MASS SPECTROMETRY ANALYSIS OF PROTEINS AND PROTEIN PHOSPHORYLATION USING FUNCTIONALIZED POLYMER MATERIALS

By

Weihan Wang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Chemistry

2011

ABSTRACT

MASS SPECTROMETRY ANALYSIS OF PROTEINS AND PROTEIN PHOSPHORYLATION USING FUNCTIONALIZED POLYMER MATERIALS

By

Weihan Wang

Posttranslational protein phosphorylation helps control many cellular processes, so identification of stimuli-induced phosphorylation events is vital to understanding cell regulation. Characterization of phosphorylated proteins by mass spectrometry (MS) typically employs small phosphopeptides (<3000 Da) generated by complete protein digestion. However, efficient phosphopeptide ionization for MS and tandem MS (MS/MS) analysis requires enrichment of phosphopeptides from complex protein digests. Alternatively, digestion of proteins to larger peptides (middle-down proteolysis) can improve protein sequence coverage in MS analysis and reveal a wider range of phosphorylation sites than complete digestion. Initially, this dissertation reviews recent studies on phosphopeptide enrichment and middle-down proteolysis. Based on these studies, the research aims to (1) improve enrichment of small phosphopeptides for MS analysis of target phosphoproteins, and (2) generate large tryptic peptides for middle-down proteomics.

In studies on enrichment, polymer brushes derivatized with nitrilotriacetate-Fe(III) complexes (NTA-Fe) or oxotitanium (oxoTi) capture phosphopeptides from protein digests. Specifically, polymer microspots derivatized with NTA-Fe and surrounded by hydrophobic poly(dimethylsiloxane) facilitate pre-concentration and enrichment of dilute phosphopeptides from only 1 μL of digest. Matrix-assisted laser desorption/ionization (MALDI)-MS directly on the 250-μm spots leads to a sub-fmol detection limit for a β-casein phosphopeptide. Compared to

films derivatized with NTA-Fe, polymer-oxoTi brushes improve the capture and detection of mono-phosphopeptides in mixtures of model protein digests. Moreover, MALDI collision-induced dissociation (CID)-MS/MS on polymer-oxoTi identified unknown phosphopeptides from several p65-associated phosphoproteins. These phosphoproteins immunoprecipitated together with p65 from the nuclear extracts of human acute monocytic leukemia cells.

Bifunctionalization of polymer brushes with both oxoTi and NTA-Fe groups leads to simultaneous low-bias enrichment of mono- and multi-phosphorylated peptides from moderately complex model protein digests. For larger samples (tens of μLs), stacking two membranes modified with polymer-oxoTi and polymer-NTA brushes, respectively, in an HPLC fitting enhances phosphopeptide enrichment. The two membranes sequentially capture mono- and multi-phosphorylated peptides from 3 phosphoprotein digests (~1 μg total mass) in the presence of ~70 μg of non-phosphoprotein digest.

To control digestion for middle-down MS analysis of proteins, pumping protein solutions through nylon membranes containing immobilized trypsin affords digestion times ranging from a few milliseconds to seconds. In modified membranes with 0.45 μm pores, a 50-ms digestion of apomyoglobin leads to peptides with MWs primarily between 2.5 and 6 kDa. Five of these peptides cover the entire protein sequence. A 10-ms β-casein residence time in modified membranes with 1.2-μm pores yields peptides with masses primarily between 3 and 20 kDa. Two large peptides constitute the entire protein sequence, and one of them covers all the five phosphorylation sites. These simple enrichment methods and new digestion techniques for middle-down proteomics should facilitate future studies of protein phosphorylation.

I dedicate this dissertation to my parents, Jiaxiang Wang and Baozhu Wang, and my wife, Jia Liu, for their love and support.

Acknowledgement

I would like to express my sincere thanks to my advisor, Prof. Merlin Bruening, for his guidance and support. I am lucky enough to join Prof. Bruening's group and receive comprehensive trainings on instrumental analysis and material synthesis. His patience and encouragement helped me overcome difficulties at the early stage of my Ph.D study; and his trust gave me confidence and freedom in exploring new areas in the last two years. I will never forget those exciting moments when we got interesting results.

I am grateful to my second reader, Prof. Gavin Reid, for his strict training on mass spectrometry and helpful discussions on several research projects. Some of my manuscripts could not have been published smoothly without his commenting and revisions.

I also thank Profs. Jetze Tepe, Dan Jones, Gregory Baker, Gary Blanchard and Scott McLuckey for their kind help and suggestions. Thanks to Prof. Jetze Tepe for providing precious samples to test my brush weapons! Thanks to Prof. McLuckey from Purdue University for his collaboration in middle-down protein analysis.

I would like to acknowledge the collaborations and discussions from many researchers in and outside Bruening group. Many thanks to Yujing Tan for our long time collaboration and friendship built from those unforgettable numbers "2061, 2089, 3122…". I will always remember Somnath Bhattacharjee and Yong Guan for their insightful discussions on organic synthesis and countless "gossip" of the organic world, and Fei Xu for her dedication on the project of membrane digestion. Because of the collaborations with Amanda Palumbo, as well as with Ken Chanthamontri from Purdue University, I was able to extend my skills to address important biological problems. I also would like to thank my friend and former labmate, Mingxi

Zhang, who shared his insights toward life and chemistry with me every week via internet. The bifunctional polymer brushes actually began growing from our MSN dialog window.

Most importantly, none of the work would have been possible without the love from my family. I owe so much to my parents, Jiaxiang Wang and Baozhu Wang, who brought their only child up, but could only hear his voice on the phone once a week in the last five years. Thanks to my wife, Jia Liu, for her constant support and understanding. Each piece of my work is soaked with her love and encouragement!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

LIST OF ABBREVIATIONS

PTM	Posttranslational modification
ATP	Adenosine triphosphate
Ser	Serine
Thr	Threonine
Tyr	Tyrosine
Lys	Lysine
Arg	Arginine
Glu	Glutamic acid
Asp	Aspartic acid
Asn	Asparagine
His	Histidine
Trp	Tryptophan
Met	Methionine
Cys	Cysteine
TNF	Tumor necrosis factor
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
THP-1	Human acute monocytic leukemia cell
BSA	Bovine serum albumin
MW	Molecular weight
PMF	Peptide mass fingerprinting
MS	Mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization

ESI	Electrospray ionization
CID	Collision-induced dissociation
ETD	Electron-transfer dissociation
TOF	Time-of-flight
HPLC	High performance liquid chromatography
FTIR	Fourier transform infrared
SEM	Scanning electron microscopy
ICP-OES	Inductively coupled plasma-optical emission spectrometry
%SCT	100% sequence coverage peak intensity threshold
IMAC	Immobilized metal ion affinity chromatography
SCX	Strong cation exchange
SAX	Strong anion exchange
IP	Immunoprecipitation
ACN	Acetonitrile
DMF	Dimethylformamide
THF	Tetrahydrofuran
TFA	Trifluoroacetic acid
FA	Formic acid
ATRP	Atom-transfer radical polymerization
PDMS	Poly(dimethylsiloxane)
PHEMA	Poly(2-hydroxyethyl methacrylate)
PGMA	Poly(glycidyl methacrylate)
PMES	Poly(2-(methacryloyloxy)ethyl succinate)

PSS	Poly(sodium 4-styrenesulfonate)
oxoTi	oxotitanium
EDTA	Ethylenediaminetetraacetic acid
SA	Succinic anhydride
DMAP	4-Dimethylaminopyridine
EDC	*N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride
NHS	*N*-hydroxysuccinimide
NTA	Nitrilotriacetate
DHB	2,5-Dihydroxybenzoic acid
DTT	Dithiothreitol
IAA	Iodoacetamide

Chapter 1. Introduction

This dissertation describes research on polymer-based techniques for phosphopeptide enrichment and size-controlled proteolysis prior to mass spectrometry (MS) analysis. To put this work in perspective, this introduction first describes protein phosphorylation and gives an overview of current strategies for MS-based identification and characterization of phosphorylation. This chapter then reviews and discusses in detail progress in phosphopeptide enrichment techniques and "middle-down" MS analysis of proteins, which both relate closely to the research in this dissertation. Subsequently, the chapter briefly introduces tandem MS and quantitative MS analysis, two rapidly growing areas that promise qualitative and quantitative protein characterization, respectively. The final section of this introduction presents a brief outline of the dissertation.

1.1 Protein phosphorylation

Posttranslational modification (PTM) of proteins by phosphorylation is an ubiquitous, highly evolved regulatory mechanism for creating diversity in both eukaryotes and prokaryotes. Phosphorylation of specific proteins provides an important control mechanism for cellular processes such as metabolic pathways, kinase cascade activation, gene transcription, membrane transport and motor mechanisms.[1-3]

Protein phosphorylation occurs most commonly at Ser, Thr and Tyr residues, which all have a side chain with a terminal –OH group that can be phosphorylated via the transfer of the γ-PO_3^{2-} from adenosine triphosphate (ATP) (Figure 1.1).[1] Protein kinases catalyze this process, and around 30% of cellular proteins are phosphorylated on at least one residue at a certain time

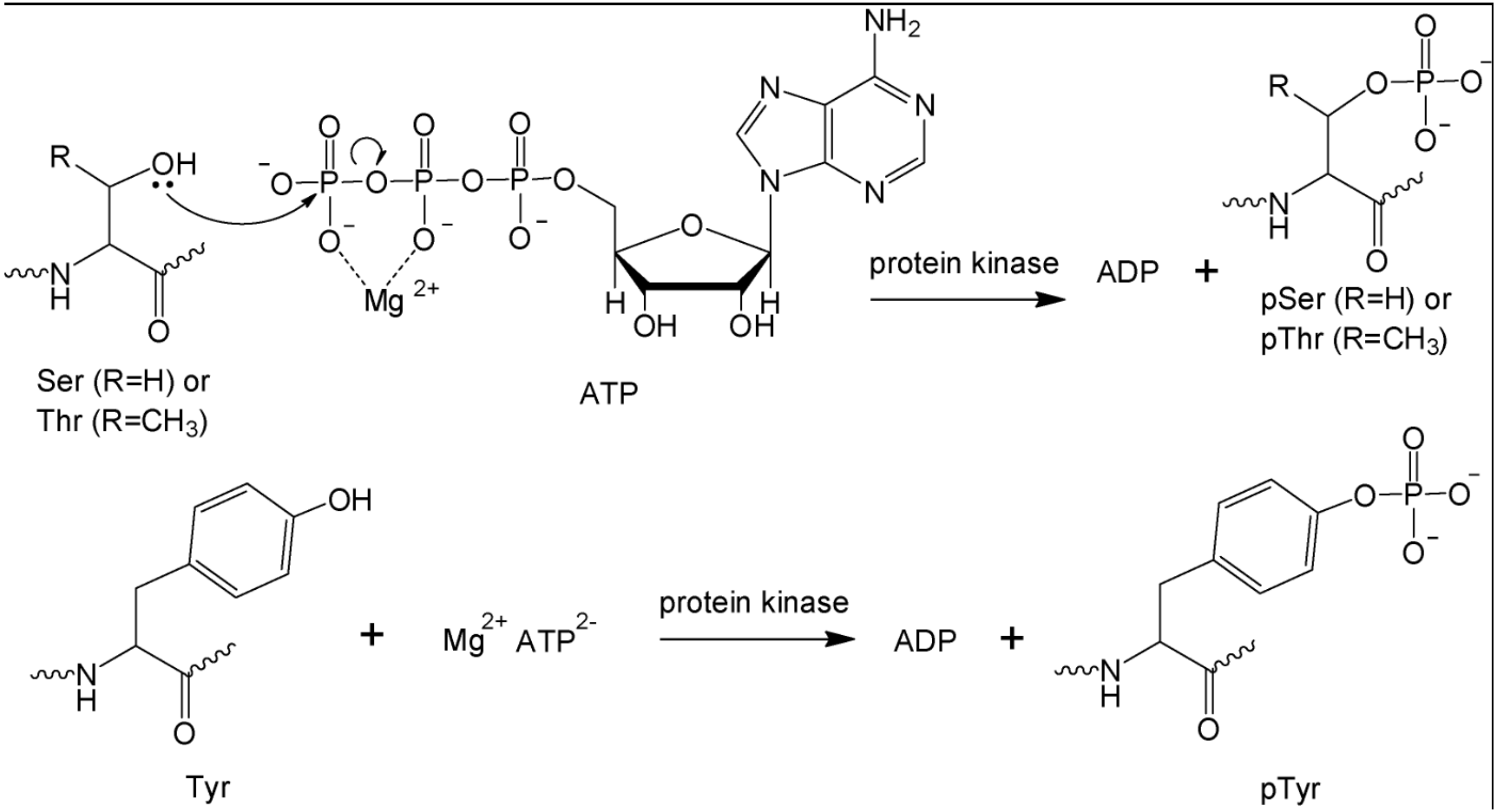


Figure 1.1. Protein phosphorylation at Ser/Thr (top) and Tyr (bottom) residues. The –OH groups of Ser and Thr attack γ-PO_3^{2-} from ATP (catalyzed by Ser/Thr specific kinases) to form pSer and pThr. A Tyr specific kinase transfers the γ-PO_3^{2-} to Tyr to form pTyr. (Redrawn from a figure in reference [1]).

during their life cycle.[4,5] Although for a wide range of species the abundances of Ser, Thr and Tyr in proteins are around 8.5%, 5.7% and 3.0% (2.8: 1.9: 1), respectively,[6] recent global mass spectrometry (MS)-based studies of human protein phosphorylation suggest that the distribution of phosphorylated Ser, Thr and Tyr (pSer, pThr and pTyr) is ~21: 4.4: 1.[7,8] The differences in the ratios between the phosphorylated residues and their non-phosphorylated forms stem from protein kinases that exhibit dramatically different specificities toward Ser, Thr and Tyr. Even though protein kinases typically share a common fold in their structure, they differ greatly in the charge and hydrophobicity of their surface residues.[9] The two main groups of protein kinases show either Ser/Thr (~80% of kinases) or Tyr specificity,[10] with the latter group having a deeper catalytic cleft that distinguishes the specificities of the two classes.[11,12]

The number of sites (Ser/Thr or Tyr) that a kinase can phosphorylate on different substrate proteins ranges from several to hundreds.[9] Catalysis of phosphorylation normally occurs after the kinase recognizes the consensus residues (amino acids situated immediately N-terminally and C-terminally to the phosphorylation site) of the substrate protein via charge, hydrogen bonding or hydrophobic interactions. For example, the consensus sequence of the substrate proteins of PKA (protein kinase A or cAMP-dependent protein kinase) is R-R-X-S/T-Φ (Φ is a hydrophobic residue and X is any amino acid residue). When PKA comes close to its substrate protein, two Glu residues near the catalytic region of PKA recognize the two Arg residues of the consensus sequence via electrostatic interactions, and the hydrophobic pocket of PKA favors binding of the hydrophobic residue (Φ).[5]

Phosphorylation controls protein function in several ways. For example, the introduction of $-OPO_3^{2-}$ (pKa_2 ~6) to Ser_{14} of glycogen phosphorylase causes intra-subunit contact of the phosphate group with Arg_{69} and Arg_{43} via charge pairing, and consequently activates the

catalytic function of this protein in cleaving glycogen to glucose-1-phosphate.[2,13] In addition to local conformational changes, many independently folding protein domains show affinity for short phosphorylated peptide sequences of certain other proteins.[14-16] These recognition domains are key elements in signal transduction in cells.

In signaling pathways, proteins play different roles in sequential steps during which they can be phosphorylated by specific kinases. Some of the kinases act in cascades: each kinase phosphorylates hundreds to thousands of proteins at each step, giving rise to dramatic signal amplifications.[17] Along with hundreds of protein kinases that work in eukaryotic cells to turn "on" signaling, there are also many protein phosphatases that turn "off" signaling pathways. The phosphatases either activate water in a bimetallic center for hydrolytic attack on pSer or pThr, or covalently release the phosphoryl group from pTyr using a nucleophilic thiolate side chain.[18]

The spatial and temporal regulation of both protein kinases and phosphatases ensure an appropriate balance of protein phosphorylation in signaling pathways.[19] Due to the crucial roles of phosphorylated proteins in cells, a comprehensive analysis of protein phosphorylation is necessary for understanding biological functions at a molecular level.

1.2 Overview of strategies for MS-based protein identification and characterization of phosphorylation

A comprehensive phosphoprotein analysis involves identification of phosphoproteins, localization of phosphorylation sites and quantitation of phosphorylation.[21] Current MS-based protein analysis mainly relies on digestion of proteins into small peptides ("bottom-up" approach, Figure 1.2) using proteases (also called proteolytic enzymes) such as trypsin. Prior to MS, the

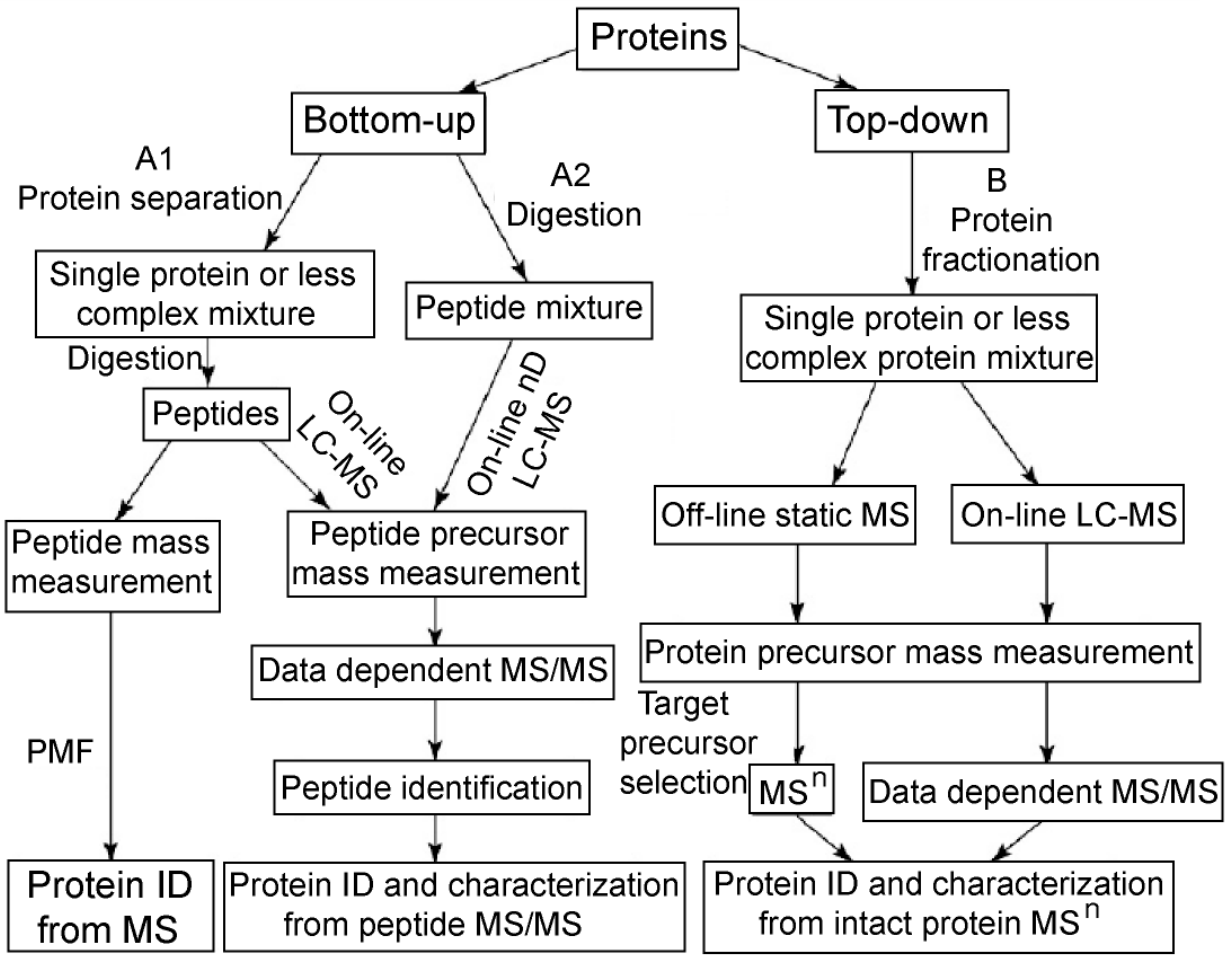


Figure 1.2. "Bottom-up" and "top-down" strategies for MS-based protein identification and characterization (redrawn according to a figure in reference [20]).

bottom-up approach typically includes multiple separation strategies, e.g. reversed phase and size-exclusion chromatography, isoelectric focusing, affinity purification, and ion exchange chromatography at the protein or peptide level.[20,22] For a sample with several proteins of specific interest, the "sort-then-break"[20] workflow (Figure 1.2, path A1) is common. In this approach, proteins are fractionated prior to proteolytic digestion of each fraction, and the resulting peptides are either directly analyzed for protein identification via "peptide mass fingerprinting" (PMF),[23] or further separated before MS/MS for protein identification and/or PTM characterization.[24,25] In recent years, a "break-then-sort" approach, sometimes known as "shotgun proteomics" (Figure 1.2, path A2), has gained popularity.[20,26] In this strategy,

protein digestion occurs without any prefractionation, and the resulting peptides are separated using multiple separation modes (normally liquid chromatography, LC, is the final step) prior to MS/MS for peptide sequencing with the assistance of database searching. This approach is attractive for large-scale and high-throughput analysis of complex samples, especially for PTM characterization on the proteome-wide scale.[27]

Despite its simple concept, shotgun proteomics requires high MS sensitivity and efficient peptide separation because of the greatly increased sample complexity (e.g. digested whole cell lysate) compared to the "sort-then-break" approach. Because proteins are expressed at a wide range of concentrations, signals for low-level proteins are often undetectable after the conversion of intact proteins into a complex mixture of peptides.[20,28] Moreover, peptides with extreme hydrophilicity, hydrophobicity, isoelectric point (pI) values or unexpected modifications (either PTM or chemical modification during sample preparation) are also easily lost during peptide separation/enrichment and MS analysis.[27,28] Nevertheless, in the case of global analysis of phosphoproteins from whole cell lysates or subcellular fractions, most studies employ a "break-then-sort" approach to identify hundreds to thousands of phosphopeptides that come from several tens to hundreds of phosphoproteins.[29] For targeted characterization of protein phosphorylation, recombinant proteins or proteins isolated using antibodies are also often digested prior to MS analysis of the phosphopeptides.[30,31]

Due to the low level of phosphopeptides relative to non-phosphopeptides, directly subjecting a protein digest to LC-electrospray ionization (ESI) MS/MS results in the loss of most phosphopeptide signals, mainly due to the low retention of phosphopeptides on C18 columns and ion suppression of phosphopeptides by non-phosphopeptides. Thus, MS-based phosphoproteomics studies in recent years inevitably included enrichment of phosphopeptides

prior to LC-MS analysis.[32] For example, using the combination of strong cation exchange (SCX) chromatography and immobilized metal ion affinity chromatography (IMAC) to separate and enrich phosphopeptides, Gygi and coworkers employed LC-MS/MS and database searching to characterize over 5500 and 13000 phosphorylation events from several mg of proteins from mouse liver and *Drosophila* embryos, respectively, in two days.[33] In analysis of phosphorylation of targeted proteins, enrichment of phosphopeptides from less complex digests (e.g. recombinant or gel electrophoresis-fractionated proteins) led to greatly improved phosphopeptide signals for MALDI-MS/MS or ESI-MS/MS characterization.[30] Section 1.3 reviews in detail the various phosphopeptide enrichment strategies for different types of samples.

An alternative to the bottom-up approach is "top-down" analysis (Figure 1.2). Instead of breaking proteins into small peptides (mainly <3 kDa), the top-down strategy introduces intact proteins into the gas phase for fragmentation inside the mass spectrometer.[34] This approach aims to examine the entire protein sequence and comprehensively identify PTMs. Because of the high masses of proteins relative to peptides, top-down methodologies require a much greater MS resolving power than bottom-up methods. Thus, expensive high resolution mass analyzers such as Fourier transform ion cyclotron resonance (FT-ICR) with electron-capture dissociation (ECD) are typically needed to provide accurate mass values and efficient ion fragmentation.[20] However, with recent advances in instrumentation the top-down approach is becoming more attractive for MS characterization of modifications of a single protein or a simple protein mixture when the protein mass is <50 kDa.[20,34] Examples of top-down MS protein characterization include comprehensive mapping of histone PTMs,[35,36] identification of a family of highly related genes encoding proteins with multifaceted mass differences,[37] analysis of membrane

proteins,[38] and revealing the alterations of phosphorylation between the truncated and full-length forms of cardiac myosin binding protein C (cMyBP-C).[31]

In spite of successes with targeted studies of single proteins, the top-down method suffers from several issues. Challenges in coupling proteome fractionation with MS analysis restrict the throughput of top-down proteome analysis,[39] and thus far only a few studies reported analysis of several hundred proteins in a short period of time. Another challenge for top-down protein characterization is the less efficient MS/MS for whole proteins compared to MS/MS of small peptides. Currently, direct MS/MS of a protein with a mass >50 kDa typically generates fragments that give the sequence information primarily for only the two tails of the protein, probably because of the folded protein confirmation in the gas phase.[31,34]

To overcome the limitations of the top-down approach in protein characterization with MS, a few recent studies proposed and implemented a so-called "middle-down" method.[20,40] This strategy produces relatively large peptides (typically 2000-10000 Da compared to <3000 Da for tryptic peptides) either via enzymes or chemicals that cleave proteins at a single type of amino acid residue (e.g. Lys, Glu or Asp), or via mild (short-time) tryptic digestion. These long peptides maintain advantages of the top-down strategy such as high coverage of PTMs on a single peptide chain and enhanced ionization for phosphopeptides, but the middle-down approach decreases the resolution requirement for analysis. Section 1.4 enumerates and discusses some strategies for "middle-down" protein analysis.

1.3 Phosphopeptide enrichment strategies

1.3.1 History of enrichment methods for phosphoprotein analysis

Traditionally, phosphoprotein analysis relied heavily on antibody-based purification. In this strategy, single phosphoproteins are isolated from a cell lysate by immunoprecipitation using immobilized antibodies with specific affinity toward the target protein, followed by gel electrophoresis separation and Western blotting to find the correct protein band for MS analysis.[29] In principle enrichments of phosphoproteins or phosphopeptides using antibodies against phosphorylated serine, threonine or tyrosine residues are possible. However, only phospho-Tyr specific antibodies are available for general phosphopeptide isolation, which hampers routine analysis of proteins containing phospho-Ser and phospho-Thr.[29,41]

An emerging awareness of the importance of protein phosphorylation expedited the development of more general phosphopeptide isolation techniques in recent years. In one of these techniques, reversible covalent binding, selective reactions convert phosphorylated residues to other functional groups (e.g. thiols) that can bind to beads modified with coupling groups (e.g. dithiopyridine). Eluting the beads with small molecules or strong acids releases the peptides for MS analysis.[42] Although very specific in principle, this approach either enriches only the peptides with pSer or pThr residues, or suffers from a high sample loss or side reactions.[42]

The most common current approach to phosphopeptide isolation involves binding phosphopeptides to supports containing immobilized metal ions or metal oxides that have specific affinity toward the phosphates on peptides (Figure 1.3). Due to the complexity of different types of biological samples, recently developed techniques using this concept cover a

wide range of innovations, including novel phospho-specific materials (both new affinity groups and new supporting materials), new phosphopeptide isolation modes, and optimized enrichment procedures using commercially available columns/beads. Several reviews of phosphopeptide enrichment techniques recently appeared.[29,32,42-45] However, these reviews mainly describe new affinity materials and separation methods and do not typically present the drawbacks of these techniques or specify the type of biological samples for which specific techniques are effective. Understanding technical limitations is vital because selection of a proper- but not necessarily the "most powerful"- method for isolating phosphopeptides depends on the sample source and complexity. With this in mind, this section mainly discusses techniques reported in the last three years, with an emphasis on how the features of these methods dictate their applicability to specific biological samples.

1.3.2 Improvement of phosphopeptide enrichment with conventional IMAC-Fe^{3+} beads and TiO_2 particles

Since the first separation of phosphorylated proteins with IMAC in 1986,[46] this technique has been widely used in phosphopeptide enrichment for MS-based characterization of protein phosphorylation.[32,42-45] Although metal oxides have been used more and more often for isolating phosphopeptides in recent years, IMAC is still often employed in large-scale multistage purification of phosphopeptides from complex protein digests.[33,47-50] Among different forms of IMAC materials, commercial resins modified with iminodiacetic acid (IDA)-Fe^{3+} or nitrilotriacetate (NTA)-Fe^{3+} are most common.[42] Resins are either packed in columns (Figure 1.3a) or dispersed in solutions (Figure 1.3b) for capturing and releasing phosphopeptides. The biggest challenge for IMAC-Fe^{3+} is the co-enrichment of acidic non-phosphorylated

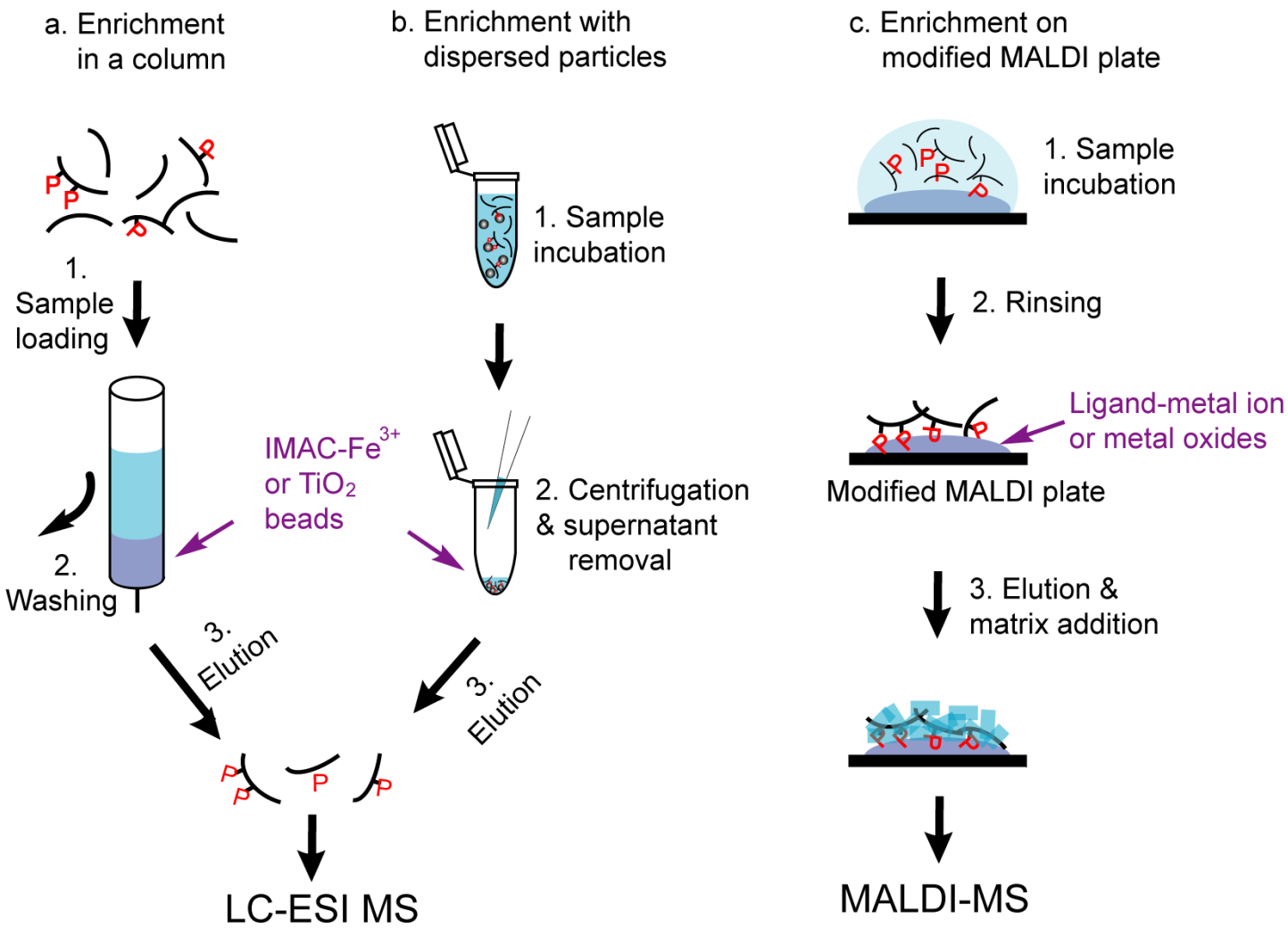


Figure 1.3. Three typical modes of phosphopeptide enrichment: (a) enrichment using a column packed with metal-ion affinity beads or TiO_2 particles, (b) enrichment in a solution using dispersed micro- or nano-size particles with their surfaces modified with metal-ion complexes or metal oxides, and (c) enrichment directly on a flat surface (often a MALDI plate) modified with metal-ion complexes or metal oxides. (For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.)

peptides, which normally contain multiple Glu or Asp residues. Although methyl esterification of carboxylic acid residues can overcome such problems,[51] this step may cause sample loss and decreased peptide solubility, and incomplete or side reactions may result in complicated mass spectra.[47,52] IMAC-Fe^{3+} also suffers from biased enrichment toward multiply phosphorylated peptides.[42] Because the binding affinity of IMAC-Fe^{3+} for mono-phosphopeptides is much weaker than the affinity for multiply phosphorylated peptides (even in mild acids such as 0.1% TFA), mono-phosphopeptides are frequently lost in the flow-through and washing solutions employed during enrichment.[47]

To improve IMAC-Fe^{3+} specificity toward phosphopeptides, Tsai et al. systematically evaluated the influence of pH and buffer concentration on the capture and release of phosphopeptides during their isolation from cell lysates.[53] Interestingly, increasing the acetic acid-sodium acetate buffer concentration while keeping the pH around 3 enhances the phosphopeptide specificity but avoids more acidic conditions that lead to high specificity but low phosphopeptide recovery. The presence of organic solvent (normally acetonitrile, ACN) also affects phosphopeptide binding and release. Jensen and co-workers found that increasing the ACN content in loading and washing solutions enhances the degree of protonation of carboxylic acids relative to protonation of the phosphate groups.[54] Based on this finding, including 60% ACN and 100 mM acetic acid in IMAC loading and washing solutions provides a moderately acidic condition (~pH 3.2) that prevents binding of most of the non-phosphopeptides from a digest mixture with 1:1000 molar ratio of phosphopeptide-to-nonphosphopeptide.

Metal oxides are beginning to displace metal-ion affinity resins for phosphoproteome studies. They are either packed in columns (Figure 1.3a) or dispersed in a solution (Figure 1.3b)

for enrichment. Due to their phosphopeptide specificity and compatibility with a wide range of buffers and pH values, TiO_2 particles are especially popular for phosphopeptide enrichment. Previous studies explored the addition of small organic acids to loading and washing solutions to suppress the binding of acidic non-phosphorylated peptides to TiO_2 particles. The presence of 2,5-dihydroxybenzoic acid, salicylic acid, phthalic acid and benzoic acid can improve phosphopeptide specificity, but some reports suggest that these compounds are sometimes incompatible with LC-ESI and may reduce the number of identified phosphopeptides.[42,55]

Recent work using TiO_2 particles for phosphopeptide enrichment addressed additional factors that influence large-scale phosphoprotein analysis. Li et al. found that the peptide-to-TiO_2 ratio affects specificity in phosphopeptide enrichment from a digest of Hela cell lysate protein.[56] As the peptide-to-TiO_2 mass ratio decreases from 4 to 0.025, the number of identified phosphopeptides and their percentage of the total number of identified peptides initially increase and then decrease. The highest value for both phosphopeptide percentage and total number of phosphopeptides occurs with a 0.5-0.125 peptide-to-TiO_2 mass ratio. Interestingly, the use of higher peptide-to-TiO_2 mass ratios leads to detection of more multiply phosphorylated peptides than the optimized peptide-to-TiO_2 mass ratio. Changing the peptide-to-TiO_2 mass ratio from 4 to 0.05 decreased the ratio of multiply phosphorylated peptides to total phosphopeptides from 90% to 5%. Sequential enrichments using low amounts of fresh beads in the same digest solution identified more and more mono-phosphopeptides, suggesting that unidentified mono-phosphopeptides in early enrichment fractions were left in the supernatant

and not bound to the particles. These findings should be useful for analysis of multiple phosphorylations along a short fragment of protein sequence, which is often related to ordered phosphorylation events.[50]

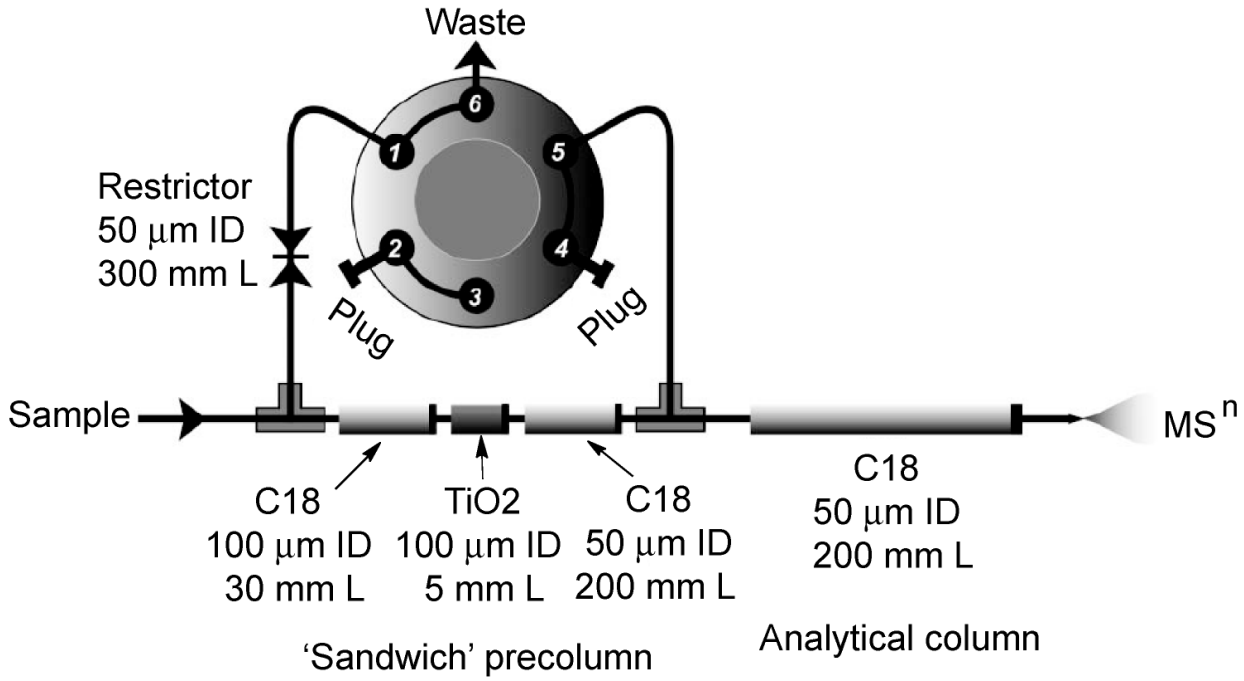


Figure 1.4. Schematic overview of a vented column system with a triple stage precolumn. The "sandwich" precolumn consists of three individual precolumns of C18, TiO_2, and C18 material, respectively. During analysis, all peptides are initially trapped during sample loading onto the first C18 column. After flow splitting, nonphosphorylated peptides are analytically separated and subsequently detected, while phosphorylated peptides are trapped on the TiO_2 precolumn at a very low flow rate. This optimal procedure allows superior binding efficiency, without increasing total analysis time. (Figure is reprinted from reference [55] with permission of American Chemical Society (ACS).)

Heck and co-workers developed a sensitive online TiO_2-based phosphoproteomics system (Figure 1.4).[55] The crucial feature of this system is the C18-TiO_2-C18 "sandwich" configuration, in which the first C18 column (0.2 µL column volume) traps all the peptides at a relatively fast flow rate (3 µL/min), followed by a much slower gradient elution (0.1 µL/min)

that carries the "concentrated" peptides onto the TiO_2 column where the phosphopeptides bind to TiO_2 particles and non-phosphorylated peptides flow through. Injection of a basic buffer into the TiO_2 column elutes phosphopeptides and carries them to the second C18 column for desalting and subsequent separation followed by ESI-MS/MS analysis. The slow flow rate (0.1 μL/min) that carries the concentrated peptides from the first C18 column to the TiO_2 column yields a greatly improved phosphopeptide recovery compared to directly injecting several tens of μL of sample into a TiO_2 column at several μL/min (a low flow rate with such a large volume would result in extended loading times), without any compromise on sample analysis time.

In MS analysis, the presence of many non- and mono-phosphopeptides often suppresses ionization of multiply phosphorylated peptides and hinders their MS/MS analysis.[29] A very recent strategy, sequential elution from IMAC (SIMAC) aims to separately elute mono- and multiply phosphorylated peptides during IMAC to increase the recovery of mono-phosphopeptides and enhance the ionization efficiency of multiply phosphorylated peptides.[47] This procedure includes incubating the protein digest with IMAC-Fe^{3+} beads in a weakly acidic (0.1% TFA) solution followed by initial elution with 1% TFA. Further enrichment of the supernatant and initial eluate on TiO_2 recovers mono-phosphopeptides; while a secondary elution on IMAC-Fe^{3+} beads with pH 11.3 ammonium hydroxide recovers multiply phosphorylated peptides. In SIMAC of digested protein from human mesenchymal stem cells, more than 96% of the identified multiply phosphorylated peptides were in the ammonium hydroxide buffer, and the number of identified mono-phosphopeptides was more than when using TiO_2 alone for enrichment. Similarly, Simon et al.[57] demonstrated that sequential elution

of TiO_2 with pH 8.5 triethylammonium bicarbonate buffer and then pH 11.5 NH_4OH solution releases mono- and multiply phosphorylated methyl esterified peptides, respectively, in the different eluates. They found that the mono-phosphopeptides recovered from the first elution are purer than from a single time elution using a pH 11.5 NH_4OH. The above results suggest an improved enrichment efficiency for both IMAC-Fe^{3+} and TiO_2 after fractionating phosphopeptides.

I should note that reported enrichment conditions or strategies for IMAC-Fe^{3+} or TiO_2 from different groups should be checked carefully when applied to a user's specific enrichment materials and samples. As the supporting materials and ligands for IMAC-Fe^{3+}, as well as the particle size and surface properties of TiO_2 particles vary from product to product, the optimized solvent, pH and sample loading capacity need to be adjusted in order to obtain the best result.

1.3.3 Novel phosphopeptide enrichment materials

Despite the success of IMAC-Fe^{3+} beads (which typically employ NTA or IDA as ligands for binding metal ions) and TiO_2 particles in phosphopeptide enrichment, a number of recent reports describe new enrichment materials with novel ligands, ions, or supporting structures to target different aspects for improving phosphopeptide enrichment. Aiming at enhancing phosphopeptide enrichment specificity, Zou and co-workers prepared a phosphate-containing polymer monolith.[58] Grinding the monolith into irregular beads followed by formation of phosphate-Ti^{4+} complexes (IMAC-OPO_3-Ti^{4+}) yielded a resin that enriched phosphopeptides from model phosphoprotein digests containing a 500 molar excess of non-

phosphoprotein digest. In the phosphoproteome analysis of mouse liver, enrichment with IMAC-OPO_3-Ti^{4+} beads led to enhanced identification of phosphopeptides compared to enrichment using ZrO_2 particles, TiO_2 particles or commercial IMAC-Fe^{3+} beads (2-5 fold increase in both the total number of phosphopeptides and the percentage of detected peptides that are phosphorylated. The authors attributed the superior enrichment with OPO_3-Ti^{4+} to its strong binding of phosphopeptides under the highly acidic condition (6% TFA) and the flexible spacing arm from the supporting polymer. Interestingly, although of all the enrichment materials tested IMAC-OPO_3-Ti^{4+} allowed identification of the largest number of phosphopeptides from a mouse liver lysate digest, the corresponding ratio of multi-phosphorylated to monosphosphorylated peptides was the least. This is consistent with prior studies that suggest a combination of enrichment materials enhances the comprehensiveness of phosphopeptide identification.[27,47,49]

Zou and co-workers also synthesized highly ordered mesoporous silica particles containing OPO_3-Ti^{4+} complexes to isolate phosphopeptides from sera of hepatocellular carcinoma patients and healthy individuals. Such modified silica facilitates the capture of phosphopeptides from human plasma based on both size exclusion of undigested proteins and the high affinity of OPO_3-Ti^{4+} toward phosphopeptides. This material identified mono-phosphorylated human fibrinogen peptide A (FPA) and its three isoforms, and quantitative analysis based on either MALDI-MS or iTRAQ (isobaric tag for relative and absolute quantitation) labeling suggested that the phosphorylated FPA was expressed differently between the cancer and healthy groups.

To overcome the issue of enriching mainly mono-phosphopeptides with Ti(IV)-based materials, Wan et al.[59] synthesized a macroporous ordered silica foam (MOSF) with the surface modified with TiOx that is rich in coordinatively unsaturated Ti(IV) species. The ~100 nm pore size in this material allows for rapid mass transfer, and enrichment of multi-phosphorylated peptides occurs even with low amounts of protein (e.g. 10-min enrichment using 40 μg of Ti-MOSF in a 50 μL solution containing 340 fmol β-casein digest) or large amounts of non-phosphopeptides (e.g. 10:1 mass ratio of BSA/β-casein). The authors attribute the high specificity for binding of multi-phosphorylated peptides to the bidentate phosphate chelation of the 4-coordinate Ti(IV) species that is rich on the material surface. On conventional TiO_2 particle surfaces phosphate binding occurs in a bidentate bridging mode with a 5-coordinated Ti(IV) moiety. As multiple phosphorylation events often occur sequentially within a short stretch of protein, sensitive detection of multi-phosphorylated peptides using such materials will help investigate ordered phosphorylation events on motifs with more than one phosphorylated residue.[50]

Although enrichment can occur with dispersed particles, column-based methods dominate the phosphoproteome field because of their convenience and potential in online analysis. In a recent study, Lin et al. employed liquid-phase deposition to form a thin layer of TiO_2 nanoparticles on the surface of a capillary.[60] This column selectively concentrated phosphopeptides from 2 pmol of α-casein digest in the presence of a 10-fold molar excess of non-phosphoprotein digest. A second deposition of TiO_2 on the capillary wall leads to much improved selectivity and S/N for phosphopeptides. Lin et al. applied this TiO_2-modified capillary to enrich substrate peptides phosphorylated by recombinant Abelson tyrosine kinase

(Abl) in the presence of kinase inhibitors.[61] Addition of an internal standard phosphopeptide that mimics the substrate phosphopeptide during enrichment allows quantitative characterization of Abl inhibitors by MALDI-MS. Because the sample is relatively simple, limited binding capacity is not an issue.

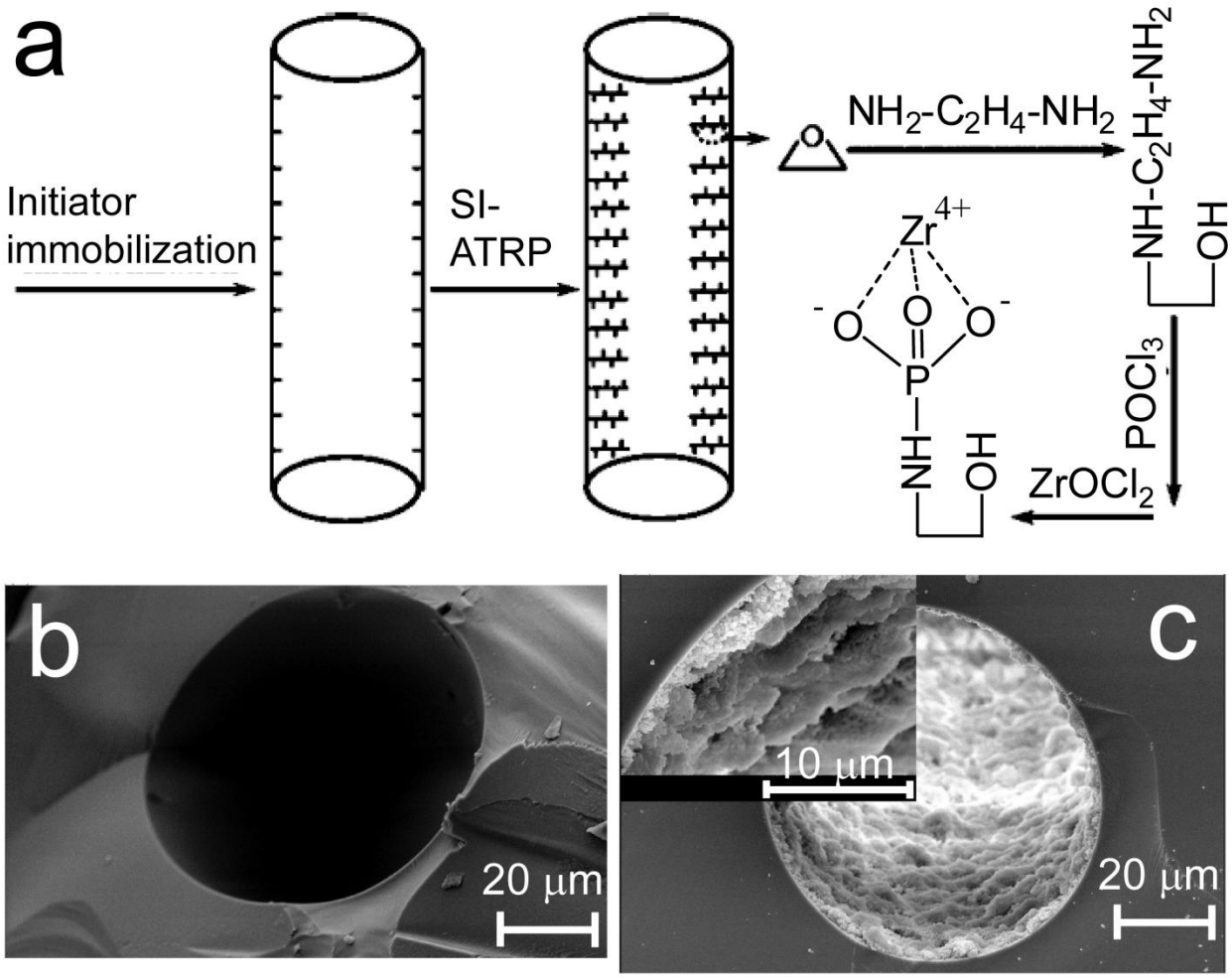


Figure 1.5. (a) synthesis of –PGMA-RPO_3-Zr^{4+} brushes in a capillary column using ATRP of glycidyl methacrylate and subsequent derivatization. (b, c) SEM images of the capillary before (b) and after (c) 20 h of polymer grafting. Figures a, b and c were adopted (with modifications) from reference [62] with permission of ACS.

To increase the binding capacity of capillary columns for large-scale phosphopeptide enrichment, Qin et al.[62] recently synthesized 3 to 5 μm-thick polymer brushes on the inner

wall of a capillary (75 μm I.D. × 50 cm). After surface-initiated atom-transfer radical polymerization (SI-ATRP) of glycidyl methacrylate to form the brush, modification of the polymer with a spacing arm and terminal zirconium phosphonate gives a stationary phase that specifically binds phosphopeptides (Figure 1.5). The loading capacity of a synthetic mono-phosphopeptide in the modified capillary is ~120 pmol (0.054 μmol/mL), more than 1 order of magnitude higher than the capacity of a conventional open tubular capillary column developed by the same group.[63] This polymer-modified capillary column can enrich 5 pmol of α-casein phosphopeptides in the presence of a 100-fold molar excess of BSA digest under mildly acidic conditions (0.1% TFA with 80% ACN). Moreover, phosphopeptide isolation with these columns after SDS-PAGE fractionation and in-gel digestion facilitated the identification of ~500 unique phosphopeptides from 400 μg of human Hepatocellular carcinoma cell (Hep G2) lysate.

Hou et al.[64] prepared an organic-inorganic hybrid silica monolith column with immobilized $-RPO_3\text{-}Ti^{4+}$ for IMAC enrichment of phosphopeptides. The porous structure provides a much higher loading capacity for a synthetic phosphopeptide (1.4 μmol/mL) than the capillary column modified with polymer brushes.[62] In addition, the organic-inorganic hybrid structure is superior over pure silica monoliths and organic monoliths because of its stability over a wide range of pH (pH 0.5-11.8) and organic solvent content (i.e. 50% ACN). This high stability permits the use of mobile phase conditions that minimize co-enrichment of non-phosphopeptides or enhance elution efficiency. In the analysis of a mitochondrial-protein-enriched fraction from 2 g of rat liver tissue, enrichment with the monolith column led to identification of 224 phosphopeptides from 148 phosphoproteins. However, multi-phosphorylated peptides accounted for only 7.1% of the total number of phosphopeptides.

The above novel materials present different advantages in isolating phosphopeptides. The dispersed particles may effectively enrich phosphopeptides from moderately complex samples, such as certain fractions of cell lysate. Rapid mass transfer in swollen polymer-based affinity materials or monoliths with small pore size may enhance the rate of phosphopeptide capture and elution relative to isolation with columns containing commercial IMAC beads or TiO_2 particles. If so, the new materials might expedite high-throughput phosphopeptide analysis, which requires short loading and elution times. However, careful comparisons with commercial columns must be performed to make any firm conclusions.

1.3.4 Ion exchange for phosphopeptide fractionation

For complex biological samples, e.g. digests from whole cell lysate or tissue extracts, direct enrichment of hundreds to thousands of phosphopeptides using metal ions or metal oxides is not efficient. More comprehensive isolation of phosphopeptides from such samples typically relies on pre-fractionation of peptides using ion exchange, particularly SCX where positively charged peptides bind to the negatively charged stationary phase via electrostatic interactions.[33,49,50] Elution with a salt and pH gradient releases peptides according to their charge.[65] A theoretical tryptic digestion of proteins in the human protein database revealed that 68% of the predicted peptides have a +2 net charge at pH 2.7.[65] Phosphopeptides generally elute earlier than non-phosphorylated peptides during gradient salt elution, and multiphosphorylated phosphopeptides may not bind to SCX columns at all due to the negative charge of peptides carrying multiple phosphate groups.[65] As SCX eluates contain both phosphopeptides and non-phosphorylated peptides, even in the early eluting fractions,

performing IMAC for each fraction typically improves phosphopeptide identification in the later LC-MS/MS analysis.[33,48]

In a phosphoproteome study of 10 mg of mouse liver tissue,[49] Gygi's group first separated tryptic peptides into 15 SCX chromatography fractions (Figure 1.6). Subsequent phosphopeptide enrichment by IMAC-Fe^{3+} for each fraction, along with a phospho-Tyr specific antibody enrichment for the total lysate, led to identification of 4038 unique phosphopeptides that stem from 2149 proteins. However, analysis of SCX fractions 4-10 revealed exclusively peptides with +1 charge states, suggesting a low resolution in SCX. Similarly, in a large-scale study of phosphoproteins from postsynaptic density samples using SCX-IMAC, Trinidad et al. noticed a large elution window for all peptides of a given charge state, and overall elution in order of charge state.[48] In another phosphoproteome analysis,[50] the SCX-IMAC strategy revealed 13720 non-redundent phosphorylation sites from 2702 proteins, although 70% of the phosphopeptides were multi-phosphorylated, suggesting that IMAC still suffers from biased enrichment of multi-phosphorylated peptides, even after pI-based peptide fractionation. Nie et al. compared several combined enrichment methods, including SCX, strong anion exchange (SAX) and TiO_2, in phosphopeptide enrichment from a Hela cell lysate digest. Enrichment of phosphopeptides using TiO_2 alone and combined SCX/TiO_2 reveal identical distributions for the calculated pI values of the enriched peptides and some bias toward basic or mono-phosphopeptides.[66] Heck and coworkers[55] showed similar pI distributions for peptides obtained from solely TiO_2 and combined TiO_2/SCX strategies, suggesting that TiO_2 and SCX provide parallel enrichment selectivity in isolating phosphopeptides.[55,66] These findings again suggest that separate IMAC-Fe^{3+} and TiO_2 phosphopeptide enrichment will increase the

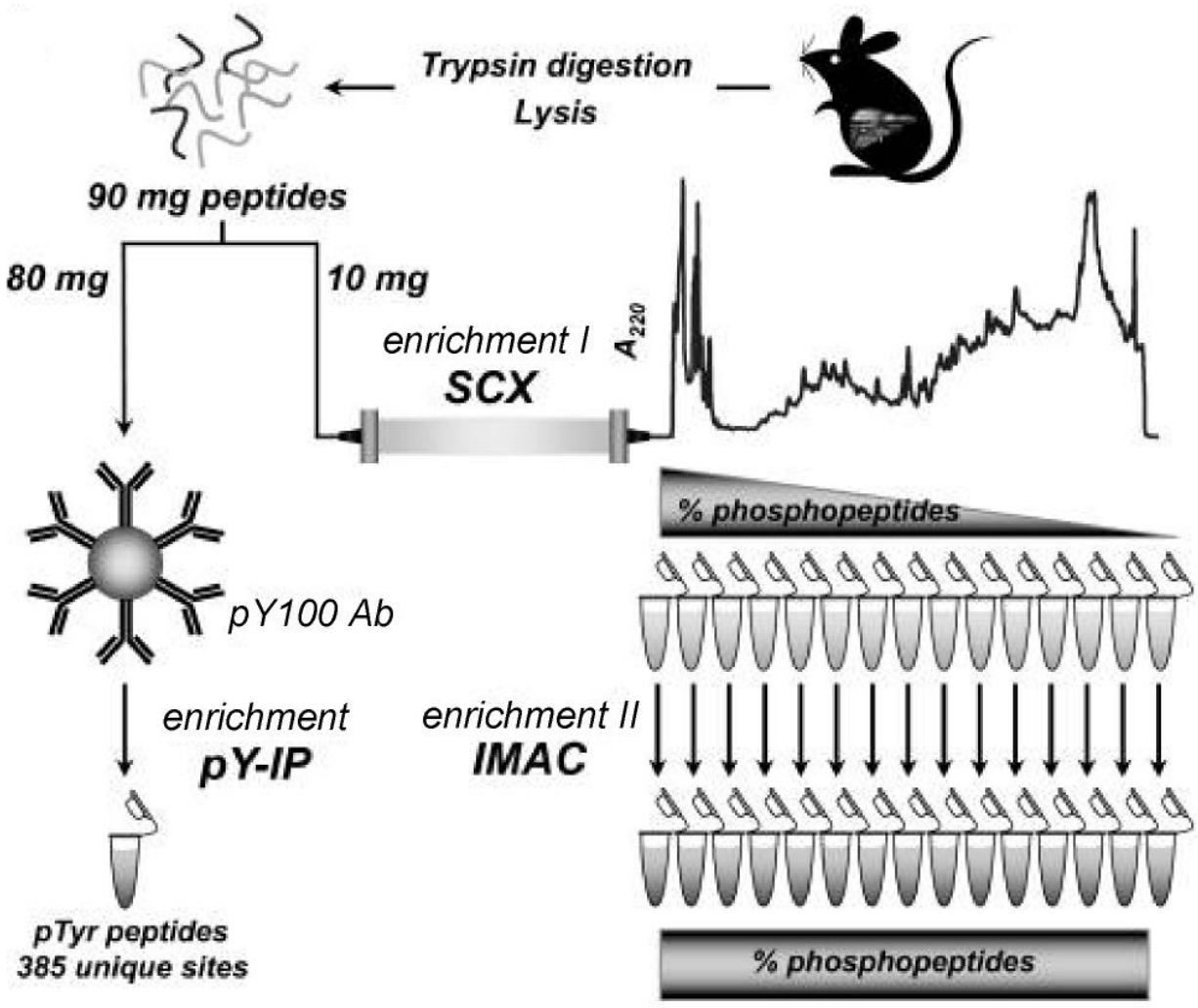


Figure 1.6. Strategy used by Gygi et al. for large-scale identification and characterization of phosphorylation sites in mouse liver proteins. After homogenization and lysis, 10 mg of protein was tryptically digested and subjected to a two-step phosphopeptide enrichment. SCX fractionation isolated many phosphopeptides in early eluting fractions, and subsequent IMAC of each fraction provided additional selective capture. In addition, 80 mg of tryptic peptides were enriched for pTyr-containing peptides by immunoaffinity purification. Figure is reprinted with permission from reference [49] with modifications.

comprehensiveness of phosphoproteome studies, even when SCX is utilized for peptide pre-fractionation.

Compared with SCX, SAX chromatography provides an alternative or even superior approach to peptide fractionation before phosphopeptide enrichment. A SAX column binds

negatively charged peptides at weakly acidic or neutral conditions and then releases these peptides during elution with salt or pH gradients.[66,67] Due to the negative charge of the phosphate groups, phosphopeptides typically elute later than non-phosphorylated peptides.[66,67] In a study of the phosphoproteome in a human liver,[67] both SAX chromatography and IMAC-Fe^{3+} enriched acidic peptides, but SAX separated phosphopeptides from acidic non-phosphorylated peptides while IMAC-Fe^{3+} enriched both species simultaneously. In addition, SAX chromatography had fewer unbound phosphopeptides than separation with SCX.[67] Two studies found that SAX typically shows better phosphopeptide fractionation resolution (the width of fractionation window for a given charge state) than SCX and enriches more multiply phosphorylated peptides than TiO_2 or SCX.[66,67] Nie et al. showed that an online continuous pH gradient elution in SAX and offline enrichment of any unbound phosphopeptides using TiO_2 leads to a very comprehensive coverage of both mono- and multiply-phosphorylated peptides.[66]

Large-scale phosphopeptide enrichment typically identifies hundreds to thousands of phosphoproteins, which allows statistical analysis of phosphorylation patterns. For example, the mouse liver phosphoproteomics study by Gygi[49] revealed both novel and known motifs for specific Ser/Thr kinases, with more frequent phosphorylation close to the protein C-termini than in other regions. They also found some cases in which only one of the two specific loci in a double phosphorylation motif existed as a mono-phosphorylated species, suggesting sequential phosphorylation of the two sites. In all, phosphopeptide isolation strategies are flexible, and combinations of different techniques could be employed in different situations.[66,68,69]

1.3.5 Phosphopeptide enrichment directly on modified MALDI plates

Enrichment on an affinity surface through incubation of a drop of sample followed by rinsing is one of the most direct methods for phosphopeptide isolation. After subsequent addition of a drop of eluent and matrix, such surfaces can serve as MALDI plates for MS or MS/MS analysis (Figure 1.3c). This "on-plate enrichment" is simple and minimizes consumption and loss of samples.[42] With recent advances in surface chemistry, a number of affinity interactions previously employed for on-column enrichment have been implemented in "on-plate" methods with sensitive and specific enrichment of phosphopeptides from digests of model proteins.[42,70] Notably, studies published in the last two years addressed enrichment conditions specific for MALDI sample preparation, and more importantly, began applying this approach to biological samples.

Hoang et al. immobilized Zr^{4+} on spots of phosphonate-terminated self-assembled monolayers for phosphopeptide enrichment and direct MALDI-MS analysis.[30] The Zr^{4+}-modified spots are surrounded by a hydrophobic monolayer that confines a 5 μL sample for analysis. A comparison study showed that the film with Zr^{4+} isolates phosphopeptides more specifically than films with Ti^{4+} and Fe^{3+} ions. Although such films can enrich phosphopeptides in samples containing 100 fmol of phosphopeptide and 1 pmol of nonphosphoprotein digest, a spot has a binding capacity of only ~650 fmol of β-casein phosphopeptides because of the limited number of affinity groups in the monolayer. On-plate phosphopeptide isolation using this film, together with MALDI collision-induced dissociation (CID)-MS/MS and Mascot database searching, identified a phosphopeptide from recombinant human MAP Kianse1, as well as several phosphoproteins digested from spots in a 2D electrophoretic gel.

Blacken et al. fabricated several metal oxide-coated steels by reactive landing of gas-phase ions produced by electrospray ionization of group IVB metal alkoxides.[71] In enrichment and MALDI-MS analysis of α-casein/BSA (1:2 molar ratio) digests on several metal oxide surfaces, ZrO_2 gave higher intensities of mono- and doubly phosphorylated α-casein peptides than HfO_2 and TiO_2 surfaces. The authors attributed the higher intensities on ZrO_2 to the optimal balance of surface coverage and metal oxide surface area compared to the other two materials. Unfortunately, the influence of the intrinsic metal oxide properties on phosphopeptide binding and elution is not clear. The ZrO_2 surface enriched a phosphopeptide from a digest of c-Src tyrosine kinase, which had been dephosphorylated and then autophosphorylated with ATP for different times. Addition of an internal standard phosphopeptide in the differentially treated digest samples and subsequent enrichment allowed monitoring of the kinetics of the autophosphorylation.

Similarly, Niklew et al.[72] fabricated spots (1-mm dia.) of anatase TiO_2 nanoparticle films on F-doped conductive SnO_2 using a screen-printing technique (Figure 1.7a-c). The advantage of such films is their high internal surface area compared to flat surfaces, as revealed by isolation of phosphopeptides from a 1-μL digests mixture containing 15 pmol of BSA and 4 pmol each of α-casein and β-casein. This TiO_2 film also enriched phosphorylated fibrinopeptide A from a human fibrinogen digest (m/z 1617 in Figure 1.7d). Moreover, dephosphorylation of bound β-casein peptides by phosphatase treatment directly on the TiO_2 surface was possible, although less efficient for the dephosphorylation of multiply phosphorylated peptides (Figure 1.7 e and f).

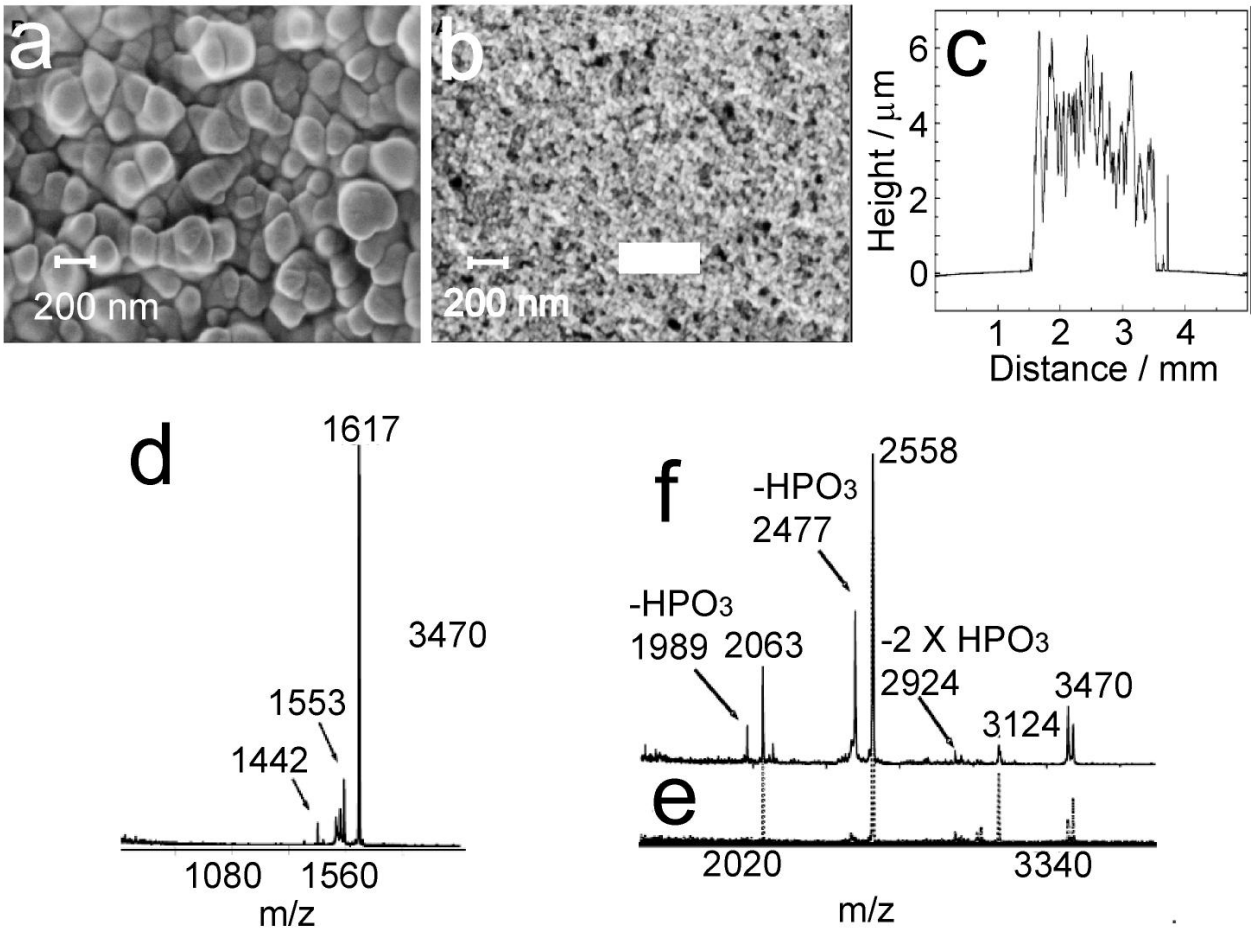


Figure 1.7. SEM images of F-doped SnO_2 film before (a) and after (b) modification by screen printing of TiO_2. (c) A profile scan of one spot by interference spectroscopy illustrates the average layer thickness on that spot. (d) MALDI mass spectra of the thrombin-digested fibrinogen (60 pmol) after sample enrichment on the TiO_2 film. (e, f) MALDI-TOF mass spectra of β-casein phosphopeptides enriched on the TiO_2 film before (e) and after (f) treatment with protein phosphatase directly on the TiO_2 film. Figures were adopted (with modifications) from reference [72] with permission of ACS.

Porous surface structures can potentially maintain fast mass transfer and increase binding capacity, and a homogeneous surface is crucial for obtaining reliable enrichment results. One of the limitations of on-plate enrichment is the difficulty of performing multi-step washing or elution (e.g. SIMAC in column-mode phosphopeptide enrichment[47]), or multi-dimensional

enrichment with more than one phospho-affinity material (e.g. tandem enrichment using IMAC and TiO_2), to improve phosphopeptide recovery. Another drawback lies in the low binding capacity (typically from low pmol to several hundred fmol of phosphopeptides) on flat surfaces compared to columns. Last, MALDI-MS/MS is usually less rich in information than ESI-MS/MS due to the low fragmentation efficiency of singly charged peptide ions in the gas phase produced by MALDI.[30] However, MALDI-MS also has several advantages over ESI-MS for phosphopeptide analysis. (1) The domination of singly charged ions greatly simplifies mass spectra compared to those produced by ESI. (2) The pulsed nature of MALDI makes possible a data-driven, selective MS/MS process. (3) Matrix/sample spots are often stable for several days and can be investigated multiple times. Considering the above disadvantages and advantages, on-plate enrichment is primarily applicable to analysis of limited amounts of target phosphoprotein(s), rather than large scale identification and characterization of unknown proteins. Possible applications of on-plate phosphopeptide enrichment include high-throughput protein kinase assays and analysis of antibody-isolated phosphoprotein(s) or protein digests from gel electrophoresis bands.

1.4 Middle-down protein analysis by MS

1.4.1 Advantages of middle-down protein analysis

Long proteolytic peptides (typically 2000-15000 Da) have several advantages for peptide/protein characterization: (1) They are generally more hydrophobic than short peptides and separate with higher resolution in LC.[73] (2) Compared to small tryptic peptides, which mainly generate y-ions during CID-MS/MS due to the basic Arg or Lys residues primarily at the

C-terminus, large peptides carry higher numbers of Asp or Lys along the peptide chains and thus give more efficient MS/MS fragmentation.[73,74] (3) For peptide sequence identification by database searching after MS/MS fragmentation, larger precursor ions have fewer potential protein candidates in the database than small ones. In addition, when using high mass accuracy instrumentation such as Fourier transform MS, setting strict precursor ion mass tolerances for database searching greatly improves the confidence of peptide assignments. In other words, identification of large peptides generally requires fewer fragment-ion assignments than identification of small peptides.[74] (4) Large peptides provide higher coverage of PTM sites than short tryptic peptides, facilitating the study of PTM coordination at nearby regions of the protein.[20,35] In all, the middle-down strategy provides many advantages of top-down or bottom-up strategies, but avoids (to some extent) the instrumental challenges of the top-down method (e.g. limited MS/MS efficiency and the demand for high mass resolution), as well as the requirement of multidimensional separation for complex mixtures of small peptides in the bottom-up method.

Several studies showed that the middle-down approach is an important complementary method to MS characterization of the whole protein.[31,36,75,76] For example, the Kelleher group[36,76] demonstrated the power of a middle-down approach in localizing multiple PTMs at histone H3 tails (1-50 residues) using Glu-C digestion followed by ECD MS/MS of the large peptide. They found that ECD of the whole histone H3 protein generates abundant fragments mainly from the unmodified part of the protein, while performing ECD of the tails improves localization of the majority of histone H3 modifications.[36] Similarly, Ge *et al.* unambiguously identified all of the phosphorylation sites in the truncated and full-length forms of cardiac myosin binding protein C (cMyBP-C) (142 kDa) using top-down and middle-down MS.[31] This

combination of approaches ensures the full sequence coverage of cMyBP-C. Other middle-down examples include: analysis of the length and linkages of polyubiquitin chains by combing partial digestion and LC-MS,[77] high-confidence identifications of ~30% of the highly basic ribosomal proteome with an average 46% sequence coverage of 70-79 possible parent proteins,[73] enhanced determination of multiple phosphorylation sites on recombinant nucleotide-binding proteins in *Escherichia coli* (*E. coli*),[78] and strengthened identification of phosphorylation and glycosylation sites of proteins using large peptides with the combination of CID, ETD, and CID of an isolated charge-reduced (CRCID) species derived from ETD.[79]

1.4.2 Middle-down proteolysis methods

1.4.2.1 Targeted proteolysis using proteases with single-residue specificity

To increase the size of proteolytic peptides, a number of studies employed proteases that cleave proteins at only one residue that appears relatively infrequently along the protein chain. The most common proteases cleave proteins at multiple types of amino acid residues and give rise to relatively short peptides. These low-specificity enzymes include: pepsin (cleaves peptide bonds at hydrophobic and preferably aromatic amino acids such as phenylalanine, tryptophan, and tyrosine)[80], chymotrypsin (cleaves at the carboxyl side of aromatic amino acid residues, including tyrosine, tryptophan and phenylalanine)[81] and trypsin (cleaves at the carboxyl side of Lys or Arg except when either is followed by Pro).[82] In contrast digestion with Glu-C (cleaves at the C-terminus of Glu),[83] Lys-C (cleaves C-terminally at Lys)[84], and Asp-N (cleaves N-terminally at aspartic and cysteic acid)[85] may lead to somewhat larger proteolytic peptides. Examinations of the abundance of amino acid residues in human proteins show similar fractions for Asp, Lys and Glu (~5%).[74,86] However, for a target protein or protein family, a

careful examination of the abundance of Glu, Lys or Arg is important for selecting the proper protease(s). For example, to circumvent the limitations of top-down analysis of multiple PTMs on histone, several recent studies focused on the characterization of residues 1-50 where most PTMs occur.[35,36,75,76] Glu-C digestion of histone provides this large PTM-containing peptide because of the sparsity of Glu in this region.

