

# Characterization of posttranslational modifications of proteins by MALDI-TOF MS: Application to the study of tubulin

V. Redeker, J. Vinh, J.P. Le Caer and J. Rossier

Laboratoire de Neurobiologie, École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris, CNRS UMR 7637, 10 rue Vauquelin, 75231 Paris Cedex 05, France

**Matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) offers a number of advantages, including the ability to analyze mixtures of peptides, a high sensitivity, a good tolerance for buffers and salts, and also the ability to perform peptide fragmentation analysis. For these reasons, MALDI-TOF MS alone or combined with Post-Source-Decay/Collision-Induced-Dissociation (PSD/CID) has been proven to be a powerful tool for the identification and localization of posttranslational modifications of proteins.**

Numerous mass spectrometry-based strategies for identification and localization of a variety of posttranslational modifications of proteins have been developed over the last few years. After translation of the genetic information, numerous proteins undergo a variety of modifications called posttranslational modifications. For example, most of eukaryotic proteins are thought to be covalently modified by enzyme-catalyzed removal or attachment of one or several groups. The most usual modifications consist in the addition of chemical moieties (such as acetate, methyl, or phosphate groups), but also in the addition of amino acid residues, or lipidic and glycosidic units [1]. These modifications are often involved in the regulation of protein function. Phosphorylation is one of the most ubiquitous mechanisms for regulating the activity of proteins. About one-third of the proteins expressed in mammalian cells are believed to be phosphorylated. Cellular functions as diverse as cell proliferation, hormonal response or metabolism are regulated by phosphorylation/dephosphorylation of the protein. Glycoproteins are also involved in a wide range of biochemical processes, as for example in cell adhesion, targeting of cells or hormones, egg fertilization. Whereas in the case of phosphorylation the modification site has only to be localized, glycosylation studies are more complex. They imply both the determination of the modification sites and the structural elucidation of the carbohydrate chain (sequence, O- or N- branching, linkage, heterogeneity).

Since all these posttranslational modifications result in the modification of the mass of the polypeptide, the combination of mass spectrometry with protein chemistry strategies such as residue-specific enzymatic reactions is an appropriate way for characterizing the presence, the location and the chemical nature of the modification. With the recent mass spectrometry techniques, that are more and more sen-

sitive and give highly accurate measurements, direct molecular mass determination by mass spectrometry provides useful information about the overall state of modification of a peptide or a protein. If the sequence of the protein or peptide is known, the measure of a difference between the experimental and the theoretical masses of the protein or the peptide indicates the presence of a posttranslational modification or a proteolytic processing. This mass difference signs for the presence of either a mutation or a posttranslational process, and is the starting point for the identification of the chemical nature of the modification. Further MS studies in tandem mass spectrometry (MS/MS) mode can be run with a high sensitivity (up to the femtomole level). They provide partial or complete amino acid sequence. Thus, MS/MS performed on modified peptides, even within mixtures, is useful to localize modification sites within the backbone sequence. MS/MS is also used to study the intrinsic structure of the modification.

## Identification of posttranslational modifications of tubulin by MALDI-TOF MS analysis

MALDI-TOF MS provides a number of advantages over other mass spectrometric techniques, including the ability to analyze complex mixtures of peptides and proteins over a very broad mass range, a high sensitivity (from the picomole to the high attomole range) and a good tolerance for contaminants (salts, buffers, and even some detergents) and for biological matrixes [2–4]. The MALDI ionization mode leads to clear spectra easy to interpret with major single charged ions. It is the dedicated technique for peptides mixtures analysis and it can be used directly after an enzymatic digestion. MALDI-TOF MS can also be combined along with automated Edman degradation chemistry after proteolytic digestion and chromatographic separations for the analysis of peptides. The power of this methodology for the identification of posttranslational modifications in proteins can be illustrated by the recent characterization of two new and original covalent modifications of tubulin: polyglutamylation in brain [5] and polyglycylation in *Paramecium* cilia [6]. These polymodifications will now be presented as an example of a complete characterization of a posttranslational modification.

