ABRF-sPRG2010: Multi-Laboratory Evaluation of a Phosphopeptide Standard for Proteomics

Christopher Colangelo[1], James Farmar[2], Alexander R. Ivanov[3], Chris Kinsinger[4], Jeffrey A. Kowalak[5], Karl Mechtler[6], Brett Phinney[7], Manfred R. Raida[8], Susan T. Weintraub[9]

[1]Yale University, New Haven, CT; [2]University of Virginia, Charlottesville, VA; [3]Harvard University HSPH, Boston, MA; [4]NIST, Gaithersburg, MD; [5]NIH, Bethesda, MD; [6]IMP Research Institute of Molecular Pathology, Vienna, AUSTRIA; [7]University of CA, Davis, Davis, CA; [8]Experimental Therapeutics Center, Singapore, Singapore; [9]University of Texas HSC, San Antonio, TX

Introduction

Phosphorylation of proteins is among the most important of the posttranslational modifications, playing a critical role in regulation of many cellular processes. Proteomics laboratories are often called upon to determine sites of phosphorylation. Although there are numerous strategies available to carry out phosphopeptide analysis; many still find these analyses a formidable challenge. After digestion of the proteins, researchers use a variety of methods to separate and characterize the phosphorylated and non-phosphorylated peptides in bottom-up proteomics approaches. For the 2010 study of the ABRF sPRG (Proteomics Standards Research Group), a standard sample was prepared that reflects the challenges associated with the post-digestion analysis of phosphopeptides. Availability of this type of standard would be helpful in both assessing established methods and developing new approaches for phosphopeptide analysis/detection.

The singly and multiply phosphorylated peptides that were included in the sample are based on two tryptic peptides that were present in the 2003 sPRG study sample and 21 from the 2009 sPRG study sample; however the sites of phosphorylation at Ser, Thr and Tyr do not necessarily exist in nature. The peptides were mixed with tryptic digests of the corresponding HPLC purified six proteins. The sample was first analyzed and validated by members of the sPRG prior to distribution to requesting laboratories. The results from participating laboratories were submitted via an on-line questionnaire and the responses subsequently tabulated by the sPRG.

Goals

The primary goal of this study is to provide each participating laboratory with an opportunity to evaluate its capabilities and approaches with regard to detecting phosphopeptides and identifying sites of phosphorylation.

Materials and Methods

Peptide synthesis

Twenty three Peptides were synthesized by automated solid-phase Fmoc chemistry using an ABI 433a peptide synthesizer (Applied Biosystems, Foster City, CA). Amino acids were purchased from Novabiochem (Laeufelfingen, Switzerland). Peptides were dissolved in 0.1% TFA and purified on a Vision HPLC (Applied Biosystems) equipped with a Phenomenex Jupiter C18 column (10 mm i.d. 250 mm length, 5 µm particle size, 300 Å pore size; Phenomenex, Torrance, CA), and their identity was confirmed by MALDI-TOF/MS analysis on a Bruker Daltonics Reflex III, Bremen, Germany .

Sample preparation

Six HPLC purified recombinant proteins (five human and one bovine) were used for the study. Each protein was reduced, alkylated, and tryptic digested separately. The phosphopeptide standard was prepared by mixing prescribed equimolar ratios of synthetic phosphopeptides and tryptic digests of the corresponding proteins. The phosphopeptides added to the standard are shown in Table 1. Each sample contained 5 pmole of each constituent dispensed (10 µL/ sample) and dried by vacuum centrifugation to completion. This amount of protein was designed to permit multiple injections or use of several different approaches. The sample was distributed to requesters who were advised to reconstitute it in a buffer suitable for downstream analysis using vortexing and sonication for 5 minutes.