In some cases separate digestion with several proteases facilitates complete sequence coverage with large peptides. Ge et al. clearly demonstrated this strategy.[31] In top-down analysis of a truncated form (C0-C4, 642 residues) of cMyBP-C using high resolution ESI/FTMS with ECD- and CID-MS/MS, they identified the sequence and PTMs of only the two tails of the whole protein. To characterize the central region of C0-C4 (covering 53% of the amino acid residues), the Ge group separately performed middle-down digestion with three different proteases, Asp-N, Glu-C and Lys-C. They sequenced 54 resulting peptides, mostly with molecular masses from 2500-10000 Da, and the peptide revealed the entire sequence and all phosphorylation sites.

Some studies employing Lys-C for protein digestion also show the advantage of large peptides in characterizing protein PTMs. Lys-C digestion of a set of recombinant nucleotide-binding proteins from *E. coli* followed by MALDI QTOF MS/MS greatly improved the identification of multiply phosphorylated peptides compared to LC-ESI-MS/MS after tryptic digestion.[78] The preserved positive charge from Arg within the peptide chain may explain the enhanced ionization of phosphopeptides.[78] Wu *et al.* showed similarly improved PTM characterization for β-casein (a 24 kDa model protein with 5 phosphorylation sites) and human epidermal growth factor receptor (EGFR, a 132 kDa protein with both phosphorylation and glycosylation).[74] They observed that the Lys-C digestion of the two proteins into big peptides

followed by sensitive online LC-LTQ-FTICR resulted in much higher sequence coverage than tryptic digestion (e.g. 97% vs. 36% for 50 fmol of β-casein, and 95% vs. 70% for 1 pmol of EGFR). In addition, compared to small tryptic peptides, the middle-down strategy improves the ionization of peptides with PTMs, as well as the b-ion intensity in CID-tandem MS. These features led to localization of all the five phosphorylation sites on β-casein, as well as 10 glycosylation sites and 3 phosphorylation sites on EGFR.

1.4.2.2 Targeted proteolysis using acid cleavage

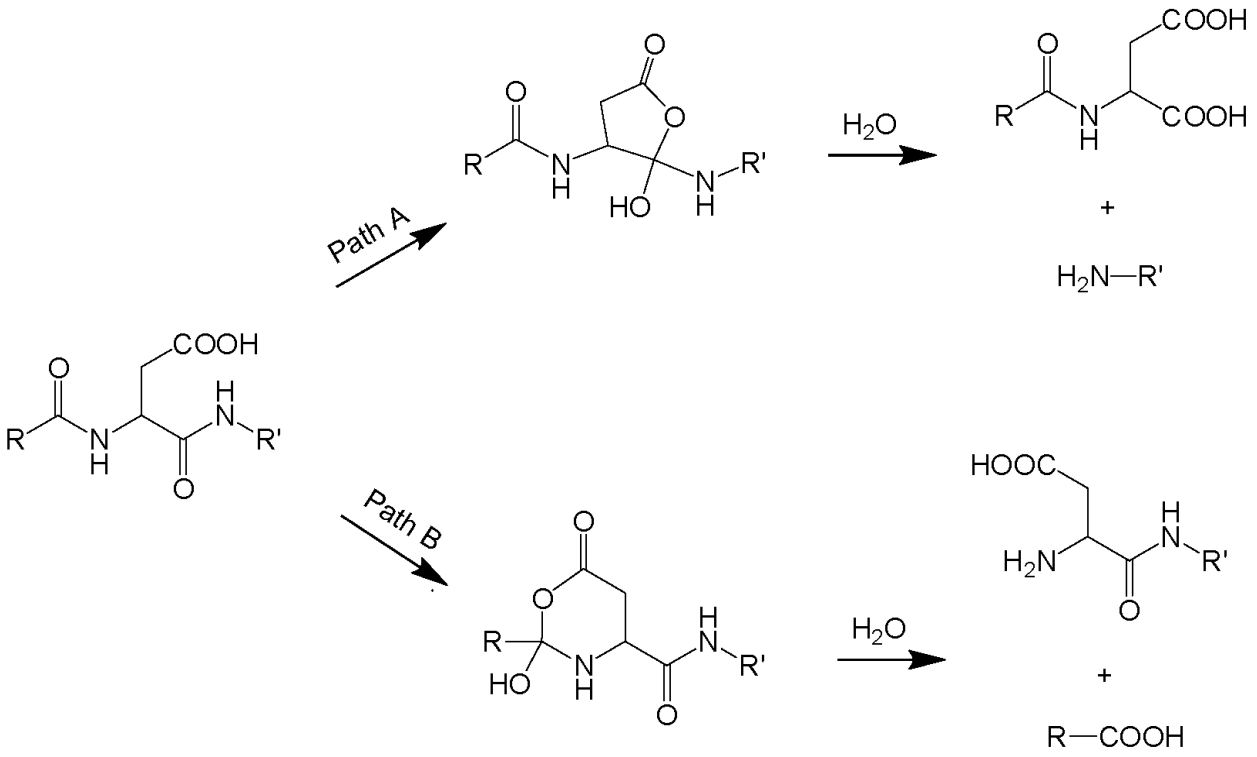


Figure 1.8. Mechanisms of the acid cleavage at Asp-X (path A) and X-Asp (path B) bonds (X represents any amino acid residue). The figure was adapted from reference [86].

In 1950, Partridge et al. found that heating proteins in solutions containing weak acids preferentially releases aspartic acid from the protein chains.[87] Based on this discovery and subsequent work,[88] several studies employed acid cleavage (e.g. using formic acid or acetic

acid) to digest proteins at either the C-terminal or N-terminal sides, or both sides, of Asp (Figure 1.8) prior to MS analysis of these peptides.[86,89,90] High temperature (>108 ℃) and low pH (<2.1) are vital for achieving rapid and efficient digestion,[89,90] and the use of microwave-assisted Asp-specific acid cleavage can shorten the digestion time from hours to 0.5-20 min.[91] In contrast to enzyme-based digestion, acid cleavage can take place in a salt-free system to avoid the need for sample desalting prior to MS analysis. Thus, in some cases acid digestion leads to a mass spectrum with less noise than in spectra obtained after enzymatic digestion.[86,92] Additionally, acidic solutions (~pH 2) readily denature many proteins to increase their solubility without adding denaturants such as urea. Finally, acid digestion avoids the production of unwanted peptides due to protease autolysis.

In human proteins, 4.9% of the amino acid residues are Asp, which is a little lower than the abundances of Lys (5.7%) and Arg (5.6%).[86] Thus Asp-cleavage on human proteins should in principle occur at 43% of the frequency of trypsin cleavage (complete digestion) to give relatively large peptides. Fenselau and coworkers studied microwave-assisted acid digestion of several types of proteins,[73,89-91,93] including digestion of highly basic ribosomal proteins for middle-down analysis.[73] In the human ribosomal database, Asp residues account for only 3.8% of total residues, in contrast to 9.3% and 12.4% for Arg and Lys, respectively.[73] *In silico* digestion of 84 human ribosomal proteins suggests that ~60% of tryptic peptides have ≤10 residues, and ~75% of Lys-C digested peptides have ≤20 residues. In contrast, both the *in silico* and experimental results show that more than two thirds of the acid-digested peptides have >20 residues.[73] These large peptides led to an average 46.2% sequence coverage for 70 ribosomal proteins after LC-LTQ-orbitrap analysis. The authors attributed the high sequence coverage to (1) the multiple internal basic residues that readily produce multiply charged precursor ions and

product ions (by CID), and (2) reliable identification of long peptides with only limited numbers of product ions when using high-resolution mass measurements of precursor and product ions.

In another work, the Fenselau group digested bacteriophage MS2 coat protein (13728 Da with 129 amino acid residues including 4 Asp) into several large peptides (9000-12500 Da) that have sequence overlaps and cover the whole protein sequence.[89] After examination by MALDI-TOF MS, these large peptides help distinguish the MS2 strain coat protein from bacteriophage R17, a protein with only a 1 Da difference from the MS2 strain protein due to a sequence reversal at two sites and an Asp/Asn replacement. Such a minor mass difference was not detectable when the whole proteins were examined with MALDI-TOF. Unfortunately, acid cleavage shows relatively low digestion efficiency (in terms of consumption of intact proteins without considering the peptide size) compared to tryptic digestion, even with the assistance of microwave radiation. However, Fenselau suggests that the parallel analysis of digest products and intact protein masses improves the specificity in protein source (e.g. microorganism) identification.[90]

Although acid digestion is attractive, it has several disadvantages. First, as noted above the protein digestion efficiency using small organic acids is generally lower than that with tryptic digestion.[90,92] Second, due to the uncertainty of the cleavage side at Asp (C-terminal or N-terminal or both sides), the number of product peptides could be higher than expected (especially when the digestion time is short).[86,93] Moreover, peptides with a single terminal Asp at a peptide N- or C-terminus have the same mass (e.g. DGEWQQVLNVWGKVEA and GEWQQVLNVWGKVEAD), complicating peptide sequence identification based on m/z values.[86,89,90] The third challenge in acid digestion is the appearance of side reactions, such as dehydration and formylation when the digestion time is beyond 2 h (no microwave), that

complicate peptide sequence identification.[86] Moreover, variable extents of removal of glycans and phosphates from proteins during acid cleavage can compromise characterization of PTMs.[91,93]

In addition to acid cleavage, other chemicals were employed to generate long peptides by cleavage at Trp residues.[94,95] Vestling *et al.*[95] reported the digestion of model proteins specifically at Trp residues using (3-bromo-3-methyl-2-(o-nitrophenylsulfeny1)indolenine. However, many side reactions occurred, including oxidation of Trp, Met and Cys (forming disulfide bonds) and substitution of bromine for hydrogen (probably on Tyr).

1.4.2.3 Partial proteolysis for generating large peptides

Partial digestion, which can occur at short digestion times or with a low enzyme-to-substrate ratio, provides another strategy for obtaining relatively large peptides.[77] A recent MS characterization of polyubiquitin (polyUb) chain structure by Xu et al. provides an excellent example of exploiting incomplete digestion.[77] Ub is a 76-residue polypeptide with 7 Lys and 4 Arg residues. In polyUb, the N-terminus (GG-NH_2) of one Ub is linked with any primary amine (from either Lys or the N-terminus) of another Ub. With a 1:3 trypsin-to-substrate ratio, ubiquitin or polyUb in its native (folded) form cleaves completely only at R74 in the C-terminal tail (RGG-NH_2), even after incubation at 37 ℃ for 9 h. Such a specific cleavage produces an almost full-length Ub (residues 1-74, UbR74) and a linkage-specific GG-tagged UbR74. Combining this digestion strategy with LC-MS, Xu et al. analyzed the length and linkage sites of human and yeast polyUb chains.

Limiting the digestion time can also lead to partial digestion. Walcher et al.[96] reported that 3-min tryptic digestion of the type I antenna protein of spinach photosystem II (Lhcb 1)

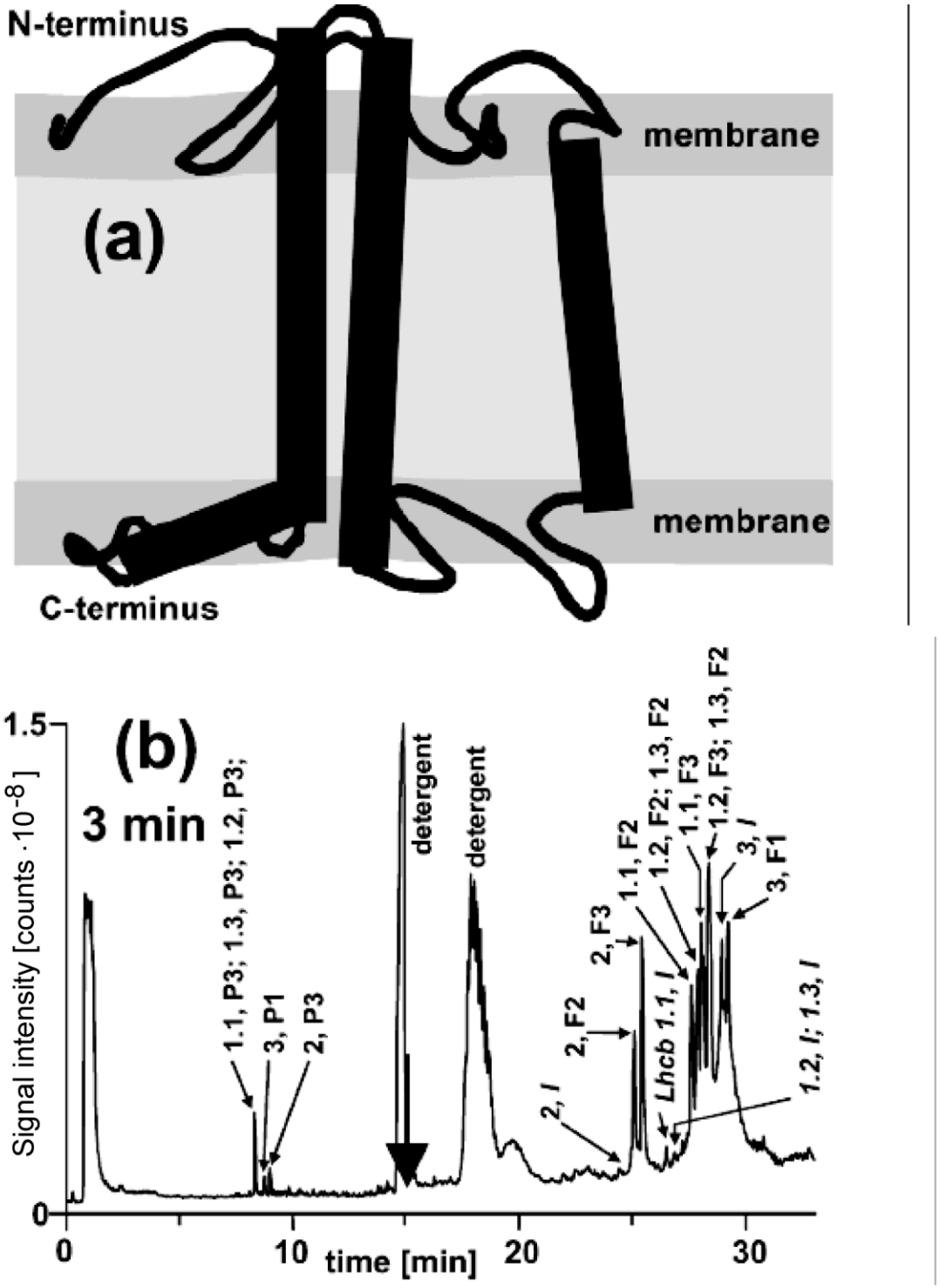


Figure 1.9. (a) Molecular structure of the pea Lhcb1 antenna protein showing the highly conserved transmembrane helices and the more variable amino terminus. (b) Liquid chromatography-ESI MS of the Lhcb antenna proteins (from a gel-electrophoresis band) after 3-min typtic digestion for simultaneous analysis of intact proteins (identified by index I), truncated proteins (identified by indexes F1-F3), and aminoterminal peptides (identified by indexes P1-P3). Figure is reprinted with permission from reference [96].

under non-denatured conditions releases only a few small peptides near the protein N-terminus. In contrast, the structure of the majority of the protein remains, probably due to restricted access of trypsin to the stable 3-D structure of the transmembrane part of the protein (Figure 1.9a). RP-HPLC (monolith)-ESI-MS and -MS/MS successfully separated and identified the tryptic products, including small peptides, truncated proteins and intact membrane proteins. This strategy enabled the identification and characterization of a new isoform of Lhcb 1 (Figure 1.9b). Short-time digestion has also been employed when using proteases with higher specificity such as Glu-C.[31]

Compared to digestions using pure acids or proteases with high cleavage specificity (e.g. Glu-C and Asp-N), partial tryptic digestion has not been systematically explored. Examples of deliberate partial digestion employ proteins in their native confirmation,[77,96] perhaps because proteolysis of unfolded protein by trypsin is uncontrolled. Limited studies suggest that tryptic digestion is a complicated process: the proteolysis occurs either in a one-by-one scheme (the protease breaks only one bond in a substrate protein before immediate dissociation) or a zipper scheme (the protease remains bound to the protein until digestion is complete).[97,98] Although less employed for middle-down protein analysis than high specificity proteases such as Glu-C, trypsin is potentially attractive for middle-down studies because it is available in mg quantities at relatively low cost. This allows for inexpensive immobilization of trypsin on different supports for applications such as online protein digestion-LC-ESI-MS analysis.[99] Thus, using trypsin for partially digesting proteins is potentially viable for routine, especially large-scale, middle-down protein MS analysis. Careful studies on tryptic digestion pathways will be essential for controlled, partial proteolysis. This dissertation describes immobilization of trypsin in membranes to allow extremely short digestion times that yield large peptides.

1.5 Tandem MS and data analysis for peptide sequencing, protein identification and phosphorylation localization

Tandem mass spectrometry is vital for peptide or protein sequencing and PTM characterization.[20,41] Collision-induced dissociation (CID), which increases the internal energy of gas-phase peptide/protein ions relatively slowly (μsec-msec) to induce backbone cleavage through the lowest-energy dissociation pathway, is the most widely used fragmentation technique for MS/MS.[41] This process breaks peptide chains at the C-N bond to give a series of b- and y-fragment ions. For phosphopeptides, a prominent neutral loss of H_3PO_4 (m/z 98), along with neutral losses of H_2O and NH_3 are also very common in CID.[20] Although the neutral loss of 98 facilitates phosphopeptide identification, it hampers backbone fragmentation for peptide sequencing and assignment of the phosphorylation site. Performing MS/MS/MS (MS^3) in an ion trap, through isolation of the neutral loss ion followed by CID and analysis of the fragment ions, may overcome the lack of backbone fragmentation in phosphopeptides. This method was employed in phosphoproteome analysis with software-controlled, neutral loss-dependent MS^3,[65] but the MS^3 efficiency is high primarily for abundant ions.[29] Pseudo MS^3, where fragmentation of the precursor ion and the ion originating from the neutral-loss occur simultaneously, can also improve database searching based on fragment ions.[100]

Two alternative fragmentation methods, ECD and electron transfer dissociation (ETD), show tremendous promise in fragmentation of large peptides and peptides with PTMs.[20,41] In ECD, capture of a thermal electron by a peptide ion results in a nonergodic dissociation that produces fragmentation largely independent on the peptide sequence.[101] Due to this advantage,

the labile phosphate groups remain on the resulting c and z fragment ions, enabling phosphorylation site assignment. However, the requirement of a static magnetic field for producing thermal electrons restricts the use of ECD to FT-ICR. Similarly, ETD employs anions with low electron affinities to reduce the charge of peptide ions and induce fragmentation via a nonergodic pathway to produce c and z ions.[102] Due to its implementation in lower-cost instruments such as quadrupole ion traps, ETD has great potential for global analysis of protein phosphorylation, especially for peptides with multiple charges (≥3+) or multiple phosphorylation sites.[41,102] Moreover, implementation of ETD on a hybrid linear ion trap-orbitrap enables sequencing large peptides and proteins.[103] Recent studies showed that the combination of CID and ETD may yield efficient fragmentation for peptides with a wide range of charge sates.[79,103]

Finally, computational approaches and software tools are vital for rapidly interpreting the data in MS/MS spectra.[104] Widely used MS/MS data analysis tools include Sequest, MASCOT, Phenyx, X!Tandem, Sonar and OMSSA, which commonly identify 5–30% of the MS/MS spectra data set, with an incomplete list of predefined protein modifications for PTM screening.[105] In a typical MS/MS data analysis using one of the above tools, the proteins in a database (e.g. Human, chosen according to the sample category) are digested into peptides *in silico* according to the cleavage rules of the protease used for sample digestion. Each peptide generates a theoretical MS/MS spectrum for comparison with the experimental MS/MS data to generate a similarity score. The highest score corresponds to the most possible peptide sequence (with modifications if the variable PTMs are pre-set) and the parent protein(s). This approach, though successful for sequencing unmodified peptides, is less efficient for peptides with PTMs. The main reasons include complications in spectra due to the modified groups, the greatly

increased number of candidate peptides, and difficulties in predicting fragmentation patterns.[104,105] Some of these issues can be overcome to some extent by narrowing down the database category, using high resolution mass spectrometers or employing efficient MS/MS fragmentation methods (e.g. ETD or ECD for phosphopeptides). In addition, employing more specific tools (e.g. predictors, network/pathway viewers)[45] may improve peptide sequencing and PTM assignments.

1.6 MS-based quantification of proteins and protein phosphorylation

Protein phosphorylation is a reversible, highly dynamic process, and changes in the extent of phosphorylation of specific residues in response to various stimuli may play important roles in cellular signaling pathways.[29,44] Thus, simply exploring the existence of phosphorylation or assigning phosphorylation sites on proteins is not sufficient to relate phosphorylation to cell growth, differentiation, invasion and apoptosis. Traditional analysis of phosphorylated proteins employed incorporation of radioactive ^{32}P (from [^{32}P]ATP or [^{32}P]phosphate) into phosphoproteins either *in vivo* or *in vitro*.[106] Although extremely sensitive, this technique involves the use of high amounts of radioactivity and does not provide identification or quantitation of specific phosphorylation sites. Other biochemistry approaches, including flow kinase activity assays, antibody arrays, cytometry, Western blotting and ELISA facilitate target phosphoprotein quantification.[107,108] However, these methods reveal changes in protein phosphorylation only at specific sites, leaving a large unexplored protein region devoid of phosphorylation quantitation. Over the last few years, advances in MS-based quantitative

proteomics have led to discovery-mode phosphorylation quantitation as well as quantitation of targeted phosphorylation sites.[41,108]

In MS, responses of peptide ions vary greatly and depend on chemical structure, including amino acid sequence and modifications (e.g. phosphorylation and glycosylation),[41] and the co-existence of other species during ionization.[109] Thus, relative quantitation of peptides from two samples frequently requires encoding peptides in one or both samples with stable isotopes (e.g. ^{2}H, ^{13}C, ^{15}N, and ^{18}O). In chemical labeling, incorporation of isotopes occurs via reaction between a mass tag and peptide amino acid residues containing amino groups, carboxylic groups, or cysteines.[110] Reaction of two or several different samples with different isotopes of the mass tag allows comparison of the intensities of the peptide isotopes, which are specific to a given sample. In principle, this method applies to both proteins and peptides, but most applications involve labelling proteolytic digests.[41] One related method, iTRAQ[111] (isobaric tag for relative and absolute quantitation), labels the primary amines of peptides from different sources (up to 8) with different isotopic tags that each consists of a reporter and balance group. The derivatized peptides are indistinguishable in MS (due to the same total mass for the tags containing the same chemical structure but encoded with different isotopes), but exhibit low-mass MS/MS signature ions (unique reporter ion for each peptide source) that enable quantitation.

Using iTRAQ and MS on an LTQ-Orbitrap with high energy collision dissociation, Boja et al. recently assessed relative changes in protein phosphorylation in the mitochondria upon physiological perturbation.[112] For example, relative quantification of phosphopeptides of the pyruvate dehydrogenase-E1α subunit from porcine heart revealed dephosphorylation at three serine sites by pyruvate dehydrogenase phosphatase in response to dichloroacetate or de-

energization, or Ca^{2+}. The advantage of iTRAQ over other chemical labeling methods (especially those comparing precursor ions in MS spectra) lies in its ability in simultaneous comparing up to 8 samples and providing quantitation spanning 2 orders of magnitude. However, iTRAQ also suffers from drawbacks such as the existence of isotopic impurities and other potential interferences in mixed MS/MS.[113] A more serious issue that faces chemical labeling is that the comparison samples must be prepared in precisely the same way before isotope labeling to achieve accurate relative quantitation in the final MS step. The multiple sample preparation steps (e.g. protein purification and digestion) before isotope labeling at the peptide level may result in quantitation errors.[114]

Metabolic labeling provides an alternative approach for relative quantitation by MS methods.[41] In stable isotope labeling by amino acids in culture (SILAC)[114,115], one population of cells grows in media containing normal amino acids, whereas the media for a second population contains amino acids labeled with stable heavy isotopes.[115] The main advantage of SILAC lies in its early-stage isotope labeling, which allows combination of different samples immediately after cell lysis to eliminate the risk of quantitation errors during the subsequent sample preparation steps.[115,116] The reliability of SILAC makes it a powerful tool for assessing protein phosphorylation. For example, in a SILAC study to understand how growth factors influence stem-cell differentiation, Mann and co-workers found that although both epidermal growth factor (EGF) and platelet-derived growth factor induce Tyr phosphorylation, only EGF strongly enhances differentiation to bone-forming cells.[117]

Thus far the majority of antibody-based and MS-based (e.g. using isotope labeling) protein quantitation studies compare relative abundances of proteins or their PTMs in different samples. Absolute quantitation of protein levels and PTM stoichiometry is even more

challenging. Inductively coupled plasma (ICP)-MS can reveal the total amount of P in either the whole protein sample or single proteins after their isolation.[118] However, because it cannot identify protein/peptide sequences, ICP-MS is only relevant to simple samples, e.g. the specific activity of a kinase.[118] For global absolute protein quantitation, Malmström et al. recently developed an approach that combines selected reaction monitoring measurements with a limited number of internal reference standards, the average MS signal intensities of the most intense three peptides per protein, and weighted MS/MS spectra counts.[119] Based on this approach, the authors generated an absolute protein abundance scale for 83% of the MS-detectable proteome from *Leptospira* (a human pathogen).

In addition to absolute protein quantities, the PTM stoichiometry of a given protein also plays a crucial role in cellular signal transduction. In typical relative MS methods, a several fold increase in a phosphopeptide signal could correspond to the same fold phosphorylation degree changes in a protein with <1% of its population phosphorylated or 20% of its population phosphorylated. The former case might be insignificant for inhibiting phosphorylation unless turnover is rapid.[44] To address this problem, Steen et al. determined relative changes in phosphorylation stoichiometry by monitoring the changes in the normalized ion currents of the phosphopeptide(s) of interest.[120]

1.7 Research aims

A full understanding of relationships between phosphorylation and protein function requires comprehensive mapping of phosphorylation sites. Previous studies suggest that either efficient isolation of phosphopeptides from complete tryptic digests (bottom-up approach) or

cleaving the proteins into large peptides (middle-down approach) facilitate MS characterization of phosphoproteins. This research aims to create improved methods for MS characterization of target proteins (especially phosphoproteins) using both of these approaches. More specifically, the dissertation will cover:

1. Preconcentration and purification of tryptic phosphopeptides using micro-arrays of functional polymer brushes on MALDI plates (Chapter 2).
2. MALDI-MS and CID-MS/MS identification of p65-associated phosphoproteins after phosphopeptide enrichment using polymer-oxotitanium films (Chapter 3).
3. Low-bias enrichment of mono- and multiply phosphorylated peptides with bifunctional polymer brushes and tandem affinity membranes (Chapter 4).
4. Size-controlled proteolysis using trypsin-modified nylon membranes to provide large peptides for middle-down protein identification and characterization (Chapter 5).

Chapter 6 summarizes the research and presents future directions.

REFERENCES

REFERENCES

1. Walsh, C. T., *Posttranslational Modification of Proteins: Expending Nature's Inventory*. Roberts and Company Publishers: Englewood, Colorado, USA, 2006.

2. Johnson, L. N.; Lewis, R. J., Structural Basis for Control by Phosphorylation. *Chem. Rev.* **2001,** 101, 2209-2242.

3. Lee, M.; Goodbourn, S., Signalling from the Cell Surface to the Nucleus. In *Regulation of Gene Expression*, Portland Press Ltd: London, 2001; Vol. 37, pp 71-85.

4. Cohen, P., The Regulation of Protein Function by Multisite Phosphorylation - a 25 Year Update. *Trends Biochem.Sci.* **2000,** 25, 596-601.

5. Pinna, L. A.; Ruzzene, M., How Do Protein Kinases Recognize Their Substrates? *Biochim. Biophys. Acta-Mol. Cell Res.* **1996,** 1314, 191-225.

6. Echols, N.; Harrison, P.; Balasubramanian, S.; Luscombe, N. M.; Bertone, P.; Zhang, Z. L.; Gerstein, M., Comprehensive Analysis of Amino Acid and Nucleotide Composition in Eukaryotic Genomes, Comparing Genes and Pseudogenes. *Nucleic Acids Res.* **2002,** 30, 2515-2523.

7. Hunter, T.; Sefton, B. M., Transforming Gene-Product of Rous-Sarcoma Virus Phosphorylates Tyrosine. *Proc. Natl. Acad. Sci. U. S. A.-Biol. Sci.* **1980,** 77, 1311-1315.

8. Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M., Global, in Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **2006,** 127, 635-648.

9. Ubersax, J. A.; Ferrell, J. E., Mechanisms of Specificity in Protein Phosphorylation. *Nat. Rev. Mol. Cell Biol.* **2007,** 8, 530-541.

10. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S., The Protein Kinase Complement of the Human Genome. *Science* **2002,** 298, 1912-1934.

11. Brown, N. R.; Noble, M. E. M.; Endicott, J. A.; Johnson, L. N., The Structural Basis for Specificity of Substrate and Recruitment Peptides for Cyclin-Dependent Kinases. *Nat. Cell Biol.* **1999,** 1, 438-443.

12. Hubbard, S. R., Crystal Structure of the Activated Insulin Receptor Tyrosine Kinase in Complex with Peptide Substrate and Atp Analog. *Embo J.* **1997,** 16, 5572-5581.

13. Barford, D.; Hu, S. H.; Johnson, L. N., Structural Mechanism for Glycogen-Phosphorylase Control by Phosphorylation and Amp. *J. Mol. Biol.* **1991,** 218, 233-260.

14. Kuriyan, J.; Cowburn, D., Modular Peptide Recognition Domains in Eukaryotic Signaling. *Annu. Rev. Biophys. Biomolec. Struct*. **1997,** 26, 259-288.

15. Pawson, T.; Nash, P., Assembly of Cell Regulatory Systems through Protein Interaction Domains. *Science* **2003,** 300, 445-452.

16. Yaffe, M. B.; Elia, A. E. H., Phosphoserine/Threonine-Binding Domains. *Curr. Opin. Cell Biol.* **2001,** 13, 131-138.

17. Huang, C. Y. F.; Ferrell, J. E., Ultrasensitivity in the Mitogen-Activated Protein Kinase Cascade. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** 93, 10078-10083.

18. Kennelly, P. J., Protein Phosphatases - a Phylogenetic Perspective. *Chem. Rev*. **2001,** 101, 2291-2312.

19. Bauman, A. L.; Scott, J. D., Kinase- and Phosphatase-Anchoring Proteins: Harnessing the Dynamic Duo. *Nat. Cell Biol.* **2002,** 4, E203-E206.

20. Han, X. M.; Aslanian, A.; Yates, J. R., Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008,** 12, 483-490.

21. Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A., Analysis of Protein Phosphorylation Using Mass Spectrometry: Deciphering the Phosphoproteome. *Trends Biotechnol.* **2002,** 20, 261-268.

22. Fournier, M. L.; Gilmore, J. M.; Martin-Brown, S. A.; Washburn, M. P., Multidimensional Separations-Based Shotgun Proteomics. *Chem. Rev*. **2007,** 107, 3654-3686.

23. Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C., Identifying Proteins from 2-Dimensional Gels by Molecular Mass Searching of Peptide-Fragments in Protein-Sequence Databases. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** 90, 5011-5015.

24. Loo, R. R. O.; Hayes, R.; Yang, Y. N.; Hung, F.; Ramachandran, P.; Kim, N.; Gunsalus, R.; Loo, J. A., Top-Down, Bottom-up, and Side-to-Side Proteomics with Virtual 2-D Gels. *Int. J. Mass Spectrom*. **2005,** 240, 317-325.

25. Wang, W. H.; Palumbo, A. M.; Tan, Y. L.; Reid, G. E.; Tepe, J. J.; Bruening, M. L., Identification of p65-Associated Phosphoproteins by Mass Spectrometry after on-Plate Phosphopeptide Enrichment Using Polymer-Oxotitanium Films. *J. Proteome Res.* **2010,** 9, 3005-3015.

26. McDonald, W. H.; Yates, J. R., Shotgun Proteomics: Integrating Technologies to Answer Biological Questions. *Curr. Opin. Mol. Ther.* **2003,** 5, 302-309.

27. Bodenmiller, B.; Mueller, L. N.; Mueller, M.; Domon, B.; Aebersold, R., Reproducible Isolation of Distinct, Overlapping Segments of the Phosphoproteome. *Nat. Methods* **2007,** 4, 231-237.

28. Delahunty, C. M.; Yates, J. R., Mudpit: Multidimensional Protein Identification Technology. *Biotechniques* **2007,** 43, 563.

29. Thingholm, T. E.; Jensen, O. N.; Larsen, M. R., Analytical Strategies for Phosphoproteomics. *Proteomics* **2009,** 9, 1451-1468.

30. Hoang, T.; Roth, U.; Kowalewski, K.; Belisle, C.; Steinert, K.; Karas, M., Highly Specific Capture and Direct MALDI MS Analysis of Phosphopeptides by Zirconium Phosphonate on Self-Assembled Monolayers. *Anal. Chem*. **2010,** 82, 219-228.

31. Ge, Y.; Rybakova, I. N.; Xu, Q. G.; Moss, R. L., Top-Down High-Resolution Mass Spectrometry of Cardiac Myosin Binding Protein C Revealed That Truncation Alters Protein Phosphorylation State. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** 106, 12658-12663.

32. Zhao, Y. M.; Jensen, O. N., Modification-Specific Proteomics: Strategies for Characterization of Post-Translational Modifications Using Enrichment Techniques. *Proteomics* **2009,** 9, 4632-4641.

33. Villen, J.; Gygi, S. P., The Scx/IMAC Enrichment Approach for Global Phosphorylation Analysis by Mass Spectrometry. *Nat. Protoc.* **2008,** 3, 1630-1638.

34. McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X. M.; Infusini, G.; Jiang, H.; Kong, X. L.; Begley, T. P., Top-Down MS, a Powerful Complement to the High Capabilities of Proteolysis Proteomics. *Febs J.* **2007,** 274, 6256-6268.

35. Garcia, B. A.; Pesavento, J. J.; Mizzen, C. A.; Kelleher, N. L., Pervasive Combinatorial Modification of Histone H3 in Human Cells. *Nat. Methods* **2007,** 4, 487-489.

36. Garcia, B. A.; Siuti, N.; Thomas, C. E.; Mizzen, C. A.; Kelleher, N. L., Characterization of Neurohistone Variants and Post-Translational Modifications by Electron Capture Dissociation Mass Spectrometry. *Int. J. Mass Spectrom*. **2007,** 259, 184-196.

37. Roth, M. J.; Forbes, A. J.; Boyne, M. T.; Kim, Y. B.; Robinson, D. E.; Kelleher, N. L., Precise and Parallel Characterization of Coding Polymorphisms, Alternative Splicing, and Modifications in Human Proteins by Mass Spectrometry. *Mol. Cell. Proteomics* **2005,** 4, 1002-1008.

38. Whitelegge, J.; Halgand, F.; Souda, P.; Zabrouskov, V., Top-Down Mass Spectrometry of Integral Membrane Proteins. *Expert Rev. Proteomics* **2006,** 3, 585-596.

39. Siuti, N.; Kelleher, N. L., Decoding Protein Modifications Using Top-Down Mass Spectrometry. *Nat. Methods* **2007,** 4, 817-821.

40. Young, N. L.; Plazas-Mayorca, M. D.; Garcia, B. A., Systems-Wide Proteomic Characterization of Combinatorial Post-Translational Modification Patterns. *Expert Rev. Proteomics* **2010,** 7, 79-92.

41. Macek, B.; Mann, M.; Olsen, J. V., Global and Site-Specific Quantitative Phosphoproteomics: Principles and Applications. *Annu. Rev. Pharmacol. Toxicol.* **2009,** 49, 199-221.

42. Dunn, J. D.; Reid, G. E.; Bruening, M. L., Techniques for Phosphopeptide Enrichment Prior to Analysis by Mass Spectrometry. *Mass Spectrom. Rev.* **2010,** 29, 29-54.

43. Paradela, A.; Albar, J. P., Advances in the Analysis of Protein Phosphorylation. *J. Proteome Res.* **2008,** 7, 1809-1818.

44. Ashman, K.; Villar, E. L., Phosphoproteomics and Cancer Research. *Clin. Transl. Oncol.* **2009,** 11, 356-362.

45. Rogers, L. D.; Foster, L. J., Phosphoproteomics-Finally Fulfilling the Promise? *Mol. Biosyst.* **2009,** 5, 1122-1129.

46. Andersson, L.; Porath, J., Isolation of Phosphoproteins by Immobilized Metal (Fe-3+) Affinity-Chromatography. *Anal. Biochem.* **1986,** 154, 250-254.

47. Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R., Simac (Sequential Elution from IMAC), a Phosphoproteomics Strategy for the Rapid Separation of Monophosphorylated from Multiply Phosphorylated Peptides. *Mol. Cell. Proteomics* **2008,** 7, 661-671.

48. Trinidad, J. C.; Specht, C. G.; Thalhammer, A.; Schoepfer, R.; Burlingame, A. L., Comprehensive Identification of Phosphorylation Sites in Postsynaptic Density Preparations. *Mol. Cell. Proteomics* **2006,** 5, 914-922.

49. Villen, J.; Beausoleil, S. A.; Gerber, S. A.; Gygi, S. P., Large-Scale Phosphorylation Analysis of Mouse Liver. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** 104, 1488-1493.

50. Zhai, B.; Villen, J.; Beausoleil, S. A.; Mintseris, J.; Gygi, S. P., Phosphoproteome Analysis of Drosophila Metanogaster Embryos. *J. Proteome Res.* **2008,** 7, 1675-1682.

51. Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M., Phosphoproteome Analysis by Mass Spectrometry and Its Application to Saccharomyces Cerevisiae. *Nat. Biotechnol.* **2002,** 20, 301-305.

52. Kokubu, M.; Ishihama, Y.; Sato, T.; Nagasu, T.; Oda, Y., Specificity of Immobilized Metal Affinity-Based IMAC/C18 Tip Enrichment of Phosphopeptides for Protein Phosphorylation Analysis. *Anal. Chem.* **2005,** 77, 5144-5154.

53. Tsai, C. F.; Wang, Y. T.; Chen, Y. R.; Lai, C. Y.; Lin, P. Y.; Pan, K. T.; Chen, J. Y.; Khoo, K. H.; Chen, Y. J., Immobilized Metal Affinity Chromatography Revisited: Ph/Acid Control toward High Selectivity in Phosphoproteomics. *J. Proteome Res.* **2008,** 7, 4058-4069.

54. Ye, J. Y.; Zhang, X. M.; Young, C.; Zhao, X. L.; Hao, Q.; Cheng, L.; Jensen, O. N., Optimized IMAC-IMAC Protocol for Phosphopeptide Recovery from Complex Biological Samples. *J. Proteome Res.* **2010,** 9, 3561-3573.

55. Pinkse, M. W. H.; Mohammed, S.; Gouw, L. W.; van Breukelen, B.; Vos, H. R.; Heck, A. J. R., Highly Robust, Automated, and Sensitive on Line Tio2-Based Phosphoproteomics Applied to Study Endogenous Phosphorylation in Drosophila Melanogaster. *J. Proteome Res.* **2008,** 7, 687-697.

56. Li, Q. R.; Ning, Z. B.; Tang, J. S.; Nie, S.; Zeng, R., Effect of Peptide-to-Tio2 Beads Ratio on Phosphopeptide Enrichment Selectivity. *J. Proteome Res.* **2009,** 8, 5375-5381.

57. Simon, E. S.; Young, M.; Chan, A.; Bao, Z. Q.; Andrews, P. C., Improved Enrichment Strategies for Phosphorylated Peptides on Titanium Dioxide Using Methyl Esterification and Ph Gradient Elution. *Anal. Biochem.* **2008,** 377, 234-242.

58. Zhou, H. J.; Ye, M. L.; Dong, J.; Han, G. H.; Jiang, X. N.; Wu, R. N.; Zou, H. F., Specific Phosphopeptide Enrichment with Immobilized Titanium Ion Affinity Chromatography Adsorbent for Phosphoproteome Analysis. *J. Proteome Res.* **2008,** 7, 3957-3967.

59. Wan, J. J.; Qian, K.; Qiao, L.; Wang, Y. H.; Kong, J. L.; Yang, P. Y.; Liu, B. H.; Yu, C. Z., Tio2-Modified Macroporous Silica Foams for Advanced Enrichment of Multi-Phosphorylated Peptides. *Chem.-Eur. J.* **2009,** 15, 2504-2508.

60. Lin, B.; Li, T.; Zhao, Y.; Huang, F. K.; Guo, L.; Feng, Y. Q., Preparation of a Tio2 Nanoparticle-Deposited Capillary Column by Liquid Phase Deposition and Its Application in Phosphopeptide Analysis. *J. Chromatogr. A* **2008,** 1192, 95-102.

61. Lu, S.; Luo, Q.; Li, X. C.; Wu, J. H.; Liu, J. A.; Xiong, S. X.; Feng, Y. Q.; Wang, F. Y., Inhibitor Screening of Protein Kinases Using MALDI-Tof MS Combined with Separation and Enrichment of Phosphopeptides by Tio2 Nanoparticle Deposited Capillary Column. *Analyst* **2010,** 135, 2858-2863.

62. Qin, W. J.; Zhang, W. J.; Song, L. N.; Zhang, Y. J.; Qian, X. H., Surface Initiated Atom Transfer Radical Polymerization: Access to Three Dimensional Wavelike Polymer Structure Modified Capillary Columns for Online Phosphopeptide Enrichment. *Anal. Chem.* **2010,** 82, 9461-9468.

63. Xue, Y. F.; Wei, J. Y.; Han, H. H.; Zhao, L. Y.; Cao, D.; Wang, J. L.; Yang, X. M.; Zhang, Y. J.; Qian, X. H., Application of Open Tubular Capillary Columns Coated with

Zirconium Phosphonate for Enrichment of Phosphopeptides. *J. Chromatogr. B* **2009,** 877, 757-764.

64. Hou, C. Y.; Ma, J. F.; Tao, D. Y.; Shan, Y. C.; Liang, Z.; Zhang, L. H.; Zhang, Y. K., Organic-Inorganic Hybrid Silica Monolith Based Immobilized Titanium Ion Affinity Chromatography Column for Analysis of Mitochondrial Phosphoproteome. *J. Proteome Res.* **2010,** 9, 4093-4101.

65. Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J. X.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P., Large-Scale Characterization of Hela Cell Nuclear Phosphoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, 12130-12135.

66. Nie, S.; Dai, J.; Ning, Z. B.; Cao, X. J.; Sheng, Q. H.; Zeng, R., Comprehensive Profiling of Phosphopeptides Based on Anion Exchange Followed by Flow-through Enrichment with Titanium Dioxide (Afet). *J. Proteome Res.* **2010,** 9, 4585-4594.

67. Han, G. H.; Ye, M. L.; Zhou, H. J.; Jiang, X. N.; Feng, S.; Jiang, X. G.; Tian, R. J.; Wan, D. F.; Zou, H. F.; Gu, J. R., Large-Scale Phosphoproteome Analysis of Human Liver Tissue by Enrichment and Fractionation of Phosphopeptides with Strong Anion Exchange Chromatography. *Proteomics* **2008,** 8, 1346-1361.

68. Motoyama, A.; Xu, T.; Ruse, C. I.; Wohlschlegel, J. A.; Yates, J. R., Anion and Cation Mixed-Bed Ion Exchange for Enhanced Multidimensional Separations of Peptides and Phosphopeptides. *Anal. Chem.* **2007,** 79, 3623-3634.

69. Dai, J.; Jin, W. H.; Sheng, Q. H.; Shieh, C. H.; Wu, J. R.; Zeng, R., Protein Phosphorylation and Expression Profiling by Yin-Yang Multidimensional Liquid Chromatography (Yin-Yang Mdlc) Mass Spectrometry. *J. Proteome Res.* **2007,** 6, 250-262.

70. Jain, P.; Baker, G. L.; Bruening, M. L., Applications of Polymer Brushes in Protein Analysis and Purification. *Annu. Rev. Anal. Chem.* **2009,** 2, 387-408.

71. Blacken, G. R.; Volny, M.; Diener, M.; Jackson, K. E.; Ranjitkar, P.; Maly, D. J.; Turecek, F., Reactive Landing of Gas-Phase Ions as a Tool for the Fabrication of Metal Oxide Surfaces for in Situ Phosphopeptide Enrichment. *J. Am. Soc. Mass Spectrom.* **2009,** 20, 915-926.

72. Niklew, M. L.; Hochkirch, U.; Melikyan, A.; Moritz, T.; Kurzawski, S.; Schluter, H.; Ebner, I.; Linscheid, M. W., Phosphopeptide Screening Using Nanocrystalline Titanium Dioxide Films as Affinity Matrix-Assisted Laser Desorption Ionization Targets in Mass Spectrometry. *Anal. Chem.* **2010,** 82, 1047-1053.

73. Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C., High-Throughput Middle-Down Analysis Using an Orbitrap. *J. Proteome Res.* **2010,** 9, 3886-3890.

74. Wu, S. L.; Kim, J.; Hancock, W. S.; Karger, B., Extended Range Proteomic Analysis (ERPA): A New and Sensitive LC-MS Platform for High Sequence Coverage of Complex

Proteins with Extensive Post-Translational Modifications-Comprehensive Analysis of Beta-Casein and Epidermal Growth Factor Receptor (Egfr). *J. Proteome Res.* **2005,** 4, 1155-1170.

75. Coon, J. J.; Ueberheide, B.; Syka, J. E. P.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F., Protein Identification Using Sequential Ion/Ion Reactions and Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** 102, 9463-9468.

76. Garcia, B. A.; Thomas, C. E.; Kelleher, N. L.; Mizzen, C. A., Tissue-Specific Expression and Post-Translational Modification of Histone H3 Variants. *J. Proteome Res.* **2008,** 7, 4225-4236.

77. Xu, P.; Peng, J. M., Characterization of Polyubiquitin Chain Structure by Middle-Down Mass Spectrometry. *Anal. Chem.* **2008,** 80, 3438-3444.

78. She, Y. M.; Xu, X. A. H.; Yakunin, A. F.; Dhe-Paganon, S.; Donald, L. J.; Standing, K. G.; Lee, D. C.; Jia, Z. C.; Cyr, T. D., Mass Spectrometry Following Mild Enzymatic Digestion Reveals Phosphorylation of Recombinant Proteins in Escherichia Coil through Mechanisms Involving Direct Nucleotide Binding. *J. Proteome Res.* **2010,** 9, 3311-3318.

79. Wu, S. L.; Huehmer, A. F. R.; Hao, Z. Q.; Karger, B. L., On-Line LC-MS Approach Combining Collision-Induced Dissociation (CID), Electron-Transfer Dissociation (Etd), and CID of an Isolated Charge-Reduced Species for the Trace-Level Characterization of Proteins with Post-Translational Modifications. *J. Proteome Res.* **2007,** 6, 4230-4244.

80. Hamuro, Y.; Coales, S. J.; Molnar, K. S.; Tuske, S. J.; Morrow, J. A., Specificity of Immobilized Porcine Pepsin in H/D Exchange Compatible Conditions. *Rapid Commun. Mass Spectrom.* **2008,** 22, 1041-1046.

81. Appel, W., Chymotrypsin - Molecular and Catalytic Properties. *Clin. Biochem.* **1986,** 19, 317-322.

82. Lee, T. D.; Shively, J. E., Enzymatic and Chemical Digestion of Proteins for Mass-Spectrometry. *Method Enzymol.* **1990,** 193, 361-374.

83. Birktoft, J. J.; Breddam, K., Glutamyl Endopeptidases. In *Proteolytic Enzymes: Serine and Cysteine Peptidases*, Academic Press Inc: San Diego, 1994; Vol. 244, pp 114-126.

84. Raijmakers, R.; Neerincx, P.; Mohammed, S.; Heck, A. J. R., Cleavage Specificities of the Brother and Sister Proteases Lys-C and Lys-N. *Chem. Commun.* **2010,** 46, 8827-8829.

85. Drapeau, G. R., Substrate-Specificity of a Proteolytic-Enzyme Isolated from a Mutant of Pseudomonas-Fragi. *J. Biol. Chem.* **1980,** 255, 839-840.

86. Li, A. Q.; Sowder, R. C.; Henderson, L. E.; Moore, S. P.; Garfinkel, D. J.; Fisher, R. J., Chemical Cleavage at Aspartyl Residues for Protein Identification. *Anal. Chem.* **2001,** 73, 5395-5402.

87. Partridge, S. M.; Davis, H. F., Preferential Release of Aspartic Acid During the Hydrolysis of Proteins. *Nature* **1950,** 165, 62-63.

88. Inglis, A. S., Cleavage at Aspartic-Acid. *Method Enzymol.* **1983,** 91, 324-332.

89. Swatkoski, S.; Russell, S.; Edwards, N.; Fenselau, C., Analysis of a Model Virus Using Residue-Specific Chemical Cleavage and MALDI-Tof Mass Spectrometry. *Anal. Chem.* **2007,** 79, 654-658.

90. Swatkoski, S.; Russell, S. C.; Edwards, N.; Fenselau, C., Rapid Chemical Digestion of Small Acid-Soluble Spore Proteins for Analysis of Bacillus Spores. *Anal. Chem.* **2006,** 78, 181-188.

91. Swatkoski, S.; Gutierrez, P.; Wynne, C.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C., Evaluation of Microwave-Accelerated Residue-Specific Acid Cleavage for Proteomic Applications. *J. Proteome Res.* **2008,** 7, 579-586.

92. Hauser, N. J.; Basile, F., Online Microwave D-Cleavage Lc-ESI-MS/MS of Intact Proteins: Site-Specific Cleavages at Aspartic Acid Residues and Disulfide Bonds. *J. Proteome Res.* **2008,** 7, 1012-1026.

93. Li, J. X.; Shefcheck, K.; Callahan, J.; Fenselau, C., Extension of Microwave-Accelerated Residue-Specific Acid Cleavage to Proteins with Carbohydrate Side Chains and Disulfide Linkages. *Int. J. Mass Spectrom.* **2008,** 278, 109-113.

94. Huang, H. V.; Bond, M. W.; Hunkapiller, M. W.; Hood, L. E., Cleavage at Tryptophanyl Residues with Dimethylsulfoxide Hydrochloric-Acid and Cyanogen-Bromide. *Method Enzymol.* **1983,** 91, 318-324.

95. Vestling, M. M.; Kelly, M. A.; Fenselau, C., Optimization by Mass-Spectrometry of a Tryptophan-Specific Protein Cleavage Reaction. *Rapid Commun. Mass Spectrom.* **1994,** 8, 786-790.

96. Walcher, W.; Timperio, A. M.; Zolla, L.; Huber, C. G., Characterization of a Variant of the Spinach Psii Type I Light-Harvesting Protein Using Kinetically Controlled Digestion and Rp-Hplc-ESI-MS. *Anal. Chem.* **2003,** 75, 6775-6780.

97. Bi, H. Y.; Qiao, L.; Busnel, J. M.; Liu, B. H.; Girault, H. H., Kinetics of Proteolytic Reactions in Nanoporous Materials. *J. Proteome Res.* **2009,** 8, 4685-4692.

98. Srividhya, J.; Schnell, S., Why Substrate Depletion Has Apparent First-Order Kinetics in Enzymatic Digestion. *Comput. Biol. Chem.* **2006,** 30, 209-214.

99. Krenkova, J.; Svec, F., Less Common Applications of Monoliths: Iv. Recent Developments in Immobilized Enzyme Reactors for Proteomics and Biotechnology. *J. Sep. Sci.* **2009,** 32, 706-718.

100. Schroeder, M. J.; Shabanowitz, J.; Schwartz, J. C.; Hunt, D. F.; Coon, J. J., A Neutral Loss Activation Method for Improved Phosphopeptide Sequence Analysis by Quadrupole Ion Trap Mass Spectrometry. *Anal. Chem.* **2004,** 76, 3590-3598.

101. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W., Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998,** 120, 3265-3266.

102. Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, 9528-9533.

103. McAlister, G. C.; Phanstiel, D.; Good, D. M.; Berggren, W. T.; Coon, J. J., Implementation of Electron-Transfer Dissociation on a Hybrid Linear Ion Trap-Orbitrap Mass Spectrometer. *Anal. Chem.* **2007,** 79, 3525-3534.

104. Nesvizhskii, A. I.; Vitek, O.; Aebersold, R., Analysis and Validation of Proteomic Data Generated by Tandem Mass Spectrometry. *Nat. Methods* **2007,** 4, 787-797.

105. Ahrne, E.; Muller, M.; Lisacek, F., Unrestricted Identification of Modified Proteins Using MS/MS. *Proteomics* **2010,** 10, 671-686.

106. Yan, J. X.; Packer, N. H.; Gooley, A. A.; Williams, K. L., Protein Phosphorylation: Technologies for the Identification of Phosphoamino Acids. *J. Chromatogr. A* **1998,** 808, 23-41.

107. Pelech, S.; Jelinkova, L.; Susor, A.; Zhang, H.; Shi, X. Q.; Pavlok, A.; Kubelka, M.; Kovarova, H., Antibody Microarray Analyses of Signal Transduction Protein Expression and Phosphorylation During Porcine Oocyte Maturation. *J. Proteome Res.* **2008,** 7, 2860-2871.

108. White, F. M., Quantitative Phosphoproteomic Analysis of Signaling Network Dynamics. *Curr. Opin. Biotechnol.* **2008,** 19, 404-409.

109. Annesley, T. M., Ion Suppression in Mass Spectrometry. *Clin. Chem.* **2003,** 49, 1041-1044.

110. Julka, S.; Regnier, F., Quantification in Proteomics through Stable Isotope Coding: A Review. *J. Proteome Res.* **2004,** 3, 350-363.

111. Ross, P. L.; Huang, Y. L. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J., Multiplexed Protein Quantitation in

Saccharomyces Cerevisiae Using Amine-Reactive Isobaric Tagging Reagents. *Mol. Cell. Proteomics* **2004,** 3, 1154-1169.

112. Boja, E. S.; Phillips, D.; French, S. A.; Harris, R. A.; Balaban, R. S., Quantitative Mitochondrial Phosphoproteomics Using Itraq on an Ltq-Orbitrap with High Energy Collision Dissociation. *J. Proteome Res.* **2009,** 8, 4665-4675.

113. Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Rehman, I.; Wright, P. C., Itraq Underestimation in Simple and Complex Mixtures: "The Good, the Bad and the Ugly". *J. Proteome Res.* **2009,** 8, 5347-5355.

114. Mann, M., Functional and Quantitative Proteomics Using Silac. *Nat. Rev. Mol. Cell Biol.* **2006,** 7, 952-958.

115. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable Isotope Labeling by Amino Acids in Cell Culture, Silac, as a Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* **2002,** 1, 376-386.

116. Schreiber, T. B.; Mausbacher, N.; Breitkopf, S. B.; Grundner-Culemann, K.; Daub, H., Quantitative Phosphoproteomics - an Emerging Key Technology in Signal-Transduction Research. *Proteomics* **2008,** 8, 4416-4432.

117. Kratchmarova, I.; Blagoev, B.; Haack-Sorensen, M.; Kassem, M.; Mann, M., Mechanism of Divergent Growth Factor Effects in Mesenchymal Stem Cell Differentiation. *Science* **2005,** 308, 1472-1477.

118. Navaza, A. P.; Encinar, J. R.; Sanz-Medel, A., Quantitative Protein Phosphorylation Analysis: The Role of Icp-MS. *J. Anal. At. Spectrom.* **2007,** 22, 1223-1237.

119. Malmstrom, J.; Beck, M.; Schmidt, A.; Lange, V.; Deutsch, E. W.; Aebersold, R., Proteome-Wide Cellular Protein Concentrations of the Human Pathogen Leptospira Interrogans. *Nature* **2009,** 460, 762-U112.

120. Steen, H.; Jebanathirajah, J. A.; Springer, M.; Kirschner, M. W., Stable Isotope-Free Relative and Absolute Quantitation of Protein Phosphorylation Stoichiometry by MS. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** 102, 3948-3953.

Chapter 2. Phosphopeptide Enrichment on Functionalized Polymer Microspots for MALDI-MS Analysis

(The results described in Chapter 2 have been published in: W.H. Wang, M.L. Bruening, Phosphopeptide enrichment on functionalized polymer microspots for MALDI-MS analysis. *Analyst*, 2009, 134, 512–518.)

2.1 Introduction

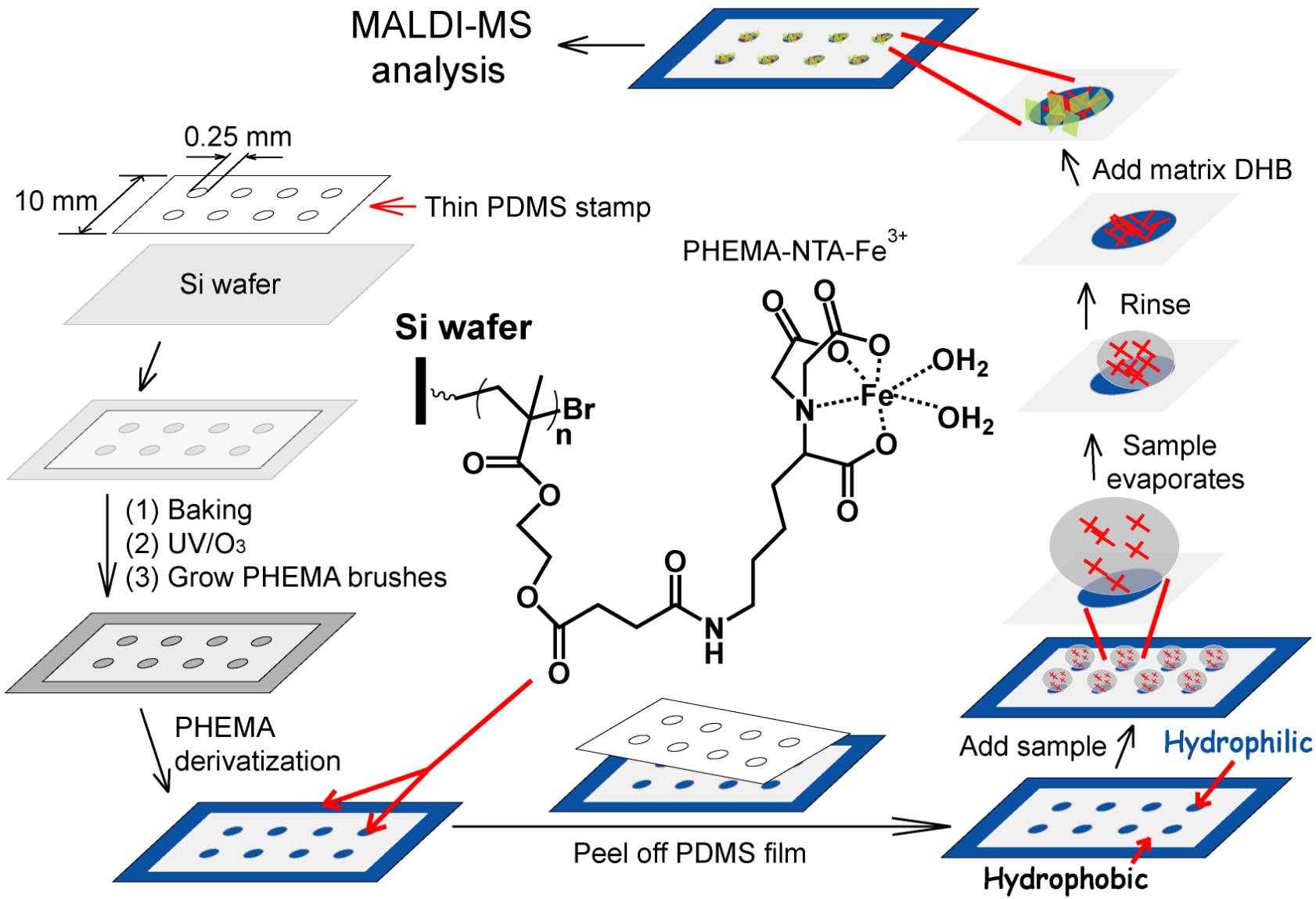
Mass spectrometry (MS) is the preferred technique for detecting phosphorylation sites on hydroxylated amino acid residues (Ser, Thr, and Tyr).[1,2] However, the stoichiometry of phosphorylation is relatively low within cells,[3] so phosphopeptide MS signals are often weak because of ion suppression in the presence of salts and nonphosphopeptides.[4] Selective enrichment of phosphopeptides in phosphoprotein digests provides one way to enhance phosphopeptide MS signals, and enrichment strategies include covalent binding of phosphopeptides to surfaces[5-7] and selective adsorption of phosphopeptides on immobilized metal ion complexes[8-10] or metal oxides.[11,12] The latter two approaches were applied directly on a matrix-assisted laser desorption/ionization (MALDI) plate prior to MALDI-MS analysis.[8,10,12] These on-plate enrichment strategies integrate binding of phosphopeptides, rinsing to remove impurities, and matrix/eluent addition on a single plate to simplify sample handling, minimize sample loss, and facilitate high-throughput analysis.

Recently, our group developed a strategy for polymer brush-based on-plate enrichment of phosphopeptides.[13] We grow poly(2-hydroxyethyl methacrylate) (PHEMA) brushes from a sample probe by atom transfer radical polymerization (ATRP) and then derivatize these brushes

with a high density of nitrilotriacetate-Fe(III) (NTA-Fe(III)) complexes that selectively bind phosphopeptides. The high binding capacity of the brushes decreases the detection limits for on-probe purification with MALDI-MS analysis compared to enrichment on monolayer and thin polymer films.[13]

This chapter describes a method for forming PHEMA-NTA-Fe(III) microspots to selectively concentrate phosphopeptides into small areas, thereby further increasing the sensitivity of phosphopeptide detection. The strategy for sample concentration (right side of Scheme 1) includes pinning a sample droplet on a 0.25-mm spot of polymer brush, allowing the droplet to evaporate to the spot size, rinsing to remove unwanted compounds, and addition of matrix to elute phosphopeptides and allow MALDI-MS analysis. This method restricts the sample size to ~0.05 mm^2, which approaches the cross-sectional area of the MALDI laser. Thus a much larger fraction of the sample is available for ionization upon irradiation than when using conventional MALDI wells that have diameters of 2-4 mm (areas of 3-13 mm^2). Similar strategies were employed with MALDI plates modified with hydrophobic surfaces such as paraffin wax[14,15] and Teflon.[16,17] However, those methods either concentrate all the species in a small sample droplet on the plate, or simply remove salts from the sample without any selectivity for enriching a particular class of biomacromolecules.[14-19] Phosphopeptide enrichment requires small spots that realize both pre-concentration and selective binding.

In previous work on enrichment of peptides and DNA,[20,21] our group used a micro-contact printing/film growth technique[22] to create small (~0.2 mm), hydrophilic spots of poly(acrylic acid) surrounded by a pre-patterned hydrophobic monolayer. When we attempted to use this method for synthesizing small spots of PHEMA brushes, however, the pre-patterned monolayer partially lost its hydrophobicity due to the adsorption of PHEMA or monomer, and



Scheme 2.1. Fabrication of micro-arrays of PHEMA-NTA-Fe(III) brushes surrounded by hydrophobic poly(dimethylsiloxane) (PDMS), and on-plate enrichment of phosphopeptides prior to MALDI-MS analysis. DHB stands for 2,5-dihydroxybenzoic acid.

this resulted in inefficient sample pre-concentration during enrichment.

This chapter presents a simple technique for transferring a patterned, hydrophobic poly(dimethylsiloxane) (PDMS) layer onto a Si wafer by heating of a PDMS stamp that is in contact with the Si (Scheme 1). The patterned stamp contains micropores (~250 μm diameter) that expose Si for growth of thick polymer brushes. After brush formation, removal of the bulk PDMS stamp leaves behind a pristine PDMS layer that is hydrophobic and capable of repelling sample droplets during enrichment. Isolation of phosphopeptides in the small area of the hydrophilic, polymer brush microspots greatly enhances MALDI-MS signals of low levels of phosphopeptides. Detection limits are typically 5-fold lower than those obtained with enrichment on a 2 mm-diameter spot of polymer brushes on a non-patterned substrate.

2.2 Experimental

2.2.1 Chemicals and materials

Most chemicals and materials employed in this work were obtained from Sigma-Aldrich. Other chemicals include succinic anhydride (Matheson Coleman & Bell), N_{α},N_{α}-bis(carboxymethyl)-L-lysine hydrate (aminobutyl NTA, Fluka), anhydrous ferric chloride ($FeCl_3$, Spectrum), Tris-HCl (Invitrogen), urea (J.T.Baker), DL-dithiothreitol (Fluka), acetonitrile (ACN, Mallinckrodt Baker), ammonium bicarbonate (Columbus Chemical) and sequencing grade modified porcine trypsin (treated with L-1-tosylamide-2-phenylethyl chloromethyl ketone, Promega). *N*,*N*-Dimethylformamide (DMF, Jade Scientific) was newly distilled before use, and deionized water was prepared with a Millipore ion-exchange system (Milli-Q, 18 MΩcm).

2.2.2 Fabrication of functional polymer brushes on Si wafers

The left of Scheme 1 shows the steps in fabricating microspots of polymer brushes on a Si wafer. To prepare the thin PDMS stamp, ~0.5 g of PDMS elastomer solution (Sylgard 184, Dow Corning) with a 10:1 polymer/curing agent ratio was poured onto a flat 8.5 cm-diameter sterile polystyrene petri dish (Fisherbrand) and dispersed such that it covered the dish uniformly. This solution was then cured in air at 50 ℃ for 2 h to form a thin (<0.1 mm) PDMS stamp, which was cut into 10×15 mm pieces. Individual pieces were peeled off the petri dish and placed on 5-mm thick, pre-cured PDMS prior to punching 250 μm-diameter holes through the thin PDMS stamp and the thick PDMS support with a blunt dispensing needle (i.d. 0.23-0.25 mm, Part No. ZT-5-026-5-X, ZEPHYRTRONICS). The thin stamp was then peeled off the thick PDMS and gently attached to a 15×20 mm piece of Si wafer that was previously sonicated in acetone and dried with N_2. After curing the PDMS on the Si wafer in air at 75 ℃ for 1 h and subsequent UV/ozone treatment for 15 min, initiator was attached to exposed Si by immersing the stamp-covered wafer overnight in an anhydrous DMF solution containing 7.4 mg/mL (16 mM) 11-(2-bromo-2-methyl)propionyloxyundecenyltrichlorosilane, which was synthesized according to reference literature procedure.[23] Synthesis of PHEMA brushes on the exposed, initiator-modified Si spots through CuCl/$CuBr_2$/2,2'-bipyridine catalyzed ATRP (1 h), and derivatization of the polymer with NTA-Fe(III) were performed as previously reported by our group.[13] Finally, the thin PDMS stamp was gently peeled off the Si wafer to leave a thin, residual PDMS coating surrounding the microspots.

For enrichment on a 2-mm polymer spot, the entire surface of a Si wafer was coated with polymer brushes in the same way as above, but no PDMS stamp was attached to the Si. To

restrict sample droplets to a diameter of ~2 mm, sample wells were made by lightly scratching 2 mm-diameter circles onto the wafer using a tungsten carbide-tipped pen and a teflon mask with arrays of 2 mm-diameter holes.

2.2.3 Polymer brush characterization

The thicknesses of polymer films on Si wafers were determined using a rotating analyzer spectroscopic ellipsometer (J. A. Woollam, M-44), assuming a film refractive index of 1.5. Transmission FTIR spectra of films on double-side polished Si wafers were collected using a Mattson Galaxy Series 3000 instrument and an air background, and static water contact angles were measured using a First Ten Ångstroms 200 contact angle analyzer. Field-emission scanning electron microscopy (SEM) and energy dispersive X-ray spectroscopy (EDS) mapping were performed with a Hitachi S-4700II microscope using an acceleration voltage of 15 kV.

To quantify Fe(III) ions in polymer brushes, five single-side polished silicon wafers (2.4 × 1.1 cm dimensions) coated with PHEMA-NTA-Fe(III) brushes (average thickness 78 nm) were sequentially immersed in the same 7.0 mL of 50 mM EDTA (pH 7.2) for 5 h with stirring (one wafer was taken out before the new one was put in). The absorbance of Fe(III) in the stripping solution was determined using a Varian Spectra AA-200 atomic absorption spectrophotometer, and the amount of Fe(III) was calculated from this absorbance using a calibration curve, which was determined by measuring the absorbance of $FeCl_3$ standard solutions in 50 mM EDTA.

2.2.4 Sample preparation and MS analysis

To digest proteins into peptides, 100 μg of each protein (β-casein, ovalbumin and bovine serum albumin (BSA)) were dissolved in 20 μL of 6 M urea, 50 mM Tris-HCl. After addition of 5 μL of 10 mM DL-dithiothreitol (for disulfide bond cleavage) the solution was heated in a water

bath at 60 °C for 1 h, followed by dilution with 160 μL of 50 mM ammonium bicarbonate and 10 μL of 100 mM iodoacetamide. The solution was then placed in the dark for 1 h for cysteine carbamidomethylation. Subsequently, the solution was mixed with 10 μL of 0.5 μg/μL trypsin prior to incubation at 37 ˚C for 16 h. To quench the tryptic digestion, 11 μL of glacial acetic acid was finally added to each solution. These digests were dried with a SpeedVac and reconstituted in 5% acetic acid (HOAc). Digests mixtures were prepared by mixing different volumes of each digest, and the amounts of salts and digest reagents decrease with decreasing total digest concentration.

For microspot enrichment of protein digests, 0.5 μL of digest was first spotted onto the microspot and allowed to evaporate (and shrink) for ~5 min. The Si wafer was then placed in a covered, water-saturated petri dish for further incubation for 1 h. (The droplet showed minimal evaporation in the humidified petri dish.) For enrichment on a 2-mm polymer spot, 0.5 μL of protein digest was spotted on the 2-mm well and incubated for 1 h. As the sample droplet evaporated quickly on the 2-mm spot, 0.5 μL aliquots of 5% HOAc were periodically added to the droplet to compensate for evaporation. The micro- and 2-mm spots were then rinsed with ~30 mL of 3:30:67 HOAc:ACN:water (either from a wash bottle or by a 3-min sonication after an initial wash-bottle rinse) to remove salts and unbound nonphosphorylated peptides, and the plate was immediately dried with N_2 to remove any remaining rinse solution. After incubation, rinsing, and drying, a 0.2-μL droplet of 5 mg/mL 2,5-dihydroxybenzoic acid (DHB, Aldrich, prepared in 1:9 ACN:1% H_3PO_4) was gently added to the microspot to form matrix crystals covering an area similar in size to the original microspot. For the 2-mm polymer spot, a 1-μL droplet of 1% phosphoric acid was added to the well immediately followed by 0.25 μL of 40 mg/mL DHB (prepared in 1:1 ACN:1% phosphoric acid). The spot of matrix crystals on the non-patterned

brushes had a diameter of ~2 mm. Using double-sided tape, the Si wafer was then secured to a modified stainless steel MALDI sample plate for subsequent MALDI-MS analysis. For conventional MALDI-MS analysis (no enrichment), 0.5 μL of sample was added to the stainless steel plate, followed by addition of 1 μL of 10 mg/mL DHB in 1:1 ACN:1% phosphoric acid.

Mass spectra were obtained using a MALDI linear ion trap mass spectrometer (Thermo model vMALDI LTQ XL, San Jose, CA) in positive ion mode. For each experiment, mass spectra were obtained from at least 3 spots (micro or 2 mm), and the mass spectra given here are representative examples. Protein Prospector (http://prospector.ucsf.edu) was used to make preliminary assignments to intact precursor ion signals in the mass spectra, and assignments were confirmed using low energy collision-induced dissociation-tandem mass spectrometry (CID-MS/MS) with standard precursor ion isolation and activation conditions.

