

**Targeted absolute quantitative proteomics with SILAC internal standards and  
unlabeled full-length protein calibrators (TAQSI)**

Xinwen Wang<sup>1</sup>, Yan Liang<sup>1,2</sup>, Li Liu<sup>1,2</sup>, Jian Shi<sup>1</sup>, and Hao-Jie Zhu<sup>1\*</sup>

<sup>1</sup>Department of Clinical Pharmacy, University of Michigan, Ann Arbor, Michigan

<sup>2</sup>The Key Laboratory of Drug Metabolism and Pharmacokinetics, China Pharmaceutical  
University, Nanjing, China

**Corresponding Authors:**

Hao-Jie Zhu Ph.D.

Assistant Professor

Department of Clinical Pharmacy

University of Michigan College of Pharmacy

428 Church Street, Room 3567A CCL

Ann Arbor, MI 48109-1065

Tel: 734-763-8449, E-mail: [hjzhu@med.umich.edu](mailto:hjzhu@med.umich.edu)

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/rcm.7482](https://doi.org/10.1002/rcm.7482)

## **ABSTRACT**

**RATIONALE:** Liquid Chromatography-Mass Spectrometry (LC-MS)-based proteomics for absolute protein quantification has been increasingly utilized in both basic and clinical research. There is a great need to overcome some major hurdles of current absolute protein quantification methods, such as significant inter-assay variability and high cost associated with the preparation of purified stable isotope-labeled peptide/protein standards.

**METHODS:** We developed a novel targeted absolute protein quantification method, named TAQSI, utilizing full-length isotope-labeled protein internal standards generated from SILAC (stable isotope labeling by amino acid in cell culture) and unlabeled full-length protein calibrators. This approach was applied to absolute quantification of carboxylesterase 1 (CES1), the primary human hepatic hydrolase, in a large set of human liver samples. Absolute CES1 quantities were derived from the standard calibration curves established from unlabeled CES1 protein calibrators and the isotope-labeled CES1 internal standards obtained from SILAC HepG2 cells.

**RESULTS:** The TAQSI assay was found to be accurate, precise, reproducible, and cost effective. Importantly, protein quantification was not affected by various protein extraction and digestion protocols, and measurement errors associated with nonsynonymous variants can be readily identified and avoided. Furthermore, the TAQSI approach significantly simplifies the procedure of identifying the best performance surrogate peptides.

**CONCLUSIONS:** TAQSI assay can be widely used for targeted absolute protein quantification in various biomedical research and clinical practice settings.

**Keywords:** absolute quantitative proteomics; LC-MS/MS; MRM; SILAC; HepG2 cells; human liver; carboxylesterase 1; protein expression

Author Manuscript

## INTRODUCTION

Mass spectrometry (MS)-based absolute targeted protein quantification is being increasingly adopted in proteomics research, and has proven its utility in both basic research and clinical biomarker discovery and validation. Several absolute quantitative proteomics approaches have been developed utilizing different stable isotope-labeled internal standards (IS), such as synthetic peptides (AQUA),<sup>[1-3]</sup> quantification concatemers (QconCATs),<sup>[4, 5]</sup> full-length protein standards (PSAQ) and absolute SILAC (stable isotope labeling by amino acid in cell culture).<sup>[6-8]</sup> A combination of isotope-labeling IS-based approaches with modern MS/MS technologies, mainly multiple reaction monitoring (MRM), permits high accuracy and precision in absolute protein quantification.

To perform AQUA, the most commonly used absolute protein quantification method, isotopically labeled synthetic peptides are added to digested protein samples as the IS, followed by peptide extraction and MS analysis. Protein quantity is determined according to the ratio of peak intensities of unlabeled natural peptides to their heavy isotope-labeled counterparts. AQUA is based on the assumptions that 1) all targeted proteins are fully recovered after protein extraction; 2) the selected reference peptides are fully digested from the proteins of interest and remain intact during enzymatic digestion.<sup>[9, 10]</sup> However, these assumptions have never been fully verified for most proteins/peptides.<sup>[6, 11]</sup> Additionally, synthesis and purification of isotope labeling peptides is often expensive.

Another commonly used absolute protein quantification method is the QconCATs.<sup>[4, 5]</sup>

QconCATs are artificial proteins made of a number of concatenated surrogate peptides for various analyte proteins. Isotope-labeled QconCATs are obtained via the expression of artificial QconCATs genes in cells cultured in a heavy isotope enriched medium.

QconCATs are added to protein samples before digestion, and the digested concatenated peptides serve as IS for different proteins. A major pitfall with this approach is that protein extraction and digestion efficiencies could vary significantly between QconCATs and native proteins due to different amino acid compositions.<sup>[7, 12]</sup>

Protein standards for absolute quantification (PSAQ) and absolute SILAC are an alternative approach to the AQUA and QconCATs. PSAQ and absolute SILAC involve biosynthesis and purification of stable isotope-labeled analyte proteins and the addition of a known quantity of labeled proteins to samples before extraction and digestion.<sup>[6, 7]</sup>

PSAQ and absolute SILAC can overcome some drawbacks of AQUA and QconCATs, such as the variations associated with protein extraction and digestion. However, PSAQ and absolute SILAC are generally low-throughput assays, and each isotope-labeled protein standard has to be generated and purified individually, which is very labor intensive and expensive.

In addition to the stable isotope labeling-based methods, so-called “label-free” absolute quantification methods have been developed using unlabeled proteins as external standard.<sup>[13, 14]</sup> However, this approach is generally considered unfavorable in terms of

accuracy and precision when compared with the methods employing isotopically labeled IS.

In view of the aforementioned limitations of current assays, in the present study, we developed a novel targeted absolute quantitative proteomics approach that employs SILAC IS and unlabeled full-length protein calibrators. This approach, named TAQSI, offers several advantages, such as lower cost and improved robustness and accuracy, over current existing methods. This method was successfully applied to a proof-of-concept study to quantify absolute expressions of carboxylesterase 1 (CES1) in human livers.

