution. *The Journal of Biological Chemistry, 242*(16), 3749–3753.

Woods, A. S., Buchsbaum, J. C., Worrall, T. A., Berg, J. M. & Cotter, R. J. (1995) Matrix-assisted laser desorption/ionization of noncovalently bound compounds. *Analytical Chemistry*, *67*, 4462-4465.

Wortmann, A., Pimenova, T., Alves, S., & Zenobi, R. (2007). Investigation of the first shot phenomenon in MALDI mass spectrometry of protein complexes. *Analyst*, *132*(3), 199.

Xu, J. (2004). Characterization of Oligomeric Human Half-ABC Transporter ATP-binding Cassette G2. *Journal of Biological Chemistry, 279*(19), 19781–19789.

Yao, C., Qi, Y., Zhao, Y., Xiang, Y., Chen, Q., & Fu, W. (2009). Aptamer-based piezoelectric quartz crystal microbalance biosensor array for the quantification of IgE. *Biosensors & Bioelectronics, 24*(8), 2499–2503.

Yin, S., & Loo, J. A. (2009). Mass spectrometry detection and characterization of noncovalent protein complexes. *Methods in Molecular Biology*, *492*, 273–282.

Young, M M, Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W. & Dollinger, G. (2000). High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proceedings of the National Academy of Sciences, 97*(11), 5802–5806.

Zehl, M., & Allmaier, G. (2004). Ultraviolet matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of intact hemoglobin complex from whole human blood. *Rapid Communications in Mass Spectrometry, 18*(17), 1932–1938.

Zehl, M., & Allmaier, G. (2005). Instrumental parameters in the MALDI-TOF mass spectrometric analysis of quaternary protein structures. *Analytical Chemistry 77*(1), 103–110.

Zenobi, R. & Knochenmuss, R. (1998) Ion formation in MALDI mass spectrometry. *Mass Spectrometry Review*, *17*, 337-366.

Zhang, D., Lu, M., & Wang, H. (2011). Fluorescence anisotropy analysis for mapping aptamer-protein interaction at the single nucleotide level. *Journal of the American Chemical Society, 133*(24), 9188–9191.

Zhang, H., Cui, W., Wen, J., Blankenship, R. E., & Gross, M. L. (2011). Native electrospray and electron-capture dissociation FTICR mass spectrometry for top-down studies of protein assemblies. *Analytical Chemistry, 83*(14), 5598–5606.

Zhang, N., Doucette, A., & Li, L. (2001). Two-Layer Sample Preparation Method for MALDI Mass Spectrometric Analysis of Protein and Peptide Samples Containing Sodium Dodecyl Sulfate. *Analytical Chemistry, 73*(13), 2968–2975.

Zhou, M., Morgner, N., Barrera, N. P., Politis, A., Isaacson, S. C., Matak-Vinkovic, D., Murata, T.,

Bernal, R. A., Stock, D. & Robinson, C. V. (2011). Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. *Science, 334*(6054), 380–385.

Zhou, M. W., Dagan, S., & Wysocki, V. H. (2013). Impact of charge state on gas-phase behaviors of noncovalent protein complexes in collision induced dissociation and surface induced dissociation. *Analyst, 138*(5), 1353–1362.

Zhou, M. W., Huang, C., & Wysocki, V. H. (2012). Surface-induced dissociation of ion mobility-separated noncovalent complexes in a quadrupole/time-of-flight mass spectrometer. *Analytical Chemistry, 84*(14), 6016–6023.

Curriculum Vitae

Fan Chen

Born August 15, 1986 in Fujian, China

Chinese Citizen

01/2011-02/2014	PhD Thesis in Analytical Chemistry Prof. Dr. Renato Zenobi, ETH Zürich, Switzerland Department of Chemistry and Applied Biosciences
08/2008-10/2010	MPhil in Analytical Chemistry Prof. Dr. Dominic Tak Wah Chan, CUHK, Hong Kong Department of Chemistry
09/2004-06/2008	BSc in Chemistry Xiamen University, China