DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Title Page

LC-MS Differential Analysis for Fast and Sensitive Determination of Biotransformation of Therapeutic Proteins

Ming Yao, Bingming Chen, Weiping Zhao, John T. Mehl, Lingjun Li, and Mingshe Zhu

Pharmaceutical Candidate Optimization, Bristol-Myers Squibb, Princeton, New Jersey, United States: MY, WZ, JTM, MZ

School of Pharmacy and Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States: BC, LL

School of Life Sciences, Tianjin University, No. 92 Weijin Road, Nankai District, Tianjin 300072, China: LL

MassDefect Technologies, Princeton, New Jersey, United States: MZ

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Running Title Page

a). Running title

Differential Analysis for Biotransformation in Therapeutic Proteins

b). Corresponding authors:

Dr. Mingshe Zhu, MassDefect Technologies, PO Box 172, Rocky Hill, NJ 08553: Phone: +1-609 694-3946. Email: mingshe.zhu@yahoo.com.

Dr. Lingjun Li, School of Pharmacy, University of Wisconsin-Madison, 777 Highland Avenue, Madison, Wisconsin 53705-2222 USA; Phone: +1-608-265-8491; Fax: +1-608-262-5345; Email: lingjun.li@wisc.edu

c). Page and word counts

Number of

Text pages: 31

Tables: 3

Figures: 5

References: 22

Number of words in

Abstract: 182

Introduction: 910

Discussion: 702

Abstract

Therapeutic biologics have become a fast-growing segment within the pharmaceutical industry during the past three decades. Although the metabolism of biologics is more predictable than small molecule drugs, biotransformation can significantly affect the activity of biologics. Unfortunately, there are only a limited number of published studies on the biotransformation of biologics, most of which are focused on one or a few types of modification. In this study, an untargeted LC-MS based differential analysis approach was developed to rapidly and precisely determine the universal biotransformation profile of biologics with the assistance of bioinformatic tools. A human monoclonal antibody (mAb) was treated with t-butyl hydroperoxide and compared with control mAb using a bottom-up proteomics approach. Thirty-seven types of post-translational modifications were identified and thirty-eight peptides were significantly changed. Moreover, while all modifications were screened and detected, only the ones related to the treatment process were revealed by differential analysis. Other modifications that co-exist in both groups were filtered out. This novel analytical strategy can be effectively applied to study biotransformation mediated protein modifications, which will streamline the process of biologics drug discovery and development.

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

Since the approval and commercialization of the first monoclonal antibody (mAb) in 1986, the development of therapeutic biologics has captured significant attention within the pharmaceutical industry during the past three decades (Waldmann, 2003; Xu and Vugmeyster, 2012). In general, therapeutic biologics refers to a substance derived from living organisms that is applicable to prevent, treat, or cure, a human disease or condition (Hamuro and Kishnani, 2012; Zhao et al., 2012). In addition to mAbs, which represent the majority of marketed biologics, other biologics modalities, such as therapeutic peptides, bi-functional biologics and antibody-drug conjugates (ADCs), have rapidly emerged (Xu and Vugmeyster, 2012). Some therapeutic proteins contain artificial amino acids, unique peptide sequences or linkage structures, or small molecule drug moieties (Hamuro and Kishnani, 2012; Katsila et al., 2012; Xu and Vugmeyster, 2012).

Like endogenous proteins, antibody therapeutics undergo catabolism such as oxidation, deamination, glycosylation, phosphorylation and isomerization in a biological system to form catabolites (peptides and amino acids). Here we define biotransformation of a therapeutic protein as a chemical or enzymatic process occurring within a biological system in vitro or in vivo, which changes the structure of the therapeutic protein. Biotransformation of a therapeutic protein can follow either predictable catabolism pathways, similar to the disposition and elimination processes of endogenous proteins, or pathways different from protein catabolism due to the presence of an artificial amino acid, peptide sequence, linkage structure of amino acids and conjugated drug (Cheloha et al., 2014; Walles et al., 2016; Zhang et al., 2016). Biotransformation-derived protein modifications may have profound impact on the biologics' activity or disposition and clearance behavior, as they can potentially lead to reduced target binding efficiency, shorter half-life, higher clearance rate, poor bioavailability, increased off-target binding/toxicity, increased

immunogenicity and altered assay detection (Liu et al., 2009; Cai et al., 2011; Liu et al., 2011; Wang et al., 2011; Hall, 2014). For example, the methionine oxidation in human IgG1 causes significant reduction in binding affinity and reduced serum half-life (Wang et al., 2011); the isomerization of a single Asp residue leads to complete loss of target-binding in a model mAb (mAb X) (Prueksaritanont and Tang, 2012).

Usually, drug metabolism and pharmacokinetic research of therapeutic proteins relies on ligand-binding assays (LBA), such as enzyme-linked immunosorbent assay (ELISA), for protein quantitative analysis during drug development. LBAs can be highly sensitive, are typically easy to operate and capable of high throughput analysis. Recently, quantitative LC-MS based methods have increasingly been employed to support pharmacokinetic characterization during drug discovery and preclinical development (An et al., 2014). However, there is no effective methodology available to study global changes in biologics upon biotransformation, and most of the published biotransformation studies focus on only one type of modification (Liu et al., 2011; Wang et al., 2011; Bults et al., 2016; Kullolli et al., 2017). For example, Bults *et al.* reported an LC-MS approach to analyze trastuzumab and its deamidation product *in vivo*, in which a selected reaction monitoring approach was used to monitor five molecular species from tryptic digested trastuzumab from human plasma without enrichment (Bults et al., 2016). This method is highly specific to monitor the deamidation at Asp55 and it is not a means to monitor global biotransformation of a therapeutic protein.

The investigation of therapeutic protein biotransformation relies on HRMS-based proteomics workflows, such as top-down, bottom-up approaches and a combination of both approaches (Domon and Aebersold, 2006). Intact protein or protein assemblies are directly analyzed in a top-down proteomics workflow, which requires ultrahigh resolving powers and

usually has limited protein sequence coverage and sensitivity (Qian et al., 2001; Kang et al., 2017). In a typical bottom-up proteomics workflow, protein samples are enzymatically digested into peptides, which are then subjected to LC-MS experiments. The resulting MS/MS spectra are compared with the protein sequence database to find the best matched peptide. To search for modified peptides, MS/MS spectra are compared with a list of pre-defined modifications on peptides in traditional database software. Some software packages, such as PEAKS, used an improved algorithm for PTM search, that are capable of screening more than 650 known PTM types in one run (Han et al., 2011). Compared to the top-down workflow, bottom-up approach offers several advantages, including improved sequence coverage, superior analytical specificity, sensitivity, precision and accuracy (Ewles and Goodwin, 2011; Xu and Vugmeyster, 2012; Kullolli et al., 2017).