Author Manuscript

## EXPERIMENTAL

### Materials

Urea, dl-dithiothreitol (DTT), trifluoroacetic acid (TFA), and acetonitrile were purchased from Fisher Scientific Co. (Pittsburgh, PA). Iodoacetamide (IAA) was the product of Acros Organics (Morris Plains, NJ). TPCCK-treated trypsin was obtained from Worthington Biochemical Corporation (Freehold, NJ). Water Osis HLB columns were from Waters Corporation (Milford, MA). Recombinant CES1 (purity > 95%) was the product of R&D system (Minneapolis, MN). Dulbecco's Modified Eagle Medium (DMEM), fetal bovine serum (FBS), and 100 × antibiotics mixture containing penicillin (100 IU/ml) and streptomycin (100 µg/ml) were the products of Invitrogen (Carlsbad, CA). Trypsin-EDTA (0.25%) and SILAC Protein Quantitation Kit-DMEM containing SILAC DMEM (deficient in arginine and lysine),  $^{13}\text{C}_6$  l-lysine-2HCl,  $^{13}\text{C}_6$   $^{15}\text{N}_4$  l-arginine-HCl, and dialyzed FBS were obtained from Thermo Scientific (Waltham, MA). Sterile syringe filters with 0.2µm sterile cellulose acetate membranes were purchased from VWR international Inc. (Bridgeport, NJ). HepG2 cells were purchased from ATCC (HB-8065<sup>TM</sup>, Manassas, VA). A total of 102 individual normal human liver samples were randomly selected from 535 biobanked samples. These liver samples were obtained from several sources including the University of Minnesota Liver Tissue Cell Distribution System, Cooperative Human Tissue Network (CHTN), and XenoTech LLC (Lenexa, KS). The donors consisted of 92 Caucasians, 6 African-Americans, 2 Hispanics, and 2

classified as others, with 46 males and 56 females. The donor ages ranged from 22 to 81 years old.

### **Cell culture and SILAC labeling**

Human hepatocellular carcinoma cell line HepG2, which exhibits similar gene expression pattern of human livers, was utilized to generate stable isotope labeling proteins. HepG2 cells were initially cultured in DMEM containing 10% FBS, 100 IU/ml penicillin and 100 µg/ml streptomycin at 37 °C under 5% CO<sub>2</sub> and 95% humidity. The SILAC culture medium consisted of SILAC DMEM supplemented with 0.1 mg/ml of <sup>13</sup>C<sub>6</sub> l-lysine-2HCl and 0.1 mg/ml <sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>4</sub> l-arginine-HCl, 10% dialyzed FBS, 100 IU/ml of penicillin and 100 µg/ml streptomycin. To obtain isotope-labeled cell culture, HepG2 cells were cultured in SILAC DMEM, and the medium was replaced every 3 days. Cells were subcultured when reaching 90% confluency. MS analysis showed that the incorporation rate of isotope-labeled arginine and lysine was more than 99% in the cell S9 fractions after cells were cultured in SILAC medium for 5 generations or more. Thus, only the HepG2 cells with SILAC culture ≥ 5 passages were utilized in the study. To avoid potential variability in protein expression between different batches of cell culture, all SILAC S9 fractions were pooled, and subsequently used throughout the entire experiment.

### **Samples preparation**



Individual human liver S9 fractions (HLS9) were prepared from about 200 mg frozen liver tissues. The liver tissues were cut into small pieces (1×1×1 mm) and homogenized in 0.5ml ice-cold phosphate buffer saline (PBS) in 1.5 ml microcentrifuge tubes using a microcentrifuge pestle (VWR International LLC, Chicago, IL). The homogenates were centrifuged at 9000 ×g at 4 °C for 20 min. Following the centrifugation, the top layer containing fats was carefully removed, and the remaining samples were centrifuged again at 9000 ×g at 4 °C for 20 min to remove the remaining fats. The resulting supernatants (S9 fractions) were collected and diluted to 2 mg/ml in PBS. HepG2 cell S9 fractions were prepared based on our previously published method.<sup>[15]</sup> Protein concentrations of human liver and HepG2 S9 fractions were determined using Pierce BCA protein quantification assay. An aliquot of 20 µl of HLS9 from each of the 102 samples prepared were mixed to make the pooled HLS9 sample. All samples were stored at -80 °C until use.

Protein aliquots of 20 µg HLS9 and 40 µg SILAC HepG2 cell S9 fractions (IS) were mixed in Eppendorf Protein LoBind tubes, followed by addition of 10-fold volume of acetonitrile. The mixtures were briefly vortexed and then centrifuged at 17000 ×g for 10 min at 4°C. The supernatants were discarded, and the precipitated proteins were air dried for 5 min at room temperature. The dried proteins were re-suspended in 100 µl of freshly prepared 4 mM DTT/8 M urea solution, and incubated at 37 °C for 45 min. Following the incubation, samples were cooled to room temperature, and then alkylated by incubation

with 100  $\mu$ l of freshly prepared 20 mM IAA solution at room temperature for 30 min in dark. After alkylation, 800  $\mu$ l water was added to dilute urea concentration to 0.8 M. Samples were then digested by trypsin at an enzyme/protein ratio of 1:500 in an incubation shaker at 200 rpm at 37°C for 16 h. The digestion was terminated by the addition of 1  $\mu$ l TFA. Digested peptides were extracted using Waters Oasis HLB columns according to the manufacture's instructions. Extracted peptides were dried in a Speed Vac SPD1010 (Thermo Scientific, Hudson, NH), and reconstituted in 80  $\mu$ l of 50% acetonitrile. The peptide solutions were centrifuged at 17,000  $\times$ g for 10 min at 4°C, and the supernatant was collected for LC-MS/MS analysis.

Unlabeled purified recombinant full-length CES1 protein was obtained from R&D system (Minneapolis, MN), and was utilized as a calibrator to prepare the calibration solutions for the quantification of CES1 expression in human livers. The CES1 calibrators were 0.59, 1.18, 2.36, 4.73 and 11.82 pmol, which cover normal range of CES1 expressions in 20  $\mu$ g proteins of human livers. Quality control (QC) samples were also prepared from the recombinant CES1 with absolute quantities of 1.18, 2.36 and 4.73 pmol. Additionally, a pooled HLS9 sample (12  $\mu$ g protein) was included in each run to further evaluate the between-run variability. Calibrators and QC samples were processed in parallel with liver samples using the aforementioned sample preparation protocol. Figure 1 is the illustration of the workflow of the TAQSI assay and its comparison to

other absolute protein quantification methods including AQUA, QconCATs, PSAQ and absolute SILAC.

