Title

A strategy for the quantification of protein polyethylene glycol (PEG) derivatized sites using iTRAQ

Author

Monk, Rebecca Helen

Publication Date

2011

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

A Strategy for the Quantification of Protein Polyethylene Glycol (PEG) Derivatized Sites using iTRAQ™

A Thesis submitted in partial satisfaction of the

requirements for the degree of Master of Science

in

Chemistry

by

Rebecca Helen Monk

Committee in charge:

Professor Elizabeth A. Komives, Chair
Professor Michael Burkart
Professor Nuno Bandeira

2011

The Thesis of Rebecca Helen Monk is approved,

and it is acceptable in quality and form for publication on

microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2011

TABLE OF CONTENTS

Signature page……..……. iii

Table of Contents…….....iv

List of Abbreviations …… vii

Lists of Figures…….……..… viii

Abstract……..……ix

Chapter I **Introduction**…….…… 1

A. ADI…….…….…….. 2

B. Protein PEGylation…….…….…….. 2

C. Quantification of PEGylation Sites……..…….…… 3

D. References……..4

Chapter II **Identification of Succinyl Modification**……. 6

A. Introduction……..……..7

B. Materials and Methods……...……..……..……...7

1. Materials…..……..……..…..…….. 7

2. Protein dePEGylation……...……....…..... 7

3. Reduction and Alkylation……..……...…..…..…….. 8

4. Trypsin Digestion…….……..…….…… 8

5. Glu-C Digestion…….……..…….……...8

6. Arg-C Digestion……...9

7. Pepsin Digestion…… 9

8. Sample Resuspension…….....…… 10

9. Mass Spectrometry Method……. 10

10. ProteinPilot Analysis……... 11

C. Results……..……...……..11

1. Experimental Workflow……..…..... 11

2. Base Treatment…… 11

3. Succinyl Modification…….....12

D. Discussion……..……..……...…….. 13

E. References……. 22

Chapter III Quantification of PEGylated Lysines…….23

A. Introduction…….……..…….24

B. Materials and Methods……..…….24

1. Materials……...……...……..….. 24

2. Protein dePEGylation…….…..…..……..……25

3. Reduction Alkylation …….……..…….……..25

4. Trypsin Digestion……...…….…...……. 25

5. Glu-C Digestion……...…….…...……26

6. Sample Resuspension……....……. 26

7. Mass Spectrometry Method…….……26

8. iTRAQ……..…….……. 27

9. ProteinPilot Analysis…….. 27

10. Normalization……. 28

C. Results……..……..……..…..28

1. iTRAQ…….……...…...……...…… 28

2. Peptide Selection Criteria…… 29

3. Calculating Percent PEGylation…….. 29

D. Discussion……...……..…….30

References…….……...……40

LIST OF ABBREVIATIONS

PEG	Polyetheylene glycol
Da	Dalton
HPLC	High Pressure Liquid Chromatography
MALDI-TOF	Matrix-Assisted Laser Desorption Ionization with Time of Flight Detection
Q-TOF MS	Quadrapole Time-of- Flight Mass Spectrometer
TCEP	tris-carboxyethylphosphine
TFA	Trifluoroacetic Acid
FA	Formic Acid

LIST OF FIGURES

Figure 2.1. Schematic diagram of experimental workflow…… 14

Figure 2.2. Base treatment for protein dePEGylation……. 15

Figure 2.3. Gel of dePEGylation…….....16

Figure 2.4. Schematic of y- and b- ion series……...17

Figure 2.5. Evidence of +100.01Da mass shift……..... 18

Figure 2.6. Map of ADI protein sequence……..……...……...19

Figure 2.7. Sequence coverage of ADI…… 20

Figure 2.8. Final list of lysines observed with the succinyl modification….....21

Figure 3.1. iTRAQ workflow……...32

Figure 3.2. Chemical description of the iTRAQ reagents…….….. 33

Figure 3.3. Results from control experiments……. 34

Figure 3.4. Results from normalization peptides……. 35

Figure 3.5. Reporter ion region for normalization peptides……36

Figure 3.6. Summary of iTRAQ quantitation using Glu-C……..37

Figure 3.7. Summary of iTRAQ quantitation using trypsin…….38

Figure 3.8. Sequence map of percent PEGylation for lysines……..39

ABSTRACT OF THE THESIS

A Strategy for the Quantification of Protein Polyethylene Glycol (PEG) Derivatized Sites using iTRAQ™

by

Rebecca Helen Monk

Master of Science in Chemistry

University of California, San Diego, 2011

Professor Elizabeth A. Komives, Chair

Polyethylene glycol (PEG)-derivatized proteins are important pharmaceuticals and the characterization of their PEGylation patterns are of great importance to regulatory agencies for their approval. Accurate quantification of the PEGylated sites are technically challenging. The PEG moiety attached to the protein is very bulky and hinders protease digestion, with the removal of PEG a succinyl linker (100.01 Da) is still attached and can be used for identification of PEGylated sites. The dePEGylated protein can also be used in comparison to the Native protein to determine the percent occupancy of PEG on lysine. Here, we present an improved mass spectrometry based method using iTRAQ™ reagents (AB SCIEX) to quantitate ADI, a PEGylated protein. By using LC coupled with tandem mass spectrometry (QTOF), PEGylated lysines are characterized by measuring

the ratio of 114:115 iTRAQ™ reagent tags, on Native and dePEGylated protein. 24 of 27 lysines were determined to have a succinyl modification. 14 of 27 lysines were quantified to determine their percent occupancy.

Chapter I

Introduction

A. ADI-PEG20

ADI is an arginine deiminase, a novel anticancer enzyme that produces depletion of arginine. Certain tumors are auxotrophic for arginine and without it their growth is inhibited. ADI has a short serum half-life and, as a microbial enzyme, is highly immunogenic (Holtsberg, Ensor et al. 2002). For many proteins including ADI, attachment of polyethylene glycol (PEG) to the protein has been found to extend the stability of therapeutic proteins in human serum and reduce their immunogenicity while maintaining their catalytic activities (Vestling, Murphy et al. 1993). Typically, a 20kDa PEG is attached to the primary amine of lysine via a succinate linker. ADI contains 27 lysines in the protein sequence, and therefore there are 27 potential PEGylation sites on ADI-PEG20.

B. Protein PEGylation

Protein PEGylation is the covalent attachment of PEG. PEG is water soluble, non-toxic, biocompatible, and FDA approved which makes it an ideal group to add to a drug (Veronese and Pasut 2005).

There are multiple variations in the attachment of PEG to a molecule which include; the molecular weight of the PEG group attached, the structure of the PEG chain, the location of the PEG on the polypeptide, and the chemistry used to attach to the polypeptide (Roberts, Bentley et al. 2002). There are various mechanisms that can be used based on these different attachments of PEG. Most importantly, the PEG has to be activated by preparing a derivative of the PEG which will have a functional group at one or both termini, allowing the reaction of an electrophilically activated PEG and a

nucleophilic amino acid. The functional group is chosen based on the type of available reactive group on the molecule that will be coupled with PEG. It is most common to attach the PEG to the amino group of lysine (Roberts, Bentley et al. 2002). This is the case for ADI when a 20kDa PEG is attached to the ε-amino group of lysine via a succinate linker. There are 27 lysines in the ADI sequence and therefore there are multiple PEGylated lysines, which results in a heterogeneous mixture. The succinate group may also react with the N-terminal amino group.

C. Quantification of PEGylation Sites

When there are multiple possible PEGylation sites, it is difficult to determine which sites, and to what extent these are PEGylated. PEGylation can vary and therefore, from site to site, the percent occupancy can differ. Mass Spectrometry based analytical methods have been developed for the characterization of the lysine PEGylation sites in proteins to support patent registration, clinical testing, and quality control in manufacturing. These methods often only identify the PEGylated sites on the protein and use inaccurate and cumbersome quantification schemes to determine the lysine-PEG occupation within the protein. One method used a biotin-PEG-N-hydroxyl succinimide derivative to find PEG conjugation sites in a PEGylated protein. Reverse-phase high-pressure liquid chromatography coupled with Matrix Assisted Laser Desorption Ionization (MALDI) was then used to determine the PEG conjugated residues as well as the preferential residues which were PEGylated (Lee and Park 2003). This method only allowed for the determination of which sites on the polypeptide were PEGylated and did not allow for quantitaion. Another method used a monoclonal IgG antibody to PEG to

detect and quantify PEG-derivatized molecules. These antibodies can bind the PEG backbone, and therefore detect how much PEG has been attached to the protein. This method only allows for the determination of how much PEG was attached in total and it does not allow for the identification of the occupied PEGylated sites (Cheng, Cheng et al. 2005). So far, only proteins with a single lysine that is PEGylated have been approved by the FDA. In the future, it will be necessary to have methods that allow quantification of the amount of PEGylation on multiply-PEGylated proteins. Here, we demonstrate the utility of iTRAQ™ reagents (AB SCIEX) to quantify the PEGylation sites of ADI-PEG20.