A key analytical challenge in determining global biotransformation of therapeutic proteins via a bottom-up workflow is the large number of low level protein modifications that are not associated with the biotransformation process. These modifications are either present in a dosing solution or formed during sample preparation, such as enzymatic hydrolysis, isolation and enrichment processes. It would require significant effort and time to detect and characterize these modified proteins and distinguish them from biotransformation-derived protein modifications. Herein, we developed and validated an untargeted LC-HRMS based differential analysis method for rapid detection and identification of modified peptides derived from the biotransformation of therapeutic mAb. A human mAb was forced-oxidized and compared with untreated control mAb using a bottom-up proteomics workflow. Generated LC-MS datasets were processed using bioinformatic tools for differential analysis. Modified peptides undergone significantly higher levels of biotransformation were highlighted and structurally characterized. Moreover, relative

quantification of biotransformation-related modified peptides was performed. To our knowledge, this is the first study that analyzes the global biotransformation of mAb using LC-HRMS and differential analysis. With relatively easy sample preparation and rapid data acquisition, this method provides a template for the drug discovery to characterize therapeutic biologics in a high throughput manner.

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Materials and Methods

Materials and Chemicals

All reagents were used without additional purifications. Mouse anti-human IgG R10Z8E9 was available in house and Dynabeads (M-280 tosylactivated) was purchased from Invitrogen (Carlsbad, CA). Male rat serum was purchased from Bioreclamation IVT (New York, NY). Trypsin used for mAb digestion was purchased from Promega (Madison, WI). The phosphate buffered saline (PBS) and all solvents used in LC mobile phases (water, acetonitrile and formic acid (FA) were purchased from GE healthcare Bio-sciences (Malborough, MA). Tert-butyl hydroperoxide (tBHP), dithiothreitol (DTT) iodoacetamide (IAA) and ammonium bicarbonate (AMBIC) were purchased from Sigma Aldrich (St. Louis, MO) and the Zwittergent 3-12 detergent (Zwitt) was purchased from Calbiochem (EMD Millipore, Billerica, MA).

Preparation of immunocapture beads

Mouse anti-human IgG antibodies were used in preparing tosylactivated beads. Dynabeads M-280 tosylativated (24 mg or 800 µL) were prepared by washing with buffer A (0.1 M borate buffer pH 9.5) twice. The total of 480 µg antibody was mixed with the beads in buffer A followed by adding 488 µL buffer C (3 M ammonium sulfate in buffer A). The mixture was incubated on a rotator in 37 °C overnight. The supernatant was removed after the incubation and the mixture was incubated in 1 mL buffer D (PBS pH 7.4 with 0.5% (w/v) BSA) on a rotator at 37 °C for 1 hr. The mixture was then washed with 1 mL buffer E (PBS pH 7.4 with 0.1% (w/v) BSA) twice and stored in 800 µL buffer E at 4 °C until use (within 10 weeks). The resulting tosylactivated beads had a concentration of 0.6 µg/µL of captured antibody.

Generation of biotransformation and control samples of biologics

A model therapeutic mAb (145.0 kDa, 20 µg/µL, 10 µL) was treated with 40 µL 0.1% tBHP at 25 °C overnight to introduce various protein modifications. The treated mAb and control mAb in solution (25 µL each with 100 µg mAb) were spiked into 75 µL rat serum to generate a biotransformation sample and a control sample, respectively. The resulting serum sample each had 100 µg mAb with a concentration of 1 µg/µL.

Sample preparations

mAb in rat serum was enriched by immunocapture using the tosylactivated beads, which were prepared according to the procedure described in the previous section. For each sample, 170 µL beads were loaded into a well of a 96-deepwell plate followed by 200 µL buffer E and 100 µL of rat serum sample. The mixture was incubated at room temperature for 1 hr on a shaker and then washed with 200 µL PBS with 0.05% Zwitt. The beads were then transferred to a fresh well using 200 µL PBS with 0.5 M NaCl and 0.05% Zwitt and washed with 200 µL 500 mM AMBIC/0.05% Zwitt, 200 µL of 2 mM AMBIC and eluted with 50 µL of 12 mM HCl. The eluted fraction was transferred to a well containing 25 µL of 100 mM AMBIC and 20 µL water. Five µL DTT (100 mM) was added to the mixture and incubated at 60 °C for 60 min. Then 11 µL of IAA (100 mM) was added to the mixture and incubated in dark for 30-45 min. Trypsin was added to each sample with a 1:25 trypsin:protein ratio (for samples with 100 µg mAb, 4 µg trypsin was added) and incubated at 37 °C overnight. Finally, 10% formic acid (1.2 µL) was added to make a final concentration of 0.1% FA to quench the reaction.

LC-MS analysis

An Agilent 1290 Infinity II LC system (Agilent Technologies, Santa Clara, US) was connected to a TripleTOF 5600 mass spectrometer (AB Sciex, Framinghan, MA) for all LC-MS

analysis. 10 µL of sample was injected onto a C_{18} column (Waters Acquity UPLC BEH C_{18}, 1.7 µm) for each run. Mobile phase A was H_2O with 0.1% FA and mobile phase B was ACN with 0.1% FA. The following gradient with a flow rate of 400 ng/mL was used (time/minute, % mobile phase B): (0, 2), (2, 2), (45, 45), (47, 90), (50, 90), (51, 2) (55, 2).

To maximize the information acquired on the mass spectrometer for each sample, a full MS scan (*m/z* 300-2000) was acquired followed by top 20 information dependent acquisition (IDA) MS/MS scans (*m/z* 100-1600) at positive ion mode. The parameters for CUR, DP, CE; Gas1, Gas2 in full MS scan mode was 30, 10, 55, and 55. The mass tolerance was 50 mDa. Source temperature was set to 450 °C and tray temperature was set to 22 °C. The criteria for the IDA scans were as follows: top 20 most intensive peaks with charge states from 2 to 5 and intensities greater than 50 were selected. Dynamic background subtraction was set for exclusion. Dynamic collision energy was enabled for multiply charge peptides.