Polyglutamylation and polyglycylation involve the formation of polyglutamate or polyglycine lateral chains branched to the  $\gamma$ -carboxyl group of at least one glutamic residue in the backbone sequence of tubulin: up to 11 glutamates and up to 32 glycines can be added, respectively. Tubulin is a heterodimer consisting of two 50 kDa subunits, reported as  $\alpha$ - and  $\beta$ -tubulin. The tubulin heterodimer polymerizes to

form the structural unit of microtubules which are involved in cell motility and cell division. Tubulin presents a high degree of heterogeneity essentially due to numerous post-translational modifications, in addition to the expression of several genes. Most of this heterogeneity, including both of these polymodifications, is located in the C-terminal domain of the protein. This region of the molecule is thought to be involved in interactions with other molecules, such as microtubule-associated proteins [7,8]. Such interactions could regulate both the dynamic of microtubules and the microtubule-mediated transport of various cellular components.

Polyglycylation has been characterized as the major modification of tubulin in the stable microtubular structures found in *Paramecium* cilia [6]. C-terminal peptides were generated by a proteolytic cleavage of purified tubulin. They were isolated by chromatographic separations and characterized by automated Edman degradation chemistry and MALDI-TOF MS. The analysis of  $\beta$ -tubulin peptides will be presented as an example. In the peptides corresponding to the deduced carboxyl-terminal amino acid sequence of *Paramecium*  $\beta$ -tubulin  $^{427}\text{DATAEEEGEFEEEEGEQ}^{442}$ , none of the last six residues,  $^{437}\text{EEEEGEQ}^{442}$ , was identified by Edman degradation. Previous works led to the suggestion that at least the glutamate at position 437 was modified [5]. MALDI-TOF MS analysis of the isolated peptides allowed the characterization of the nature and the level of the modification. The difference between the theoretical mass, calculated from the encoded sequence, and each of the experimental masses (Fig. 1), corresponded to a multiple of 57 Da,

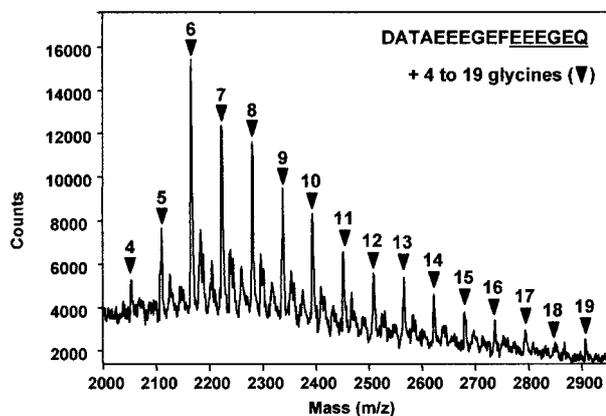


Figure 1. MALDI-TOF MS analysis of C-terminal peptides of  $\beta$ -tubulin isolated from *Paramecium* cilia. Mass increments of 57 Da provide evidence for of polyglycylation. Addition of 4 to 19 glycines to the linear sequence  $^{427}\text{DATAEEEGEFEEEEGEQ}^{442}$  are observed. The underlined residues were not identified by Edman degradation.

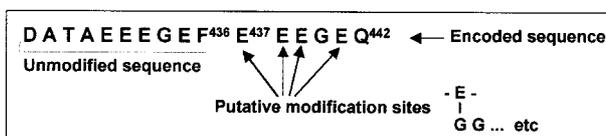


Figure 2. Sequence of the C-terminal peptide of  $\beta$ -tubulin from *Paramecium* with the four potential modification sites.

the mass of one glycine residue. The amino acid composition analysis confirmed the presence of additional glycine residues. MALDI-TOF MS of all the isolated peptides revealed that the C-terminal peptide of  $\beta$ -tubulin from *Paramecium* cilia can bear up to 32 additional glycine residues [6]. The results suggested that there are four potential sites of glycylation –  $\text{Glu}^{437}$ ,  $\text{Glu}^{438}$ ,  $\text{Glu}^{439}$ , and  $\text{Glu}^{441}$  – (Fig. 2), but the actual modification sites could not be identified at this step.

