November 6, 2014

Dear Colleagues,

Please find enclosed the ABRF 2015 PSRG samples that you requested from the ABRF Protein Sequence Research Group. This is the 27th study in an annual series designed to assist laboratories in evaluating their abilities to obtain and interpret protein sequencing data. Thank you for your interest and participation!

N-terminal sequencing is in the midst of a technology transition from classical Edman sequencing to mass spectrometry-based terminal identification. For core laboratories, the ultimate goal in the coming years will be to have a well-defined protocol for terminal sequence analysis by mass spectrometry that has the same level of maturity as Edman sequencing. Knowing the types of samples, sample preparation protocols and expected results are critical for core laboratories so that they can provide the most reliable data to their customers.

To help in development and establishment of such techniques, the PSRG is conducting a study using chemical derivitization to identify the N-terminal peptide of a standard protein by mass spectrometry. The current 2015 study entails **dimethyl labeling of purified myoglobin. The goal for this year is to test the abilities of participating laboratories to a) successfully derivitize 5pmol, 1pmol, and 300fmol of myoglobin, b) digest and identify the derivitized peptide fragments by mass spectrometry, and c) obtain terminal peptide sequence information**.

Study restrictions: Dimethyl labeling of protein N-termini is fast and inexpensive, but requires some reagents that the PSRG is not able to ship to participants. Formaldehyde and sodium cyanoborohydride must be purchased by participants and can be obtained by vendors such as Sigma Aldrich for less than $50 each. This study is not designed for Edman sequence analysis; participants must use either top-down or bottom-up mass spectrometry. Analysis must include the use of bioinformatics tools to derive terminal sequences.

Participating laboratories will receive an aliquot of purified myoglobin and protocols for dimethyl labeling: **(A)** in-solution labeling with SDS-PAGE or cut-off filter cleanup, or **(B)** SDS-PAGE separation and in-gel labeling. Participants may choose which workflow to perform and will follow the sample preparation with trypsin digestion and mass spectrometry using the MS system and operating parameters of their choice.

(A) Workflow "in-solution labeling":

1) Reductive amidation (dimethyl labeling) of protein N-terminus
2) SDS-PAGE or on-membrane cut-off filter trypsin digestion
3) MS analysis
4) Data analysis for identification of labeled N-terminal amino acid

(B) Workflow "in-gel labeling and digest"

1) SDS-PAGE, excision of bands and reductive amidation (dimethyl labeling) of protein in-gel
2) In-gel trypsin digestion
3) MS analysis
4) Data analysis for identification of labeled N-terminal amino acid

Materials supplied in this study:

Protein	Catalog Number	Quantity
Myoglobin	Sigma M1882	100 ug

Reagents required to be purchased by the participants:

Chemical	Catalog Number	Quantity
Formaldehyde	Sigma F8775	25 mL
Sodium cyanoborohydride	Sigma 156159	10 g

Data submission:
Please provide data and analysis results directly to Amanda McGinnis at apmcg@umich.edu. You will receive an identification number via email from Ms. McGinnis, who will act as the anonymizer and is the only person who will know the relation between your email and your identification number. In order to ensure anonymity, Ms. McGinnis will remove all identifying marks prior to forwarding your data to the PSRG committee for analysis. The sequencing results will be presented at the ABRF 2015 meeting March 28-31, 2015, in St. Louis, MO and subsequently posted on the ABRF website, and will also help guide future studies and tutorial sessions.

Please include in your email to Ms. McGinnis Power Point files with mass spectra containing the digested labeled fragments and report the N-terminal sequence as determined by bottom-up (or top-down) mass spectrometry. You may also include relevant spectra, SDS-PAGE images, database search results, or chromatographic data supporting your analysis and sequence calls. **Please include your laboratory identification number and Workflow type in the title of the supporting document.** If alternative methods are used, PSRG requests a brief description of the protocol used and literature references as appropriate. Recommendations for successful analytical and bioinformatics methods are included in the Appendix.

In addition to submission of data, the PSRG is requesting that each participant complete a survey about the study on Survey Monkey. Your participation will be recorded using the identification number assigned to you by Ms. McGinnis.

The results survey can be accessed here:
https://www.surveymonkey.com/s/X9NC8GC

If you are unable to open the link, email Amanda McGinnis at apmcg@umich.edu and she will fax you a paper copy. If your sample arrived damaged, or if you have questions about the study, please contact co-chair Sara McGrath (sara.mcgrath@fda.hhs.gov).

***** Equipment failures and "no data obtained" analyses are as important to us as data from "successful" runs. Please send us your results and fill out the online survey, regardless of your success.**

*****The deadline for receiving data for inclusion in the study is February 6, 2015.**

Thank you for your valuable participation in this year's study!