Data processing

A customized version of PEAKS 7.5 (Bioinformatics Solution Inc., Waterloo, ON, Canada) was used for peptide mapping, PTM searching and differential analysis (label free quantification). The following parameters were used for peptide mapping against the tocilizumab database: parent mass error tolerance of 50 ppm, fragment mass error tolerance of 1.0 Da and up to 3 missed cleavages were allowed. Deamidation (NQ), and oxidation (M) were set as variable modifications. Carbamidomethylation was set as a fixed modification. All 53 common and 420 uncommon PTMs were included for PTM searching. Up to 3 PTMs were allowed for each peptide. False discovery rate was set as 1%. Each data file was normalized to total ion count before differential analysis. One of the control samples was used as a reference sample for normalization in order to compare the ratio for each peptide in different samples. After database searching and PTM analysis, the

data files were processed with differential analysis. The resulting heat map represents the $\log_2$ ratio to the average peak area across different samples. Different filtering criteria were used for data interpretation.

Results

A bottom-up proteomics-based differential analysis workflow

In the current study, a bottom-up proteomics-based differential analysis strategy was developed and evaluated for untargeted detection and characterization of global biotransformation products on a therapeutic mAb (Figure 1). In this analysis, a biotransformation sample and a corresponding control sample are compared. The biotransformation sample can either be from an *in vitro* incubation of a therapeutic protein or collected from an animal after dosing a therapeutic protein. The control was either a pre-incubation or a pre-dosing sample. Immunocapture and/or other methods were employed to recover and enrich the therapeutic protein and its modification products from the biological matrix. The recovered proteins were further reduced, alkylated and digested by trypsin, followed by peptide analyses via LC-HRMS. Resulting full scan MS and MS/MS datasets of the biotransformation and control samples, which could be acquired by a data-dependent or data-independent method, were subjected to peptide mapping, modification analysis and differential analysis using customized proteomics data-processing software. Peptide mapping was achieved by comparing the acquired MS^2 spectra with theoretically calculated fragment ion *m/z* values from the mAb amino acid residue sequence. Peptide modification analysis was achieved by searching the mass differences of known modifications and compared with the acquired MS^2 spectrum. Differential analysis was performed by comparing the peak area of extracted ion chromatograms of detected features from the biotransformation and control samples in triplicate. Peptide mapping coverage was determined to ensure the quality of protein hydrolysis and peptide mapping processes. Peptides modified via biotransformation, which were either absent in the control sample or had significantly higher levels in the test sample, were determined using differential analysis. Furthermore, sequences of modified peptides were elucidated automatically

using the software, or by manual interpretation for uncommon peptide modifications. In the current study, simplified biotransformation and control samples of a human mAb were generated to test the effectiveness of the workflow (Figure 1).

The coverage of peptide mapping and identification of protein modifications

The sequence coverage and peptide modifications detected in the light chain and heavy chain of the control mAb sample were presented in Supplemental Figure 1. The sequence coverage was calculated using the number of amino acid residues detected divided by the total number of amino acid residues in the known mAb sequence. By using the untargeted LC-HRMS approach, 93% of the light chain and 92% of the heavy chain were covered when the false discovery rate was set to be 1%. The result indicated that the quality of protein processes and LC/HRMS analysis are good. Sixty and 129 unique peptides were detected for the light chain and heavy chain of the mAb, respectively, recovered from the biotransformation sample. Common modifications, such as carbamidomethylation, deamidation and oxidation, were labeled in the figures. Additionally, less common protein modifications such as sulfation and pyro-Glu from Glu (Table 1 and Supplemental Figure 1) could be found via data-processing using the customized software. In total, 37 types of modifications were determined and characterized by the PEAKS PTM search algorithm (Supplemental Table 1). The corresponding peptide spectrum match (#PSM), delta mass of each modification (ΔMass), position and -10logP values were listed on Table 1. Carbamidomethyl, deamidation, oxidation and dihydroxy were among the common popular modifications; other less common ones, such as point mutation of single amino acid, sulfation and carboxylation were also detected (Supplemental Figure 1).

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Identification of modified peptides derived by biotransformation

Modified peptides derived from biotransformation were determined by differential analysis (Figure 2). These peptides were either only present or had a significantly higher level in the biotransformation sample. Table 2 summarized a partial list of these peptides with significance values (-10logP) above 5. Most of these peptides were associated with Methionine oxidation. In the sample profile and group profile columns, the left panel represented the biotransformation samples while the right panel represented the control samples. All the unoxidized peptides (row 1, 3, 5 & 7) had much higher intensities in the control samples while all the oxidized peptides (row 2, 4, 6 & 8) had much higher intensities in the biotransformation samples. Some data points were not shown in the sample profile due to the corresponding peptide intensity was below detection.

As an example, the extracted ion chromatogram of the doubly charged unmodified and oxidized forms of a peptide (TTAMDYWGQGSLVTVSSASTK) are shown in Figure 2. High abundance of the unmodified peptide was observed in all three of the control samples with intensities above 1E4, while the signals were barely detected in the biotransformation sample (Figure 2A). On the other hand, high abundance of the oxidized form of this peptide was observed in the biotransformation sample, but barely any signal was detected in the control sample (Figure 2B). A 2-minute shift in retention time was observed between these peptide forms, which was associated with the oxidation of Methionine that increased the hydrophilicity of the peptide.

The corresponding MS^2 spectra of these peptides (Figure 3) were acquired by the IDA scans during the elution of the corresponding peptides. For both peptides, y_2 to y_{15} ions (except for y_{11} ion of the oxidized form) and some of the b ions were detected with high abundance. Based on the fragment ion information, the sequence of these peptides could be confidently assigned by database searching algorithm of PEAKS. By comparing the spectra of unmodified (Figure 3A) and

oxidized (Figure 3B) forms, the mass difference between Methionine and oxidized Methionine (+15.99) was observed in the b_7 ions (shown in dotted blue lines) as well as some other b ions. The mass difference in the MS^2 spectra allow the database searching algorithm as well as the PTM analysis algorithm to detect the type and location of potential modifications existing in each peptide.

