Dear Colleagues,

Please find enclosed the ABRF 2014 PSRG samples that you requested from the ABRF Protein Sequence Research Group. This is the 26th 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 (MS)-based terminal sequencing. 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 enhance N-terminal sequencing of proteins by mass spectrometry. The current 2014 study entails terminal sequencing and identification of 3 purified proteins. The goal for this year is to test the abilities of participating laboratories to a) successfully derivitize the provided proteins, b) digest and identify the derivitized peptide fragments by mass spectrometry, and c) obtain terminal sequence information.

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 three known proteins provided separately, the labeling reagent (TMPP), and protocols for TMPP 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. Participants may also choose to analyze the provided protein standards with a mass spectrometry based N-terminal identification method of their choice such as dimethyl labeling by reductive amination.

(A) Workflow “in-solution labeling”:
1) TMPP labeling of proteins at N-terminus
2) SDS-PAGE and in-gel tryptic digestion or on-membrane cut-off filter tryptic digestion
3) MS analysis including data analysis
4) Identification of N-termini

(B) Workflow “in-gel labeling and digest”
1) SDS-PAGE, excision of bands and in-gel TMPP labeling
2) In-gel tryptic digestion and cleanup
3) MS analysis including data analysis
4) Identification of N-termini

Proteins supplied in this study include:

<table>
<thead>
<tr>
<th>Protein</th>
<th>Catalog Number</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin</td>
<td>Sigma H7379</td>
<td>1 mg</td>
</tr>
<tr>
<td>Myoglobin</td>
<td>Sigma M1882</td>
<td>1 mg</td>
</tr>
<tr>
<td>BSA</td>
<td>Sigma A0281</td>
<td>1 mg</td>
</tr>
<tr>
<td>TMPP</td>
<td>Sigma 29208</td>
<td>1 mg (required for Workflow A)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 mg (required for Workflow B)</td>
</tr>
</tbody>
</table>

Participants are asked to a) derivitize the proteins using one of the suggested protocols, and b) analyze samples for terminal sequencing using mass spectrometry.
Data submission:
Please provide data and analysis results directly to Amanda McGinnis at apmcg@umich.edu. You will be receiving 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 2014 meeting March 22-25, 2014, in Albuquerque, NM, 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 survey can be accessed here: https://www.surveymonkey.com/s/GMY372W

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-chairs Robert English (rdenglis@utmb.edu) or Sara McGrath (sara.mcgrath@fda.hhs.gov).

*** Due to high cost, this year’s study has limited participants. Please use your sample carefully, as there is no more TMPP available!

*** 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 1, 2014.

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
Mark Garfield – Research Technologies Branch NIAID/NIH
Pegah Jalili – Sigma-Aldrich
Ejvind Mortz – Alphalyse
Henriette Remmer – University of Michigan (ad-hoc)
William Hendrickson – EB liaison
Appendix

Sample Preparation
1. Record date sample is received and store sample in -80
2. Sample sent as approximately 1mg dried protein
3. Solubilize the proteins as follows to make 100pmol/uL solutions:
   a. To the hemoglobin tube add 208ul H2O and 100ul acetonitrile
   b. To the myoglobin tube add 490ul H2O
   c. To the BSA tube add 151ul H2O

Workflow “A” N-Terminal Labeling in Solution
1. Dissolve 1mg TMPP in 1.3mL of acetonitrile to make a 1mM solution
   a. Solution is stable for 2 days at -20°C
2. Make a 100mM Hapes buffer and pH to 8.0
3. Label the desired amount of each protein in separate tubes according to Table 1.
4. Alternatively or additionally, you may dimethyl-label proteins according to Table 2.
   a. **Note**: Formaldehyde and sodium cyanoborohydrde are toxic and must be handled with appropriate safety precautions.

<table>
<thead>
<tr>
<th>Table 1: TMPP In-solution Labeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of Protein in Reaction (pmol)</td>
</tr>
<tr>
<td>Reagent</td>
</tr>
<tr>
<td>100pmol/uL protein in H2O</td>
</tr>
<tr>
<td>0.1M Hepes pH 8.0</td>
</tr>
<tr>
<td>1mM TMPP in acetonitrile</td>
</tr>
<tr>
<td>Reaction Conditions</td>
</tr>
<tr>
<td>Final Concentration (pmol/uL)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: (Alternative Method) Dimethyl In-Solution Labeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of Protein in Reaction (pmol)</td>
</tr>
<tr>
<td>Reagent</td>
</tr>
<tr>
<td>10pmol/uL protein in 100mM sodium acetate, pH 5</td>
</tr>
<tr>
<td>4% Formaldehyde</td>
</tr>
<tr>
<td>260mM sodium cyanoborohydride</td>
</tr>
<tr>
<td>Reaction Conditions</td>
</tr>
<tr>
<td>Final Concentration (uM)</td>
</tr>
</tbody>
</table>