2.3 Results and discussion

2.3.1 Fabrication and characterization of PHEMA-NTA-Fe(III) microspots surrounded by a hydrophobic film

Figure 2.1 shows an image of a patterned plate prepared using the procedure in Scheme 1. In the patterning process, a PDMS template stamp with 0.25 mm-diameter holes is first attached onto a native Si wafer such that no air bubbles form. The PDMS stamp protects the Si surface, so only the uncovered regions (microspots and the margin outside the PDMS stamp) are exposed to UV/ozone for cleaning and to chemicals for initiator attachment, polymerization, and derivatization. The template stamp should be as thin as possible (but strong enough to be peeled off from the petri dish) so that the microwell does not affect the UV/ozone cleaning and

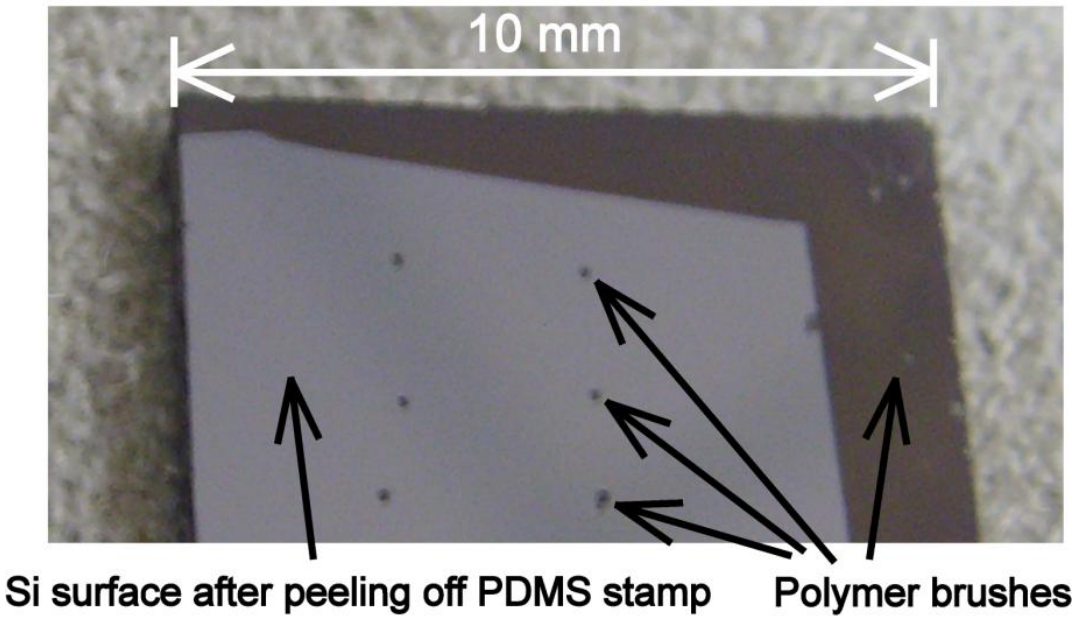


Figure 2.1. Si wafer coated with residual PDMS (light gray) and microspots of PHEMA-NTA-Fe(III) brushes (dark gray). The PDMS stamp did not cover the entire wafer, so brushes also grew outside the light gray area on the wafer.

subsequent chemical reactions. Initially, we did not cure the stamp on the Si wafer at 75 °C, but the wafer became hot after UV/ozone treatment, and this allowed a thin film of PDMS to adhere to the oxide layer on the Si after removing the template PDMS stamp. Because the chemical composition of SiO_2 at the surface is somewhat similar to that of PDMS (-$Si(CH_3)_2$-O-$Si(CH_3)_2$-), some low-molecular-weight PDMS molecules might transfer from the PDMS template surface to the wafer by covalent bonding. To verify the effect of temperature on this procedure, we heated a PDMS-coated Si wafer at 75 °C for 1 h without UV/ozone treatment. After peeling the PDMS stamp off the wafer, a thin hydrophobic layer remains, and the static water contact angle on this surface is 98 °(Figure 2.2c), only ~10 °less than on bulk PDMS (110 °, Figure 2.2b). This angle is higher than that after only UV/ozone treatment and removal of the PDMS stamp (77 °), and much higher than that on a native Si wafer (24 °, Figure 2.2a) or on a

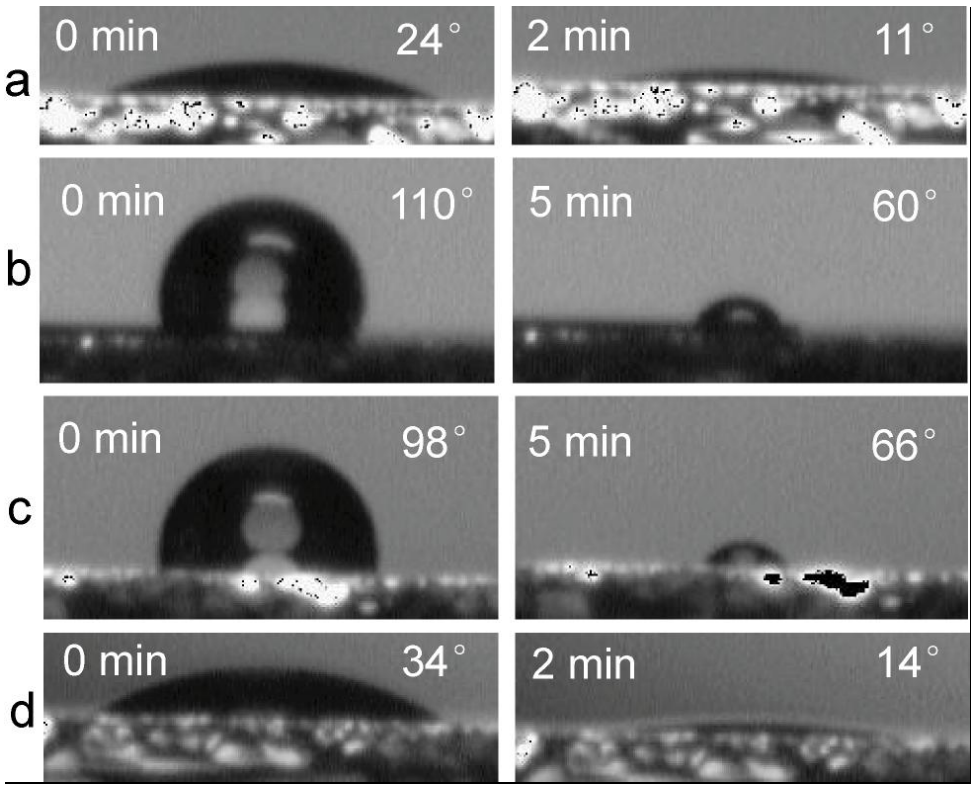


Figure 2.2. Contact angles measured 0, 2, or 5 min after placing water droplets (0.5 μL) on different surfaces: (a) native Si wafer; (b) native PDMS; (c) Si wafer modified with PDMS by curing a PDMS/Si wafer at 75 ℃ for 1 h and peeling off the bulk PDMS; (d) Si wafer with its entire surface coated with PHEMA-NTA-Fe(III) brushes.

wafer with its entire surface coated with polymer brushes (34 º; Figure 2.2d). The ellipsometric thickness of the residual film on the Si wafer after heating and removal of the PDMS stamp is ~1.8 nm. Longer heating times do not enhance the hydrophobicity of the Si wafer, and too high a temperature, e.g. 120 ℃, results in permanent bonding of the entire PDMS stamp to the Si wafer. Adhesion of the stamp to Si is unfavorable for this work because the sample wells are likely too deep for efficient enrichment and MALDI-MS.

For sample concentration by evaporation, the PDMS should remain hydrophobic when exposed to water, and the residual PDMS surface shows a relatively high contact angle (66 º; Figure 2.2c) after 5 min of evaporation. Moreover, the contact area between the drop and the residual PDMS surface decreases to ~5% of the initial area after 5 min of drop evaporation,

which is essential for concentrating samples. In contrast, for the wafer entirely coated with polymer brushes (Figure 2.2d), the area of the droplet-substrate contact has a diameter of ~1.8 mm after 2 min of water evaporation. The hydrophobicity of the residual PDMS film is also stable to sonication in toluene, THF, and acetone for five minutes, in spite of the fact that bulk PDMS stamps swell in these solvents.

Prior to removal of the PDMS stamp from the Si wafer, immobilization of the silane initiator on the uncovered Si regions occurs from anhydrous DMF. We use DMF rather than toluene to avoid swelling of the bulk PDMS template. Growth of PHEMA from immobilized initiators for 1 h yields ~40 nm-thick polymer brushes on the uncovered Si surface. (The thickness of the brush was determined ellipsometrically on the uncovered area outside of the PDMS stamp. See Figure 2.1.) Derivatization of the PHEMA brushes with NTA-Fe(III) increases the film thickness to ~80 nm. After removing the PDMS stamp, the hydrophobic region covered with residual PDMS shows the original color (light gray) of the Si, while the uncovered regions is a dark gray to blue color (Figure 2.1) due to the thickness of the polymer brushes. The color distinction is important because it allows rapid location of the polymer brush spots.

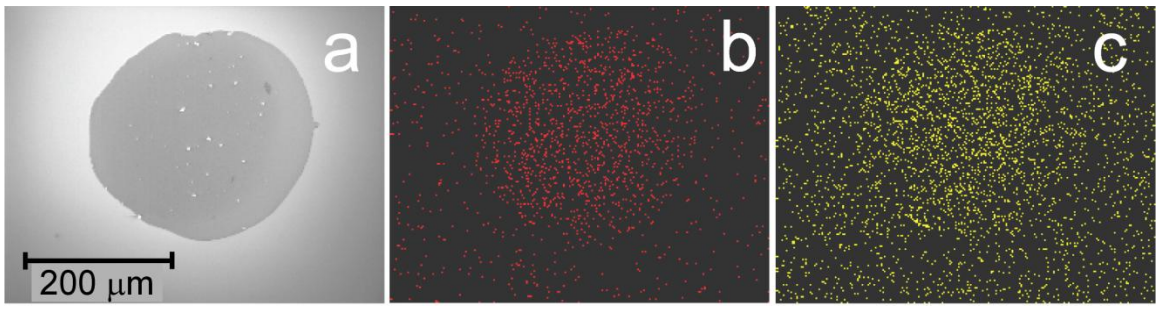


Figure 2.3. (a) SEM image of a microspot of PHEMA-NTA-Fe(III) brushes surrounded by a residual PDMS layer on Si, (b) EDS carbon mapping of the spot, and (c) EDS oxygen mapping of the spot.

SEM images (Figure 2.3a) show that the microspots are well-defined with an average diameter of 280 μm, which is close to the inner diameter (250 μm) of the blunt dispensing tip used to form the pores in the PDMS stamp. EDS mapping shows higher distributions of carbon and oxygen within the microspot, which is due to the relatively thick PHEMA-NTA-Fe(III) brushes (Figure. 3b, c). Unfortunately, the signal for Fe is too weak for imaging.

The PHEMA synthesis and subsequent chemical modifications were also performed on the entire surface of Si wafers to compare enrichment on different size spots (see below) and to allow FTIR spectroscopic characterization of the brush synthesis. The IR of polymer film after each reaction step is consistent with previous results. Briefly, the ester carbonyl absorbance at 1725 cm^{-1} in Figure 2.4a confirms the formation of PHEMA. The doubled carbonyl absorption near 1740 cm^{-1} in Figure 2.4b reveals ester functionality. The pronounced succinimde ester peaks at 1817 and 1786 cm^{-1} in Figure 2.4c confirms the activation of the carboxylic acid with NHS. The broad peak near 1650 cm^{-1} is probably due to the mixed adsorption of amide bond and carboxylate groups of NTA (Figure 2.4d). Lastly, the small absorption at 1590 cm^{-1} indicates the complexation of Fe(III) to NTA. Atomic absorption spectroscopy shows that the amount of Fe(III) chelated in PHEMA-NTA on fully coated wafers is ~14 nmol/cm^2. This high density of complexes provides a high capacity for phosphopeptide binding compared to monolayers of Fe(III) complexes. (A monolayer of alkanethiols has a density around 0.8 nmol/cm^2, and a monolayer of NTA-Fe(III) on Si will have an even lower surface coverage.)[24]

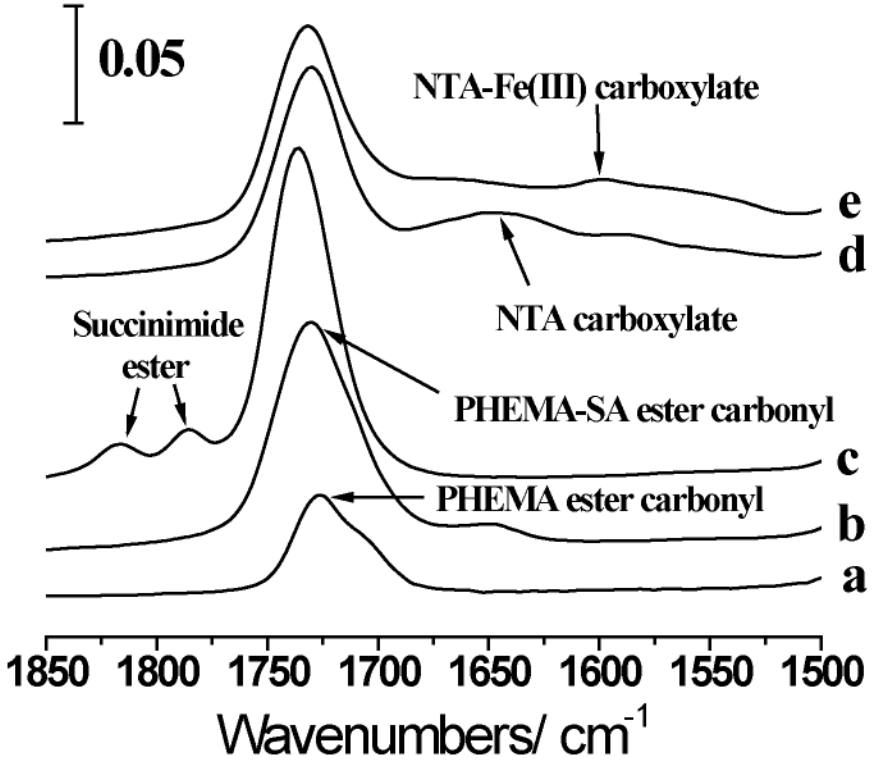


Figure 2.4. Transmission FTIR absorbance spectrum of a double-side polished Si wafer coated with PHEMA brushes on both sides (a), and absorbance spectra of the same sample after the following modification steps: (b) reaction with succinic anhydride (SA), (c) activation with 0.1 M *N*-hydroxysuccinimide/*N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (NHS/EDC), (d) derivatization with aminobutyl NTA and (e) immersion in 0.1 M $FeCl_3$ followed by rinsing with 250 mM HOAc/30% acetonitrile/70% water. Reference 13 provides a detailed explanation of similar spectra of PHEMA films on Au.

2.3.2 MALDI-MS analysis of phosphoprotein digests enriched on PHEMA-NTA-Fe(III) microspots

We initially examined phosphopeptide enrichment from digests of β-casein and ovalbumin. Table 2.1 provides the sequences of tryptic phosphopeptides from these two model phosphoproteins. Without enrichment of phosphopeptides, 1 pmol of β-casein digest deposited on a steel MALDI plate gives rise to a complex mass spectrum (Figure 2.5a). The three phosphopeptides, 1P2062, 4P2966, and 4P3122, have weaker signals than other nonphosphopeptides. After enrichment of the same amount of β-casein digest on a 2-mm

PHEMA-NTA-Fe(III) spot, however, 1P2062 dominates the spectrum, and the two tetra-phosphopeptides, 4P2966 and 4P3122, also give signals strong enough for MS/MS (Figure 2.5b). Moreover, many small peaks correspond to phosphopeptides from α-casein impurities and phosphopeptide-related species.

Table 2.1. Phosphopeptides that result from tryptic digestion of β-casein and ovalbumin. The amino acids in parentheses represent cleavage sites and are not part of the peptide sequence.

Phospho-protein	$[M+H]^+$ m/z	Phosphory-lation sites	Amino acid sequence
β-casein	2062	1	(K)FQpSEEQQQTEDELQDK
β-casein	2966	4	(R)ELEELNVPGEIVEpSLpSpSpSEESITR
β-casein	3122	4	RELEELNVPGEIVEpSLpSpSpSEESITR*
ovalbumin	2089	1	(R)EVVGpSAEAGVDAASVSEEFR
ovalbumin	2511	1	(K)LPGFGDpSIEAQCGTSVNVHSSLR
ovalbumin	2901	1	(R)FDKLPGFGDpSIEAQCGTSVNVHSSLR*

* 4P3122 and 1P2901 are generated by missed cleavage of the phosphoprotein by trypsin.

At protein amounts ≤1 pmol, enriching the sample on the microspot greatly increases the signals due to the 4P2966 and 4P3122 peptides. As Figure 2.5c shows, after the enrichment of 1 pmol of β-casein digest on a microspot, the tetra-phosphorylated peptides exhibit some of the strongest signals in the mass spectrum. The signal-to-noise ratios (S/N) for 4P3122 and 4P2966 are ~9-fold and ~3-fold greater with microspot enrichment than with 2-mm-spot enrichment (compare Figure 2.5 b and c). The reproducibly weak signal for 1P2062 in Figure 2.5c likely results from saturation of binding sites on the microspot with tetra-phosphopeptides (see below). With 2-mm-spot enrichment, tetra-phosphopeptide signals dominate the MALDI mass spectrum only when the amount of β-casein digest is above 7 pmol (data not shown), whereas at 250 fmol of digest, signals due to 4P2966 and 4P3122 are already at the detection limit (Figure 2.5d). In contrast, microspot enrichment still yields a high S/N for multiply phosphorylated peptides in 250 fmol of digested protein (Figure 2.5e). Unfortunately, when the amount of β-casein digest is

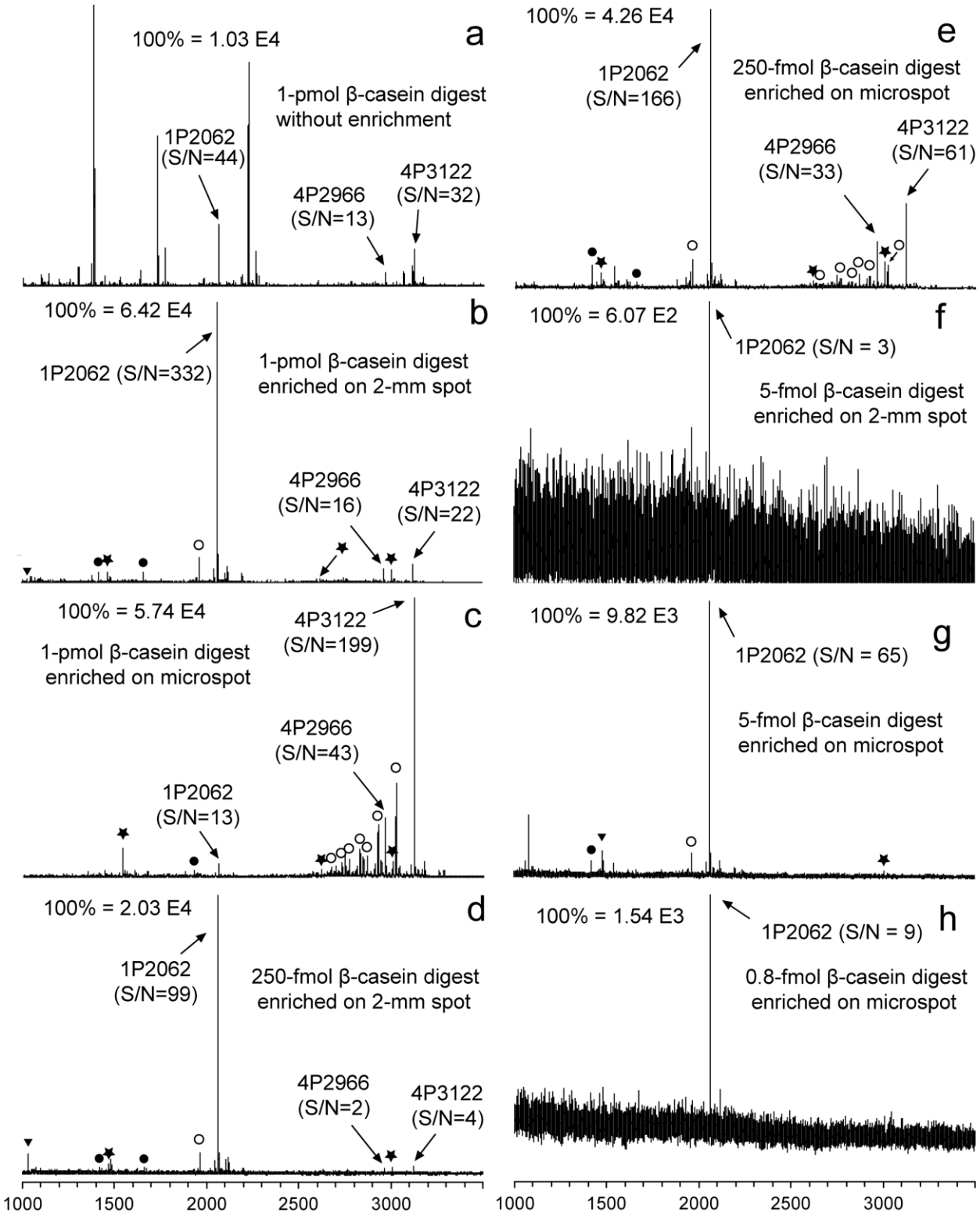


Figure 2.5. MALDI mass spectra of β-casein digests obtained using: (a) a conventional stainless steel MALDI plate without enrichment, and (b-h) enrichment on wafers coated with PHEMA-NTA-Fe(III) brushes. The amount of β-casein digest and the size of enrichment spots are noted

in each spectrum. Mono- and tetra-phosphorylated peptides are labeled as 1P and 4P, respectively, with their S/N in parentheses. Other phosphopeptide-related peaks are labeled with different symbols: ▼ doubly protonated phosphopeptide, ● α-S1-casein phosphopeptide, ★ α-S2-casein phosphopeptide, ○ phosphopeptide signal after loss of 98 Da (H_3PO_4). Peak identification was facilitated by CID-MS/MS.

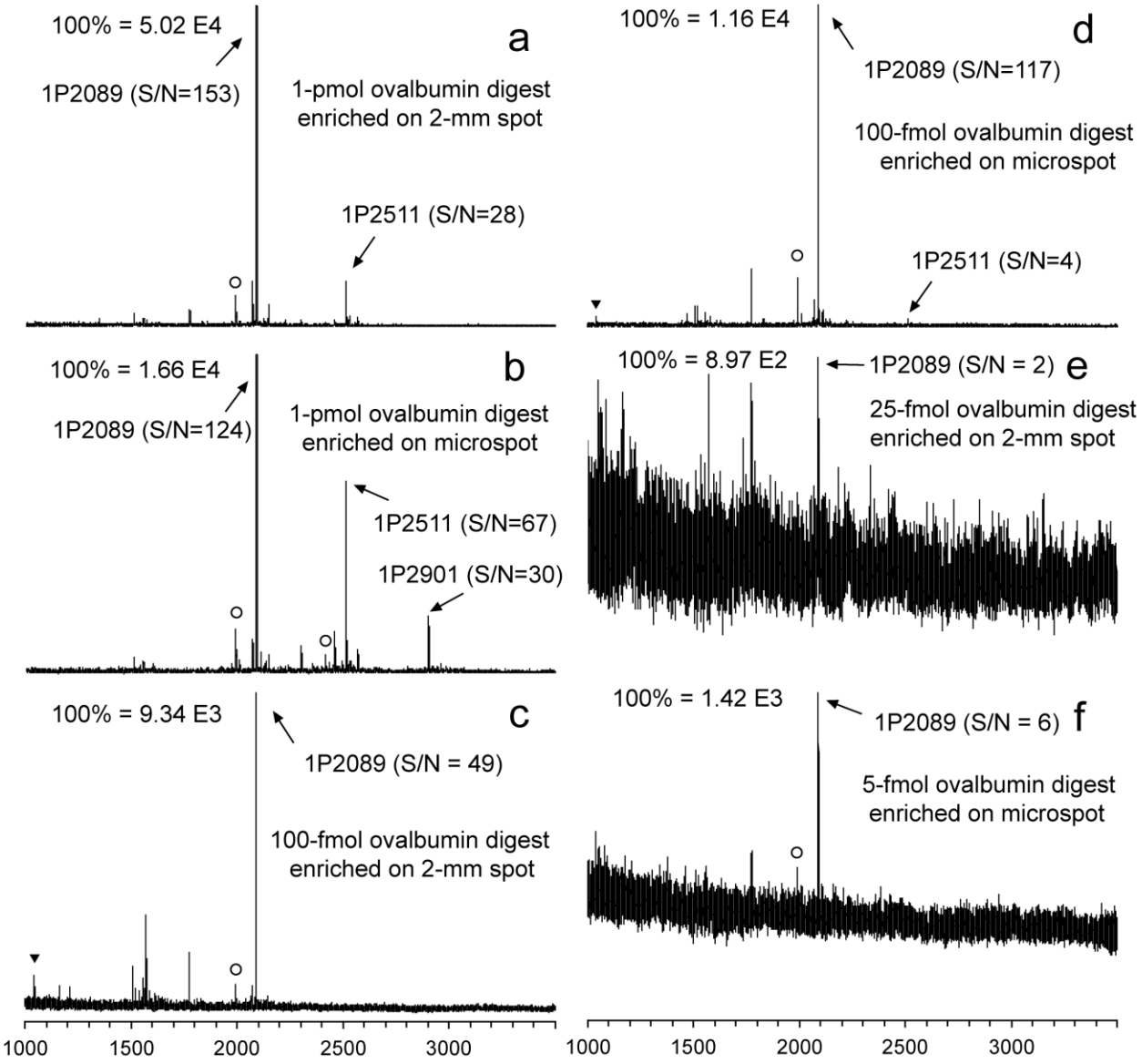


Figure 2.6. MALDI mass spectra of ovalbumin digests after enrichment on PHEMA-NTA-Fe(III) brushes. The amount of ovalbumin digest and the size of enrichment spot are noted in each mass spectrum. The mono-phosphopeptides are labeled as 1P, with their S/N ratios in parentheses. Phosphopeptides formed by loss of 98 Da (H_3PO_4) are labeled as ○; doubly protonated phosphopeptides are labeled as ▼. Peak identification was facilitated by CID-MS/MS.

below 100 fmol, enrichment on the microspot does not allow detection of 4P2966 and 4P3122.

The ability of the microspot to enrich mono-phosphopeptides from very small amounts of digest is especially remarkable. Even 5 fmol of β-casein digest generates a very strong peak for 1P2062 (S/N=65) with microspot enrichment, whereas the corresponding signal with 2-mm-spot enrichment is at the detection limit (Figure 2.5 f and g). The lowest level for detecting 1P2062 after microspot enrichment is around 0.8 fmol (S/N=9, Figure 2.5h), which is significantly lower than the approximate detection limits for 2-mm-spot enrichment (5 fmol, Figure 2.5f).

Enrichment on microspots rather than 2-mm spots also enhances the detection of phosphopeptides in ovalbumin digests (Figure 2.6). With 1 pmol of digest, enrichment on the 2-mm spots followed by MALDI-MS reveals the 1P2089 and 1P2511 peptides, whereas enrichment on the microspots allows detection of a third phosphopeptide, 1P2901. Moreover, the lowest detection limit for 1P2089 by MALDI-MS after microspot enrichment is ~5 fmol (S/N=6), which is 5-fold lower than the approximate detection limit after 2-mm-spot enrichment (25 fmol, S/N=2, see Figure 2.6 e and f).

2.3.3 Enrichment of phosphopeptides from mixtures of protein digests

Enrichment of a simple mixture of β-casein and ovalbumin digests (25 fmol of each) yields signals of phosphopeptides from both phosphoproteins (Figure 2.7a). However, in many biological samples, the phosphoprotein represents only a small fraction of the total protein, so we also tested phosphopeptide enrichment on microspots of PHEMA-NTA-Fe(III) using a mixture that contains digested β-casein, ovalbumin and BSA with a molar ratio of 1:1:10. (BSA is not a phosphoprotein.) Because there are many more nonphosphopeptides than phosphopeptides in the sample, enrichment of this complex mixture on a 2-mm spot yields relatively large nonphosphopetide signals and only one strong phosphopeptide signal that can be distinguished

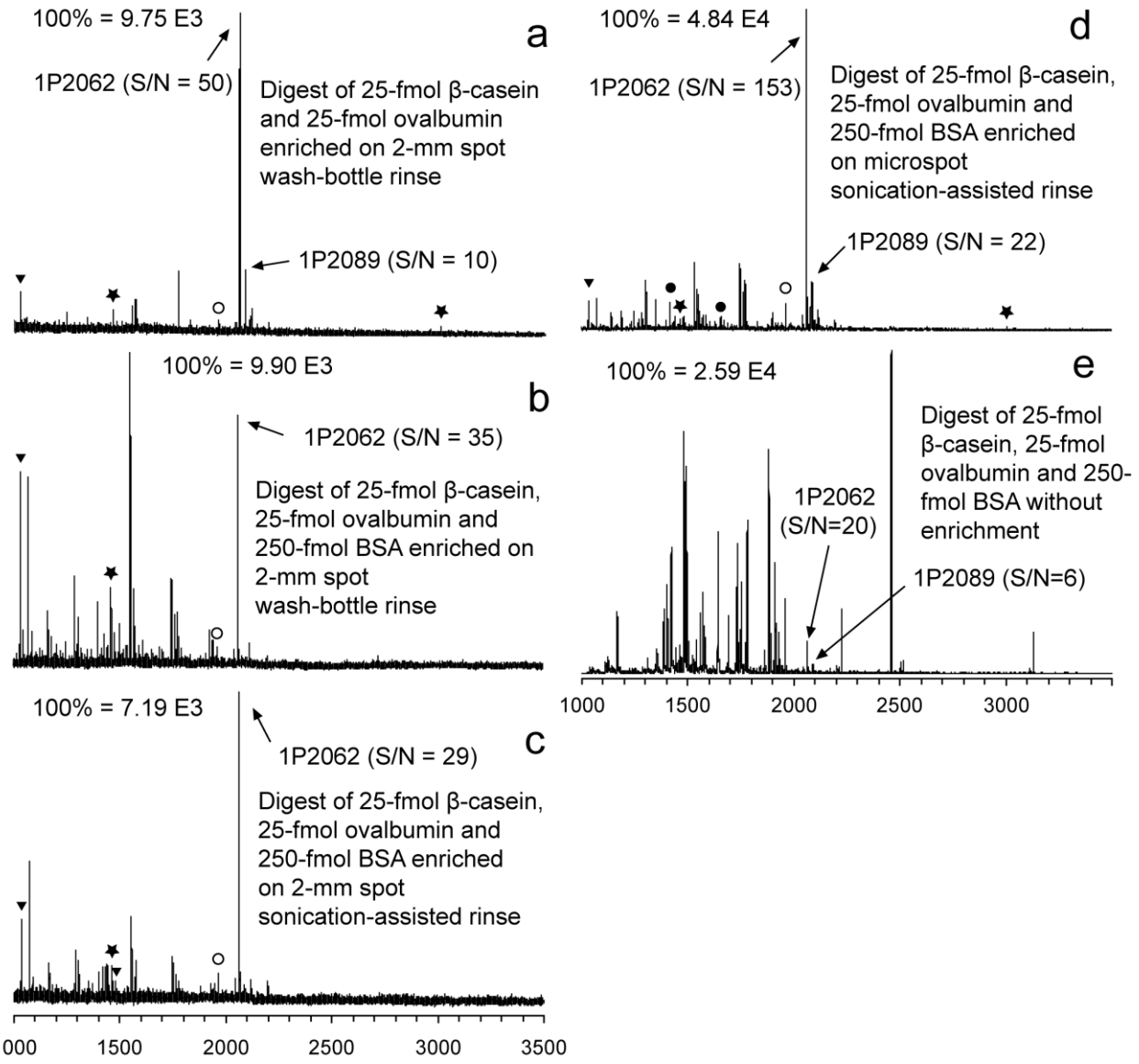


Figure 2.7. MALDI mass spectra of mixtures of protein digests with (a-d) and without phosphopeptide enrichment (e). The digest contents, size of enrichment spot (a-d) and method of rinsing (a-d) are noted in the spectra. Mono-phosphopeptides are labeled as 1P, with their S/N ratios in parentheses. Other phosphopeptide-related peaks are labeled with different symbols: ▼ doubly protonated phosphopeptide, ● α-S1-casein phosphopeptide, ★ α-S2-casein phosphopeptide, and ○ phosphopeptide signal after loss of 98 Da (H_3PO_4). Peak identification was facilitated by CID-MS/MS.

from adjacent nonphosphopeptides (Figure 2.7b). To improve the efficiency of rinsing, we immerse the enriched samples in a sonicated HOAc/ACN/water solution. This results in smaller signals for nonphosphopeptides, but 1P2062 is still the only phosphopeptide (generated from β-

casein and ovalbumin) that can be identified on the 2-mm spot (Figure 2.7c). No signals from multiply phosphorylated peptides of β-casein or mono-phosphopeptides of ovalbumin are detectable. Using microspot enrichment with a sonication-assisted rinse, however, the S/N of 1P2062 is 5-fold higher than after 2-mm-spot enrichment, and the mono-phosphopeptide 1P2089 from ovalbumin, as well as three more phosphopeptides from α-casein, give detectable signals (Figure 2.7d). The S/N for 1P2062 and for 1P2089 are also greater than in conventional MALDI-MS (Figure 2.7e).

2.3.4 Mechanism of the enrichment of phosphopeptides on PHEMA-NTA-Fe(III) microspots

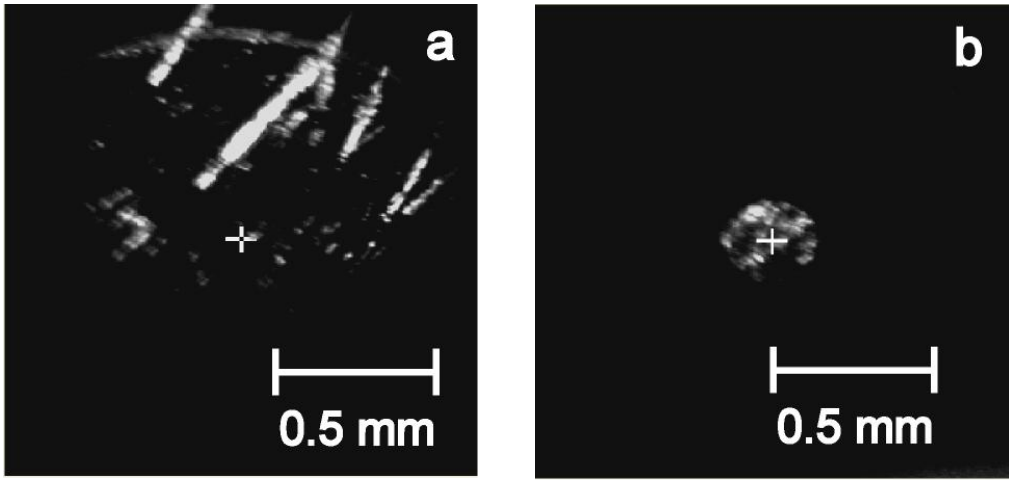


Figure 2.8. Charge-coupled device images of sample enriched on (a) a 2-mm spot and (b) a microspot. The images were obtained after addition of DHB matrix. (Image (a) shows only a portion of the 2-mm spot. The cross is not at the center of the spot.)

Because of differences in sample area, the surface concentration of phosphopeptides on the microspot should ideally be 64 times higher than on the 2-mm spot. Figure 2.8 shows that the matrix on the microspot covers a much smaller area and is more homogeneous than on the 2-mm spot. However, the S/N obtained with the microspot is only 3- to 20-fold greater than the S/N

with the 2-mm spot. Other studies also reported that MALDI-MS signals are not directly proportional to sample surface area.[14-16] When a peptide/matrix solution deposited on parafilm formed a layer of sample/matrix crystals covering an area ~1/10 of that on a stainless steel MALDI plate, the S/N was only 3- to 5-fold higher than on the stainless steel plate.[15] Moreover, in our work, the improvement in S/N for phosphopeptides using the microspot approach is much greater for low-femtomole amounts of phosphoprotein digest than for amounts >100 fmol. There are likely two reasons for these findings.

First, there is already some concentration of the sample during matrix crystallization on 2-mm spots. After adding the DHB matrix solution onto 2-mm spots, the phosphopeptides dissolve and then partition into certain matrix crystals. Mass spectra typically come from "sweet spot" crystals that likely contain a disproportionate fraction of the total phosphopeptides, whereas MALDI-MS signals from other regions of the well are very weak. Addition of 6.4-fold more matrix to achieve the same amount of matrix per area on the 2-mm spot as on the microspot does not improve the homogeneity of the 2-mm sample and gives a ~25% decrease in S/N.

Second, there is some loss of phosphopeptides in microspot enrichment, especially at high levels of phosphopeptide because the binding capacity of the microspot is not sufficient to capture all of the phosphopeptides in a sample. According to our previous work,[13] the binding capacity of ~60-nm thick PHEMA-NTA-Fe(III) brushes is 0.6 μg of phosphoangiotensin per cm^2. Considering that the thickness of polymer brushes in the present work is ~80 nm, and that the number of phosphopeptides is 2 per digested β-casein or ovalbumin molecule (the third phosphopeptide in both cases is due to unspecific or missed cleavage), as a first approximation the binding capacity of a 2-mm spot should be ~5 pmol for a β-casein or ovalbumin digest. (This calculation assumes that the binding capacity in μg is independent of molar mass.) The binding

capacity of a 0.25-mm spot is only ~80 fmol for a β-casein or ovalbumin digest. The low binding capacity of the microspot likely explains why the overall S/N improvement after microspot enrichment is more evident when the amount of phosphoprotein digest is below 100 fmol. At phosphopeptide levels above the binding capacity of the polymer microspot, the excess phosphopeptides will be removed with the nonphosphopeptides in the rinsing step.

When there is a high excess of phosphopeptides on the microspot, multiply phosphorylated peptides are more likely to bind to the available affinity sites than mono-phosphopeptides.[25,26] As Figure 2.5c shows, after enrichment of 1 pmol of β-casein digest on the microspot (80-fmol binding capacity), the sum of the S/N of the tetra-phosphopeptides is 19 fold higher than the S/N of the mono-phosphopeptide. In contrast, enrichment of the same amount of sample on a 2-mm spot (~5-pmol binding capacity) results in a sum of the S/N of the tetra-phosphopeptides that is 9-fold lower than the S/N of the mono-phosphopeptide. In samples where the amount of phosphopeptide is significantly less than the binding capacity, the brushes bind both tetra- and mono-phosphopeptides, but the signals of tetra-phosphopeptides are lower, probably because of inefficient elution or ionization of the tetra-phosphopeptides. When trying to estimate relative levels of phosphopeptides, it is important to spot dilute samples whose phosphopeptide amount does not exceed the binding capacity.

2.4 Conclusions

This work presents a simple method for fabricating polymer brush microspots (250 μm diameter) surrounded with a hydrophobic PDMS layer. The use of these spots for on-plate enrichment of femtomole amounts of phosphopeptides allows pre-concentration of the sample droplet on the plate and gives an improved phosphopeptide MALDI-MS S/N compared to

enrichment on larger spots. Moreover, the polymer microspot can enrich low-femtomole amounts of phosphopeptides even when the sample contains an order of magnitude excess of non-phosphoprotein digest. Enhancement of the MALDI-MS S/N for mono-phosphopeptides is especially evident in the enrichment of low-femtomole amounts of phosphoprotein digests because the amount of phosphopeptide in the sample is less than the binding capacity. Thus, this microspot enrichment method may be particularly attractive for analysis of samples such as immunoprecipitates, where a limited amount of material is available. Moreover, high-throughput MALDI-MS analysis can be realized by fabricating arrays of microspots on a Si wafer. The small size of the polymer microspot also enhances the detection of multiply phosphorylated peptides, but these species still have higher detection limits than mono-phosphorylated peptides.

Because the microspot fabrication does not employ any special instruments, it might be a substitute for microcontact printing in some cases. Moreover, the principle of this method could be employed to synthesize microspots of other polymers on Si wafers for the enrichment of different species.

REFERENCES

REFERENCES

1. Paradela, A.; Albar, J. P., Advances in the Analysis of Protein Phosphorylation. *J. Proteome Res.* **2008,** 7, 1809-1818.

2. Thingholm, T. E.; Jensen, O. N.; Larsen, M. R., Analytical Strategies for Phosphoproteomics. *Proteomics* **2009,** 9, 1451-1468.

3. Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A., Analysis of Protein Phosphorylation Using Mass Spectrometry: Deciphering the Phosphoproteome. *Trends Biotechnol.* **2002,** 20, 261-268.

4. Steen, H.; Jebanathirajah, J. A.; Rush, J.; Morrice, N.; Kirschner, M. W., Phosphorylation Analysis by Mass Spectrometry - Myths, Facts, and the Consequences for Qualitative and Quantitative Measurements. *Mol. Cell. Proteomics* **2006,** 5, 172-181.

5. Goshe, M. B.; Conrads, T. P.; Panisko, E. A.; Angell, N. H.; Veenstra, T. D.; Smith, R. D., Phosphoprotein Isotope-Coded Affinity Tag Approach for Isolating and Quantitating Phosphopeptides in Proteome-Wide Analyses. *Anal. Chem*. **2001,** 73, 2578-2586.

6. McLachlin, D. T.; Chait, B. T., Improved Beta-Elimination-Based Affinity Purification Strategy for Enrichment of Phosphopeptides. *Anal. Chem*. **2003,** 75, 6826-6836.

7. Warthaka, M.; Karwowska-Desaulniers, P.; Pflum, M. K. H., Phosphopeptide Modification and Enrichment by Oxidation-Reduction Condensation. *ACS Chem. Biol.* **2006,** 1, 697-701.

8. Dunn, J. D.; Watson, J. T.; Bruening, M. L., Detection of Phosphopeptides Using Fe(III)-Nitrilotriacetate Complexes Immobilized on a MALDI Plate. *Anal. Chem*. **2006,** 78, 1574-1580.

9. Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M., Phosphoproteome Analysis by Mass Spectrometry and Its Application to Saccharomyces Cerevisiae. *Nature Biotechnol.* **2002,** 20, 301-305.

10. Zhou, H. J.; Xu, S. Y.; Ye, M. L.; Feng, S.; Pan, C.; Jiang, X. G.; Li, X.; Han, G. H.; Fu, Y.; Zou, H., Zirconium Phosphonate-Modified Porous Silicon for Highly Specific Capture of Phosphopeptides and MALDI-Tof MS Analysis. *J. Proteome Res.* **2006,** 5, 2431-2437.

11. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D., Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Mol. Cell. Proteomics* **2005,** 4, 873-886.

12. Qiao, L.; Roussel, C.; Wan, J. J.; Yang, P. Y.; Girault, H. H.; Liu, B. H., Specific on-Plate Enrichment of Phosphorylated Peptides for Direct MALDI-Tof MS Analysis. *J. Proteome Res.* **2007,** 6, 4763-4769.

13. Dunn, J. D.; Igrisan, E. A.; Palumbo, A. M.; Reid, G. E.; Bruening, M. L., Phosphopeptide Enrichment Using MALDI Plates Modified with High-Capacity Polymer Brushes. *Anal. Chem*. **2008,** 80, 5727-5735.

14. Hung, K. C.; Rashidzadeh, H.; Wang, Y.; Guo, B. C., Use of Paraffin Wax Film in MALDI-Tof Analysis of DNA. *Anal. Chem*. **1998,** 70, 3088-3093.

15. Wang, J. H.; Chen, R. B.; Ma, M. M.; Li, L. J., MALDI MS Sample Preparation by Using Paraffin Wax Film: Systematic Study and Application for Peptide Analysis. *Anal. Chem*. **2008,** 80, 491-500.

16. Hung, K. C.; Ding, H.; Guo, B. C., Use of Poly(Tetrafluoroethylene)S as a Sample Support for the MALDI-Tof Analysis of DNA and Proteins. *Anal. Chem*. **1999,** 71, 518-521.

17. Schuerenbeg, M.; Luebbert, C.; Eickhoff, H.; Kalkum, M.; Lehrach, H.; Nordhoff, E., Prestructured MALDI-MS Sample Supports. *Anal. Chem*. **2000,** 72, 3436-3442.

18. Chen, X. X.; Murawski, A.; Kuang, G.; Sexton, D. J.; Galbraith, W., Sample Preparation for MALDI Mass Spectrometry Using an Elastomeric Device Reversibly Sealed on the MALDI Target. *Anal. Chem*. **2006,** 78, 6160-6168.

19. Jia, W. T.; Wu, H. X.; Lui, H. J.; Li, N.; Zhang, Y.; Cai, R. F.; Yang, P. Y., Rapid and Automatic on-Plate Desalting Protocol for Mald-MS: Using Imprinted Hydrophobic Polymer Template. *Proteomics* **2007,** 7, 2497-2506.

20. Xu, Y. D.; Bruening, M. L.; Watson, J. T., Use of Polymer-Modified MALDI-MS Probes to Improve Analyses of Protein Digests and DNA. *Anal. Chem*. **2004,** 76, 3106-3111.

21. Xu, Y. D.; Watson, J. T.; Bruening, M. L., Patterned Monolayer/Polymer Films for Analysis of Dilute or Salt-Contaminated Protein Samples by MALDI-MS. *Anal.l Chem*. **2003,** 75, 185-190.

22. Xia, Y. N.; Whitesides, G. M., Soft Lithography. *Angew. Chem.-Int. Ed.* **1998,** 37, 551-575.

23. Matyjaszewski, K.; Miller, P. J.; Shukla, N.; Immaraporn, B.; Gelman, A.; Luokala, B. B.; Siclovan, T. M.; Kickelbick, G.; Vallant, T.; Hoffmann, H.; Pakula, T., Polymers at Interfaces: Using Atom Transfer Radical Polymerization in the Controlled Growth of Homopolymers and Block Copolymers from Silicon Surfaces in the Absence of Untethered Sacrificial Initiator. *Macromolecules* **1999,** 32, 8716-8724.

24. Ulman, A., Formation and Structure of Self-Assembled Monolayers. *Chem. Rev*. **1996,** 96, 1533-1554.

25. Li, Q. R.; Ning, Z. B.; Tang, J. S.; Nie, S.; Zeng, R., Effect of Peptide-to-Tio2 Beads Ratio on Phosphopeptide Enrichment Selectivity. *J. Proteome Res.* **2009,** 8, 5375-5381.

26. Qiao, L. A.; Bi, H. Y.; Busnel, J. M.; Hojeij, M.; Mendez, M.; Liu, B. H.; Girault, H. H., Controlling the Specific Enrichment of Multi-Phosphorylated Peptides on Oxide Materials: Aluminium Foil as a Target Plate for Laser Desorption Ionization Mass Spectrometry. *Chem. Sci.* **2010,** 1, 374-382.

Chapter 3. Identification of p65-associated Phosphoproteins by Mass Spectrometry after On-plate Phosphopeptide Enrichment Using Polymer-oxotitanium Films

(The work described in this chapter was performed in collaboration with Dr. Amanda M. Palumbo from Prof. Gavin E. Reid's and Prof. Jetze J. Tepe's groups at Michigan State University. This work has been published in: W.H. Wang, A.M. Palumbo, Y.J. Tan, G.E. Reid, J.J. Tepe, M.L. Bruening, Identification of p65-Associated Phosphoproteins by Mass Spectrometry after On-Plate Phosphopeptide Enrichment Using Polymer-oxotitanium Films. *J. Proteome Res.* 2010, 9, 3005–3015.)

3.1 Introduction

Nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) is an ubiquitous transcription factor that regulates the transcription of genes by binding to discrete DNA sequences.[1] The NF-κB signaling pathway impacts many aspects of cellular adaptation such as activation of immune cell function[2] and programmed cell death[3]. Seven NF-κB family proteins exist in mammalian cells, including RelA (p65), RelB, c-Rel, p50, p105, p52 and p100,[4] and these proteins form combinations of homo- and heterodimers. The different NF-κB proteins share an N-terminal NF-κB/Rel homology domain (RHD), which mediates DNA binding, dimerization, nuclear translocation and interaction with the inhibitory IκB family proteins.[4] In resting cells, NF-κB is inactive due to association with IκB proteins. Stimulation of cells with agonists results in phosphorylation, ubiquitination and degradation of IκBs, which

releases NF-κB for nuclear translocation and activation of gene transcription.[5] Well-known stimuli that induce NF-κB activity include cytokines, γ radiation, and DNA damage.[4]

Within the NF-κB protein family, the heterodimmer p50/p65 is the most abundant and prototypical form, and only the p65 subunit contains a C-terminal transcriptional activation domain. Many biochemical and genetic experiments suggest an important role of post-translational modifications of NF-κB subunits in its function. For example, different protein kinases phosphorylate p65 at Thr254, Ser276, Ser311, Thr435, Ser468, Ser529 and Ser536 residues.[1,4,6,7] These modifications enhance the transcriptional activity of p65 and its ability to interact with co-activators.[1] In addition, the phosphorylation of other proteins may control their interactions with p65 and alter the regulation of p65-mediated transcriptional activity.[8] Thus, a full understanding of p65 transcriptional regulation requires examination of the post-translational modifications of both p65 and its associated proteins in response to different stimuli. A rapid, sensitive analysis technique that requires low sample volumes would facilitate such studies.

This work describes an analytical approach to elucidate the phosphorylation status of a target protein and its associated proteins in small samples (μg quantities), with a specific emphasis on proteins associated with immunoprecipitated p65. Currently, profiling of protein phosphorylation with MS typically requires enrichment of phosphopeptides from proteolytic digests using immobilized metal ion affinity chromatography (IMAC) with Fe(III) or Ga(III) complexes or metal oxide affinity chromatography.[9-12] Several recent studies suggest that TiO_2 particles or IMAC-Ti(IV) frequently exhibit greater specificity for phosphopeptide adsorption than IMAC-Fe(III) or –Ga(III).[13-19] However, most enrichment methods utilize column-based separations in which the sample volume is more than that needed for MS,

especially for MALDI-MS. Moreover, column-based procedures often require extensive sample handling that decreases throughput and risks sample loss due to adsorption on columns and container walls.[9]

One strategy to minimize sample consumption and simplify enrichment is capture of the desired species (~1 μL sample volume) directly on a MALDI plate. Compared to electrospray ionization (ESI), MALDI allows for lower sample consumption, as well as decreased complexity in the MS spectrum due to the formation of primarily singly protonated ions. Moreover, the stability of sample/matrix crystals on MALDI plates allows multiple MS or MS/MS experiments to be performed on the same spot. Several groups reported on-plate enrichment of model phosphopeptides with TiO_2-coated surfaces[16,20-22] or other substrates.[9] However, most studies did not employ a chaotropic denaturing agent (e.g. urea)[23] during sample preparation, which is often necessary for complete protein digestion, especially for cell lysates.[24] High concentrations of urea (up to 6 M) are often deleterious to enrichment procedures and MALDI-MS.[25] Moreover, as far as we know, none of the previous on-plate enrichment methods were used to examine the phosphorylation status of proteins immunoprecipitated from cell extracts. In the case of column-based techniques, the majority of research efforts aimed at enrichment of hundreds of phosphopeptides from whole cell extracts,[11] rather than addressing biological questions related to one or several target phosphoproteins that may be present at low levels in cells.

This chapter describes a polymer brush-oxotitanium (P-oxoTi) hybrid material synthesized on Au wafers for on-plate enrichment of phosphopeptides in digested immunoprecipitates that contain a number of proteins and high levels of urea and salts. Selective binding to titania nanoclusters in P-oxoTi brushes results in more MS signals from

phosphopeptides than on-plate capture with IMAC or enrichment with commercial TiO_2 particles. Moreover, enrichment on Au-P-oxoTi facilitates the identification of several phosphopeptides from digests of p65-associated proteins immunoprecipitated together with p65 from human acute monocytic leukemia cell (THP-1) nuclear extracts. MALDI-MS/MS of the enriched phosphopeptides and database searching allow the elucidation of phosphopeptide sequences, phosphorylation sites, and the parent phosphoproteins.

3.2 Experimental

3.2.1 Chemicals and materials

Most reagents for polymer brush synthesis, sample preparation and phosphopeptide enrichment were purchased from Sigma-Aldrich. Other chemicals and materials used for protein digestion, enrichment, and MS include: sequencing grade modified porcine trypsin (treated with L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK), Promega), Tris-HCl (Invitrogen), urea (J.T.Baker), DL-dithiothreitol (Fluka), ammonium bicarbonate (Columbus Chemical), trifluoroacetic acid (TFA, EMD Chemicals), acetonitrile (ACN, Mallinckrodt Baker), propionic anhydride-d_{10} (98%, C/D/N Isotopes), TiO_2 particles (Titansphere, 5 μm, GL Science, Japan), and desalting pipette tips (ZipTipC18, Millipore). Tetrahydrofuran (THF, Mallinckrodt Baker) was newly distilled before use, and deionized water was prepared with a Millipore ion-exchange system (Milli-Q, 18 MΩ cm). Au-coated Si wafers (200 nm of sputtered Au on 20 nm of sputtered Cr on Si wafers) were prepared by LGA Thin Films.

Table 3.1. *N*-Propionylated standard phosphopeptides employed for determination of enrichment recoveries.*

Peptide sequence	# $-PO_3H_2$	# $-NHCOCH_2CH_3$	m/z MH^+ after propionylation
LPQEpTAR	1	1	950.5
RDpSLGTYSSR	1	1	1277.6
EVQAEQPSSpSSPR	1	1	1537.7
LGPGRPLPTFPpTSE(carbamidomethyl)TSDVEPDTR	1	1	2765.6
RYpSpSRSR	2	1	1127.5
pSRSPpSSPELNNK	2	2	1587.7
IHRLARpYpYKTK	2	3	1777.0

* The peptides listed in this table are those with complete N-propionylation (propionyl-derivatization occurs on both the N-terminus and the primary amine of Lys residues).

A mixture of 10 standard phosphopeptides (donated by Sigma Genosys (The Woodlands, TX) was *N*-propionylated for quantitative evaluation of phosphopeptide recovery on Au-P-oxoTi wafers. The peptide mixture (200 pmol of total phosphopeptides) was dissolved in 20 μL of 50 mM NH_4HCO_3. Half of the solution was mixed with 20 μL of propionic anhydride solution (5 μL of propionic anhydride in 45 μL of methanol) and allowed to react for 2 h at room temperature. The sample was dried with a SpeedVac, and the resultant phosphopeptides were dissolved in 10 μL of 0.1% TFA. The other half of the standard mixture solution was derivatized using deuterated propionic anhydride in the same way. With this protocol, all the standard phosphopeptides were propionylated on at least one primary amine (no unmodified phosphopeptides were found in MALDI-mass spectra). The majority of phosphopeptides were completely propionylated on N-terminus and Lys residues, but some peptides were over-propionylated on other residues. Only phosphopeptides with complete propionylation (Table 3.1) were used for determination of the percent recovery. Procedures for synthesizing the single standard phosphopeptide, CD_3CD_2CO-LFTGHPEpSLEK, which was used as an internal

standard for the semi-quantitative analysis of p65-associated phosphoproteins, were reported before.[26]

THP-1 cells were obtained from ATCC (Manassas, VA), and primary p65 polyclonal antibodies (C-20, C-20 agarose conjugate, C22B4), protein A/G PLUS agarose, and normal goat IgG were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Phosphate buffered saline (PBS), tumor necrosis factor (TNF)-α, and cellular growth media along with its supplements were acquired from Invitrogen (Carlsbad, CA).

3.2.2 Preparation of polymer-oxotitanium Films on Au Wafers (Au-P-oxoTi)

The synthesis of poly(2-hydroxyethyl methacrylate) (PHEMA) brushes on Au-coated wafers (2.4 × 1.1 cm) by atom transfer radical polymerization, and the subsequent reaction with succinic anhydride to introduce carboxylic acid groups were described previously.[27] In a glove bag, such polyacid-modified (polymer thickness of ~50 nm) Au wafers were immersed in a stirred, 10-mL THF solution containing 50 mM titanium(IV) isopropoxide ($Ti(i\text{-}PrO)_4$) for 30 min to form the P-oxoTi hybrid material (Figure 3.1a). The wafer was finally rinsed with THF, dried with nitrogen, and removed from the glove bag.

3.2.3 Characterization of Au-P-oxoTi films

The thicknesses of polymer brushes on Au wafers were determined using a rotating analyzer spectroscopic ellipsometer (J. A. Woollam, M-44), assuming a polymer refractive index of 1.5. Reflectance FT-IR spectra of the modified Au wafers were collected on a Nicolet Magna 560 spectrophotometer with a Pike grazing angle (80˚) accessory.

The amount of Ti on wafer surfaces was quantified by inductively coupled plasma optical emission spectrometry (ICP-OES) as follows. First a Au/aqua regia stock solution was prepared

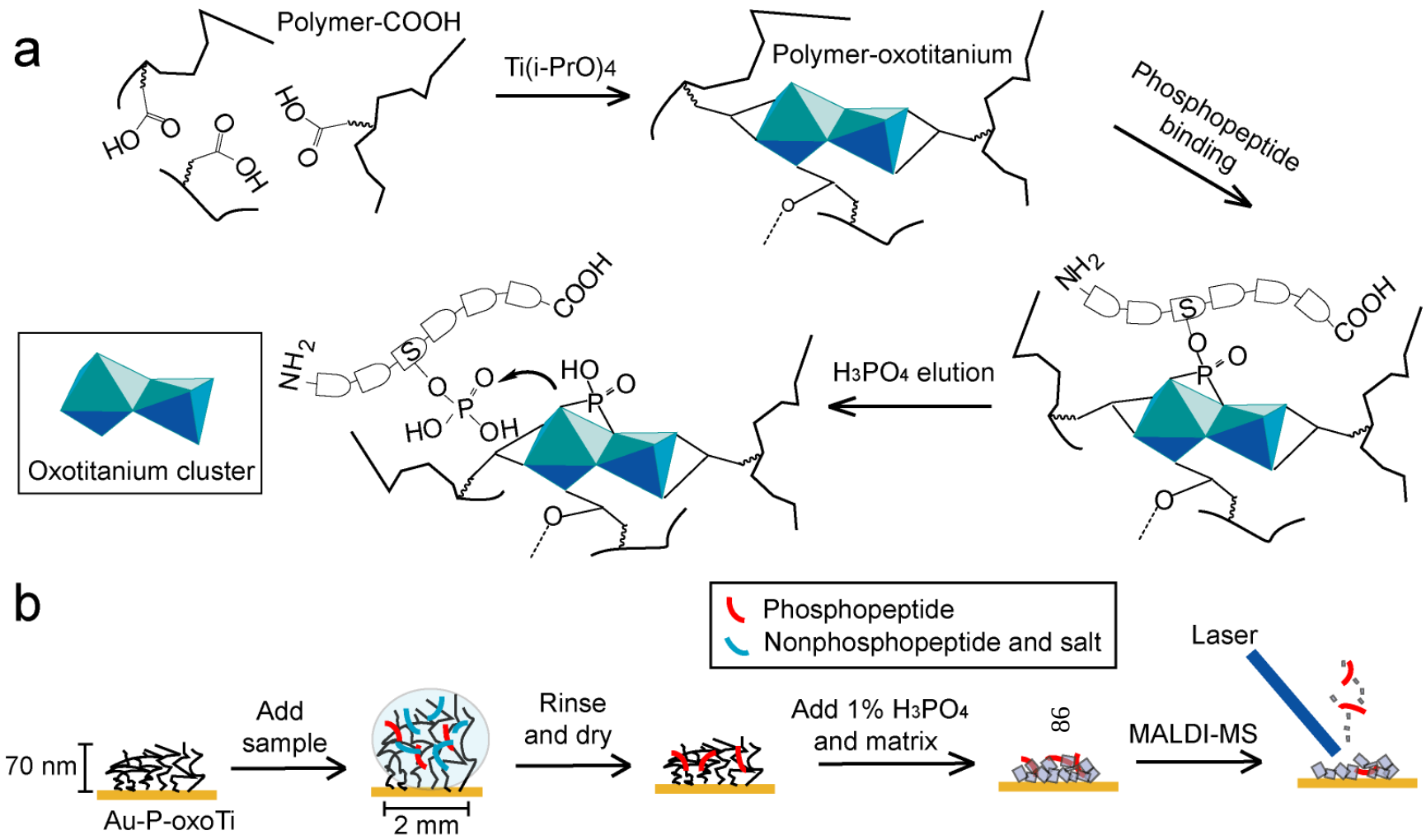


Figure 3.1. (a) Schematic drawing of the formation of the P-oxoTi hybrid structure, and the binding and elution of phosphopeptides in this material. (b) Procedures for on-plate phosphopeptide enrichment using Au-P-oxoTi.

by dissolving Au from 20 pieces of 2.4 × 1.1 cm Au-coated wafers into 80 mL of aqua regia solution. This stock solution was added to some Ti solutions to maintain a constant concentration of Au in all analysis solutions. To test the stability of P-oxoTi films in loading (0.1% TFA or 2% TFA) and elution solutions (1% H_3PO_4), 0.5 mL of the test solution was spread on the entire surface of a Au-P-oxoTi wafer (2.4 × 1.1 cm) for either 30 min (for TFA solution) or 5 min (for 1% H_3PO_4 solution). The solution was then collected, and the wafer was rinsed with 5 mL of water. The combined eluate and rinsing solutions were added to 4 mL of Au/aqua regia stock solution and diluted to 50 mL with deionized water. The remaining Ti on each wafer was removed together with polymer and Au by immersing the entire wafer in 4 mL of pure aqua regia for 5 min. This solution, combined with a 5-mL washing solution, was diluted to 50 mL with deionized water. Triplicate tests were carried out for exposure of the wafers to each type of solution. A series of Ti standard solutions were prepared by adding specific volumes of a Ti ICP standard solution (Sigma) to a 50-mL flask containing 4 mL of Au/aqua regia stock solution, and finally diluting the solution to 50 mL with D.I. water. Using these standards, a Ti calibration curve was obtained using ICP-OES, and the amount of Ti in different eluate solutions was determined using the calibration curve. The total amount of Ti on a single intact Au-P-oxoTi wafer was calculated by averaging the amount of Ti completely removed from 10 Au-P-oxoTi wafers (70±5 nm polymer-oxotitanium film thickness). The amount of Ti per unit area was calculated by dividing the amount of Ti on a single Au-P-oxoTi wafer by 2.64 cm^2 (the area of the 2.4 × 1.1 cm Au wafer).

3.2.4 Immunoprecipitation (IP) of p65 and associated proteins from THP-1 nuclear extracts

This part was performed by Amanda M. Palumbo from Prof. Gavin E. Reid's group. Briefly, nuclei were isolated from THP-1 cells with and without TNF-α treatment (10 ng/mL TNF-α for 30 min at 37 ℃ with 5% CO_2 and ambient oxygen), respectively, by centrifugation. These nuclei were lysed followed by determination of protein concentration using Bradford assay (IgG as standards). Samples for IP were prepared with 1 mg of total nuclear protein in 800 μL of PBS buffer. After preclearance of endogenous IgGs with normal goat IgG and Protein A/G PLUS agarose, the nuclear extracts were immunoprecipitated onto agarose cross-linked with p65 antibody (overnight incubation at 4 ℃). The agarose was centrifuged to discard supernatant, washed, drained and finally stored at -20 ℃ until use. Blank samples (800 μL PBS) were also "immunoprecipitated" as a control. Experimental details can be found in the Ph.D dissertation of Dr. Amanda M. Palumbo.[28]

3.2.5 Preparation of tryptic digests

The digestion of standard proteins followed a literature procedure.[29] Briefly, 100 μg of each phosphoprotein (α-casein, β-casein, and ovalbumin) were dissolved in 20 μL of 6 M urea containing 50 mM Tris-HCl. Five μL of 10 mM DL-dithiothreitol was added to all the protein solutions followed by heating in a water bath at 60 °C for 1 h. Ammonium bicarbonate (160 μL of 50 mM solution) and 10 μL of 100 mM iodoacetamide were added to the protein solutions, which were then placed in the dark for 1 h. Subsequently, 10 μL of 0.5 μg/μL sequencing grade TPCK modified trypsin was added to each solution prior to incubation at 37 ˚C for 16 h. To quench the tryptic digestion, 11 μL of glacial acetic acid was finally added to each solution. As a

relatively large amount of bovine serum albumin (BSA) was used as an excess nonphosphoprotein, TPCK-treated bovine pancreas trypsin with an activity ≥10,000 N_{α}-benzoyl-L-arginine ethyl ester units/mg protein was used for BSA digestion, with the digestion procedures given above. All the digest samples were dried with a Speed Vac (Savant SC110) and dissolved in TFA solutions (0.1% - 2% TFA, depending on sample complexity). The standard protein digests mixtures were prepared by mixing different amount of each protein digest. For enrichment using TiO_2 particles (Titansphere), 50% ACN was included in the sample solutions.

For tryptic digestion of immunoprecipitated proteins, the IP samples (agarose) described above were concentrated to dryness with a Speed Vac, denatured, reduced, alkylated and tryptically digested (using sequencing grade TPCK-modified trypsin) directly on the IP beads (the same procedures as for 100 μg of standard proteins) without protein elution. The digested samples were then centrifuged for 1 min at maximum speed and the supernatant was collected. The residual resin was washed with 20 μL water, centrifuged for 1 min, and the wash solution was collected and combined with the initial supernatant. The combined supernatant was concentrated to near dryness and stored at -20 °C. Ten μL of 0.1% TFA was added before enrichment.

3.2.6 Phosphopeptide enrichment

3.2.6.1 On-plate enrichment using modified wafers

Arrays of 2-mm circles were scratched on the modified wafer surface to form sample wells.[26] Figure 3.1b shows the enrichment procedure. One μL of sample solution was added to each sample well, followed by incubation in a covered, humid petri dish for 30 min. The wafer

was then washed with a solution containing 50% ACN and the same concentration of TFA as in the sample solution, and dried with nitrogen. Phosphopeptides were eluted from the polymer brush by adding 1.5 μL of 1% H_3PO_4 on the sample well and waiting for 5 min. Finally, 0.25 μL of 40 mg/mL DHB (prepared in 1% H_3PO_4 in 50% ACN) was directly added on top of the 1% H_3PO_4 droplet, and the entire solution was allowed to dry. The wafer with phosphopeptide/matrix spots was attached to a modified steel MALDI plate for MS analysis.

To determine the recovery of standard phosphopeptides (Table 3.1) on Au-P-oxoTi films, 0.4 μL of the 10-μL mixture of CH_3CH_2CO-phosphopeptides, in which the concentration of each peptide was 1 μM, was spotted on Au-P-oxoTi. The droplet was immediately mixed with 1 μL of a solution containing specific amounts of BSA digest (0, 1, or 10 pmol) in a TFA solution to give a final sample droplet with the desired concentration of TFA (0.1%, 1%, or 2%) and BSA digest (0, 0.71, or 7.1 μM) on the wafer. After incubation, rinsing, and elution, 0.4 μL of the deuterated analog mixture (CD_3CD_2CO-phosphopeptides) was added on top of the 1% H_3PO_4 eluate, followed by addition of DHB matrix. MALDI-MS spectra were taken after allowing the peptide/matrix solution to dry. A calibration curve of unenriched CH_3CH_2CO-phosphopeptide/CD_3CD_2CO-phosphopeptide MS signal ratios as a function of their molar ratios was obtained by conventional MALDI-MS on a steel plate. The quantitative analysis of each sample was carried out three times to obtain the average recovery and its standard deviation.

3.2.6.2 Phosphopeptide enrichment using commercial TiO_2 particles

Phosphopeptide enrichment using commercial TiO_2 followed a literature procedure[30] with some modification. Briefly, TiO_2 particles (Titansphere, 15 mg) were dispersed in 300 μL of a solution containing 0.1% TFA and 50% ACN (vortex mixing for 3 min) to form a slurry. Immediately, 60 μL of the slurry was transferred to a tube containing 200 μL of loading solution (0.1% or 2% TFA in 50% ACN). After 3 min of vortex mixing and 5 min of centrifugation with a microcentrifuge, the supernatant was discarded and the TiO_2 particles were redispersed by vortex mixing in 60 μL of loading solution to form the final TiO_2 slurry. Ten μL of this slurry was added to 80-100 μL of sample for enrichment. The TiO_2-sample mixture was treated with vortex mixing for 3 min after every 5-min of incubation, and the total time for enrichment was 30 min. After spinning down the TiO_2 particles, the supernatant was discarded, and 150 μL of loading solution was added to wash the particles (3 min of vortex mixing, 5 min of centrifugation, and discarding the supernatant). This washing step was performed three times.

For phosphopeptide elution from TiO_2 particles, our initial studies showed that an eluent containing 1% H_3PO_4 and 50% ACN results in better matrix crystallization and often higher MS phosphopeptide signals than basic elution using 1.6% NH_4OH in 50% ACN. The decrease in signal with the basic eluent may occur because the residual acidic washing solution partially neutralizes the NH_4OH, especially when the eluent volume is low. In addition, the formation of salts due to the addition of the acidic matrix solution to the basic eluate may also decrease the

MS signals. Thus, phosphopeptides were eluted from TiO_2 by adding 50-75 μL of 1% H_3PO_4 in 50% ACN to the tube followed by 3 min of vortex mixing and 5 min of centrifugation. We used this volume of eluent to keep the concentrations of phosphopeptides at a similar level as in on-plate enrichment, assuming the recovery is 100%. A 1-μL aliquot of the eluate was deposited on the steel MALDI plate prior to addition of matrix (0.25 μL of 40 mg/mL DHB in 1% H_3PO_4, 50% ACN) for MS analysis.

For protein digests containing a high level of nonphosphopeptides (e.g. phosphoprotein/nonphosphoprotein ratio of 1/100), we found that introducing DHB in the sample and washing solutions results in a high loss of phosphopeptides (both the number of phosphopeptides and S/N of phosphopeptides in mass spectra). Instead, simply increasing the TFA concentration helps suppress the binding of nonphosphopeptides on TiO_2 particles.

3.2.7 MALDI-MS and CID-MS/MS analysis

For conventional MALDI-MS analysis of protein digests (no enrichment), 1 μL of the digest sample was spotted on the stainless steel MALDI plate, followed by matrix addition (0.25 μL of 40 mg/mL DHB in 1% H_3PO_4, 50% ACN).

Mass spectrometry analysis was performed using an LTQ XL ion trap mass spectrometer equipped with a vMALDI source (Thermo Fisher Scientific, San Jose, CA). Ion trap CID-MS/MS and MS^3 experiments were performed on mass-selected precursor ions. Isolation and activation of the precursor ion was limited to either the monoisotopic ion or the monoisotopic ion and the ^{13}C isotope (the first heavy isotope), to decrease the amount of signal in the product ion

spectrum resulting from activation of extraneous ions with similar m/z. All MS/MS spectra shown are typically the average of 500-3000 scans (3 microscans/scan).

MS/MS spectra were transformed to DTA files by Bioworks 3.3 (Thermo Fisher Scientific), and searched using the Mascot algorithm (Matrix Science) against the NCBInr database with the taxonomy as metazoa (animals) for standard proteins and homo sapiens (human) for IP proteins. The allowed mass tolerance was 0.5 Da for both precursor ions and fragment ions, and the peptide charge state was 1+. The experimental mass values were monoisotopic and the instrument was set as ESI-Trap. The enzyme was set as semi-trypsin with the allowance of up to three missed cleavages. Initial variable modifications used included: carbamidomethylation of Cys (+57 Da), oxidation of Met (+16), and phosphorylation of Ser, Thr, and Tyr (+80). All results were verified manually with the MS/MS data. If no result could be manually verified, then other variable modifications were considered for searching, including: sodium adduct (+22), carbamidomethylation of Glu or Asp (+57), carbamylation of N-term or Lys (+43), dethiomethylation of Met (-48), and pyro-glu from Gln (-17). In some instances, peptides could not be identified by Mascot, but their MS/MS spectra showed product ions similar to those from other identified peptides. For these cases, sequences were identified manually.

3.3 Results and discussion

3.3.1 Preparation and characterization of polymer-oxotitanium (P-oxoTi) hybrid materials

The goal of preparing P-oxoTi hybrid films is to provide immobilized oxotitanium clusters that present a high surface area for phosphopeptide enrichment. After forming a

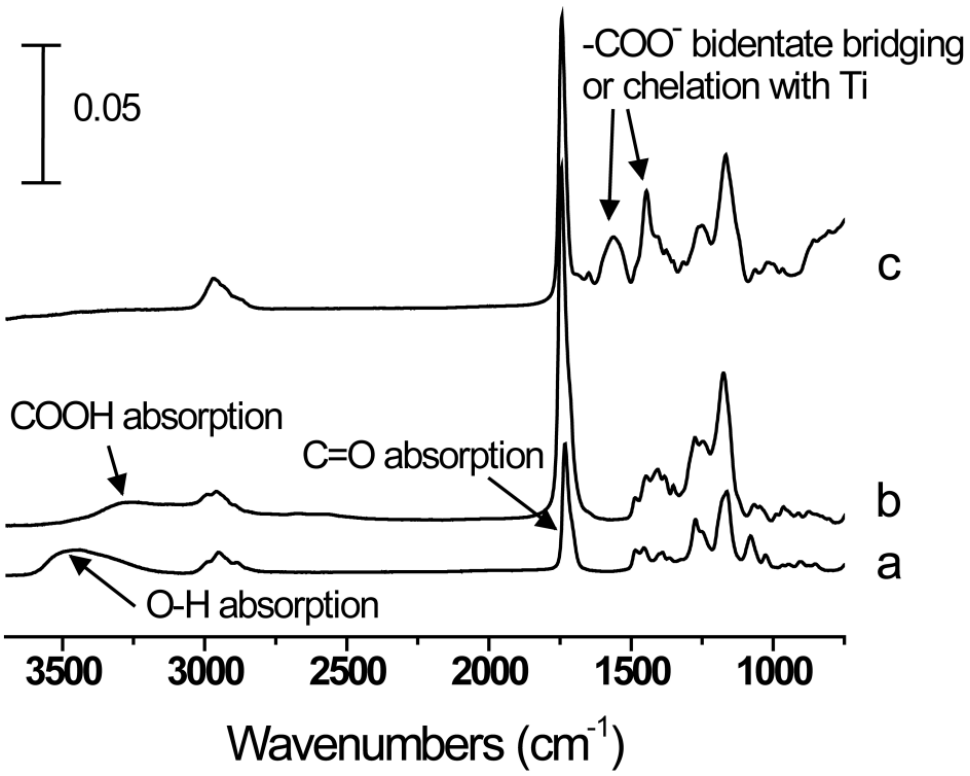


Figure 3.2. Reflectance FT-IR spectra of a PHEMA film before (a) and after reaction with succinic anhydride (b), and subsequent immersion in titanium isopropoxide (c).

polyacid film on the Au-coated Si wafer (confirmed by the COOH absorption at ~3300 cm^{-1} and the enhanced C=O absorption in Figure 3.2b compared to 3.2a), the carboxylic acid containing polymer brushes react with titanium isopropoxide to form oxotitanium clusters (Figure 3.1a). Figure 3.2c confirms the formation of P-oxoTi species. After the reaction, absorption maxima appear at 1560 cm^{-1} (asymmetric stretch) and 1443 cm^{-1} (symmetric stretch) due to the bidentate bridging or bidentate chelation between carboxylate and Ti (Figure 3.2c).[31,32] The small difference between the wave numbers of the two stretching modes rules out the presence of unidentate coordination, which normally gives rise to a much larger peak separation.[33] The broad –COOH absorption near 3300 cm^{-1} also disappears after the reaction with $Ti(i\text{-}PrO)_4$. Elemental analysis by ICP-OES yielded a Ti surface coverage of 1.4±0.2 μg (30±4 nmol) /cm^2

wafer area in a 70 nm-thick P-oxoTi film. Assuming that the molecular weight of the polyacid repeating unit is 230. and the polymer density is 1.0 g/cm^3 in the dry form, the number of polymer repeating units per Ti (or COOH/Ti ratio) is ~0.7.