Quantitative estimation of modified peptides derived by biotransformation

Percentage of modified peptides presented in each sample was calculated and summarized in Figure 4 (common modifications of heavy chain) and Table 3 (other modifications). By examining the oxidation profiles, a large percentage of Methionine residues in the biotransformed samples have been oxidized (Figure 4A). The highly oxidized (> 95%) residues included M70o, M254o and M430o of the heavy chain. About 50% of the M106o had also been oxidized. On the other hand, most of the Methionine residues in the control samples remained unmodified (Figure 4B). Only 5%-10% of the M106 were oxidized in the control sample, which could be modified during the storage or sample preparation process. Additionally, the disappearance of unmodified peptides in the biotransformation and control samples was quantitatively estimated. As shown in Figure 4, some unmodified peptides were absent such as M106o and M254o, and some remained a small % after biotransformation, such as M70o. Furthermore, some unmodified peptides were absent in both biotransformation and control samples, such as carbamidomethylated residues C228c and C263c due to sample processing rather than biotransformation.

Summary of biotransformation-derived protein modifications

The modified percentages of other types of modifications including deamidation, carbamidomethylation and dimethylation, were consistent in both sample groups (Figure 4 C&D,

Table 3). The sources of these modifications were mostly from production, storage or sample preparation, which were not directly related to tBHP treatment.

Figure 5 summarized the results of differential analysis with a significance (-10logP) of at least 15 and fold changes at least 2 folds. Peptides confidently detected in at least 3 samples (Figure 5A) and peptides confidently detected in all 6 samples (Figure 5B) were listed. Figure 5C represents the relationship amongst these datasets. As some peptides were undetectable in the control or biotransformed group, features detected in both sets (n=3) samples (in control group or biotransformed group) and all 6 samples were monitored.

Discussion

A rapid and sensitive LC-HRMS based strategy was developed in this study to characterize the biotransformation-mediated modifications in an artificially oxidized mAb model system. Due to the similarity in property and sample handling procedures, this approach is also suitable for other classes of therapeutic biologics, such as therapeutic peptides, ADC and bi-functional biologics.

Most of the previous studies on biotransformation focused on one particular type of modification, such as oxidation (Wang et al., 2011) or N-terminal glutamate to pyroglutamate conversion (Liu et al., 2011). While these *in vivo* studies allow in depth structure elucidation, they are not as informative as a high throughput study to screen all modifications in one run. By using the untargeted LC-HRMS based approach, different modifications can be detected in a high throughput fashion to generate valuable information, while the sources of modifications (storage, sample preparation or treatment) could be potentially unveiled by differential analysis. The uniqueness of this workflow is mainly presented in the following two aspects: 1. Instead of focusing on one modification, all modifications can be examined at one time with the use of bioinformatic tools; 2. By comparing with control samples, the shared modifications were filtered away and only the modifications of interest were retained and thus reported.

The conserved Methionine residues (M254 and M430) located in the FnRn binding site and play important roles in FcRn binding were found to be oxidized in the biotransformation samples. This result is consistent with previous report that oxidizing these two Methionine residues at high level could significantly reduce the serum circulation half-life of mAb (Wang *et al.*, 2011).

Due to the large size of therapeutic biologics and large number of potential modification sites, there exists a large amount of low level modifications widely present in all samples. Most of

these modifications are introduced during production, storage and sample preparations, which are not directly related to the biotransformation process. By performing database searching and PTM analysis using bioinformatic tools, we observed a relatively high level of basal modifications in the control samples. Moreover, most of the modification types and modified peptides are shared by the control and biotransformation samples. Some random modifications were observed in only one replicate of either the control or biotransformation samples. Therefore, it is difficult to pick out the biotransformation-mediated modifications that were induced by the tBHP treatment in this case without using differential analysis. Furthermore, if conventional LBAs are used, these modifications can impact drug detection and are convoluted with the actual biotransformation-mediated modifications results, leading to overestimation of the degree of biotransformation within biologics. In contrast, LC-HRMS based differential analysis enables direct analysis of the biotransformation-mediated modifications without including unspecific modifications, which generates more accurate results and saves time and effort from analyzing irrelevant data.

Moreover, the possible source of modifications can be determined by this approach. If the modifications are induced by certain treatment, significant differences can be observed between biotransformed and control samples; if the modifications are related to storage, differences can be observed between old and fresh samples; if the modifications are associated with sample preparation, it will be universally presented in all sample groups.

It is worth pointing out that there could be discrepancies during the immune-based enrichment process between control and biotransformation samples, as the targeted epitope in the biotransformed sample could be altered during tBHP treatment. Other enrichment methods will be compared in future studies. Also, the quantitative analysis of stoichiometry of modification was based on peak area ratio of the modified and unmodified peptides. However, there could be

differences in ionization efficiency between the modified and unmodified peptide. To precisely determine the actual stoichiometry, an internal standard-based quantification approach should be used in addition to the method developed in this study.

In summary, a novel approach was developed to systematically study the universal biotransformation of therapeutic biologics using LC-HRMS based untargeted differential analysis. With this approach, sequence coverage of the human mAb was above 92% and 37 types of modifications were identified (all point mutations were treated as one type of modification). As many as 38 peptides were significantly changed, up to 21 of them have modifications and up to 18 of the modified peptides with modifications increased in the biotransformed samples compared to the control samples. In this method, different modifications could be rapidly detected, and the sources of modification could be potentially unveiled by differential analysis, as co-existing modifications occurring in both samples are filtered out, revealing the modifications specifically due to biotransformation. This method will be used for studying *in vivo* biotransformation of test proteins in future study.

Due to similarity in property and sample handling procedures, this approach is also suitable for other classes of therapeutic biologics, such as therapeutic peptides, ADCs and bi-functional biologics. This new analytical strategy can be effectively applied to study biotransformation-mediated protein modifications in the process of drug discovery and development.

Acknowledgments

The authors would like to thank Bioinformatics Solution Inc. (Waterloo, ON, Canada) for the technical support on the PEAKS software. MZ would like to thank Pharmaceutical Candidate Optimization, Bristol-Myers Squibb Research and Development for supporting this study. The authors declare that they have no conflict of interest.

Authorship Contributions

Participated in research design: Yao and Zhu

Conducted experiments: Yao and Zhao

Interpreted data and results: Yao, Chen, Mehl and Zhu

Wrote or contributed to the writing of the manuscript: Yao, Chen. Zhao, Mehl, Li and Zhu

Yao and Chen contributed equally.