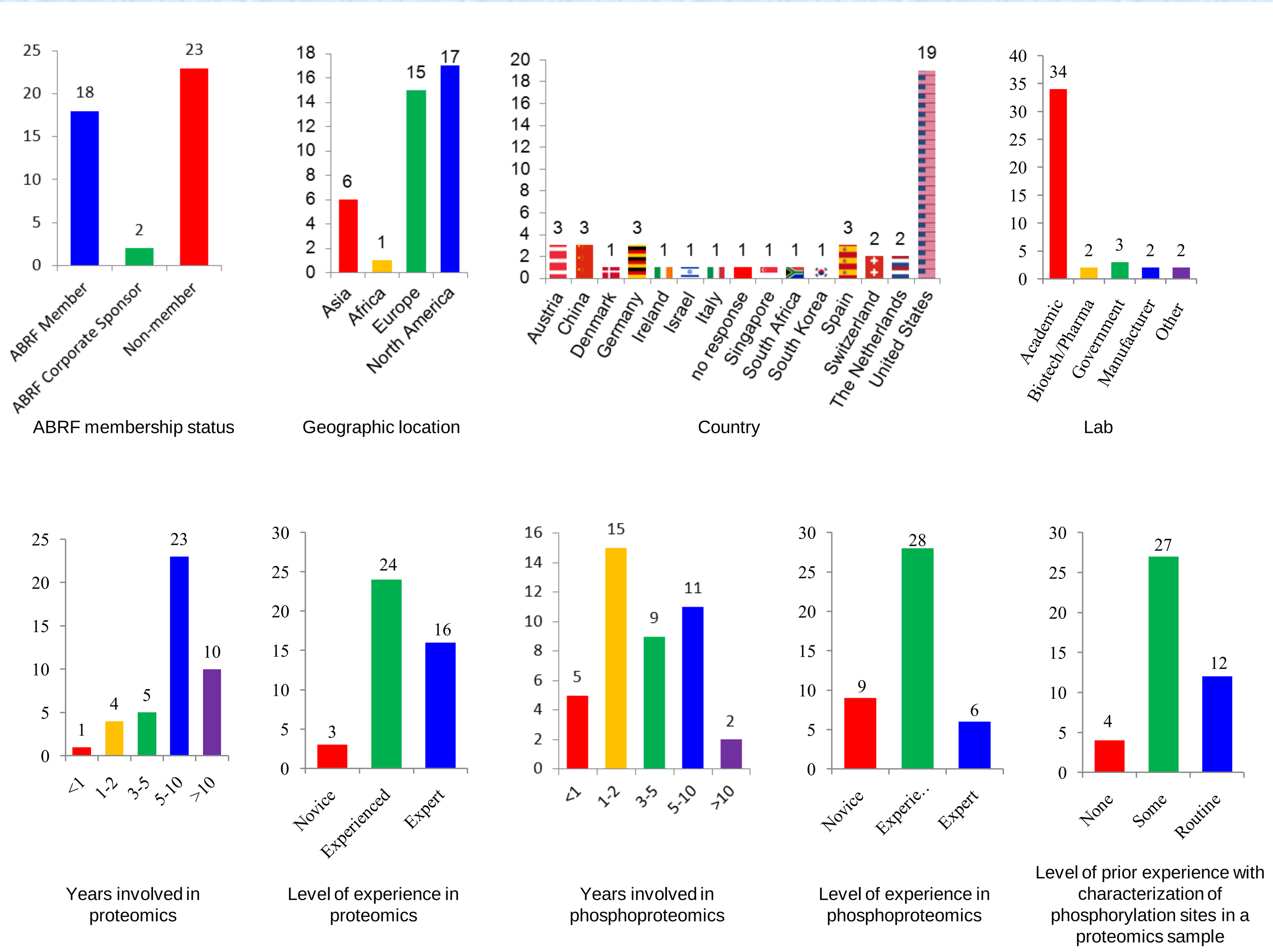

Figure 1. Demographics and Level of Experience of Study Participants

Table 1. Results for Study Participants

Sequence	# of p sites	# of times ID	% ID	11480	11821	12010	12166	12178s	12206	13037	13711	14997	16317	17171	19552	20036	20118	20566	20814	23123	23300	24107	24601	25616
Country				China	Greece	USA	USA	Austria	Netherlands	Denmark	China	Spain	USA	Austria	Germany	Israel	USA	Switzerland	USA	Germany	Spain	South Africa	Germany	USA
Enrichment				IMAC TiO2	TiO2	IMAC TiO2	TiO2	TiO2	TiO2	IMAC TiO2	None	TiO2	None	TiO2	TiO2	IMAC TiO2	IMAC	TiO2	TiO2	IMAC	TiO2	TiO2	ERLIC	TiO2
Media Used				AB		BC		B	B	BC		B		D	B	BC	F	B	B	C	B	B		B
Instrument				B	D	D	E	C	E	A	A	E	B	F	A	A	AF	A	A	E	E	F	D	A
Database Search Engine				E	B	B	F	A	G	H	B	C	A	A	A	B	A	G	A	D	C	F	A	I
# of correct identifications				16	0	7	19	21	16	19	14	19	11	11	20	12	13	21	11	6	19	5	17	10
% corrected identified				69.6%		30.4%	82.6%	91.3%	69.6%	82.6%	60.9%	82.6%	47.8%	47.8%	87.0%	52.2%	56.5%	91.3%	47.8%	26.1%	82.6%	21.7%	73.9%	43.5%
# of false positives				0	0	0	0	0	4	10	1	0	0	1	0	0	4	4	0	2	0	3	5	1
# of"Endogenous" IDs				0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	1	0	0	2	1
Years in Proteomics				5	5	5	0	10	5	10	5	3	5	5	5	10	5	10	5	5	5	3	1	5
Years in Pi-Proteomics				5	3	1	0	10	1	10	3	1	1	5	5	3	3	1	5	3	1	0	1	1
1st, Sometime, Routine				R	S	S	1	R	S	R	R	S	S	S	S	S	S	S	S	S	S	1	S	S
AEFAEVpSK (ALBU)	1	34	81%	x			x	x		x	x	x	x		x	x	x	x	x		x		x	x
LVNEVpTEFAK (ALBU)	1	40	95%	x			x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x
NVAVDELpSR (NQO2)	1	35	83%	x		x	x	x	x		x	x	x		x	x	x	x	x		x		x	x