Several studies report that polymer-oxotitanium network materials can form by polymerization of monomer-capped oxotitanium clusters,[34-36] but the preparation of such clusters typically requires careful control of hydrolysis and several weeks of reaction.[34,37] In our approach, the P-oxoTi structure forms during simple immersion of the polymer-modified wafer in a $Ti(i\text{-}PrO)_4$/THF solution for 30 min (with stirring). The carboxylate groups of the polymer brushes likely replace some of the isopropoxy groups of $Ti(i\text{-}PrO)_4$, and a small amount of water may hydrolyze other isopropoxy groups to promote the formation of oxo clusters.[36] Although the composition of the carboxylate-coordinated oxotitanium in the final hybrid material is still not clear, the molar ratio of polymer repeating units (-COOH groups) to Ti atoms (~0.7) suggests the formation of small carboxylate-containing oxotitanium clusters. Some reported carboxylate-substituted (six-coordinate) titanium polyoxoalkoxy clusters also show a low degree of condensation ($0.33 < O^{2-}/Ti < 1$) along with 2~9 Ti atoms per cluster and carboxylate/Ti < 2.[35] O^{2-}/Ti ratios less than 2 could favor the enrichment of phosphopeptides by presenting many labile Ti species on the cluster surface. Although the Ti precursor, $Ti(i\text{-}PrO)_4$, hydrolyzes rapidly in aqueous solutions, the as-formed polymer-oxotitanium film is stable in a 50 mM EDTA solution at pH 7 (no change in IR spectra after a 30-min immersion).

Because protein digests are usually acidified prior to phosphopeptide enrichment, we also tested the short-term stability of the hybrid films in several acidic solutions. Covering the Au-P-

oxoTi surface for 30 min with 0.1% TFA removes 7±4% of the immobilized Ti. Unfortunately, exposure to 2% TFA for 30 min removes 66±2% of the Ti from the film, showing that there is a risk of losing phosphopeptides under this acidic condition. Coverage of the surface with 1% H_3PO_4 for 5 min (mimicking the elution process) removes 16±3% of the Ti from the Au-P-oxoTi brush. Thus, the majority of the Ti will remain in the polymer brush during elution.

3.3.2 Enrichment of model phosphopeptides on Au-P-oxoTi plates

We first examined phosphopeptide enrichment on Au-P-oxoTi using tryptic digests of model phosphoproteins. This on-plate enrichment requires only 1 μL of sample, and rinsing removes salts and nonphosphopeptides that can otherwise suppress the MS signals of phosphopeptides. In initial trials, basic elution with pH 11 NH_4OH produced MS signals similar to those from elution with 1% H_3PO_4. (Note that the eluate remains on the wafer for subsequent on-plate analysis.) However, the matrix crystallization is inhomogeneous after basic elution, probably because ammonium salts form upon mixing the acidic DHB matrix solution (which contains H_3PO_4 to prevent rebinding of phosphopeptides to the modified wafer after neutralizing the basic eluate) with the NH_4OH eluate. Thus, we chose 1% H_3PO_4 for phosphopeptide elution in all subsequent work.

After Au-P-oxoTi enrichment of a mixture of digested α-casein, β-casein, and ovalbumin (200 fmol of each) in 0.1% TFA, the MALDI mass spectrum of the captured species reveals 24 phosphopeptides and almost no nonphosphopeptides (Figure 3.3b). In contrast, direct MALDI-MS of this mixture without any enrichment yields detectable signals for many nonphosphopeptides and only 5 phosphopeptides (Figure 3.3a). The signal-to-noise ratios (S/N)

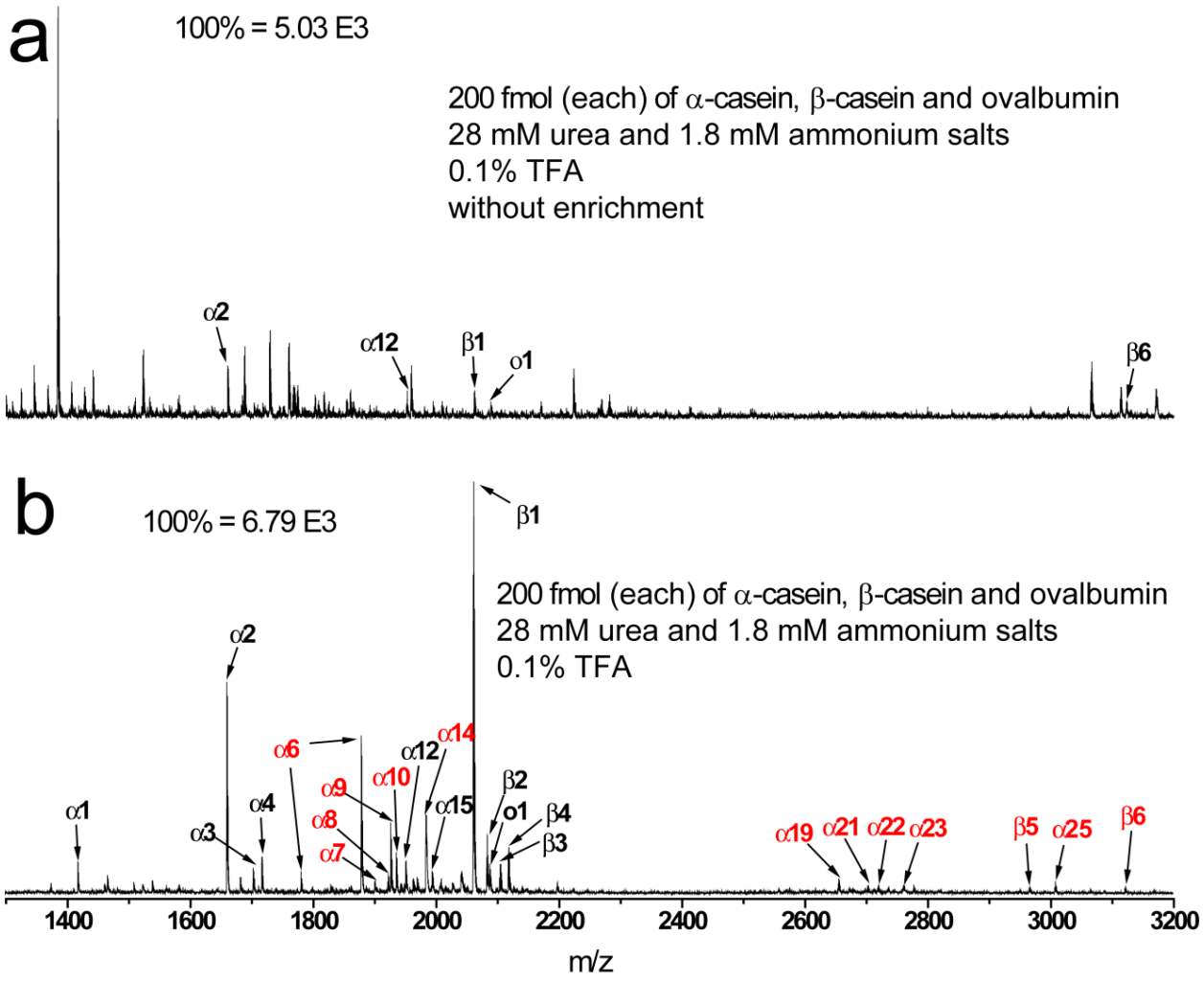


Figure 3.3. MALDI mass spectra of model protein digests without phosphopeptide enrichment (a) and after enrichment on Au-P-oxoTi MALDI plates (b-d). The amounts of digested proteins, and concentrations of urea and ammonium salts in the sample (before enrichment) are given in each spectrum. The TFA concentrations in the sample and washing (b-d only) solutions are also given. Note that for spectrum d, the sample contains a 100-fold molar excess of BSA. MS peaks labeled with α, β, and o represent phosphopeptides generated from α-casein, β-casein and ovalbumin, respectively, and red symbols denote multi-phosphorylated peptides. Table 3.2 presents sequence information for each phosphopeptide. Peptides due to loss of one or more H_3PO_4 units or (H_2O+HPO_3) during/after ionization are labeled together with the intact phosphopeptides using multiple arrows.

Figure 3.3 cont'd

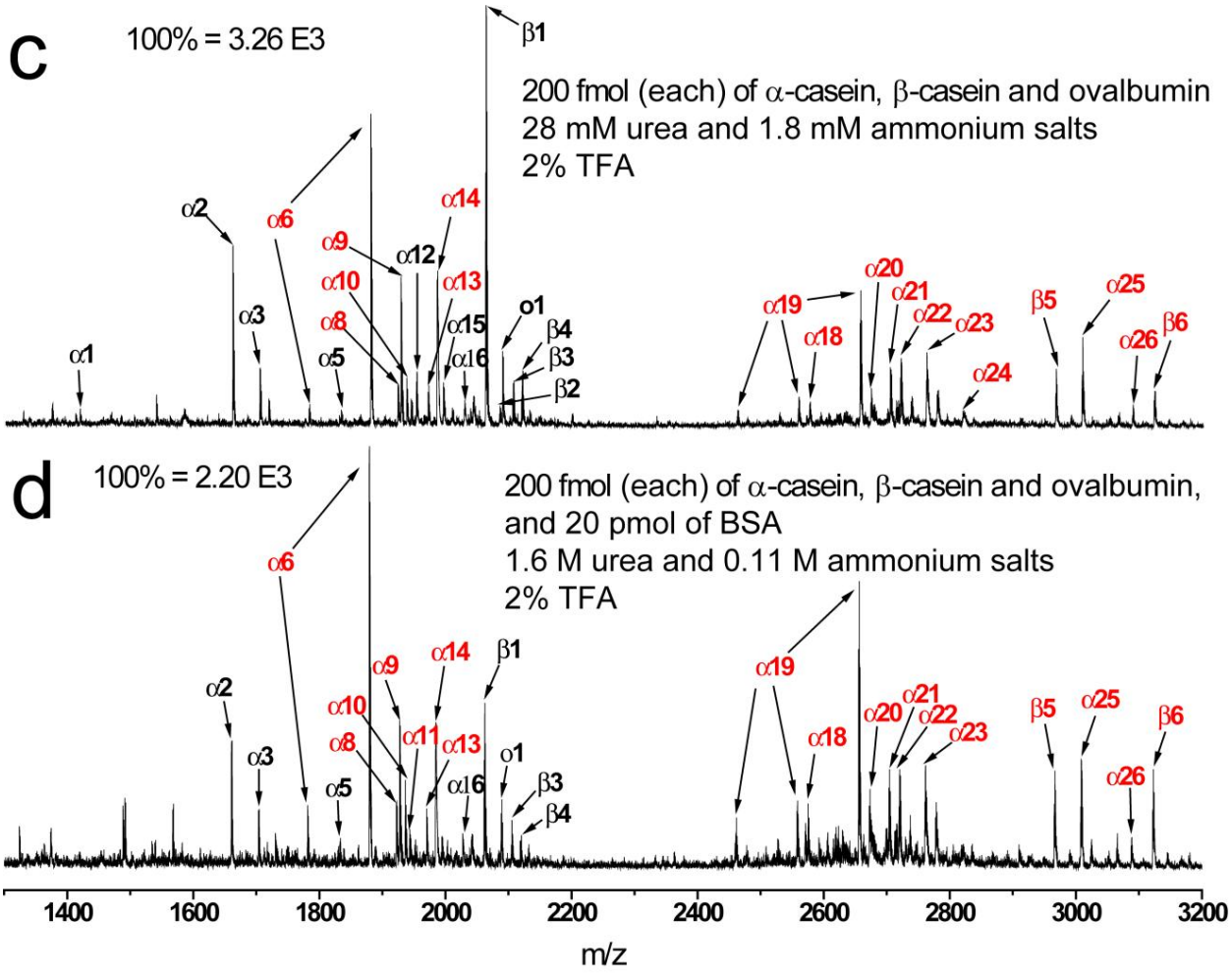


Table 3.2. Phosphopeptides observed in the positive-ion mode MALDI-MS spectra of tryptic α-casein, β-casein, and ovalbumin digests. The phosphopeptide sequence* (if identified), number of phosphoryl groups, and monoisotopic m/z of $[M+H]^+$ are listed.

Protein		Phosphopeptide sequence**	Sequence number	# – PO_3H_2	m/z MH^+
α-casein	α1	K.TVDME**pS**TEVFTK.K M(cam)-105	S2 153-164	1	1418.6
	α2	K.VPQLEIVPN**pS**AEER.L	S1 121-134	1	1660.8
	α3	K._VPQLEIVPN**pS**AEER.L Carbamyl (N-term)	S1 121-134	1	1703.8
	α4	K.VPQLEIVPN**pS**AEER.L E(cam)	S1 121-134	1	1717.8
	α5	YLGEYLIVPN**pS**AEER	S1***	1	1832.8
	α6	K.DIG**pS**E**pS**TEDQAMEDIK.Q M(cam)-105	S1 58-73	2	1879.7
	α7	K.DIG**pS**E**pS**TEDQAMEDIK.Q sodium adduct, M(cam)-105	S1 58-73	2	1901.7
	α8	K._DIG**pS**E**pS**TEDQAMEDIK.Q Carbamyl (N-term), M(cam)-105	S1 58-73	2	1922.7
	α9	K.DIG**pS**E**pS**TEDQAMEDIK.Q	S1 58-73	2	1927.7
	α10	K.DIG**pS**E**pS**TEDQAMEDIK.Q M(cam)-105, D(cam)	S1 58-73	2	1936.7
	α11	K.DIG**pS**E**pS**TEDQAMEDIK.Q M(ox)	S1 58-73	2	1943.7
	α12	K.YKVPQLEIVPNpSAEER.L	S1 119-134	1	1952.0
	α13	K.DIG**pS**E**pS**TEDQAMEDIK.Q Carbamyl (K)	S1 58-73	2	1970.7
	α14	K.DIG**pS**E**pS**TEDQAMEDIK.Q D(cam)	S1 58-73	2	1984.7
	α15	K._YKVPQLEIVPN**pS**AEER.L Carbamyl (N-term)	S1 119-134	1	1995.0
	α16	N/A		1	2027
	α17	N/A		1	2037
	α18	N/A		5	2575
	α19	K.QMEAE**pS**I**pS**p**S**p**S**EEIVPN**pS**VEQK.H M(cam)-105, pyro-glu from Q	S1 74-94	5	2655.9
	α20	N/A		≥2	2672
	α21	K.QMEAE**pS**I**pS**p**S**p**S**EEIVPN**pS**VEQK.H Pyro-glu from Q	S1 74-94	5	2702.9
	α22	K.QMEAE**pS**I**pS**p**S**p**S**EEIVPN**pS**VEQK.H	S1 74-94	5	2720.9
	α23	N/A		4	2760
	α24	N/A		≥3	2819
	α25	R.NANEEEYSIG**pS**p**S**p**S**EE**pS**AEVATEEVK.I	S2 61-85	4	3008.0
	α26	N/A	S2 61-85****	5	3088.0

Table 3.2 cont'd

βcasein	β1	K.FQ**pS**EEQQQTEDELQDK.I	48-63	1	2061.8
	β2	K.FQ**pS**EEQQQTEDELQDK.I Sodium adduct	48-63	1	2083.8
	β3	K._FQ**pS**EEQQQTEDELQDK.I Carbamyl (N-term)	48-63	1	2104.8
	β4	K.FQ**pS**EE̲QQQTEDELQDK.I E(cam)	48-63	1	2118.8
	β5	R.ELEELNVPGEIVE**pS**L**pS**p**S**p**S**EESITR.I	17-40	4	2966.2
	β6	A.RELEELNVPGEIVE**pS**L**pS**p**S**p**S**EESITR.I	16-40	4	3122.3
ovalbumin	o1	R.EVVG**pS**AEAGVDAASVSEEFR.A	341-360	1	2088.9

* Most of these sequences were identified by Mascot, and all of the sequences were manually verified. Some phosphopeptides share the same sequence, but are differently modified during protein denaturing, alkylation, or digestion. To treat all digests equally, even though α-casein (composed of S1 and S2 chains) and β-casein do not contain disulfide linkages, each protein was digested separately using the entire digestion procedure. The overalkylation on Asp, Glu, and Met were reported previously.[38,39]
** Some of the phosphopeptides were not sequenced (labeled as 'N/A') due to their weak signals in the MS full scan. However, their phosphorylation was confirmed according to their dominant losses of 98 Da in tandem MS, and the number of phosphates that can be confirmed is noted.
*** This is a new sequence variant of the α-S1 casein in the region 104–119.[15] The sequence was manually verified.
**** This is probably a penta-phosphorylated peptide.[15]

of the five phosphopeptides also increase 2-20 fold after enrichment. However, with 0.1% TFA in the protein digest and washing solutions, MS signals of multi-phosphorylated peptides (labeled in red) are much weaker than signals of mono-phosphorylated peptides (labeled in black, Figure 3.3b). Increasing the TFA level to 2% in both sample and washing solutions greatly enhances signals of multi-phosphorylated peptides, especially those with high m/z, but the S/N for some mono-phosphorylated peptides decrease by up to 80% (Figure 3.3c). The lower signals of mono-phosphorylated peptides stem from both the weakened binding of phosphopeptides (especially mono-phosphorylated peptides) and the loss of Ti in 2% TFA, whereas the higher signals of multi-phosphorylated peptides likely result from higher ionization efficiency in the presence of fewer mono-phosphorylated peptides.

In biological samples, e.g., whole cell extracts, phosphoproteins are generally much less abundant than nonphosphoproteins. Even after IP, many proteins (including the antibody) are typically present. Although electrophoresis can further fractionate such mixtures, the amount of targeted phosphopeptide obtained after purification will be low due to both the limited amount of sample loaded for gel electrophoresis and incomplete recovery after in-gel digestion. Digestion of targeted phosphoproteins directly on antibody beads can potentially increase the amount of available phosphoprotein, but subsequent enrichment procedures must isolate phosphopeptides from a sample containing more nonphosphopeptides than would be present following electrophoresis and in-gel digestion. Moreover, the salt and urea concentrations in the on-bead digest solution could be high to facilitate digestion of the greater amounts of proteins compared to in-gel digestion.

To examine the effect of a large excess of nonphosphopeptides on enrichment, we captured phosphopeptides from digest mixtures containing 200 fmol each of α-casein, β-casein, and ovalbumin along with 20 pmol of BSA, a nonphosphoprotein. In this case, the presence of 0.1% TFA in the sample and washing solution is not sufficient to prevent capture of nonphosphopeptides on the wafer (Figure 3.4a). Several studies suggest that addition of DHB to sample and washing solutions improves TiO_2-based enrichment specificity.[9,15,16,22] With the P-oxo-Ti, however, the presence of DHB during the enrichment of a β-casein/BSA mixture (molar ratio of 1/10) increased the number of nonphosphopeptide signals in the mass spectrum of the enriched sample (Figures 4b and 4c). This suggests that phosphopeptides and DHB may exhibit different binding modes on P-oxoTi films and TiO_2 particles.

Although the addition of DHB does not enhance enrichment selectivity for phosphopeptides over nonphosphopeptides, increasing the TFA concentration effectively

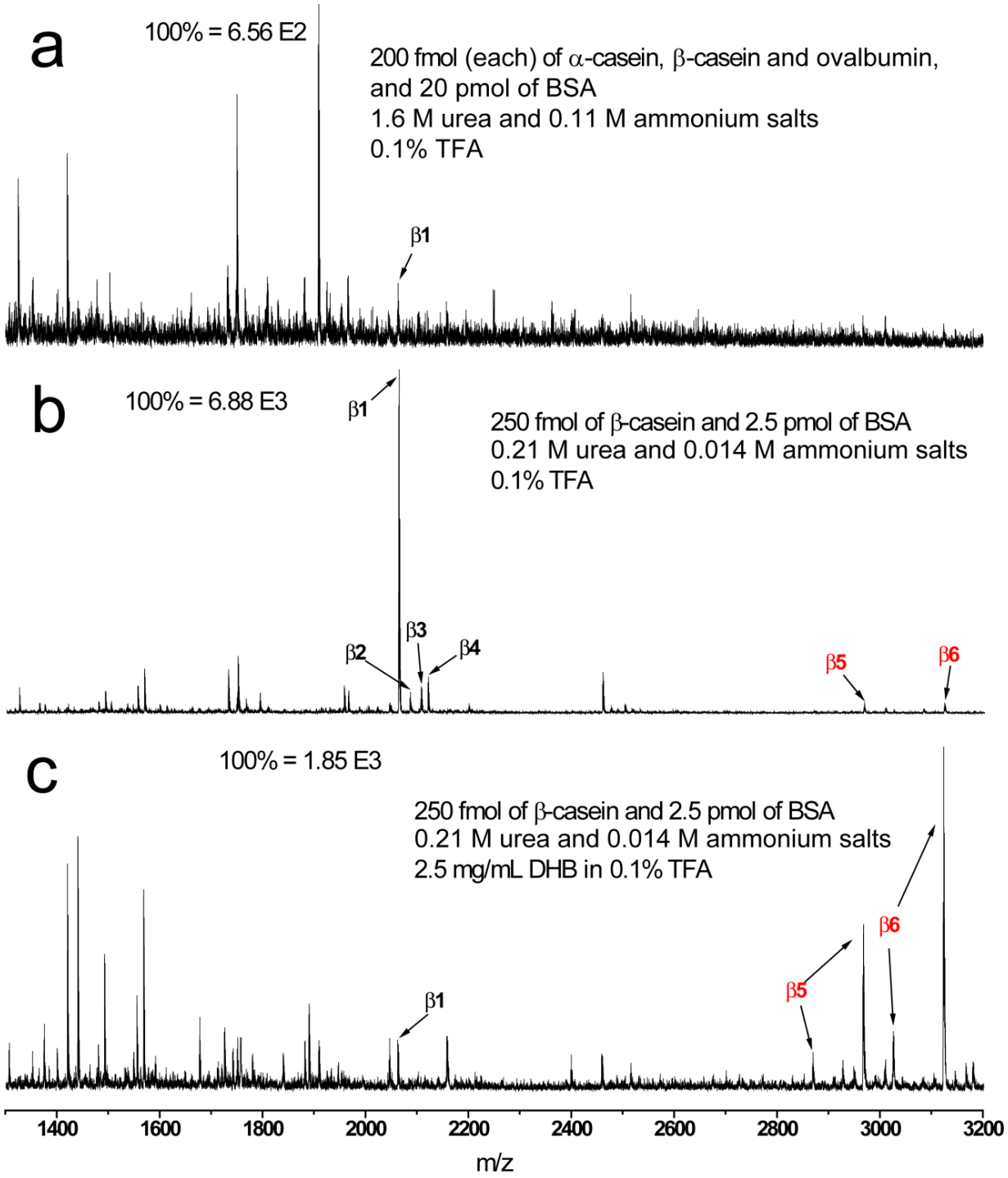


Figure 3.4. MALDI mass spectra of model protein digests that were enriched on Au-P-oxoTi wafers. The amount of protein digests and concentrations of urea, ammonium salts, TFA and DHB (c only) during sample incubation are noted in each spectrum. MS peaks labeled with β represent phosphopeptides generated from β-casein, and red symbols denote multi-phosphorylated peptides. Table 3.2 presents sequence information for each phosphopeptide. Peptides due to loss of one or more H_3PO_4 units or (H_2O+HPO_3) during/after ionization are labeled together with the intact phosphopeptides using multiple arrows.

suppresses the binding of nonphosphopeptides to Au-P-oxoTi. As Figure 3.3d shows, after the enrichment of phosphopeptides from digest mixtures containing 200 fmol of α-casein, β-casein, ovalbumin, 20 pmol of BSA, and 2% TFA, phosphopeptides dominate the mass spectrum, and only a few weak signals of nonphosphopeptides are present. Despite the removal of around 65% of the Ti from the wafer in 2% TFA, signals from 25 phosphopeptides appear in the mass spectrum. Remarkably, even though the 1-μL sample droplet contains only 23 ng of phosphoprotein digests out of 1.4 μg total protein digests, phosphopeptides dominate the mass spectrum. Moreover, the concentrations of urea and ammonium salts in this sample are 1.6 M and 0.11 M, respectively. In conventional MALDI-MS, inhibition of matrix crystallization by such high concentrations of small compounds typically results in minimal peptide signals. After enrichment from this relatively complex solution (Figure 3.3d), the mass spectrum is similar to that after enrichment of the same amount of phosphoprotein digests in the absence of BSA digest (using 2% TFA during incubation, Figure 3.3c). Moreover, for the sample that gives rise to Figure 3c, the amounts of total protein (in mass), urea, and salts are also 60 times lower than when the excess BSA digest is present (Figure 3.3d).

Even in digests containing 6 M urea and 0.4 M ammonium salts, enrichment is possible. After Au-P-oxoTi enrichment, the MS spectrum of a 500-fmol β-casein digest containing 6 M urea, 0.4 M ammonium salts, and 0.1% TFA reveals both mono- and multi-phosphorylated peptides (Figure 3.5a). The limit for detecting at least one phosphopeptide from β-casein digests containing such high concentrations of urea and ammonium salts is around 50 fmol in 1 μL (S/N= 8, Figure 3.5b). However, the same detection limit for β-casein digests containing 14-fold diluted urea and ammonium decreases to about 15 fmol in 1 μL (S/N= 17, Figure 3.5c).

To quantify phosphopeptide capture and elution efficiencies under different sample

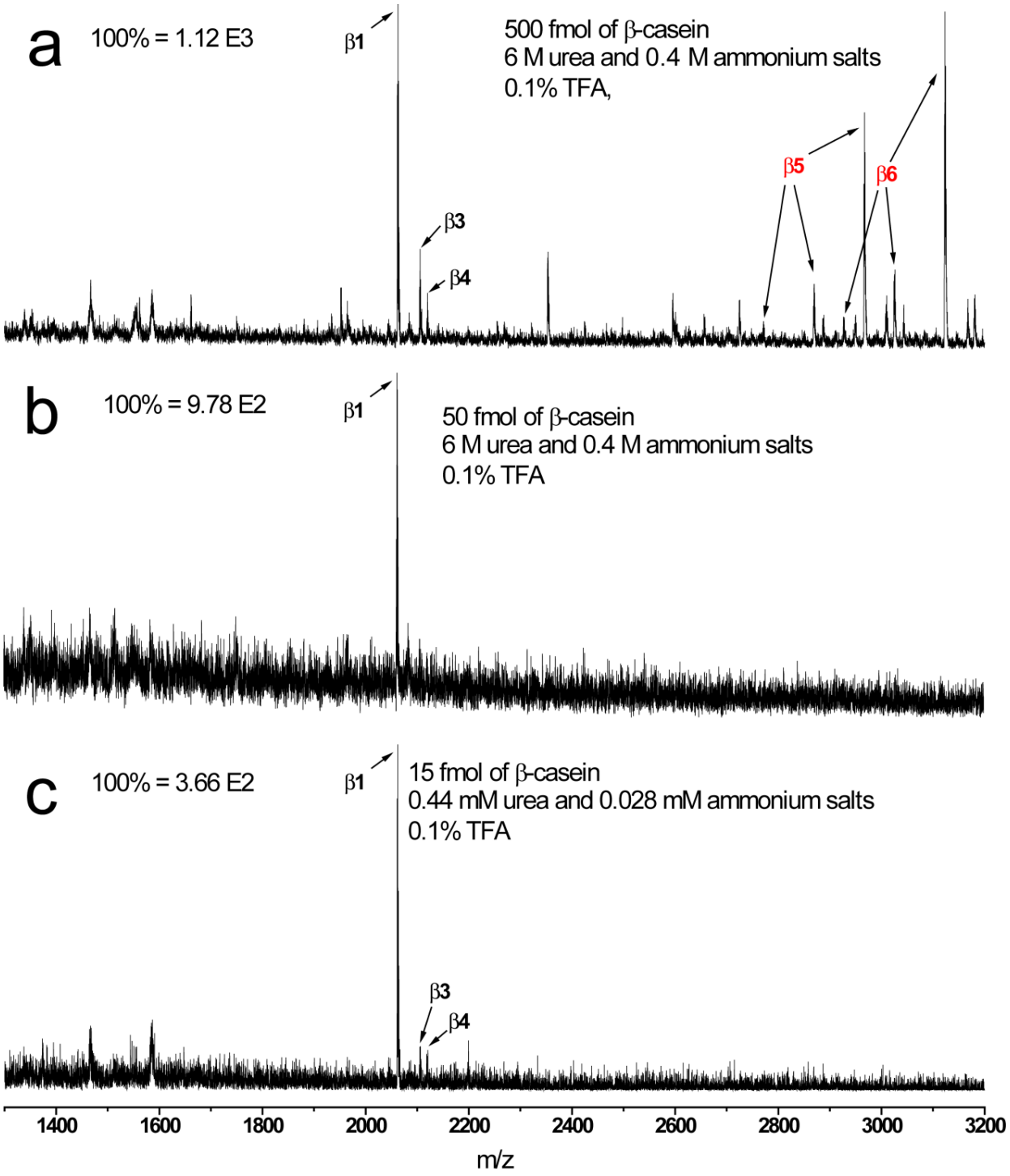


Figure 3.5. MALDI mass spectra of β-casein digests that were enriched on Au-P-oxoTi wafers. The amount of β-casein digest and concentrations of urea, ammonium salts and TFA during sample incubation are noted in each spectrum. MS peaks labeled with β represent phosphopeptides generated from β-casein, and red symbols denote multi-phosphorylated peptides. Table 3.2 presents sequence information for each phosphopeptide. Peptides due to loss of one or more H_3PO_4 units or (H_2O+HPO_3) during/after ionization are labeled together with the intact phosphopeptides using multiple arrows.

complexities and acidities, we enriched a mixture of N-propionylated standard phosphopeptides (CH_3CH_2CO-phosphopeptides, see Table 3.1 for peptide sequences) under several sample conditions, and added these peptides' deuterated analogs, CD_3CD_2CO-phosphopeptides, to the elution droplet (prior to matrix addition) as an internal standard. Using a calibration curve, the percent recovery of each CH_3CH_2CO-phosphopeptide was calculated from the ratio of the MS signals of the nondeuterated and deuterated phosphopeptides. Figure 3.6 shows that in a simple solution of 0.1% TFA, the on-plate enrichment yields a recovery of 32-95% for the seven phosphopeptides. Adding 1 pmol of BSA digest to the phosphopeptide mixture results in a 10:1 mass ratio of total nonphosphopeptides/total phosphopeptides in the sample and decreases recovery by no more than 1/3. One mono-phosphorylated peptide and one multi-phosphorylated peptide show an unexpected increase in recovery upon the addition of BSA digest. In 1% TFA, all of the peptides show detectable signals, but increasing the TFA concentration to 2% and the level of nonphosphoprotein (BSA) digest to 10 pmol causes the loss of nearly all the signals of three mono-phosphorylated peptides. Compared to mono-phosphorylated peptides, the bi-phosphorylated peptides in Figure 3.6 show much higher recoveries under very acidic conditions or in the presences of a large excess of BSA. We note that, the total amount of phosphopeptides on the 2-mm-diameter enrichment spot is 4 pmol, which may be close to the binding capacity of polymer brushes.[26] Thus after the 30-min incubation in 0.1 or 1% TFA, the multi-phosphorylated peptides may occupy the majority of affinity sites of P-oxoTi and decrease the binding of mono-phosphorylated peptides. A similar phenomenon also occurred on the IMAC-Fe polymer micro-spot as shown in Chapter 1. Future studies should examine recoveries as a function of the amount of phosphopeptides in the sample.

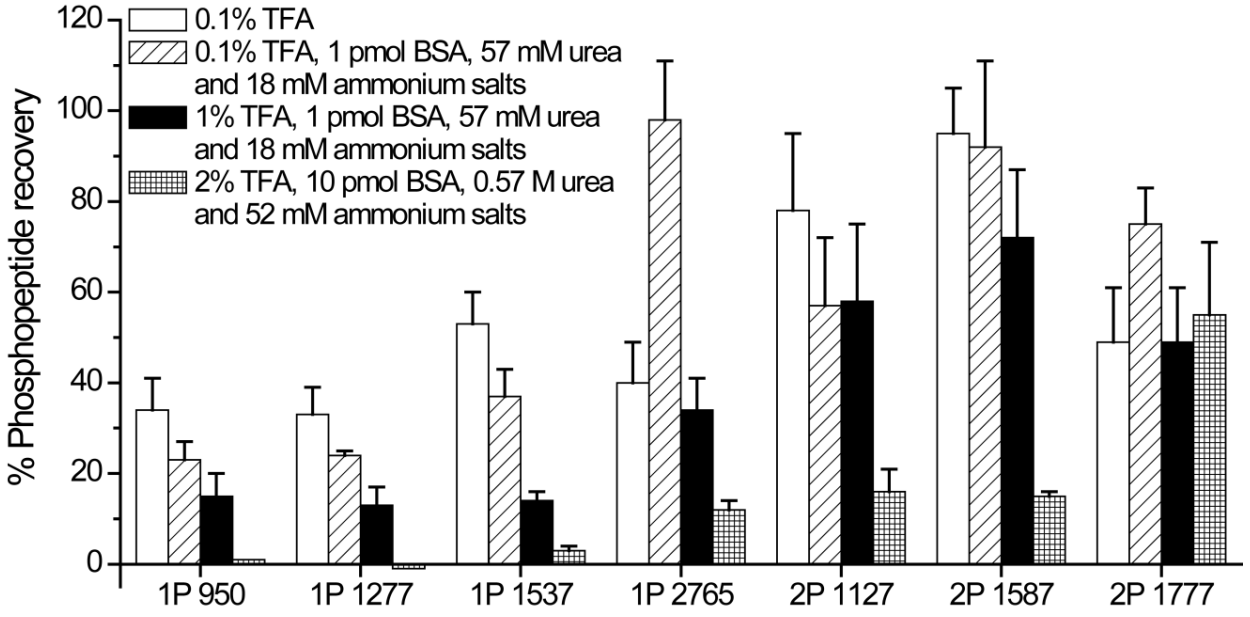


Figure 3.6. Recoveries of CH_3CH_2CO-phosphopeptides from solutions with different TFA concentrations and various amounts of BSA digest, urea, and ammonium salts. The number of phosphates and the MH^+ m/z value (mono-isotopic) of each CH_3CH_2CO-phosphopeptide are noted below the columns. The sample contained 400 fmol of each CH_3CH_2CO-phosphopeptide. For the enrichment of samples containing BSA digest, nonphosphopeptide signals were not detected in the mass spectra. Table 3.1 give phosphopeptide sequences.

As the above results show, the main disadvantage of increasing the TFA concentration (especially to above 1%) in the sample and washing solutions is the loss of signal from mono-phosphorylated peptides. If the drop in signal in 2% TFA is largely due to removal of oxotitanium clusters from the film, synthesis of polymer brushes that bind the clusters more strongly should increase recoveries. Such brushes could contain phosphate groups or hydroxamic acids, or other hard ligands. Fortunately, 0.1% TFA is sufficient for excluding nonphosphopeptides from a modestly complex protein digest sample (e.g. ~0.1 μg/μL nonphosphopeptides). For unknown protein digests, a good enrichment should provide not only a 'purified' mass spectrum, but also sufficient phosphopeptide signals for MS/MS sequencing.

Thus, for complex samples, we recommend on-plate enrichment with a series of TFA concentrations from 0.1%-2% to optimize conditions for specific peptides. Because on-plate enrichment requires only 1 μL of sample and the enrichment process can occur in parallel for a series of samples, this procedure will not greatly increase the analysis time, and the required amount of sample will still be relatively small.

3.3.3 Comparison with other enrichment methods

IMAC is probably the most widely used technique for isolating phosphopeptides, so we compared enrichment on the Au-P-oxoTi wafer with a previously developed IMAC on-plate enrichment technique that uses gold surfaces coated with polymer brushes derivatized with nitrilotriacetate-Fe(III) complexes, Au-P-NTA-Fe.[26] Synthesis of this IMAC material starts from the same polymer used for Au-P-oxoTi but ends with functionalization with NTA-Fe(III). A previous study by our group[26] and Chapter 2 showed that Au-P-NTA-Fe can enrich both mono- and multi-phosphorylated peptides from a simple β-casein tryptic digest. When tested with a mixture of three model phosphoprotein digests, however, the IMAC wafer reveals 17 phosphopeptides, but all of them are multiply phosphorylated (Figure 3.7a). In contrast, of the 24 phosphopeptides found on Au-P-oxoTi using the same sample and the same enrichment conditions, 11 are mono-phosphorylated peptides (Figure 3.3b). For a more complex mixture of three phosphoprotein digests (200 fmol of each) and BSA digest (20 pmol), increasing the TFA concentration to 1% in the sample and washing solutions can exclude most of the nonphosphopeptides from the Au-P-NTA-Fe wafer (Figure 3.7b), but nearly all of the detected phosphopeptides are multiply phosphorylated. The number of detected phosphopeptides under these conditions, however, is still fewer than with enrichment on Au-P-oxoTi using 2% TFA (compare Figures 7b and 3d, 19 vs. 25 total phosphopeptides, and 1 vs. 8 mono-phosphorylated

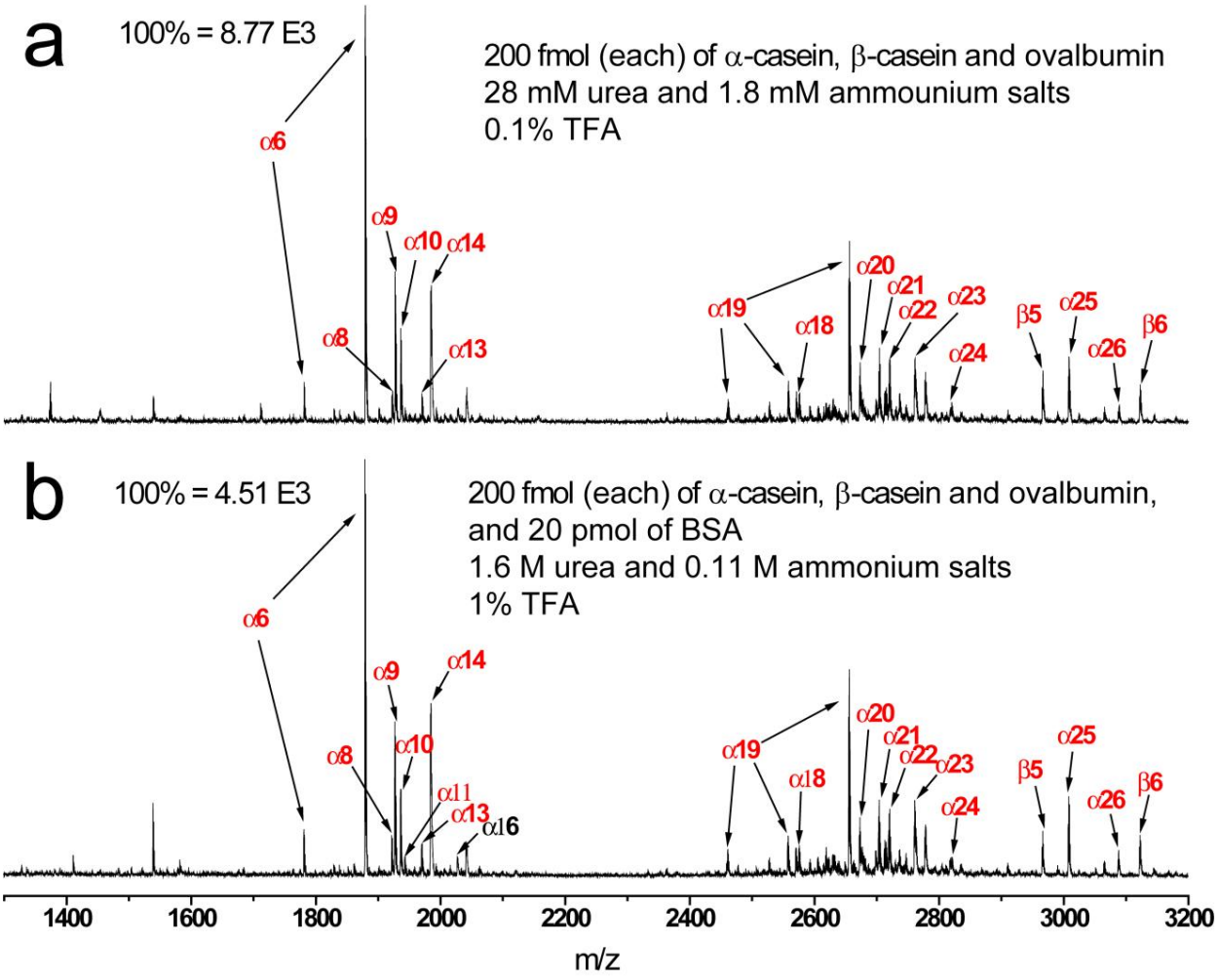
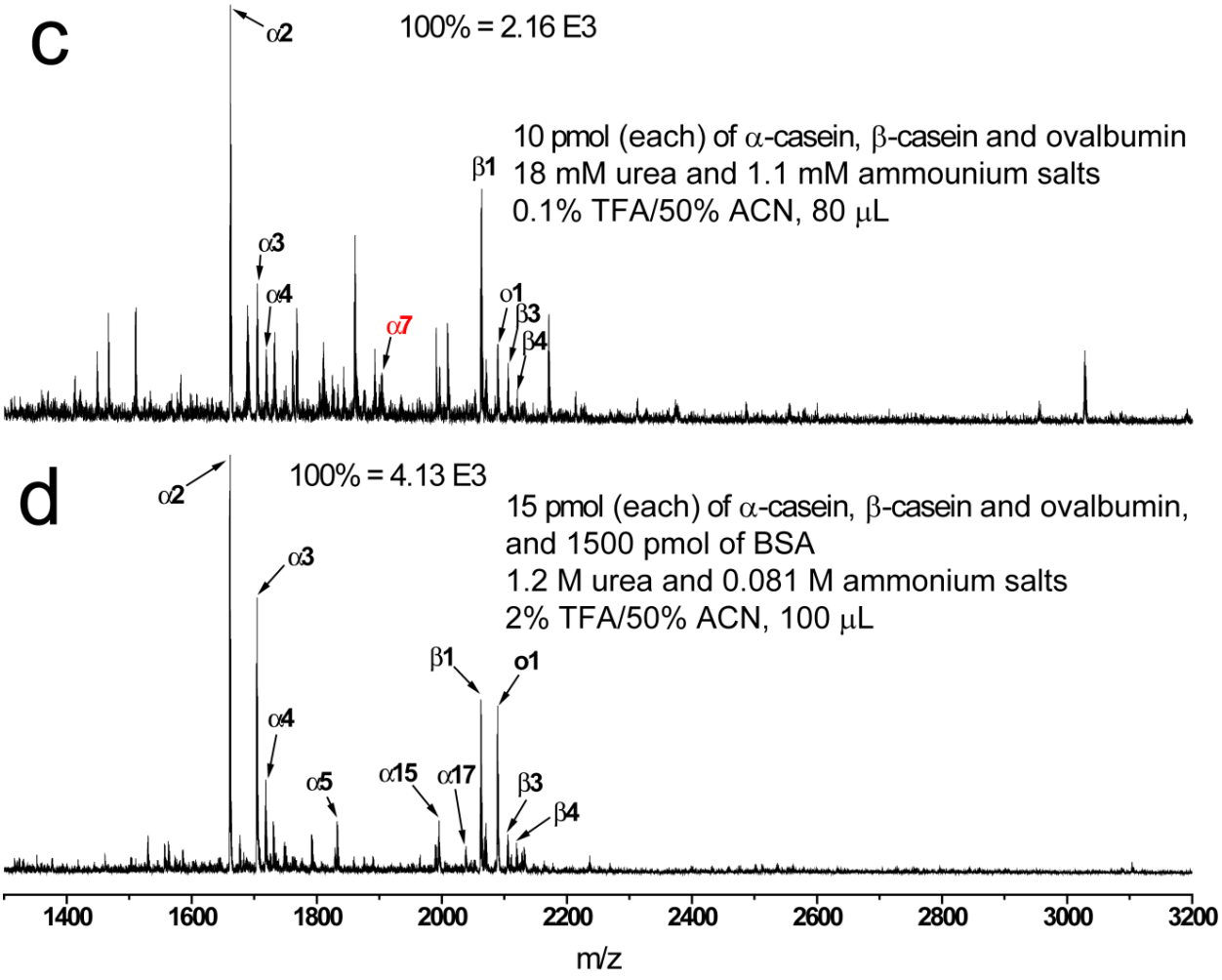


Figure 3.7. MALDI mass spectra of protein digest mixtures after on-plate IMAC enrichment using Au-P-NTA-Fe (a, b) and in-solution enrichment on TiO_2 particles (c, d). The amounts of digested proteins, sample volumes (TiO_2 enrichment only), and the TFA concentrations in sample and washing solutions are noted in each spectrum. In TiO_2 enrichment, a 1 μL aliquot of the 1% H_3PO_4 eluate (out of 50 μL and 75 μL total eluate for c and d, respectively) was spotted on a stainless steel MALDI plate for MS analysis, but the initial solution contained correspondingly more phosphoprotein digests than the on-plate enrichment to compensate for the small fractional amount spotted. MS peaks labeled with α, β, and o represent phosphopeptides generated from α-casein, β-casein and ovalbumin, respectively, and red symbols denote multi-phosphorylated peptides. Table 3.2 presents sequence information for each phosphopeptide. Peptides due to loss of one or more H_3PO_4 units or (H_2O+HPO_3) during/after ionization are labeled together with the intact phosphopeptides using multiple arrows.

Figure 3.7 cont'd



peptides for Au-P-NTA-Fe and Au-P-oxoTi, respectively).

We also tested the enrichment of phosphopeptides with commercial TiO_2 particles (Titansphere, 5 μm particle size). Dispersing the micron-size TiO_2 particles in 80 μL of solution containing 0.1% TFA, 50% ACN, and three digested phosphoproteins (10 pmol of each) followed by washing and elution allows MALDI-MS detection of 7 mono-phosphorylated peptides, 1 multi-phosphorylated peptide, and some non-phosphopeptides (Figure 3.7c). For samples containing higher levels of nonphosphopeptides, increasing the TFA concentration to 2% without using DHB as a non-phosphopeptide excluder (which may cause the loss of phosphopeptides), again, gives mass spectra dominated by phosphopeptides (Figure 3.7d). However, the detected 10 phosphopepitdes are all mono-phosphorylated. Under similar conditions, the Au-P-oxoTi enrichment gives signals from 25 phosphopeptides, 17 of which are multiply phosphorylated (Figure 3.2d).

3.3.4 Analysis of immunoprecipitated p65-associated phosphoproteins.

Building on our success in phosphopeptide enrichment from tryptic digests of standard proteins, we turned to the study of phosphorylation of p65-associated proteins in the THP-1 cell nucleus. THP-1 cells activate the NF-κB pathway upon stimulation with TNF-α,[40] and Western blot analysis of p65-specific immunoprecipitates shows that the TNF-α treatment significantly increases the amount of nuclear p65 (data not shown).

After IP of p65 from the nuclear extract (Western blotting shows that the IP effectively pulls down p65, Figure 3.8), we denatured and digested p65 and associated proteins directly on the antibody beads, and then enriched the phosphopeptides from these digests on Au-P-oxoTi plates. Figure 3.9 shows a representative MALDI mass spectrum of a phosphopeptide-enriched

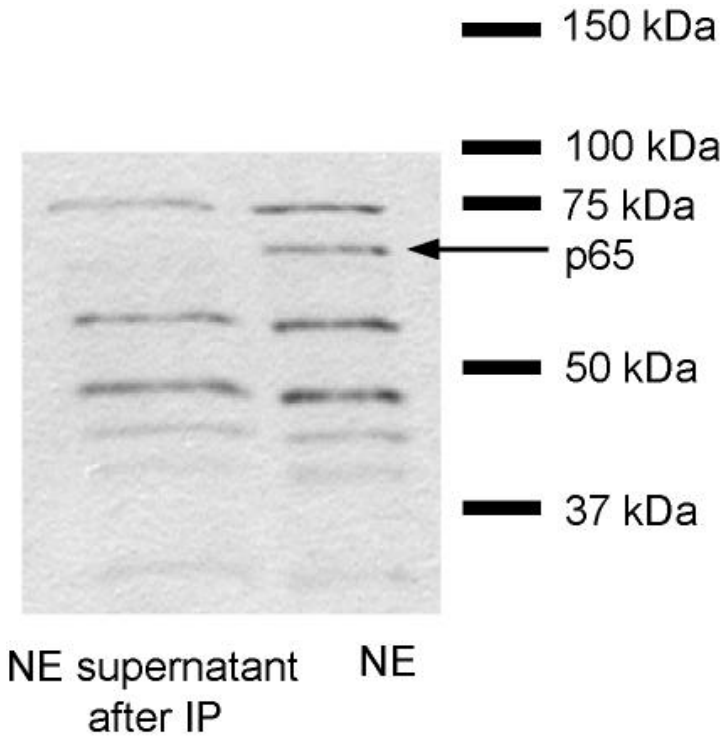


Figure 3.8. p65 Western blotting (using C-22B4 rabbit polyclonal antibody) of 25 μg of THP-1 cell nuclear extract (NE) before (right lane) and after (left lane) IP.

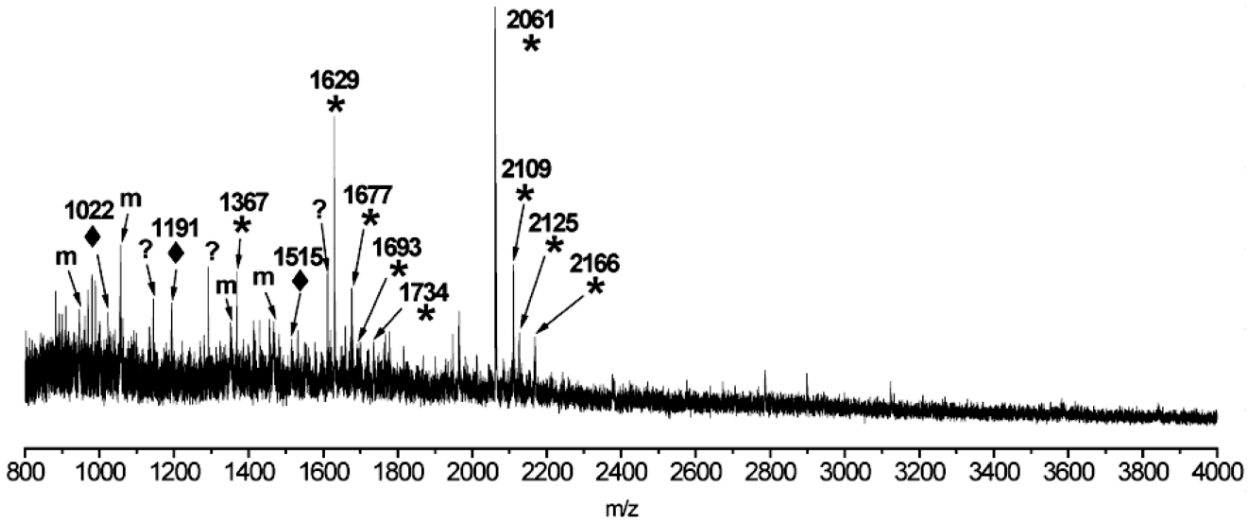


Figure 3.9. MALDI mass spectrum of a Au-P-oxoTi-enriched trypsin-digest of p65-immunoprecipitated nuclear extract from TNF-α-treated THP-1 cells, where * = identified phosphopeptides; ♦ = identified nonphosphopeptides; ? = unidentified peptides; and m = matrix ions. The sample contained 6 M urea, 0.4 M ammonium salts, and 0.1% TFA.

tryptic digest from the immunoprecipitate of a nuclear extract from TNF-α-treated THP-1 cells. Under the 0.1% TFA enrichment condition, nearly all of the signals in the mass spectrum stem from peptides (indicated by CID-MS/MS product ion spectra). Moreover, phosphopeptides (indicated by the predominant neutral loss of 98 Da in CID-MS/MS[41]) dominate the spectrum. The only exceptions include ions corresponding to matrix clusters (labeled as "m"), which are also present in the mass spectra of some standard protein digests (data not shown). Because the nonphosphopeptide signals are fewer and weaker than phosphopeptide signals in the mass spectrum, we did not further increase the TFA concentration in enrichment solutions.

Table 3.3. Peptides identified from the Au-P-oxoTi enriched digests of the nuclear extracts from TNF-α-treated THP-1 cells.*

m/z (mono)	Peptide sequence	Protein	Mascot score
1629	KEEpSEEpSDDDM(cam-105)GF	P1 or P2 isoform of the 60S acidic ribosomal protein	18
1677	KEE**pS**EE**pS**DDDMGF	as above	28
1693	KEE**pS**EE**pS**DDDM(ox)GF	as above	N/A
1734	KEE**pS**EE**pS**DDDM(cam)GF	as above	N/A
2061	KEE**pS**EE**pS**DDDM(cam-	as above	10
2109	KEE**pS**EE**pS**DDDMGFGLFD	as above	11
2125	KEE**pS**EE**pS**DDDM(ox)GFGLFD	as above	N/A
2166	KEE**pS**EE**pS**DDDM(cam)GFGLFD	as above	N/A
1367	**pS**R**pS**FDYNYR	human translocation liposarcoma-associated serine-arginine protein	15
1191	HMYHSLYLK	ribosomal protein L19	36
1022	AAIDWFDGK	translocation liposarcoma protein	39
1515	IWHHTFYNELR	actin (β isoform)	73

* Some of these sequences were identified by Mascot, and all of the sequences were manually verified. M(cam) corresponds to S-carbamidomethylmethionine, M(cam-105) corresponds to the

neutral loss of S-(methylthio)acetamide from M(cam), and M(ox) corresponds to methionine sulfoxide.

Table 3.3 lists the sequences of the peptides that we identified using the information obtained by Figure 3.9, CID-MS/MS product ion spectra (Figure 3.10), guided Mascot database search, and manual spectra interpretation. Among the three identified phosphopeptides (presented in different modified forms in Table 3.2), KEEpSEEpSDDDMGF is a truncated version of KEEpSEEpSDDDMGFGLFD. This truncated peptide may result from chymotrypsin proteolysis during sample preparation, incomplete translation of the C-terminus of the protein, or proteolytic processing occurring in the cells prior to extraction. It should be noted that, all three distinct phosphopeptides (and their modified isoforms resulting from sample preparation) are doubly phosphorylated, and two of those phosphopeptides are highly acidic. Thus these peptides should have especially high affinity for the oxo-Ti clusters. Other phosphopeptides, particularly mono-phosphorylated peptides, might be present in the sample but not enriched. However, under the same salt and TFA conditions, enrichment of phosphopeptides from a β-casein digest revealed both mono- and multi-phosphorylated peptides with similar MS intensities (Figure 3.5a). We also identified several non-phosphopeptides. Capture of these peptides on Au-P-oxoTi likely stemmed from non-specific adsorption, but the corresponding parent proteins are probably specifically associated with p65.

Because the proteins identified in the immunoprecipitate have other known phosphorylation sites, and the phosphorylation of p65 may also occur, we searched for other tryptic phosphopeptides of these proteins using the ProteinProspector MS-Digest Program. However, we could not identify additional phosphopeptides from this IP digest sample. Possible explanations for the absence of other phosphopeptides include insufficient levels of

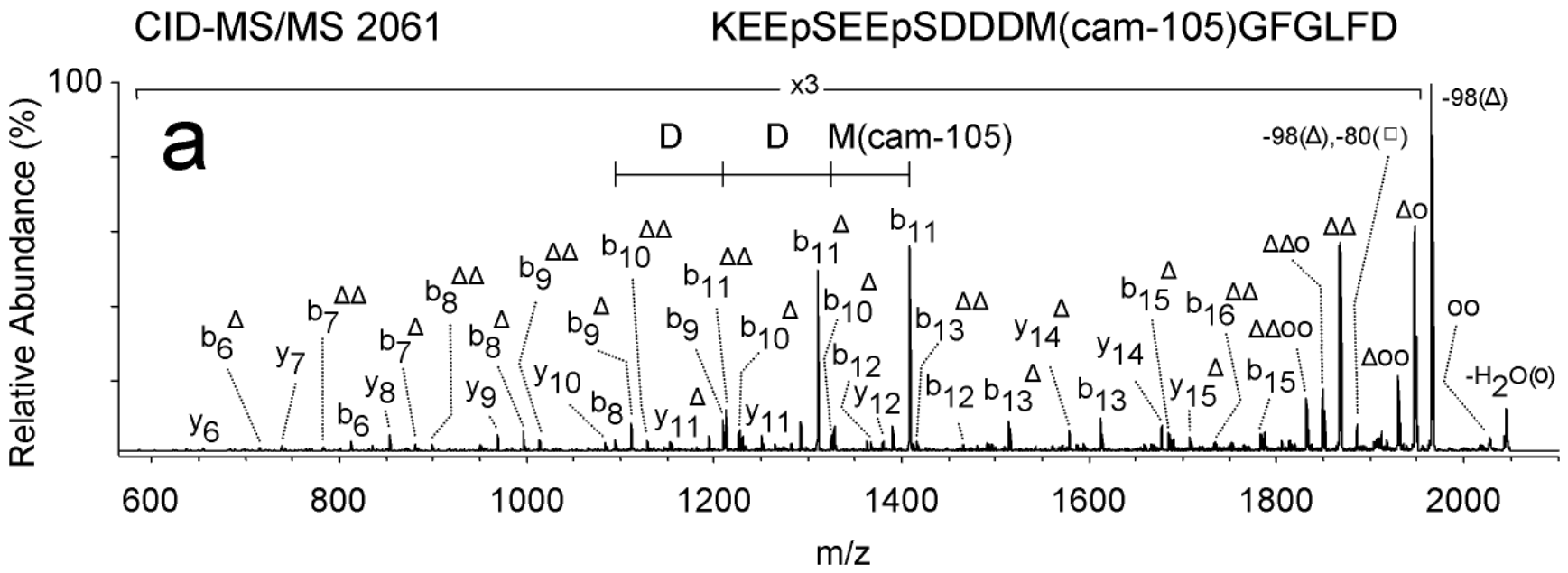


Figure 3.10. CID-MS/MS product ion spectra of m/z 2061 (a), 2109 (b), 2125 (c), 2166 (d), 1629 (e), 1677 (f), 1693 (g), 1734 (h), and 1367 (i) (from the MALDI mass spectrum shown in Figure 3.9). These peptides were identified as various isoforms of KEEpSEEpSDDDMGFGLFD (a-d), KEEpSEEpSDDDMGF (e-h), and pSRpSFDYNYR (i). Δ = –98 Da ($–H_3PO_4$ or $–(H_2O+HPO_3)$); □ = –80 Da ($–HPO_3$); ^{o} = –18 Da ($–H_2O$); ‡ = –64 Da ($–CH_3SOH$ or methane sulfenic acid).

Figure 3.10 cont'd

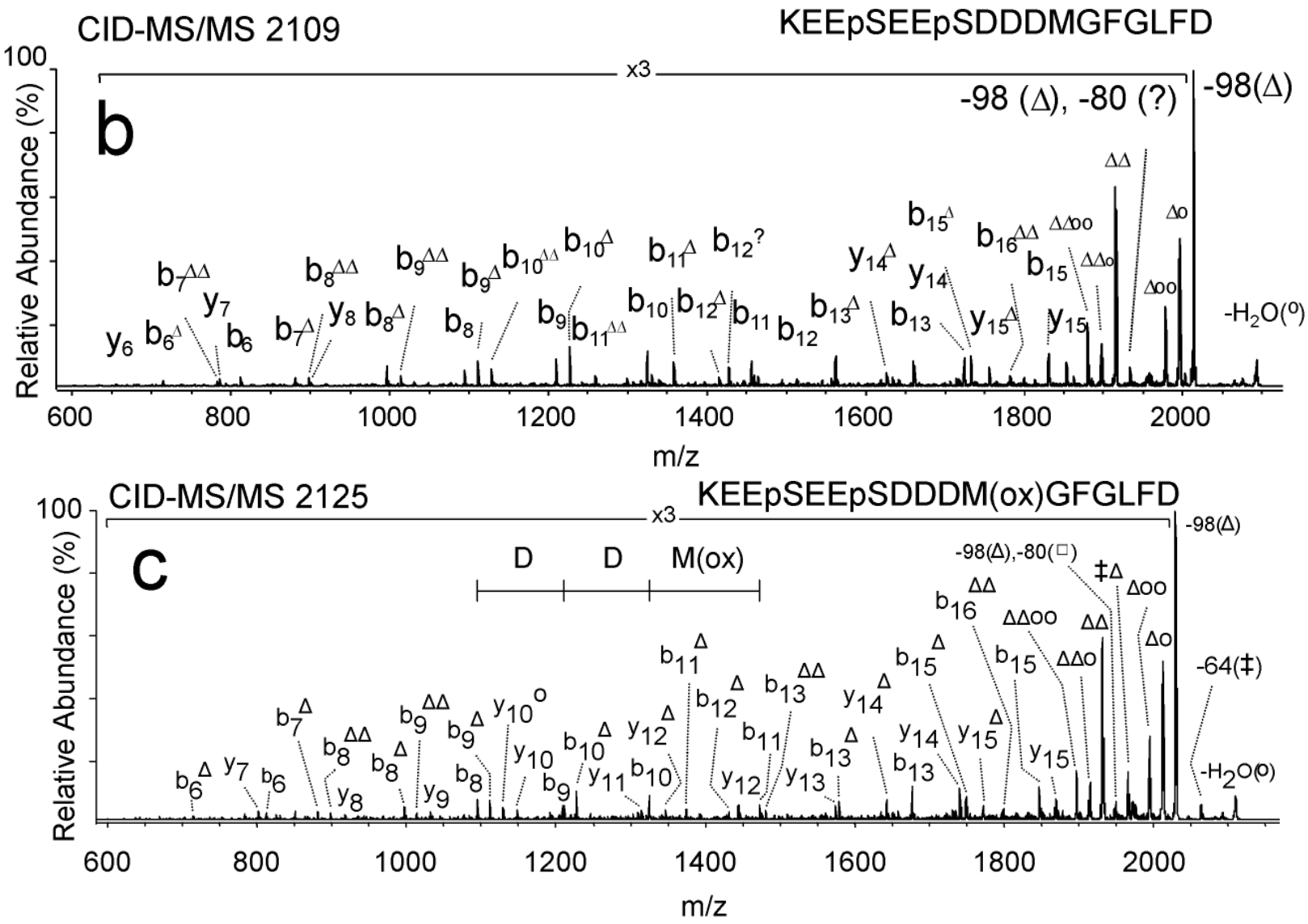


Figure 3.10 cont'd

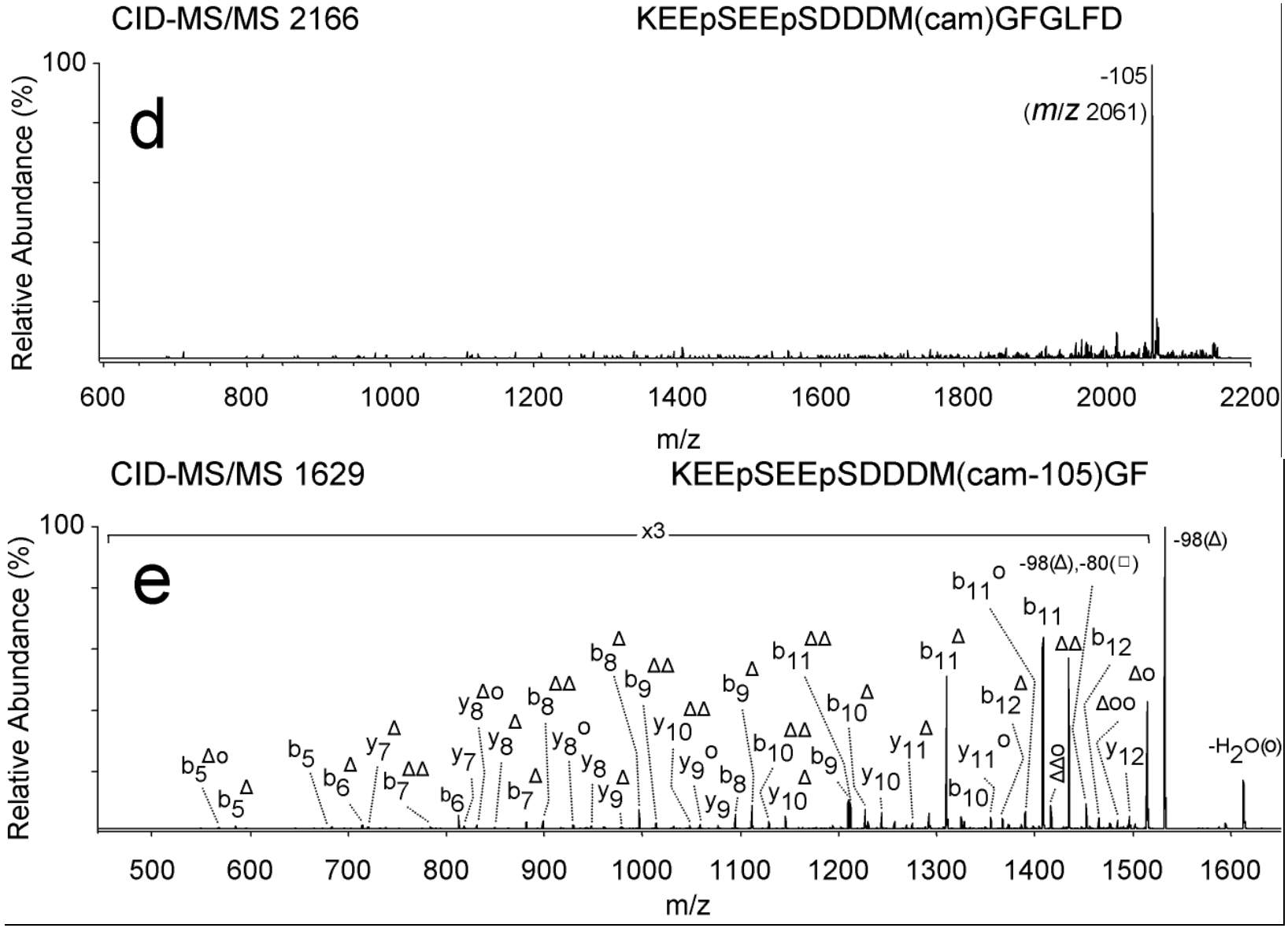


Figure 3.10 cont'd

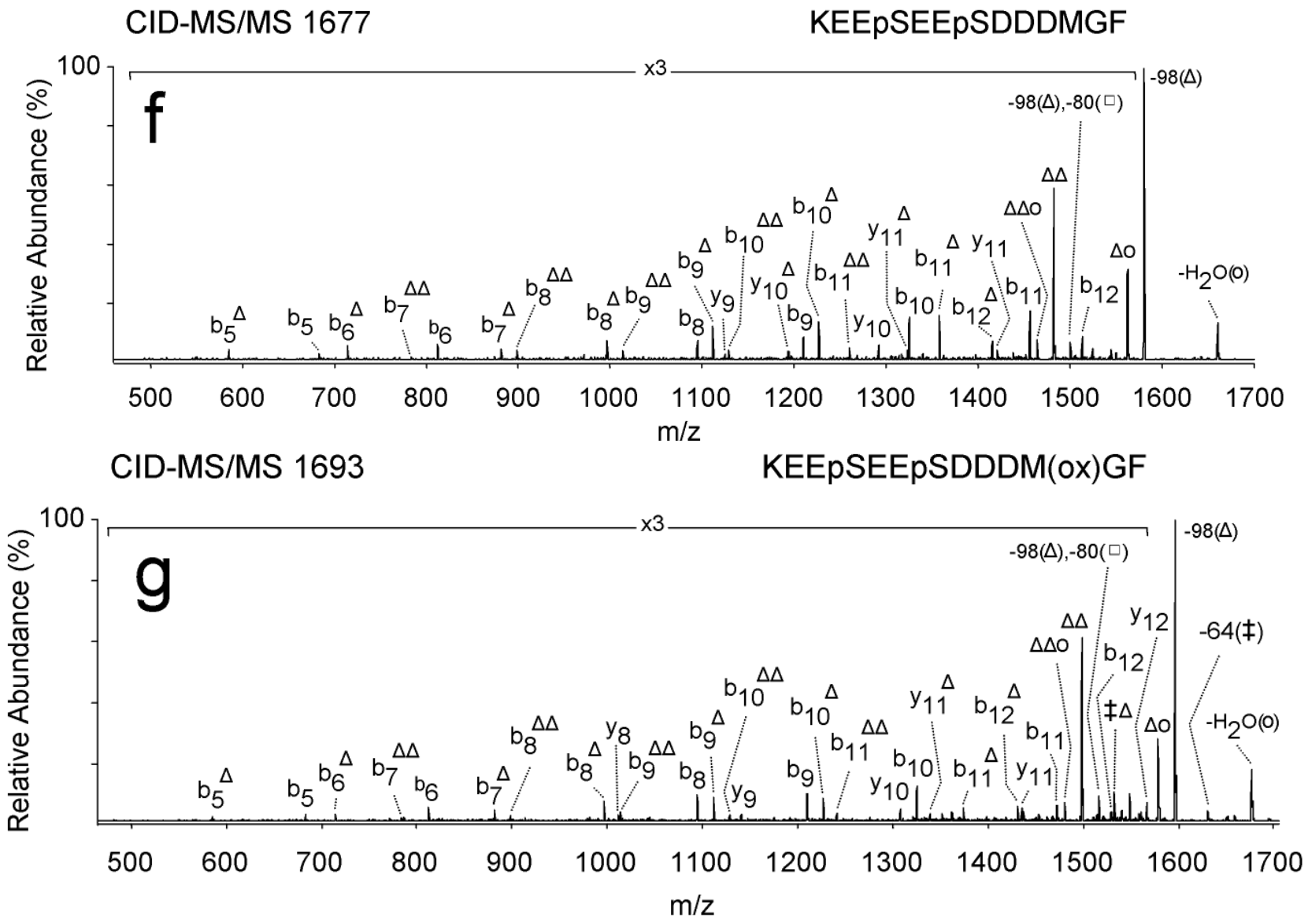
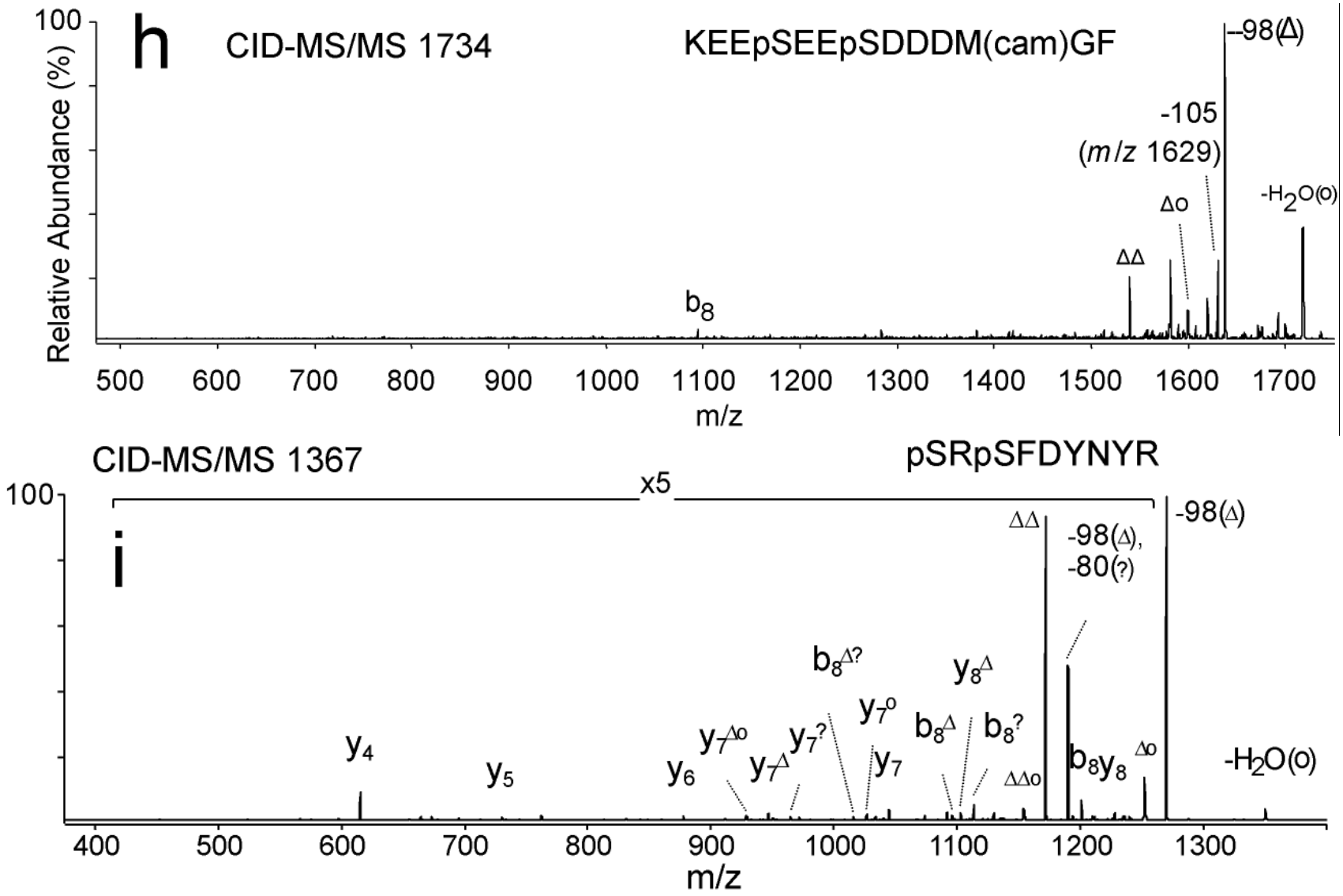


Figure 3.10 cont'd



phosphorylation, ineffective enrichment, elution, or ionization, and masses above or below the user-defined mass range (*m/z* 800-4000). The latter reason likely explains the lack of phosphopeptides from p65 itself. Among the p65 tryptic peptides that contain experimentally verified phosphorylation sites (we generated digested peptides using the ProteinProspector MS-Digest Program with no missed cleavages), only one has a mass between 800-4000 Da. Ryo et al. showed that this peptide is phosphorylated at Thr254 using immunoblotting.[6]

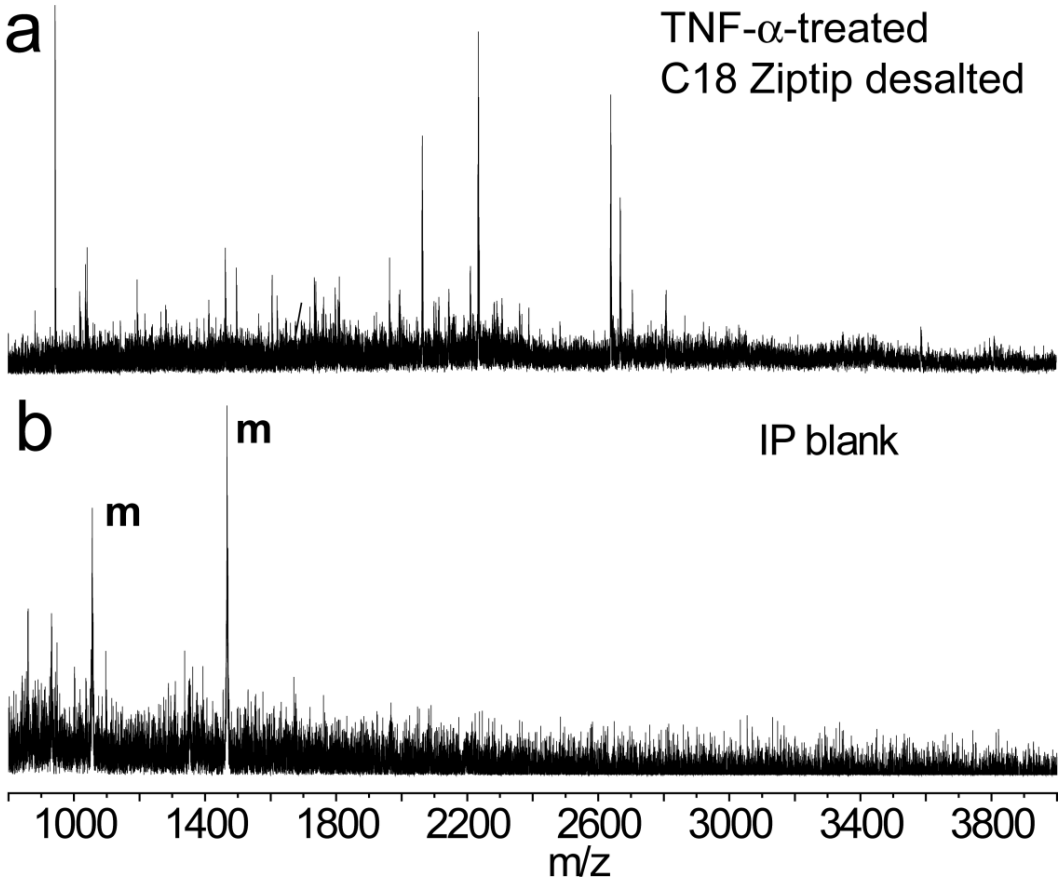


Figure 3.11. MALDI mass spectra of the samples resulting from (a) the trypsin-digested, ZipTipC18-desalted, p65-immunoprecipitated nuclear extracts from TNF-α-treated THP-1 cells (20 μL of the IP digest sample was used for desalting, and 2/5 of the tip eluate was spotted on MALDI steel plate for conventional analysis), and (b) the Au-P-oxoTi enriched trypsin digest of the p65-IP beads exposed to a PBS blank. None of the observable strong signals in (a) and (b) showed dominant neutral loss of 98 upon CID-MS/MS (spectra not shown). Signals due to matrix clusters are noted as "m".