Workflow “A” In-Solution Labeling: In-Gel Digestion
1. For TMPP labeled protein, add a volume of sample corresponding to 100pmol protein to SDS-PAGE load buffer
2. For Dimethyl labeled protein, add a volume of sample corresponding to 10pmol protein to SDS-PAGE load buffer.
3. Separate each labeled protein in a separate lane of a 10 or 12% gel.
4. Stain with colloidal coomassie or other mass spectrometry compatible stain.
5. Excise protein bands and proceed with an in-gel digest protocol of choice
   a. You may use the in-gel protocol found below under Workflow “B” or use the protocol described by Shevchenko et al (1).
6. 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 Digest

Table 3. Trypsin Digest of TMPP Labeled Protein

<table>
<thead>
<tr>
<th>TMPP Labeled Protein</th>
<th>Volume of 20ng/µL trypsin in 100mM Hepes pH 8 (µL)</th>
<th>Volume of 100mM Hepes</th>
<th>Final Concentration (pmol/µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin 500pmol (20µL)</td>
<td>32.5</td>
<td>47.5</td>
<td>5</td>
</tr>
<tr>
<td>Hemoglobin 100pmol (3µL)</td>
<td>6.5</td>
<td>10.5</td>
<td>5</td>
</tr>
<tr>
<td>Myoglobin 500pmol (20µL)</td>
<td>17.0</td>
<td>63.0</td>
<td>5</td>
</tr>
<tr>
<td>Myoglobin 100pmol (3µL)</td>
<td>3.4</td>
<td>13.6</td>
<td>5</td>
</tr>
<tr>
<td>BSA 500pmol (20µL)</td>
<td>66.4</td>
<td>13.6</td>
<td>5</td>
</tr>
<tr>
<td>BSA 100pmol (3µL)</td>
<td>13.3</td>
<td>3.7</td>
<td>5</td>
</tr>
</tbody>
</table>

Incubate overnight at 37°C

Table 4. (Alternative Method) Trypsin Digest of Dimethyl Labeled Protein

<table>
<thead>
<tr>
<th>Dimethyl Labeled Protein</th>
<th>Volume of 20ng/µL trypsin in 100mM Hepes pH 8 (µL)</th>
<th>Volume of 100mM Hepes</th>
<th>Final Concentration (pmol/µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin 100pmol (12µL)</td>
<td>32.5</td>
<td>47.5</td>
<td>1</td>
</tr>
<tr>
<td>Hemoglobin 10pmol (3µL)</td>
<td>6.5</td>
<td>10.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Myoglobin 100pmol (12µL)</td>
<td>17.0</td>
<td>63</td>
<td>1</td>
</tr>
<tr>
<td>Myoglobin 10pmol (3µL)</td>
<td>3.4</td>
<td>13.6</td>
<td>0.5</td>
</tr>
<tr>
<td>BSA 100pmol (12µL)</td>
<td>66.4</td>
<td>13.6</td>
<td>1</td>
</tr>
<tr>
<td>BSA 10pmol (3µL)</td>
<td>13.3</td>
<td>3.7</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Incubate overnight at 37°C

Note: 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. (2).

WorkFlow “B” In-Gel Labeling with TMPP followed by In-Gel Digest

1. Add a volume of unlabeled protein as prepared in Table 1, corresponding to 50pmol or 10pmol of each protein with SDS-PAGE loading buffer.
2. Separate each unlabeled protein 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 200µL 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.
   a. This is a potential stopping point where sample may be placed in 4°C overnight.
9. Dissolve 2mg TMPP in 10uL acetonitrile then dilute further with 110uL of 0.5M ammonium bicarbonate.
10. Add TMPP solution to dry gel piece without delay and incubate 1 hour RT. Then add 400uL 0.5M ammonium bicarbonate and incubate 4 hours or overnight.
11. Mix 10 minutes then wash 3X with 300uL 100mM ammonium bicarbonate removing liquid after each wash.
12. Dehydrate gel pieces 2x with 200 µL acetonitrile.
13. 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 37°C.
14. 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.
15. Withdraw solution avoiding gel pieces. Place solution in a 500uL tube and take to dryness using a 
SpeedVac. Store -20°C until analysis.
16. Add an appropriate volume of 0.1%TFA to the dry peptides and sonicate 2-5 minutes.
17. Centrifuge the sample 15minutes at 10,000rpm and transfer to a injection vial for LC-MS.

Workflow “B” In-Gel Dimethyl Labeling followed by In-Gel Digest
1. Add a volume of unlabeled protein prepared in Table 2, corresponding to 10pmol or 1pmol of each 
protein with SDS-PAGE loading buffer.
2. Proceed with steps 2-9. As described above.
3. Just prior to adding to the dehydrated gel pieces vortex mix 100uL of 100mM sodium acetate pH 5 
buffer with 1uL 4%formaldehyde and 1uL 260mM sodium cyanoborohydride.
4. Incubate the gel pieces for 20min RT then wash 3X with 300uL 100mM ammonium bicarbonate 
removing liquid after each wash.
5. Proceed with steps 13-18 above.