ADEGIpSFR (PRDX1)	1	39	93%	x			x	x	x	x	x	x	x	x	x	x	x	x	x		x		x	x
DISLSDpYK (PRDX1)	1	34	81%	x		x	x	x	x	x	x		x		x	x	x	x	x			x	x	
DIpSLSDpYK (PRDX1)	2	25	60%	x			x	x	x	x		x		x	x			x			x			
DISLpSDpYK (PRDX1)	2	20	48%	x			x	x	x	x		x			x			x			x		x	
DIpSLpSDpYK (PRDX1)	3	10	24%					x	x			x		x	x			x			x		x	
LVQAFQFpTDK (PRDX1)	1	40	95%	x		x	x	x		x	x	x	x		x	x	x	x	x	x	x	x	x	x
DQGGELLpSLR (SYHC)	1	38	90%	x		x	x	x	x	x	x	x	x		x	x	x	x	x	x	x		x	x
IFpSIVEQR (SYHC)	1	38	90%	x			x		x	x	x	x	x		x	x	x	x	x	x	x	x	x	x
ESpTLHLVLR (UBIQ)	1	25	60%	x			x	x	x	x	x	x	x		x		x	x			x		x	
EpSpTLHLVLR (UBIQ)	2	15	36%				x	x	x	x		x		x	x			x			x			
EpSTLHLVLR (UBIQ)	1	23	55%	x			x	x	x	x	x			x	x	x	x	x		x		x		
TLSDpYNIQK (UBIQ)	1	31	74%	x				x		x	x	x	x		x	x		x	x	x	x		x	x
pTLSDYNIQK (UBIQ)	1	31	74%	x			x	x		x	x	x		x	x	x	x		x		x		x	x
pTLpSDYNIQK (UBIQ)	2	25	60%	x			x	x			x	x		x	x			x			x		x	
TITLEVEPpSDpTIENVK (UBIQ)	2	26	62%			x	x	x	x	x	x	x		x	x		x	x			x		x	
pTITLEVEPpSDpTIENVK (UBIQ)	3	10	24%					x	x	x		x		x				x			x			
pTIpTLEVEPpSDpTIENVK (UBIQ)	4	9	21%			x	x		x	x				x				x						
pTLpSDpYNIQK (UBIQ)	3	16	38%					x		x		x		x	x			x			x		x	
SVpSDYEGK (PDIA1)	1	17	40%				x	x		x					x	x	x		x					
THILLFLPKpSVSDYEGK (PDIA1)	1	26	62%	x		x	x	x	x			x	x					x			x		x	x
"Endogenous Peptide IDs"																								
TLpSDYNIQK(UBIQ)	1	8	19%							x			x	x									x	
DISLpSDYK (PRDX1)	1	5	12%							x										x				
DIpSLSDYK (PRDX1)	1	2	5%																				x	
DIpSLpSDYK (PRDX1)	2	2	5%																				x	
THILLFLPKSVpSDYEGK (PDIA1_BOVIN)	1	2	5%																					x
pSVSDYEGK (PDIA1_BOVIN)	1	3	7%							x														