In a control experiment, we desalted the TNF-α-treated IP sample using ZipTipC18 pipette tips and performed conventional MALDI-MS (Figure 3.11a, no phosphopeptide enrichment). As expected, this approach revealed no phosphopeptides. A separate control experiment carried out by performing IP from a PBS blank solution also gave no significant peptide signals in the mass spectrum after on-bead digestion and Au-P-oxoTi enrichment (Figure 3.11b).

3.3.5 Semi-quantitative analysis of p65-associated phosphoproteins upon TNF-α treatment

For comparison with the TNF-α-treated cell sample, we performed the entire sample preparation and purification process with untreated THP-1 cells. Since p65 is present in significantly lower abundance in the nucleus without TNF-α treatment, any protein that has specific interactions with p65 (rather than nonspecific interactions with the p65 antibody) should be present in lower abundance in the nucleus of an untreated sample than in the treated sample, assuming similar enrichment efficiencies. After parallel IP, digestion, on-plate enrichment, and elution of both TNF-α-treated and untreated cells, 100 amol (in 0.5 μL) of an internal standard phosphopeptide was added to each sample spot followed by addition of matrix. Figure 3.12 presents the resulting MS spectra for the two samples and shows that the relative signal intensities (the S/N of phosphopeptide divided by the S/N of internal standard) of the phosphopeptide ions, m/z 2061 and 2109, identified from the TNF-α-treated sample are 13 and 4 times higher, respectively, than in the untreated sample. Thus, the proteins from which these phosphopeptides come are most likely specifically associated with p65 in the cell.

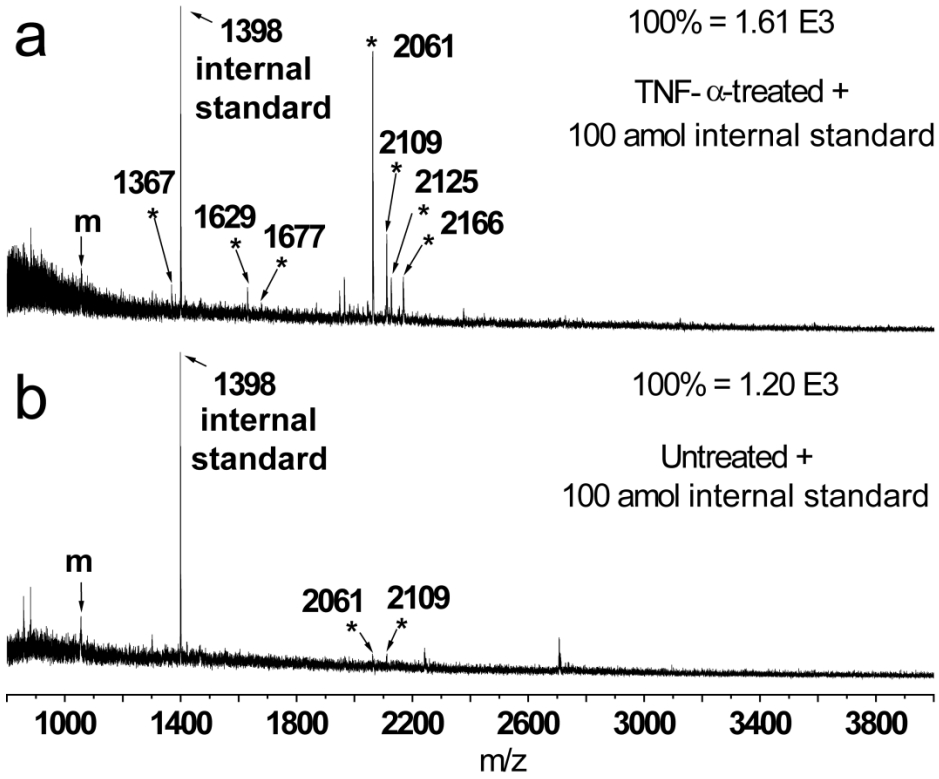


Figure 3.12. MALDI mass spectra of the (a) TNF-α-treated THP-1 cell and (b) untreated THP-1 cell samples resulting from nuclear extraction, p65 IP, trypsin digestion, phosphopeptide enrichment, and addition of 100 amol of an internal standard phosphopeptide (CD_3CD_2CO-LFTGHPEpSLEK), where * = identified phosphopeptides, and m = matrix ions.

3.3.6 Relevance of the identified proteins co-immunoprecipitated with p65

Table 3.3 lists the proteins we identified from the p65 immunoprecipitate. Ribosomal proteins are typically involved in cytoplasmic translational processes, but a number of these proteins have "extra-ribosomal" functions, including regulation of gene transcription, mRNA translation outside of the ribosome, and DNA repair.[42,43] This work identified 60S acidic ribosomal protein from the P1 or P2 isoforms as a phosphoprotein associated with p65. In a previous study, Bouwmeester and coworkers co-purified nine ribosomal proteins and four other proteins involved in ribosomal function with p65.[44] The large subunit (60S) of eukaryotic ribosomes contains the P0, P1, and P2 isoforms of the 60S acidic ribosomal protein (collectively

known as the "P-proteins"), whose sequences are highly conserved across species.[45] Both the P1 and P2 proteins participate in peptide synthesis and interact with elongation factors 1 and 2, which facilitate translational elongation.[46-48] Importantly, the P2 isoform of the 60S acidic ribosomal protein co-immunoprecipitated with key players of the NF-κB pathway, including IκB kinase (γ subunit), p52, NF-κB-inducing kinase (NIK), and RelB, indicating that it too plays a potentially important role in this pathway.[44] However, no reports described the association of the P-proteins directly with p65. Consequently, the association between the P1 or P2 isoform of the 60S acidic ribosomal protein and p65 demonstrated here substantiates a function for these proteins in the NF-κB pathway. RPL19, another identified protein in this study, is part of the eukaryotic 60S ribosomal subunit and its exact function has yet to be elucidated. As far as we are aware, this study provides the first evidence for interaction between RPL19 and p65. Similar to the association between the ribosomal P-protein and p65, this interaction could be cytoplasmic or nuclear.

A recent study shows that human translocation liposarcoma (TLS)-associated serine-arginine protein (TASR) enhances NF-κB mediated transactivation induced by TNF-α, interleukin-1β, and overexpression of NIK by its interaction with the p65 subunit.[49] This is consistent with the TASR and TLS co-immunoprecipitation with nuclear p65 that we observe. TASR serves as a splicing factor, processing pre-mRNA, and interacts with TLS.[50,51] Specifically, the *C*-terminal region (residues 266 to the *C*-terminal residue 526) of TLS links gene transcription and RNA splicing by recruitment of TASR to RNA polymerase II.[50,51] However, when TLS constitutes part of an oncogenic fusion protein, where its *C*-terminal region is replaced with another protein, it cannot recruit TASR to RNA-polymerase II.[50,52] In our study, the TLS peptide constitutes residues 349-357, indicating that TLS is likely present in its

wild-type form. Collectively these results suggest that the TASR protein co-immunoprecipitated with p65 may result from a higher order complex of p65, TLS, TASR, and RNA polymerase II.

The last identified protein is actin, which is a highly conserved major component of the cytoskeleton and plays a critical role in a variety of essential cellular processes.[53,54] Consistent with our results, several studies demonstrated interactions between p65 and actin.[55-57] However, since actin is also a nuclear protein that interacts with transcriptional machinery,[58] and the samples prepared here are a nuclear fraction, the p65-actin interaction proposed in this study may suggest a nuclear role for p65 regulation.

3.4 Conclusion

IP of target proteins from a cell nuclear extract, followed by tryptic digestion directly on antibody beads and enrichment of phosphopeptides using Au-P-oxoTi reveals the association of a number of proteins with p65. The on-plate phosphopeptide enrichment strategy consumes only 1 μL of protein digest for MALDI-MS analysis, and provides signals from both mono- and multi-phosphorylated peptides of complex model protein digests. Overall, the Au-P-oxoTi wafer successfully enriched nine phosphopeptides from trypsin-digested, p65-immunoprecipitated nuclear extracts from TNF-α-treated THP-1 cells. CID-MS/MS enabled the identification of the sequences and possible phosphorylation sites of the nine phosphopeptides along with the sequences of three nonphosphopeptides. The phosphopeptides result from five p65-associated proteins, two of which (RPL19 and the 60S acidic ribosomal protein from the P1 or P2 isoforms) were not previously known to interact with p65. This study also confirms previous results that implicate a role for TASR, TLS, and actin, in the NF-κB pathway. Additionally, the use of

internal standards allows for the comparative analysis of the extent of protein association with p65 in treated and untreated cells. Collectively, these results broaden the perspective on the processes involved in the NF-κB pathway.

REFERENCES

REFERENCES

1. Perkins, N. D., Integrating Cell-Signalling Pathways with Nf-Kappa B and Ikk Function. *Nat. Rev. Mol. Cell Biol.* **2007,** 8, 49-62.

2. Baeuerle, P. A.; Henkel, T., Function Amd Activation of Nf-Kappa-B in the Immune-System. *Annu. Rev. Immunol.* **1994,** 12, 141-179.

3. Wang, C. Y.; Mayo, M. W.; Baldwin, A. S., Tnf- and Cancer Therapy-Induced Apoptosis: Potentiation by Inhibition of Nf-Kappa B. *Science* **1996,** 274, 784-787.

4. Schmitz, M. L.; Mattioli, I.; Buss, H.; Kracht, M., Nf-Kappa B: A Multifaceted Transcription Factor Regulated at Several Levels. *ChemBioChem* **2004,** 5, 1348-1358.

5. Scherer, D. C.; Brockman, J. A.; Chen, Z. J.; Maniatis, T.; Ballard, D. W., Signal-Induced Degradation of Iκbα Requires Site-Specific Ubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* **1995,** 92, 11259-11263.

6. Ryo, A.; Suizu, F.; Yoshida, Y.; Perrem, K.; Liou, Y. C.; Wulf, G.; Rottapel, R.; Yamaoka, S.; Lu, K. P., Regulation of Nf-Kb Signaling by Pin1-Dependent Prolyl Isomerization and Ubiquitin-Mediated Proteolysis of P65/Rela. *Mol. Cell* **2003,** 12, 1413-1426.

7. Yeh, P. Y.; Yeh, K. H.; Chuang, S. E.; Song, Y. C.; Cheng, A. L., Suppression of Mek/Erk Signaling Pathway Enhances Cisplatin-Induced Nf-Kappa B Activation by Protein Phosphatase 4-Mediated Nf-Kappa B P65 Thr Dephosphorylation. *J. Biol. Chem.* **2004,** 279, 26143-26148.

8. Shen, C. H.; Stavnezer, J., Interaction of Stat6 and Nf-Kb: Direct Association and Synergistic Activation of Interleukin-4-Induced Transcription. *Mol. Cell. Biol.* **1998,** 18, 3395-3404.

9. Dunn, J. D.; Reid, G. E.; Bruening, M. L., Techniques for Phosphopeptide Enrichment Prior to Analysis by Mass Spectrometry. *Mass Spectrom. Rev.* **2010,** 29, 29-54.

10. Paradela, A.; Albar, J. P., Advances in the Analysis of Protein Phosphorylation. *J. Proteome Res.* **2008,** 7, 1809-1818.

11. Rogers, L. D.; Foster, L. J., Phosphoproteomics-Finally Fulfilling the Promise? *Mol. Biosyst.* **2009,** 5, 1122-1129.

12. Schmidt, S. R.; Schweikart, F.; Andersson, M. E., Current Methods for Phosphoprotein Isolation and Enrichment. *J. Chromatogr. B* **2007,** 849, 154-162.

13. Cantin, G. T.; Shock, T. R.; Park, S. K.; Madhani, H. D.; Yates, J. R., Optimizing TiO_2-Based Phosphopeptide Enrichment for Automated Multidimensional Liquid Chromatography Coupled to Tandem Mass Spectrometry. *Anal. Chem.* **2007,** 79, 4666-4673.

14. Kyono, Y.; Sugiyama, N.; Imami, K.; Tomita, M.; Ishihama, Y., Successive and Selective Release of Phosphorylated Peptides Captured by Hydroxy Acid-Modified Metal Oxide Chromatography. *J. Proteome Res.* **2008,** 7, 4585-4593.

15. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D., Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Mol. Cell. Proteomics* **2005,** 4, 873-886.

16. Qiao, L.; Roussel, C.; Wan, J.; Yang, P.; Girault, H. H.; Liu, B., Specific on-Plate Enrichment of Phosphorylated Peptides for Direct MALDI-Tof MS Analysis. *J. Proteome Res.* **2007,** 6, 4763-4769.

17. Rainer, M.; Sonderegger, H.; Bakry, R.; Huck, C. W.; Morandell, S.; Huber, L. A.; Gjerde, D. T.; Bonn, G. K., Analysis of Protein Phosphorylation by Monolithic Extraction Columns Based on Poly(Divinylbenzene) Containing Embedded Titanium Dioxide and Zirconium Dioxide Nano-Powders. *Proteomics* **2008,** 8, 4593-4602.

18. Wan, J. J.; Qian, K.; Qiao, L.; Wang, Y. H.; Kong, J. L.; Yang, P. Y.; Liu, B. H.; Yu, C. Z., TiO_2-Modified Macroporous Silica Foams for Advanced Enrichment of Multi-Phosphorylated Peptides. *Chem.-Eur. J.* **2009,** 15, 2504-2508.

19. Zhou, H.; Ye, M.; Dong, J.; Han, G.; Jiang, X.; Wu, R.; Zou, H., Specific Phosphopeptide Enrichment with Immobilized Titanium Ion Affinity Chromatography Adsorbent for Phosphoproteome Analysis. *J. Proteome Res.* **2008,** 7, 3957-3967.

20. Bi, H. Y.; Qiao, L.; Busnel, J. M.; Devaud, V.; Liu, B. H.; Girault, H. H., TiO_2 Printed Aluminum Foil: Single-Use Film for a Laser Desorption/Ionization Target Plate. *Anal. Chem.* **2009,** 81, 1177-1183.

21. Blacken, G. R.; Voln, M.; Diener, M.; Jackson, K. E.; Ranjitkar, P.; Maly, D. J.; Turecek, F., Reactive Landing of Gas-Phase Ions as a Tool for the Fabrication of Metal Oxide Surfaces for in Situ Phosphopeptide Enrichment. *J. Am. Soc. Mass Spectrom.* **2009,** 20, 915-926.

22. Torta, F.; Fusi, M.; Casari, C. S.; Bottani, C. E.; Bachi, A., Titanium Dioxide Coated MALDI Plate for on Target Analysis of Phosphopeptides. *J. Proteome Res.* **2009,** 8, 1932-1942.

23. Bennion, B. J.; Daggett, V., The Molecular Basis for the Chemical Denaturation of Proteins by Urea. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** 100, 5142-5147.

24. Villen, J.; Gygi, S. P., The Scx/IMAC Enrichment Approach for Global Phosphorylation Analysis by Mass Spectrometry. *Nat. Protoc.* **2008,** 3, 1630-1638.

25. Jensen, S. S.; Larsen, M. R., Evaluation of the Impact of Some Experimental Procedures on Different Phosphopeptide Enrichment Techniques. *Rapid Commun. Mass Spectrom.* **2007,** 21, 3635-3645.

26. Dunn, J. D.; Igrisan, E. A.; Palumbo, A. M.; Reid, G. E.; Bruening, M. L., Phosphopeptide Enrichment Using MALDI Plates Modified with High-Capacity Polymer Brushes. *Anal. Chem.* **2008,** 80, 5727-5735.

27. Sun, L.; Dai, J.; Baker, G. L.; Bruening, M. L., High-Capacity, Protein-Binding Membranes Based on Polymer Brushes Grown in Porous Substrates. *Chem. Mater.* **2006,** 18, 4033-4039.

28. Palumbo, A. M. Development and Application of Tandem Mass Spectrometry Methods for Phosphoprotein Analysis (Ph.D Dissertation). Michigan State University, East Lansing, 2009.

29. MacCoss, M. J.; McDonald, W. H.; Saraf, A.; Sadygov, R.; Clark, J. M.; Tasto, J. J.; Gould, K. L.; Wolters, D.; Washburn, M.; Weiss, A.; Clark, J. I.; Yates, J. R., Shotgun Identification of Protein Modifications from Protein Complexes and Lens Tissue. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** 99, 7900-7905.

30. Wilson-Grady, J. T.; Villen, J.; Gygi, S. P., Phosphoproteome Analysis of Fission Yeast. *J. Proteome Res.* **2008,** 7, 1088-1097.

31. Doeuff, S.; Dromzee, Y.; Taulelle, F.; Sanchez, C., Synthesis and Solid-State and Liquid-State Characterization of a Hexameric Cluster of Titanium(Iv): $Ti_6(\mu_2\text{-}O)_2(\mu_3\text{-}O)_2(\mu_2\text{-}Oc_4h_9)_2(Oc_4h_9)_6(Ococh_3)_8$. *Inorg. Chem.* **1989,** 28, 4439-4445.

32. Sui, R. H.; Thangadurai, V.; Berlinguette, C. P., Simple Protocol for Generating TiO_2 Nanofibers in Organic Media. *Chem. Mater.* **2008,** 20, 7022-7030.

33. Deacon, G. B.; Phillips, R. J., Relationships between the Carbon-Oxygen Stretching Frequencies of Carboxylato Complexes and the Type of Carboxylate Coordination. *Coord. Chem. Rev.* **1980,** 33, 227-250.

34. Moraru, B.; Husing, N.; Kickelbick, G.; Schubert, U.; Fratzl, P.; Peterlik, H., Inorganic-Organic Hybrid Polymers by Polymerization of Methacrylate- or Acrylate-Substituted Oxotitanium Clusters with Methyl Methacrylate or Methacrylic Acid. *Chem. Mater.* **2002,** 14, 2732-2740.

35. Rozes, L.; Steunou, N.; Fornasieri, G.; Sanchez, C., Titanium-Oxo Clusters, Versatile Nanobuilding Blocks for the Design of Advanced Hybrid Materials. *Mon. Chem.* **2006,** 137, 501-528.

36. Schubert, U., Chemical Modification of Titanium Alkoxides for Sol-Gel Processing. *J. Mater. Chem*. **2005,** 15, 3701-3715.

37. Kickelbick, G.; Holzinger, D.; Brick, C.; Trimmel, G.; Moons, E., Hybrid Inorganic-Organic Core-Shell Nanoparticles from Surface-Functionalized Titanium, Zirconium, and Vanadium Oxo Clusters. *Chem. Mater*. **2002,** 14, 4382-4389.

38. Boja, E. S.; Fales, H. M., Overalkylation of a Protein Digest with Iodoacetamide. *Anal. Chem*. **2001,** 73, 3576-3582.

39. Lapko, V. N.; Smith, D. L.; Smith, J. B., Identification of an Artifact in the Mass Spectrometry of Proteins Derivatized with Iodoacetamide. *J. Mass Spectrom*. **2000,** 35, 572-575.

40. Lu, Z. G.; Liu, H. S.; Yamaguchi, T.; Miki, Y.; Yoshida, K., Protein Kinase Cδ Activates Rela/P65 and Nuclear Factor-Kb Signaling in Response to Tumor Necrosis Factor-A. *Cancer Res.* **2009,** 69, 5927-5935.

41. Mann, M.; Ong, S.-E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A., Analysis of Protein Phosphorylation Using Mass Spectrometry: Deciphering the Phosphoproteome. *Trends Biotechnol.* **2002,** 20, 261-268.

42. Lindstrom, M. S., Emerging Functions of Ribosomal Proteins in Gene-Specific Transcription and Translation. *Biochem. Biophys. Res. Commun*. **2009,** 379, 167-170.

43. Wool, I. G., Extraribosomal Functions of Ribosomal Proteins. *Trends Biochem. Sci.* **1996,** 21, 164-165.

44. Bouwmeester, T.; Bauch, A.; Ruffner, H.; Angrand, P. O.; Bergamini, G.; Croughton, K.; Cruciat, C.; Eberhard, D.; Gagneur, J.; Ghidelli, S.; Hopf, C.; Huhse, B.; Mangano, R.; Michon, A. M.; Schirle, M.; Schlegl, J.; Schwab, M.; Stein, M. A.; Bauer, A.; Casari, G.; Drewes, G.; Gavin, A. C.; Jackson, D. B.; Joberty, G.; Neubauer, G.; Rick, J.; Kuster, B.; Superti-Furga, G., A Physical and Functional Map of the Human Tnf-A/Nf-Kb Signal Transduction Pathway. *Nat. Cell Biol.* **2004,** 6, 97-105.

45. Rich, B. E.; Steitz, J. A., Human Acidic Ribosomal Phosphoproteins P0, P1, and P2: Analysis of Cdna Clones, in Vitro Synthesis, and Assembly. *Mol. Cell. Biol.* **1987,** 7, 4065-4074.

46. Gomez-Lorenzo, M. G.; Spahn, C. M. T.; Agrawal, R. K.; Grassucci, R. A.; Penczek, P.; Chakraburtty, K.; Ballesta, J. P. G.; Lavandera, J. L.; Garcia-Bustos, J. F.; Frank, J., Three-Dimensional Cryo-Electron Microscopy Localization of Ef2 in the Saccharomyces Cerevisiae 80s Ribosome at 17.5 Å Resolution. *EMBO J.* **2000,** 19, 2710-2718.

47. Macconnell, W. P.; Kaplan, N. O., The Activity of the Acidic Phosphoproteins from the 80 S Rat-Liver Ribosome. *J. Biol. Chem*. **1982,** 257, 5359-5366.

48. Sanchezmadrid, F.; Reyes, R.; Conde, P.; Ballesta, J. P. G., Acidic Ribosomal-Proteins from Eukaryotic Cells - Effect on Ribosomal Functions. *Eur. J. Biochem.* **1979,** 98, 409-416.

49. Uranishi, H.; Tetsuka, T.; Yamashita, M.; Asamitsu, K.; Shimizu, M.; Itoh, M.; Okamoto, T., Involvement of the Pro-Oncoprotein Tls (Translocated in Liposarcoma) in Nuclear Factor-Kb P65-Mediated Transcription as a Coactivator. *J. Biol. Chem.* **2001,** 276, 13395-13401.

50. Yang, L.; Embree, L. J.; Hickstein, D. D., Tls-Erg Leukemia Fusion Protein Inhibits Rna Splicing Mediated by Serine-Arginine Proteins. *Mol. Cell. Biol.* **2000,** 20, 3345-3354.

51. Yang, L.; Embree, L. J.; Tsai, S.; Hickstein, D. D., Oncoprotein Tls Interacts with Serine-Arginine Proteins Involved in Rna Splicing. *J. Biol. Chem.* **1998,** 273, 27761-27764.

52. Yang, L.; Chansky, H. A.; Hickstein, D. D., Ews·Fli-1 Fusion Protein Interacts with Hyperphosphorylated Rna Polymerase Ii and Interferes with Serine-Arginine Protein-Mediated Rna Splicing. *J. Biol. Chem.* **2000,** 275, 37612-37618.

53. Carlier, M. F.; Le Clainche, C.; Wiesner, S.; Pantaloni, D., Actin-Based Motility: From Molecules to Movement. *Bioessays* **2003,** 25, 336-345.

54. Chen, H.; Bernstein, B. W.; Bamburg, J. R., Regulating Actin-Filament Dynamics in Vivo. *Trends Biochem. Sci.* **2000,** 25, 19-23.

55. Are, A. F.; Galkin, V. E.; Pospelova, T. V.; Pinaev, G. P., The P65/Rela Subunit of Nf-Kb Interacts with Actin-Containing Structures. *Exp. Cell Res.* **2000,** 256, 533-544.

56. Fazal, F.; Bijli, K. M.; Minhajuddin, M.; Rein, T.; Finkelstein, J. N.; Rahman, A., Essential Role of Cofilin-1 in Regulating Thrombin-Induced Rela/P65 Nuclear Translocation and Intercellular Adhesion Molecule 1 (Icam-1) Expression in Endothelial Cells. *J. Biol. Chem.* **2009,** 284, 21047-21056.

57. Fazal, F.; Minhajuddin, M.; Bijli, K. M.; McGrath, J. L.; Rahman, A., Evidence for Actin Cytoskeleton-Dependent and -Independent Pathways for Rela/P65 Nuclear Translocation in Endothelial Cells. *J. Biol. Chem.* **2007,** 282, 3940-3950.

58. Zheng, B.; Han, M.; Bernier, M.; Wen, J. K., Nuclear Actin and Actin-Binding Proteins in the Regulation of Transcription and Gene Expression. *FEBS J.* **2009,** 276, 2669-2685.

Chapter 4. Low-bias Enrichment of Mono- and Multi-phosphorylated Peptides Using Bifunctional Polymer Brushes

4.1 Introduction

Although several materials extract phosphopeptides from complex protein digests, these materials recover unique, partially overlapping sets of phosphopeptides, and no single enrichment method is comprehensive enough for phosphoproteome analysis.[1] The two most commmon phosphopeptide enrichment techniques, immobilized metal ion affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC), often show high specificity for multi-phosphorylated and mono-phosphorylated peptides, respectively.[2] Many enrichment strategies attempt to simultaneously capture and elute mono- and multi-phosphorylated peptides using a single affinity material (e.g. TiO_2 or NTA-Fe(III)),[2] but such protocols typically lead to biased recoveries for either mono- or multi-phosphorylated peptides in the following MS analysis.

A few recent studies with MOAC and IMAC show that the binding capacity of the affinity materials also affects the distribution of recovered phosphopeptides observed by MS. For example, although previous studies suggest that TiO_2 preferentially enriches mono-phosphorylated peptides,[3] Li et al. found that a high peptide-to-TiO_2 ratio (insufficient binding capacity) greatly enhances identification of multi-phosphorylated peptides.[4] These authors

recommend pre-examination of samples to find the digest-to-TiO_2 ratio that maximizes phosphopeptide identifications.[4] Al_2O_3 foils with a relatively small loading capacity also specifically capture multi-phosphorylated peptides in on-target phosphopeptide enrichment.[5] Similarly, Moser et al. increased the capacity of IMAC columns to reduce the bias toward detection of multi-phosphorylated peptides to some extent.[6] In Chapter 2, we showed that the limited binding capacity of IMAC-Fe(III) micro-polymer spots leads to preferential enrichment of tetra-phosphopeptides in β-casein digests. Although new elution protocols, e.g. sequential elution of mono- and multi-phosphorylated peptides from IMAC resins,[7] enhance identification of different types of phosphopeptides, the intrinsic properties of an affinity material may still prohibit comprehensive binding of different subsets of phosphopeptides. Enrichment with both IMAC and TiO_2[8] or sequential capture of phosphopeptides with these two types of resins can greatly enhance phosphopeptide identification.[9,10]

The studies in this chapter aim to reduce biases in phosphopeptide enrichment by using bifunctional polymer brushes as the affinity material. Chapter 3 shows that under mild acidic conditions (≤ 1% TFA), Au-P-oxoTi shows high specificity toward mono-phosphorylated peptides, while Au-P-NTA-Fe favors binding of multi-phosphorylated peptides. Considering that the molar ratio of the polymer repeating unit to Ti in Au-P-oxoTi is close to that of polymer repeating unit to Fe on Au-P-NTA-Fe (0.7 vs. 1), we modified poly(glycidyl methacrylate) (PGMA) with equal amounts of NTA-Fe(III) and oxoTi species to simultaneously enrich mono- and multi-phosphorylated peptides directly on a modified MALDI plate. Phosphopeptide enrichment on these bifunctional coatings results in similar signal intensities for mono- and multi-phosphorylated peptides in a single MALDI mass spectrum. For larger scale

phosphopeptide enrichment, we synthesized polymer brushes in the micron-size pores of nylon membrane discs that serve as enrichment media. Stacking two differentially functionalized membrane discs in an HPLC fitting allows sequential capture of both mono- and multi-phosphorylated peptides from samples that contain a high non-phosphoprotein-to-phosphoprotein ratio. A single elution leads to recovery and MS-identification of these phosphopeptides.

4.2 Experimental

4.2.1 Chemicals and materials

Most chemicals and reagents used in this study were obtained from Sigma-Aldrich. Other chemicals and materials include *N*-hydroxysulfosuccinimide sodium salt (sulfo-NHS, ProteoChem), $FeCl_3$ (anhydrous, 98%, Spectrum), dimethylsulfoxide (DMSO, EMD Chemicals), and acetonitrile (ACN, EMD Chemicals). Tetrahydrofuran (THF, Mallinckrodt Baker) was newly distilled before use, and deionized water was obtained from a Millipore ion-exchange system (Milli-Q, 18 MΩ cm). Au-coated Si wafers (200 nm of sputtered Au on 20 nm of sputtered Cr on Si) were obtained from LGA Thin Films. Loprodyne® LP nylon membranes with a 0.45 μm pore size were a gift from PALL Corporation. Materials for protein digestion include: sequencing grade modified porcine trypsin (treated with L-1-tosylamide-2-phenylethyl chloromethyl ketone, Promega), Tris-HCl (Invitrogen), urea (J.T.Baker), DL-dithiothreitol (Fluka), ammonium bicarbonate (Columbus Chemical) and trifluoroacetic acid (TFA, EMD Chemicals).

4.2.2 Synthesis of bifunctionalized polymer brushes on Au wafers

Synthesis of bifunctional polymer brushes (oxoTi-PGMA-Fe) on Au-coated wafers started with self-assembly of a disulfide initiator monolayer during immersion of a cleaned (UV/O_3 for 15 min) wafer in an ethanol solution containing 1 mM $(BrC(CH_3)_2COO(CH_2)_{10}S)_2$ for 24 h and rinsing with ethanol. (Nishotha Anuraj of our research group prepared the disulfide initiator following a literature procedure.[11]) ATRP of glycidyl methacrylate (GMA) on the initiator-modified Au surface followed a modified literature procedure.[12] Typically, a solution containing glycidyl methacrylate (5 mL, 5.21 g, 36.7 mmol), methanol (4 mL), water (1 mL) and 2,2'-dipyridyl (141 mg, 0.904 mmol) was degassed with three freeze-pump-thaw cycles. CuCl (36.4 mg, 0.368 mmol) and $CuBr_2$ (3.9 mg, 0.017 mmol) were added to the frozen solution prior to three additional freeze-pump-thaw cycles. In a N_2-filled glove bag, the initiator-modified Au wafers were immersed in the degassed monomer solution for 1 to 4 h to yield PGMA brushes with different thicknesses on the Au surface. The wafers were finally removed from the monomer/catalyst solution, rinsed with ethanol, and dried with N_2.

Figure 4.1 shows the PGMA functionalization steps, which were accomplished by sequentially immersing the Au-PGMA wafer in the following solutions: (1) 1.5 M NaN_3 in water/DMF (1:1, v/v) at 50 ℃ for 24 h; (2) 0.1 M succinic anhydride (SA) and 0.16 M 4-dimethylaminopyridine (DMAP) in anhydrous DMF at 50 ℃ for 8 h; (3) 0.1 M sulfo-NHS and 0.1 M *N,N'*-diisopropylcarbodiimide (DIC) in DMSO for 12 h; (4) 0.1 M N_α,N_α-bis(carboxymethyl)-L-lysine hydrate (aminobutyl-NTA) in pH 7.4 water/DMSO (1:1, v/v) for 4 h (the solution was prepared by initially adding 0.1 M NaOH until the aminobutyl NTA

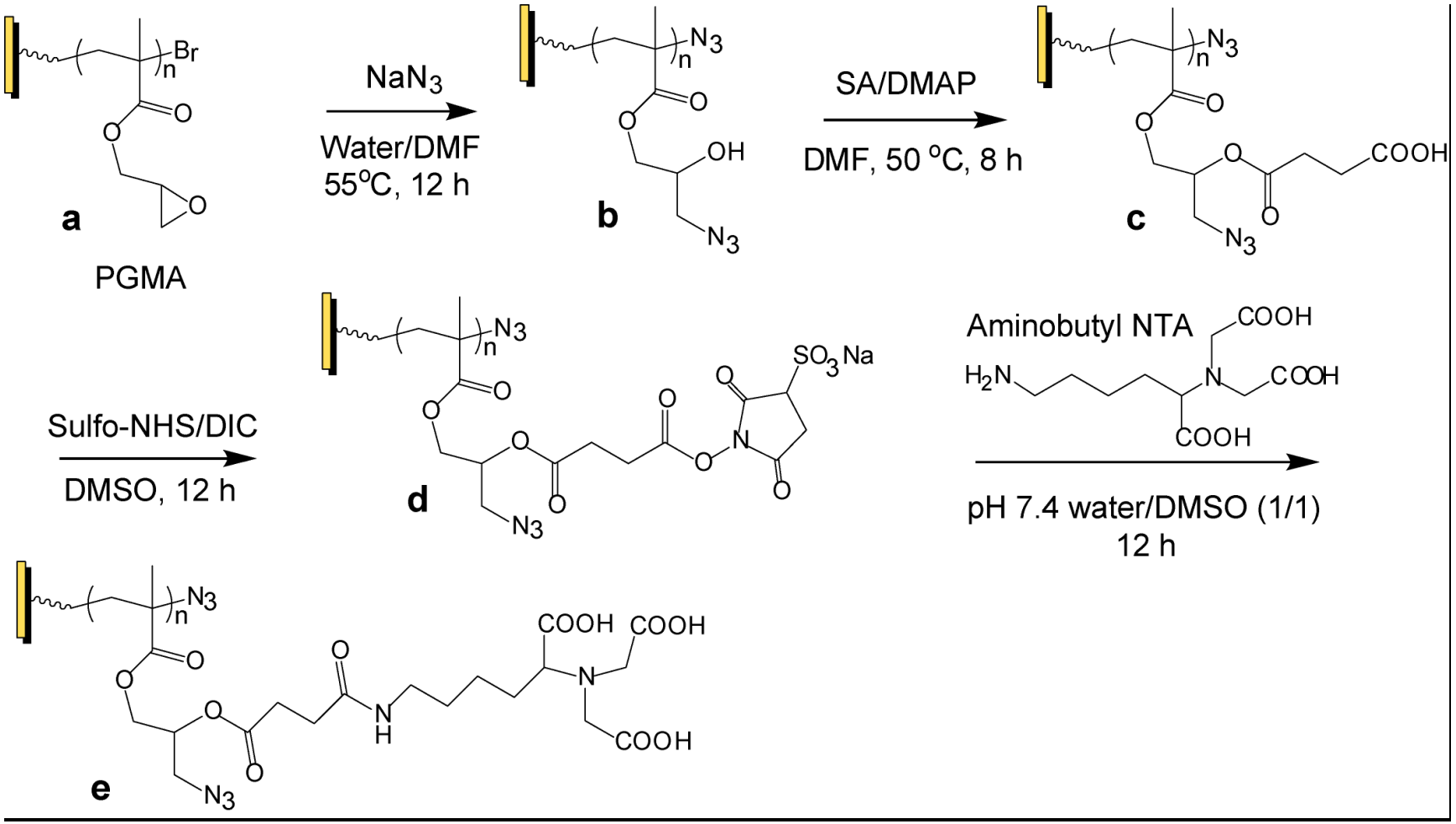
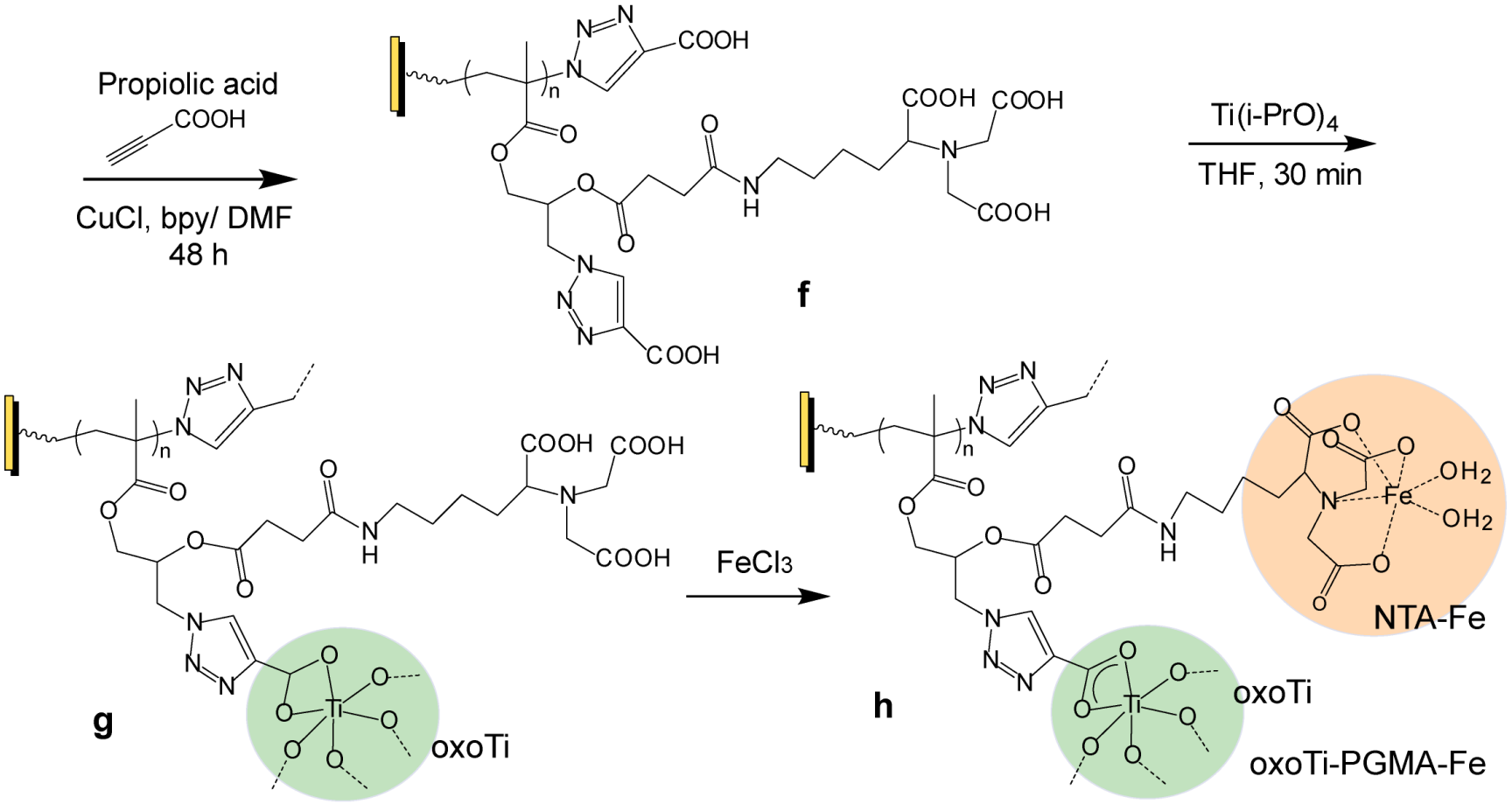
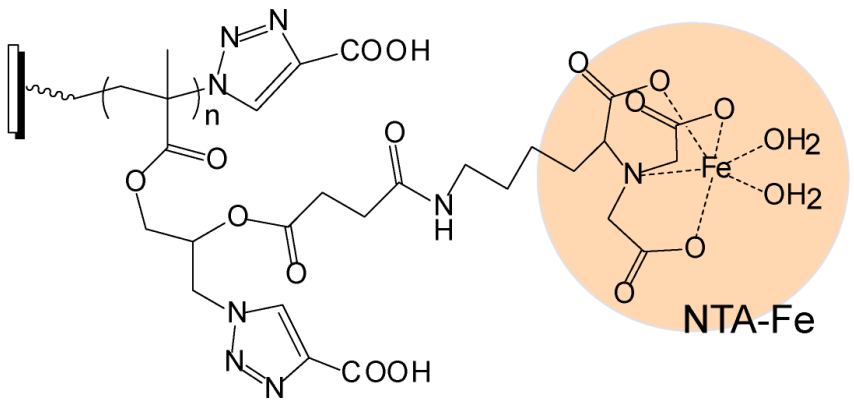


Figure 4.1. Scheme for orthogonal modifications of PGMA brushes to obtain the bifunctional oxoTi-PGMA-Fe.

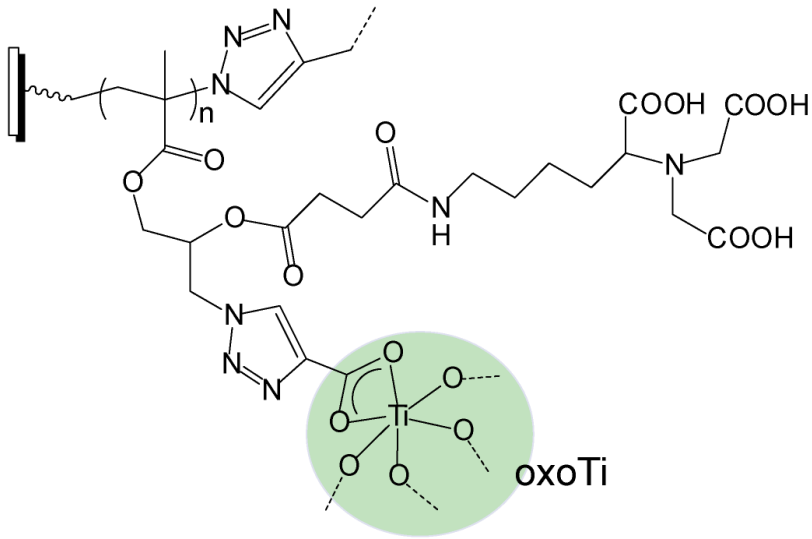
Figure 4.1 cont'd



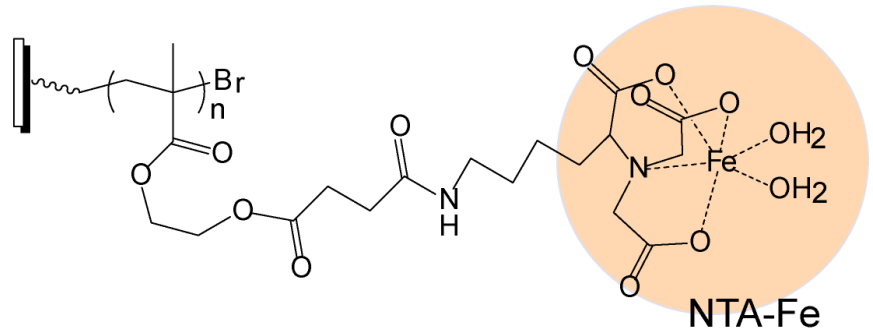
Control polymer 1: PGMA-Fe



Control polymer 2: oxoTi-PGMA



Control polymer 3: PHEMA-Fe



Control polymer 4: oxoTi-PHEMA

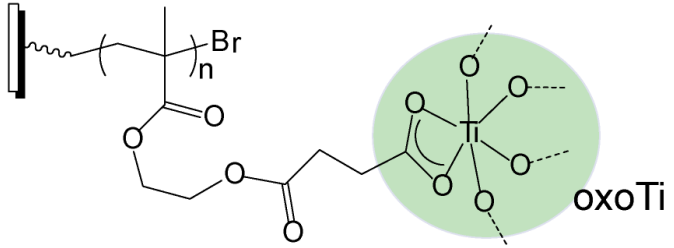


Figure 4.2. Structures of control polymer brushes for phosphopeptide enrichment.

dissolved. Then 0.1 M HCl was added to lower the solution pH to 7.4 prior to dilution with water and DMSO); (5) DMF containing 0.1 M propiolic acid, 5 mM CuBr and 8.75 mM 2,2'-dipyridyl (the solution was degassed with three freeze-pump-thaw cycles) for 24 h, with modest stirring to promote mass transfer, followed by rinsing the wafer with 0.1% TFA containing 50% ACN to protonate the carboxylic acids; (6) 0.1 M pH 7.4 EDTA for 30 min to remove any chelated Cu (I) or Cu (II) ions in the polymer; (7) 0.1 M $Ti(i\text{-}PrO)_4$ in anhydrous THF for 30 min followed by rinsing with anhydrous THF in the N_2-filled glove bag; (8) 0.1 M $FeCl_3$ for 30 min followed by rinsing with ethanol. Unless otherwise stated, each of the above steps includes rinsing the wafer first with the solvent employed for reaction and then with ethanol. The wafer was always dried with N_2 before the next step.

We also synthesized four control materials (Figure 4.2) on Au wafers for comparison with the bifunctionalized oxoTi-PGMA-Fe brushes. The PGMA-based control polymers, oxoTi-PGMA and PGMA-Fe, were prepared following the bifunctionalization route except omitting the reaction with $Ti(i\text{-}PrO)_4$ or Fe^{3+} chelation, respectively. Preparation of PHEMA-based control polymers, oxoTi-PHEMA and PHEMA-Fe, followed the procedures in Chapters 3.

4.2.3 Synthesis of functionalized polymer brushes in nylon membranes

Synthesis of poly(2-(methacryloyloxy)ethyl succinate) (PMES) brushes in the pores of nylon membranes followed our previous method.[13] Briefly, a UV/ozone cleaned nylon membrane (25 mm diameter disc, 0.45 μm pore size) was supported in a home built Teflon cell that allows flow through the membrane. Immobilization of a silane initiator occurred by circulating a THF solution containing 16 mM (11-(2-bromo-2-

methyl)propionyloxy)undecyltrichlorosilane (synthesized as reported[14]) through the membrane for 2 h followed by rinsing with 20 mL of THF. A mixture of 2-(methacryloyloxy)ethyl succinate monomer and 1 M aqueous NaOH (1:1, v/v), and a DMF solution containing 2 mM CuBr, 1 mM $CuBr_2$, and 6 mM 1,1,4,7,10,10-hexamethyltriethylenetetramine were degassed separately with three freeze-pump-thaw cycles. In a N_2 bag, the two solutions were mixed at a 10:1 volume ratio, and the resulting solution was circulated through the membrane for 15 min to accomplish ATRP of MES. Finally, 10 mL of ethanol and 10 mL of water were sequentially passed through the membrane to remove physisorbed monomer, polymer, or catalyst.

Introduction of oxoTi clusters in the PMES-modified membrane to form Mem-P-oxoTi occurred by passing ~40 mL of a 50 mM $Ti(i\text{-}PrO)_4$/THF solution through a membrane modified with PMES for 30 min and rinsing with 20 mL of THF. To immobilize NTA groups in the MES-modified membrane, 10 mL of an aqueous solution containing 0.1 M *N*-hydroxysuccinimide (NHS) and 0.1 M *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride was circulated through the membrane for 30 min, followed by 10 mL of 0.1 M aminobutyl NTA (adjusted to pH 10.0 with 0.1 M NaOH) for 1 h. Ten mL of water was passed through the membrane for rinsing after each step. Finally, the Fe(III)-NTA complex was formed by first circulating10 mL of 0.1 M $FeCl_3$ aqueous solution through the membrane for 30 min and then sequentially passing 10 mL of 0.1% TFA/50% ACN (v/v) and 10 mL of ethanol through the membrane to remove unbound Fe(III) ions. Both types of functionalized membrane discs were stored in a desiccator until use. Unless specified, the flow rate in the above procedures was ~10 mL/min except in the ATRP process, where the rate of monomer solution dropped from ~8 mL/min to ~2 mL/min after 15 min.

4.2.4 Characterization of polymer brushes

The thicknesses of polymer brushes on Au-coated Si wafers were determined using a rotating analyzer spectroscopic ellipsometer (J. A. Woollam, M-44), assuming a film refractive index of 1.5. Reflectance FTIR spectra of the modified Au surfaces were collected on a Nicolet Magna 560 spectrophotometer with a Pike grazing angle (80 °) accessory, and polymer growth on nylon membranes was verified by attenuated total reflectance (ATR) FTIR spectroscopy (Perkin-Elmer Spectrum One Instrument) using an air background. For reflectance IR spectra, a UV/O_3-cleaned, Au-coated Si wafer served as the background.

Inductively coupled plasma optical emission spectrometry (ICP-OES) was employed to determine the amount of Ti and Fe eluted from the bifunctional polymer brushes. For this determination, each polymer-modified Au-coated Si wafer was immersed in 4 mL of aqua regia for 5 min to remove the entire polymer along with Au from the Si wafer, and the aqua regia was diluted to 50 mL with water. To prepare standard solutions (for calibration) with the same background as the sample solution, 20 pieces of 2.4 $\times$ 1.1 cm Au-coated wafers were immersed in 80 mL of aqua regia to form the Au/aqua regia solution, and 4 mL of this solution was added to a series of 50-mL volumetric flasks. Certain volumes of 998 ppm Ti and 100 ppm Fe stock solutions were then added to each flask prior to dilution to 50 mL with water. Using these standards, calibration curves for Ti and Fe were obtained with ICP-OES, and the amounts of Ti and Fe in different sample solutions were determined using the calibration curve. The surface coverages of Ti and Fe were calculated by dividing the metal amounts by the wafer area.

4.2.5 Phosphopeptide enrichment

Protein digestion and on-plate enrichment of phosphopeptides using modified Au-coated wafers followed the procedures in Chapter 3. For enrichment using modified membranes, 6-mm-diameter discs were cut from the 25-mm-diameter membranes using a hole punch, and one or two small discs were inserted in a flangeless fitting system (A-424, Upchurch Scientific) with a frit between the membrane and the downstream Teflon tubing (I.D. 1/32"). The solution was loaded in a Hamilton glass syringe and pushed through the membrane disc using a syringe pump (Fisher Scientific). The dry protein digests were dissolved in 0.1% or 1% TFA containing 50% ACN. Before enrichment, 30 μL of blank loading solution (the solution used for dissolving samples) was passed through the membrane (100 μL/min) for equilibration. Phosphopeptide isolation occurred by passing the digest solution through the membrane (50 μL/min) and washing with blank loading solution (100 μL/min, 200 μL for samples containing only α-casein, β-casein and ovalbumin digests, and 1 mL for samples containing a mixture of digests of three phosphoproteins and bovine serum albumin (BSA)). Before elution, ~500 μL of air was loaded into the syringe and gently pushed through the membrane by hand (~10 s) to remove residual washing solution. Finally, phosphopeptides were eluted by passing 1% H_3PO_4 solution containing 50% ACN through the membrane at 10 μL/min, and 10 μL of eluate was collected in an Eppendorf tube. One μL of the eluate was spotted on the steel MALDI plate followed by addition of 2,5-dihydroxybenzoic acid (DHB) for MALDI-MS analysis.

4.2.6 MS analysis of phosphopeptides and data analysis

MS analysis was performed in positive ion mode using an LTQ XL ion trap mass spectrometer equipped with a vMALDI source (Thermo Fisher Scientific, San Jose, CA). For

each experiment, mass spectra were obtained from at least 3 enrichment spots, and the mass spectra given here are representative examples. Confirmation of phosphopeptide sequences by CID-MS/MS and database search were the same as in Chapter 3. Table 4.2 gives peptide sequence information.

4.3 Results and discussion

4.3.1 Polymer synthesis on Au wafers and in nylon membranes

PGMA is attractive as a versatile affinity chromatography material because opening of the epoxy ring allows facile polymer functionalization.[13-15] Thus, formation of bifunctionalized polymer brushes on Au surfaces starts with surface-initiated ATRP of GMA. Varying the polymerization time from 1 to 4 h yields homogeneous PGMA brushes with thicknesses that range from 20 to 70 nm, and control over film thickness allows variation of the phosphopeptide binding capacity of functionalized films.

Reflectance FTIR spectra (Figure 4.3) provide evidence for the synthesis and derivatization of PGMA to form bifunctional brushes. The ester carbonyl absorbance at 1740 cm^{-1} in Figure 2a confirms the formation of PGMA (**a**, Figure 4.1) on the Au surface. Subsequent opening of the epoxide by reaction of **a** with NaN_3 generates adjacent azide and hydroxyl groups (**b**, Figure 4.1), as confirmed by the antisymmetric N=N=N stretch at 2100 cm^{-1} and the broad hydroxyl stretch at 3100-3650 cm^{-1} in Figure 4.3b. Functionalization of the hydroxyl groups follows the same route used to form PHEMA-SA-NTA in Chapter 2, with some changes in reaction conditions or reagents. Reaction of **b** with SA takes longer (8 h) than

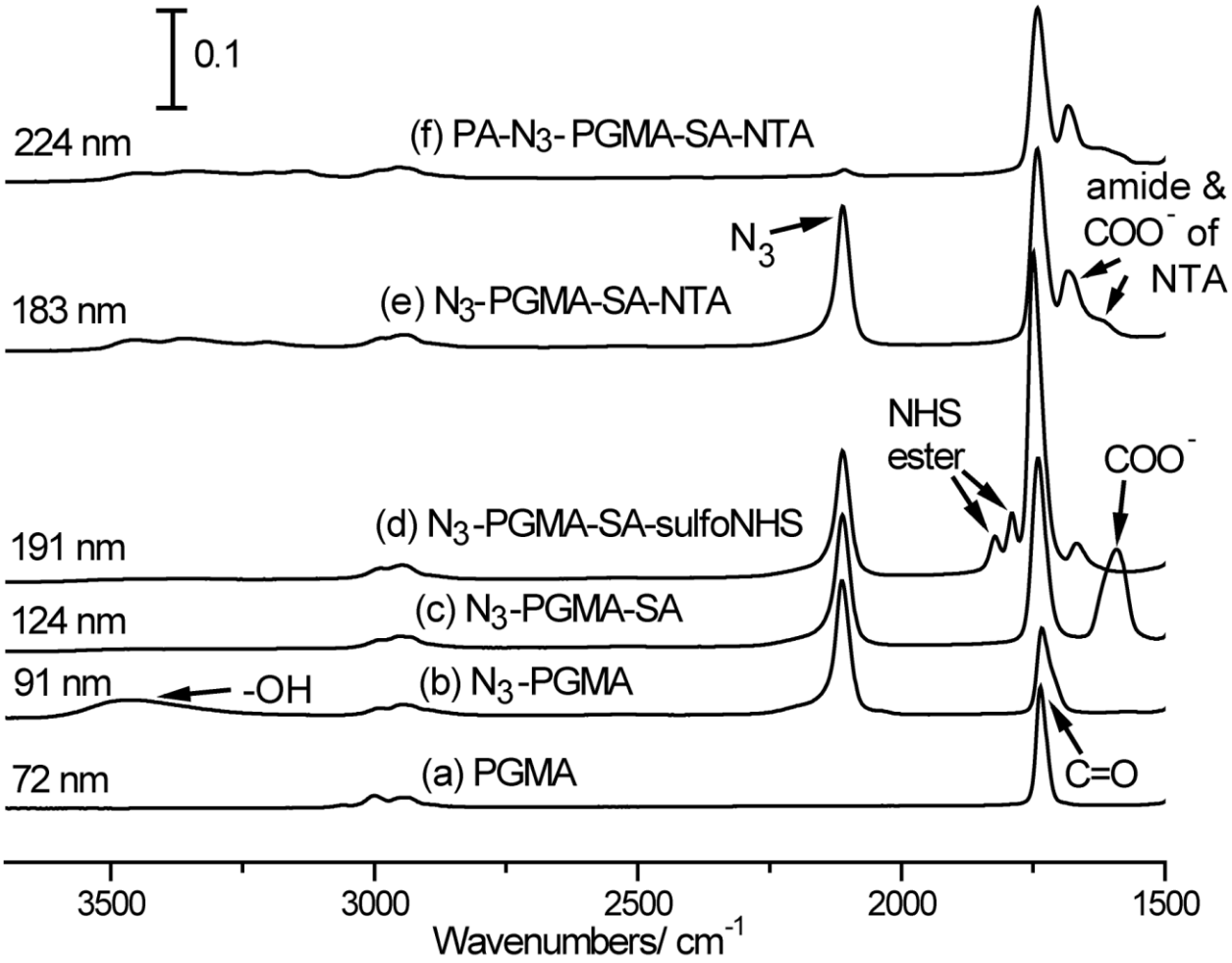


Figure 4.3. Reflectance FTIR spectra of a PGMA film on a Au-coated wafer before (a) and after the following sequential steps: (b) reaction with NaN_3; (c) reaction with SA followed by 1-min deprotonation in 0.01 M NaOH and rinsing with acetone; (d) activation with sulfo-NHS/DIC; (e) reaction with aminobutyl NTA followed by 1-min deprotonation in 1 mM NaOH and rinsing with acetone; (f) reaction with propiolic acid. The thickness of the film (dry form) after each step is noted above each spectrum.

reaction of PHEMA with SA (3 h), probably due to the less swollen structure of PGMA-N_3 compared to PHEMA. (We monitored the completeness of this step by occasionally removing the wafer from the reaction solution, deprotonating the carboxylic acid groups with a 1-min immersion in 0.01 M NaOH, rinsing with acetone, and examining the intensity of the antisymmetric COO^- stretch at 1590 cm^{-1}. We stopped the reaction when this absorbance ceased

to increase with reaction time.) After activation of the newly formed polyacid (**c**, Figure 4.1) with sulfo-NHS/DIC (NHS ester peaks at 1825 cm^{-1} and 1793 cm^{-1} in Figure 4.3d confirm the activation), the polymer (**d** in Figure 4.1) reacts with aminobutyl NTA to yield **e** (Figure 4.1). Both the disappearance of the NHS ester peaks and the appearance of a peak at 1685 cm^{-1} provide evidence for derivatization with NTA (Figure 4.3e). The absorbance around 1685 cm^{-1} and its broad tail likely stem from the newly formed amide bond and the carboxylate groups of NTA. The introduction of another carboxylic acid group occurs through a Cu(I)-catalyzed (3 + 2) azide-alkyne cycloaddition, also known as "click chemistry".[16] We chose this reaction because of its high efficiency and tolerance to a variety of functional groups.[17] Reaction of 180 nm thick **e** (Figure 4.1) with propiolic acid (catalyzed by CuBr/2,2'-dipyridyl) is nearly quantitative, as shown by the minimal residual N=N=N stretch at 2100 cm^{-1} (Figure 4.3f) after a 48 h "click reaction". When **e** is below 100 nm thick, the N=N=N stretch completely disappears after 24 h of reaction. A control experiment with "click reaction" conditions in the absence of propiolic acid for yielded no change in the IR spectrum of polymer **e**, providing further evidence for the formation of new carboxylic acid groups in the polymer brushes. Immersing a wafer coated with **f** (Figure 4.1) in $Ti(i\text{-}PrO)_4$/THF solution leads to formation of oxoTi clusters (see Chapter 3) coordinated with the carboxylic acid groups attached to the triazole rings (**g**, Figure 4.1). The thickness of the polymer increases ~10% after this step. Unfortunately, IR spectra show no evident changes after this reaction. A control experiment using a Au-PHEMA-SA-NTA wafer suggested that the NTA groups do not bind a significant amount of Ti(IV) species, as the polymer thickness does not increase (<1 nm) after exposure to $Ti(i\text{-}PrO)_4$. Finally, immersion of

polymer **g** (Figure 4.1) in $FeCl_3$ solution gives the NTA-Fe(III) complex, **h** (Figure 4.1). Binding of Fe(III) also does not lead to significant changes in the IR spectrum of the film.

Table 4.1. The amount of Fe and Ti in functional polymer brushes.

		oxoTi-PGMA-Fe					PHEMA-Fe-1**	PHEMA-Fe-2***
		1	2	3	4	5		
Base polymer* thickness (nm)		71.4	76.9	120	151	193	148	143
Fe	μg/cm2	1.15	1.13	1.73	2.43	3.32	2.68	2.90
	nmol/cm2-nm	Average 0.28±0.02					0.32	0.36
Ti	μg/cm2	1.45	1.50	2.67	2.76	2.89	0.04	0.56
	nmol/cm2-nm	Average 0.40±0.06					0.01	0.08

* Base polymer is the polymer structure before modification with either Ti(IV) or Fe(III) species.
** The NTA-derivatized PHEMA film was directly immersed in 0.1 M $FeCl_3$.
*** The NTA-derivatized PHEMA film was immersed in Ti(*i*-PrO)$_4$/THF solution for 30 min prior to rinsing and immersion in 0.1 M $FeCl_3$.

To demonstrate the presence of Fe and Ti in oxoTi-PGMA-Fe films, we eluted Ti and Fe from modified wafers using aqua regia and performed ICP-OES to determine Ti and Fe concentrations in the eluate. Table 4.1 shows that surface coverages of Fe and Ti generally increase with the thickness of the base polymer film (**f**). The metal coverages normalized to base film thickness are 0.28±0.02 nmol/(cm^2-nm) Fe and 0.40±0.06 nmol/(cm^2-nm) Ti in oxoTi-PGMA-Fe films. Assuming the density of the film is 1 g/cm^3 and using a repeating-unit mass of 599.6 g/mol (structure **f** in Figure 4.1), these metal densities correspond to 1.6±0.1 Fe atoms and 2.4±0.4 Ti atoms per repeating unit. Attainment of >1 Fe atom per repeating unit could occur because there was some polymerization on the back side of the wafer, which was not coated with

Au. (The back side of the wafer appears to change color after polymerization.) In the case of Ti, there could be several Ti atoms per clusters, and some Ti clusters may also form on the back side of the wafer.

We also examined the extent of reaction of NTA groups with $Ti(i\text{-}PrO)_4$. Immersing an NTA-derivatized PHEMA brush in $Ti(i\text{-}PrO)_4$/THF solution for 30 min followed by Fe(III) chelation (PHEMA-Fe-2 in Table 4.1) leads to a much smaller loading of Ti in the film compared to in polymer **f** (0.08 vs. 0.40 nmol Ti atoms/cm^2-nm). Elemental analysis of a PHEMA-Fe film (PHEMA-Fe-1 in Table 4.1) that was not exposed to $Ti(i\text{-}PrO)_4$ shows 0.1 nmol Ti/cm^2-nm, which is essentially the experimental uncertainty of the measurement.

To perform larger scale phosphopeptide enrichment, we synthesize and functionalize PMES brushes in nylon membranes. Although peaks due to the amide groups (1632 cm^{-1}) in the base membrane dominate the ATR-IR spectra of PMES-modified nylon, the carbonyl absorbance at ~1740 cm^{-1} (Figure 4b) confirms the formation of PMES on the membrane. The succinimide ester absorptions at 1820 cm^{-1} and 1790 cm^{-1} in Figure 4c show activation of the PMES brushes, and the disappearance of these peaks in Figure 4d suggests the substitution of NHS by aminobutyl NTA, although hydrolysis could also occur. The peaks due to the base membrane mask the amide and NTA absorbances that should appear around 1685 cm^{-1} in Figure 4c. Subsequent passage of $FeCl_3$ solutions through the membrane yields the NTA-Fe(III) complex. Similar to the synthesis of poly acid-oxoTi on Au wafers (Chapter 3), oxoTi clusters also form when passing a $Ti(i\text{-}PrO)_4$/THF solution through PMES-modified nylon membranes. Formation

of PMES-NTA-Fe in membranes also occurs using the chemistry previously employed on gold wafers, but solutions are passed through the membrane.

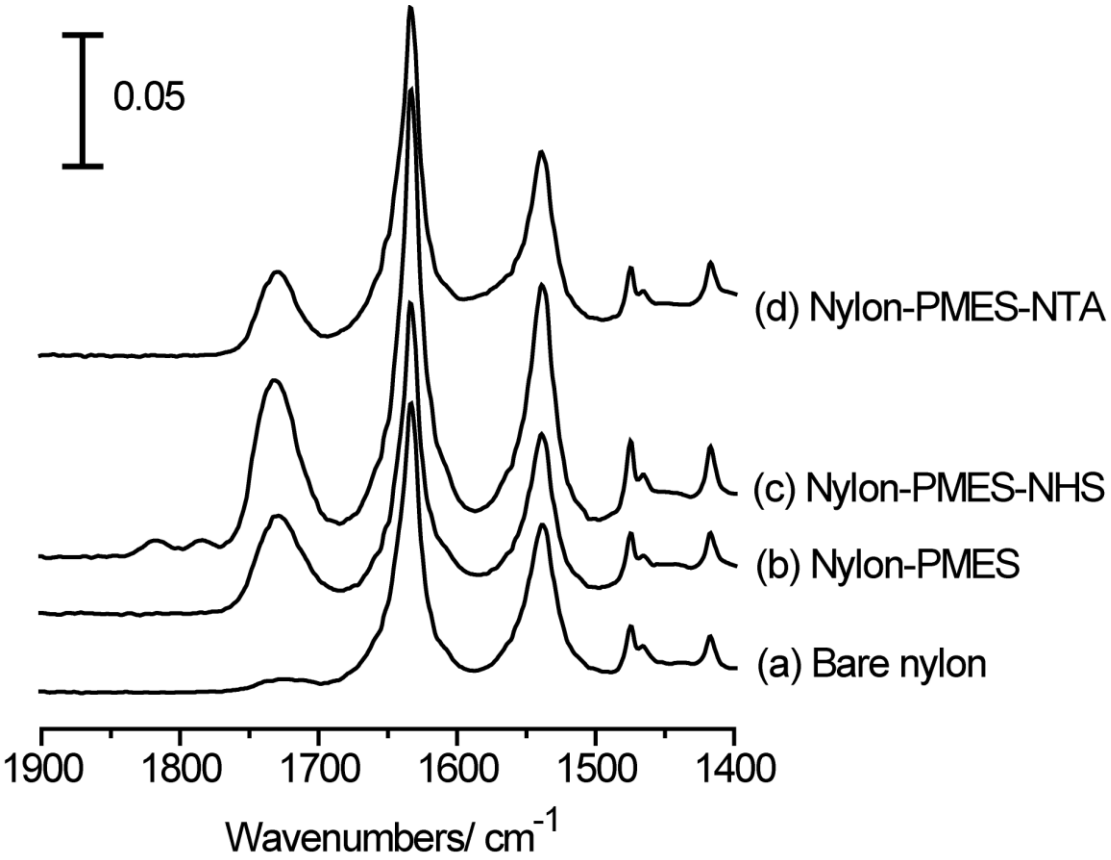


Figure 4.4. ATR-IR spectra of nylon membranes before (a) and after (b) modification with PMES, (c) activation with NHS/1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and (d) reaction with aminobutyl NTA (pH 10).

4.3.2 On-plate phosphopeptide enrichment using oxoTi-PGMA-Fe brushes

The previous chapters and a number of literature studies show that preferences for binding mono- and multi-phosphorylated peptides stem from (1) the intrinsic affinities of metal-ion complexes or metal oxide materials,[7,18] (2) the binding capacity of the substrates (or peptide-to-affinity sites ratio),[4,5] and (3) the acidity of the sample solution.[7] With these considerations in mind, we performed on-plate phosphopeptide enrichment using oxoTi-PGMA-Fe brushes with different thicknesses. Similar studies with the four control materials in Figure

4.2 separately verify the phosphopeptide enrichment specificities of the oxoTi and NTA-Fe. Among the four control polymers, PHEMA-Fe was used in Chapters 2 and 3, and oxoTi-PHEMA was described in Chapter 3.

4.3.2.1 Influence of film thickness on the specificity of affinity groups

We first tested phosphopeptide enrichment from a 1-μL mixture of three phosphoprotein digests (200 fmol each of α-casein, β-casein and ovalbumin) dissolved in 0.1 % TFA. Phosphopeptide isolation occurred in 60 nm-thick polymer films functionalized with NTA-Fe(III) and/or oxoTi complexes (the thickness corresponds to the base polymer prior to modification with Fe(III) or oxoTi). Enrichment on PGMA-Fe and PHEMA-Fe leads to spectra dominated by multi-phophorylated peptides (Figures 5a and 6a). In contrast, the two strongest signals in the mass spectrum of sample enriched on oxoTi-PHEMA stem from mono-phosphorylated peptides (Figure 4.6b). In the case of oxoTi-PGMA, signals from both types of phosphopeptides appear, but the spectral pattern (Figure 4.4c) is closer to that of oxoTi-PHEMA than the PHEMA-Fe control. We should note that although the four control films have thicknesses of 60 nm, PHEMA-based brushes typically show 2- to 6-fold greater signal-to-noise ratios (S/N) than PGMA-based systems. This suggests a lower binding capacity for the PGMA-based control polymers, which is reasonable because unlike derivatized PHEMA, the functionalized PGMA contains two branched side chains, and only one of them contains a binding site in the control polymers.