The same strategy was applied to study the polyglutamylation of some mammalian brain tubulins. Initially, polyglutamylolation has been characterized by Fast-Atom-Bombardment (FAB) MS after methyl esterification of the acidic C-terminal peptides of tubulin [5]. With MALDI-TOF MS it is now possible to obtain a comprehensive view of the glutamylation levels in the subpicomole range, directly from a mixture of peptides isolated from rat brain tubulin. In figure 3, the mass spectrum exhibits two series of ions with increments of 129 Da, the mass of one glutamate residue. The amino acid sequence of the peptides contained in this fraction was  $^{431}\text{DYEEVGVDSVEGEGXEEG}(\text{EE})^{450}$ , indicating that it was the C-terminal peptide of the  $\alpha 1/2$  tubulin isotype with a modified glutamate residue at position X. At this position no amino acid was identified by Edman degradation. The two C-terminal residues were also not identified. The first series of ion observed in figure 3 corresponds to the addition of one to eight glutamate residues to the sequence  $^{431}\text{DYEEVGVDSVEGEGEEEEGEE}^{450}$  and the second series matches with a peptide modified by the addition of one to five glutamate residues and containing a C-terminal tyrosine. Tyrosylation/detyrosylation consists in the removal and readdition of the C-terminal tyrosine residue. It is an other tubulin specific posttranslational modification [9]. This example shows the amount of information that can be obtained concerning the nature and the level of modifications of a peptide, in one direct MALDI-TOF MS experiment. Recently, MALDI-TOF MS showed that one molecule of  $\alpha$  tubulin from the flagella of sea urchin

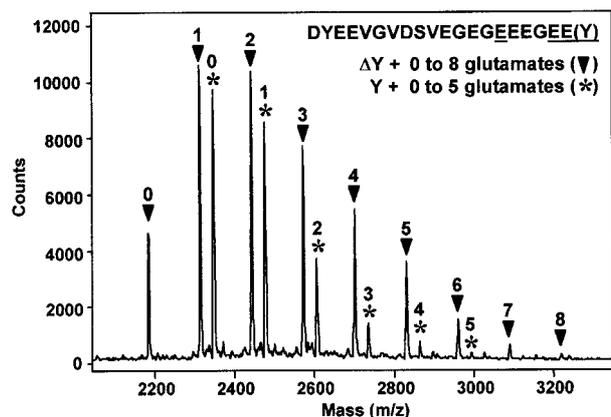


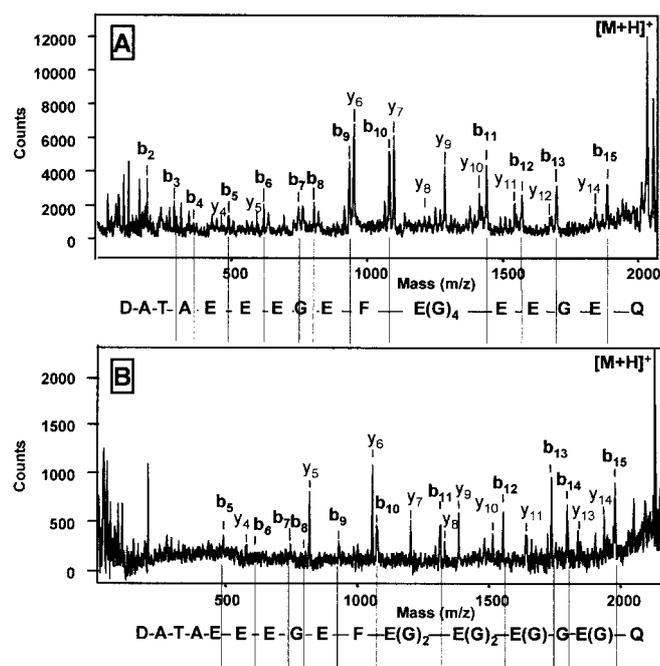
Figure 3. MALDI-TOF MS analysis of C-terminal peptides of  $\alpha$ -tubulin isolated from rat brain. Mass increments of 129 Da provide evidence for polyglutamylolation. Two series of peptides corresponding to the sequence  $^{431}\text{DYEEVGVDSVEGEGEEEEGEE}(\text{Y})^{450}$  of  $\alpha 1/2$  tubulin are observed. The first matches with the non tyrosinated peptide ( $\Delta\text{Y}$ ) bearing 0 to 8 additional glutamates, and the second matches with the tyrosinated peptide ( $\text{Y}$ ) bearing 0 to 5 additional glutamates.

spermatozoa can be simultaneously polyglycylated and polyglutamylated [10]. Further MS analyses by tandem mass spectrometry are required to determine the exact structure of these complex polymodifications, involving the localization of the modification sites within the sequence, the number of residues on each modification site, and, for a molecule that is both polyglutamylated and polyglycylated, the distribution of these residues.