To evaluate the impact of digestion completion on protein quantification, we quantified CES1 after 1, 2, 4, 8, 12, 16, 24 h trypsin digestion using the same sample preparation method described above.

### **LC-MS/MS analysis**

The workflow of LC-MS/MS analysis of absolute CES1 expression was established with the assistance of the Skyline software (University of Washington, Seattle, WA). A list of 26 candidate peptides were generated based on the MS/MS spectrum currently available in NIST and human ISB Plasma spectral libraries (**Supplemental Table 1**). The peptides containing any nonsynonymous variants with a minor allele frequency  $> 0.1\%$  were excluded, resulting in four peptides being removed from the original list. A table containing the top three most intensive precursor-product ion transitions of each of the remaining 22 peptides and the corresponding MS/MS instrumental parameters (e.g. dwell time, ionspray voltage, declustering potential, collision energy) was generated from Skyline, and was integrated into the LC-MS/MS method. The method was applied to the analysis of CES1 quantities in three pooled HLS9 samples. A total of 6 unique CES1 peptides that exhibited the highest intensity of MS/MS spectra were targeted in the final method (**Supplemental Table 2**).

The LC-MS/MS analysis was performed on a triple quadrupole mass spectrometer (API 4000 QTRAP<sup>®</sup>, AB SCIEX, Concord, ON) coupled with a Shimadzu Prominence HPLC system (Shimadzu Corporation, Kyoto Japan). Analytes were separated on an Agilent ZORBAX 300SB-C18 column (5  $\mu$ M, 150 $\times$ 2.1mm, Agilent Technologies, Santa Clara, CA) set at 40°C. The mobile phase consisted of acetonitrile (phase A) and water (phase B), both containing 0.1% formic acid (v/v), and was delivered at a flow rate of 0.2 ml/min. A gradient elution was applied for the separation with the following time program: phase A was set at 5% from 0 to 1 min and increased to 37% from 1 to 20 min, and further increased to 100% from 20 to 25 min, and maintained 100% from 25 to 35 min, then returned to 5% from 35 to 36 min, and maintained at 5% to the end of elution (40 min). The MS was operated in positive electrospray using a Turbolon Spray ion source.

Digested bovine serum albumin (BSA) was utilized to generate the retention time predictor for the estimation of retention times of CES1 peptides. Standard curves were established based on the peak area ratios of the light (unlabeled) peptides from the recombinant CES1 to the heavy (isotope-labeled) peptides from the SILAC HepG2 S9 fractions. Assay accuracy and precision were assessed by analyzing the four QC samples in each run.

## RESULTS

### Consistency of protein quantification among the six reference peptides

We determined absolute CES1 quantities in the three QCs and pooled HLS9 samples using the six selected CES1 signature peptides. No significant interference peaks were observed for any of the peptides (**Figure 2**). The quantifications were consistent across the six peptides with the relative standard deviation (RSD) of 2.0%, 2.5%, 2.0%, and 2.3% for 1.18, 2.36, 4.73 pmol QCs, and pooled human liver S9, respectively (**Figure 3**). The results suggest that the variability of CES1 quantification among the different signature peptides is very small, and any of the six surrogate peptides could be used to reliably quantify CES1 expression in human livers.

### Linearity, accuracy, and precision

The standard calibration curves, established by plotting peptide quantity versus the peak area ratio of unlabeled-to-isotope labeled peptides, were found to be linear within the tested range (0.59 – 11.82 pmol) for all six peptides (**Supplemental Table 3**). This range was anticipated to bracket possible CES1 quantities in 20 µg protein of HLS9 samples. The correlation coefficients of all six peptides were greater than 0.9999 under our experimental conditions (**Supplemental Table 3**). The precision and accuracy of the assay were determined utilizing the three QC samples (1.18, 2.36, 4.73 pmol). As shown in **Table 1**, the inter- and intra-day precision measured as RSD were equal or less than

9.0% and 5.5%, respectively. The inter- and intra-day accuracy were within the ranges of 96.1%-102.2% and 94.6%-106.1%, respectively. Pooled HLS9 samples with 20 µg of total proteins were included in each run to serve as an additional QC. The between-run RSDs of the pooled HLS9 sample were between 0.7% and 4.4% for the six selected peptides.

### **Digestion time**

To evaluate the effect of digestion completion on CES1 protein quantitation, a mixture containing 20 µg protein of pooled HLS9 samples and 40 µg protein of SILAC HepG2 cell S9 fractions (IS) was digested by trypsin at an enzyme/protein ratio of 1:500 for 1, 2, 4, 8, 12, 16, and 24 h. The intensity of MS/MS spectra of both the six unlabeled surrogate CES1 peptides and their isotope-labeled counterparts were increased from 1 to 8 h following digestion, and reached a plateau after 8 h. However, the ratios of the light to heavy peptides that we used for CES1 quantitation remained consistent during the whole period time of digestion, indicating that incomplete digestion should not affect protein quantitation results of the TAQSI method (**Figure 4**).

### **LC-MS/MS quantification of CES1 protein in human liver samples**

CES1 expressions and activities vary significantly among individual human livers. We applied the newly developed TAQSI assay to quantifying absolute CES1 expression in 102 individual human liver samples. CES1 expression was found to vary significantly

among those liver samples, ranging from  $42.0 \pm 2.3$  to  $477.9 \pm 21.9$  pmol/mg protein. The mean expression level was  $176.08 \pm 75.6$  pmol/mg protein ( $10.0 \pm 4.2$   $\mu\text{g}/\text{mg}$  protein), which is approximately 1% of total proteins expressed in human liver. To evaluate the reproducibility of the assay, a total of 24 samples were analyzed for CES1 expression by two independent experiments. The RSD was found to range from 0.1% to 3.8% between the two runs among those samples, indicating that the method is highly reproducible (**Figure 5**).

Author Manuscript

## DISCUSSION

In the present study, we established a novel LC-MS/MS approach for absolute quantification of targeted proteins with very high accuracy and precision. Compared with current absolute protein quantification assays, our method is advantageous in the following three major aspects.