D. References

Cheng, T.-L., C.-M. Cheng, et al. (2005). "Monoclonal Antibody-Based Quantitation of Poly(ethylene glycol)-Derivatized Proteins, Liposomes, and Nanoparticles." Bioconjugate Chemistry **16**(5): 1225-1231.

Holtsberg, F. W., C. M. Ensor, et al. (2002). "Poly(ethylene glycol) (PEG) conjugated arginine deiminase: effects of PEG formulations on its pharmacological properties." Journal of Controlled Release **80**(1-3): 259-271.

Lee, H. and T. G. Park (2003). "A novel method for identifying PEGylation sites of protein using biotinylated PEG derivatives." Journal of Pharmaceutical Sciences **92**(1): 97-103.

Roberts, M. J., M. D. Bentley, et al. (2002). "Chemistry for peptide and protein PEGylation." Advanced Drug Delivery Reviews **54**(4): 459-476.

Veronese, F. M. and G. Pasut (2005). "PEGylation, successful approach to drug delivery." Drug Discovery Today **10**(21): 1451-1458.

Vestling, M. M., C. M. Murphy, et al. (1993). "A strategy for characterization of polyethylene glycol-derivatized proteins. A mass spectrometric analysis of the

attachment sites in polyethylene glycol-derivatized superoxide dismutase." Drug Metabolism and Disposition **21**(5): 911-917.

Chapter II

Identification of Succinyl Modification

A. Introduction

ADI-PEG20 contains 27 lysines and in order to quantitate the percent PEGylation of lysines it is important to first map out which sites are PEGylated. To do this we chemically removed the PEG groups using high pH treatment leaving a succinyl linker (Vestling, Murphy et al. 1993). A 5 hour 0.1M NaOH reaction removes the bulky PEG moiety while still leaving a succinyl linker, a modification of $C_4H_5O_3$. This results in a mass shift of +100.01Da, which will be seen in the MS^1 of the parent ion as well as in the y- and b- ion series of MS^2. Observing the mass shift on a lysine containing peptide confirms that site to be PEGylated.

B. Materials and Methods

1. Materials

RapiGest™ SF Surfactant (Waters), TCEP (Thermo), Iodoacetamide (Sigma), endoproteinase trypsin (Roche), Glu-C (Roche), (desalting pipette tips (Thermo), Native protein and PEGylated protein (Polaris Pharmaceuticals).

2. Protein dePEGylation

Chemical dePEGylation proceeds efficiently at pHs above 8. Samples (1000µg) of either Native ADI (1.8µg/µL) or PEG-ADI (11µg/µL) were mixed with 0.2 M NaOH so that the final concentration was 0.1M NaOH (pH 13). This reaction was left for 5hrs at room temperature. After incubation, 1mL of 1M Tris-HCl was added to lower the pH to ~9. The reaction mixture was then buffer exchanged (x2) in an (Amicon 10K cut-off).

The final pH was 8. A BCA assay was then preformed to obtain protein concentration and a gel was run to confirm dePEGylation.

3. Reduction and Alkylation

A sample (10µg) of Native (1.4µg/µL) or dePEGylated ADI (1.5µg/µL) was combined with H_2O (112µL total) and 12.5µL RapiGest (5% in 10x TNE), and then boiled at 100°C for 5min. The mixture was then diluted with 165µL of H_2O to reduce the RapiGest concentration, and 0.05M TCEP was added and incubated at 37°C for 30min. Iodoacetamide at double the concentration of TCEP was added and incubated at 37°C for 30min. Finally, additional TCEP was added to reduce the concentration of IAA.

4. Trypsin Digestion

After reduction and alkylation of the Native and dePEG ADI, trypsin (10µg) was added for protein digestion. Both were incubated O/N at 37°C. To stop the reaction, 1M HCl was added and the reaction was incubated at 37°C for 30min. The sample was then cooled on ice for 4min and centrifuged at 4°C for 30min at max RPM. Supernatant was removed and ammonium hydroxide (1.45M) was added to neutralize the pH, and finally the reaction was dried in the speedvac. The dried samples were redissolved in 20% acetonitrile/80% 2% TFA and then desalted using Thermo Desalting Pipette Tips and dried again. Importantly, trypsin never cleaves after a succinyl-modified lysine, so the dePEGylated protein is expected to be cleaved at fewer positions than the Native protein.

Asp-N digestions were performed in the same manner as the trypsin digestions.

5. Glu-C Digestion

In order to avoid reduction and alkylation, which was determined to be problematic in combination with the Glu-C digestion, samples of 10µg of Native or dePEG ADI were dissolved to a final concentration of 100mM ammonium bicarbonate and boiled for 3 minutes to denature the proteins. Glu-C (10µg) was added and the proteins were digested O/N at 37°C. Thermo Desalting Pipette Tips were used to stop the reaction and the samples were dried down.

6. Arg-C Digestion

A different digestion was used for proteins digested with Arg-C. Arg-C (5µg) was combined with 10µg Protein in a total of 85µL of incubation buffer. Incubation buffer consisted of 100mM Tris-HCl, 10mM $CaCl_2$ at pH 7.6). Activation solution (10µL) was then added and incubated O/N at 37°C. The reaction was stopped by adding 25µL 2%TFA 20%ACN and desalting pipette tips used to clean up the reaction. Samples were dried until ready to iTRAQ™.

7. Pepsin Digestion

Immobilized Pepsin beads (50µL, Pierce/Thermo) were activated by addition of 1mL 0.1% FRESH TFA with vortexing, and the beads were collected by centrifugation. The TFA was removed and the process was repeated once more. The supernatant was removed and the pepsin beads were kept on ice until ready to use (not more than 2 hrs). The protein solution (100µL containing 50µg of either Native or DePEGylated ADI) was dissolved to a final concentration of 0.1% TFA, and the pH was checked to ensure it was less than 2.5. The protein was digested by incubation at room temperature for 5min

while vortexing every min. The supernatant was collected, desalted, and analyzed on the QSTAR Elite.

8. Sample Resuspension

Samples were resuspended in 20μL of 5% ACN/ 2% FA and run on a QSTAR Elite.

9. Mass Spectrometry Method

After protein digestion, High Pressure Liquid Chromatography (HPLC) was coupled with tandem mass spectrometry (QSTAR Elite, AB SCIEX) for data collection. Peptides were loaded onto a reverse phase (180 ID, 365 OD) capillary column packed with C18 and capped with 0.2 mm filters on either end connected via 4 cm (25 ID, 365 OD) capillary to the nanospray source. The spray needle was a pulled capillary (180 ID, 365 OD) tip packed with C18. This "post-column" C18 is critical for preventing clogging of the spray tip over many samples. Peptides were slowly eluted off the column and into the mass spectrometer with increasing organic solvent. The acquisition method used is an information dependent acquisition (IDA) method with charge state 2-4 which exceeds 30 counts and former target ions were dynamically excluded for 15 sec and switch after 1 spectrum. A time of flight (TOF) survey scan was first obtained to select for high intensity parent ions that were being eluted off the column. Based on this TOF survey scan, peptides were selected for fragmentation in the collision cell and the fragment ion masses were measured in the TOF. These fragments from the MS^2 scan are used to determine the peptide sequence. This tandem mass-spectrometry method was used to obtain the most data.

10. ProteinPilot Analysis

All data were analyzed using ProteinPilot software version 3.0. In ProteinPilot a succinyl modification (100.01 Da) was added to the search algorithm. The software was then able to search for the modification of $C_4H_5O_3$ on the ε-amino group of lysine. This allowed the software to fit the correct y- and b- ion series to the +100.01 Da mass shift of the MS^2 fragmentation pattern.

C. Results

1. Experimental Workflow

Figure 2.1. is a schematic of the experimental workflow. First the Native and PEGylated proteins were base treated and proteins were digested with trypsin and Glu-C proteases to obtain peptide fragments. These samples were run on the QSTAR Elite (QTOF) to obtain sequence coverage of the protein and identify which lysines contained the succinyl linker modification and therefore were PEGylated. After digestion the same samples were also labeled with iTRAQ reagents for quantification (described in Chapter 3).