### Localization of modification sites by CID/PSD MALDI-TOF MS

The structural study of polyglycylation has been realized for *Paramecium*  $\beta$ -tubulin with a MALDI-TOF instrument. The delayed fragmentation/neutralization of the analyte ions during the flight has been used in the so called Post Source Decay (PSD) and Collision Induced Dissociation (CID) MALDI-TOF MS to localize the modification sites within a peptide sequence [11,12]. This kind of fragmentation, spontaneous or activated by the increase of the residual gas pressure in the collision cell, has been well described with a specific nomenclature [13]. In MALDI-TOF MS, the PSD or the CID of the protonated peptides can be analyzed. Precursor together with fragment ions are selected with a pulsed filter synchronous with the laser source. It is localized on the trajectory of ions, that are finally separated in an electrostatic mirror (reflector) before their detection for time of flight (TOF) analysis. The spectral pattern are closely related to CID pattern obtained from conventional instruments. Previous applications of PSD MALDI-TOF MS have been reported on modified peptides to localize disulfide bonds [14], phosphorylation [15] or glycosylation [16] sites. We have recently reported comparative studies of PSD and CID/PSD for the structural elucidation of polyglycylation on C-terminal peptides of  $\beta$ -tubulin [17]. Synthetic peptides were designed to match different possible biological structures (main sequence, level and site of glycylation). They were used to establish some fragmentation rules with major *b*- and *y*-type ions fragments. The spectral pattern obtained in CID mode provided us additional information in the low mass range [18] compared to the analysis in PSD.

For an unmodified peptide, two successive peaks of the same series present a mass difference corresponding to the mass of the residue lost from the main peptidic chain. In this way, the sequence can be easily found out from the spectrum. In our fragmentation study, the glycylation residues behave like non modified residues. One modification will result in a mass shift from the expected mass difference between two given peaks. The identification of the two peaks indicates the site of the modification. The value of the mass shift gives the size of the modification. Figure 4 presents the MALDI-TOF spectrum in PSD/CID mode of a synthetic tetraglycylated peptide (Fig. 4A) and of a biological hexaglycylated peptide (Fig. 4B). Both of them have the main sequence described in figure 2. The tetraglycylated peptide bears a tetraglycine chain on the glutamate in position 437. In figure 4A, the mass of the modified residue  $-E(G)_4-$  is found between fragments  $b_{10}$  and  $b_{11}$  in one hand, and fragments  $y_5$  and  $y_6$  in the other hand. The unique modification site can be identified unambiguously. This strategy has been applied on about 600 fmol of the biological hexaglycylated peptide (Fig. 4B). The mass of a biglycylated glutamate residue  $-E(G)_2-$  was found twice between fragments  $b_{10}$  and  $b_{11}$  and between fragments  $b_{11}$  and  $b_{12}$ . In the same manner, the



**Figure 4.** PSD/CID MALDI-TOF MS analysis of A) a synthetic peptide with the sequence of a C-terminal peptide of *Paramecium*  $\beta$ -tubulin and tetraglycylated on the glutamate in position 437, and B) a biological hexaglycylated peptide purified from *Paramecium*  $\beta$ -tubulin. This analysis shows that the four potential modification sites are glycylation in *Paramecium*. The signature ions observed in the low mass range (not indicated here) confirm the structure identified [18].

mass of a monoglycylated glutamate residue  $-E(G)-$  was also found twice, between fragments  $b_{12}$  and  $b_{13}$  and between fragments  $b_{14}$  and  $b_{15}$ . Results on *y*-type fragments were observed accordingly. Moreover, in the low mass range, two internal fragments were observed as signatures for monoglycylated and biglycylated glutamate residues [18]. The whole set of results lead to the structure described in figure 4B. The polyglycylation appeared to be distributed on the four potential glycylation sites, leading to a multi-site glycylation structure.