References

An B, Zhang M, and Qu J (2014) Toward sensitive and accurate analysis of antibody biotherapeutics by liquid chromatography coupled with mass spectrometry. *Drug metabolism and disposition: the biological fate of chemicals* **42:**1858-1866.

Bults P, Bischoff R, Bakker H, Gietema JA, and van de Merbel NC (2016) LC-MS/MS-Based Monitoring of In Vivo Protein Biotransformation: Quantitative Determination of Trastuzumab and Its Deamidation Products in Human Plasma. *Analytical chemistry* **88:**1871-1877.

Cai B, Pan H, and Flynn GC (2011) C-terminal lysine processing of human immunoglobulin G2 heavy chain in vivo. *Biotechnology and bioengineering* **108:**404-412.

Cheloha RW, Maeda A, Dean T, Gardella TJ, and Gellman SH (2014) Backbone modification of a polypeptide drug alters duration of action in vivo. *Nature biotechnology* **32:**653-655.

Domon B and Aebersold R (2006) Mass spectrometry and protein analysis. *Science* **312:**212-217.

Ewles M and Goodwin L (2011) Bioanalytical approaches to analyzing peptides and proteins by LC--MS/MS. *Bioanalysis* **3:**1379-1397.

Hall MP (2014) Biotransformation and in vivo stability of protein biotherapeutics: impact on candidate selection and pharmacokinetic profiling. *Drug metabolism and disposition: the biological fate of chemicals* **42:**1873-1880.

Hamuro LL and Kishnani NS (2012) Metabolism of biologics: biotherapeutic proteins. *Bioanalysis* **4:**189-195.

Han X, He L, Xin L, Shan BZ, and Ma B (2011) PeaksPTM: Mass Spectrometry-Based Identification of Peptides with Unspecified Modifications. *J Proteome Res* **10:**2930-2936.

Kang L, Camacho RC, Li W, D'Aquino K, You S, Chuo V, Weng N and Jian W (2017) Simultaneous catabolite identification and quantitation of large therapeutic protein at the intact lby immunoaffinity capture liquid chromatography−high-resolution mass spectrometry. *Analytical Chemistry* **89**: 6065-6075.

Katsila T, Siskos AP, and Tamvakopoulos C (2012) Peptide and protein drugs: the study of their metabolism and catabolism by mass spectrometry. *Mass spectrometry reviews* **31:**110-133.

Kullolli M, Rock DA and Ma J (2016) Immuno-affinity capture followed by TMPP N-terminus tagging to study catabolism of therapeutic proteins. *Journal of Proteome Research* **16**: 911-919.

Keck RG (1996) The use of t-butyl hydroperoxide as a probe for methionine oxidation in proteins. *Analytical biochemistry* **236:**56-62.

Liu YD, Goetze AM, Bass RB, and Flynn GC (2011) N-terminal glutamate to pyroglutamate conversion in vivo for human IgG2 antibodies. *The Journal of biological chemistry* **286:**11211-11217.

Liu YD, van Enk JZ, and Flynn GC (2009) Human antibody Fc deamidation in vivo. *Biologicals : journal of the International Association of Biological Standardization* **37:**313-322.

Mao Y, Valeja SG, Rouse JC, Hendrickson CL, and Marshall AG (2013) Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-Fourier transform ion cyclotron resonance-mass spectrometry. *Analytical chemistry* **85:**4239-4246.

Prueksaritanont T and Tang CY (2012) ADME of Biologics-What Have We Learned from Small Molecules? *Aaps Journal* **14:**410-419.

Qian R, Ji QC, Arnold ME, Humphreys WG and Zhu M (2011) A Strategy and its implications of protein bioanalysis utilizing high resolution mass spectrometric detection of intact protein. *Analytical Chemistry* **83**:8937-8944.

Waldmann TA (2003) Immunotherapy: past, present and future. *Nature medicine* **9:**269-277.

Walles M, Rudolph B, Wolf T, Bourgailh J, Suetterlin M, Moenius T, Peraus G, Heudi O, Elbast W, Lanshoeft C, and Bilic S (2016) New Insights in Tissue Distribution, Metabolism, and Excretion of [3H]-Labeled Antibody Maytansinoid Conjugates in Female Tumor-Bearing Nude Rats. *Drug metabolism and disposition: the biological fate of chemicals* **44:**897-910.

Wang W, Vlasak J, Li Y, Pristatsky P, Fang Y, Pittman T, Roman J, Wang Y, Prueksaritanont T, and Ionescu R (2011) Impact of methionine oxidation in human IgG1 Fc on serum half-life of monoclonal antibodies. *Molecular immunology* **48:**860-866.

Xu X and Vugmeyster Y (2012) Challenges and Opportunities in Absorption, Distribution, Metabolism, and Excretion Studies of Therapeutic Biologics. *Aaps Journal* **14:**781-791.

Yates JR, 3rd and Kelleher NL (2013) Top down proteomics. *Analytical chemistry* **85:**6151.

Zhang D, Yu SF, Ma Y, Xu K, Dragovich PS, Pillow TH, Liu L, Del Rosario G, He J, Pei Z, Sadowsky JD, Erickson HK, Hop CE, and Khojasteh SC (2016) Chemical Structure and

Concentration of Intratumor Catabolites Determine Efficacy of Antibody Drug Conjugates. *Drug metabolism and disposition: the biological fate of chemicals* **44:**1517-1523.

Zhang H, Cui W, and Gross ML (2014) Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. *FEBS letters* **588:**308-317.

Zhao L, Ren TH, and Wang DD (2012) Clinical pharmacology considerations in biologics development. *Acta pharmacologica Sinica* **33:**1339-1347.

Footnotes

Yao, M. and Chen, B. contributed equally. L.L. acknowledges financial support from National Institutes of Health [R56MH110215], a Vilas Distinguished Achievement Professorship and a Janis Apinis Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

Figure Legends

Figure 1. Workflow of LC-HRMS based differential analysis for detecting biotransformation-mediated modifications.

Figure 2. Representative chromatograms that show significant differences between biotransformed and control groups of samples.

Figure 3. Representative MS/MS spectra of *m/z* 1095.530 (A) and *m/z* 1103.523 (B), z=2 for both ions. Peptide sequences and fragment ions were annotated on the spectra.