2. Base Treatment

Native and dePEGylated ADI were digested and analyzed on MS to determine the sequence coverage of multiple proteases (Arg-C, Asp-N, Glu-C, trypsin, and pepsin). Mass spectrometric analysis of the digested proteins revealed significant differences between the peptides obtained from the digest of the Native protein vs. the digest of the PEGylated protein. We concluded that the accessibility of the cleavage site was

restricted by the PEG moiety. Protein dePEGylation removed the bulky PEG moiety while leaving a succinyl group that could then be used for identification of PEGylation site. Figure 2.2. shows the chemical scheme of the base treatment reaction to remove the PEG moiety. After dePEGylation, SDS polyacrylamide gel electrophoresis showed that the PEG moiety had been removed (Figure 2.3.). Sequence coverage of the dePEGylated ADI was now comparable to the Native protein and could be used for further analysis.

3. Succinyl Modification

MS^1 spectra showed a mass shift of +100.01 Da for several of the lysine-containing peptides suggesting that these peptides contained a succinyl linker. To confirm succinyl modification, the MS^2 fragmentation was used to sequence the peptide. Under high energy collision, peptides fragment in a predictable manner creating the y- and b-ion series (Figure 2.4.). Figure 2.5. shows an MS^2 spectrum of a Glu-C peptide with an internal lysine. The y- and b- ion series confirm the sequence in the native protein as well as the mass shift for the succinyl modification. This modification was added into the ProteinPilot search algorithm and the entire data was searched for peptides containing the succinyl modification. ProteinPilot found 24 of 27 lysines to have a succinyl linker. As expected, the succinyl modification was never found at the C-terminus of a tryptic peptide. Fourteen of these succinyl-lysines were found in multiple peptides from multiple protease digests. Seven additional succinyl-lysines were only found in one or two peptides but with good y-and b- ion series and are included in the total succinyl modifications found. Three succinyl-lysines were not found even though peptides containing these lysines were observed suggesting that these lysines might not be PEGylated. Thus, these three lysines probably have less than 1% PEGylation resulting in

low abundance which is beyond the detection limit. Figure 2.6. is a map of the protein sequence highlighting the PEGylated lysines found.

D. Discussion

Within the ADI protein sequence there are a total of 27 lysines. With 95% sequence coverage and 100% coverage of lysines, a total of 24 lysines were found to have the succinyl modification (Figure 2.7.). This was surprising since it had previously been determined that the PEG to Protein molar ratio was 5 to 1, meaning that for every protein there are 5 PEG molecules. We see that 24 sites are PEGylated, not 5, which means that there is a distribution of PEG. In the dePEGylated protein we observe the same peptide both with and without the succinyl modification. If the lysine were PEGylated 100% of the time, we would expect to see 100% of the succinyl modified peptide and none of the "native" peptide. However, this is not the case, in the population of peptides there is more of the "native" peptide than the succinyl modified peptide. Thus, none of the lysines were observed to be 100% PEGylated, instead many lysines were partially PEGylated.

We only identified 24 PEGylated lysines (Figure 2.8.). The sites that were never detected with the modifications suggest that these lysines might not be PEGylated. Now that the succinyl- modified lysines (PEGylated lysines) have been identified, quantification can be done to determine percent PEGylation of theses sites.

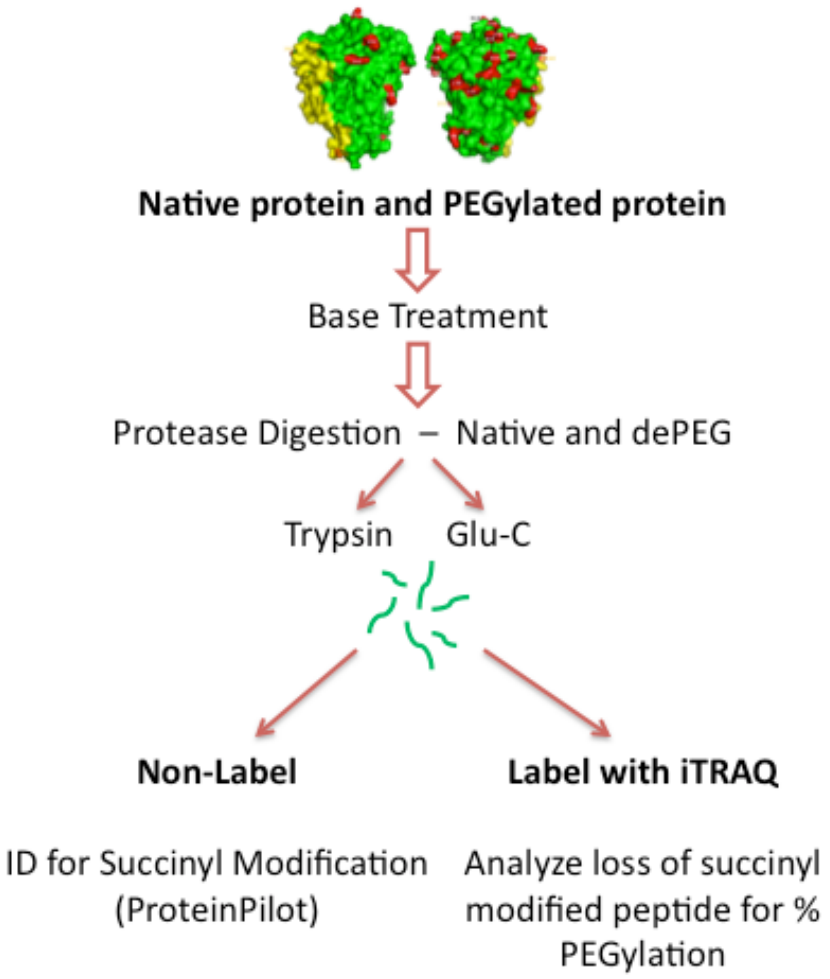


Figure 2.1. Schematic of experimental workflow. Native and PEGylated protein were both base treated and digested with various proteases. First, the identification of succinyl-modified lysines was determined. Then, quantification (explained in Chapter III) for the loss of the succinyl-modified peptide determined the percent PEGylation. All steps are done in parallel to yield direct comparison of Native and PEGylated ADI.

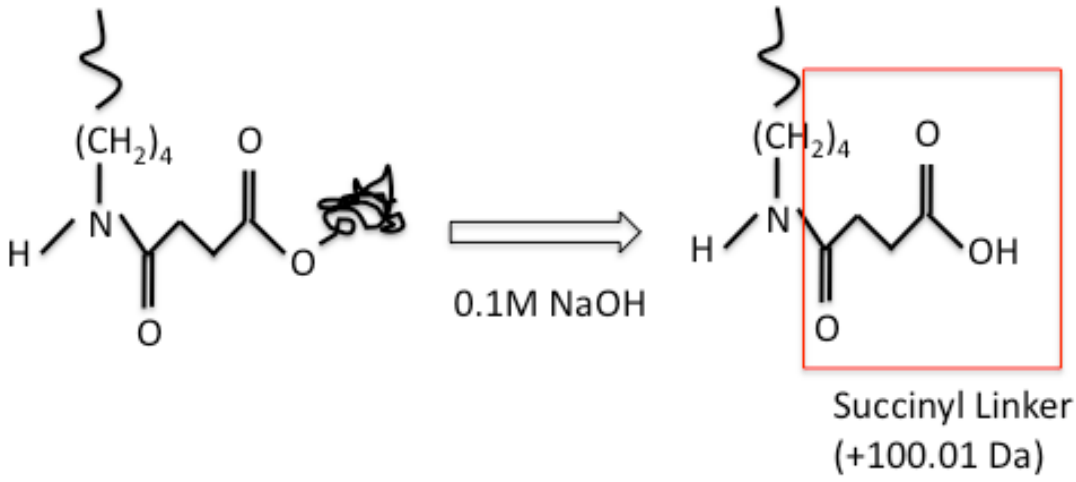


Figure 2.2. Base treatment for protein dePEGylation. 0.1M NaOH hydrolyzes off the PEG moiety leaving a succinyl linker ($C_4H_5O_3$) with mass 100.01 Da.

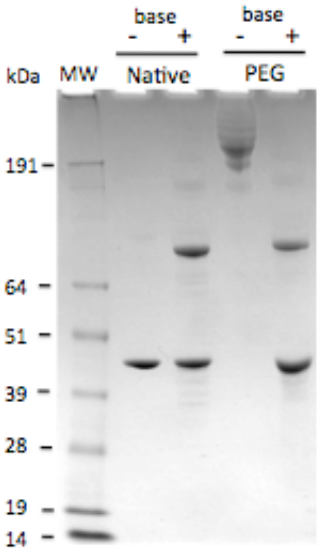


Figure 2.3. Native and PEGylated ADI with and without base treatment. The first lane is the molecular weight marker. 2nd and 4th lanes show no base treatment confirming the 46kDa native ADI as well as the heterogeneous PEGylated ADI. 3rd and 5th lanes show native and PEGylated ADI after base treatment. The top band shows a dimmer and the bottom band a monomer.