### Study of other posttranslational modifications

Related strategies based on MALDI-TOF MS analysis are performed for the elucidation of other more common modifications such as phosphorylation and glycosylation. If the sequence is known, the comparison of the MALDI-TOF spectra obtained for the protein digest, without and after phosphatase (or glycosidase) treatment allows the location of the phosphate (or the glycan groups) in specific proteolytic fragments within the protein [19]. Specific separations for purification of modified proteins or peptides can simplify the analysis, as for example immobilized metal-ion affinity chromatography (IMAC) for phosphopeptides [20]. The phosphorylation implies a 80 Da mass increment for each phosphorylated amino acid compared to the calculated masses of the corresponding unphosphorylated peptides. Interpretation errors, due to the presence of an unphosphorylated peptide with the same mass than the phosphorylated

one, or the occurrence of more than one modification within the peptide of interest, should be avoided. This can be realized with PSD MALDI-TOF MS [21] that produces specific fragments with informations about the presence and the nature of phosphorylated residues. In this MS mode, phosphorylation is identified by the loss of  $\text{H}_3\text{PO}_4$  (98 Da) and/or  $\text{HPO}_3$  (80 Da). It can also distinguish phosphorylation on tyrosine from that on serine and threonine for peptides containing a single phosphate group [21]. In phosphotyrosine containing peptides the loss of  $\text{HPO}_3$  is major, while the loss of  $\text{H}_3\text{PO}_4$  is major when serine or threonine are phosphorylated. Phosphopeptides can be also identified with a MALDI ion trap MS [22], which offers tandem MS that combines the advantageous features of MALDI with the ease and the speed of ion trap analyzer. With this technique only the loss of 98 Da, due to the loss of  $\text{H}_3\text{PO}_4$  and/or  $\text{HPO}_3$  and  $\text{NH}_3/\text{H}_2\text{O}$ , is observed. For unknown sequences, CID MS/MS analysis of phosphopeptides should be used to obtain the amino acid sequence and the site(s) of modification. The combination of nano-electrospray ion source (NanoESI) [23] with MS/MS using a triple quadrupole, has proven also to be an excellent tool for sequencing peptides. Scans for the parents of specific fragmentation products allow the detection of posttranslationally modified peptides (such as glycosylated and phosphorylated peptides) in unseparated protein digest and with a very high sensitivity [24,25]. Due to the low flow rate (in the nL/min range) of the nano-electrospray ion source, it is possible to detect (by parent ion scan) and localize (by product ion scan) post-translational modifications in one experiment with about 1  $\mu\text{L}$  of sample.

The analysis of protein glycosylation is much more challenging because it requires the determination of the complete carbohydrate structure as well as the site(s) of attachment of the glycan chains to the protein backbone. The direct MALDI-TOF MS analysis after enzymatic-specific digestions (with proteases, exoglycosidases, N-Glycosidase F) is the method of choice for the structural elucidation of glycosylation. Sequence informations of the glycan chains can be performed by fragmentation using PSD MALDI-TOF MS or ESI MS/MS. The advantage of MALDI remains the easiest desorption of glycans, particularly because of the presence of salts. But PSD MALDI MS is a time consuming method, and ESI MS/MS competes with this technique because it gives more rapidly clearer sequences information. Ion trap MS also enables rapid identification of glycopeptides with a high sensitivity. In MALDI ion trap MS, fragmentation of the glycan chains also provides informations for the determination of the types of sugar moieties (e.g. hexose, 162 Da; N-acetylhexosamine, 203 Da; and sialic acid, 291 Da). Recently, it has been reported that the structure of glycosylation can be studied with very small quantities of glycoproteins that are directly processed on a MALDI sampling plate prior to their MS analysis [26]. Digestion with N-glycanase, performed on the MALDI plate, released the corresponding oligosaccharides in a very short time. The following treatment with a set of exoglycosidase enzymes, specific for a sugar and its linkage type, allowed the rapid determination of the glycan structure with a minimal substrate consumption.

## Conclusion

MALDI-TOF MS is particularly well adapted to the characterization of posttranslational modifications of proteins because samples can be analyzed even in mixtures or in presence of buffers or biological matrixes. It is one of the most sensitive techniques commercially available at present time. Moreover, it is possible to run MS experiments in PSD mode with a minimal sample consumption on the same sample spot preparation, even some days later, to gain further structural information. But some other instruments have been designed and are worthy to note. The alternative to MALDI source in the same sensitivity range is the nano-electrospray source working with a flow rate of about a few nL/min. Its advantage is the direct LC-MS analysis feasibility. Coupled to quadrupole analyzers the sensitivity is lower than for the MALDI-TOF configuration. It is not anymore the case when coupled to a TOF analyzer. The nanoESI-quadrupole-TOF configuration is highly sensitive and the MS/MS mode is much faster and accurate than the PSD mode. Recently, appeared also MALDI ion trap mass spectrometers combining the advantages of the MALDI source and those of the ion trap. Anyway, MS techniques have to be followed very carefully, because they are still in rapid progress.