### **Reducing inter-assay variability and improving accuracy and precision**

AQUA is the most commonly used method for absolute protein quantification, and has been adopted in many basic and clinical proteomics studies.<sup>[9, 16, 17]</sup> To perform AQUA, isotope-labeled synthetic peptide(s) are added to samples as IS after protein digestion. As a result, the isotope-labeled peptide IS(s) are not able to correct for systematic and accidental errors that may have occurred during protein extraction and enzymatic digestion. It is also noted that proteotypic peptide yields are often protein/peptide dependent, which makes it very difficult to establish a digestion protocol that is ideal for all targeted proteins. As shown in the **Figure 4**, different digestion protocols can significantly affect the yields of targeted peptides. Furthermore, the stability of peptides during digestion could introduce additional variability to the recovery of peptides. Consequently, protein quantification results could differ significantly when different isotope-labeled IS peptides or extraction and digestion protocols are utilized. For



instance, the use of different isotope-labeled IS peptides resulted in more than 7-fold differences of UGT1A4 quantification in the same human liver samples.<sup>[18]</sup>

Distinct from AQUA, in which synthetic isotope-labeled peptides IS are used, our TAQSI approach employs full-length isotope-labeled native proteins obtained from SILAC as the IS. The full-length labeled proteins are added to the lysate of tissues or cells before protein extraction and digestion. Therefore, the variabilities introduced from protein extraction and digestion as well as other downstream procedures can be largely eliminated. Our study demonstrated that, even though peptide recoveries varied depending on the time of digestion, the unlabeled-to-labeled ratios of all six selected CES1 peptides remained consistent throughout the digestion process (**Figure 4**). Thus, the TAQSI method is robust in terms of being resistant to the inter-assay variability caused by imperfect protein extraction and/or digestion protocols and the selection of different surrogate peptides.

Similar to our TAQSI approach, PSAQ and absolute SILAC assays can significantly reduce the variability associated with protein extraction and digestion.<sup>[6, 7, 9]</sup> However, each PSAQ and absolute SILAC isotope standards need to be individually generated from cell free systems or SILAC, and further purification and quantification are required before the proteins can be utilized as the standards for protein quantification. The whole procedure of PSAQ and absolute SILAC standard preparation is often very expensive and time consuming. Full-Length Expressed Stable Isotope-labeled Proteins for

Quantification (FLEXIQuant) is a further development of PSAQ assay,<sup>[19]</sup> in which isotope-labeled protein standards tagged with a FLEX-peptide are generated from cell culture. Relative to PSAQ, the major advantage of FLEXIQuant is that the FLEX-peptide allows for absolute protein quantification without the need for the purification of the isotope-labeled FLEXIQuant IS. However, each FLEXIQuant standard has to be generated individually through customized expression constructs, which hinders the applicability of this assay in simultaneous quantification of multiple proteins. With comparison to PSAQ, absolute SILAC and FLEXIQuant, our TAQSI method utilizes unpurified multiple isotope-labeled IS proteins simultaneously generated from SILAC. Additionally, unlabeled full-length proteins, which serve as the external standards for the construction of calibration curves, are often commercially available, or can be readily prepared in house through standard molecular biological assays. Thus, the TAQSI method appears to be an efficient and cost effective alternative approach of the established PSAQ and FLEXIQuant assays.

### **Reducing measurement errors caused by nonsynonymous variants**

Measurement errors in protein quantification can occur when samples contain nonsynonymous variants that reside in the selected signature peptide(s). As a general principle, peptides containing high-frequency nonsynonymous variants should be avoided when selecting surrogate peptides for protein quantification. However, availability of suitable peptides for protein quantification is often limited due to that the selected

peptides also need to satisfy other criteria, such as specificity, MS/MS sensitivity, and digestion efficiency, etc. It is not uncommon that peptides with reported nonsynonymous variants are adopted as surrogate peptides for protein quantification. The TAQSI approach allows to apply multiple signature peptides to protein quantification without increasing overall costs and assay complexity. The data from our proof-of-concept study of absolute CES1 quantification demonstrated very consistent quantitative results across the different signature peptides (**Figure 3**). Missing peaks of an unlabeled peptide would indicate a homozygous nonsynonymous variant in the peptide whereas it would suggest the presence of a heterozygous nonsynonymous variant if the calculated protein quantity based on a selected peptide is approximately 50% of that determined from other reference peptides (**Supplemental Figure 1**). Thus, cross examination of the data derived from different peptides can not only enhance the assurance and reliability of quantification data but also allow to identify the peptide(s) containing nonsynonymous variants and exclude it from the final data analysis.

### **Simplifying the selection of optimal performance proteotypic peptides**

Ideal reference peptides for protein quantification should meet several criteria, including 1) be unique to targeted proteins; 2) generate good MS/MS responses; 3) be chemically stable; 4) be unlikely affected by genetic variants. Some bioinformatics tools have been developed to aid in signature peptides selection.<sup>[20, 21]</sup> However, it remains a challenging task to predict the best performance peptides as MS/MS profiles of a peptide can be

affected by many experimental conditions, such as MS instrumentation and sample preparation methods. AQUA assays usually do not test all candidate peptides due to the high costs of obtaining synthetic isotope-labeled peptides. Therefore, the peptide(s) selected from limited candidate peptides might not be the best performance peptides for AQUA protein quantification. Our TAQSI method employs full-length unlabeled and labeled native proteins, which allows us to readily explore the performance of all digested peptides unbiasedly during method development. In the present proof-of-concept study, we evaluated a total of 22 CES1 candidate peptides and identified six peptides with the best performance in terms of MS/MS responses and chromatographic selectivity (**Figure 2**).

### **Limitations**

Several limitations have been recognized for the TAQSI assay. First of all, purified unlabeled whole length proteins, utilized as the calibrators to establish standard curve for absolute quantification, are associated with the added costs and efforts related to the purchase or preparation of those protein calibrators. Secondly, given that this method relies on SILAC internal standard, potential applications of this approach may be limited by the availability of SILAC culture medium. Although commercial SILAC kits currently available already contain those most commonly used medium such as DMEM and RPMI 1640, it could be difficult to obtain some less commonly used medium in SILAC form. Finally, the use of whole protein lysates of SILAC cells as the internal standard could

increase the complexity of protein samples and may enhance the background noise, which could pose a challenge to quantify the proteins with low abundance.