Figure 4. Partial PTM profiles of biotransformation (A,C) and control (B,D) samples. Panels a and b compared the oxidation profiles and panels c and d compared all modifications. The column in red represented the modified peptides and the column in blue represented unmodified peptides. In the x-axis, the capital letter represented corresponding amino acid residues, the number represented the positions of these modifications and the last letter represented the type of modifications: d = deamination; c = carbamidomethylation; o = oxidation. Y-axis represented the percentage of unmodified or modified peptides related to the sum of this peptide with all modifications. In panel C, D314r, L330o and Q349d were not detected in the biotransformation sample.

Figure 5. Features detected by differential analysis with significance (-10logP) of 15 in at least 3 samples (A) and all 6 samples (B) and a Venn diagram (C) representing the relationships among these datasets. The X-axis represents the fold changes of 2, 5, 10, 20, 50 and 100.

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Tables

Table 1. Partial List of PTMs detected by PEAKS PTM algorithm.

Name	#PSM	ΔMass	Position	-10lgP
Carbamidomethyl	1638	57.02	C	200
Carbamidomethyl	1407	57.02	DEHK,N-term	132.5
Deamidation	964	0.98	NQ	200
Oxidation	223	15.99	M	155.8
Dihydroxy	143	31.99	CFKPRW	126.5
Oxidation	129	15.99	HW	122.7
Ubiquitin	80	114	CKST	84.31
Ammonia loss	59	-17.03	N	130.8
Ammonium	56	17.03	DE,C-term	53.12
Hydroxylation	39	15.99	DKNPY	72.49
Pyro-glu from Q	34	-17.03	N-term	105.1
Dehydration	33	-18.01	DSTY,C-term	87.02
DehydroalaC	26	-33.99	C	104.9
Trifluoro	23	53.97	L	43.29
Cysteic acid	22	47.98	C	83.85
Amidation	19	-0.98	C-term	56.34
Dethiomethyl	18	-48	M	43.77
Methyl ester	15	14.02	DE,C-term	38.11
Methylation	14	14.02	S	42.68
Carbamylation	13	43.01	N-term	35.57
Sulfation	13	79.96	ST	133.3
Carboxylation	12	43.99	DK	55.7
Pyro-cmC	10	39.99	N-term	87.79
Carbamidomethyl DTT	10	209	C	62.44

Table 2. Partial list of significantly changed peptides by differential analysis. In the sample profile and group profile columns, the left panel represented the biotransformed samples and the right panel represented the control samples. In the sample profile and group profile, green represented that the peptide was detected at lower concentration and red represented that the peptide was detected at higher concentration Sample 1, 2 and 3 represented the biotransformation samples and sample 4, 5, and 6 represented the control samples.

Peptide	Significance (-10lgP)	m/z	z	Avg. Area	Sample Profile	Group Profile	PTM
DTLMISR	29.63	418.2209	2	1.40E+04			
DTLM(+15.99)ISR	18.98	426.2198	2	2.70E+04			Oxidation (M)
SVMHEALHNHYTQK	16.39	565.6129	3	660.4			
SVM(+15.99)HEALHNHYTQK	7.72	428.4561	4	180.6			Oxidation (M)
SC(+57.02)SVMHEALHNHYTQK	17.57	486.2236	4	513.1			Carbamidomethylatio
SC(+57.02)SVM(+15.99)HEALHNHYTQK	12.71	490.2209	4	238.2			Carbamidomethylatio Oxidation (M)
TTAMDYWGQGSLVTVSSASTK	12.94	730.6826	3	666.4			
TTAM(+15.99)DYWGQGSLVTVSSASTK	17.17	1103.523	2	6.20E+03			Oxidation (M)

Table 3. Summary of peptides that underwent other significant biotransformation. The mass difference of each PTM was shown in the peptide sequence: deamidation (+0.98), carbamidomethylation (+57.02) and dimethylation (+28.03).

Native peptides	Modified peptides	% Modified	
		Biotransformed	Control
TVLHQDWLNGK	TVLHQDWLN(+0.98)GK	45.5	36.6
VVSVLTVLHQDWLNGK	VVSVLTVLHQDWLN(+0.98)GK	76.3	39.6
C(+57.02)PAPELLGGPSVFLFPPKPK	C(+57.02)(+28.03)PAPELLGGPSVFLFPPKPK	55.5	44.3
EEQYNSTYR	EEQYN(+0.98)STYR	42	10.5
FNWYVDGVEVHNAK	FNWYVDGVEVHN(+0.98)AK	3.3	4.8
HQDWLNGK	HQDWLN(+0.98)GK	64	51
IC(+57.02)NVNHKPSNTK	IC(+57.02)NVN(+0.98)HKPSNTK	6.4	0

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

Figures

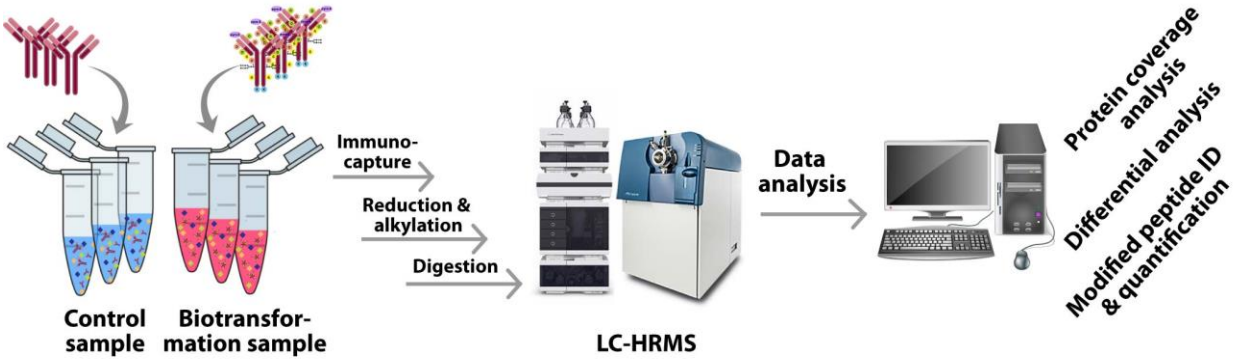

Figure 1

Figure 1.