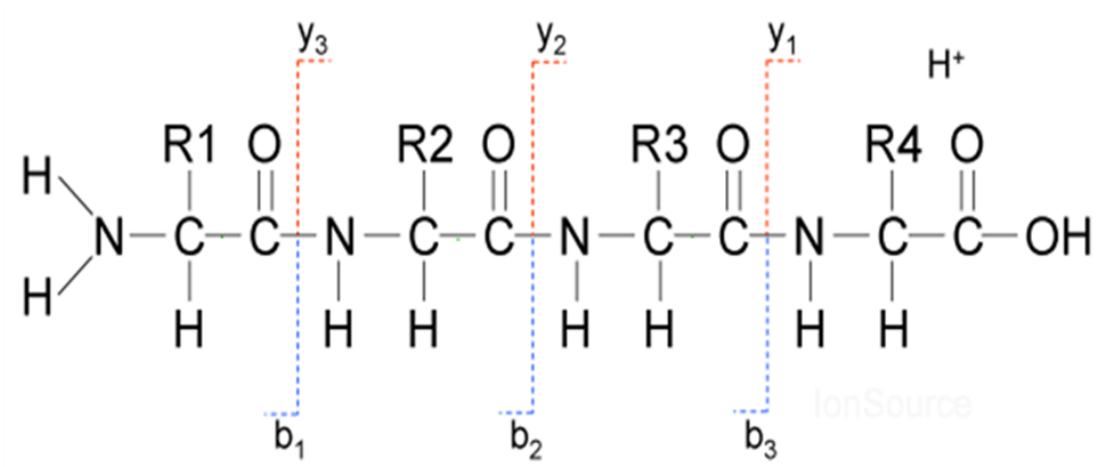


Figure 2.4. Schematic of the y- and b- ion series generated by high energy fragmentation of a peptide.

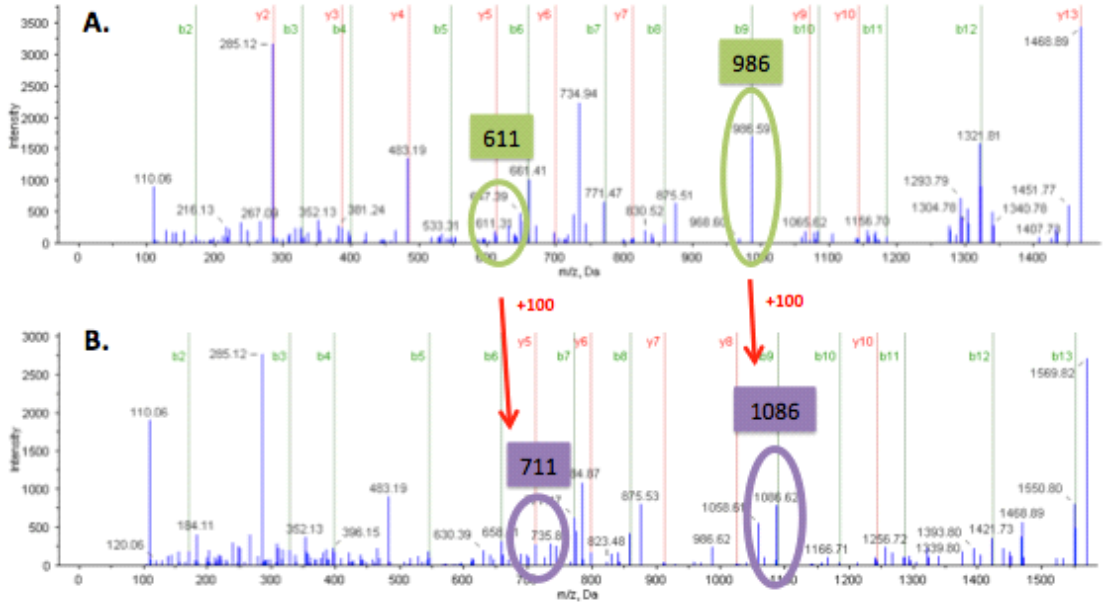


Figure 2.5. Evidence of +100.01 Da mass shift identifying the succinyl linker. ESI data from a +2 charged Glu-C peptide (AVRAFLLSKPTHE). **A.** This spectrum is the MS^2 fragmentation from native ADI with MS^1 parent mass of 1467. The y- and b- ions correlate to the correct sequence ID of AVRAFLLSKPTHE. Highlighted at 611 is the y-ion for K and 986 is the b- ion for K. **B.** This spectrum is the MS^2 fragmentation from dePEG ADI with MS^1 parent mass of 1567. The y- and b- ions correlate to the +100 mass shift confirmed that this lysine has the succinyl modification.

```
        10         20         30         40         50         60
MSVFDSKFNG IHVYSEIGEL ETVLVHEPGR EIDYITPARL DELLFSAILE SHDARKEHQS
        70         80         90        100        110        120
FVKIMKDRGI NVVELTDLVA ETYDLASKAA KEEFIETFLE ETVPVLTEAN KEAVRAFLLS
       130        140        150        160        170        180
KPTHEMVEFM MSGITKYELG VESENELIVD PMPNLYFTRD PFASVGNGVT IHFMRYIVRR
       190        200        210        220        230        240
RETLFARFVF RNHPKLVKTP WYYDPAMKMP IEGGDVFIYN NETLVVGVSE RTDLDTITLL
       250        260        270        280        290        300
AKNIKANKEV EFKRIVAINV PKWTNLMHLD TWLTMLDKNK FLYSPIANDV FKFWDYDLVN
       310        320        330        340        350        360
GGAEPQPQLN GLPLDKLLAS IINKEPVLIP IGGAGATEME IARETNFDGT NYLAIKPGLV
       370        380        390        400
IGYDRNEKTN AALKAAGITV LPFHGNQLSL GMGNARCMSM PLSRKDVKW
```

Figure 2.6. Map of ADI protein sequence indicating lysines with succinyl modification.

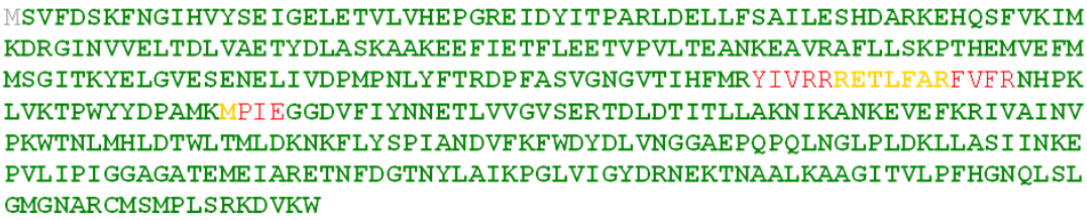


Figure 2.7. Sequence coverage of ADI from all of the different protease digestions. Green denotes >95% confidence in the identity of the peptide, yellow is 50 – 95% confidence, and red is <50% confidence.

K	Trypsin	Glu-C	Arg-C	Pepsin
7	✓	✓	✓	✓
56		✓		✓
(63		✓		
66)		✓		✓
88	✓	✓		✓
91		✓		
111	✓	✓		
(121	✓			
136)	✓	✓		
195	✓		✓	
198			✓	
208	✓			✓
242	✓	✓	✓	
245		✓	✓	
248		✓		
253			✓	
262		✓		✓
(278				✓
280)				✓
292	✓	✓		
316	✓			✓
324	✓	✓		
356	✓	✓		✓
368				
374	✓	✓	✓	
(405				
408)				

K found and ID'd as a succinyl modification
K found with LOW CONFIDENCE and FEW Peptides found with modification
NO Succinyl Modification found = may NOT be PEGylated (low abundance)

Figure 2.8. Final list of which lysines were observed to be modified with the succinyl group under the various protease digestion conditions used.

E. References

Vestling, M. M., C. M. Murphy, et al. (1993). "A strategy for characterization of polyethylene glycol-derivatized proteins. A mass spectrometric analysis of the attachment sites in polyethylene glycol-derivatized superoxide dismutase." Drug Metabolism and Disposition **21**(5): 911-917.