## **CONCLUSIONS**

We have developed a novel LC-MS/MS-based MRM absolute targeted protein quantification method, named TAQSI, which is accurate, precise, reproducible, and cost effective relative to other existing methods. This approach has the potential to be widely used in the study of protein expression in various biomedical research settings and absolute quantification of protein therapeutics in pharmaceutical industry. Additionally, the reproducibility of the TAQSI makes it a preferred tool for the discovery and validation of clinical protein biomarkers for quantitative proteomics-based diagnostics and precision medicine.

Author Manuscript

## **Acknowledgment**

Research reported in this publication was supported in part by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number 2UL1TR000433 (Hao-Jie Zhu), the National Institute on Aging (R21AG048500) (Hao-Jie Zhu), the American Association of Colleges of Pharmacy (AACP) 2015 New Investigator Award (Hao-Jie Zhu), and the National Nature Science Foundation of China (81273589) (Guangji Wang). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Manuscript

**Table 1.** Intra-day and inter-day assay precision and accuracy of three QC samples of the six unique peptides for CES1 quantification.

Nominal CES1 quantities (pmol)	Intra-day (n=4)		Inter-day (n=4)	
	Accuracy (%)	Precision (%)	Accuracy (%)	Precision (%)
<b>AISESGVALTSVLVK</b>				
1.18	101.9	2.4	96.1	5.5
2.36	95.1	3.9	98.9	3.3
4.73	99.2	4.5	101.4	3.1
<b>FLSLDLQG DPR</b>				
1.18	100.5	2.8	97.8	1.1
2.36	96.8	3.2	98.2	4.3
4.73	94.7	2.6	98.2	2.3
<b>TAMSLWK</b>				
1.18	106.1	2.4	100.0	9.0
2.36	104.2	4.1	100.2	4.1
4.73	102.3	4.0	102.2	4.4
<b>SYPLVC[CAM]IAK*</b>				
1.18	97.6	4.1	98.1	5.7
2.36	94.6	1.8	96.8	5.0
4.73	99.1	4.8	101.5	0.8
<b>ELIPEATEK</b>				
1.18	95.8	4.3	98.7	6.0
2.36	98.4	4.8	101.2	4.7
4.73	99.6	3.3	102.2	3.2
<b>FWANFAR</b>				
1.18	99.4	5.5	99.7	3.2
2.36	100.4	2.7	100.8	3.5
4.73	100.6	2.3	97.6	3.3

\*[CAM]: Cysteine was alkylated with iodoacetamide during sample preparation.

## Figure Legends

**Figure 1.** The workflow of the TAQSI assay and its comparison to other absolute protein quantification methods including AQUA, QconCATs, PSAQ and absolute SILAC. The TAQSI method employs a dilution series of unlabeled target proteins as calibrators and SILAC preparations containing target proteins as the internal standards (I.S.) to generate calibration curves for absolute protein quantification. Equal amount of SILAC I.S. is spiked in calibrators and unknown samples. The mixers are then processed according to the procedures detailed in the Methods section. Following LC-MS/MS analysis, absolute amounts of target proteins are determined by comparing the ratios of unlabeled-to-labeled signature peptides to the established calibration curves. The major differences between the PSAQ/Absolute SILAC and the TAQSI assays are that the PSAQ/Absolute SILAC methods require purified isotope-labeled I.S., and protein concentrations are calculated directly from the ratios of unlabeled-to-labeled peptides without using calibration curves. Unlike the PSAQ/Absolute SILAC and TAQSI assays in which full-length isotope-labeled proteins I.S. are used, the QconCATs assay utilizes artificial proteins made of a number of concatenated surrogate peptides as the I.S.. For AQUA, the I.S. are isotopically labeled synthetic individual peptides. It should be noted that the I.S. of the PSAQ/Absolute SILAC, the TAQSI, and the QconCATs assays are added before sample processing while the AQUA I.S. are spiked after protein digestion.

**Figure 2.** Representative LC-MS/MS chromatograms of the six selected unique CES1



peptides (A: AISESGVALTSVLVK, B: FLSLDLQGDPR, C: TAMSLW, D: SYPLVC[CAM]IAK, E: ELIPEATEK, F: FWANFAR) and their corresponding isotope-labeled peptides. The data were obtained from the analysis of the mixture of pooled HLS9 samples and SILAC HepG2 S9 fractions.

**Figure 3.** Absolute CES1 quantities in three QCs and pooled HLS9 samples determined using the six selected CES1 signature peptides. Nominal CES1 quantities in the low, medium, and high QC samples were 1.18, 2.36, 4.73 pmol, respectively. The pooled HLS9 sample contained 20 µg of total proteins.

**Figure 4.** The impact of digestion time on the MS/MS intensities of the six unlabeled and isotope-labeled CES1 signature peptides as well as their ratios. The MS/MS peak intensities of both labeled and unlabeled peptides increased along with the time of digestion, and reached plateau after 8h tryptic digestion. However, the ratios of unlabeled-to-labeled peptides remained consistent throughout the period of digestion. Data are mean ± S.D. from three independent experiments.

**Figure 5.** Absolute CES1 quantities (pmol/mg total protein) in 24 HLS9 samples. CES1 expression was determined based on the averages of the CES1 quantities calculated from the six surrogate CES1 peptides. Data are presented as means from two independent studies with error bars representing S.D..