The Protein Sequencing Research Group:

Robert English	–	Shimadzu Scientific Instruments (co-chair)
Sara McGrath	–	FDA Center for Food Safety and Applied Nutrition (co-chair)
Greg Cavey	–	Launch MI Lab, Southwest Michigan Innovation Center
Hediye Erdjument-Bromage	–	Memorial Sloan-Kettering Cancer Center
Xuemei Luo	–	University of Texas Medical Branch
Henriette Remmer	–	University of Michigan (*ad-hoc*)
Brett Phinney	–	UC Davis Genome Center (EB liaison)

Appendix

Sample Preparation

1. Record date sample is received and store sample in -20°C freezer.
2. Sample contains 100 ug dried protein
3. Solubilize the myoglobin with 588 uL H2O to make a 10 pmol/uL solution
4. Dilute the solution to make a 1 pmol/uL solution by adding 10 uL of myoglobin solution to 90 uL of 100mM sodium acetate, pH 5

Workflow "A" N-Terminal Labeling in Solution

1. Label the desired amount of each protein in separate tubes according to Table 1

   Note: Formaldehyde and sodium cyanoborohyride are toxic and must be handled with appropriate safety precautions.

Table 1: Dimethyl In-Solution Labeling

	Amount of Protein in Reaction		
Reagent	5 pmol	1 pmol	0.3 pmol
1 pmol/uL protein in 100mM Na acetate, pH 5	5 uL	1 uL	0.3 uL
100mM Na acetate, pH	-	4	4.7
4% Formaldehyde	1uL	1uL	1 uL
260mM sodium cyanoborohydride	1uL	1uL	1 uL
Reaction Conditions	vortex 5min RT	vortex 5min RT	vortex 5min RT
4% ammonium hydroxide (quench reaction) *optional, but required for in-solution digest	1 uL, vortex	1 uL, vortex	1 uL, vortex
Final Volume	8 uL	8 uL	8 uL

Workflow "A" In-Solution Labeling + In-Gel Digestion

1. After derivatization, add 5x SDS-PAGE sample buffer to each tube.
2. To ensure neutral pH, 1 uL (or more) of 1M TrisHCL, pH 8.3, is added to each tube
3. Warm samples at 95 degrees for 5 min.
4. Separate each labeled protein in a separate lane of a 10 or 12% gel.
5. Stain with colloidal coomassie or other mass spectrometry compatible stain.
6. Excise protein bands and proceed with an in-gel digest protocol of choice
   a. You may usethe in-gel protocol found below under Workflow "B" or use the protocol described by Shevchenko et al (4).
7. Analyze the resulting peptide mixture using a mass spectrometry system set up for protein identification. (See Data Analysis section below.)

WorkFlow "A" In-Solution Labeling + In-Solution Digestion

Table 2. Trypsin Digest of Dimethyl Labeled Protein

Dimethyl Labeled Protein	Reaction volume	Volume of 1M Hepes pH 8 (rebuffer after quenching)	100mM Hepes pH 8 (uL)	Volume of 1 ng/uL trypsin in 100mM Hepes pH 8	Final Concentration (fmol/uL)
5 pmol	8 uL	3 uL	5	5 uL	250
1 pmol	8 uL	3 uL	9	1 uL	50
0.3 pmol	8 uL	3 uL	9	1 uL	16.67
Incubate overnight at 37°C					

Note: Because myoglobin does not contain cystiene residues, reduction and alkylation steps prior to digestion are not necessary. For participants preparing samples for MALDI-TOF MS/MS analysis it is recommended to follow the Filter Aided Sample Preparation protocol as described by Wiśniewski, et al (5).

WorkFlow "B" In-Gel Labeling followed by In-Gel Digest

1. Add a volume of unlabled myglobin corresponding to Table 1 to SDS-PAGE loading buffer.
2. Separate each protein sample in a separate lane of a 10 or 12% gel.
3. Stain with colloidal coomassie or other mass spectrometry compatible stain.
4. Excise protein bands and destain.
5. Shrink 2X with 200uL acetonitrile.
6. Add 100 µL of 10 mM DTT and incubate for ~40 min at 56 °C. Remove liquid
7. Add 100 µL 55 mM iodoacetamide and incubate in dark for ~45 min at RT. Remove liquid
8. Dehydrate gel pieces 2x with 200 µL acetonitrile.
   ***This is a potential stopping point where samples may be placed at 4°C overnight**.*