Films with both oxoTi and Fe(III) complexes enrich a wider range of phosphopeptides than films containing one of these functionalities. Enrichment on the 60 nm-thick oxoTi-PGMA-Fe (Figure 4.5c) allows MS identification of 7 more mono-phosphorylated peptides and the same number of multi-phosphorylated peptides as enrichment on PGMA-Fe (Figure 4.5a), as well as 3

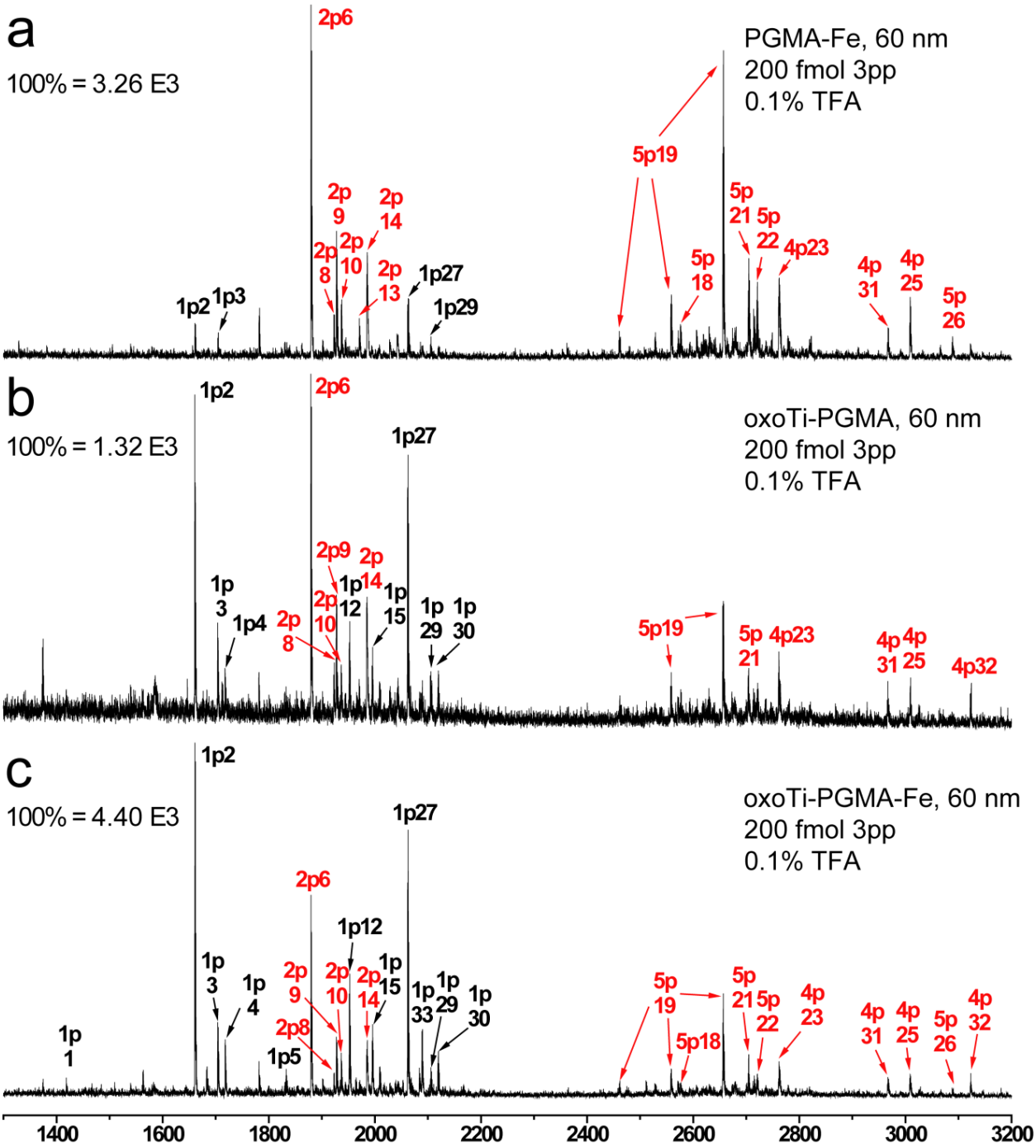


Figure 4.5. MALDI mass spectra after on-plate phosphopeptide enrichment from a 1-μL 0.1% TFA solution containing 200 fmol each of three phosphoprotein digests, α-casein, β-casein and ovalbumin (3pp). The MALDI plates were coated with 60 nm-thick (prior to addition of affinity groups) polymer films, whose structures after functionalization are noted in each spectrum. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

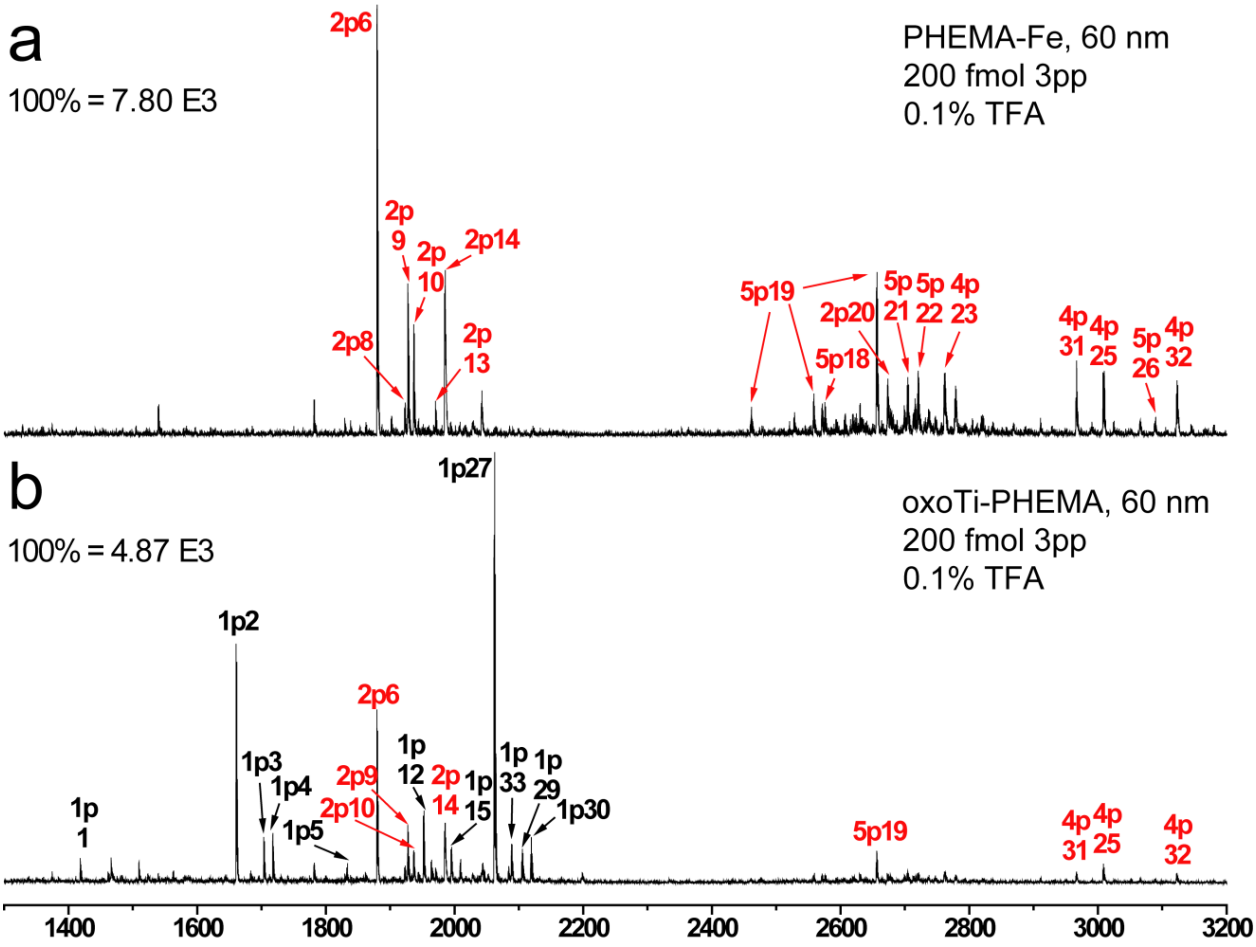


Figure 4.6. MALDI mass spectra after on-plate phosphopeptide enrichment from a 1-μL 0.1% TFA solution containing 200 fmol each of three phosphoprotein digests, α-casein, β-casein and ovalbumin (3pp). The MALDI plates were coated with 60 nm-thick (prior to addition of affinity groups) polymer films, whose structures after functionalization are noted in each spectrum. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

more mono-phosphorylated peptides and 3 more multi-phosphorylated peptides than enrichment on oxoTi-PGMA (Figure 4.5b). In addition, the MS noise after oxoTi-PGMA-Fe (Figure 4.5c) enrichment is similar to that from PGMA-Fe (Figure 4.5a) and only 25% of the noise with oxoTi-PGMA (Figure 4.5b). The presence of two binding functionalities and the resulting increased binding capacities likely lead to the superior performance of oxoTi-PGMA-Fe relative to enrichment on oxoTi-PGMA or PGMA-Fe.

Table 4.2. Phosphopeptides detected in positive-ion mode MALDI-MS spectra. The list contains the phosphopeptide sequence* (if identified), the range of amino acids, the number of phosphoryl groups, and the monoisotopic m/z of $[M+H]^+$.

Peptide #		Phosphopeptide sequence	Sequence Number**	# – PO_3H_2	Mono. m/z MH^+
α-casein	1	K.TVDME**pS**TEVFTK.K M(cam)-105	S2 153-164	1	1418.6
	2	K.VPQLEIVPN**pS**AEER.L	S1 121-134	1	1660.8
	3	K._VPQLEIVPN**pS**AEER.L Carbamyl (N-term)	S1 121-134	1	1703.8
	4	K.VPQLEIVPN**pS**AEER.L E(cam)	S1 121-134	1	1717.8
	5	YLGEYLIVPN**pS**AEER	S1***	1	1832.8
	6	K.DIG**pS**E**pS**TEDQAMEDIK.Q M(cam)-105	S1 58-73	2	1879.7
	7	K.DIG**pS**E**pS**TEDQAMEDIK.Q sodium adduct, M(cam)-105	S1 58-73	2	1901.7
	8	K._DIG**pS**E**pS**TEDQAMEDIK.Q Carbamyl (N-term), M(cam)-105	S1 58-73	2	1922.7
	9	K.DIG**pS**E**pS**TEDQAMEDIK.Q	S1 58-73	2	1927.7
	10	K.DIG**pS**E**pS**TEDQAMEDIK.Q M(cam)-105, D(cam)	S1 58-73	2	1936.7
	11	K.DIG**pS**E**pS**TEDQAMEDIK.Q M(ox)	S1 58-73	2	1943.7
	12	K.YKVPQLEIVPN**pS**AEER.L	S1 119-134	1	1952.0
	13	K.DIG**pS**E**pS**TEDQAMEDIK.Q Carbamyl (K)	S1 58-73	2	1970.7
	14	K.DIG**pS**E**pS**TEDQAMEDIK.Q D(cam)	S1 58-73	2	1984.7
	15	K._YKVPQLEIVPN**pS**AEER.L Carbamyl (N-term)	S1 119-134	1	1995.0
	16	N/A		1	2027
	17	N/A		1	2037
	18	N/A		5	2575
	19	K.QMEAE**pS**I**pS**p**S**p**S**EEIVPN**pS**VEQK.H M(cam)-105, pyro-glu from Q	S1 74-94	5	2655.9
	20	N/A		≥2	2672
	21	K.QMEAE**pS**I**pS**p**S**p**S**EEIVPN**pS**VEQK.H Pyro-glu from Q	S1 74-94	5	2702.9
	22	K.QMEAE**pS**I**pS**p**S**p**S**EEIVPN**pS**VEQK.H	S1 74-94	5	2720.9
	23	N/A		4	2760
	24	N/A		≥3	2819
	25	R.NANEEEYSIG**pS**p**S**p**S**EE**pS**AEVATEEVK.I	S2 61-85	4	3008.0
	26	N/A	S2 61-85****	5	3088.0

Table 4.2 cont'd

β-casein	27	K.FQ**pS**EEQQQTEDELQDK.I	48-63	1	2061.8
	28	K.FQ**pS**EEQQQTEDELQDK.I Sodium adduct	48-63	1	2083.8
	29	K._FQ**pS**EEQQQTEDELQDK.I Carbamyl (N-term)	48-63	1	2104.8
	30	K.FQ**pS**EEQQQTEDELQDK.I E(cam)	48-63	1	2118.8
	31	R.ELEELNVPGEIVE**pS**L**pS**p**S**p**S**EESITR.I	17-40	4	2966.2
	32	A.RELEELNVPGEIVE**pS**L**pS**p**S**p**S**EESITR.I	16-40	4	3122.3
Ovalbu-min	33	R.EVVG**pS**AEAGVDAASVSEEFR.A	341-360	1	2088.9

* Some phosphopeptides share the same sequence, but are modified differently during protein denaturing, alkylation, or digestion. "Carbamyl" corresponds to carbamylation, (cam) corresponds to carbamidomethylation, M(cam-105) corresponds to the neutral loss of S-(methylthio)acetamide from M(cam), and M(ox) corresponds to methionine sulfoxide.
** S1 and S2 represent S1- and S2-chains of α-casein, respectively.
*** According to Larsen, M. R. et al.,[19] this is a new sequence variant of the α-S1 casein in the region 104–119.
**** According to Larsen, M. R. et al.,[19] this is probably a penta-phosphorylated peptide.

Increasing the thickness of PGMA films from 60 to 120 nm should enhance the phosphopeptide binding capacity of functionalized films. As comparison of Figure 4.5 and Figure 4.7 shows, enrichment on the thicker (120 nm) oxoTi-PGMA-Fe and PGMA-based control materials gives 1.2- to 4-fold higher S/N ratios for phosphopeptides. The high specificity of oxoTi toward mono-phosphorylated peptides is more evident after the increase in film thickness (compare Figures 7b and 5b). For specific quantitative comparison of different enrichments, we chose the four most frequently detected phosphopeptides, 1p2, 1p27, 2p6 and 5p19, which represent a range of degrees of phosphorylation. Signals for all of these phosphopeptides appear in Figure 4.7c with >44% relative intensity. However, 1p2 and 1p27 show intensities below 37% in Figure 4.7a, and 2p6 and 5p19 show intensities below 23% in Figure 4.7b. This suggests that oxoTi-PGMA-Fe enriches phosphopeptides with less bias toward the number of phosphate groups on the peptide than do polymers with only one type of affinity

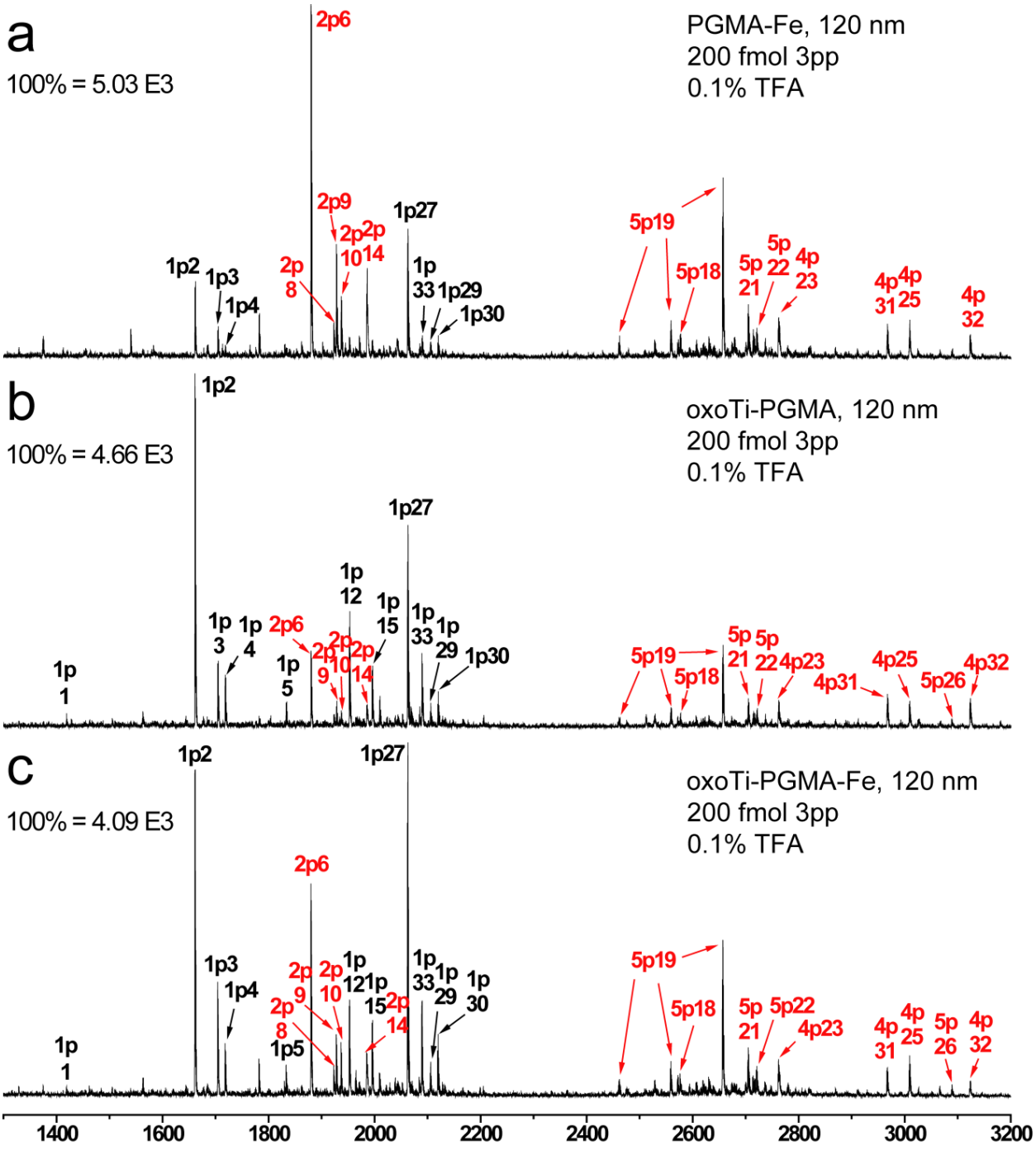


Figure 4.7. MALDI mass spectra after on-plate phosphopeptide enrichment from 1-μL of 0.1% TFA containing 200 fmol each of three phosphoprotein digests, α-casein, β-casein and ovalbumin (3pp). The MALDI plates were coated with 120 nm-thick (prior to addition of affinity groups) polymer films, whose structures after functionalization are noted in each spectrum. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

group. In addition, the MS signal pattern after enrichment on oxoTi-PGMA-Fe remains almost the same after the polymer thickness doubles (compare Figures 5c and 7c), suggesting that after surpassing a certain thickness (e.g. 60 nm for 200 fmol of three phosphoprotein digests mixture in 0.1% TFA) the bifunctionalized polymer maintains its low-bias enrichment of mono- and multi-phosphorylated peptides.

4.3.2.2 Influence of sample acidity on phosphopeptide enrichment

Sample acidity also affects phosphopeptide enrichment. With PGMA-Fe, increasing the TFA level in the sample and rinsing solutions from 0.1 to 1% raises the S/N of tetra- and penta-phosphorylated peptides 2- to 3-fold (compare Figures 8a and 7a). Similarly, higher acidity increases the signals of doubly phosphorylated peptides enriched by oxoTi-PGMA (2.3- to 3.4 -fold increase, compare Figures 8b and 7b). Enrichment of the same sample on the 120-nm thick oxoTi-PGMA-Fe yields a mass spectrum that is an "average" of the spectra that result from enrichment on the two control polymers PGMA-Fe and oxoTi-PGMA (compare Figures 8a-c). Nevertheless, the doubly phosphorylated peptide 2p6 and the penta-phosphopeptide 5p19 now dominate the spectrum (Figure 4.8c). As mono-phosphorylated peptides generally ionize better than multi-phosphorylated peptides, such a result suggests a biased enrichment of multi-phosphorylated peptides by the 120 nm thick bifunctional polymer film in 1% TFA. Because binding of mono-phosphorylated peptides to metal affinity sites decreases with higher solution acidity, the inefficient enrichment of mono-phosphorylated peptides likely stems from insufficient binding capacity of oxoTi-PGMA-Fe in 1% TFA.

Further increasing the thickness of oxoTi-PGMA-Fe to 220 nm for enrichment of 200 fmol of three phosphoprotein digests in 1% TFA brings the MS pattern (Figure 4.9c) back to that in Figure 4.7c, which was generated from the same amount of sample using the same polymer

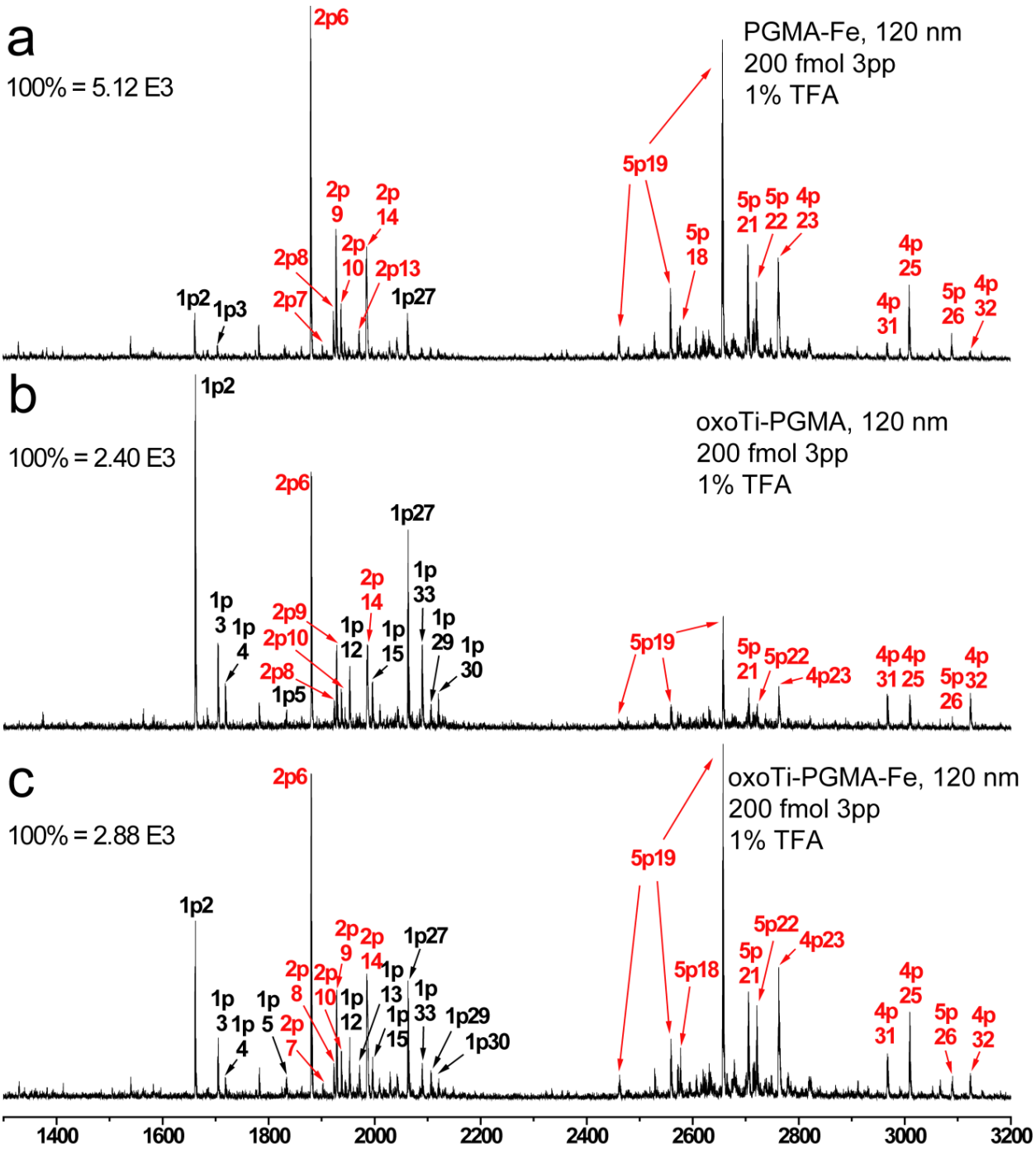


Figure 4.8. MALDI mass spectra after on-plate phosphopeptide enrichment from 1-μL of 1% TFA containing 200 fmol each of three phosphoprotein digests, α-casein, β-casein and ovalbumin (3pp). The MALDI plates were coated with 120 nm-thick (prior to addition of affinity groups) polymer films, whose structures after functionalization are noted in each spectrum. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

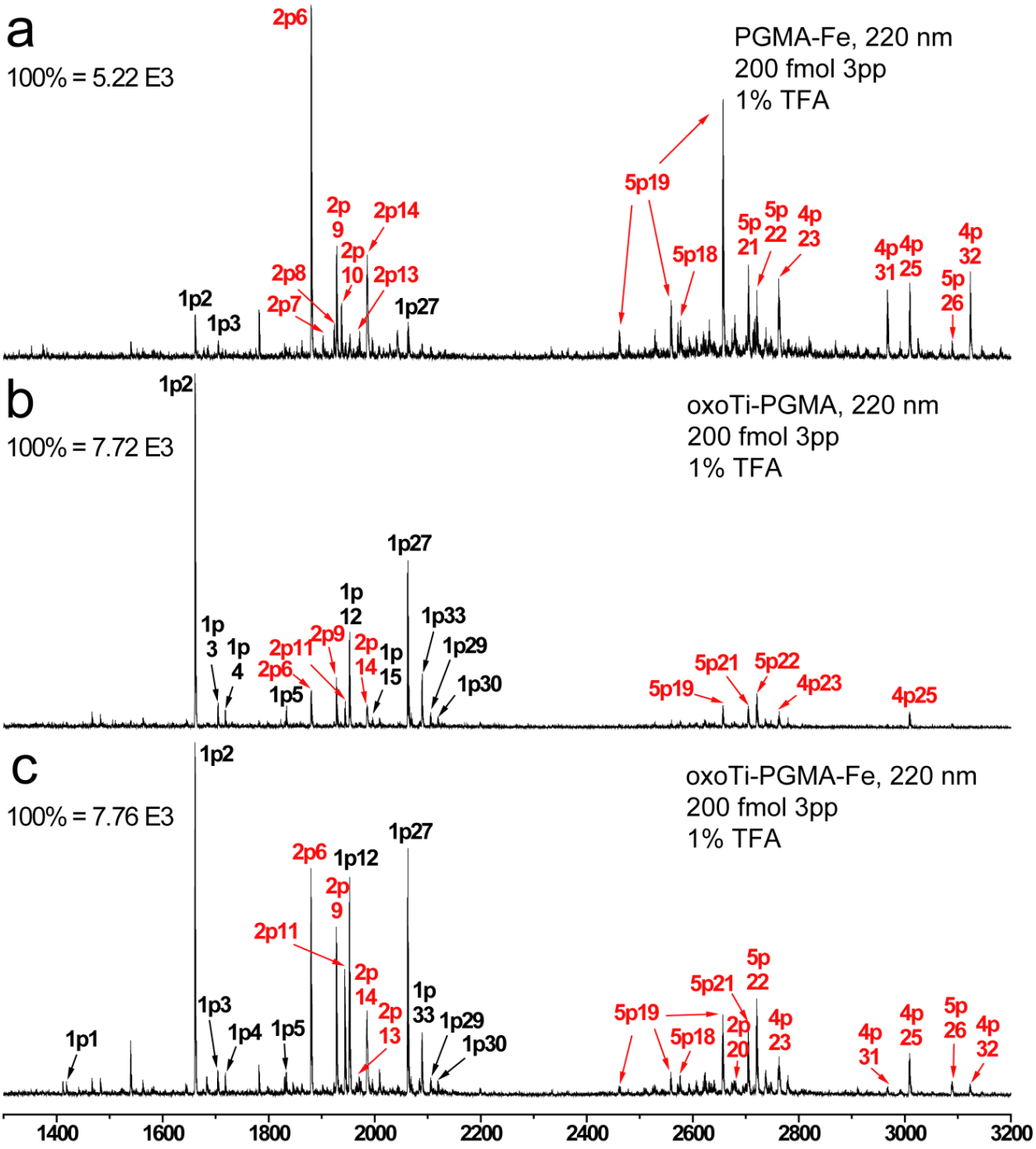


Figure 4.9. MALDI mass spectra after on-plate phosphopeptide enrichment from 1 μL of 1% TFA containing 200 fmol each of three phosphoprotein digests, α-casein, β-casein and ovalbumin (3pp). The MALDI plates were coated with 220 nm-thick (prior to addition of affinity groups) polymer films, whose structures after functionalization are noted in each spectrum. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

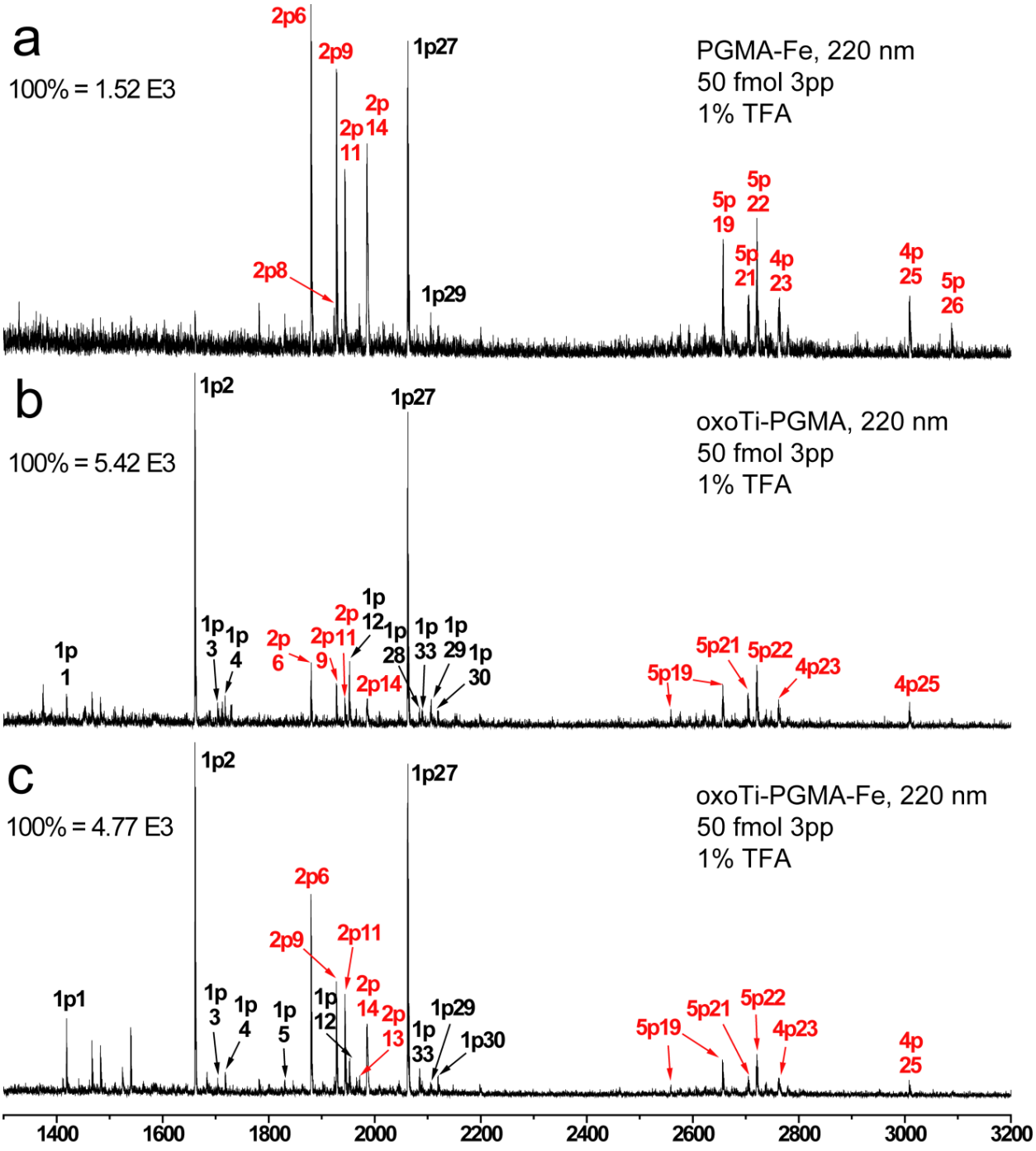


Figure 4.10. MALDI mass spectra after on-plate phosphopeptide enrichment from a 1 μL of 1% TFA containing 50 fmol each of three phosphoprotein digests, α-casein, β-casein and ovalbumin (3pp). The MALDI plates were coated with 220 nm-thick (prior to addition of affinity groups) polymer films, whose structures after functionalization are noted in each spectrum. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

structure, but with a 120 nm film and weaker acidity for enrichment. In addition, Figure 4.9c presents ~50% less noise than Figure 4.7c, as well as a nearly perfect combination of the spectra generated from its PGMA-based control polymers (Figure 4.9 a and b).

Finally we tested samples with lower amounts of phosphoprotein or and excess of non-phosphoprotein. With only 50 fmol each of three phosphoprotein digests in 1% TFA, enrichment on the 220 nm thick oxoTi-PGMA-Fe yields strong signals for mono- and doubly phosphorylated peptides due to the combined affinities of oxoTi and NTA-Fe groups (Figure 4.10). Enrichment on this film also leads to strong signals of mono- and multi-phosphorylated peptides from a mixture containing 200 fmol of three phosphoprotein digests and 2 pmol of BSA (non-phosphoprotein) digest . In the presence of excess BSA (0.16 mg total protein digest) peak patterns and S/N are similar for samples with and without BSA (compare Figure 4.11 and Figure 4.9c).

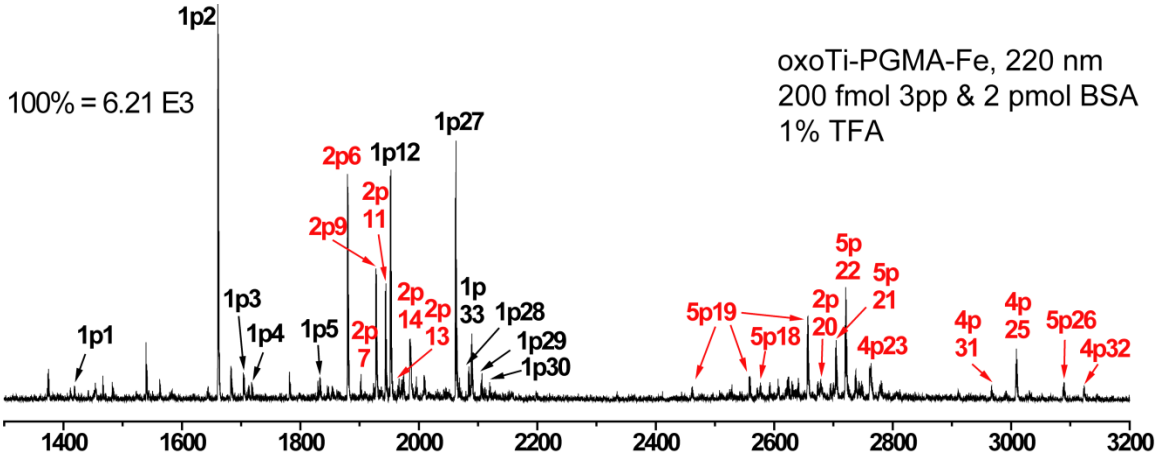


Figure 4.11. MALDI mass spectrum after phosphopeptide enrichment from 1 μL of 1% TFA containing 200 fmol each of three phosphoprotein digests (3pp), α-casein, β-casein and ovalbumin, and 2 pmol of BSA digest, The MALDI plates were coated with 220 nm-thick (prior to addition of affinity groups) oxoTi-PGMA-Fe. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

To summarize these two sections, the dual affinities of high-capacity, bifunctional polymer brushes allow low-bias enrichment of mono- and multi-phosphorylated peptides from samples with moderate acidity (0.1-1% TFA) and complexity (e.g. <5-20 μg/mL of several phosphoproteins and < 0.5 mg/mL non-phosphoproteins). The bifunctional film is attractive for enrichment of target phosphoproteins from low-volume samples such as those that result immuno-purifications, gel electrophoresis, and high-throughput kinase assays. In addition, the orthogonal brush derivatization strategy can create bifunctional films for other applications that require two adjacent functional groups.

4.3.3 Enrichment of phosphopeptides using functional membranes

To enrich dilute phosphopeptides from 10-500 μL samples, we modified the micron-size pores of nylon membranes with PMES-oxoTi (Mem-PMES-oxoTi) or with PMES-NTA-Fe(III) (Mem-PMES-Fe). In an initial test with Mem-PMES-oxoTi, we enriched phosphopeptides from a digest mixture containing 3 pmol each of α-casein, β-casein and ovalbumin in 200 μL of 0.1% TFA in 50% ACN. After passing this mixture through the membrane and eluting with 1% H_3PO_4 in 50% ACN, mono-phosphorylated peptide signals dominate the MALDI mass spectrum of 1 μL of eluate (Figure 4.12a). Increasing the amount of phosphoprotein digests to 10 pmol (without changing the sample volume) does not significantly improve the detection of multi-phosphorylated peptides, but the S/N of most of the mono-phosphorylated peptides increases ~2 fold (Figure 4.12b). At these protein amounts, the PMES-oxoTi brushes in the nylon membrane evidently have sufficient binding capacity to maintain the specificity of oxoTi for mono-phosphorylated peptides. We should emphasize that only 1/10 of the eluate was spotted on the MALDI plate in all cases, which leads to lower S/N than might be expected for these relatively

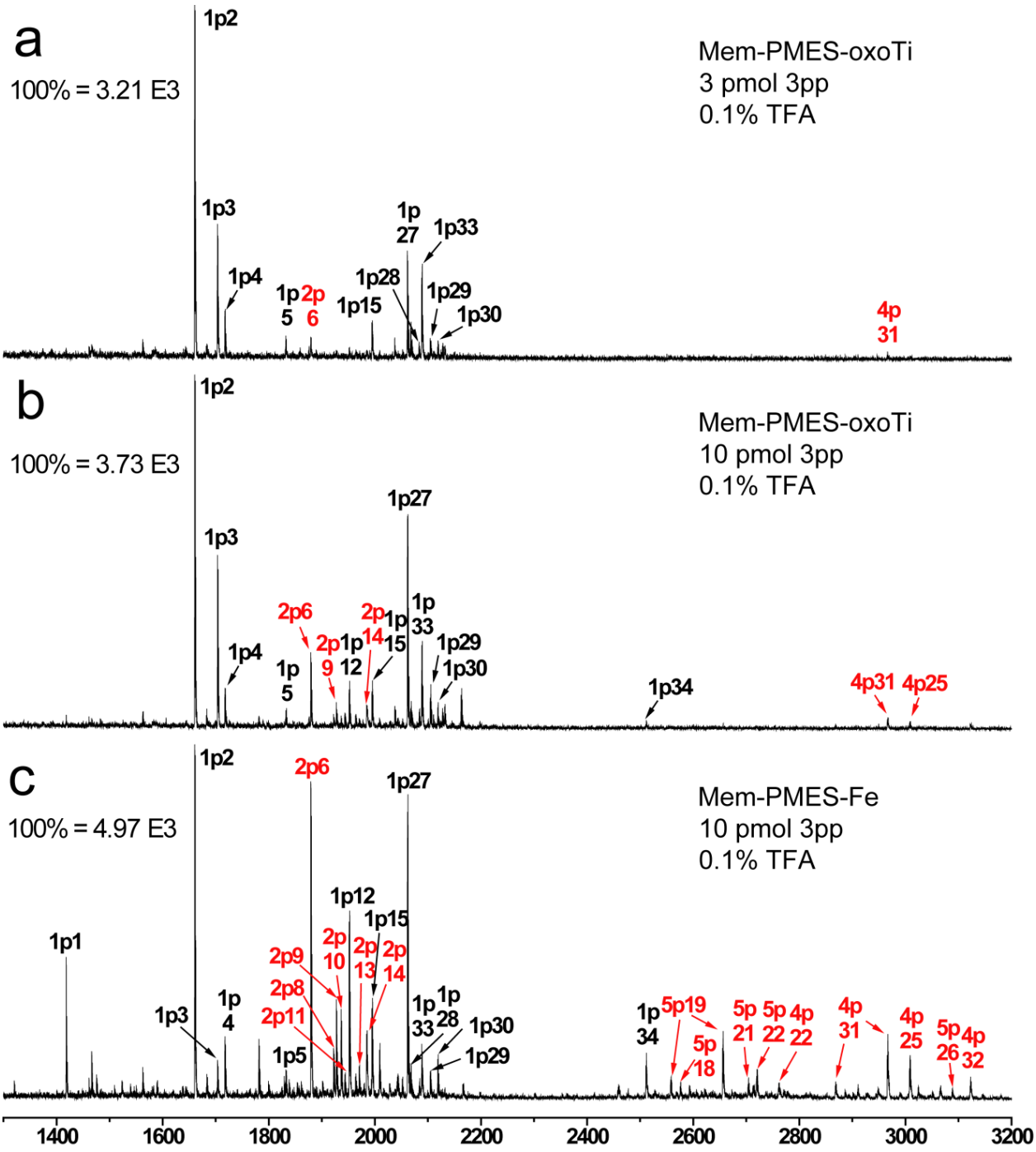


Figure 4.12. MALDI mass spectra of phosphopeptides enriched from 200 μL of 0.1% TFA, 50% ACN containing three phosphoprotein digests (3pp), α-casein, β-casein and ovalbumin. In all cases, only 1 μL of the 10 μL eluate was spotted on the sample plate. The membrane modifications and amount of protein in each sample are noted above the spectra. Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

large amounts of protein.

Compared with the Mem-PMES-oxoTi system, enriching the 10-pmol sample with Mem-PMES-Fe yields a more comprehensive mass spectrum (Figure 4.12c) with strong signals from both mono- and multi-phosphorylated peptides, including signals from 9 tetra- or penta-phosphopeptides. The high binding capacity likely increases the enrichment of monophosphopeptides by the NTA-Fe(III) complexes relative to capture in a film on a MALDI plate.

When enriching a sample containing 10 pmol (each) of α-casein, β-casein and ovalbumin digests, and 1 nmol of BSA digest (1:1:1:100 molar ratio), we employed 1% TFA in 50% ACN to suppress the binding of non-phosphorylated peptides. Under these conditions, however, Mem-PMES-Fe enriches only multi-phosphorylated peptides, especially tetra- and penta-phosphopeptides (Figure 4.13a). In contrast, Mem-PMES-oxoTi still enriches mono-phosphorylated peptides (Figure 4.13b) but does not reveal significant signals from tetra- and penta-phosphorylated species. To realize a more comprehensive enrichment of phosphopeptides in one sample, we stacked Mem-PMES-oxoTi and Mem-PMES-Fe in series in a single fitting system. (The loading solution first encountered the membrane with oxoTi.) The tandem membrane enrichment leads to a mass spectrum that essentially combines the signals from the first two spectra (compare Figure 4.13a-c), although signals are not as strong as after enrichment with the separate membranes, perhaps because of ion suppression.

We also mixed the eluates generated separately from Mem-PMES-Fe and Mem-PMES-oxoTi. MALDI-MS of the mixture shows dominant signals for tetra- and penta-phosphorylated peptides (Figure 4.13d), suggesting a lower recovery of mono-phosphorylated peptides with Mem-PMES-oxoTi than for multi-phosphorylated peptides with Mem-PMES-Fe. Separation of

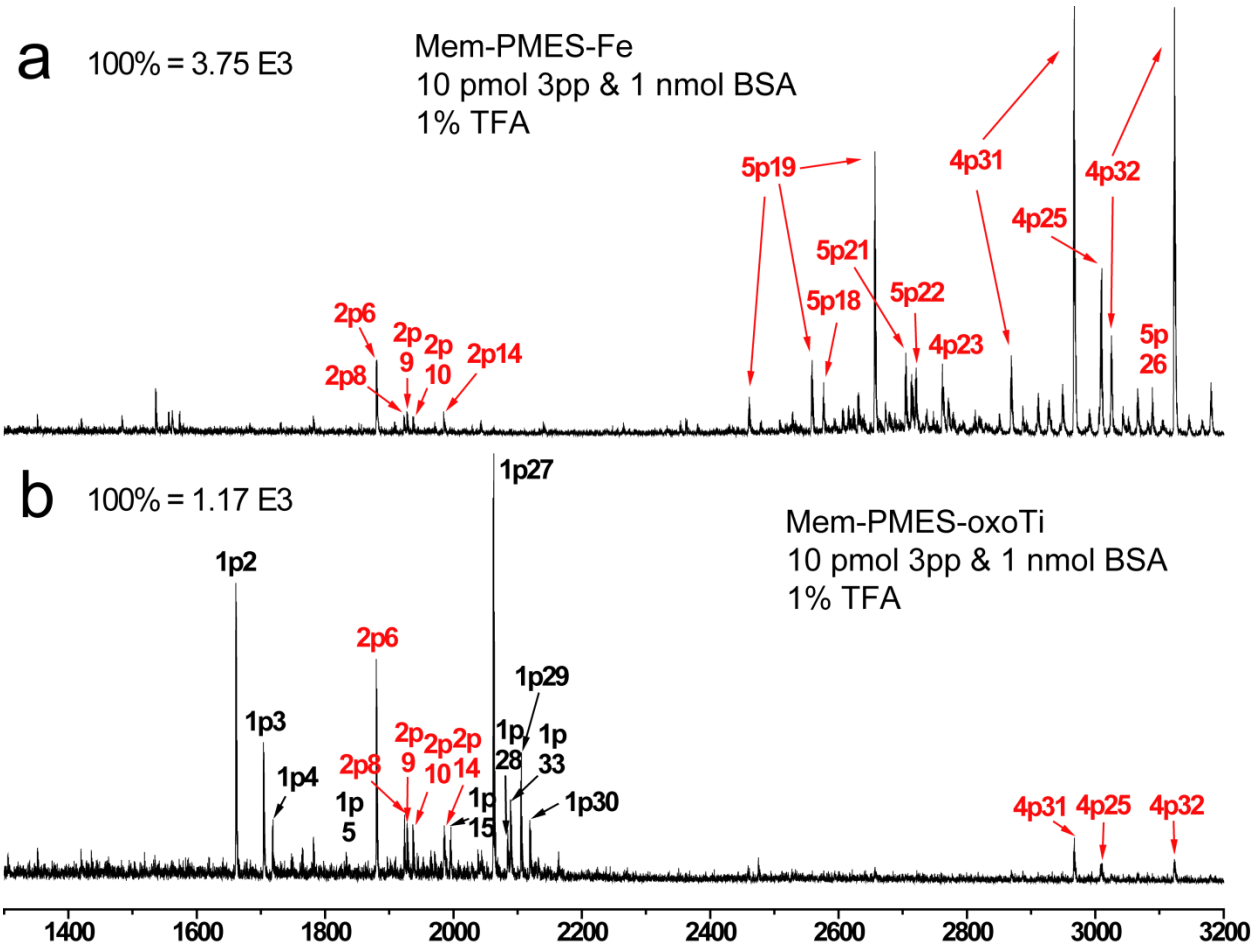
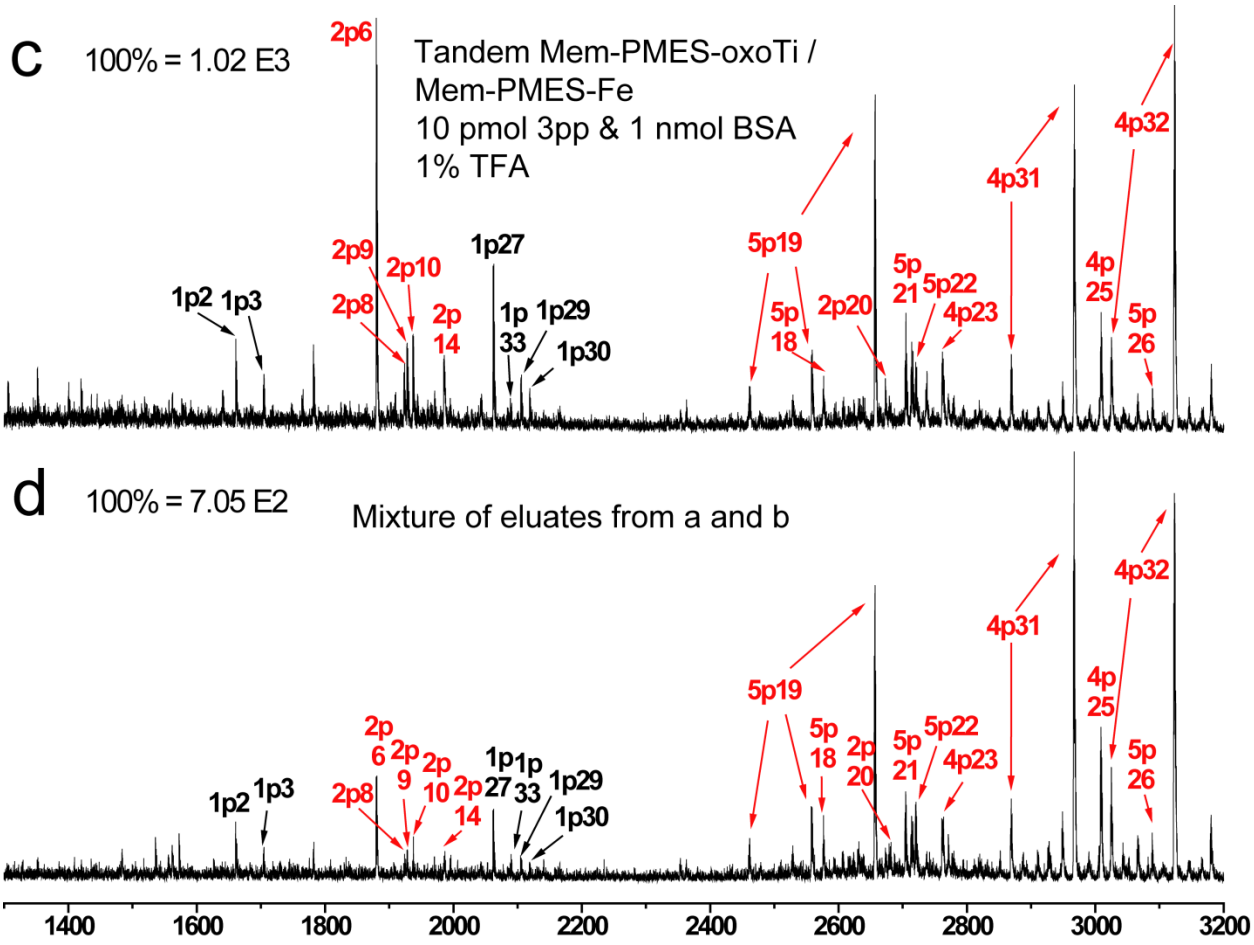


Figure 4.13. (a-c) MALDI mass spectra after phosphopeptide enrichment from 200-μL of 1% TFA, 50% ACN containing 10 pmol each of three phosphoprotein digests (α-casein, β-casein and ovalbumin), and 1 nmol of BSA digest. The membrane used for enrichment is noted in each spectrum. Figure (d) is a mixture of 1 μL of eluate from (a) and 1 μL of eluate from (b). Each phosphopeptide is labeled as #p#, where the first # denotes the number of phosphates in that peptide and the second # denotes the peptide number in Table 4.2. Black and red colors represent mono- and multi-phosphorylated peptides, respectively.

Figure 4.13 cont'd



the recovered mono- and multi-phosphorylated peptides prior to MS analysis would probably improve analysis of the two classes of phosphopeptides. In contrast, the tandem enrichment leads to more evenly distributed signals for mono- and multi-phosphorylated peptides (Figure 4.13c), probably due to the strong binding of a portion of multi-phosphorylated peptides to Mem-PMES-oxoTi (the first membrane), which leads to enrichment of only a portion of the multiply phosphopeptides in Mem-PMES-Fe (the second membrane). Final elution with 1% H_3PO_4 is probably sufficient to release multi-phosphorylated peptides from NTA-Fe groups but not from oxoTi. Considering that NTA-Fe(III) does not capture mono-phosphorylated peptides efficiently during sample loading and washing with highly acidic solutions(1% TFA), reversing the membrane sequence in the set-up might improve the tandem enrichment.

In summary, tandem membrane enrichment separates the capture and the elution of phosphopeptides from the two different materials in space and time. The recovery of mono- and multi-phosphorylated peptides may change as the flow rate or membrane stacking sequence changes. Such a tandem strategy should provide a more comprehensive and efficient enrichment of phosphopeptides from complex protein digests samples than enrichment using a single affinity material. However, the tandem enrichment strategy will produce a more complex mixture of phosphopeptides, and separation (e.g. RP-LC) prior to MS may assist analysis.

REFERENCES

REFERENCES

1. Bodenmiller, B.; Mueller, L. N.; Mueller, M.; Domon, B.; Aebersold, R., Reproducible Isolation of Distinct, Overlapping Segments of the Phosphoproteome. *Nat. Methods* **2007,** 4, 231-237.

2. Dunn, J. D.; Reid, G. E.; Bruening, M. L., Techniques for Phosphopeptide Enrichment Prior to Analysis by Mass Spectrometry. *Mass Spectrom. Rev.* **2010,** 29, 29-54.

3. Thingholm, T. E.; Jensen, O. N.; Larsen, M. R., Analytical Strategies for Phosphoproteomics. *Proteomics* **2009,** 9, 1451-1468.

4. Li, Q. R.; Ning, Z. B.; Tang, J. S.; Nie, S.; Zeng, R., Effect of Peptide-to-Tio2 Beads Ratio on Phosphopeptide Enrichment Selectivity. *J. Proteome Res.* **2009,** 8, 5375-5381.

5. Qiao, L. A.; Bi, H. Y.; Busnel, J. M.; Hojeij, M.; Mendez, M.; Liu, B. H.; Girault, H. H., Controlling the Specific Enrichment of Multi-Phosphorylated Peptides on Oxide Materials: Aluminium Foil as a Target Plate for Laser Desorption Ionization Mass Spectrometry. *Chem. Sci.* **2010,** 1, 374-382.

6. Moser, K.; White, F. M., Phosphoproteomic Analysis of Rat Liver by High Capacity IMAC and LC-MS/MS. *J. Proteome Res.* **2006,** 5, 98-104.

7. Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R., Simac (Sequential Elution from IMAC), a Phosphoproteomics Strategy for the Rapid Separation of Monophosphorylated from Multiply Phosphorylated Peptides. *Mol. Cell. Proteomics* **2008,** 7, 661-671.

8. Wilson-Grady, J. T.; Villen, J.; Gygi, S. P., Phosphoproteome Analysis of Fission Yeast. *J. Proteome Res.* **2008,** 7, 1088-1097.

9. Carrascal, M.; Ovefletro, D.; Casas, V.; Gay, M.; Ablan, J., Phosphorylation Analysis of Primary Human T Lymphocytes Using Sequential IMAC and Titanium Oxide Enrichment. *J. Proteome Res.* **2008,** 7, 5167-5176.

10. Choi, S.; Kim, J.; Cho, K.; Park, G.; Yoon, J. H.; Park, S.; Yoo, J. S.; Ryu, S. H.; Kim, Y. H.; Kim, J., Sequential Fe3O4/Tio2 Enrichment for Phosphopeptide Analysis by Liquid Chromatography/Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2010,** 24, 1467-1474.

11. Shah, R. R.; Merreceyes, D.; Husemann, M.; Rees, I.; Abbott, N. L.; Hawker, C. J.; Hedrick, J. L., Using Atom Transfer Radical Polymerization to Amplify Monolayers of Initiators Patterned by Microcontact Printing into Polymer Brushes for Pattern Transfer. *Macromolecules* **2000,** 33, 597-605.

12. Edmondson, S.; Huck, W. T. S., Controlled Growth and Subsequent Chemical Modification of Poly(Glycidyl Methacrylate) Brushes on Silicon Wafers. *J. Mater. Chem*. **2004,** 14, 730-734.

13. Slater, M.; Snauko, M.; Svec, F.; Frechet, J. M. J., "Click Chemistry" In the Preparation of Porous Polymer-Based Particulate Stationary Phases for Mu-Hplc Separation of Peptides and Proteins. *Anal. Chem*. **2006,** 78, 4969-4975.

14. Luo, Q. Z.; Zou, H. F.; Xiao, X. Z.; Guo, Z.; Kong, L.; Mao, X. Q., Chromatographic Separation of Proteins on Metal Immobilized Iminodiacetic Acid-Bound Molded Monolithic Rods of Macroporous Poly(Glycidyl Methacrylate-Co-Ethylene Dimethacrylate). *J. Chromatogr. A* **2001,** 926, 255-264.

15. Roberts, M. W. H.; Ongkudon, C. M.; Forde, G. M.; Danquah, M. K., Versatility of Polymethacrylate Monoliths for Chromatographic Purification of Biomolecules. *J. Sep. Sci.* **2009,** 32, 2485-2494.

16. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem.-Int. Edit*. **2001,** 40, 2004-2021.

17. Hein, C. D.; Liu, X. M.; Wang, D., Click Chemistry, a Powerful Tool for Pharmaceutical Sciences. *Pharm. Res.* **2008,** 25, 2216-2230.

18. Trost, M.; Bridon, G.; Desjardins, M.; Thibault, P., Subcellular Phosphoproteomics. *Mass Spectrom. Rev*. **2010,** 29, 962-990.

19. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D., Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Mol. Cell. Proteomics* **2005,** 4, 873-886.

Chapter 5. Controlled Proteolysis with Trypsin-modified Membranes to Obtain Large Peptides for Mass Spectrometry

(The work described in this chapter was performed in collaboration with Ken Chanthamontri from Prof. Scott McLuckey's group in the Department of Chemistry at Purdue University, and with Dr. Scott Smith from Prof. Reid's group in the Department of Chemistry at Michigan State University.)

5.1 Introduction

Protein analysis using large (2000-15000 Da) proteolytic peptides, also known as middle-down proteomics, has several advantages over bottom-up and top-down approaches (see Chapter 1). One of the key challenges in exploiting the assets of middle-down proteomics, however, is controlled generation of the large peptides. Although acid digestion produces large peptides and often increases protein solubility, it suffers from low digestion efficiency and removal of protein modifications.[1,2] Milder and more widely used methods to generate large peptides employ proteases that cleave proteins at only one residue that appears relatively infrequently along the protein chain. Such proteases include Glu-C, Lys-C, and Asp-N (see Chapter 1 for details.) Unfortunately, these high-specificity enzymes are expensive for routine large-scale protein digestion. Moreover, not all proteins have sequences that will yield large peptides.

Trypsin is a relatively inexpensive protease that cleaves proteins at the carboxylic side of Lys and Arg, except when either is followed by Pro.[3] Because the total abundance of Lys and Arg in proteins is relatively high, however, tryptic digestion generates mostly small peptides (typically below 3000 Da), which are useful for bottom-up protein analysis by MS.[4] However,

recent studies suggest that trypsin also generates large peptides with uncleaved Lys or Arg within the peptide sequences when digestion is properly controlled (see Chapter 1 and below).[5,6] In addition, the availability of trypsin in mg quantities allows facile immobilization of this enzyme on a variety of porous materials for applications such as online protein digestion-LC-ESI-MS.[7] This chapter describes a membrane-based kinetic approach for generating large tryptic peptides from substrate proteins. Development of such a method requires some understanding of the tryptic proteolysis mechanism and digestion pathway.

Trypsin is a member of the chymotrypsin clan,[8,9] which is a part of the serine protease family.[9] Based on their three-dimensional structure, serine proteases can be grouped into several clans. The proteases of the chymotrypsin, subtilisin and carboxypeptidase C clans all contain a "catalytic triad" that consists of Ser (nucleophile), Asp (electrophile) and His (base). These clans exhibit similar spatial arrangements in different proteases to assist hydrolysis of the substrate protein.[9,10]

Figure 5.1 shows the mechanism of tryptic digestion. Asp189, Ser195 and His57 constitute the catalytic triad of trypsin (and other chymotrypsin-like proteases), which facilitates the cleavage of the amide bond of substrate proteins by water. molecules.[8,10,12] The specificity of trypsin in cleaving certain sites of a substrate protein stems primarily from Asp189, which sits at the base of the protease subsite most proximate to the cleavage site.[12,13] During digestion, the negatively charged Asp189 attracts and stabilizes the positively charged Lys or Arg of the substrate protein near the catalytic pocket where the hydrolysis occurs.

Although the catalytic role of proteases in cleavage of substrate proteins at specific sites is well documented,[9,10,12] the protein digestion pathway (how a single protein molecule gradually degrades to small peptides) is less explored. Nevertheless, controlling the digestion

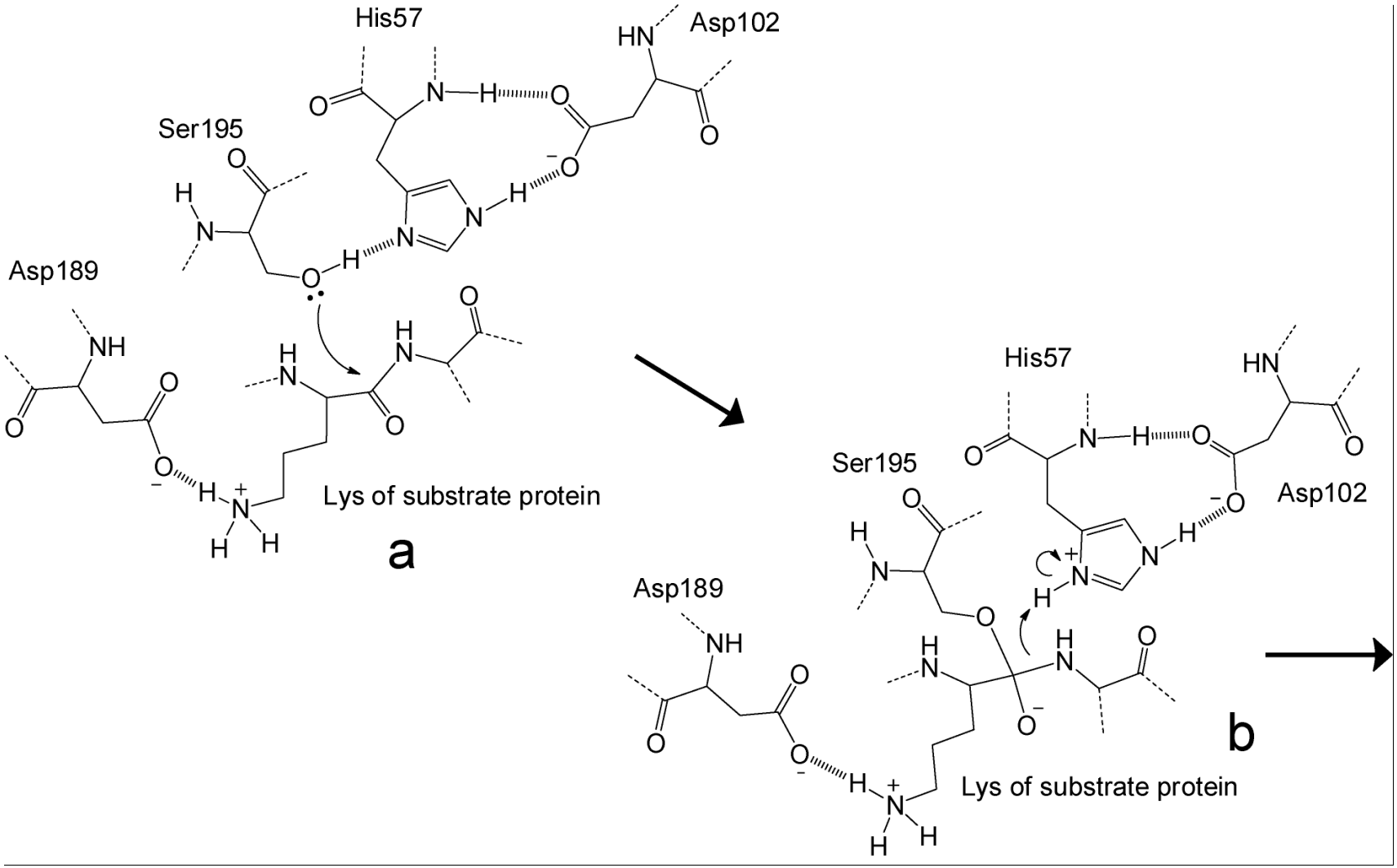


Figure 5.1. Scheme of trypsin-catalyzed proteolysis. Asp189, Ser195 and His57 constitute the catalytic triad of trypsin, and Asp189 of trypsin stabilizes the substrate protein or peptide in the catalytic pocket. The figure is drawn according to references [9,10,12].

Figure 5.1 cont'd

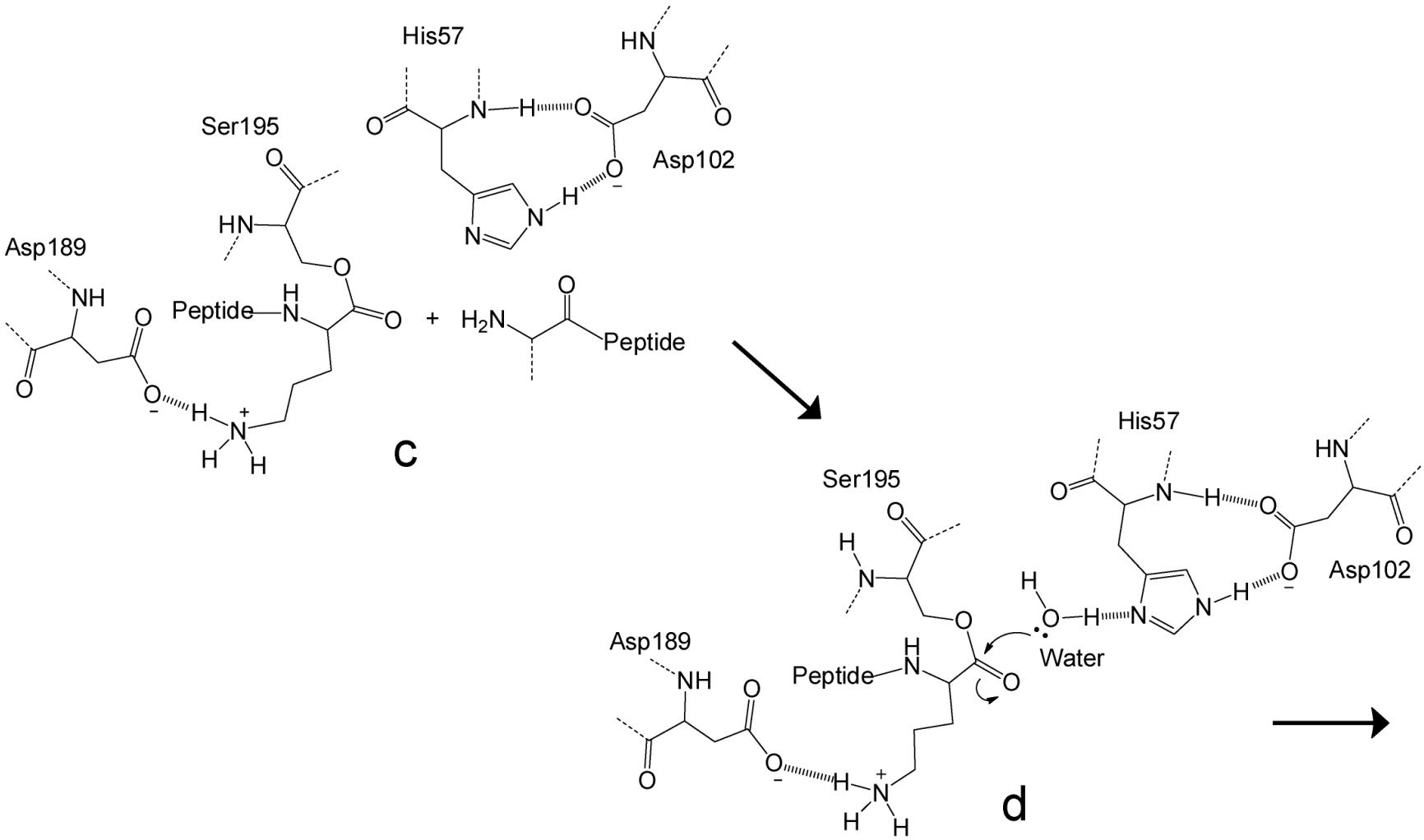
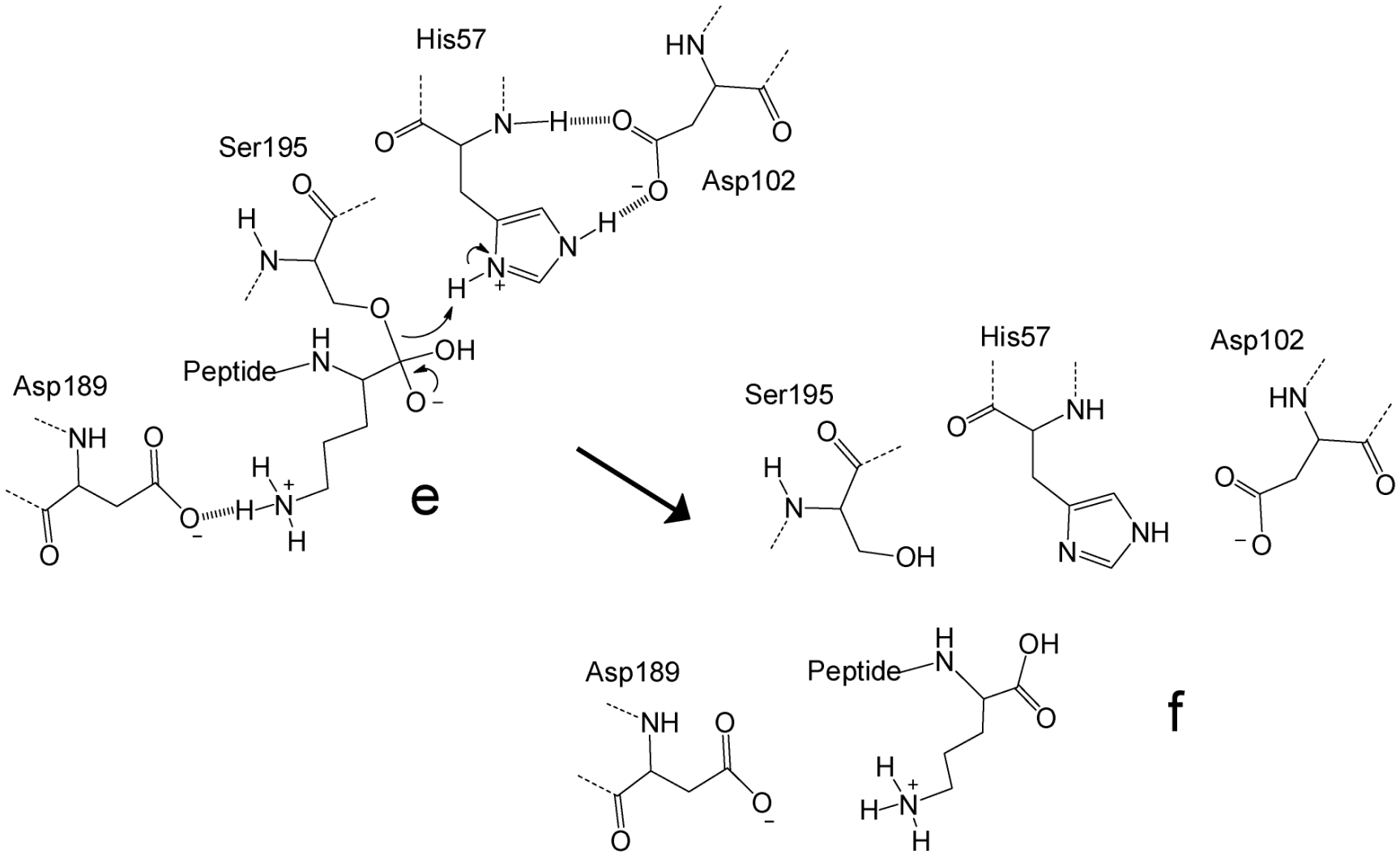
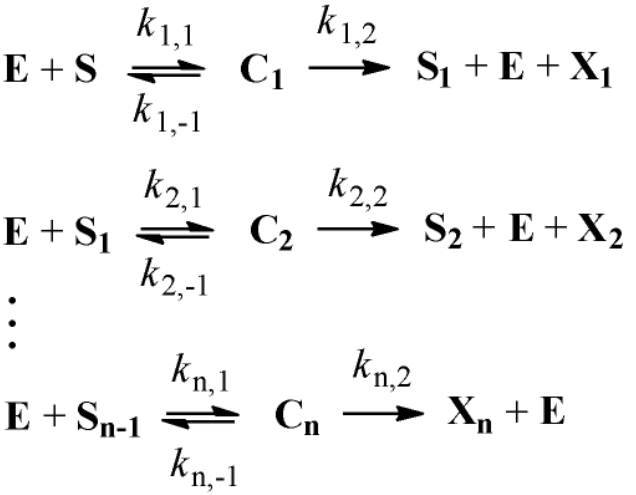


Figure 5.1 cont'd



a. One-by-one scheme



b. Zipper scheme

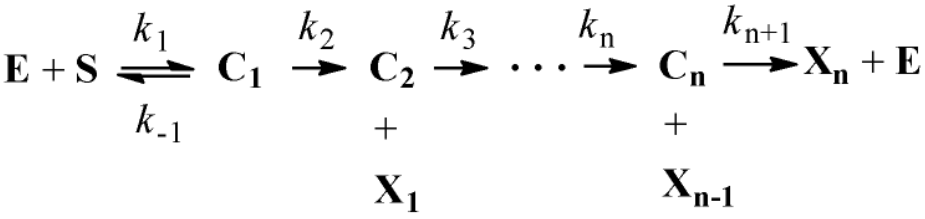


Figure 5.2. Enzyme-catalyzed protein digestion pathways: (a) one-by-one scheme and (b) zipper scheme. **E** represents a protease, **S** and $\mathbf{S_1}$…$\mathbf{S_n}$ represent the substrate protein and its degraded large peptides, respectively. $\mathbf{C_1}$…$\mathbf{C_n}$ represent the complexes of protease with the substrate protein or large peptide, ***k*** represents a rate constant at each step, and $\mathbf{X_1}$…$\mathbf{X_n}$ represent peptides that either contain no missed cleavage sites (final products), or contain one or more missed cleave sites that can undergo further proteolysis. The figure is redrawn according to reference.[11]

pathway is vital to obtaining large peptides for middle-down protein analysis. Previous studies proposed two possible degradation pathways for enzyme-catalyzed proteolysis: a one-by-one scheme and a zipper scheme (Figure 5.2).[14-16]

In the one-by-one pathway (Figure 5.2a), the protease (**E**) combines with the substrate protein (**S**) to form a complex $\mathbf{C_1}$ (e.g. trypsin with the substrate protein stabilized as shown in Figure 5.1a and b). Proteolysis occurs at a single site on the substrate protein, and the resulting peptides ($\mathbf{S_1}$ and $\mathbf{X_1}$) dissociate from the protease. Subsequently the peptide ($\mathbf{S_1}$ **or possibly** $\mathbf{X_1}$) that contains one or more additional cleavage sites (e.g. Lys or Arg for tryptic digestion), undergoes similar degradation to form smaller peptides ($\mathbf{S_2}$ and $\mathbf{X_2}$). Eventually, all the large peptides degrade into small peptides that contain no amino acid residues susceptible to proteolysis.[11]

In the zipper pathway (Figure 5.2b), the complex **C** formed between protease (**E**) and substrate protein (**S**) undergoes sequential proteolysis at different cleavable sites. This process releases intermediate products $\mathbf{X_1}$, $\mathbf{X_2}$,…$\mathbf{X_n}$ without dissociating the remaining part of the protein from the complex. The intermediate peptides could contain missed-cleavage sites.

A first glance at the two digestion pathways suggests that the one-by-one scheme is more favorable for generating large peptides through quenching of proteolysis at an early stage. In principle, arresting digestion after a large peptide dissociates from an enzyme and before formation of a second peptide-enzyme complex will lead to large peptides. In the zipper digestion scheme, however, a substrate protein molecule may undergo multiple cleavages before dissociating from the protease. In a completely zipper mechanism, short-time digestion should generate mixtures of large and small peptides and intact proteins that did not form complexes with enzymes. Overall, the zipper mechanism should provide more small peptides than the one-by-one mechanism because the interaction between the substrate and the enzyme will lead to a longer residence time in the membrane and more cleavage for many peptides.

Choisnard et al. suggest that at pH 4.5 and 23 ℃, pepsin-catalyzed hydrolysis follows the one-by-one scheme for native hemoglobin and the zipper scheme for denatured hemoglobin.[17] As would be expected, the denatured (unfolded) hemoglobin produces a target peptide twice as fast as the one-by-one scheme with the native protein. Denaturation also speeds up hydrolysis in tryptic digestion.[18] The above studies suggest that the rigid, native (folded) proteins may undergo one-by-one proteolysis , in which only a few cleavable sites of the protein are exposed to solution and catalytic pockets. In contrast, denatured proteins have a higher degree of conformational freedom that leads to easier incorporation into the protease-substrate protein complex and a higher probability for the zipper digestion pathway, i.e., digestion at multiple sites without complete dissociation of the enzyme-substrate complex. Although denaturation may favor generation of smaller peptides, considering that biological samples contain proteins with varying degrees of folding, digestion after denaturation (unfolding) may yield greater uniformity of proteolytic peptides from different proteins. Thus, developing novel methods that "quench" the zipper scheme by forcing the isolation of briefly digested protein pieces (including both the released peptides, $\mathbf{X_n}$ in Figure 5.2b, and the unreleased residual protein, not shown in Figure 5.2b, bound to protease **E**) may yield large peptides. However, digestion pathways can also be the mix of the two schemes.[11]

A few experimental studies examined peptide digestion pathways. In a study of peptide hydrolysis in human plasma, quantitative MALDI-MS analysis showed a gradual degradation pathway for fibrinogen peptide A (ADSGEGDFLAEGGGVR). The largest degraded form, DSGEGDFLAEGGGVR, shows its highest abundance at ~1 h; whereas a smaller peptide, EGDFLAEGGGVR. shows its highest abundance at ~4 h).[19] Bi et al. noticed similar gradual degradation of model peptides upon tryptic digestion.[20] Unfortunately, these kinetics studies

employed only small peptides and unknown enzymes that may not be relevant to tryptic digestion of proteins with a MW > 10 kDa.