## References

1. Krishna, R. G.; Wold, F. *Adv. Enzymol. Rel. Areas Mol.* **1993**, *67*, 265-298.
2. Carroll, J. A.; Beavis, R. C. In: Laser Desorption and Ablation, Miller, J. C.; Haglund, R. F. Eds., Academic Press: San Diego, 1996, Chapter 7.
3. Mann, M.; Talbo, G. *Curr. Opin Biotechnol.* **1996**, *7*, 11-19.
4. Hühmer, A. F. R.; Aced, G. I.; Perkins, M. D.; Gürsoy, R. N.; Jois, D. S. S.; Larive, C.; Siahaan, T. J.; Schöneich, C. *Anal. Chem.* **1997**, *69*, 38R-41R.
5. Eddé, B.; Rossier, J.; Le Caer, J. P.; Desbruyeres, E., Gros, F.; Denoulet, P. *Science* **1990**, *247*, 83-85.
6. Redeker, V.; Levilliers, N.; Schmitter, J. M.; Le Caer, J. P.; Rossier, J.; Adoutte, A.; Bré, M. H. *Science* **1994**, *266*, 1688-1691.
7. Serrano, L.; Avila, J.; Maccioni, R. B. *Biochemistry* **1984**, *23*, 4675-4681.
8. Hagiwara, H.; Yorifuji, H.; Sato-Yoshitake, R.; Hirokawa, N.; *J. Biol. Chem.* **1994**, *4*, 3581-3589.
9. Raybin, D. Flavin M., Modifications of tubulin by tyrosylation in cells and extracts and its effect on assembly in vitro. *J. Cell. Biol.* **1977**, *73*, 492-504.
10. Mary, J.; Redeker, V.; Le Caer, J. P.; Rossier, J.; Schmitter, J. M. *J. Biol. Chem.* **1996**, *271*, 9928-9933.
11. Kaufmann, R.; Spengler, B.; Lützenkirchen, F. *Rapid Commun Mass Spectrom.* **1993**, *7*, 902-910.
12. Spengler, B.; Kirsch, D.; Kaufmann, R.; Jaeger, E. *Rapid Commun Mass Spectrom.* **1992**, *6*, 105-108.
13. Biemann, K. *Methods Enzymol.* **1990**, *193*, 886-887.
14. Crimmins, D. L.; Saylor, M., Rush, J., Thoma, R. S. *Anal. Biochem.* **1995**, *226*, 355-361.
15. Talbo, G.; Mann, M. In: Techniques in Protein Chemistry, Crabb J. W., Ed., Academic Press: New-York, 1994; Vol. V, pp 105-114.
16. Huberty, M. C.; Vath, J. E.; Yu, W.; Martin, S. A. *Anal. Chem.* **1993**, *65*, 2791-2800.
17. Vinh, J.; Loyaux, D.; Redeker, V.; Rossier, J. *Anal. Chem.* **1997**, *69*, 3979-3985.

18. Vinh, J.; Redeker, V.; Loyaux, D.; Rossier, J. *Advances in MS; Vol. 14. Proceedings of the 14<sup>th</sup> IMSC, 25-29 August 1997, Tampere, Finland*, in press.
  19. Liao, P. C.; Leykam, J.; Andrews, P. C.; Gae, D. A.; Allison, J. *Anal. Biochem.* **1994**, *219*, 9-20.
  20. Neville, D. C. A.; Rozanas, C. R.; Price, A. M.; Gruis, D. B.; Verkman, A. S.; Twonsend, R. R. *Protein Sci.* **1997**, *6*, 2436-2445.
  21. Annan, R. S.; Carr, S. A. *Anal. Chem.* **1996**, *68*, 3413-3421.
  22. Qin, J.; Chait, B. T. *Anal. Chem.* **1997**, *69*, 4002-4009.
  23. Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.
  24. Carr, S. A.; Huddleston, M. J.; Annan, R. S. *Anal. Biochem.* **1996**, *239*, 180-192.
  25. Wilm, M.; Neubauer, G.; Mann, M. *Anal. Chem.* **1996**, *68*, 527-533.
  26. Mechref, Y.; Novotny, M. V. *Anal. Chem.* **1998**, *70*, 455-463.
-