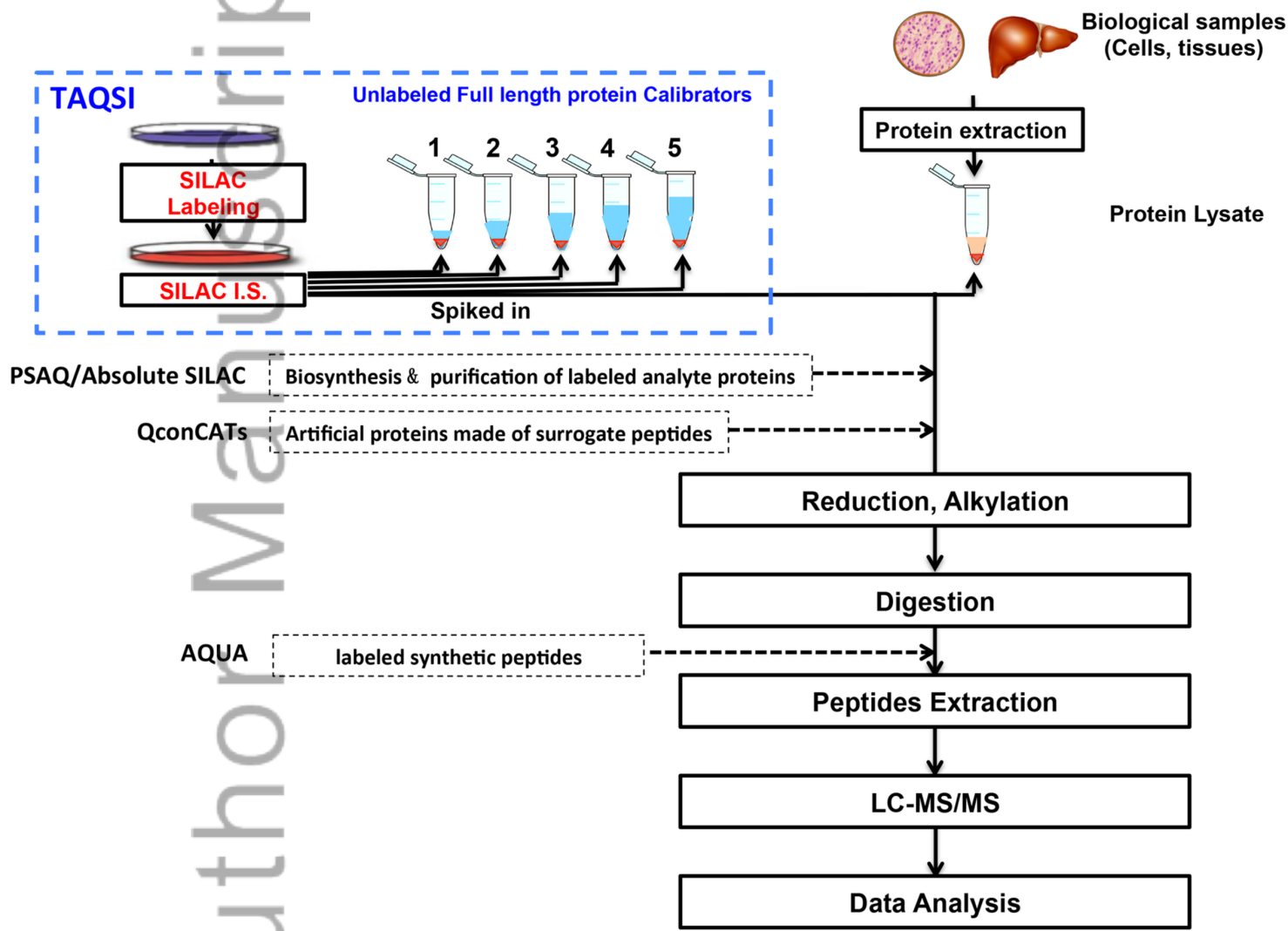
## References

- [1] S. A. Gerber, J. Rush, O. Stemman, M. W. Kirschner, S. P. Gygi, *Proc Natl Acad Sci U S A* **2003**, *100*, 6940.
- [2] H. Keshishian, T. Addona, M. Burgess, D. R. Mani, X. Shi, E. Kuhn, M. S. Sabatine, R. E. Gerszten, S. A. Carr, *Mol Cell Proteomics* **2009**, *8*, 2339.
- [3] M. A. Kuzyk, D. Smith, J. Yang, T. J. Cross, A. M. Jackson, D. B. Hardie, N. L. Anderson, C. H. Borchers, *Mol Cell Proteomics* **2009**, *8*, 1860.
- [4] R. J. Beynon, M. K. Doherty, J. M. Pratt, S. J. Gaskell, *Nat Methods* **2005**, *2*, 587.
- [5] J. M. Pratt, D. M. Simpson, M. K. Doherty, J. Rivers, S. J. Gaskell, R. J. Beynon, *Nat Protoc* **2006**, *1*, 1029.
- [6] S. Hanke, H. Besir, D. Oesterhelt, M. Mann, *J Proteome Res* **2008**, *7*, 1118.
- [7] V. Brun, A. Dupuis, A. Adrait, M. Marcellin, D. Thomas, M. Court, F. Vandenesch, J. Garin, *Mol Cell Proteomics* **2007**, *6*, 2139.
- [8] M. Zeiler, W. L. Straube, E. Lundberg, M. Uhlen, M. Mann, *Mol Cell Proteomics* **2012**, *11*, O111 009613.
- [9] J. Villanueva, M. Carrascal, J. Abian, *J Proteomics* **2014**, *96*, 184.
- [10] H. Al Feteisi, B. Achour, J. Barber, A. Rostami-Hodjegan, *AAPS J* **2015**, *17*, 438.
- [11] E. Kuhn, J. Wu, J. Karl, H. Liao, W. Zolg, B. Guild, *Proteomics* **2004**, *4*, 1175.
- [12] J. Rivers, D. M. Simpson, D. H. Robertson, S. J. Gaskell, R. J. Beynon, *Mol Cell Proteomics* **2007**, *6*, 1416.

- [13] B. M. Mayr, O. Kohlbacher, K. Reinert, M. Sturm, C. Gropl, E. Lange, C. Klein, C. G. Huber, *J Proteome Res* **2006**, *5*, 414.
- [14] L. Arike, K. Valgepea, L. Peil, R. Nahku, K. Adamberg, R. Vilu, *J Proteomics* **2012**, *75*, 5437.
- [15] H. J. Zhu, K. S. Patrick, H. J. Yuan, J. S. Wang, J. L. Donovan, C. L. DeVane, R. Malcolm, J. A. Johnson, G. L. Youngblood, D. H. Sweet, T. Y. Langaee, J. S. Markowitz, *Am J Hum Genet* **2008**, *82*, 1241.
- [16] X. Wang, G. Wang, J. Shi, J. Aa, R. Comas, Y. Liang, H. J. Zhu, *Pharmacogenomics J* **2015**.
- [17] O. Chahrour, D. Cobice, J. Malone, *J Pharm Biomed Anal* **2015**.
- [18] V. Brun, C. Masselon, J. Garin, A. Dupuis, *J Proteomics* **2009**, *72*, 740.
- [19] J. K. Fallon, H. Neubert, R. Hyland, T. C. Goosen, P. C. Smith, *J Proteome Res* **2013**, *12*, 4402.
- [20] S. Singh, M. Springer, J. Steen, M. W. Kirschner, H. Steen, *J Proteome Res* **2009**, *8*, 2201.
- [21] P. Mallick, M. Schirle, S. S. Chen, M. R. Flory, H. Lee, D. Martin, J. Ranish, B. Raught, R. Schmitt, T. Werner, B. Kuster, R. Aebersold, *Nat Biotechnol* **2007**, *25*, 125.
- [22] G. Alves, A. Y. Ogurtsov, Y. K. Yu, *J Proteomics* **2011**, *74*, 199.

