December 3, 2013

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[2]), 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[2].

(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:

Protein	Catalog Number	Quantity
Hemoglobin	Sigma H7379	1 mg
Myoglobin	Sigma M1882	1 mg
BSA	Sigma A0281	1 mg
TMPP	Sigma 29208	1 mg (required for Workflow A)
		6 mg (required for Workflow B)

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. Disssolve 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 Hepes 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 cyanoborohyride are toxic and must be handled with appropriate safety precautions.

Table 1: TMPP In-solution Labeling

	Amount of Protein in Reaction (pmol)	
Reagent	500	100
100pmol/uL protein in H2O	5uL	1uL
0.1M Hepes pH 8.0	10uL	1uL
1mM TMPP in acetonitrile	5uL	1uL
Reaction Conditions	30min RT	30min RT
Final Concentration (pmol/uL)	25	33.3

Table 2: (Alternative Method) Dimethyl In-Solution Labeling

	Amount of Protein in Reaction (pmol)	
Reagent	100	10
10pmol/uL protein in 100mM sodium acetate, pH 5	10uL	1uL
4% Formaldehyde	1uL	1uL
260mM sodium cyanoborohydride	1uL	1uL
Reaction Conditions	vortex, 5min RT	vortex, 5min RT
Final Concentration (uM)	8.33	3.33

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 usethe 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

TMPP Labeled Protein	Volume of 20ng/uL trypsin in 100mM Hepes pH 8 (uL)	Volume of 100mM Hepes	Final Concentration (pmol/uL)
Hemoglobin 500pmol (20uL)	32.5	47.5	5
Hemoglobin 100pmol (3uL)	6.5	10.5	5
Myoglobin 500pmol (20uL)	17.0	63.0	5
Myoglobin 100pmol (3uL)	3.4	13.6	5
BSA 500pmol (20uL)	66.4	13.6	5
BSA 100pmol (3uL)	13.3	3.7	5
Incubate overnight at 37°C			

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

Dimethyl Labeled Protein	Volume of 20ng/uL trypsin in 100mM Hepes pH 8 (uL)	Volume of 100mM Hepes	Final Concentration (pmol/uL)
Hemoglobin 100pmol (12uL)	32.5	47.5	1
Hemoglobin 10pmol (3uL)	6.5	10.5	0.5
Myoglobin 100pmol (12uL)	17.0	63	1
Myoglobin 10pmol (3uL)	3.4	13.6	0.5
BSA 100pmol (12uL)	66.4	13.6	1
BSA 10pmol (3uL)	13.3	3.7	0.5
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 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.
    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 37C.
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%formaldehye 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 NO.6 pp.2856-2860. 2006.
2. *Universal sample preparation method for proteome analysis*. Jacek R Wiśniewski, Alexandre Zougman, Nagarjuna Nagaraj & Matthias Mann. *Nature Methods* 6, 359-362 (2009).

Results for Identifying the N-Terminal Peptide of Labeled Proteins using TMPP

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

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

Protein/Starting amount (pmol)	Amount of sample analyzed by MS (fmol)	Dimethy Labeling in solution	Dimethyl Labeling In-Gel	Dimethyl Digest in Solution	Dimethyl Digest In-Gel	Positive Identification
Hemoglobin 100pmol						α
						β
Hemoglobin 10pmol						α
						β
Myoglobin 100pmol						
Myoglobin 10pmol						
BSA 100pmol						
BSA 10pmol						

Example Results

Protein/Starting amount (pmol)	Amount of sample analyzed by MS (fmol)	TMPP Labeling in solution	TMPP Labeling In-Gel	TMPP Digest in Solution	TMPP Digest In-Gel	Positive Identification
Hemoglobin 500pmol						α
						β
Hemoglobin 100pmol						α
						β
Myoglobin 500pmol						
Myoglobin 100pmol	1200	X	-	X	-	X
BSA 500pmol						
BSA 100pmol						