Chapter III

Quantification of PEGylated Lysines

A. Introduction

To quantify the percent PEGylation of the PEG modified lysines we used an isobaric tag for relative and absolute quantitation (iTRAQ™) (Figure 3.1.). This reagent covalently labels the primary amine group of the N-termini and lysines. The same mass is added to peptides from different samples, however each sample is given a different reporter mass. A balance group, which is fragmented in the mass spectrometer's collision cell, is used to keep the tag isobaric (Figure 3.2.). In MS^2 the different reporter masses are identified and the area under the peak is used for comparison of the two samples.

In this experiment we are measuring for the loss of the succinyl-modified lysine. The iTRAQ reagent will not label the amine group of lysine if it has been modified. Peptides labeled at the N-terminus and lysine will co-elute at one time, while, succinyl-modified lysines will elute at another time. The ratio of peptides labeled at the N-terminus and lysine residues will allow us to determine the percent PEGylation based on the lower ratio of iTRAQ labeling on the succinyl-modified lysine.

B. Materials and Methods

1. Materials

RapiGest™ SF Surfactant (Waters), TCEP (Thermo), Iodoacetamide (Sigma), endoproteinase Trypsin (Roche), Glu-C (Roche), desalting pipette tips (Thermo), Native protein and PEGylated protein (Polaris Pharmaceuticals).

2. Protein dePEGylation

Chemical dePEGylation proceeds efficiently at pHs above 8. Samples (1000µg) of either Native ADI (1.8µg/µL) or PEG-ADI (11µg/µL) were mixed with 0.2 M NaOH so that the final concentration was 0.1M NaOH (pH 13). This reaction was left for 5hrs at room temperature. After incubation, 1mL of 1M Tris-HCl was added to lower the pH to ~9. The reaction mixture was then buffer exchanged (x2) in an (Amicon 10K cut-off). The final pH was 8. A BCA assay was then preformed to obtain protein concentration and a gel was run to confirm dePEGylation.

3. Reduction and Alkylation

A sample (10µg) of Native (1.4µg/µL) or dePEGylated ADI (1.5µg/µL) was combined with H_2O (112µL total) and 12.5µL RapiGest (5% in 10x TNE), and then boiled at 100°C for 5min. The mixture was then diluted with 165µL of H_2O to reduce the RapiGest concentration, and 0.05M TCEP was added and incubated at 37°C for 30min. Iodoacetamide at double the concentration of TCEP was added and incubated at 37°C for 30min. Finally, additional TCEP was added to reduce the concentration of IAA.

4. Trypsin Digestion

After reduction and alkylation of the Native and dePEG ADI, Trypsin (10µg) was added for protein digestion. Both were incubated O/N at 37°C. To stop the reaction, 1M HCl was added and the reaction was incubated at 37°C for 30min. The sample was then cooled on ice 4min and centrifuged at 4°C for 30min at max RPM. Supernatant was removed and ammonium hydroxide (1.45M) was added to neutralize the pH, and finally the reaction was dried in the speedvac. The dried samples were redissolved in 20%

acetonitrile/80% 2% TFA and then desalted using Thermo Desalting Pipette Tips and dried again. Asp-N digestions were performed in the same manner as the trypsin digestions.

5. Glu-C Digestion

In order to avoid reduction and alkylation, which was determined to be problematic in combination with the Glu-C digestion, samples of 10μg of Native or dePEG ADI were dissolved to a final concentration of 100mM ammonium bicarbonate and boiled for 3 minutes to denature the proteins. Glu-C (10μg) was added and the proteins were digested O/N at 37°C. Thermo Desalting Pipette Tips were used to stop the reaction and the samples were dried down.

6. Sample Resuspension

Samples were resuspended in 12μL of 5% ACN/ 2% FA and run on a QSTAR Elite (AB/Sciex) to check for identification of the protein. iTRAQ™ was then done for quantitation.

7. Mass Spectrometry Method

After protein digestion, High Pressure Liquid Chromatography (HPLC) was coupled with tandem mass spectrometry (QSTAR Elite) for data collection. Peptides were loaded onto a reverse phase (180 ID, 365 OD) capillary column packed with C18 and capped with 0.2 mm filters on either end connected via 4 cm (25 ID, 365 OD) capillary to the nanospray source. The spray needle was a pulled capillary (180 ID, 365 OD) tip packed with C18. This “post-column” C18 is critical for preventing clogging of the spray tip over many samples. Peptides were slowly eluted off the column and into the

mass spectrometer with increasing organic solvent. The acquisition method used is an information dependent acquisition (IDA) method with charge state 2-4 which exceeds 30 counts and former target ions were dynamically excluded for 120 sec and switch after 1 spectrum. A time of flight (TOF) survey scan was first obtained to select for high intensity parent ions that were being eluted off the column. Based on this TOF survey scan, peptides were selected for fragmentation in the collision cell and the fragment ion masses were measured in the TOF. These fragments from the MS^2 scan are used to determine the peptide sequence. The method also enhanced the low mass fragment intensity multiplier specifically for detection of the iTRAQ reporter ions.

8. iTRAQ™

Two iTRAQ™ reagents with reporter ion masses of 114 and 115 were used. The vials of reagents were warmed to room temperature and 70µL ethanol was added to dissolve each. Dissolution buffer (30µL) was added to each digested sample (10 µg total protein) to ensure a pH > 7, and an iTRAQ™ reagent was combined to each digest. The Native protein digest was labeled with the 114 reporter ion reagent and dePEGylated protein digest was labeled with 115. Samples were then incubated at room temperature for 1hr. Contents of each iTRAQ™ reagent sample were combined (114+115 Trypsin digest and 114+115 Glu-C digest). Samples were then dried in the speed vac to remove ethanol, and 2% TFA 20% ACN was added. The samples were purified using desalting pipette tips, dried, resuspended, and run on the QSTAR Elite (AB SCIEX) using the method described above.

9. ProteinPilot Analysis

Data were analyzed using ProteinPilot version 3.0. The reporter ion ratio of 114:115 was used to determine the percent PEGylation of lysine containing peptides as described further in the Results section. The spectra were all manually checked in the Analyst software (QSTAR version 2.0).

10. Normalization

Due to slight differences in the protein concentration, normalization is required for iTRAQ quantitation. This procedure accounts for errors in concentration and pipetting errors throughout the experiment. Only non-lysine containing peptides can be used for normalization. Their reporter ion ratio should always be 1:1 since there should be no difference in the native and dePEGylated proteins. The ratios of the normalization peptides were averaged and used to correct the ratios of all other lysine-containing peptides.

C. Results

1. iTRAQ

We first analyzed the normalization peptides to make sure the iTRAQ reactions were working as expected. Although all of the non-lysine containing peptides from the Tryptic digest always had a 1:1 ratio, this was not the case for the non-lysine containing peptides from the Glu-C digest. The normalization peptides from the Glu-C digest had a wide range of ratios. Using only trypsin as a protease did not allow for enough coverage of PEGylated ADI. Control experiments for iTRAQ labeling using a 1:1 ratio in Native/Native or a 1:1 ratio in dePEG/dePEG always gave a 1:1 ratio for the

normalization peptides from different protease digestions (Figure 3.3.). This result suggested that it was something in the Glu-C digestion, and not the difference in Native and dePEG proteins that were yielding varying results. Since ADI has only one cysteine in the sequence, we decided to simplify the digest conditions by removing the reduction and alkylation step before the addition of Glu-C. Surprisingly, the simpler digestion conditions resulted in a 1:1 ratio for all of the normalization peptides in Native/dePEG experiment enabling us to use Glu-C digestion to determine percent PEGylation (Figure 3.4., 3.5.).

2. Peptide Selection Criteria

After the iTRAQ samples were analyzed on the mass spectrometer, the data was searched through ProteinPilot to determine the reporter ion 114:115 ratio. This search yielded a lot of data which had to be scrutinized before confirming results. First, we established a cut-off of 90% confidence in the MS^2 spectra for y- and b- ions to make sure the correct peptide was correctly identified. Next, we eliminated any peptides with missed cleavages. We also took only data from peptides that were labeled with iTRAQ reagent at the N-terminus for all peptides and at the lysine in lysine containing peptides. Lastly, we used only the data with sufficient ion counts in MS^1 (by manually checking each spectrum in the Analyst software) to ensure high quality MS^2 spectra of the reporter region.

3. Calculating Percent PEGylation

As previously explained we are quantifying based on the loss of the succinyl lysine. The difference in the reporter ion peaks determines the PEGylated amount. If the

114 and 115 peak areas are equal, then there was no PEGylation. If the lysine was PEGylated, the presence of the succinyl modification on that lysine decreases the amount of 115 reporter ion modification that is possible and this leads to a lower peak area for the 115 reporter ion compared to the 114 reporter ion from the non-PEGylated sample. We therefore calculate the percent PEGylation as 100 – ((114/115) x 100).