# Author Manuscript

Author Manuscript

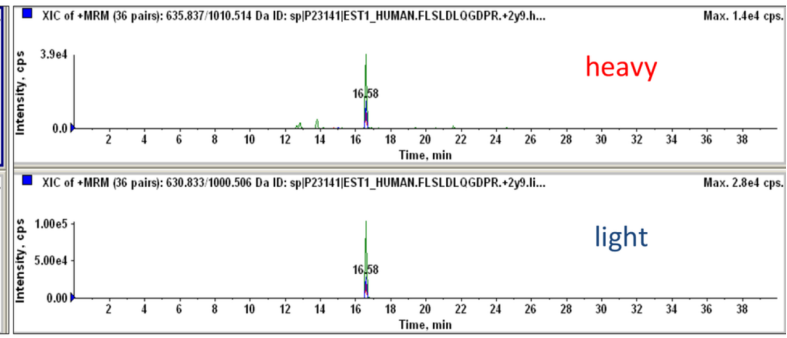
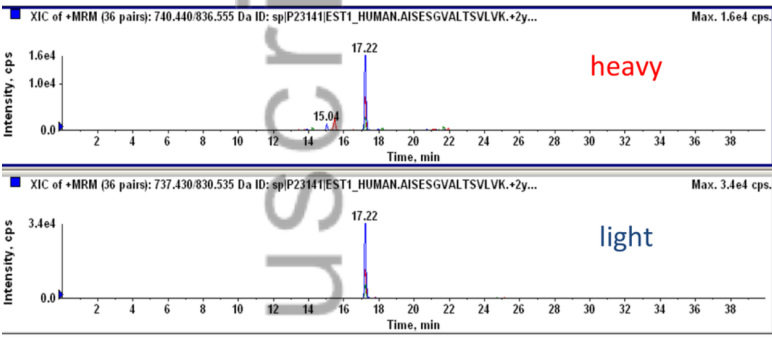


RCM\_7482\_F1.tif

Figure 2

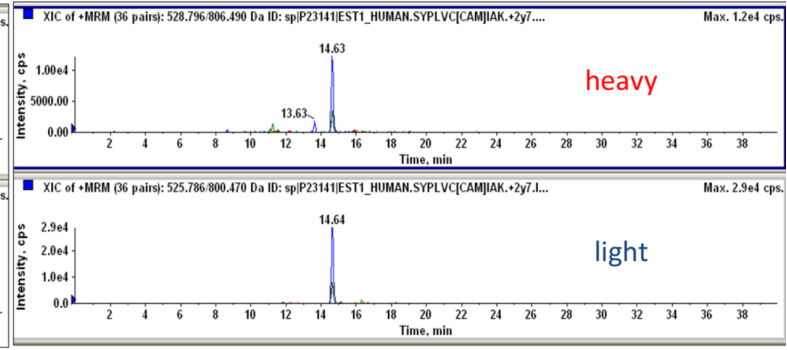
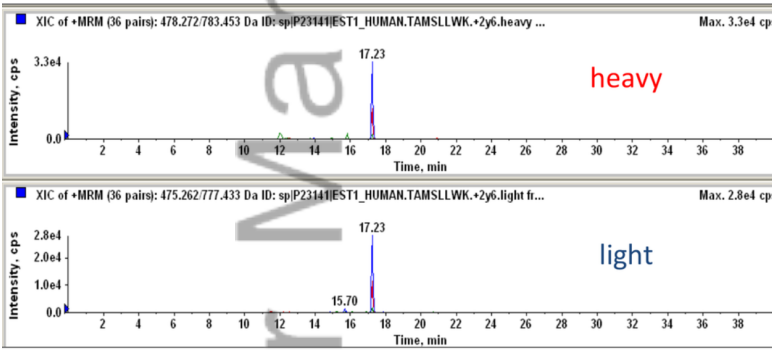
**A: AISESGVALTSVLVK**

**B: FLSLDLQGDPR**



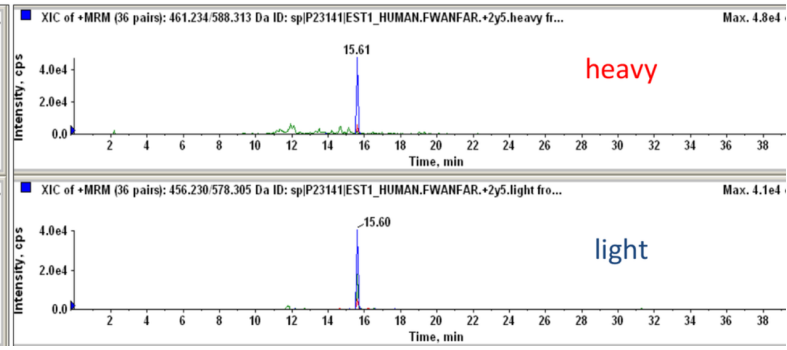
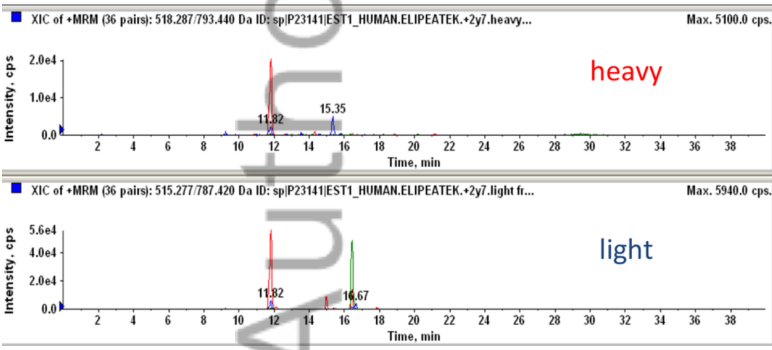
**C: TAMSLLWK**