GPM32100048168: peptide model: 831.1.1 of sp|MYG_HORSE|

| model | protein | homologues | XML | gpmDB |

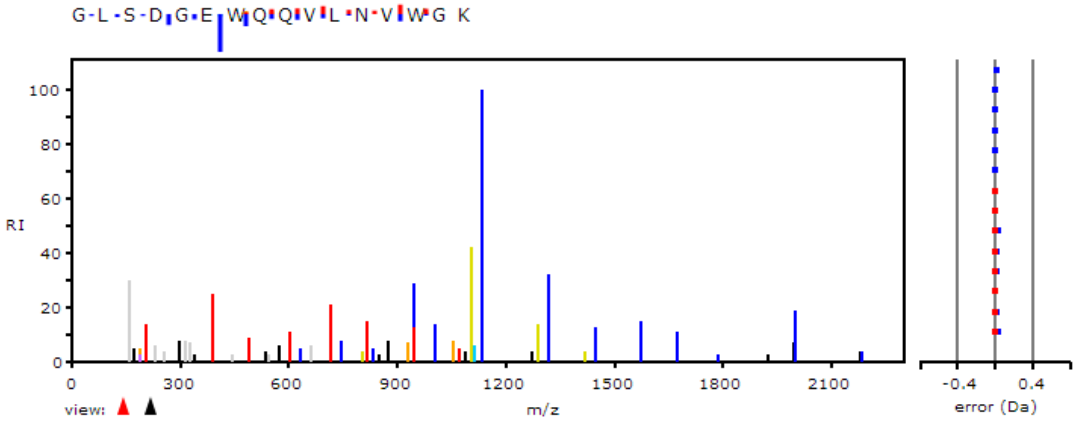
sp|MYG_HORSE|: Myoglobin;

Sample information

#	log(e)	log(l)	m+h	delta	ζ	sequence | validate | studio | mgf | mrm | details |
831	-14.4	5.33	2388.084	0.035	1/2	[m^{2} GLSDGEWQQVLNVWGK^{17}vead (8475) 0.084

mods: G2+572.18109

no description



matched/total:		# ions: 76%		intensity: 90%		μ: 0.00, σ: 0.01 Da
bond	^{+1}y	$^{+1}y^{-17}$	$^{+1}y^{-18}$	^{+1}b	$^{+1}b^{-17}$	$^{+1}b^{-18}$
$^{G}1$	1758.881	1741.854	1740.870	630.210	613.183	612.199
$^{L}2$	1645.797	1628.770	1627.786	743.294	726.267	725.283
$^{S}3$	1558.765	1541.738	1540.754	830.326	813.299	812.315
$^{D}4$	1443.738	1426.711	1425.727	945.353	928.326	927.342
$^{G}5$	1386.716	1369.690	1368.706	1002.374	985.348	984.364
$^{E}6$	1257.674	1240.647	1239.663	1131.417	1114.390	1113.406
$^{W}7$	1071.595	1054.568	1053.584	1317.496	1300.470	1299.486
$^{Q}8$	943.536	926.509	925.525	1445.555	1428.528	1427.544
$^{Q}9$	815.477	798.451	797.467	1573.613	1556.587	1555.603
$^{V}10$	716.409	699.382	698.398	1672.682	1655.655	1654.671
$^{L}11$	603.325	586.298	585.314	1785.766	1768.739	1767.755
$^{N}12$	489.282	472.255	471.271	1899.809	1882.782	1881.798
$^{V}13$	390.214	373.187	372.203	1998.877	1981.851	1980.867
$^{W}14$	204.134	187.108	186.124	2184.957	2167.930	2166.946
$^{G}15$	147.113	130.086	129.102	2241.978	2224.951	2223.967

Other observed ions:

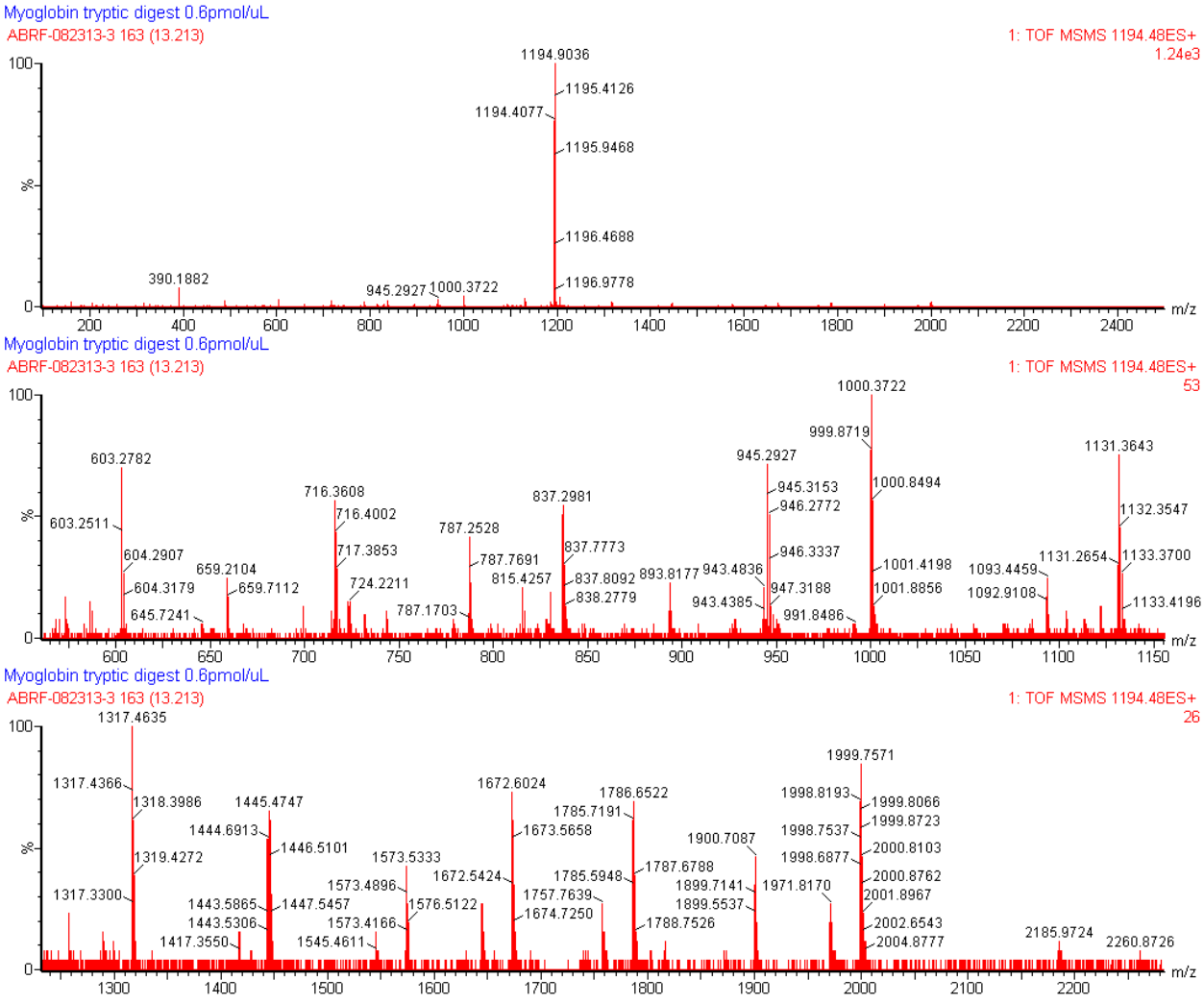
Ion type	m/z	Ion type	m/z
ay[5-6]	159.076	ay[12-13]	186.124
by[15-16]	186.124	by[5-6]	187.071
by[9-10]	228.134	by[11-12]	228.134
by[8-9]	257.124	by[7-8]	315.145
ay[13-15]	315.182	by[10-12]	327.203
by[11-13]	327.203	by[7-9]	443.204
ay[13-16]	443.277	ay[6-9]	544.251
ay[2-7]	660.299	ay[4-9]	716.300
a[3]	802.326	ay[4-10]	815.368
by[5-13]	1054.532	by[7-14]	1054.547
a[6]	1103.417	a[7]	1289.496
a[8]	1417.555		

Residue modification sets tested:

Complete mods:
i. none

Potential mods:
i. 572.1811@G
ii. Oxidation@M, Oxidation@W, Deamidated@N, Deamidated@Q
iii. Dioxidation@M, Dioxidation@W

N-terminal:
i. Ammonia-loss@Q, Dehydrated@E (peptide)
ii. ragged, Acetyl (protein)


Myoglobin tryptic digest 0.6pmol/uL
ABRF-082313-3 163 (13.213)
1: TOF MSMS 1194.48ES+
1.24e3
1194.9036
1195.4126
1194.4077
1195.9468
1196.4688
390.1882
945.2927
1000.3722
1196.9778
m/z
Myoglobin tryptic digest 0.6pmol/uL
ABRF-082313-3 163 (13.213)
1: TOF MSMS 1194.48ES+
53
1000.3722
999.8719
603.2782
945.2927
1131.3643
945.3153
1000.8494
716.3608
716.4002
837.2981
946.2772
1132.3547
603.2511
787.2528
604.2907
717.3853
837.7773
946.3337
1131.2654
1133.3700
659.2104
787.7691
1001.4198
1093.4459
604.3179
659.7112
724.2211
815.4257
837.8092
893.8177
943.4836
947.3188
1001.8856
1092.9108
645.7241
787.1703
838.2779
943.4385
991.8486
1133.4196
m/z
Myoglobin tryptic digest 0.6pmol/uL
ABRF-082313-3 163 (13.213)
1: TOF MSMS 1194.48ES+
26
1317.4635
1999.7571
1317.4366
1672.6024
1786.6522
1998.8193
1318.3986
1445.4747
1785.7191
1999.8066
1444.6913
1673.5658
1998.7537
1999.8723
1446.5101
1900.7087
1319.4272
1573.5333
1998.6877
2000.8103
1573.4896
1672.5424
1787.6788
1785.5948
2000.8762
1317.3300
1899.7141
1971.8170
1757.7639
2001.8967
1443.5865
1447.5457
1576.5122
1899.5537
1674.7250
1443.5306
1573.4166
1788.7526
2002.6543
1417.3550
1545.4611
2004.8777
2185.9724
2260.8726
m/z