Since Glu-C cleaves at Asp and Glu residues, digestion with Glu-C gave many peptides with an internal lysine for which PEGylation could be quantitatively determined (Figure 3.6.). The data from tryptic digestions was more difficult to interpret because trypsin does not cleave after a modified lysine, and therefore only two peptides with internal lysines were obtained (Figure 3.7.). In total, the PEGylation on 12 lysines could be quantified from the Glu-C digest data and PEGylation on two additional lysines could be quantified from the trypsin digest data (Figure 3.8.)

D. Discussion

Within the ADI sequence there are 27 lysines. We were able to quantify the percent PEGylation in 14 of the 27 lysines. The high number of lysines resulted in many peptides containing more than one lysine. This makes the task of quantifying the PEGylated lysines in the sequence more difficult. However, the high number of lysines is not the only hurdle to jump over in this task. The use of proteases is more important. Only being able to use Trypsin and Glu-C limited the amount of peptide fragments that could be obtained throughout the sequence. We obtained excellent coverage of the sequence with pepsin, however the pepsin digestion resulted in many overlapping peptides. Overlapping peptides can not be used for iTRAQ quantitation because the

peptides have to be compared one-on-one and the amount of each overlapping peptide was not consistent for the Native vs. dePEGylated protein. A specific protease with different substrate specificity would yield more peptide fragments and therefore more lysines could be analyzed quantitatively for PEGylation.

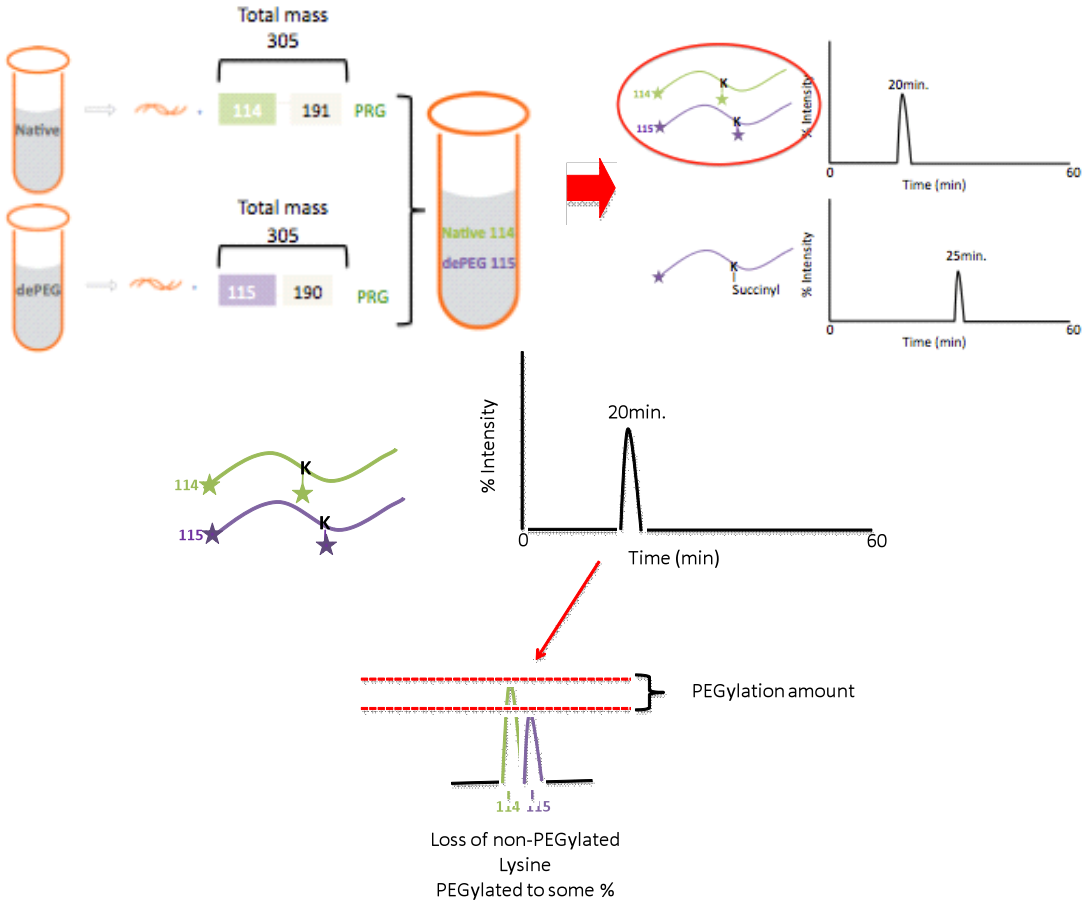


Figure 3.1. iTRAQ workflow to determine percent PEGylation. Native peptides are labeled with iTRAQ reagent 114 and dePEG peptides are labeled with 115. After 1 hour incubation samples are mixed together and analyzed based on the ratio of 114 to 115. Quantitation is based on the loss of the succinyl-modified peptide.