9. Post SDS-PAGE, in-gel derivatization is carried out for 20 min, using the same in-solution derivitizaion protocol in Workflow "A."
10. At the end of the reaction, excess reactants are removed from each gel slice.
    a. To quench any remaining reactants on gel surface, rinse individual gel slices with ~10ul of 4% ammonium hydroxide.
    b. Remove ammonium hydroxide, and wash gel slices with 300ul of 100mM ammonium bicarbonate a total of 5 times, removing liquid after each wash.
11. Dehydrate gel pieces 2x with 200 µL acetonitrile.
12. Add 30uL 12.5ng/uL trypsin in 100mM ammonium bicarbonate to gel piece and incubate on ice or in refrigerate for 2 hours. Then add 100mM ammonium bicarbonate to just cover gel pieces and incubate overnight at 37C.
13. To extract peptides add 100ul of extraction buffer (1:2 (vol/vol) 5% formic acid:acetonitrile) to each tube and incubate for 15 min at 37°C in a shaker.
14. Withdraw solution avoiding gel pieces. Place solution in a 500uL tube and take to dryness using a SpeedVac. Store -20°C until analysis.
15. Add an appropriate volume of 0.1%TFA to the dry peptides and sonicate 2-5 minutes.
16. Centrifuge the sample 15minutes at 10,000rpm and transfer to a injection vial for LC-MS.

Sample Analysis and Data Analysis

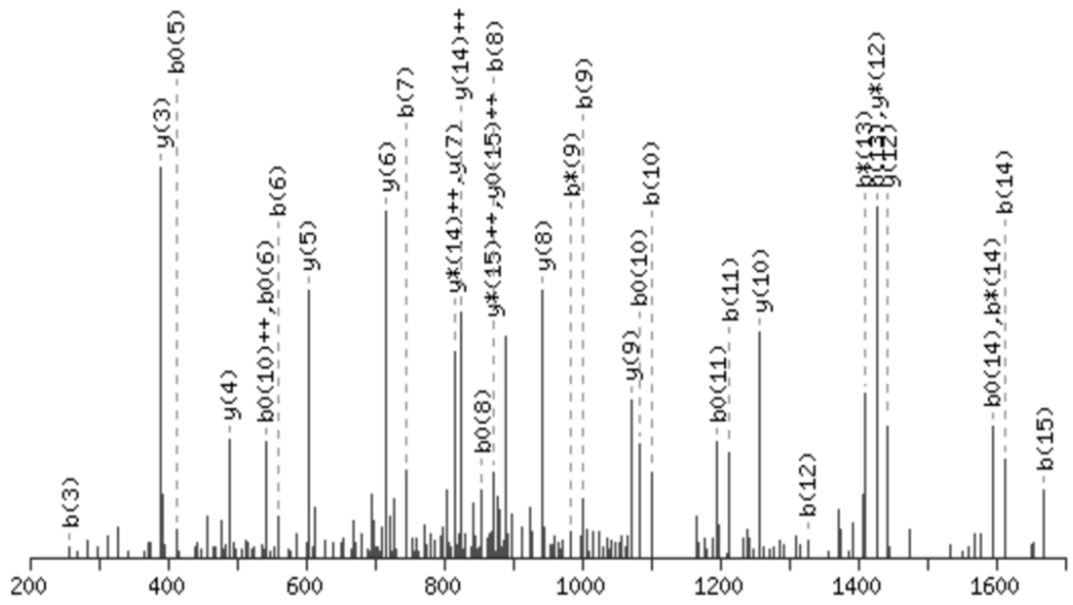
1. Samples are to be analyzed by MS/MS mass spectrometry typically used for protein identification. LC-MS with vented column/trap loading has been used to successfully remove excess reagent and identify labeled proteins. MALDI-TOF users may wish to perform the digest using FASP as described in Reference 5, or use manual desalting if the dimethyl labeling was performed in solution.
2. Data analysis is performed similar to protein identification analysis with trypsin as the proteolytic enzyme, carbamidomethyl Cys set as a fixed modification, pyro-Glu and Met oxidation as variable modifications, and maximum number of missed cleavages should be set to 4. Software for data analysis such as GPM, Mascot, PLGS, etc., must allow for custom modifications and preferably to the N-terminal amino acid of a peptide. For Dimethyl labeling, participants should include a mass addition of 28.0313 (C_2H_4) applied to the N-terminus and lysine amino acids as a variable modification.
3. Participants are requested to complete the appropriate table below, provide MS/MS spectra of the N-terminal trypsin peptide labeled by dimethylation, and an output file or screen capture from software used to make the N-terminal identification. In addition to data on the N-terminal trypsin peptide, participants are asked to provide a list of additional peptides that show evidence of n-terminal label (not lysine side chain label) on newly formed n-terminal amines as a result of trypsin digestion. Please fill out the accompanying analysis survey accessible by the link on Page 2.