A recent experimental analysis of peptide degradation in human plasma suggests an apparent first-order depletion of the substrate peptide.[19] In a simulation, Srividhya et al. showed that protein digestion under the one-by-one pathway presents apparent first-order depletion for the substrate protein regardless of the protease-to-substrate protein ratio. In contrast, the zipper pathway allows the apparent first-order protein depletion only when the protease-to-substrate protein ratio is > 1.[11] This important difference suggests a possible change of proteolysis pathway upon increasing the protease concentration in the digestion environment, which may help in the design of novel methods for generating large peptides.

In traditional enzymatic digestion in solution, the protease to substrate protein ratio is low (1/20 ~1/50) to prevent autolysis. If the digestion follows the zipper scheme (Figure 5.2b), a small fraction of the substrate protein binds to the protease and undergoes gradual proteolysis. The rate of protein digestion is low due to the limited availability of trypsin. The problem of slow digestion could also exist in the one-by-one digestion scheme if the reaction rate constant k_2 in the first one or two steps is very low. Another limitation of in-solution digestion in producing large peptides lies in the co-existence of large peptides and intact proteins. Generally, the large peptides produced from parent proteins are more susceptible to further proteolysis than the parent proteins. Thus, enzymatic digestion with a limited amount of protease should favor complete digestion of the pre-formed large peptides instead of producing large peptides from intact proteins. After quenching of the digestion process at an early stage, the majority of the protein molecules will still be undigested. However, extending digestion time will preferentially decrease the size of large peptides rather than forming many new large peptides.

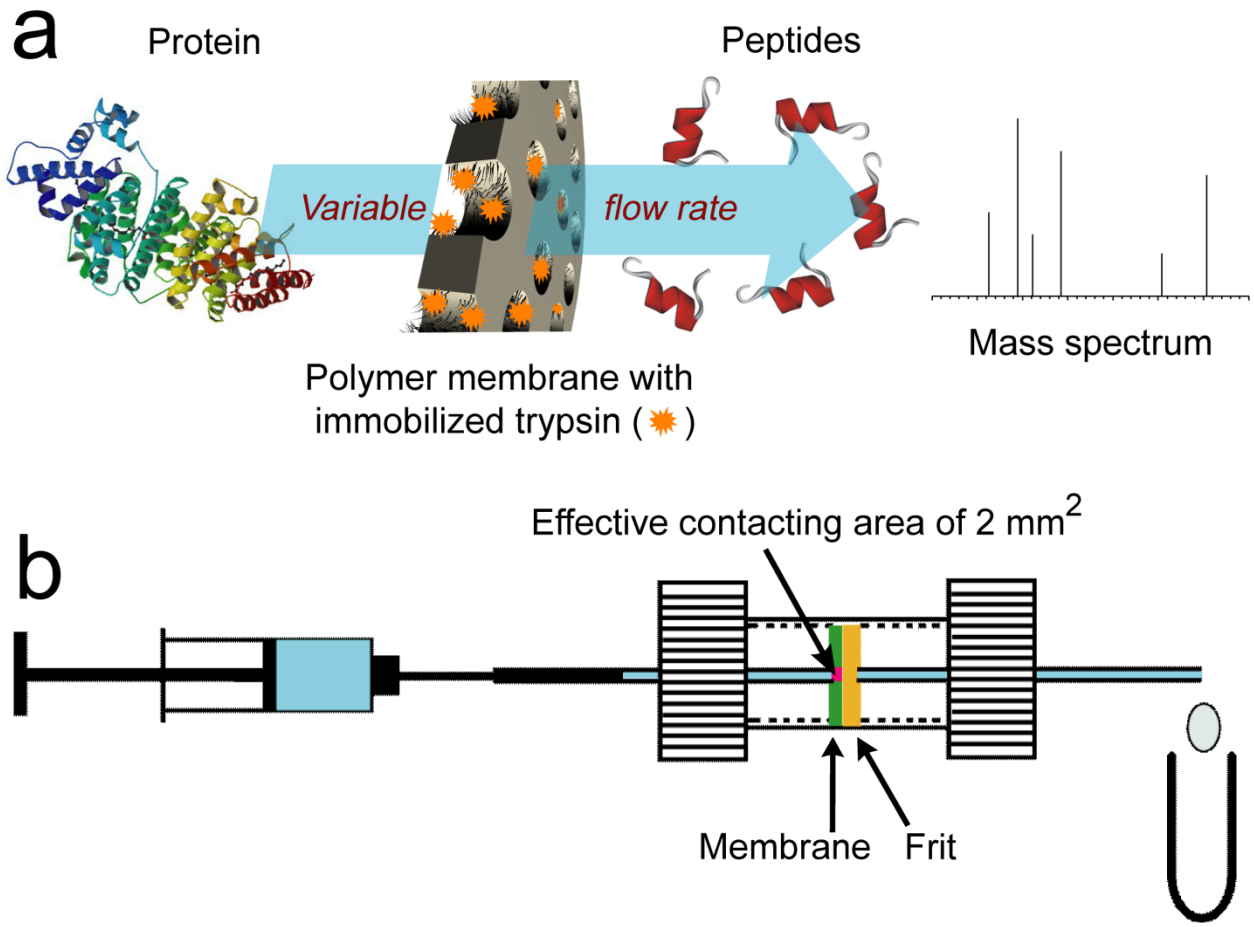


Figure 5.3. Scheme for (a) protein digestion using a nylon membrane containing immobilized trypsin and (b) the assembled setup that performs protein digestion. Both figures are adopted with permission from figures of reference [21] with modifications.

Porous materials with immobilized proteases normally have a much higher local enzyme "concentration" than that employed for in-solution digestion. In fact, enzyme immobilization may provide a protease "concentration" higher than that of the substrate protein. Such a catalytic environment should increase the proteolysis rate, as most of the substrate protein can undergo proteolysis simultaneously. This is also more likely to give large peptides and relatively little intact protein at short times because a large fraction of proteins may be bound to the enzyme.

Whichever pathway (or combination of the two pathways) proteolysis follows, isolation of the large proteolytic peptides from the enzyme can prevent further degradation into smaller

peptides. Our trypsin-modified, microporous nylon membranes provide an ideal platform for generating large peptides (Figure 5.3) because convection can limit contact between substrates and proteases. The trypsin concentration in membrane pores is 450 times higher than that in a typical in-solution digestion, and 9 times higher than the 0.1 mg/mL of substrate proteins passed though the membrane.[21] Moreover, the micron pore size provides a short diffusion distance that allows substrate proteins to encounter trypsin molecules in a short time. Previous bottom-up protein MS analysis showed that α-casein and denatured, alkylated BSA undergo complete tryptic digestion in only 6 s upon passing through the membrane.[21]

This chapter explores proteolysis when the residence time in the membrane is less than a few hundred ms. Such short exposure times are not feasible with in-solution tryptic digestion because of the time required for mixing and quenching of digestion. Thus, many of the large peptides obtained in rapid, membrane-based digestion are not accessible with in-solution digestion. Moreover, employing ion/ion reactions in the gas phase reduces peptide charge states to +1 or +2 and greatly simplifies ESI-MS spectra. The combination of the two novel methods may lead to improved protein identification based on database searching using peptide mass fingerprinting (PMF).

5.2 Experimental

5.2.1 Chemicals and materials

Chemicals and proteins purchased from Sigma-Aldrich include: poly(sodium 4-styrenesulfonate) (PSS) (average molecular weight ~70,000), trypsin (Type I, ~10,000 N_{α}-benzoyl-L-arginine ethyl ester hydrochloride units/mg protein, treated with L-1-tosylamide-2-

phenylethyl chloromethyl ketone, T8802), α-casein from bovine milk (70%, C6780), β-casein from bovine milk (≥98%, C6905), apomyoglobin from horse skeletal muscle (protein sequencing standard, A8673), lysozyme from chicken egg white (≥90%, L6876), dl-1,4-dithiothreitol (DTT) and iodoacetamide (IAA). Acetonitrile (ACN) was obtained from Mallinckrodt Baker. Hydrophilic nylon membrane discs (25 mm diameter, 170 μm thickness, 0.45 μm pore size, HNWP02500, and 1.2 μm pore size, RNWP02500) were obtained from Millipore. Dialysis cassettes (7 kDa MW cut-off, 0.5 mL capacity, Part No. 66373) were obtained from Thermo Scientific. Sep-Pak® Vac 1cc C18 cartridges (100 mg, 55-92 μm beads, Part No. WAT023590) were purchased from Waters. The low-pressure inline filter system (A-424) and Teflon tubing (I.D. 1/32" × 1/64") were acquired from Upchurch Scientific.

5.2.2 Trypsin immobilization in nylon membranes

A UV-ozone cleaned (15 min) nylon membrane disc was inserted in an Amicon cell (Model 8010, 10 mL, Millipore), and 10 mL of 0.02 M PSS, 0.5 M NaCl and 30 mL of water were sequentially pulled through the membrane at 3 mL/min using a peristaltic pump. Trypsin immobilization occurred by circulating 3 mL of 0.6 mg/mL trypsin (prepared in 2.7 mM HCl) through the membrane at 3 mL/min for 1 h followed by rinsing with 30 mL of 1 mM HCl and drying in air.

5.2.3 Protein solutions for tryptic digestion

α-casein (average MW 23694.88 for the S1 chain with 9 reported phosphorylation sites, and average MW 25308.51 for the S2 chain with 12 reported phosphorylation sites), β-casein

(average MW 23983.39, including 5 phosphorylation sites) and apomyoglobin (average MW 16951.62) were dissolved in 10 mM NH_4HCO_3 at 0.1 mg/mL without further treatment.

Lysozyme (average MW 14305.22) was denatured, reduced and alkylated prior to digestion. Typically, 3 mg of lysozyme was dissolved in 600 μL of 10 mM NH_4HCO_3 containing 6 M urea, followed by addition of 150 μL of 10 mM DTT (prepared in 10 mM NH_4HCO_3). The solution was placed in a 55 ℃ water bath for 30 min to cleave the disulfide bonds, mixed with 300 μL of 100 mM IAA, and placed in the dark for 1 h to carbamidomethylate Cys residues.

Part of the above alkylated lysozyme was dialyzed to remove small reagents. In this procedure, 420 μL of the lysozyme solution after alkylation (containing 1.2 mg protein) was mixed with 30 μL of 10 mM NH_4HCO_3 and injected into the dialysis cassette. The cassette was suspended in 300 mL of 10 mM NH_4HCO_3 for 14 h with moderate stirring. The buffer was changed after 2 and 4 h. The resulting solution containing 2.67 mg/mL (assuming no sample loss) of purified, alkylated lysozyme was collected from the cassette and stored in an Eppendorf tube at 4 ℃. Fifty μg of the purified lysozyme was dried and stored at -20 ℃ until ESI-MS analysis. The rest of the dialyzed lysozyme was diluted to 0.14 mg/mL with 10 mM NH_4HCO_3 for tryptic digestion.

The undialyzed, alkylated lysozyme (containing urea, DTT and IAA) was diluted to 0.1 mg protein/mL with 10 mM NH_4HCO_3 for digestion. To examine the completeness of lysozyme alkylation with ESI-MS, 0.5 mL of the 0.1 mg/mL undigested lysozyme was acidified (0.5% final formic acid (FA) concentration) and desalted using a C18 cartridge. The C18 beads

were first washed and equilibrated with acidic solutions by sequentially passing 3 mL of 0.05% FA in 99.95% ACN and 3 mL of aqueous 0.05% FA through the cartridge. The acidified protein solution was then forced through the cartridge followed by 0.5 mL of 0.05% FA for washing. Protein bound to the solid phase was recovered by sequentially passing 0.2 mL of 0.05% FA buffers containing 40%, 60%, 80% and 99.95% ACN through the cartridge. All the above procedures were performed manually with a ~0.7 mL/min flowrate. The combined elution fractions were finally dried with a SpeedVac.

5.2.4 Protein digestion

To perform tryptic digestion using membranes, a 4-mm-diameter membrane was cut from the 25-mm-diameter membrane modified with trypsin, and inserted in an Upchurch low-pressure inline filter system. A frit between the membrane and the downstream tubing (Figure 5.3b) served as a support. The effective filtration area was ~2 mm^2.[21] Prior to digestion 20 μL of 10 mM NH_4HCO_3 was passed through the membrane at 20 μL/min to activate the immobilized trypsin. The protein solution (0.5 mL) was then passed through the membrane at a specific flowrate for digestion, and the digest was collected in an Eppendorf tube and dried with a SpeedVac.

Tryptic digestion of different protein samples was also performed in Eppendorf tubes (in-solution digestion) by addition of 0.5 mg/mL trypsin (prepared in 2.7 mM HCl) to the protein solution at a 1:50 enzyme-to-substrate protein mass ratio. Most of the in-solution digestions were performed at room temperature, but the 16-h digestion occurred at 37 ℃. The digestion was quenched after a certain period of time by addition of glacial acetic acid to achieve 1% acetic acid by volume. The acidified digests were then dried with a SpeedVac. For digestion of

lysozyme without dialysis, the denatured protein was dissolved at 0.1 mg/mL in 0.5% FA, desalted with a C18 cartridge as described above, and finally dried with a SpeedVac.

5.2.5 Gel electrophoresis

To evaluate the completeness of membrane-based protein digestion, 0.5 mL of undigested and digested α-casein (in 10 mM NH_4HCO_3) was dried and reconstituted in gel loading buffer at 1 mg/mL.[22] Ten to twenty μL of each sample (7 μg) was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (gel prepared as a 4% stacking gel and a 12% separation gel) with Coomassie blue staining to reveal any undigested a-casein and large peptides.

5.2.6 ESI-MS analysis and peptide identification

Infusion ESI-MS of protein or protein digests was performed on a Q-TOF mass spectrometer (Agilent G6510) or a modified QqTOF tandem mass spectrometer (QSTAR XL, AB/SCIEX) that allows ion/ion reactions. The mass of each detectable ion that did not overlap with other ions in the spectrum was calculated according to the m/z of the ion and its charge state. For preliminary sequence assignment, those masses were compared with the theoretical average masses of the tryptic peptides generated *in silico* by the ProteinProspector MS-Product program (m/z ±0.5 tolerance). Accession numbers for obtaining sequences of horse myoglobin, bovine β-casein and chick lysozyme in Swiss-Prot are P68082, P02666 and P00698, respectively (signal peptide or initiator methionine removed). For *in silico* peptide generation, the protease was set to trypsin, the maximum missed cleavages was 99, the peptide mass range was 300-30000 Da, and the minimum peptide length was 1. Additionally, carbamidomethylation at Cys was allowed as a constant modification of lysozyme, and phosphorylation at Ser and Thr was set as a variable

modification of β-casein. (Peptides with phosphorylation(s) at sites other than Ser15, Ser17, Ser18, Ser19 and Ser35 were not included in the theoretical peptide candidates).

For ESI-MS performed with the Agilent G6510 (collaborated with Dr. Scott Smith from Prof. Gavin Reid's group at Michigan State University), the ESI source capillary voltage was 3.5 kV in positive ionization mode. Digests of lysozyme (after desalting with C18 cartridges), β-casein and apomyoglobin were reconstituted to 30, 5 and 5 μM, respectively, in 1% FA in 50% methanol for infusion ESI-MS analysis at 3 μL/min. Because infusion ESI mass spectra of digests are often complicated by peptide ions with different charge states, some signals from different peptide ions may overlap and give incorrect m/z values for peptide assignments. To further confirm the "purity" of each peptide ion in the spectra, the theoretical peptide peak with specific amino acid composition, charge states and mass spectrometer resolving power was simulated using IsoPro 3.0. Each experimental peptide peak was compared with the simulated peak (see examples in Figures 7, 11, 20 and 22), and only those with ±0.05 m/z accuracy and similar isotopic peak distributions were confirmed as "identified".

To simplify ESI-MS spectra, ion/ion reactions were performed in a modified QqTOF tandem mass spectrometer to reduce the charge state of peptide ions to mainly +1 (with minor peaks due to +2 species). The experiments and peptide assignment were performed by Ken Chanthamontri from Prof. Scott McLuckey's group in the Department of Chemistry at Purdue University. Briefly, the ESI spectra of ~20-50 μM protein digests containing 50/50/1 (v/v/v) of methanol/water/acetic acid were directly formed by nano-ESI, followed by ion-ion proton transfer reactions in Q2 to facilitate the interpretation of spectra. To examine the sequence coverage of the parent proteins based on the simplified tryptic peptide ions, all the ions were compared to the *in silico* generated peptide ions (as described above). Experimental ions were

considered as +1 charge states for peptide sequence assignment, with the exception of unresolved peaks above m/z 5000, where were considered as +1 or +2 charge states. For protein identification, all the ions in the post-ion/ion-reaction mass spectra with S/N ≥ 5 were considered as peptides generated from an unknown protein, and subjected to MASCOT PMF searching. Parameters for database searching included: database (combination of SwissProt, NCBI and MSDB), protease (trypsin), allowed missed cleavages (9, the maximum available number), taxonomy (mammals for horse myoglobin and bovine milk β-casein, and animals for chick lysozyme), peptide mass tolerance (1 Da), m/z type (average) and protein mass (provided if detectable in the mass spectrum). In addition, fixed modification was set as carbamidomethylation at Cys for lysozyme digests, and variable modifications were set as phosphorylation at Ser/Thr for β-casein digests.

5.3 Results and discussion

5.3.1 Short residence times in trypsin-modified membranes yield large tryptic peptides

In our previous study on α-casein digestion within trypsin-modified membranes, gel electrophoresis showed no intact protein, even with a membrane residence time of only 0.8 sec (residence times are calculated assuming 50% membrane porosity).[21] Moreover, the resulting tryptic peptides generally have <10 amino acids. Nevertheless, because the nylon membranes are only 170 μm thick, a simple syringe pump can force the protein solution through the membrane rapidly enough to achieve a residence time of a few ms. Such a short digestion time

together with the resulting immediate isolation of the digest from trypsin may result in large peptides.

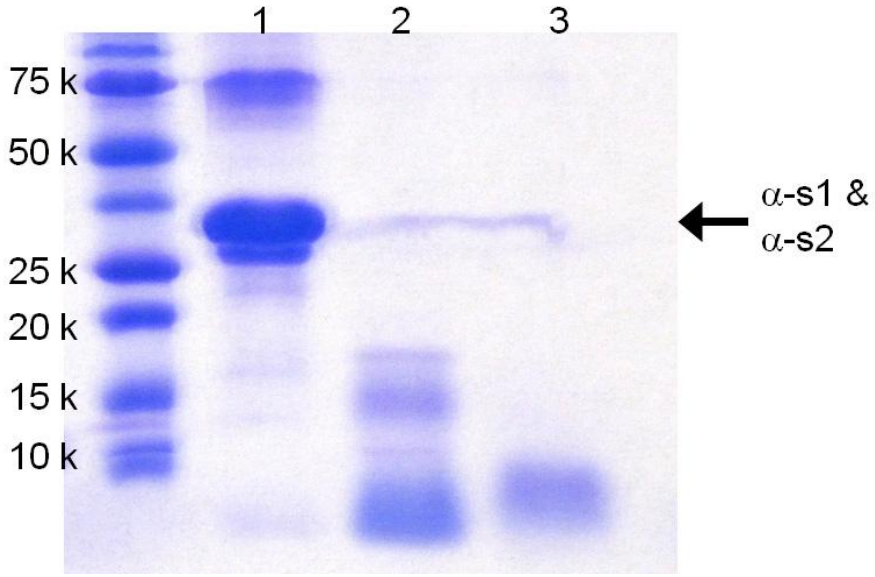


Figure 5.4. SDS-PAGE of intact α-casein (lane 1) and α-casein digests generated by passing 0.1 mg/mL α-casein (in 10 mM NH_4HCO_3) through a trypsin membrane (0.45 μm pore size) using residence times of 1.5 ms (lane 2) and 3.6 ms (lane 3). α-S1 and α-S2 represent the S1 and S2 chains of α-casein, respectively. The loading of protein or protein digests for lanes 1-3 is 7 μg. Gels were stained with Coomassie blue.

Using a trypsin-modified membrane (0.45 μm pore size), we performed α-casein digestion at flowrates of 7.2 and 3 mL/min, which give 1.5-ms and 3.6 ms residence times, respectively, in the membrane. Figure 5.4 suggests that the small difference in residence times leads to a significant difference in peptide size pattern. With a 3.6 ms residence time, most of the S-1 and S-2 chains of α-casein (both ~25 kDa) hydrolyzed to peptides with a MW of ~5 kDa (Figure 5.4, lane 3). In contrast, a 1.5-ms residence time generated several peptides with a MW between 10 and 20 kDa (Figure 5.4, lane 2) without giving rise to an increase in undigested protein (note the similar amount of residual α-casein in lanes 2 and 3 at ~25 kDa). Furthermore, the strong band at ~5 kDa in lane 2 is a little darker than that in lane 3, suggesting that more

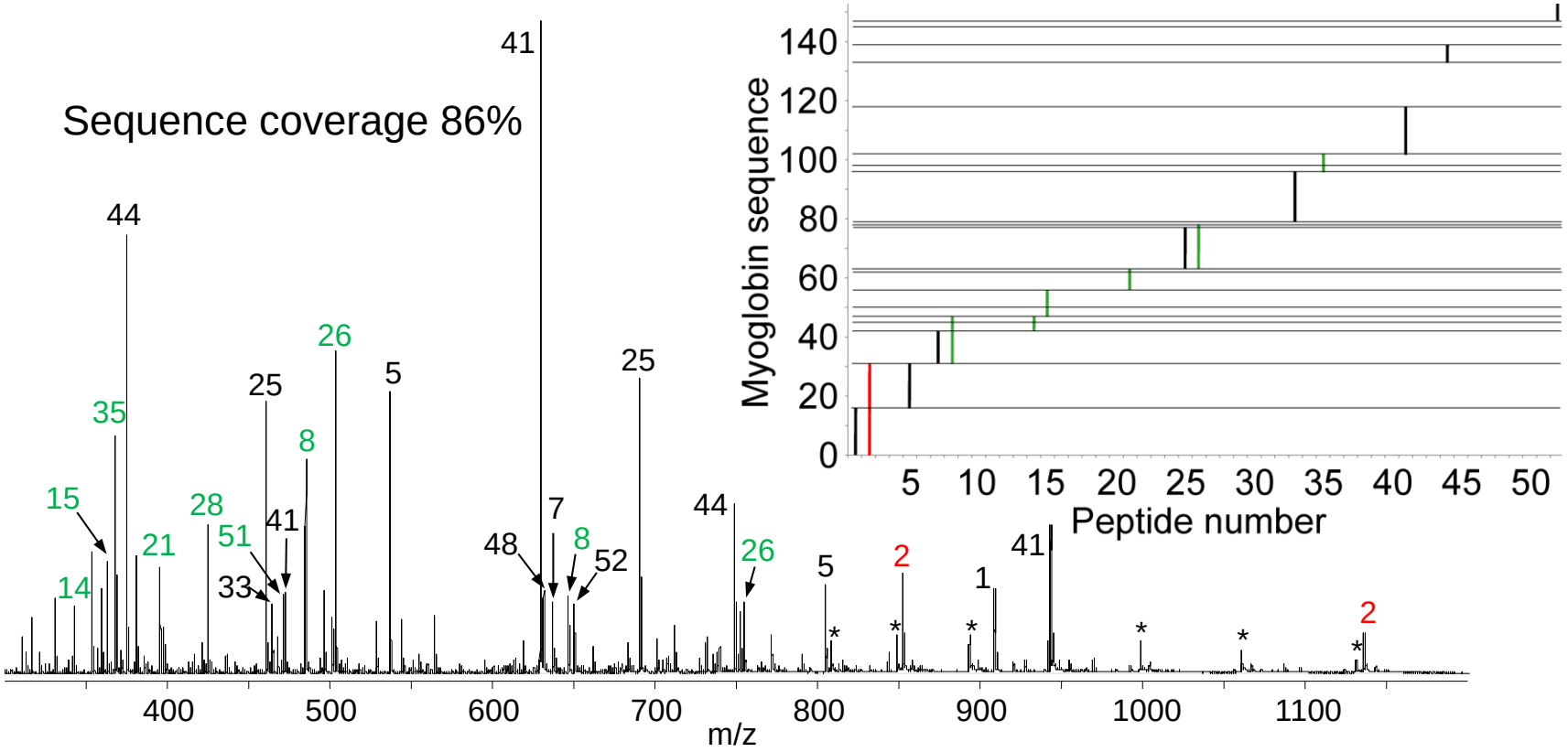


Figure 5.5. ESI-Q-TOF mass spectrum of apomyoglobin digested for 2 h in solution. The black numbers refer to peptides with no missed cleavage sites, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide number refer to Table 5.1.)

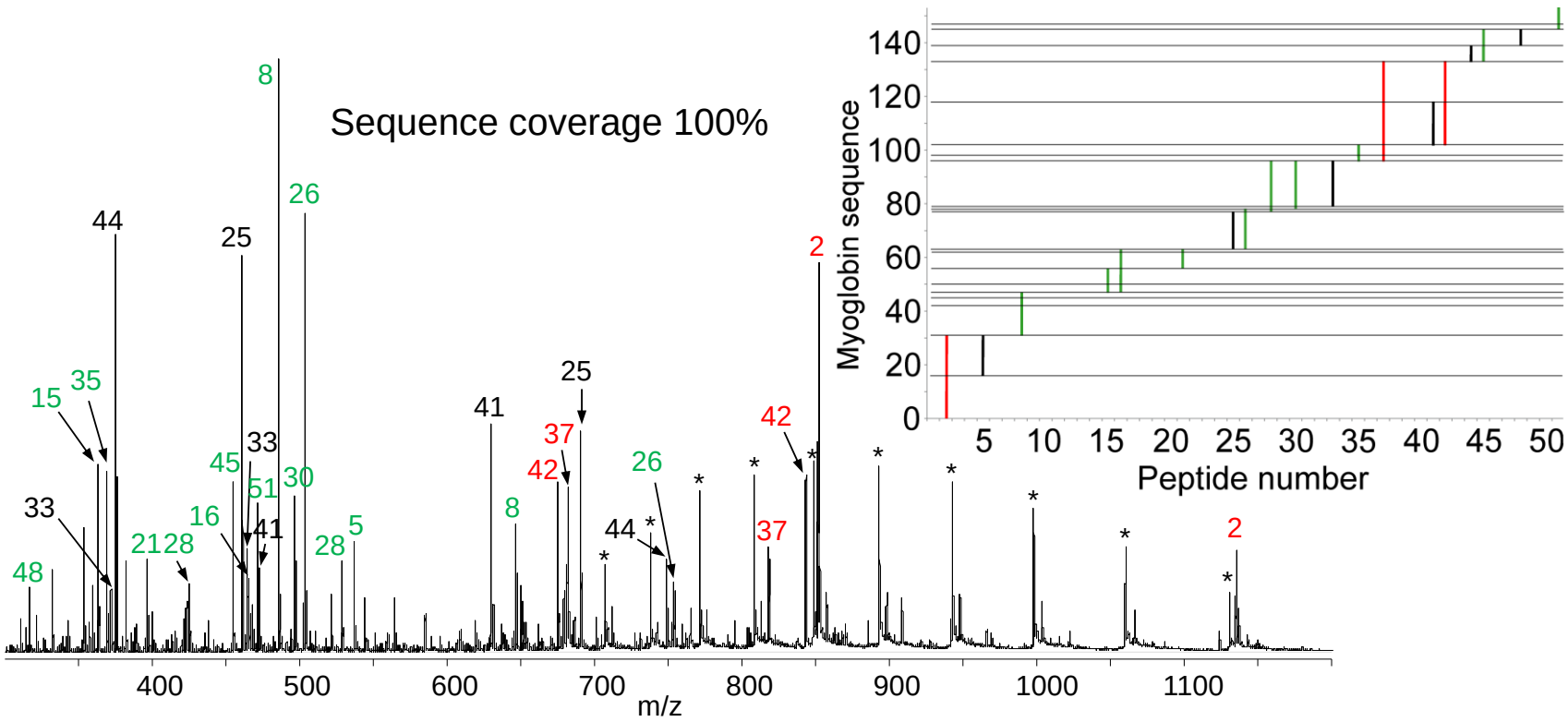


Figure 5.6. ESI-Q-TOF mass spectrum of apomyoglobin digested for 30 min in solution. The black numbers refer to peptides with no missed cleavage site, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.1.)

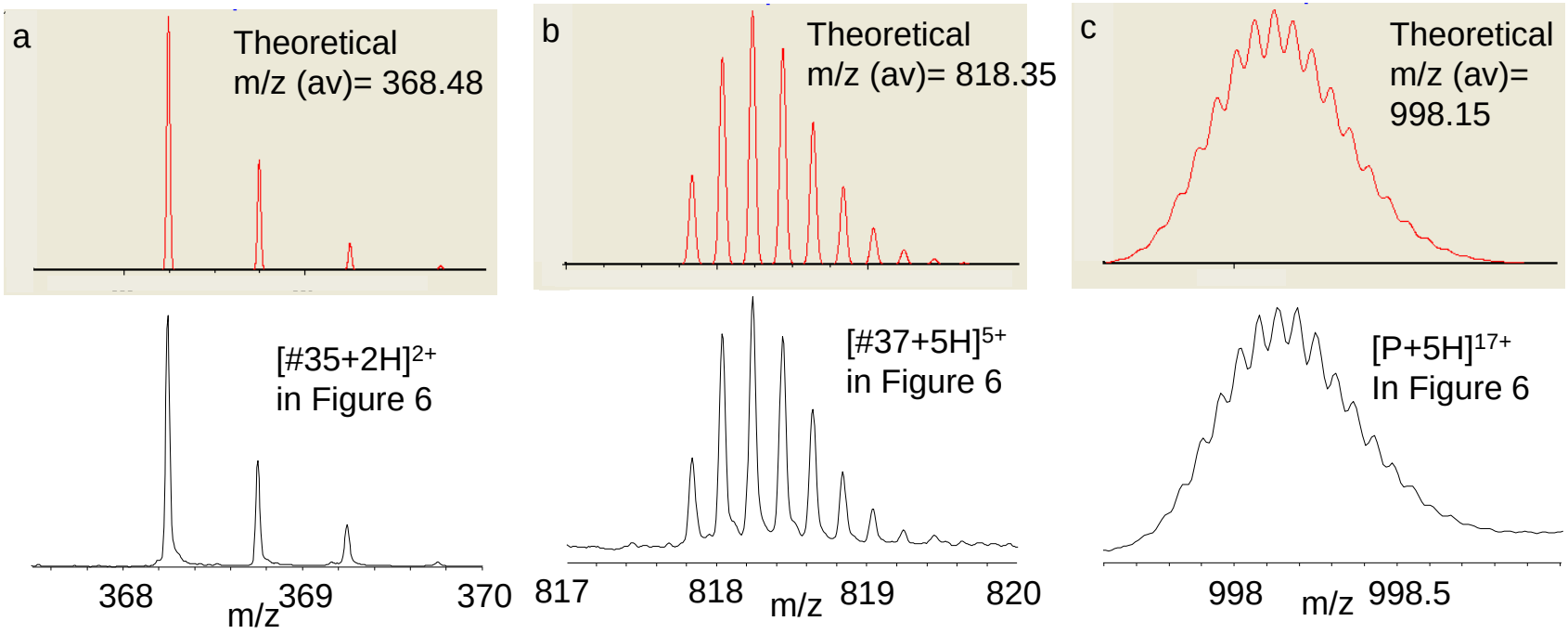


Figure 5.7. Comparison of typical experimental (bottom) and simulated (top) spectra of identified apomyoglobin peptides or protein. Simulations were performed using peptide (or protein) composition and charge state, with the mass spectrometer resolving power set as 18000. The peptide number in Table 5.1 (or protein, labeled as P), theoretical m/z, and the complete mass spectrum for the experimental ion are given in each spectrum.

α-casein degraded to peptides below 5 kDa (not visible in SDS-PAGE) after a 3.6-ms digestion compared to a 1.5-ms digestion. The above results demonstrate that the trypsin-modified membrane digests intact α-casein very efficiently even when the digestion time is a few ms. In addition, variation of the protein residence time in the membrane affords some control over peptide size.

Unfortunately, due to the low purity of the α-casein (70%) and the existence of two chains with similar MWs, infusion ESI-MS gave very complex spectra for α-casein digests. Because many of the large peptides (> 3000 Da) overlapped in the spectra, we could not determine the exact m/z of each peptide. An accurate m/z value is crucial to determine the peptide sequence by comparing the experimental m/z values with the theoretical m/z values of tryptic peptides generated *in silico* from the known sequence of the parent protein.

We then chose sequencing grade apomyoglobin for further examinations of digestion in both solution and membranes. Figure 5.5 shows that 2-h, in-solution digestion extensively hydrolyzes apomyoglobin (17 kDa) and results in a mass spectrum dominated by small peptides with few, if any, missed cleavages. The 15 detected peptides cover only 86% of the protein sequence. Decreasing the in-solution digestion time to 30 min leads to several large peptides, along with 100% sequence coverage from the combination of peptides (Figure 5.6). However, the intensity for signals due to small peptides (black numbers) and undigested protein(*'s) are quite strong, suggesting a low efficiency in producing large peptides. Figure 5.7 shows the isotopic distributions of typical apomyoglobin peptide/protein ions that show no evident overlap with other ions. The high similarity of the experimental and simulated isotope distributions demonstrates the feasibility of using infusion ESI-mass spectra for assigning apomyoglobin peptide sequences.

In contrast to solution digestion, passing the apomyoglobin solution through a trypsin-modified membrane using a 50-ms residence time cleaves the 17 kDa protein into mainly peptides with at least one missed cleavage sites (green and red numbers), many of which are above 2500 Da (red numbers) (Figure 5.8). The combination of only 5 of these large peptides covers the entire protein sequence, in contrast to the need for 8 peptides to cover the whole sequence using 30-min in-solution digestion (Figure 5.6). Moreover, compared to in-solution digestion, proteolysis in the modified membrane gives lower signals for undigested protein.

The above result suggests different apomyoglobin digestion pathways in the trypsin-modified membrane and in solution. For the 30-min in-solution digestion, the combination of signals from highly abundant ions of small peptides, large peptides and intact protein (Figure 5.6) suggests either a partial zipper digestion pathway (Figure 5.2b) or increasing rates of digestion as the sizes of the substrates decrease. Extending the time for in-solution digestion to 2 h depletes most of the intact apomyoglobin. However, the large peptides also hydrolyzed to small peptides (Figure 5.5). In contrast, membrane digestion for 50 ms depletes the majority of apomyoglobin, and the prompt isolation of the large peptides from the digestion system prevents their further hydrolysis (Figures 8).

Because infusion ESI produces peptide and protein ions with different charge states, ESI mass spectra are often complicated, and comparison of the relative intensities of different peptides is difficult. McLuckey's group at Purdue University recently developed gas-phase ion-ion reactions to reduce multiply charged peptide ions to singly or doubly charged ions.[23] Using this technique, the membrane-digested apomyoglobin shows predominantly peptides with missed cleavages, and most of the strong signals come from peptides with masses above 3000 Da (Figure 5.9).

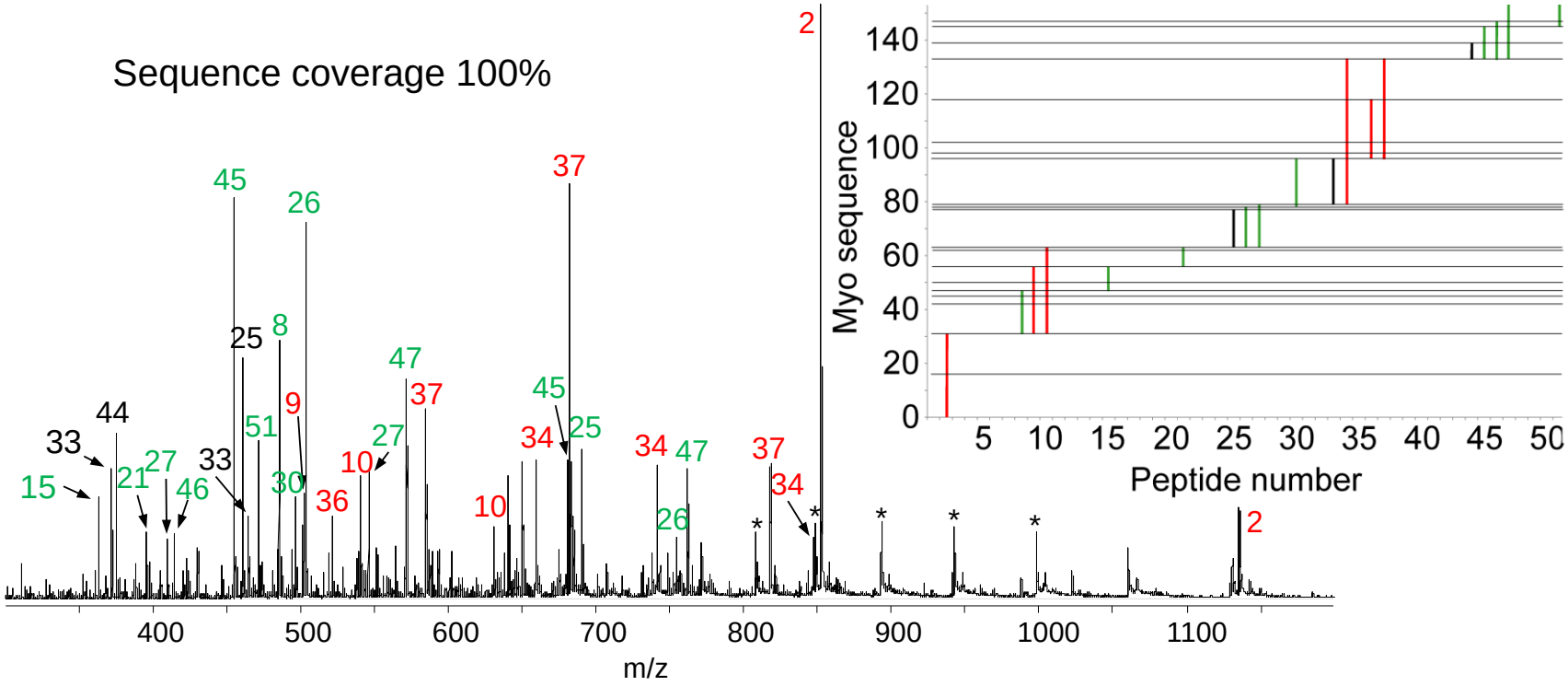


Figure 5.8. ESI-Q-TOF mass spectrum of apomyoglobin digest generated in a trypsin-modified membrane (50-ms residence time, 0.45 μm pore size). The black numbers refer to peptides with no missed cleavage site, the green number refer to peptides with≥1 missed cleavage site but MW < 2500 kDa, the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.1.)

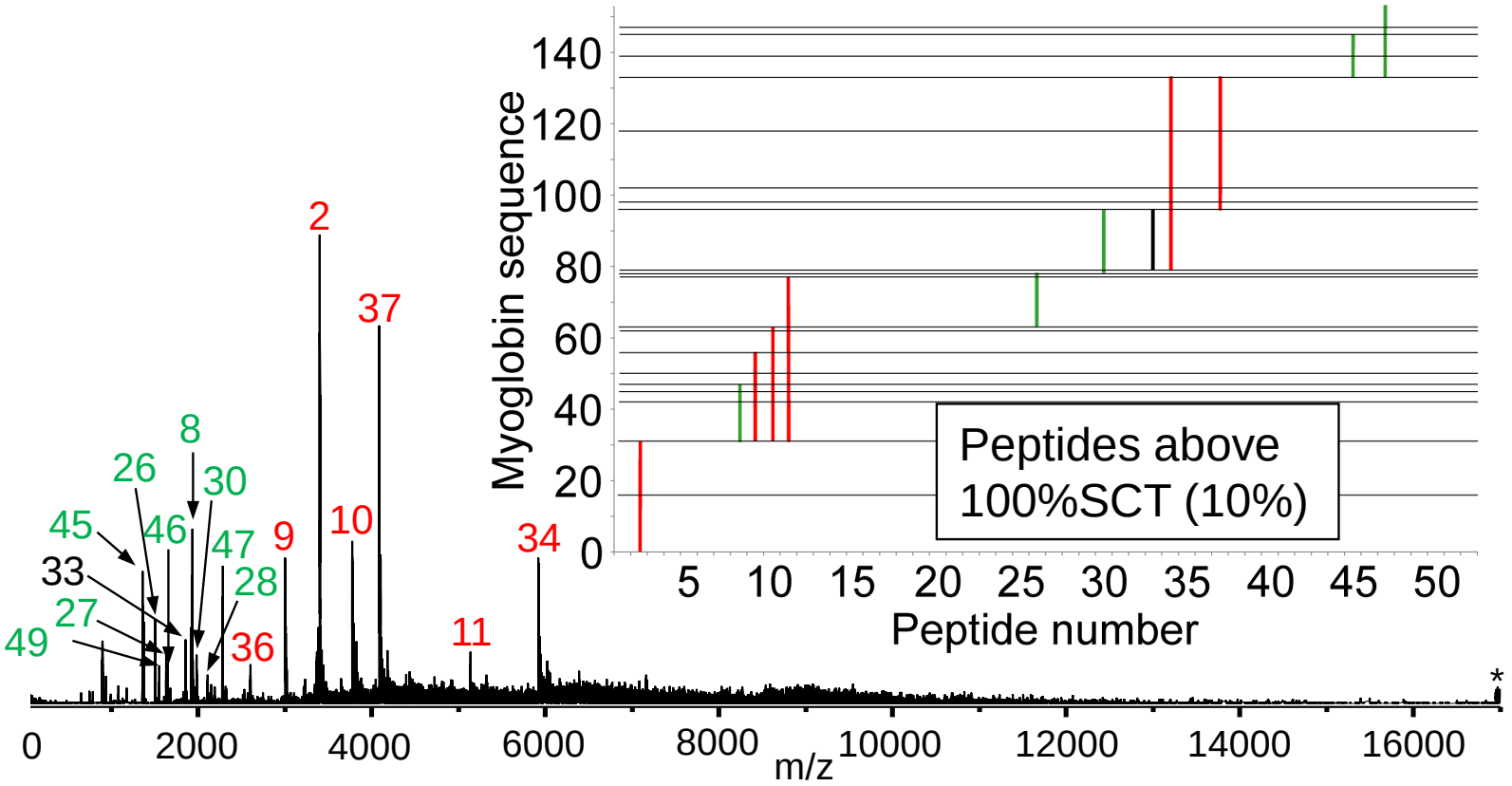


Figure 5.9. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of apomyoglobin digested in a trypsin-modified membrane (50 ms residence time, 0.45 μm pore size). The black numbers refer to peptides with no missed cleavage sites, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptides detected above the 100% sequence coverage threshold (SCT), with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.1)

Table 5.1. Apomyoglobin (average MW 16951 Da) tryptic peptides detected by ESI-MS.

Peptide #	Missed cleavage #	Sequence	Av. m/z $[MH]^+$	Peptide #	Missed cleavage #	Sequence	Av. m/z $[MH]^+$
1	0	1-16	1815.9	27	2	64-79 or 63-78	1635.0
2	1	1-31	3403.7	28	2	78-96	2110.1
3	12	1-96	10611.6	29	6	78-133	6176
4	16	1-133	14677	30	1	79-96	1982.0
5	0	17-31	1606.8	31	5	79-133	6048
6	2	17-45	3249	32	6	79-139	6777
7	0	32-42	1271.6	33	0	80-96	1853.9
8	2	32-47	1937.0	34	4	80-133	5920.0
9	4	32-56	3004.5	35	1	97-102	735.4
10	6	32-63	3775.9	36	2	97-118	2601.4
11	7	32-77	5135.8	37	3	97-133	4085.1
12	8	32-78	5263	38	7	97-153	6349.3
13	18	32-153	13557	39	2	99-133	3819.9
14	1	43-47	684.3	40	6	99-153	6084
15	1	48-56	1086.5	41	0	103-118	1885.0
16	3	48-63	1857.9	42	1	103-133	3368.6
17	5	48-78	3345	43	0	119-133	1502.6
18	7	48-96	5308	44	0	134-139	748.4
19	5	51-79	3095	45	1	134-145	1360.7
20	8	51-102	5647	46	2	134-147	1651.9
21	1	57-63	790.4	47	3	134-153	2283.2
22	2	57-77	2150.2	48	0	140-145	631.3
23	3	57-78	2278.3	49	2	140-153	1553.7
24	6	63-102	4314	50	0	146-147	310.1
25	0	64-77	1378.8	51	1	146-153	941.4
26	1	64-78 or 63-77	1506.9	52	0	148-153	650.3

A good middle-down digestion should hydrolyze the substrate protein to large peptides and avoid generating large quantities of small peptides. Smaller peptides are generally less informative in providing sequence and PTM information if their sequences are already covered by a larger peptide. Moreover, loss of some small peptides during HPLC-ESI-MS results in gaps in protein sequence coverage if only those small peptides cover that portion of the sequence. To evaluate middle-down digestion, we define a "100% sequence coverage peak-intensity threshold"

(100%SCT) for the post-ion/ion reaction mass spectra and determine the number of peptides with intensities above this threshhold. At this intensity cutoff, the detected tryptic peptides just cover the entire protein sequence. For example, in Figure 5.9, the 12 tryptic peptides detected above the 100%SCT (10% relative MS intensity) cover the entire apomyoglobin sequence. Setting the selection level above 10% intensity leads to the loss of at least one tryptic peptide and consequently <100% protein sequence coverage. The fewer tryptic peptides above this threshold, the more efficient the middle-down digestion is.

Similarly, we performed tryptic digestion of β-casein, a 24 kDa protein with 5 phosphorylation sites. A standard 16-h, in-solution digestion produces only 8 small peptides that are detectable by ESI-Q-TOF (Figure 5.10). (Similar isotopic distributions of the identified peptide ions and simulated ions, such as those in Figure 5.11, suggest the feasibility of using this ESI-MS for assigning small β-casein peptide sequences.) These small peptides cover only 20% of the protein sequence, and none of them include the phosphorylation sites. Shortening the in-solution digestion time to 5 min generates much larger peptides and no residual protein ions. However, many of the large peptides with a MW over 10 kDa overlap in the ESI-Q-TOF mass spectrum and cannot be unambiguously assigned. Thus, we performed ion/ion reactions to reduce the charge states of these peptides and generate a mass spectrum with only singly charged peptide ions (Figure 5.12). Although most of the strong signals appear in the 2500-8000 m/z region, the sequence coverage of the total detected peptides is still 98%. The missing piece is (R)INK(K), which consists of amino acid residues 26-28 and has a MW of 374 Da. Interestingly, this small peptide is not detectable in the mass spectrum generated from the 16-h digestion either (Figure 5.10), suggesting either a low ionization efficiency for this small peptide or further degradation during digestion.

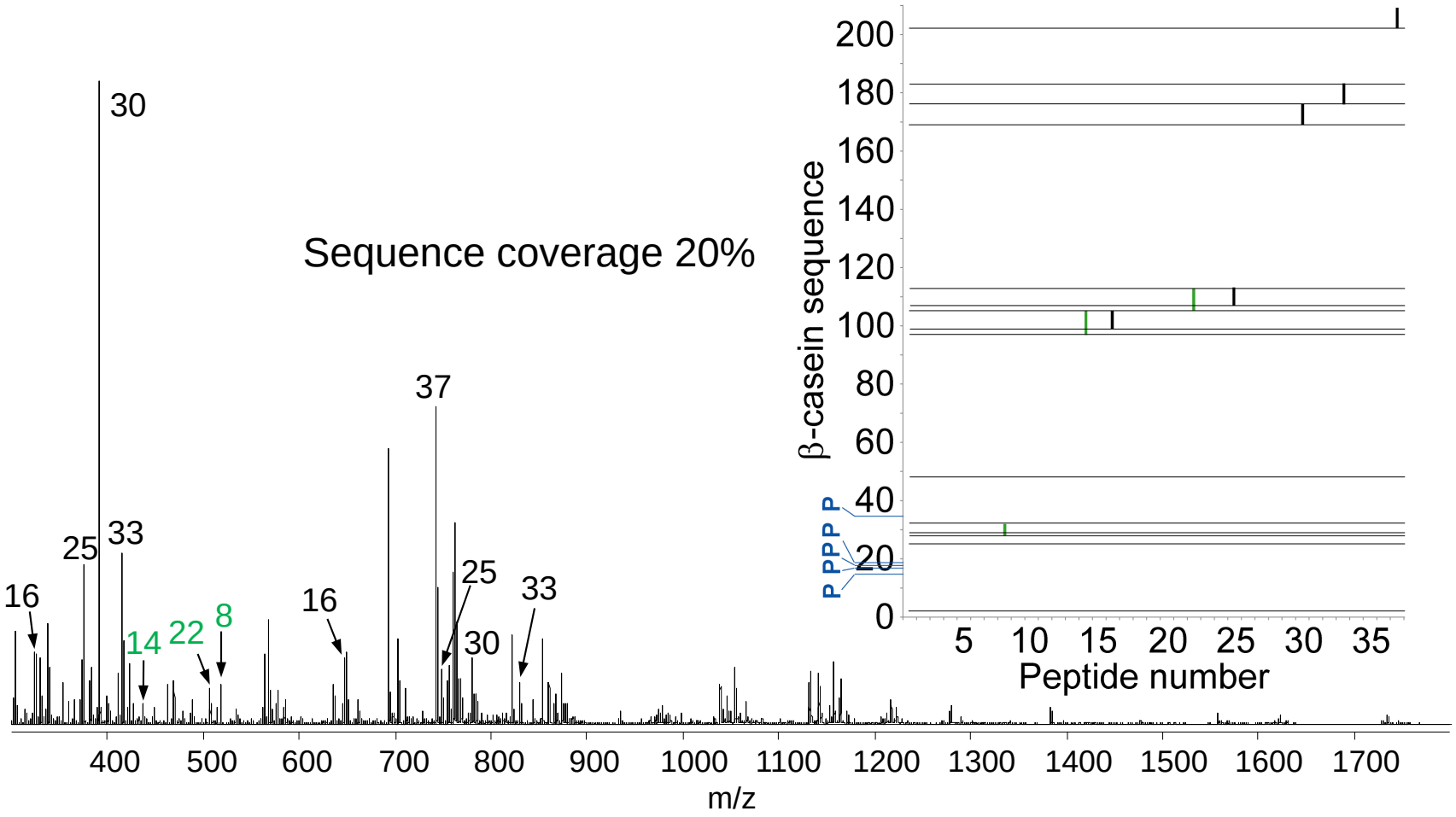


Figure 5.10. ESI-Q-TOF mass spectrum of a β-casein tryptic digest generated by 16-h in-solution digestion. The black numbers refer to peptides with no missed cleavage site, and the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites, and the blue P's showing the phosphorylation sites. (Peptide numbers refer to Table 5.2)

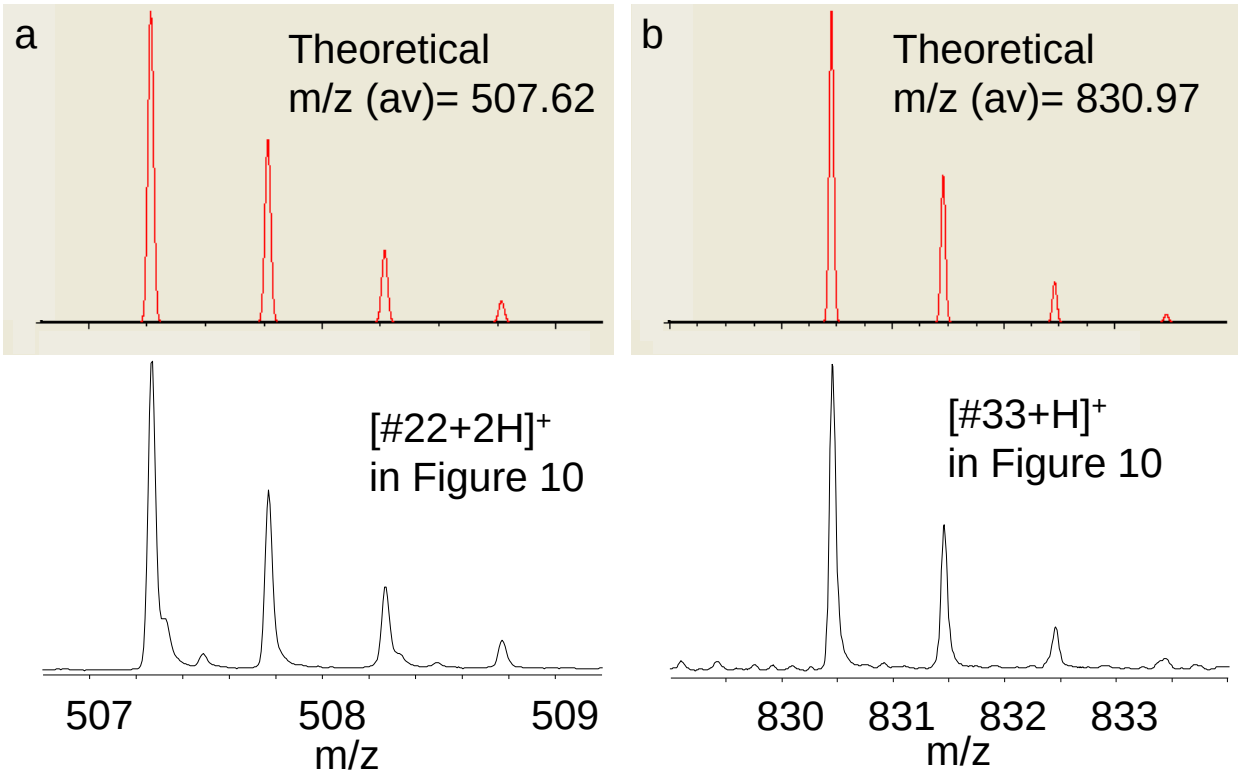


Figure 5.11. Comparison of typical experimental (bottom) and simulated (top) spectra of identified β-casein peptide ions. Simulations were performed using peptide composition and charge state, with the mass spectrometer resolving power set as 18000. The peptide number, theoretical m/z, and the complete mass spectrum for the experimental ion are given in each spectrum.

Further decreasing the in-solution digestion time to 1 min gives 100% sequence coverage, as the post-ion/ion reaction mass spectrum shows (Figure 5.13). Residues 26-28 now appear as part of larger peptides 1-28 (peptide #2) and 26-99 (peptide #7). We expected that the shorter digestion time would yield fewer but larger peptides. However, this brief digestion still generates 14 detectable peptides above 100%SCT, although only 4 peptides can constitute the entire protein sequence.

The minimum time required to perform in-solution digestion in our hands is 20 s. Even after this short period, little intact β-casein remains in the digest. The peptide pattern moves

toward even higher masses (compare Figures 5.13 and 5.14), but the number of tryptic peptides above 100%SCT drops to 13, and only 3 peptides cover the entire protein sequence.

We then compared the short-time, in-solution digestion with membrane digestion. A 2-ms residence time for β-casein in the 0.45 μm membrane generates a similar peptide pattern as the 20-s in-solution digestion (compare Figures 5.14 and 5.15), suggesting the same digestion pathway in both methods. The 2-ms membrane digestion gives 12 peptides above 100%SCT, among which 4 cover the entire protein sequence.

In addition to expanding sequence coverage, another advantage of middle-down protein analysis lies in the simultaneous characterization of many PTMs in a single peptide sequence ("combinatorial" PTM characterization).[24] β-casein is phosphorylated at Ser15, Ser17, Ser18, Ser19 and Ser35. Previous analysis of β-casein phosphorylation based on small tryptic phosphopeptides inevitably required phosphopeptide enrichment from the complete β-casein digest.[25] Due to the existence of one Arg and three Lys residues between Ser35 and the other four phosphorylation sites, simultaneous characterization of the five phosphorylation sites on a single peptide is almost impossible after 16-h, in-solution tryptic digestion. Even with the 20-s, in-solution digestion and 2-ms membrane digestion, post-ion/ion reaction mass spectra without prior phosphopeptide enrichment do not yield detectable peptides that extend across the five phosphorylation sites (Figures 14 and 15). Further increasing the flowrate through the 0.45 μm membrane in an effort to generate such a large peptide was not possible with the syringe pump that we used.

The similar patterns of predominantly large β-casein peptides generated in the membrane and short-time solution digestion suggest that the proteolysis follows the one-by-one scheme in both digestion formats. The high efficiency of intact protein depletion in the two methods

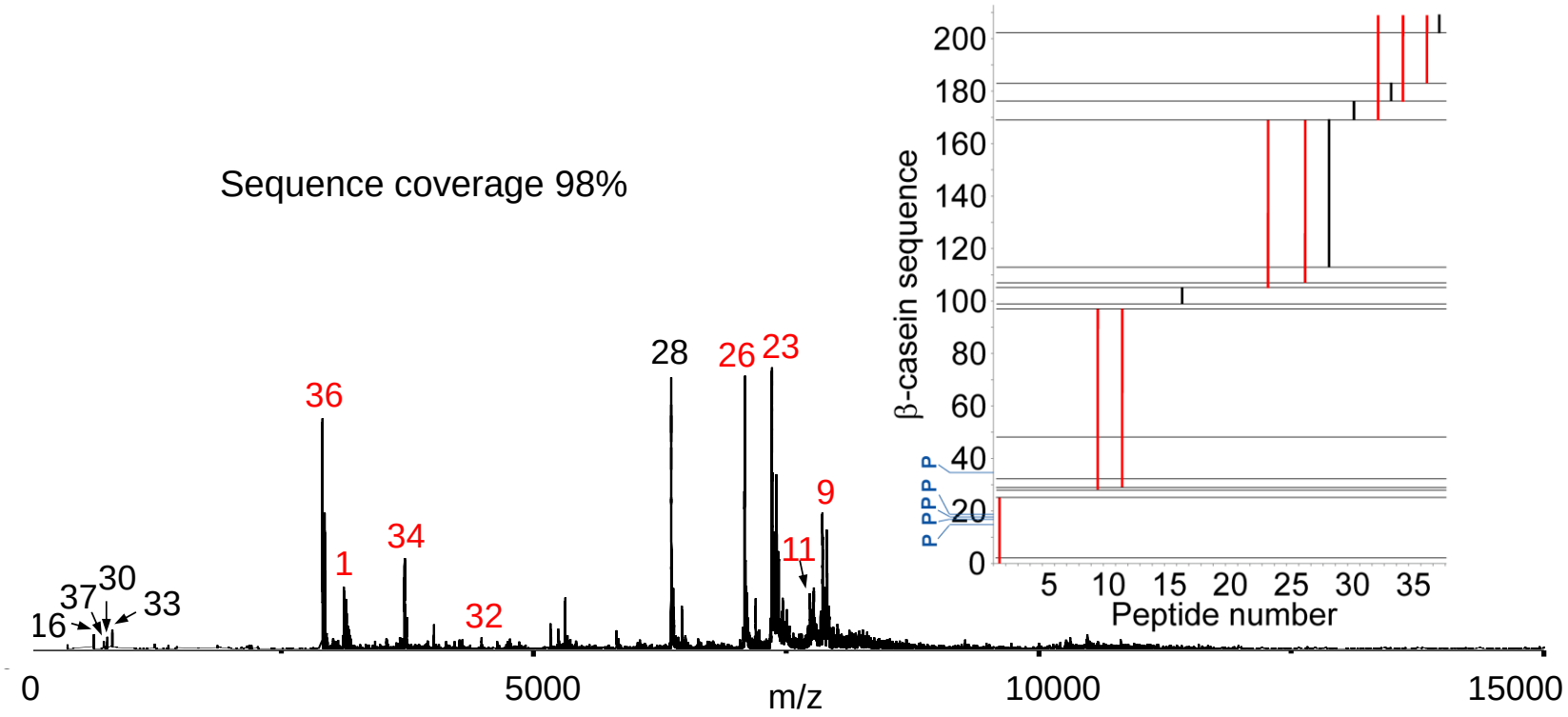


Figure 5.12. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of β-casein digested in solution for 5 min. The black numbers refer to peptides with no missed cleavage site, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites, and the blue P's showing the phosphorylation sites. (Peptide numbers refer to Table 5.2.)

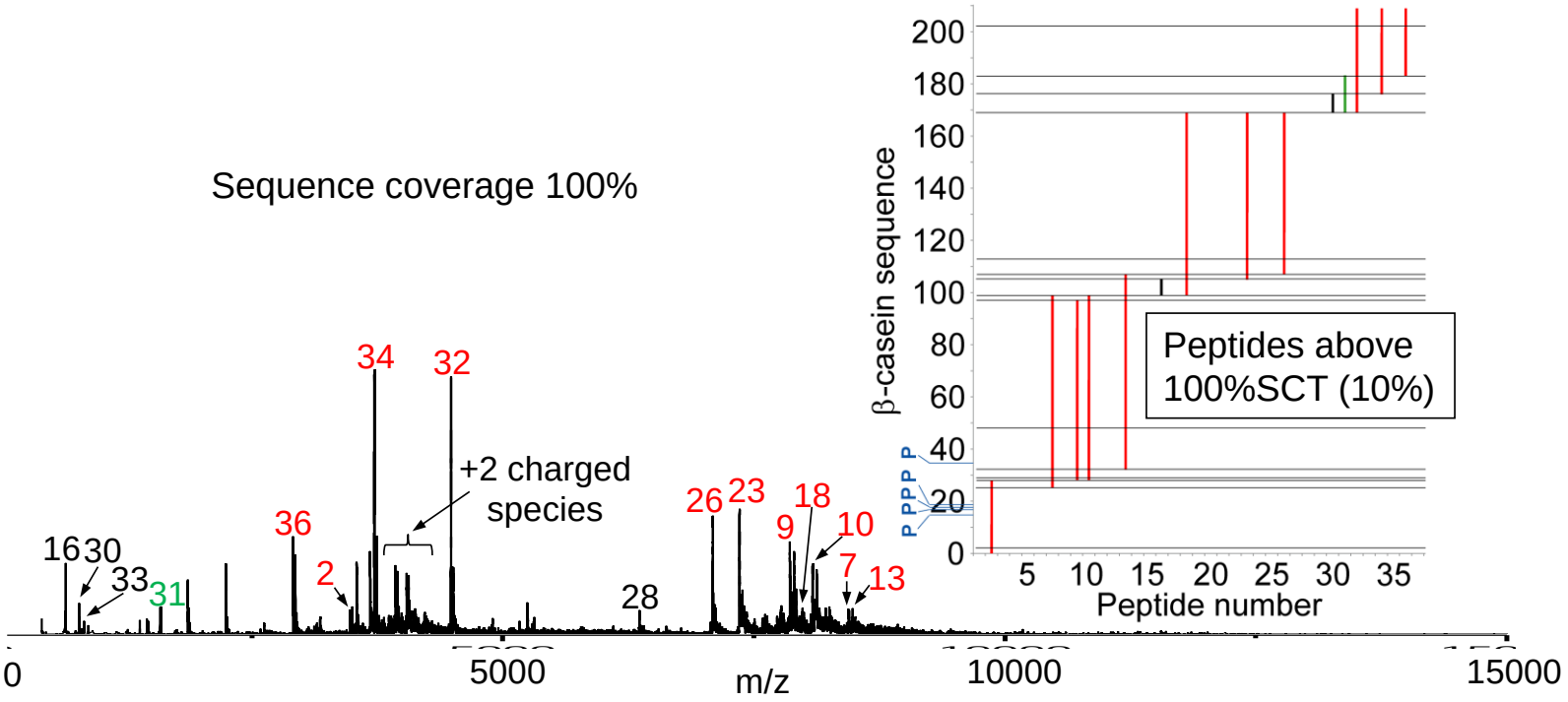


Figure 5.13. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of β-casein digested in solution for 1 min. The black numbers refer to peptides with no missed cleavage site, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa. The inset presents the detected peptides above 100%SCT, with the horizontal lines showing the possible tryptic cleavage sites, and the blue P's showing the phosphorylation sites. (Peptide numbers refer to Table 5.2.)

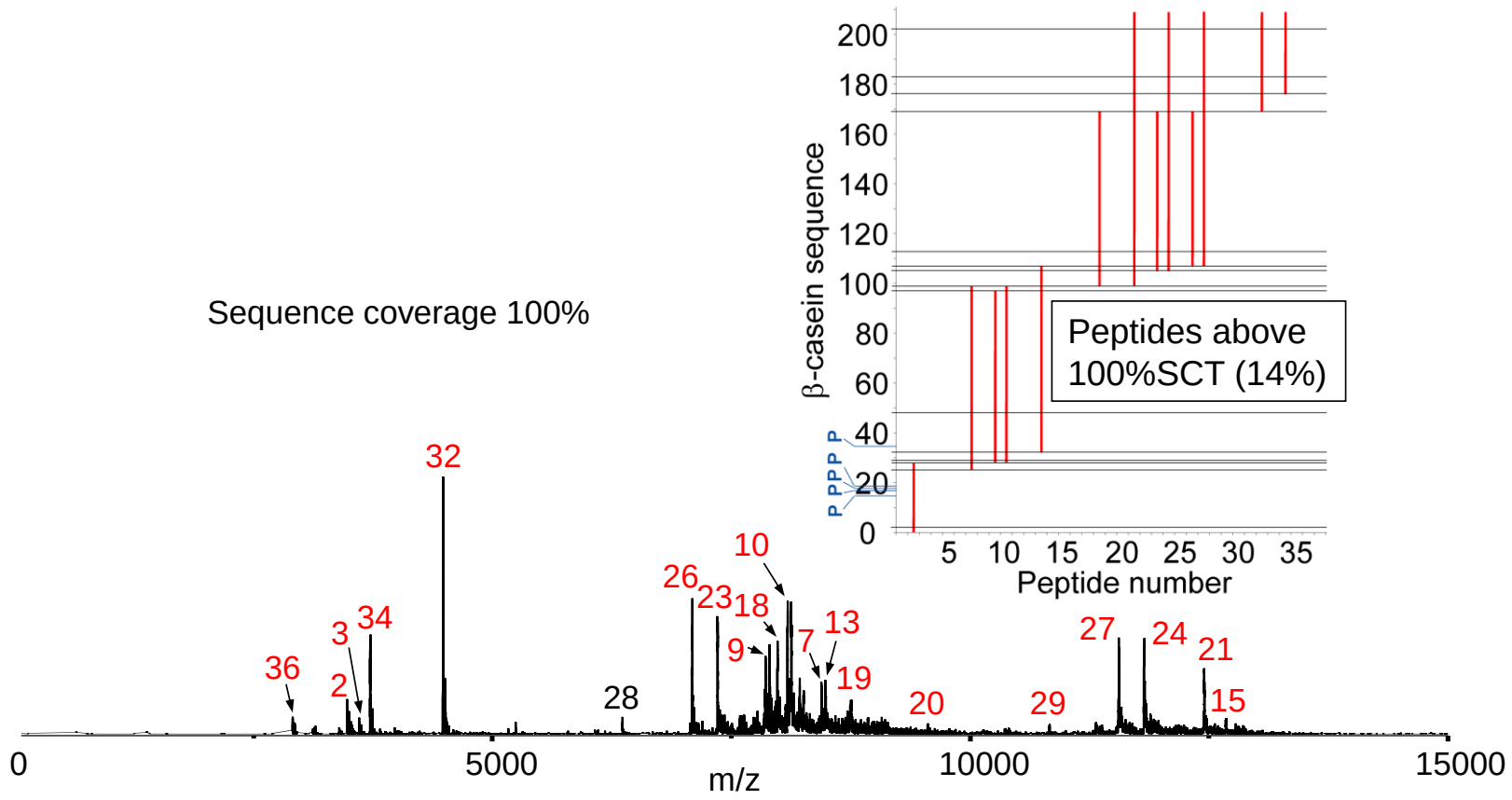


Figure 5.14. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of β-casein digested in solution for 20 s. The black numbers refer to peptides with no missed cleavage site, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa. The inset presents the detected peptides above 100%SCT, with the horizontal lines showing the possible tryptic cleavage sites, and the blue P's showing the phosphorylation sites. (Peptide numbers refer to Table 5.2.)

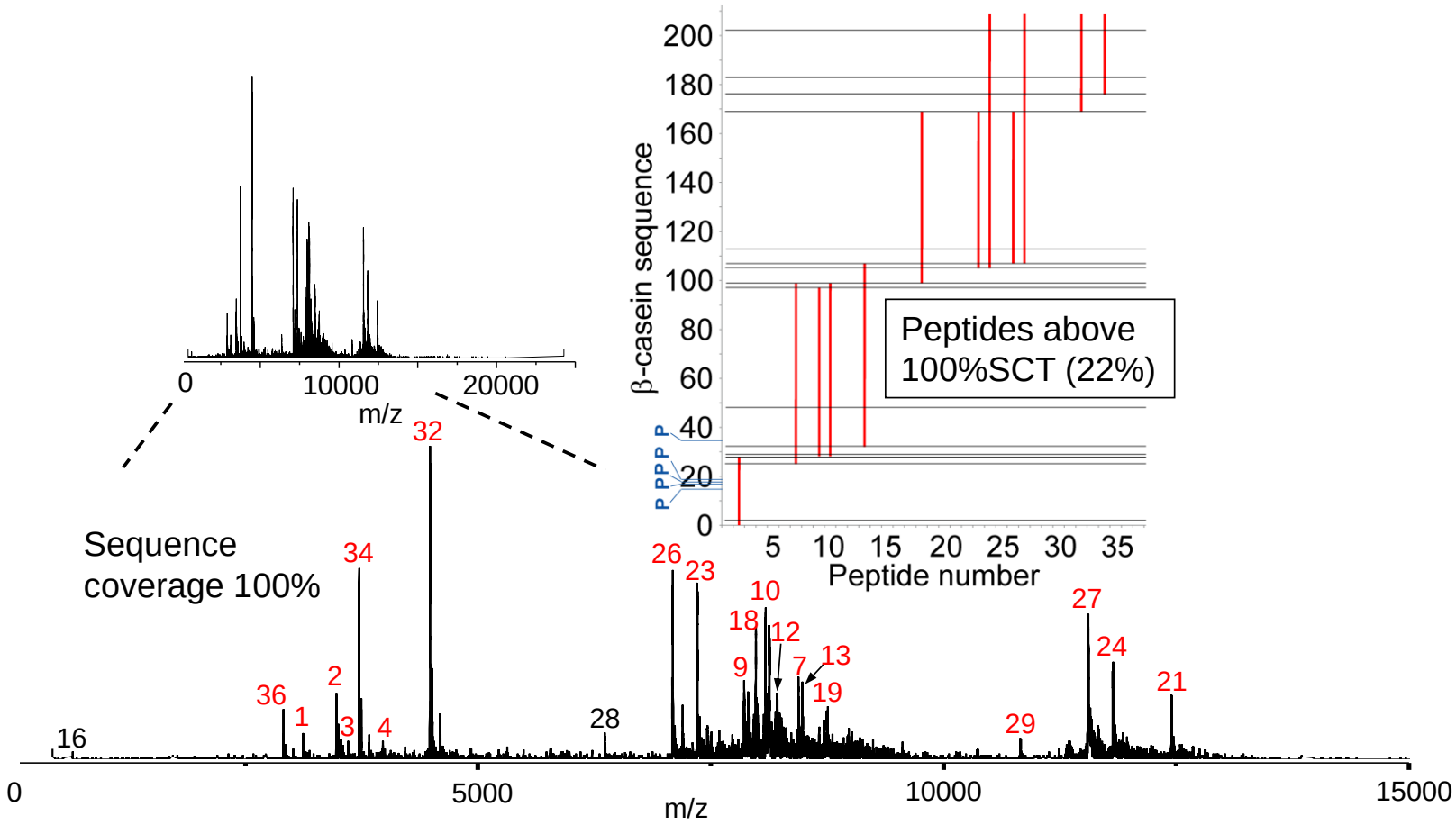


Figure 5.15. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of β-casein digested in a trypsin-membrane membrane (2 ms residence time, 0.45 μm pore size). The black numbers refer to peptides with no missed cleavage site, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa. The inset presents the peptides detected above 100%SCT, with the horizontal lines showing the possible tryptic cleavage sites, and the blue P's showing the phosphorylation sites. (Peptide numbers refer to Table 5.2.)

Table 5.2. β-casein (average MW 23983 Da) tryptic peptides detected by ESI-MS.

Peptide #	# Missed cleavages	Sequence	m/z $[MH]^+$	Peptide #	# Missed cleavages	Sequence	m/z $[MH]^+$
1	1	1-25	3122	20	5	100-183	9559
2	2	1-28	3477	21	7	100-209	12452
3	3	1-29	3605	22	1	106-113	1013
4	4	1-32	3976	23	2	106-169	7358
5	11	1-169	19518	24	6	106-209	11825
6	1	26-29	502	25	0	108-113	748
7	5	26-99	8445	26	1	108-169	7093
8	1	29-32	517	27	5	108-209	11559
9	3	29-97	7863	28	0	114-169	6363
10	4	29-99	8090	29	4	114-209	10829
11	2	30-97	7734	30	0	170-176	780
12	3	33-105	8214	31	1	170-183	1591
13	4	33-107	8485	32	3	170-209	4482
14	1	98-105	873	33	0	177-183	830
15	8	98-209	12680	34	2	177-209	3721
16	0	100-105	646	35	0	184-202	2186
17	1	100-107	911	36	1	184-209	2909
18	3	100-169	7986	37	0	203-209	742
19	4	100-176	8748				

implies both fast digestion of certain labile sites and rapid dissociation of β-casein from trypsin after each single-site proteolysis.