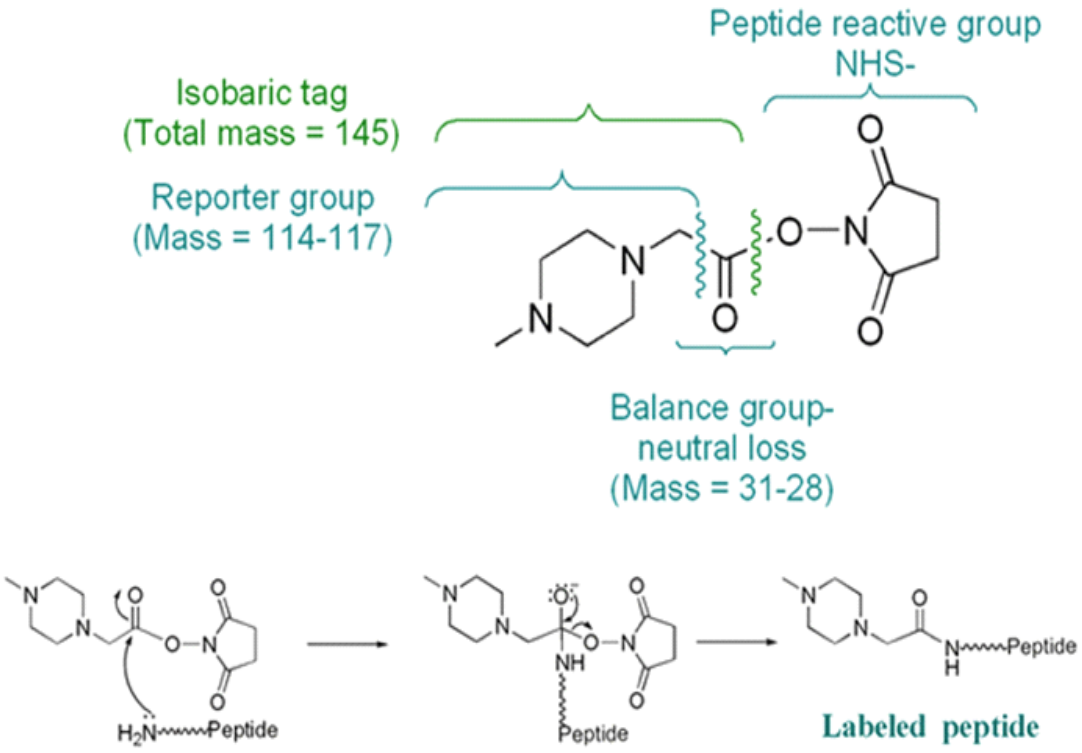
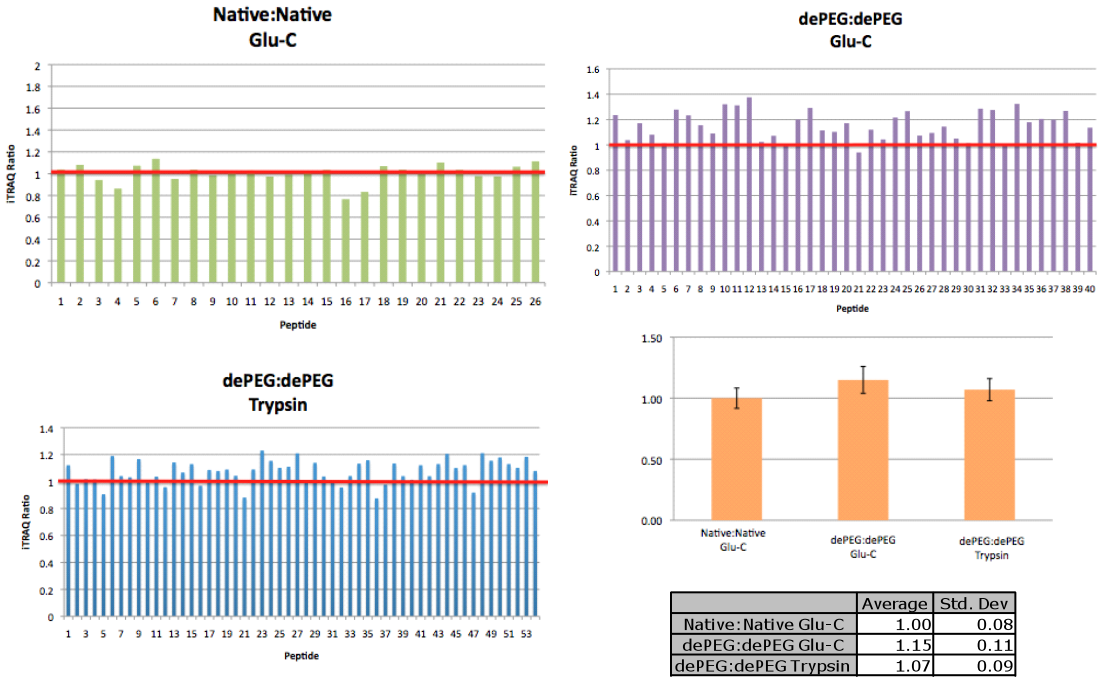
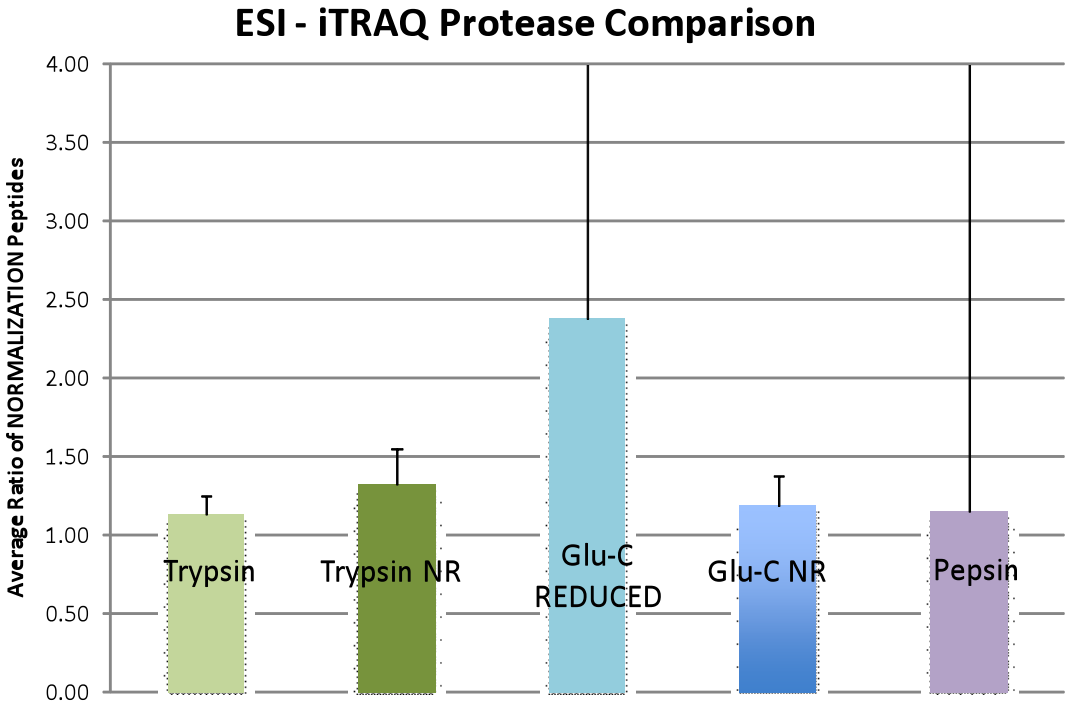


Figure 3.2. Chemical description of the iTRAQ reagents. Each peptide is labeled with a different reporter group. Each reagent also contains a balance group so that the entire label has the same mass even though the mass of the reporter ion is different. During high energy fragmentation, the bond between the reporter ion and the balance group breaks releasing the reporter ion.



	Average	Std. Dev
Native:Native Glu-C	1.00	0.08
dePEG:dePEG Glu-C	1.15	0.11
dePEG:dePEG Trypsin	1.07	0.09

Figure 3.3. Results from control experiments in which the same protein sample was labeled with the 114 and the 115 reporter ion iTRAQ reagent and ratios were obtained. As expected, the ratios were 1:1.



Protease	Average	Standard Deviation
Trypsin	1.1295	0.12
Trypsin NR	1.3236	0.23
Glu-C	2.3791	3.56
Glu-C NR	1.1873	0.20
Pepsin	1.1546	3.99

Figure 3.4. Results from normalization peptides from the various protease digestions of Native (labeled with the 114 reporter ion iTRAQ reagent) compared to dePEGylated (labeled with the 115 reporter ion iTRAQ reagent) proteins. NR denotes no reduction/alkylation was done on the proteins prior to digestion. Simplifying the digestion conditions markedly improved the results from Glu-C digestion, but the results from pepsin digestion remained problematic due to the many partially overlapping peptides obtained.

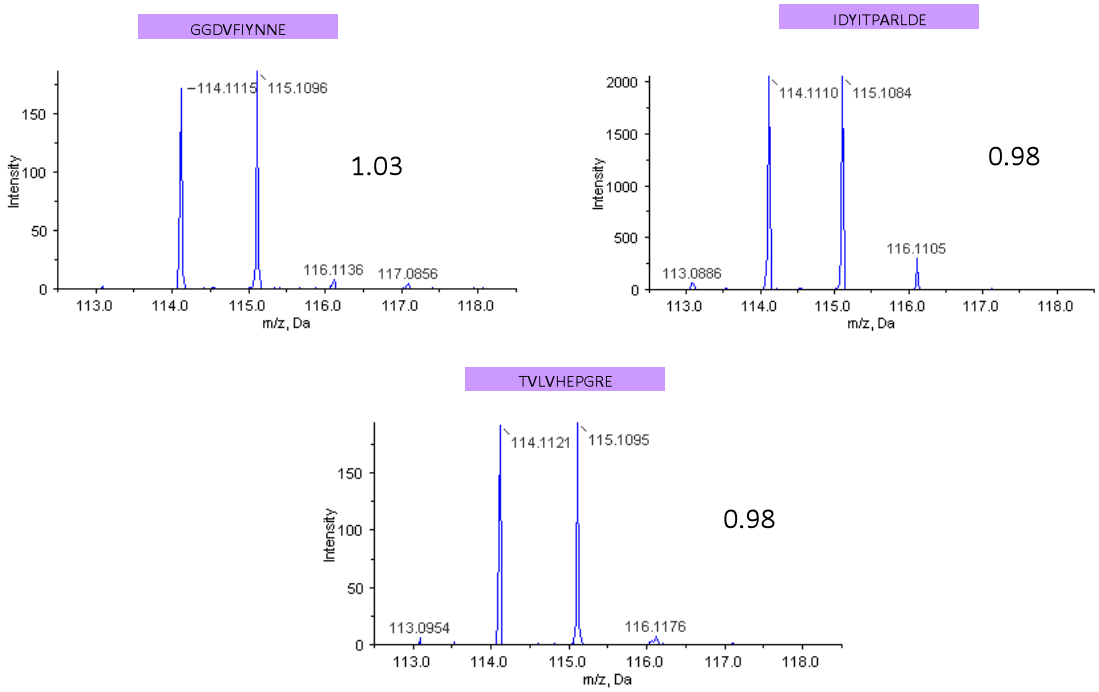


Figure 3.5. Reporter ion region showing the 1:1 ratio of the reporter ions from normalization peptides.