Results for Identifying the N-Terminal Peptide of Dimethyl Labeled Proteins

Myoglobin starting amount	Amount of sample analyzed by MS (fmol)	Dimethy Labeling Workflow		Dimethyl Digest Workflow		Positive Identification?
		In solution	In-Gel	In Solution	In-Gel	
5 pmol						
1 pmol						
0.3 pmol						
Lower (specify)						

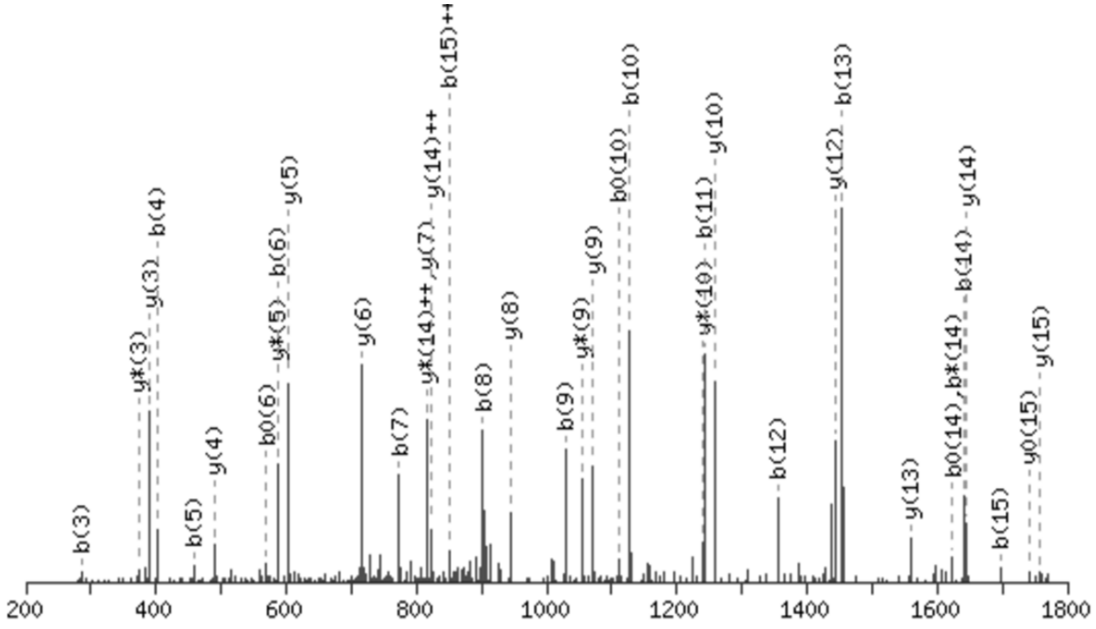
Example Data from Myoglobin Analysis

GLSDGEWQQVLNVWGK ***m/z*** **908.4539 (2+)**



#	b
1	58.0287
2	171.1128
3	*258.1448*
4	373.1718
5	430.1932
6	*559.2358*
7	*745.3151*
8	*873.3737*
9	*1001.4323*
10	*1100.5007*
11	*1213.5848*
12	*1327.6277*
13	*1426.6961*
14	*1612.7754*
15	*1669.7969*

DiMethyl-GLSDGEWQQVLNVWGK ***m/z*** **922.4694 (2+)**



#	b
1	86.0600
2	199.1441
3	*286.1761*
4	*401.2031*
5	*458.2245*
6	*587.2671*
7	*773.3464*
8	*901.4050*
9	*1029.4636*
10	*1128.5320*
11	*1241.6161*
12	*1355.6590*
13	*1454.7274*
14	*1640.8067*
15	*1697.8282*

References

1. Hennrich, M. L., Mohammed, S., Altelaar, A. F., & Heck, A. J. (2010). Dimethyl isotope labeling assisted de novo peptide sequencing. *Journal of the American Society for Mass Spectrometry*, *21*(12), 1957-1965..
2. Hsu, J. L., Huang, S. Y., Chow, N. H., & Chen, S. H. (2003). Stable-isotope dimethyl labeling for quantitative proteomics. *Analytical chemistry*, *75*(24), 6843-6852.
3. O Kleifeld, A Doucet, A Prudova, U auf dem Keller, M Gioia, J Kizhakkedathu, CM Overall. Identifying and quantifying proteolytic events and the natural N –terminome by terminal amine isotopic labeling of substrates. (2011), *Nature Protocols*; 6 (10), 1578-1611 doi:10.1038/nprot2011.382
4. Shevchenko A, Tomas H, Havlis J, Olsen J & Mann M; *In-gel digestion for mass spectrometric characterization of proteins and proteomes*.(2006) Nature Protocols, 1(6), 2856-2860.
5. Jacek R Wiśniewski, Alexandre Zougman, Nagarjuna Nagaraj & Matthias Mann. Universal sample preparation method for proteome analysis (2009) *Nature Methods* 6, 359-362.