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

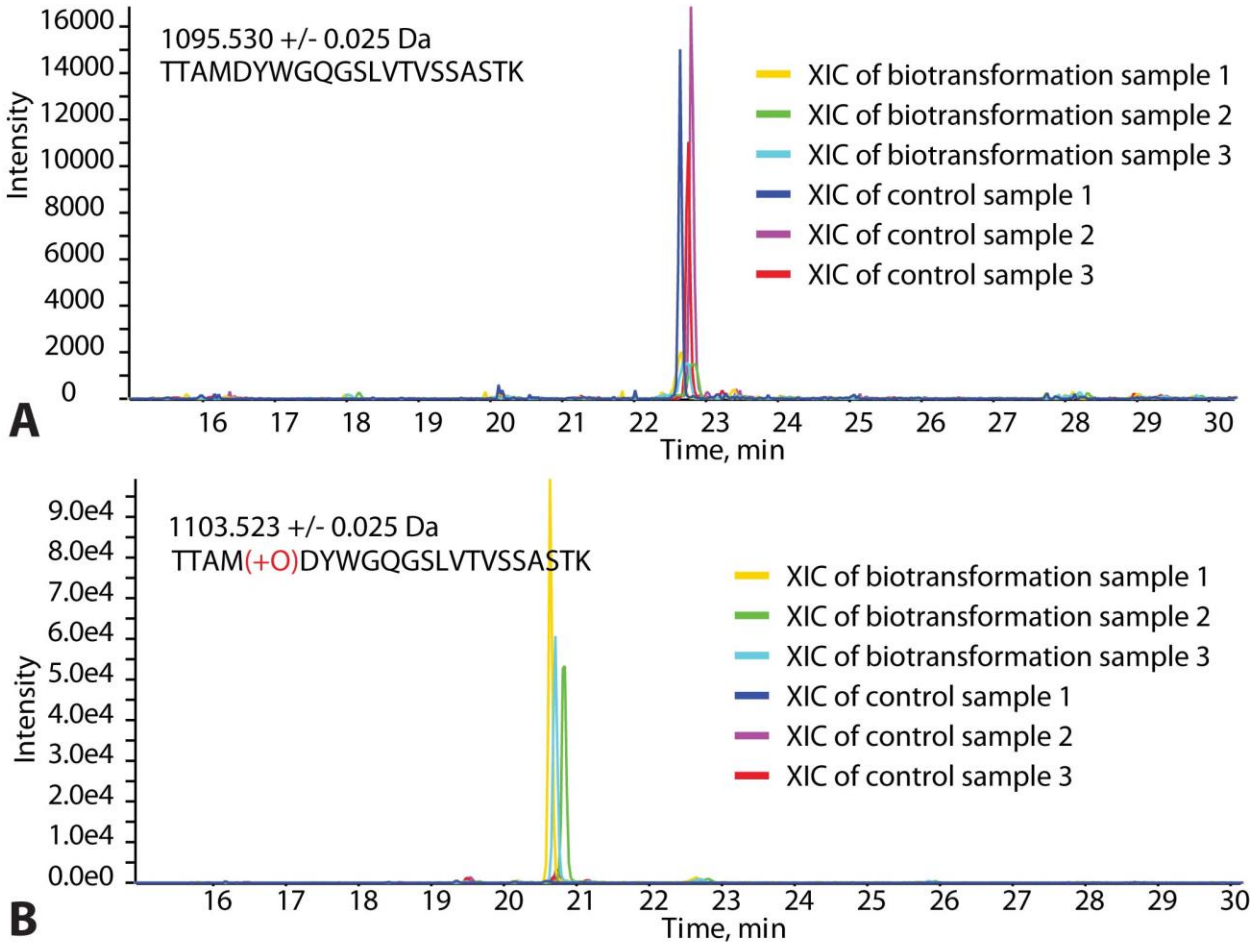

Figure 2.