Sequence	% PEG	Std. Dev.	K in Sequence
E.SVFDSKFNGIHVYSE	41	5	7
E.LLFSAILESHDARKE	21		56
E.HQSFVKIMKDRGINVVE	4		63, 66
E.TYDLASKAAKEE	40	7	88, 91
E.TYDLASKAAKEE	8		91
E.TVPVLTEANKE	7		111
E.AVRAFLLSKPTHE	14		121
E.RTDLDTITLLAKNIKANKE	54	4	242, 245, 248
D.GTNYLAIKPGLVIGYDRNE	<5%		356

M.SVFDSKFNGIHVYSE.IGELE.TVLVHEPGRE.IDYITPARLDE.LLFSAILESHDARKE.
HQSFVKIMKDRGINVVE.LTDLVAE.TYDLASKAAKEE.FIETFLEE.TVPVLTEANKE.AVR
AFLLSKPTHE.MVE.FMMSGITKYE.LGVESENE.LIVDPMPNLYFTRDPFASVGNGVTIHFM
RYIVRRRE.TLFARFVFRNHPKLVKTPWYYD.PAMKMPIEGGDVFIYNNE.TLVVGVSE.RT
DLDTITLLAKNIKANKE.VE.FKRIVAINVPKWTNLMHLD.TWLTMLDKNKFLYSPIANDVF
KFWDYDLVNGGAEPQPQLNGLPLDKLLASIINKEPVLIPIGGAGATEME.IARETNFD.GTNY
LAIKPGLVIGYDRNE.KTNAALKAAGITVLPFHGNQLSLGMGNARCMSMPLSRKDVKW

Figure 3.6. Summary of the results from iTRAQ quantitation of lysine modification using Glu-C digestion. The peptides that were quantified are tabulated and shown on the sequence of ADI. Green denotes the peptides from which lysine modification could be obtained. Purple denotes sequence we did not observe but could have based on the expected cleavage by Glu-C.

Sequence	% PEG	Std. Dev.	K in Sequence
R.GINVVELTDLVAETYDLAS**K**	32	6	88
R.ETNFDGTNYLAI**K**PGLVIGYDR	27	4	356
K.FNGIHVYSEIGELETVLVHEPGR	25	6	7
K.YELGVESENELIVDPMPNLYFTR	31	3	136
K.EPVLIPIGGAGATEMEIAR	13	6	324
K.AAGITVLPFHGNQLSLGMGNAR	33	6	374

Only 1 K in sequence
Can use for % PEGylation

K.EEFIETFLEETVPVLTEAN**K**	13	4	91, 111
R.AFLLS**K**PTHEMVEFMMSGIT**K**	28		121, 136
K.TPWYYDPAM**K**	47		198, 208
K.FLYSPIANDVF**K**	84		280, 292
K.FWDYDLVNGGAEPQPQLNGLPLD**K**	47		292, 316
K.LLASIIN**K**	40		316, 324

Multiple K in sequence
None are 1:1
Cannot use in % PEGylation

MSVFDSK.**FNGIHVYSEIGELETVLVHEPGR.EIDYITPAR.LDELLFSAILESHDAR.**KEHQS
FVKIMKD**R.GINVVELTDLVAETYDLASK.**AAK.**EEFIETFLEETVPVLTEANK.**EAVR.**AFLL
SKPTHEMVEFMMSGITK.YELGVESENELIVDPMPNLYFTR.DPFASVGNGVTIHFMR.**YIVRR
RETLFARFVFRNHPKLVK.**TPWYYDPAMK.**MPIEGGDVFIYNNETLVVGVSER.**TDLDTITLLA
K.**NIKANKEVEFKR.**IVAINVPK.**WTNLMHLDTWLTMLDK.NK.**FLYSPIANDVFK.FWDYDLV
NGGAEPQPQLNGLPLDK.LLASIINK.EPVLIPIGGAGATEMEIAR.ETNFDGTNYLAIKPGLV
IGYDR.**NEKTNAALK.**AAGITVLPFHGNQLSLGMGNAR.**CMSMPLSR.KDVKW

Figure 3.7. Summary of the results from iTRAQ quantitation of lysine modification using trypsin digestion. The peptides that were quantified are tabulated and shown on the sequence of ADI. Green denotes the peptides from which lysine modification could be obtained.

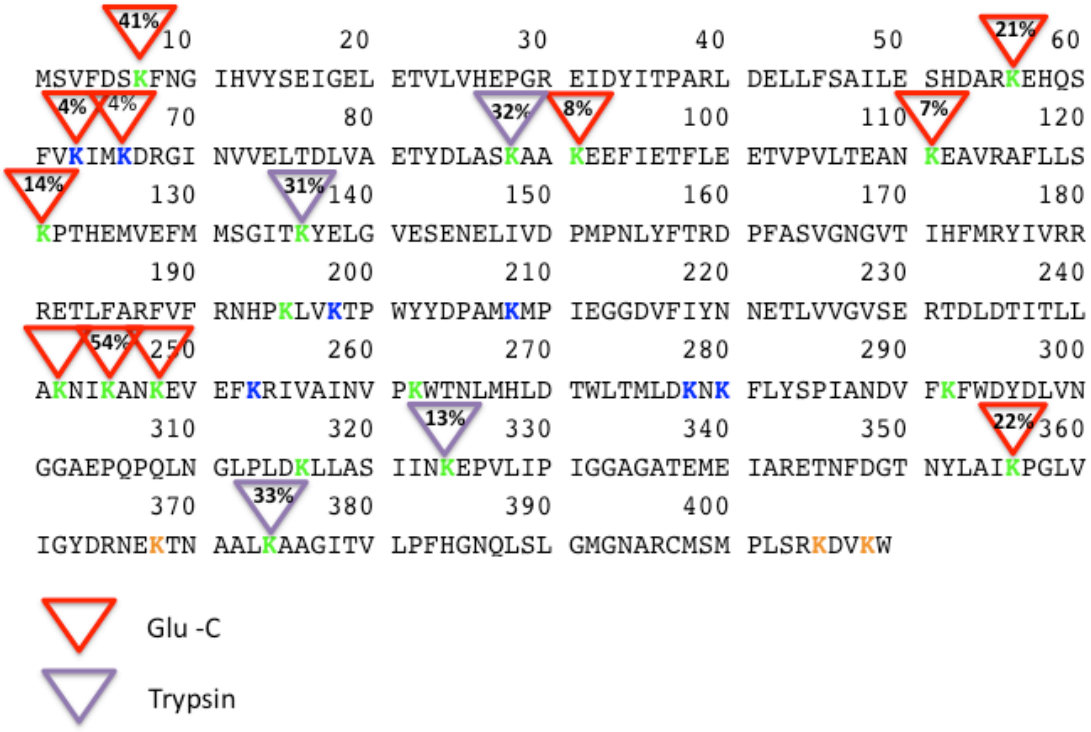


Figure 3.8. Sequence map showing the determination of percent PEGylation for lysines

References

Abuchowski, A., J. R. McCoy, et al. (1977). "Effect of covalent attachment of polyethylene glycol on immunogenicity and circulating life of bovine liver catalase." Journal of Biological Chemistry **252**(11): 3582-3586.

Abuchowski, A., T. van Es, et al. (1977). "Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol." Journal of Biological Chemistry **252**(11): 3578-3581.

Cheng, T.-L., C.-M. Cheng, et al. (2005). "Monoclonal Antibody-Based Quantitation of Poly(ethylene glycol)-Derivatized Proteins, Liposomes, and Nanoparticles." Bioconjugate Chemistry **16**(5): 1225-1231.

Feun, L. and N. Savaraj (2006). "Pegylated arginine deiminase: a novel anticancer enzyme agent." Expert Opinion on Investigational Drugs **15**(7): 815-822.

Fu, J., J. Fiegel, et al. (2002). "New polymeric carriers for controlled drug delivery following inhalation or injection." Biomaterials **23**(22): 4425-4433.

Holtsberg, F. W., C. M. Ensor, et al. (2002). "Poly(ethylene glycol) (PEG) conjugated arginine deiminase: effects of PEG formulations on its pharmacological properties." Journal of Controlled Release **80**(1-3): 259-271.

Lee, H. and T. G. Park (2003). "A novel method for identifying PEGylation sites of protein using biotinylated PEG derivatives." Journal of Pharmaceutical Sciences **92**(1): 97-103.

Roberts, M. J., M. D. Bentley, et al. (2002). "Chemistry for peptide and protein PEGylation." Advanced Drug Delivery Reviews **54**(4): 459-476.

Veronese, F. M. and G. Pasut (2005). "PEGylation, successful approach to drug delivery." Drug Discovery Today **10**(21): 1451-1458.

Vestling, M. M., C. M. Murphy, et al. (1993). "A strategy for characterization of polyethylene glycol-derivatized proteins. A mass spectrometric analysis of the attachment sites in polyethylene glycol-derivatized superoxide dismutase." Drug Metabolism and Disposition **21**(5): 911-917.

Wang, Y.-S., S. Youngster, et al. (2000). "Identification of the Major Positional Isomer of Pegylated Interferon Alpha-2b." Biochemistry **39**(35): 10634-10640.