5.3.2 Larger membrane pores allow briefer tryptic digestion

To circumvent the flux limitation in the modified membrane with 0.45 μm pores, we prepared trypsin-modified membranes with 1.2-μm pores. With the larger pore size, a 10-ms residence time generates several peptides with MWs from 4-20 kDa, and undigested β-casein is also visible. The number of peptides above 100%SCT is only 7, and two large peptides can cover the entire sequence (Figure 5.16). Moreover, peptide #5 extends over all the five phosphorylation sites. Diffusion limitations could result in the significant signals from intact

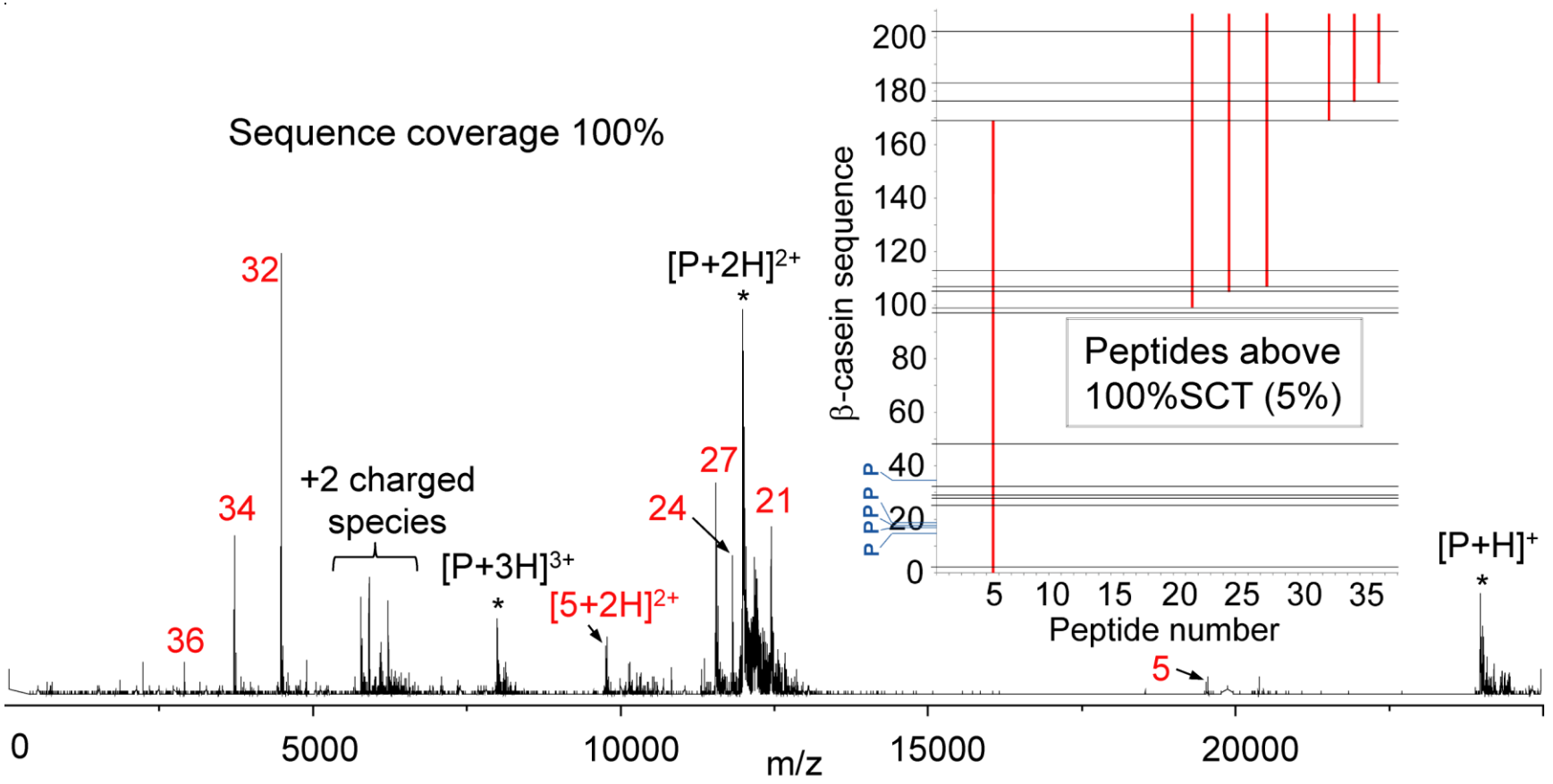


Figure 5.16. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of β-casein digested in a trypsin-membrane membrane (10 ms residence time, 1.2 μm pore size). The red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptides detected above 100%SCT, with the horizontal lines showing the possible tryptic cleavage sites, and the blue P's showing the phosphorylation sites. (Peptide numbers refer to Table 5.2.)

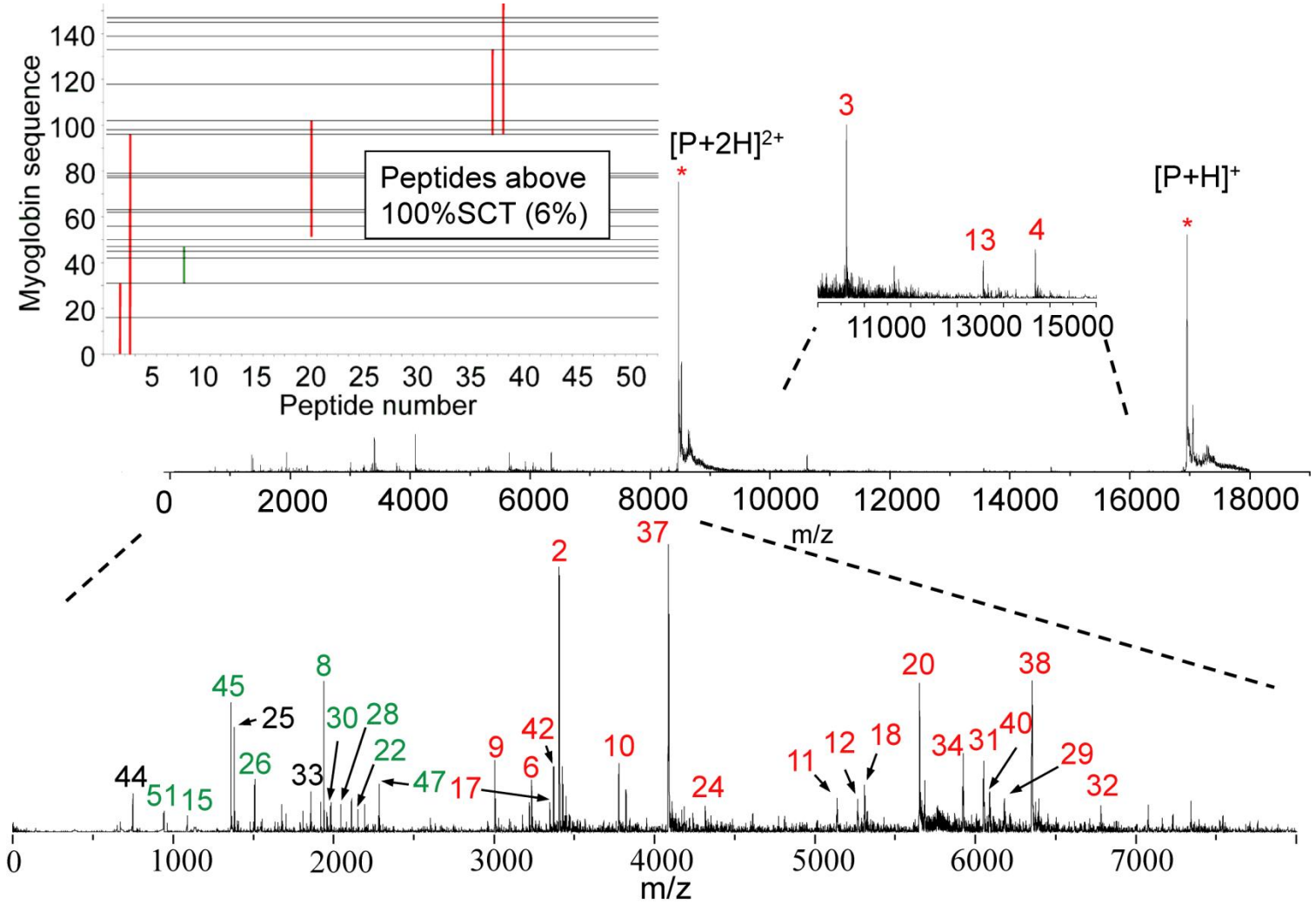


Figure 5.17. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of apomyoblobin digested in a trypsin-membrane membrane (50 ms residence time, 1.2 μm pore size). The black numbers refer to peptides with no missed cleavage site, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptides detected above 100%SCT, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.1.)

protein. In 10 ms, a protein with a diffusion coefficient of 10^{-7} cm^2/sec would diffuse on average around 0.4 μm. Thus some protein molecules may not have time to diffuse from the solution to the enzyme immobilized at the pore walls. This may be especially true for regions with pore diameters on the higher side of the distribution. Additionally, the surface area to volume ratio decreases as the pore size increases, so the 1.2 μm membrane will also contain a lower amount of enzyme than the 0.45 μm membrane. This could lead to more intact protein and larger fragments in the 1.2 μm system.

Similar to β-casein digestion, 50-ms digestion of apomyoglobin in the modified membrane with 1.2 μm pores generates only 6 peptides above 100%SCT (Figure 5.17), in contrast to 12 peptides for the same protein and residence time using a membrane with 0.45 μm pores (Figure 5.9). Among the large peptides generated with the 1.2 μm pores, just two peptides can cover the entire protein sequence, while digestion in the membrane with 0.45 μm pores requires at least 6 peptides for complete coverage in the post ion-ion reaction mass spectrum.

5.3.3 Influence of denaturation on tryptic digestion of lysozyme

Lysozyme is a 14 kDa protein with 4 disulfide bonds. Direct digestion of this protein in 10 mM NH_4HCO_3 using trypsin-modified membranes (6.5 s residence time) resulted in identification of only a few small peptides (data not shown) by MALDI-MS. Reduction of the 4 disulfide bonds of lysozyme followed by carbamidomethylation of the Cys residues unfolds the protein and facilitates protein digestion. The mass spectra in Figure 5.18 show the expected mass increase for alkylation of all 8 cysteines. (For the 8+ species the increase in m/z should be equivalent to the added mass for one alkylation or 58.08 m/z units.)

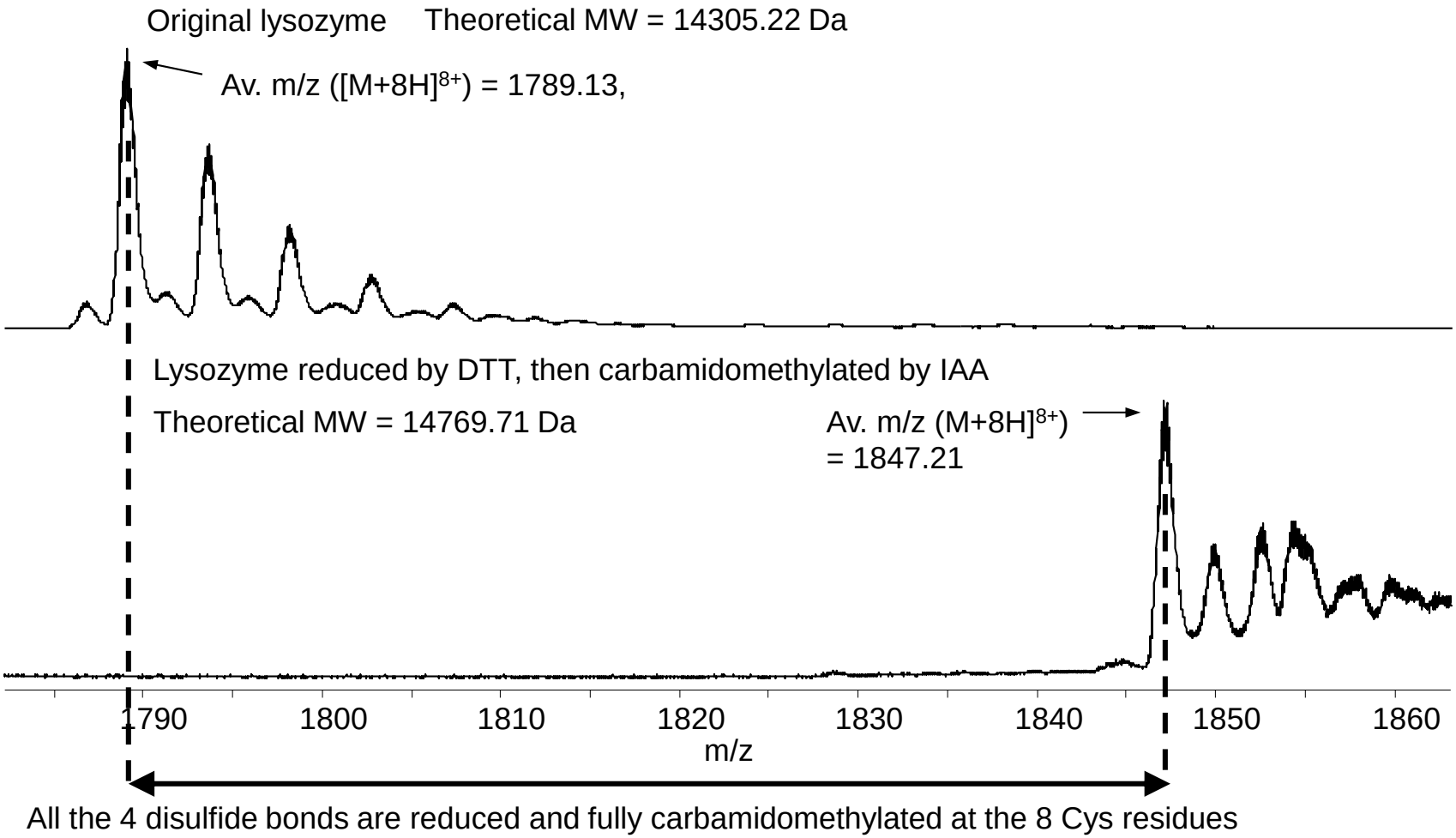


Figure 5.18. ESI-Q-TOF mass spectrum of lysozyme before (top view) and after (bottom view) carbamidomethylation of Cys residues.

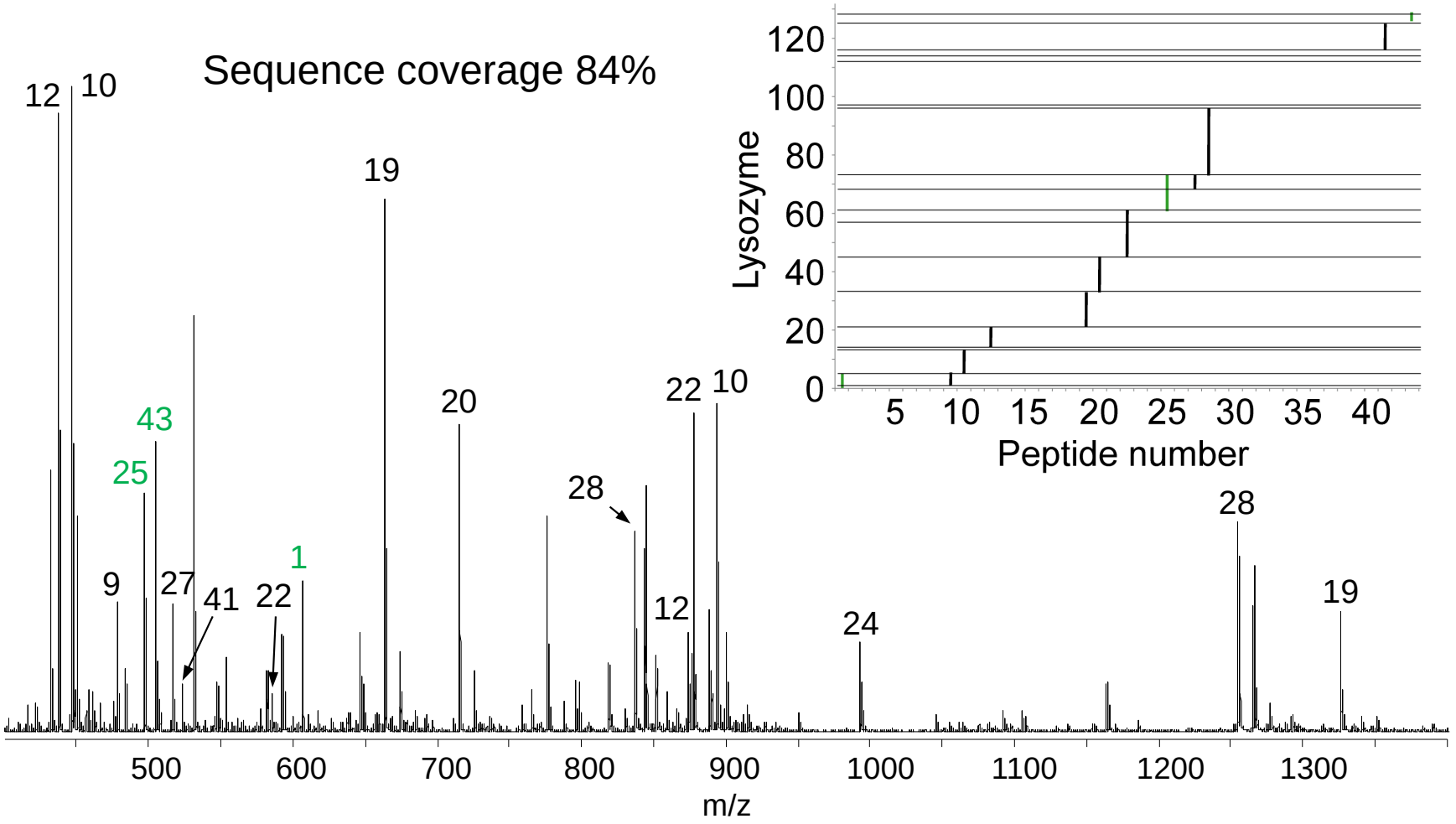


Figure 5.19. ESI-Q-TOF mass spectrum of alkylated lysozyme (without dialysis) digested for 16 h in solution. The black numbers refer to peptides with no missed cleavage sites, and the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.3.)

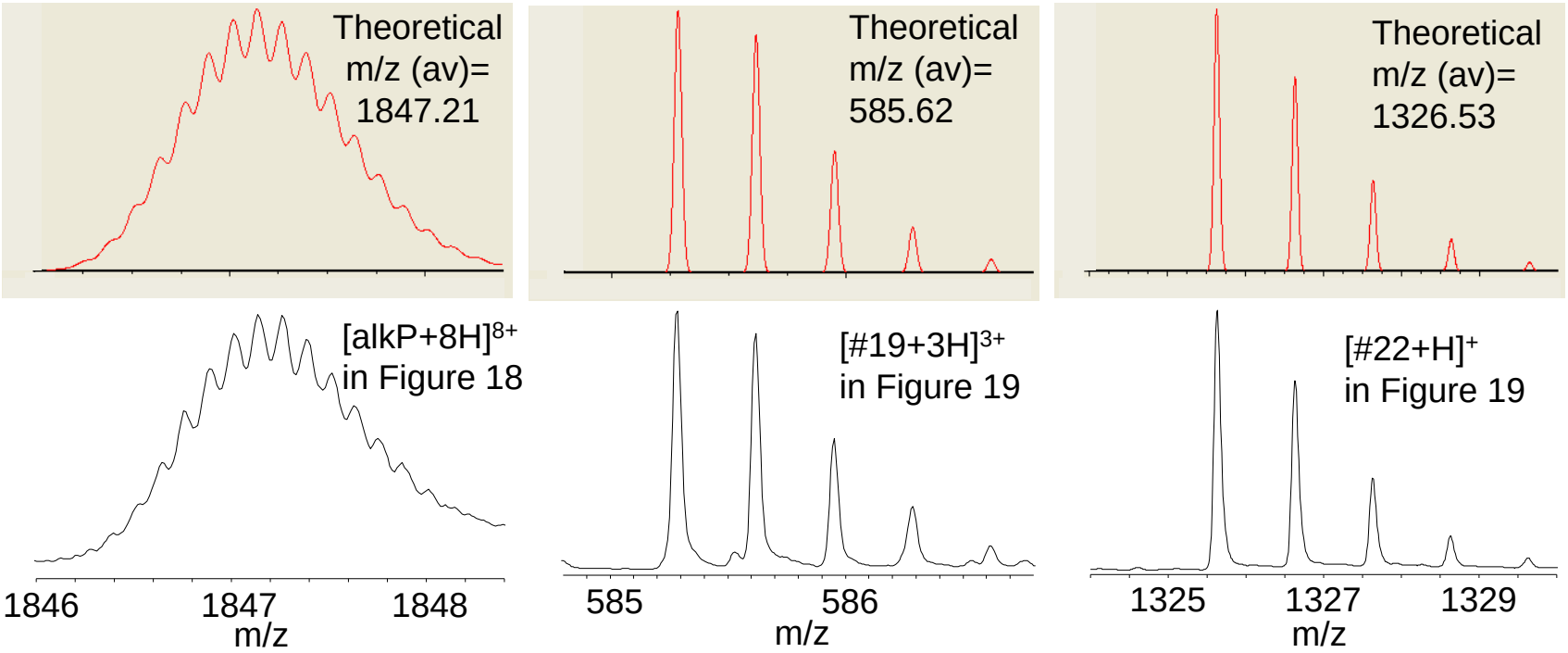


Figure 5.20. Comparison of typical experimental (bottom) and simulated (top) spectra of identified lysozyme peptides (or protein). Simulations were performed using peptide (or protein) composition and charge state, with the mass spectrometer resolving power set as 16000. The peptide number (or protein, labeled as P), theoretical m/z, and the complete mass spectrum for the experimental ion are given in each spectrum.

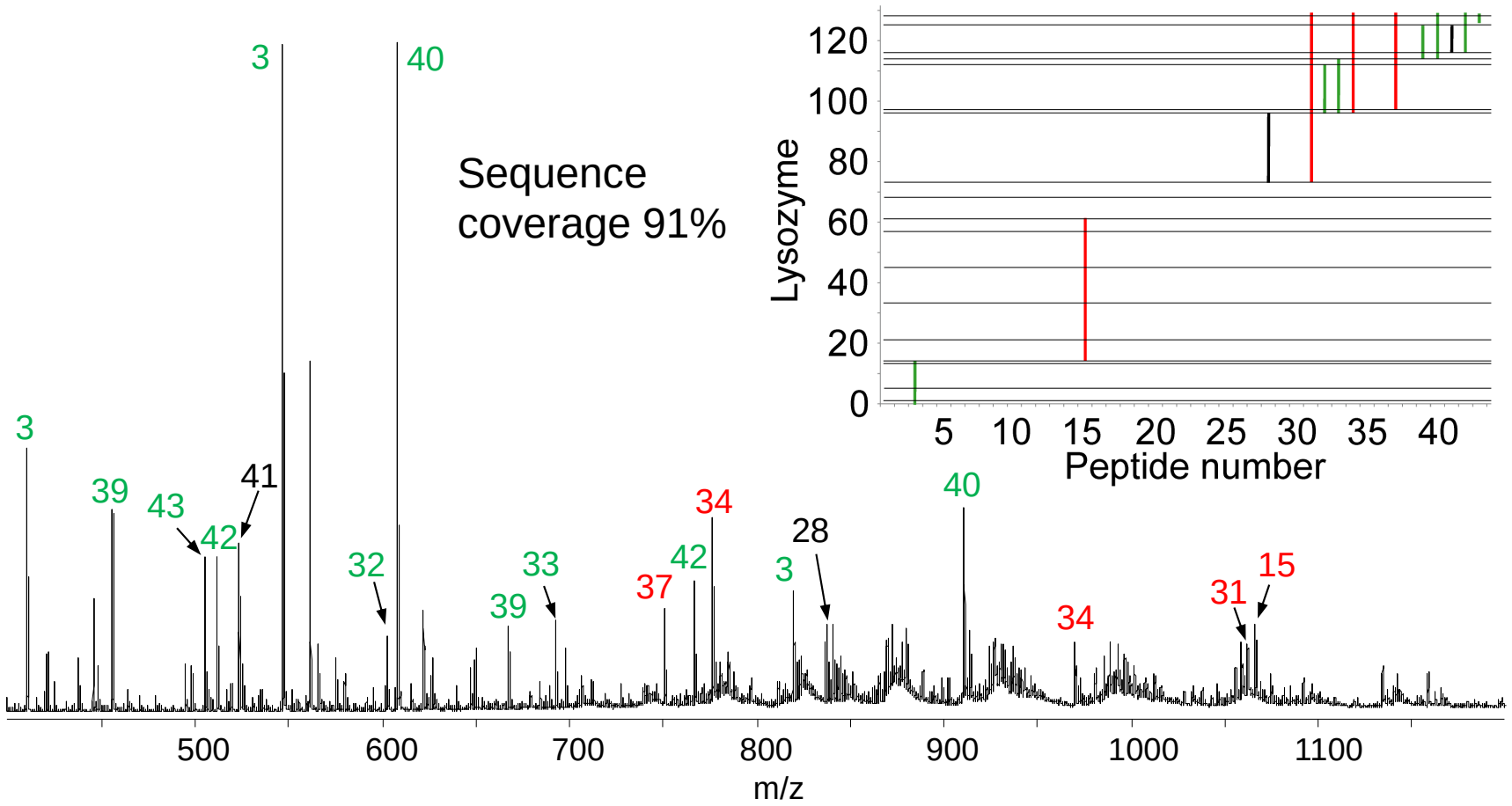


Figure 5.21. ESI-Q-TOF mass spectrum of alkylated lysozyme (without dialysis) digested for 20 s in solution. The black numbers refer to peptides with no missed cleavage site, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa.. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.3.)

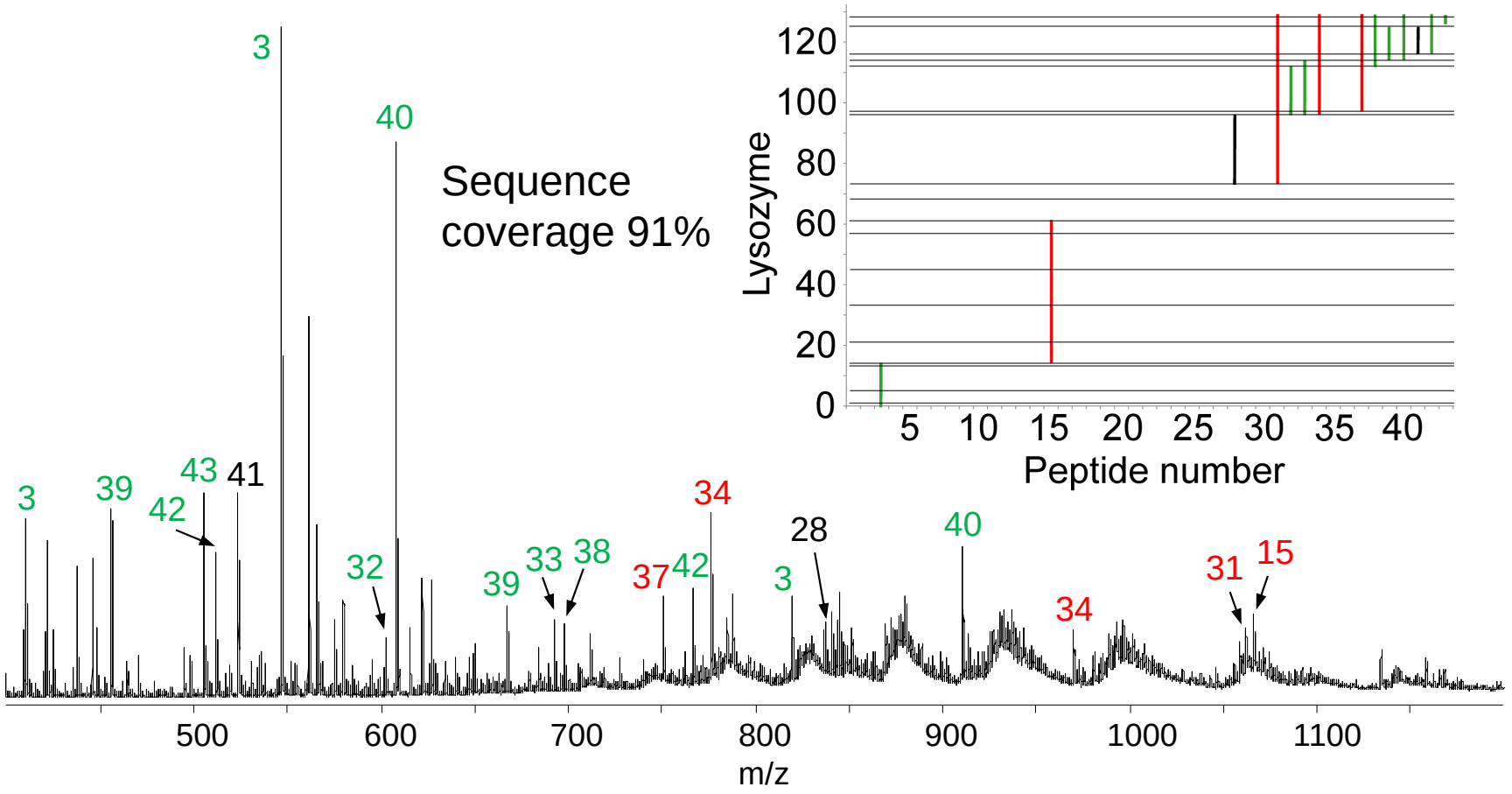


Figure 5.22. ESI-Q-TOF mass spectrum of alkylated lysozyme (without dialysis) digested in a trypsin-modified membrane (2 ms residence time, 0.45 μm pore size). The black numbers refer to peptides with no missed cleavage site, the green number refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide number refer to Table 5.3.)

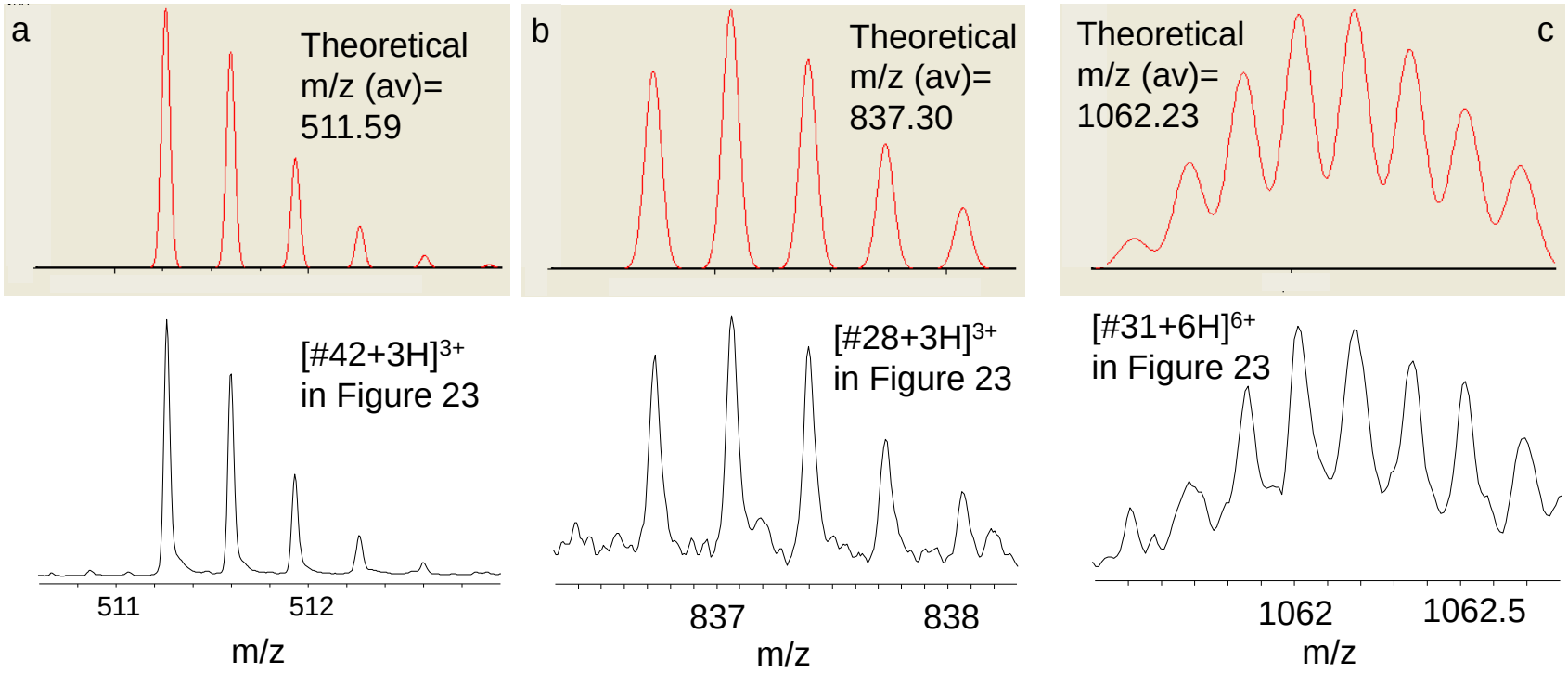


Figure 5.23. Comparison of typical experimental (bottom) and simulated (top) spectra. Simulations were performed using peptide composition and charge state, with the mass spectrometer resolving power set as 10000. The peptide number, theoretical m/z, and the complete mass spectrum for the experimental ion are given in each spectrum.

Similar to apomyoglobin and β-casein, changing the in-solution digestion time for alkylated lysozyme from 16 h to 20 s increases the protein sequence coverage from 84% to 91%. Moreover, the tryptic peptides change from predominantly small, completely cleaved peptides to large peptides with at least one missed cleavage (compare Figures 19 and 21). Figures 20 and 23 compare spectra of several lysozyme peptide/protein ions with their corresponding simulated spectra. The similar isotopic distributions suggest the feasibility of using these spectra for peptide sequence assignment.

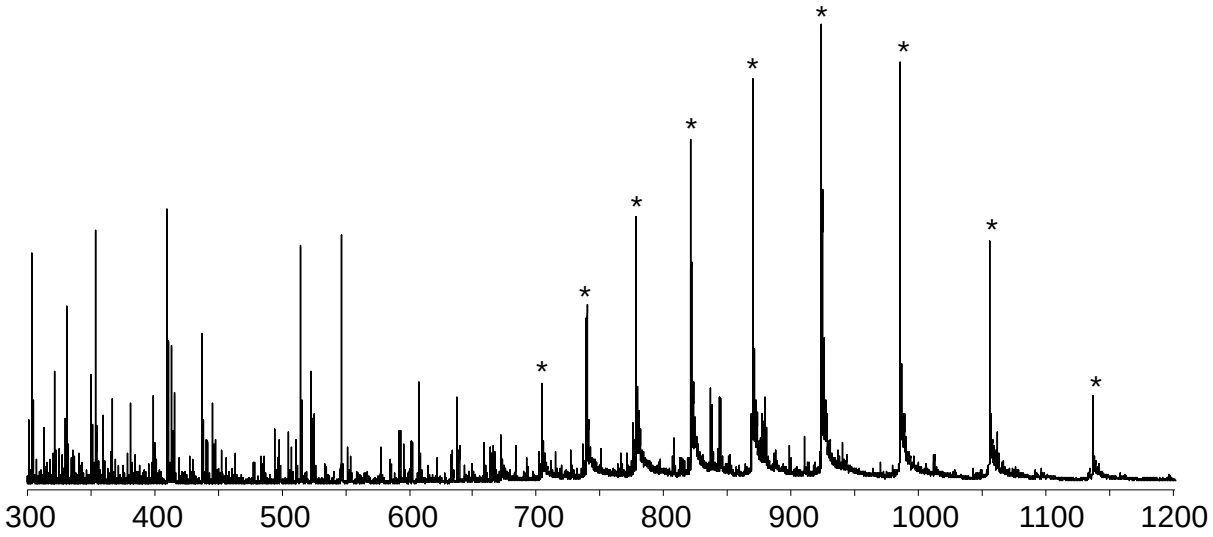


Figure 5.24. ESI-Q-TOF mass spectrum of alkylated, dialyzed lysozyme after digestion in a trypsin-modified membrane (0.32-s residence time, 0.45 μm pore size). * refers to undigested lysozyme.

Membrane digestion (0.45 μm pore size, 2 ms residence time) of the same lysozyme sample produces a tryptic peptide pattern similar to that after a 20-s in-solution digestion (compare Figures 21 and 22), but many large peptides with different charge states overlap and prevent peptide identification. Unfortunately, using ion/ion reactions to reduce the charge state of the peptides failed, probably due to the presence of small reagents (e.g. urea, DTT and IAA

used for protein denaturation, reduction and alkylation) that were not completely removed during C18 desalting

To circumvent such problems in obtaining post-ion/ion reaction mass spectra, we dialyzed the alkylated lysozyme to remove the small reagents prior to membrane digestion. Interestingly, after removing those reagents, digestion of the alkylated lysozyme in the 0.45 μm-pore size membrane slows dramatically. Even with a 0.32 s residence time, undigested lysozyme dominates the normal ESI-Q-TOF mass spectrum and interferes with the detection of large peptides (Figure 5.24). After ion/ion reactions to reduce the charge states of ions to primarily +1, a number of large peptides appear in the spectrum, although their abundances are much lower than that of lysozyme (Figure 5.24). The number of tryptic peptides detected above 100%SCT is only 4, and 2 peptides cover the entire protein sequence. However, one of the 2 peptides includes almost the entire protein. Although Figure 5.24 shows a much stronger signal for undigested lysozyme than for other peptides, the detected peptides are mainly large ones with MW above 2500 Da. The dramatically different hydrolysis pathway in the digestion samples with and without dialysis suggests a conformational change of the alkylated lysozyme upon removal of the small reagents, most notably urea. Interestingly, 4-fold longer digestion in the same membrane (1.3-s residence time) leads to a complete depletion of lysozyme, and generates mainly small peptides (Figure 5.25). Thus the peptide lifetime in the membrane must be around 1 s.

5.3.4 Protein identification by peptide mass fingerprinting (PMF)

Large tryptic peptides could potentially prove very useful in protein identification. Thus, we submitted all the detected ions (not only identified peptides) from different protein digests to MASCOT for PMF searching, and Table 5.4 shows the results. Membrane digestion of

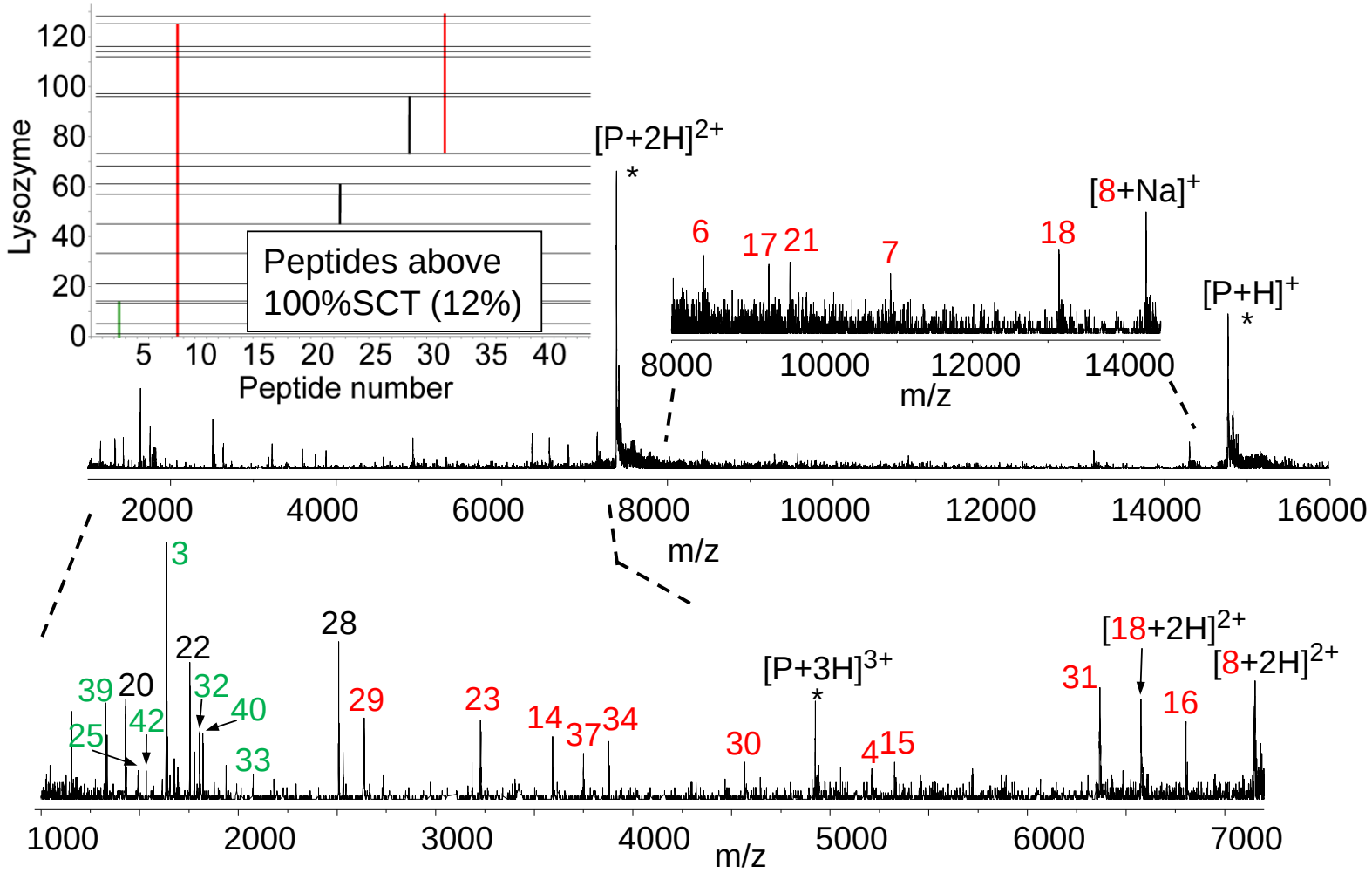


Figure 5.25. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of alkylated, dialyzed lysozyme digested in a trypsin-membrane membrane (0.32-s residence time, 0.45 μm pore size). The black numbers refer to peptide with no missed cleavage site, the green numbers refer to peptide with ≥1 missed cleavage site but MW < 2500 kDa, the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa, and * refers to undigested protein. The inset presents the peptides detected above 100%SCT, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.3.)

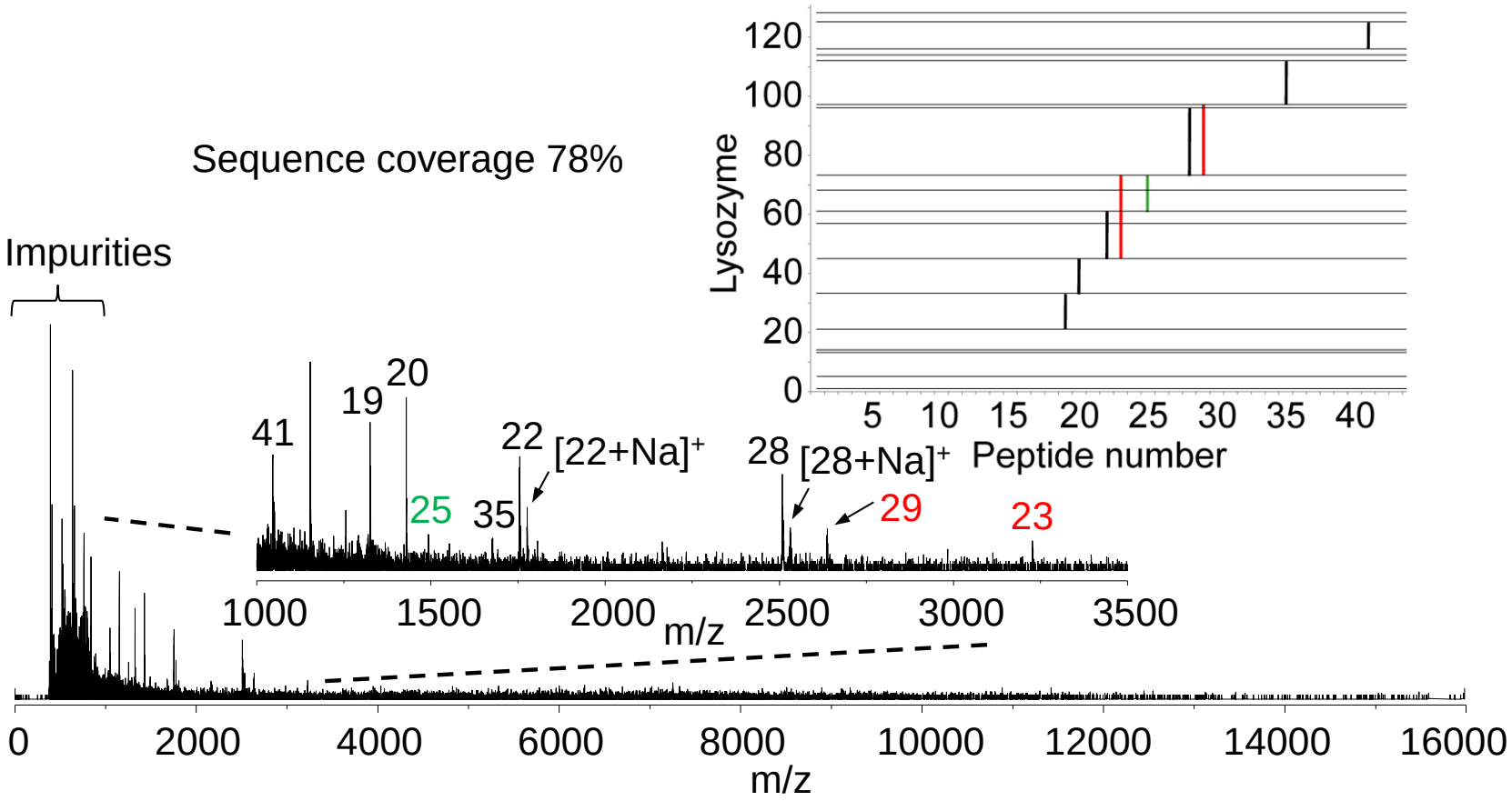


Figure 5.26. ESI-Q-TOF mass spectrum (after charge reduction by ion/ion reactions) of alkylated, dialyzed lysozyme digested in a trypsin-membrane membrane (1.3 s residence time, 0.45 μm pore size). The black numbers refer to peptides with no missed cleavage site, the green numbers refer to peptides with ≥1 missed cleavage site but MW < 2500 kDa, and the red numbers refer to peptides with ≥1 missed cleavage site and MW ≥ 2500 kDa. The inset presents the peptide pattern, with the horizontal lines showing the possible tryptic cleavage sites. (Peptide numbers refer to Table 5.3.)

Table 5.3. Tryptic peptides of lysozyme after Cys carbamidomethylation (average MW 14770 Da after carbamidomethylation) detected by ESI-MS.

Peptide #	# Missed cleavages	Sequence	m/z $[MH]^+$	Peptide #	# Missed cleavages	Sequence	m/z $[MH]^+$
1	1	1-5	606	23	2	46-73	3228
2	2	1-13	1480	24	0	62-68	994
3	3	1-14	1636	25	1	62-73	1491
4	6	1-45	5212	26	2	62-96	3983
5	7	1-61	6947	27	0	69-73	517
6	9	1-73	8421	28	0	74-96	2509
7	10	1-96 or 2-97	10912	29	1	74-97	2636
8	15	1-125	14284	30	3	74-114	4566
9	0	2-5	478	31	7	74-129	6368
10	0	6-13	893	32	1	97-112	1803
11	1	6-14	1049	33	2	97-114	2074
12	0	15-21	874	34	6	97-129	3874
13	1	15-33	2181	35	0	98-112	1676
14	2	15-45	3590	36	1	98-114	1945
15	3	15-61	5325	37	5	98-129	3746
16	5	15-73	6802	38	4	113-129	2090
17	6	15-96	9293	39	1	115-125	1333
18	13	15-129	13151	40	3	115-129	1819
19	0	22-33	1325	41	0	117-125	1045
20	0	34-45	1428	42	2	117-129	1531
21	8	34-116	9474	43	1	126-129	505
22	0	46-61	1754				

apomyoglobin "identified" the parent protein with high scores, although the mass of identified horse myoglobin is a little different from the experimental mass. In-solution digestion of β-casein for 5 and 1 min failed in protein identification based on the resulting peptides. The 20-s in-solution digestion gives a little higher score than membrane digestion. Note that the allowed number of missed cleavage sites is 9. However, the two membrane digestions both generated large peptides that covered more than 9 missed cleavage sites (Figures 15 and 16). These large peptides should contribute greatly to PMF searching if a higher number of missed cleavages is allowed for PMF searching.

Brief lysozyme digestion using modified membranes results in accurate protein identification (exact matched protein MS), while a more complete digestion (0.32 s membrane digestion) yields a lower score and less accurate protein identification. With longer digestion, loss of part of the protein sequence likely leads to uncertainty in identifying the parent protein from a group of proteins where the majority of the sequence identical. Overall, these results suggest that rapid membrane digestion has the potential to greatly facilitate protein indentification.

Table 5.4. Protein identification based on MASCOT PMF searching using ions generated by post-ion/ion reaction ESI-MS.

"Unknown" protein		Digestion method		Identified protein		
Name	av. MW (Da)	Type	Time	Name	MW (Da)	Score ***
Myoglobin (horse)	16952	Mem 0.45 μm	50 ms	Myoglobin (horse)	16940	235
		Mem 1.2 μm	50 ms			382
β-casein (bovine)	23583*	In-Sln	5 min	N/A	N/A	N/A
			1 min	N/A	N/A	N/A
			20 s	β-casein (bovine)	23583 *	122
		Mem 0.45 μm	2 ms			91
		Mem 1.2 μm	10 ms			86
Lysozyme (chick)	14770**	Mem 0.45 μm	1.3 s	Lysozyme, chain L-chicken	14500	90
			0.32 s	Chain A, structure of N59d hen egg-white lysozyme	14770 **	137

* This mass does not include increased mass due to phosphorylation.
** This mass includes the increased mass due to full carbamidomethylation.
*** Score= -10×Log(P), where P is the probability that the observed match is a random event. All the scores are the highest scores in each PMF searching, but may be underestimated due to up to 9 missed cleavages in MASCOT searching.

5.4 Conclusion

This chapter presents a novel strategy for the generation of large tryptic peptides for middle-down MS analysis of proteins. Passing substrate proteins through a trypsin-modified membrane allows tuning of proteolysis by varying the flow rate or membrane pore size. Digestion of apomyoglobin using the modified membrane yields mainly large tryptic peptides, whereas in-solution digestion for 30 min leads to the coexistence of equivalent amounts of small and large ions and undigested proteins. Passing β-casein though a modified 1.2 μm membrane allows a much briefer proteolysis than is available with in-solution digestion for only 20 s and leads to a few large peptides. As expected, denaturing and alkylation of lysozyme facilitates its tryptic digestion. Moreover, removal of ~0.1 M urea from a solution of alkylated lysozyme greatly decreases the rate of proteolysis. Furthermore, the high sequence coverage achieved with fast protein digestion generally leads to identification of proteins with higher scores and more accurate assignments for parent protein from protein family members containing similar sequences.

Control of the size of proteolytic peptides through membrane-based tryptic digestion may be useful for high-throughput generation of large peptides for protein identification and middle-down PTM characterization by MS. This approach is especially attractive when the online middle-down digestion is coupled with post ion-ion reaction MS.

REFERENCES

REFERENCES

1. Li, J. X.; Shefcheck, K.; Callahan, J.; Fenselau, C., Extension of Microwave-Accelerated Residue-Specific Acid Cleavage to Proteins with Carbohydrate Side Chains and Disulfide Linkages. *Int. J. Mass Spectrom.* **2008,** 278, 109-113.

2. Swatkoski, S.; Gutierrez, P.; Wynne, C.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C., Evaluation of Microwave-Accelerated Residue-Specific Acid Cleavage for Proteomic Applications. *J. Proteome Res.* **2008,** 7, 579-586.

3. Lee, T. D.; Shively, J. E., Enzymatic and Chemical Digestion of Proteins for Mass-Spectrometry. *Methods Enzymol.* **1990,** 193, 361-374.

4. Ahrens, C. H.; Brunner, E.; Qeli, E.; Basler, K.; Aebersold, R., Generating and Navigating Proteome Maps Using Mass Spectrometry. *Nat. Rev. Mol. Cell Biol.* **2010,** 11, 789-801.

5. Walcher, W.; Timperio, A. M.; Zolla, L.; Huber, C. G., Characterization of a Variant of the Spinach Psii Type I Light-Harvesting Protein Using Kinetically Controlled Digestion and Rp-Hplc-ESI-MS. *Anal. Chem.* **2003,** 75, 6775-6780.

6. Xu, P.; Peng, J. M., Characterization of Polyubiquitin Chain Structure by Middle-Down Mass Spectrometry. *Anal. Chem.* **2008,** 80, 3438-3444.

7. Krenkova, J.; Svec, F., Less Common Applications of Monoliths: Iv. Recent Developments in Immobilized Enzyme Reactors for Proteomics and Biotechnology. *J. Sep. Sci.* **2009,** 32, 706-718.

8. Huber, R.; Bode, W., Structural Basis of Activation and Action of Trypsin. *Accounts Chem. Res.* **1978,** 11, 114-122.

9. Rawlings, N. D.; Barrett, A. J., Families of Serine Peptidases. In *Proteolytic Enzymes: Serine and Cysteine Peptidases*, Academic Press Inc: San Diego, 1994; Vol. 244, pp 19-61.

10. Polgar, L., The Catalytic Triad of Serine Peptidases. *Cell. Mol. Life Sci.* **2005,** 62, 2161-2172.

11. Srividhya, J.; Schnell, S., Why Substrate Depletion Has Apparent First-Order Kinetics in Enzymatic Digestion. *Comput. Biol. Chem.* **2006,** 30, 209-214.

12. Perona, J. J.; Craik, C. S., Structural Basis of Substrate-Specificity in the Serine Proteases. *Protein Sci.* **1995,** 4, 337-360.

13. Schechte.I; Berger, A., On Size of Active Site in Proteases .I. Papain. *Biochem. Biophys. Res. Commun.* **1967,** 27, 157-&.

14. Bull, H. B.; Currie, B. T., Peptic Hydrolysis of Egg Albumin .1. Kinetic Studies. *J. Am. Chem. Soc.* **1949,** 71, 2758-2760.

15. Hollands, T. R.; Fruton, J. S., Kinetics of Hydrolysis of Synthetic Substrates by Pepsin and by Acetyl-Pepsin. *Biochemistry* **1968,** 7, 2045-&.

16. Sachdev, G. P.; Fruton, J. S., Kinetics of Action of Pepsin on Fluorescent Peptide Substrates. *Proc. Natl. Acad. Sci. U. S. A.* **1975,** 72, 3424-3427.

17. Choisnard, L.; Froidevaux, R.; Nedjar-Arroume, N.; Lignot, B.; Vercaigne-Marko, D.; Krier, F.; Dhulster, P.; Guillochon, D., Kinetic Study of the Appearance of an Anti-Bacterial Peptide in the Course of Bovine Haemoglobin Peptic Hydrolysis. *Biotechnol. Appl. Biochem.* **2002,** 36, 187-194.

18. Pace, C. N.; Barrett, A. J., Kinetics of Tryptic Hydrolysis of the Arginine Valine Bond in Folded and Unfolded Ribonuclease-T1. *Biochem. J.* **1984,** 219, 411-417.

19. Yi, J. Z.; Liu, Z. X.; Craft, D.; O'Mullan, P.; Ju, G.; Gelfand, C. A., Intrinsic Peptidase Activity Causes a Sequential Multi-Step Reaction (Smsr) in Digestion of Human Plasma Peptides. *J. Proteome Res.* **2008,** 7, 5112-5118.

20. Bi, H. Y.; Qiao, L.; Busnel, J. M.; Liu, B. H.; Girault, H. H., Kinetics of Proteolytic Reactions in Nanoporous Materials. *J. Proteome Res.* **2009,** 8, 4685-4692.

21. Xu, F.; Wang, W. H.; Tan, Y. J.; Bruening, M. L., Facile Trypsin Immobilization in Polymeric Membranes for Rapid, Efficient Protein Digestion. *Anal. Chem.* **2010,** 82, 10045-10051.

22. http://www.gelifesciences.com/aptrix/upp00919.nsf/Content/elpho _ applications ~elpho _applications _2d_protein_analysis~elpho_detection~3.+Staining+gels+wit h+Coomassie + Brilliant+Blue (06/03/2010)

23. Huang, T. Y.; McLuckey, S. A., Top-Down Protein Characterization Facilitated by Ion/Ion Reactions on a Quadrupole/Time of Flight Platform. *Proteomics* **2010,** 10, 3577-3588.

24. Young, N. L.; Plazas-Mayorca, M. D.; Garcia, B. A., Systems-Wide Proteomic Characterization of Combinatorial Post-Translational Modification Patterns. *Expert Rev. Proteomics* **2010,** 7, 79-92.

25. Dunn, J. D.; Reid, G. E.; Bruening, M. L., Techniques for Phosphopeptide Enrichment Prior to Analysis by Mass Spectrometry. *Mass Spectrom. Rev.* **2010,** 29, 29-54.

Chapter 6. Summary and Future Directions

6.1 Research summary

This dissertation describes the development of several polymer-based functional materials for protein isolation and digestion prior to MS analysis. The studies in chapters 2-4 aim at creating simple methods to enrich phosphopeptides from modestly complex samples such as immunoprecipitates. Chapter 2 describes a novel method for fabricating microarrays of functional polymer brushes surrounded by a hydrophobic thin film on Si wafers. A 1-μL sample droplet pinned on a 0.25-mm polymer spot shrinks during evaporation to concentrate peptides within the droplet. Additionally, during sample incubation, the aminobutyl NTA-Fe(III) complexes in the polymer brushes selectively capture phosphopeptides in an area (~0.05 mm^2) that approaches the cross-sectional area of the MALDI laser. Elution of the phosphopeptides with 1% H_3PO_4 followed by matrix addition allows sensitive MALDI-MS detection of phosphopeptides from as little as 1 μL of 0.8 nM β-casein digest.

The drawbacks of NTA-Fe(III)-derivatized polymer brushes include co-enrichment of some non-phosphopeptides and biased enrichment toward multi-phosphorylated peptides. In an effort to overcome these drawbacks, Chapter 3 presents an oxotitanium-derivatized polymer brush (P-oxoTi) that captures mainly mono-phosphopeptides under weakly acidic conditions (0.1% TFA). Increasing the TFA concentration to 2% during enrichment on P-oxoTi leads to MS signals from both mono- and multi-phosphorylated peptides, but with decreased S/N. The P-oxoTi polymer brush facilitated the identification of several phosphopeptides from digests of

p65-associated proteins immunoprecipitated together with p65 from human acute monocytic leukemia cell nuclear extracts. MALDI-MS/MS of the enriched phosphopeptides and database searching elucidated phosphopeptide sequences, phosphorylation sites, and the parent phosphoproteins.

To further improve phosphopeptide enrichment on polymer-modified wafers, we orthogonally derivatized poly(glycidyl methacrylate) brushes with oxoTi and NTA-Fe groups (Chapter 4). Such bifunctionalized polymer brushes reveal both mono- and multi-phosphorylated peptides. Thick polymer films (220 nm) with a high binding capacity enrich a wide variety of phosphopeptides over a range of sample acidity and complexity (0.1-1% trifluoroacetic acid, 5-20 μg/mL of three phosphoproteins and < 0.5 mg/mL of non-phosphoproteins). The above three generations of polymer brushes synthesized on wafers are attractive for MALDI-MS and -MS/MS analysis of target phosphoproteins in small volume-samples with moderate complexity, e.g., immunoprecipitates, single bands from gel electrophoresis, and solutions from kinase assays.

Chapter 4 also presents a tandem enrichment approach that captures different subsets of phosphopeptides sequentially in two nylon membranes derivatized with P-oxoTi and P-NTA-Fe(III), respectively. Compared to the bifunctional polymer brushes on wafers, this method provides a higher binding capacity that allows enrichment of phosphopeptides from more complex samples. The functionalized membranes could in principle facilitate phosphopeptide sequencing when coupled to HPLC-ESI-MS/MS.

Although enrichment of phosphopeptides from tryptic digests enhances the intensity of phosphopeptide signals in MS, complete tryptic digestion prior to enrichment often decreases phosphopeptide signals due to peptide loss in HPLC or inefficient ionization of the

phosphorylated species. Moreover, short tryptic peptides do not allow determination of the "crosstalk" (or correlation) of PTM sites along the entire protein chain. In an effort to overcome such issues, Chapter 5 presents a proteolysis strategy that employs brief (several ms) digestion in a trypsin-modified nylon membrane to obtain large proteolytic peptides. For example, limiting the protein residence time in a modified membrane (1.2 μm pores) to 10 ms yields 100% sequence coverage of β-casein (24 kDa, 15 potential trypsin cleavage sites) with as few as 2 peptides. Furthermore, one of the resulting large peptides (19 kDa) covers all the five phosphorylation sites of β-casein. This size-controlled proteolysis enhances the sequence coverage of target proteins and provides larger PTM coverage on a single peptide than traditional bottom-up methods.

6.2 Future work

6.2.1 Improving on-plate phosphopeptide enrichment by decreasing sample complexity

Much of the work presented in this dissertation (Chapters 2-4) focuses on phosphopeptide enrichment on wafers modified with functional polymer brushes. These materials are specifically designed to facilitate MALDI-MS analysis of low-volume protein samples. Although sensitive and specific for detecting phosphopeptides from known model protein digests with moderate complexity and acidity, this technique still shows limited successes in identifying unknown phosphorylated proteins.[1] Chapter 3 shows that even after immunoprecipitation, binding of the target phosphoprotein p65 with many other proteins leads to identification of phosphopeptides from p65-associated proteins instead of p65 itself. The excess of peptides from the co-

immunoprecipitated proteins might suppress the signals of p65 phosphopeptides and decrease the efficiency of phosphopeptide enrichment. Thus, further protein purification such as SDS-PAGE or ion-exchange prior to protein digestion is needed to improve phosphopeptide enrichment. Moreover, the tryptic digestion of p65 and its associated proteins directly on the antibody beads is probably not efficient. The high concentration of urea used for protein denaturation prior to digestion might also affect phosphopeptide binding during sample incubation on the Au-P-oxoTi wafer. Using degradable surfactants[2] for protein elution from the antibody beads followed by multiple enzymatic digestions (trypsin, Glu-C, Lys-C, etc.) and phosphopeptide enrichment may give enhanced enrichment.

6.2.2 Large-scale phosphopeptide enrichment using tandem membranes

Chapter 4 presents the possibility of capturing multiple subsets of phosphopeptides using stacked membranes with different functionalities. Because polymer brushes grown in membranes provide higher binding capacities than films on wafers,[3] functionalized membranes could enrich phosphopeptides from larger volume samples, perhaps with greater complexity. Before applying the membranes to any complex protein digest, membrane binding capacities should be evaluated using a group of synthetic mono- and multi-phosphorylated peptides.

6.2.3 Quantitative evaluation of phosphopeptide recovery

In Chapter 4, comparisons of enrichment of mono- and multi-phosphopeptides rely on MS peak patterns and S/N. Because the ionization efficiencies for mono- and multi-phosphorylated peptides are likely very different, a more careful recovery study on different subsets of phosphopeptides will require isotope-labeled phosphopeptides for MS quantification (see Chapter 3 for an example). Although previous studies described several optimization

strategies that enhance phosphopeptide enrichment efficiency in TiO_2 or IMAC columns,[4-8] none of them used quantitative MS methods for recovery evaluation. The consensus explanation for biased enrichment of mono-phosphorylated peptides on TiO_2 is strong binding of multi-phosphorylated peptides but inefficient elution from the TiO_2 surface.[9] However, the tandem phosphopeptide enrichment using stacked membranes suggests that multi-phosphopeptides may not bind well to oxoTi.

Verification of this assumption requires quantitative MS measurements. More specifically, these studies should include enrichment of an isotopically labeled protein digest (with known phosphoproteins) using TiO_2 columns or polymer-oxoTi and collection of the flow-through, washing, and elution fractions. Mixing each fraction with the same amount of starting digests with differentially labeled isotopes followed by a second phosphopeptide enrichment should allow quantitative measurement of different phosphopeptides from each fraction. Because the recovery of multi-phosphorylated peptides from the Ti-based materials is low, quantification of multi-phosphorylated peptides using the second-time enrichment will require an Fe(III)-based affinity material. To avoid the effect of isotopic tags on phosphopeptide binding and elution, isotope labeling of the flow-through, washing, and eluate fractions performed after the first enrichment may give better quantification results.

6.2.4 Tuning flowrate, membrane pore size, trypsin local "concentration", and substrate protein concentration to control membrane-based digestion

Chapter 5 shows that decreasing digestion times for both in-solution digestion and membrane-based digestion increases the size of proteolytic peptides. However, apomyoglobin

digestion gives rise to very different proteolytic peptides in brief membrane digestion than brief in-solution digestion. The membrane provides an excess of trypsin that may allow simultaneous, limited protein cleavage for most apomyoglobin molecules. However, the early apomyoglobin cleavage in solution is a slow process because there is much less trypsin. In this case, during the time needed to achieve significant depletion of intact apomyoglobin, the initially formed peptides degrade further to give a mixture of large peptides, undigested protein, and small peptides. Examination of a wider class of standard proteins that digest slowly in solution (reflected by the existence of a large portion of undigested protein after several minutes of in-solution digestion) should provide further support for differences between solution and membrane-based digestion. The membrane-based digestion should occur at a series of flowrates because different proteins will present different resistances toward tryptic digestion. If hydrolysis of such proteins in trypsin-modified membranes gives more efficient depletion of intact protein but larger peptides than in-solution digestion, we can conclude that membrane tryptic digestion provides a different hydrolysis pathway for digestion-resistant proteins.

For more rapidly digested proteins like β-casein, short in-solution digestion (<1 min) already degrades almost all of the protein. Chapter 5 shows that a 20-s, in-solution digestion depletes most of the β-casein and produces large peptides similar to those from a 2-ms membrane digestion (0.45 μm pore size). This suggests a similar digestion pathway for both methods. Using modified membranes with 1.2 μm pores gives larger tryptic β-casein peptides, but a significant portion of the β-casein is undigested. Although the intact protein may assist in protein identification by peptide mass fingerprint database searching, converting a majority of the proteins to the peptide form is necessary for middle-down MS analysis, especially PTM characterization via tandem MS. Future work on such "easily digested" proteins needs fine

tuning of membrane parameters to both degrade intact protein and provide peptides with sizes larger than those produced by the 20-s in-solution digestion. There are several possible ways to achieve this aim: (1) Minimizing the protein residence in the 0.45 μm pores to sub-millisecond times with a stronger syringe pump. (2) Shortening the trypsin modification time so that less trypsin will adsorb in the 0.45 μm pores after depositing poly(sodium 4-styrenesulfonate) (PSS) in the membrane. However, this method may lead to binding of positively charged substrate proteins to the exposed PSS in the membrane pores. (3) Depositing less PSS in the 0.45-μm membrane pores to immobilize less trypsin. To achieve efficient protein depletion, diffusion distances <1 μm promise encounters of each substrate protein with immobilized trypsin. However, decreasing the local trypsin "concentration" (it should still be close to the substrate protein concentration) could further minimize proteolysis of the resulting large peptides.

The substrate protein concentration also affects the trypsin-substrate protein ratio. When the substrate protein concentration is too high, the available trypsin in the membrane pores may be insufficient for simultaneous digestion of all the substrate molecules. Varying the substrate concentration over a large range and examining the resulting peptide patterns in mass spectra may give clues to the digestion pathway.

6.2.5 Immobilization of pepsin in nylon membranes for protein proteolysis in acidic conditions

Pepsin is another inexpensive protease, but this enzyme is rarely used for traditional bottom-up protein analysis because its low specificity in protein cleavage leads to short peptides. Possible protein hydrolysis sites with pepsin include hydrophobic and preferably aromatic amino acids such as phenylalanine, tryptophan, and tyrosine.[10] Nevertheless, pepsin is also

advantageous for protein digestion in some cases because of its acidic working condition. Acidic proteolysis with pepsin facilitates digestion of proteins that are insoluble in neutral buffers.[11] As an example, membrane proteins often require surfactant for solublization prior to tryptic digestion, but the surfactant complicates downstream peptide separation and MS analysis.[12] These proteins may, however, dissolve in acidic conditions. Pepsin is also widely used in studies of protein confirmation by hydrogen/deuterium exchange, where the proteolysis of target proteins requires an acidic condition to inhibit back exchange of deuterium with hydrogen.[13] More importantly, fast proteolysis also favors the inhibition of deuterium back exchange.[14]

Previous studies reported immobilization of pepsin for fast proteolysis.[11,14] Membranes provide a platform for achieving ms residence times, so digestion with pepsin-modified membrane could generate larger peptides than in-solution pepsin digestion. The limited digestion time with pepsin may lead to more specific cleavage than longer digestion with free pepsin in solution. For example, cleavage may occur only at aromatic residues instead of other hydrophobic residues. Narrowing the digestion specificity for database searching can improve protein identification.[15] Because of its compatibility with acidic buffers, digestion with immobilized pepsin may be attractive for integration of protein analysis procedures in an online system, which could include protein separation, pepsin digestion, peptide separation or enrichment, and finally ESI-MS.

6.3 Potential impact

The results in this dissertation lay the ground work for potential new devices that isolate small phophopeptides or produce large peptides for bottom-up or middle-down MS analysis of phosphoproteins. The metal-affinity modified polymer brushes grown on wafers greatly simplify

phosphopeptide purification processes and facilitate MALDI-MS analysis. These modified wafers have potential utility in analysis of target phosphoproteins isolated from complex samples by up-stream purifications, such as immuno-affinity fractionation and gel electrophoresis. Extending the growth of the functional polymer brushes to membranes give a higher binding capacity that allows larger scale phosphopeptide isolation from more complex digests. The phosphopeptide enrichment platforms may improve the bottom-up MS analysis of phosphoproteins. As a complement to small peptide analysis, middle-down analysis requires generation of large peptides for MS characterization. Immobilization of proteases in membranes makes possible a kinetic-controlled proteolysis. Millisecond digestion times produce large peptides with higher coverages of protein sequences and PTMs than with in-solution digestion. Further systematic studies with a large group of proteins may give direction for generating peptides with desired sizes by controlling the flowrate of the substrate protein through the membrane. Furthermore, the short-time digestion in membranes could be coupled with protein- or peptide-level separation platforms and MS for online protein analysis.

REFERENCES

REFERENCES

1. Hoang, T.; Roth, U.; Kowalewski, K.; Belisle, C.; Steinert, K.; Karas, M., Highly Specific Capture and Direct MALDI MS Analysis of Phosphopeptides by Zirconium Phosphonate on Self-Assembled Monolayers. *Anal. Chem*. **2010,** 82, 219-228.

2. Proteasemax™ Surfactant, Trypsin Enhancer. http://www.promega.com/catalog/catalogproducts.aspx?categoryname=productleaf_2163 (03/04/2011)

3. Jain, P.; Baker, G. L.; Bruening, M. L., Applications of Polymer Brushes in Protein Analysis and Purification. *Annu. Rev. Anal. Chem*. **2009,** 2, 387-408.

4. Aryal, U. K.; Ross, A. R. S., Enrichment and Analysis of Phosphopeptides under Different Experimental Conditions Using Titanium Dioxide Affinity Chromatography and Mass Spectrometry. *Rapid Commun. Mass Spectrom*. **2010,** 24, 219-231.

5. Li, Q. R.; Ning, Z. B.; Tang, J. S.; Nie, S.; Zeng, R., Effect of Peptide-to-Tio2 Beads Ratio on Phosphopeptide Enrichment Selectivity. *J. Proteome Res.* **2009,** 8, 5375-5381.

6. Simon, E. S.; Young, M.; Chan, A.; Bao, Z. Q.; Andrews, P. C., Improved Enrichment Strategies for Phosphorylated Peptides on Titanium Dioxide Using Methyl Esterification and Ph Gradient Elution. *Anal. Biochem*. **2008,** 377, 234-242.

7. Tsai, C. F.; Wang, Y. T.; Chen, Y. R.; Lai, C. Y.; Lin, P. Y.; Pan, K. T.; Chen, J. Y.; Khoo, K. H.; Chen, Y. J., Immobilized Metal Affinity Chromatography Revisited: Ph/Acid Control toward High Selectivity in Phosphoproteomics. *J. Proteome Res.* **2008,** 7, 4058-4069.

8. Ye, J. Y.; Zhang, X. M.; Young, C.; Zhao, X. L.; Hao, Q.; Cheng, L.; Jensen, O. N., Optimized IMAC-IMAC Protocol for Phosphopeptide Recovery from Complex Biological Samples. *J. Proteome Res.* **2010,** 9, 3561-3573.

9. Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R., Simac (Sequential Elution from IMAC), a Phosphoproteomics Strategy for the Rapid Separation of Monophosphorylated from Multiply Phosphorylated Peptides. *Mol. Cell. Proteomics* **2008,** 7, 661-671.

10. Hamuro, Y.; Coales, S. J.; Molnar, K. S.; Tuske, S. J.; Morrow, J. A., Specificity of Immobilized Porcine Pepsin in H/D Exchange Compatible Conditions. *Rapid Commun. Mass Spectrom*. **2008,** 22, 1041-1046.

11. Ma, J. F.; Hou, C. Y.; Sun, L. L.; Tao, D. Y.; Zhang, Y. Y.; Shan, Y. C.; Liang, Z.; Zhang, L. H.; Yang, L.; Zhang, Y. K., Coupling Formic Acid Assisted Solubilization and Online Immobilized Pepsin Digestion with Strong Cation Exchange and Microflow Reversed-Phase Liquid Chromatography with Electrospray Ionization Tandem Mass Spectrometry for Integral Membrane Proteome Analysis. *Anal. Chem*. **2010,** 82, 9622-9625.

12. Lin, Y.; Liu, Y.; Li, J. J.; Zhao, Y.; He, Q. Z.; Han, W. J.; Chen, P.; Wang, X. C.; Liang, S. P., Evaluation and Optimization of Removal of an Acid-Insoluble Surfactant for Shotgun Analysis of Membrane Proteome. *Electrophoresis* **2010,** 31, 2705-2713.

13. Nazabal, A.; Hornemann, S.; Aguzzi, A.; Zenobi, R., Hydrogen/Deuterium Exchange Mass Spectrometry Identifies Two Highly Protected Regions in Recombinant Full-Length Prion Protein Amyloid Fibrils. *J. Mass Spectrom*. **2009,** 44, 965-977.

14. Liuni, P.; Rob, T.; Wilson, D. J., A Microfluidic Reactor for Rapid, Low-Pressure Proteolysis with on-Chip Electrospray Ionization. *Rapid Commun. Mass Spectrom*. **2010,** 24, 315-320.

15. Ahrne, E.; Muller, M.; Lisacek, F., Unrestricted Identification of Modified Proteins Using MS/MS. *Proteomics* **2010,** 10, 671-686.