Sample Analysis and Data Analysis
1. Samples are to be analyzed using the mass spectrometry MS/MS method typically used for protein 
identification. LC-MS with vented column/trap loading has been used to successfully remove excess 
TMPP and identify TMPP labeled proteins. Tandem MALDI-TOF users may wish to perform the 
digest using FASP as noted above or use manual desalting if the TMPP 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 needs to allow for custom modifications and preferably to 
the N-terminal amino acid of a peptide. For TMPP labeled samples participants need to include a 
mass addition of 572.1811(C28H33O10P) for the N-Terminus. For Dimethyl labeling the participants 
need to include a mass addition of 28.0313 (C2H4) applied to the N-terminus and lysine amino acids 
as a viable modification. Participants are requested to complete the appropriate table below,
MS/MS spectra of N-terminal trypsin peptide labeled with TMPP or by dimethylation, and an output 
file or screen capture from software used to make the N-Terminal identification.

References
1. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Andrej 
Shevchenko, Henrik Tomas, Jan Havlis, Jesper V Olsen & Matthias Mann Nature Protocols, VOL.1 
2. Universal sample preparation method for proteome analysis. Jacek R Wiśniewski, Alexandre 
### Results for Identifying the N-Terminal Peptide of Labeled Proteins using TMPP

<table>
<thead>
<tr>
<th>Protein/Starting amount (pmol)</th>
<th>Amount of sample analyzed by MS (fmol)</th>
<th>TMPP Labeling in solution</th>
<th>TMPP Labeling In-Gel</th>
<th>TMPP Digest in Solution</th>
<th>TMPP Digest In-Gel</th>
<th>Positive Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin 500pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>Hemoglobin 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β</td>
</tr>
<tr>
<td>Myoglobin 500pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>Myoglobin 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β</td>
</tr>
<tr>
<td>BSA 500pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSA 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Protein/Starting amount (pmol)</th>
<th>Amount of sample analyzed by MS (fmol)</th>
<th>Dimethyl Labeling in solution</th>
<th>Dimethyl Labeling In-Gel</th>
<th>Dimethyl Digest in Solution</th>
<th>Dimethyl Digest In-Gel</th>
<th>Positive Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin 100pmol</td>
<td></td>
<td></td>
<td></td>
<td>α</td>
<td>β</td>
<td></td>
</tr>
<tr>
<td>Hemoglobin 10pmol</td>
<td></td>
<td></td>
<td></td>
<td>α</td>
<td>β</td>
<td></td>
</tr>
<tr>
<td>Myoglobin 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myoglobin 10pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSA 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSA 10pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Example Results

<table>
<thead>
<tr>
<th>Protein/Starting amount (pmol)</th>
<th>Amount of sample analyzed by MS (fmol)</th>
<th>TMPP Labeling in solution</th>
<th>TMPP Labeling In-Gel</th>
<th>TMPP Digest in Solution</th>
<th>TMPP Digest In-Gel</th>
<th>Positive Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin 500pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>Hemoglobin 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β</td>
</tr>
<tr>
<td>Myoglobin 500pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>α</td>
</tr>
<tr>
<td>Myoglobin 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β</td>
</tr>
<tr>
<td>BSA 500pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSA 100pmol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myoglobin 100pmol</td>
<td>1200</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
</tbody>
</table>
Sample information

<table>
<thead>
<tr>
<th>model</th>
<th>protein</th>
<th>homologues</th>
<th>XML</th>
<th>gpmDB</th>
</tr>
</thead>
<tbody>
<tr>
<td>sp</td>
<td>MYG_HORSE</td>
<td>Myoglobin</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

no description

G-L-S-Dg5G+E WQ+QIVL+*VWEG K

m/z: 8475 (8475.0084)
mod: G2=572.18109
<table>
<thead>
<tr>
<th>ion type</th>
<th>m/z</th>
</tr>
</thead>
<tbody>
<tr>
<td>b(15-16)</td>
<td>186.124</td>
</tr>
<tr>
<td>b(9-10)</td>
<td>228.134</td>
</tr>
<tr>
<td>b(8-9)</td>
<td>257.124</td>
</tr>
<tr>
<td>a(15-13)</td>
<td>327.203</td>
</tr>
<tr>
<td>a(13-6)</td>
<td>443.204</td>
</tr>
<tr>
<td>a(7-2)</td>
<td>504.211</td>
</tr>
<tr>
<td>a(2)-[5]</td>
<td>660.329</td>
</tr>
<tr>
<td>a(4)-[4]</td>
<td>726.300</td>
</tr>
<tr>
<td>b(5-3)</td>
<td>825.368</td>
</tr>
<tr>
<td>a(1)-[6]</td>
<td>1103.417</td>
</tr>
<tr>
<td>a(1)-[8]</td>
<td>1289.496</td>
</tr>
</tbody>
</table>

Residue modification sets tested:

- Complete mods: i. none
- Potential mods: i. 572.1811@Q  ii. Oxidation@M, Oxidation@W, Deamidated@N, Deamidated@Q  iii. Deamidation@M, Deamidation@W
- N-terminal: i. Ammonia-loss@Q, Dehydrated@E (peptide)  ii. Acetylated, Acetyl (protein)