**D: SYPLVC[CAM]IAK**



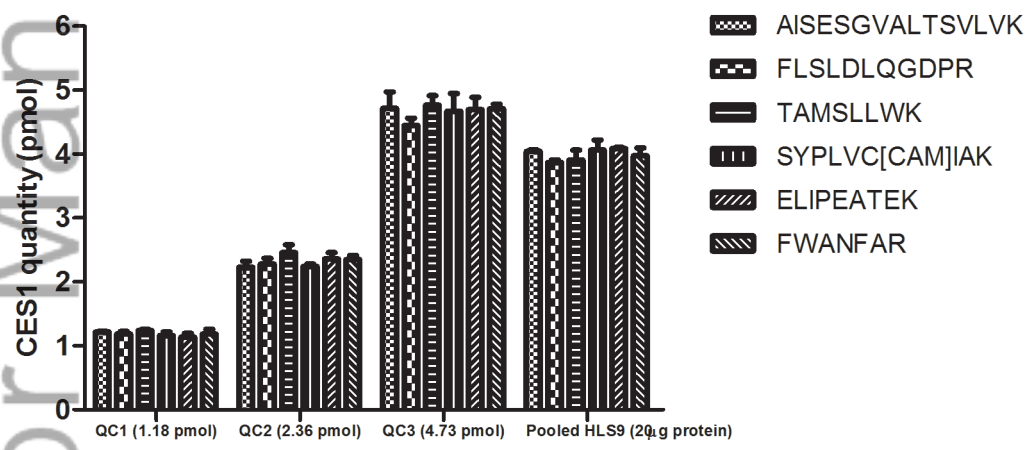
**E: ELIPEATEK**

**F: FWANFAR**



RCM\_7482\_F2.tif

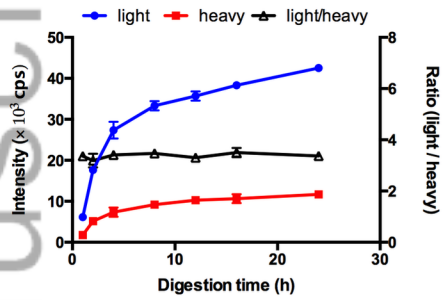
Figure 3



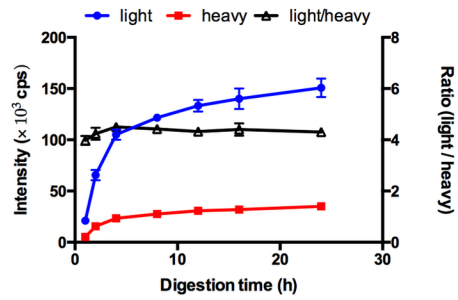
RCM\_7482\_F3.tif

**Figure 4**

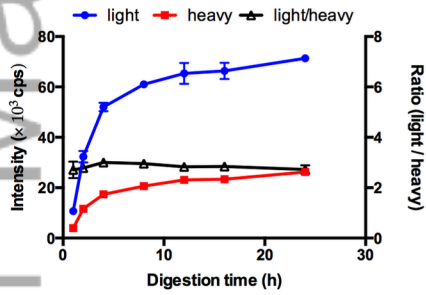
**A: AISESGVALTSVLVK**



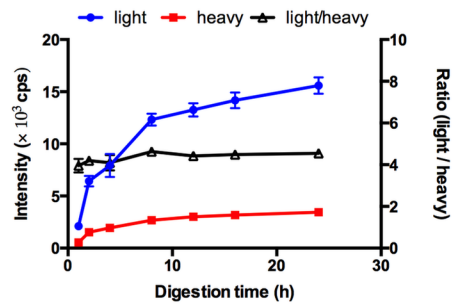
**B: FLSLDLQGDPR**



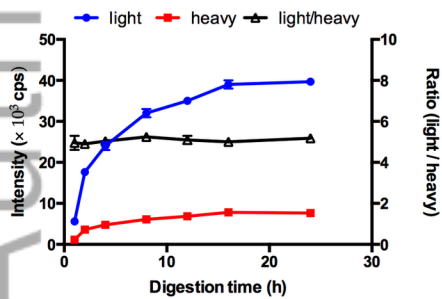
**C: TAMSLWK**



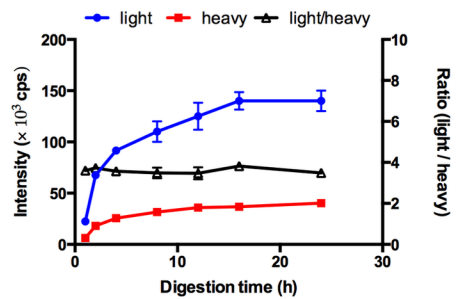
**D: SYPLVC[CAM]IAK**



**E: ELIPEATEK**



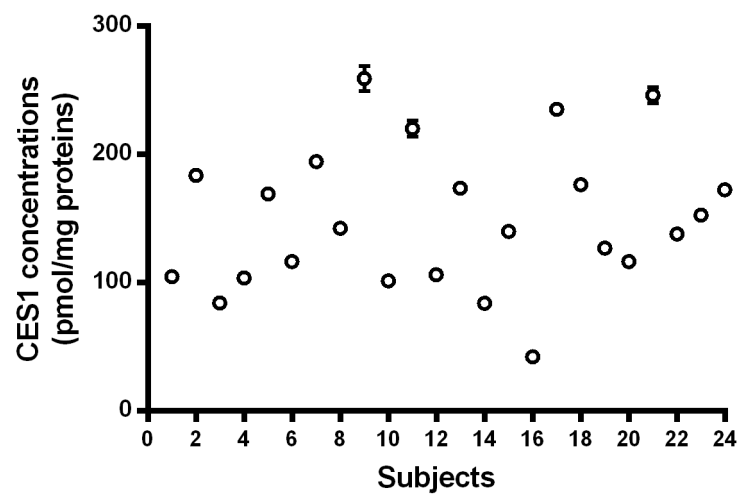
**F: FWANFAR**



RCM\_7482\_F4.tif



Figure 5



RCM\_7482\_F5.tif