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

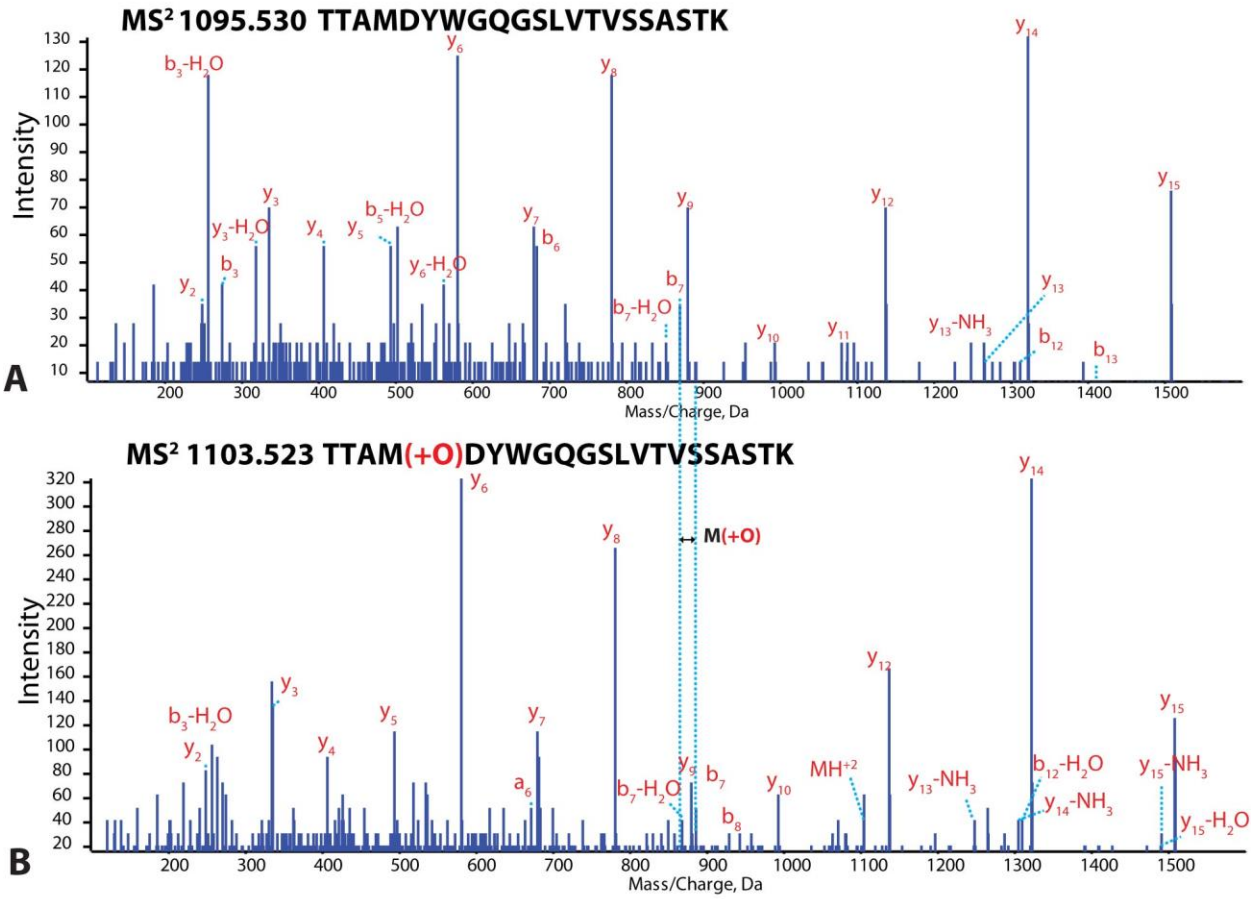

Figure 3

Figure 3.

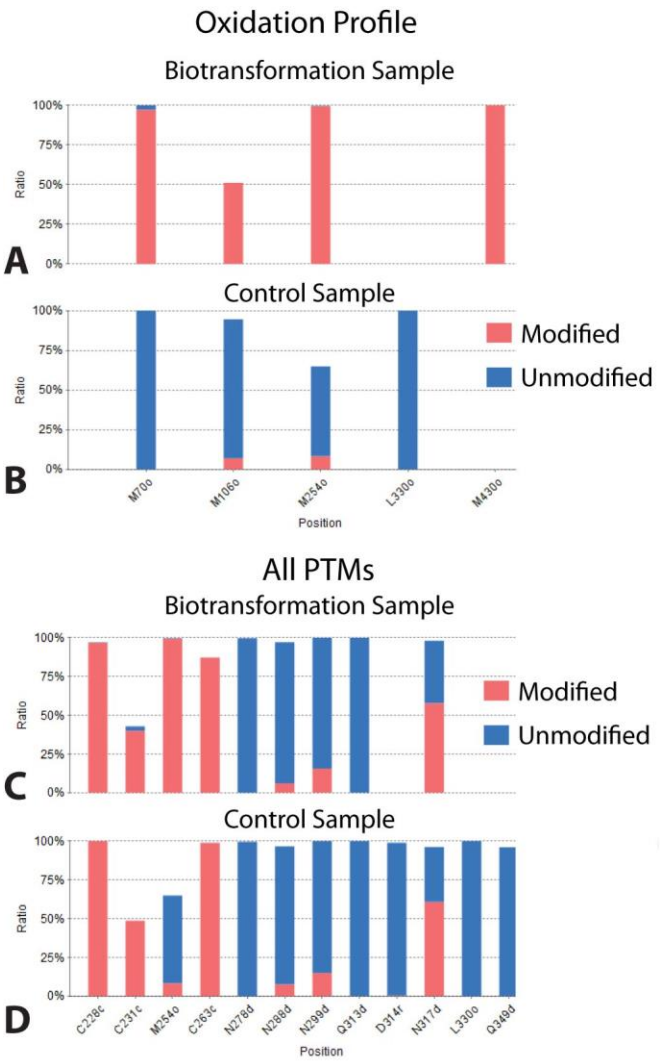

Figure 4

Figure 4.

DMD Fast Forward. Published on January 31, 2018 as DOI: 10.1124/dmd.117.077792
This article has not been copyedited and formatted. The final version may differ from this version.

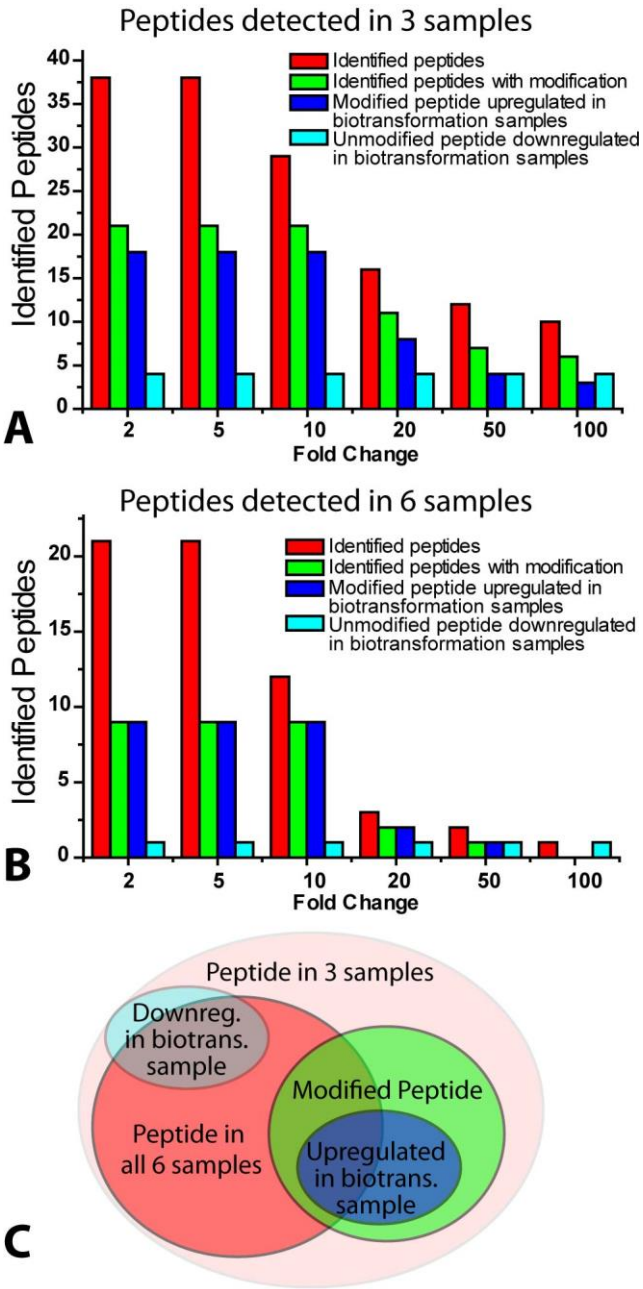

Figure 5

Figure 5.