Sequence	26402	27516	31610	32677v	33033	33168	40480	63103	63122	63490	64768	66514s	69464v	77681	77760s	79650	82972s	84573	85285s	87101	91231	92024	95031
Country	Spain	USA	South Korea	USA	USA	Italy	USA		USA	USA	Netherlands	USA	USA	USA	USA	Switzerland	USA	Ireland	USA	China	Sinapore	USA	Austria
Enrichment	IMAC	TiO2	TiO2	None	TiO2	TiO2	TiO2	IMAC TiO2	TiO2	TiO2	IMAC TiO2	IMAC TiO2	None	IMAC TiO2	None	TiO2	TiO2	None	None	None	ERLIC	IMAC	TiO2
Media Used	C	B	B		B	B	B	AB	B	B	AB	A		B		B	B				E	C	D
Instrument	D	D	A	E	E	B	A	B	N	F	A	A	C	E	D	A	A	E	B	NA	B	A	D
Database Search Engine	A	A	B	D	A	A	A	A	A	A	A	B	H	A	A		A	J	K		A	B	B
# of correct identifications	15	17	13	20	17	19	11	10	11	10	18	16	20	14	14	11	22	0	16	16	12	12	6
% corrected identified	65.2%	73.9%	56.5%	87.0%	73.9%	82.6%	47.8%	43.5%	47.8%	43.5%	78.3%	69.6%	87.0%	60.9%	60.9%	47.8%	95.7%		69.6%	69.6%	52.2%	52.2%	26.1%
# of false positives	1	1	2	0	0	2	5	0	0	2	1	1	0	2	0	0	0	29	0	2	4	0	5
# of"Endogenous" IDs	0	1	2	0	0	2	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Years in Proteomics	5	5	3	5	5	5	5	3	10	10	5	5	10	5	10	5	10	1	1	1	10	5	3
Years in Pi-Proteomics	1	5	1	3	5	1	3	1	3	5	5	N	5	3	0	5	5	1	0	1	1	3	0
1st, Sometime, Routine	S	R	S	S	R	S	S	S	S	S	S	S	R	S	S	S	R	R	1	S	R	R	S
AEFAEVpSK (ALBU)	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x		x	x		x	
LVNEVpTEFAK (ALBU)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x
NVAVDELpSR (NQO2)	x	x	x	x	x	x	x		x		x	x	x	x	x	x	x		x	x	x	x	
ADEGIpSFR (PRDX1)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x
DISLSDpYK (PRDX1)	x	x	x	x	x	x	x	x		x	x	x	x	x	x		x		x	x	x	x	
DIpSLSDpYK (PRDX1)	x		x	x	x	x			x		x		x	x	x		x		x	x	x		x
DISLpSDpYK (PRDX1)	x	x		x	x	x					x		x				x		x		x		
DIpSLpSDpYK (PRDX1)													x				x						
LVQAFQFpTDK (PRDX1)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x
DQGGELLpSLR (SYHC)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		x	x	x	x	
IFpSIVEQR (SYHC)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	
ESpTLHLVLR (UBIQ)	x	x		x	x	x	x				x	x	x			x	x				x		
EpSpTLHLVLR (UBIQ)		x		x	x												x			x		x	
EpSTLHLVLR (UBIQ)		x		x	x	x					x	x	x	x		x				x			
TLSDpYNIQK (UBIQ)	x	x	x	x	x	x	x		x	x	x	x	x			x	x		x	x		x	
pTLSDYNIQK (UBIQ)	x	x	x		x	x	x	x			x	x	x	x	x		x		x	x	x	x	
pTLpSDYNIQK (UBIQ)	x	x		x	x	x				x	x	x	x	x	x	x	x		x	x			
TITLEVEPpSDpTIENVK (UBIQ)	x	x		x		x		x	x			x	x	x			x		x	x			x
pTITLEVEPpSDpTIENVK (UBIQ)				x											x		x						
pTIpTLEVEPpSDpTIENVK (UBIQ)				x											x		x						
pTLpSDpYNIQK (UBIQ)				x		x					x	x	x	x			x			x			
SVpSDYEGK (PDIA1)			x			x		x	x	x	x		x				x		x			x	
THILLFLPKpSVSDYEGK (PDIA1)		x	x	x	x	x					x	x	x	x	x	x	x		x		x		x
"Endogenous Peptide IDs"																							
TLpSDYNIQK(UBIQ)				x			x			x		x											
DISLpSDYK (PRDX1)			x				x				x												
DIpSLSDYK (PRDX1)			x																				
DIpSLpSDYK (PRDX1)		x																					
THILLFLPKSVpSDYEGK (PDIA1_BOVIN)						x																	
pSVSDYEGK (PDIA1_BOVIN)						x	x																

Table 1 Key

Key	Used	Instrument	Search Engine
A	IMAC	Orbitrap	MASCOT
B	TiO2	Trap-FT	Sequest
C	select	Velos	VEMS
D	TiZr	3D Trap	PLGS
E	ERLIC	QTOF	pFind
F	Ni-NTA	TOF/TOF	ProteinPilot
G			Phenyx
H			Discover
I			Inspect
J			Biotools
K			X!Tandem
N		None listed	

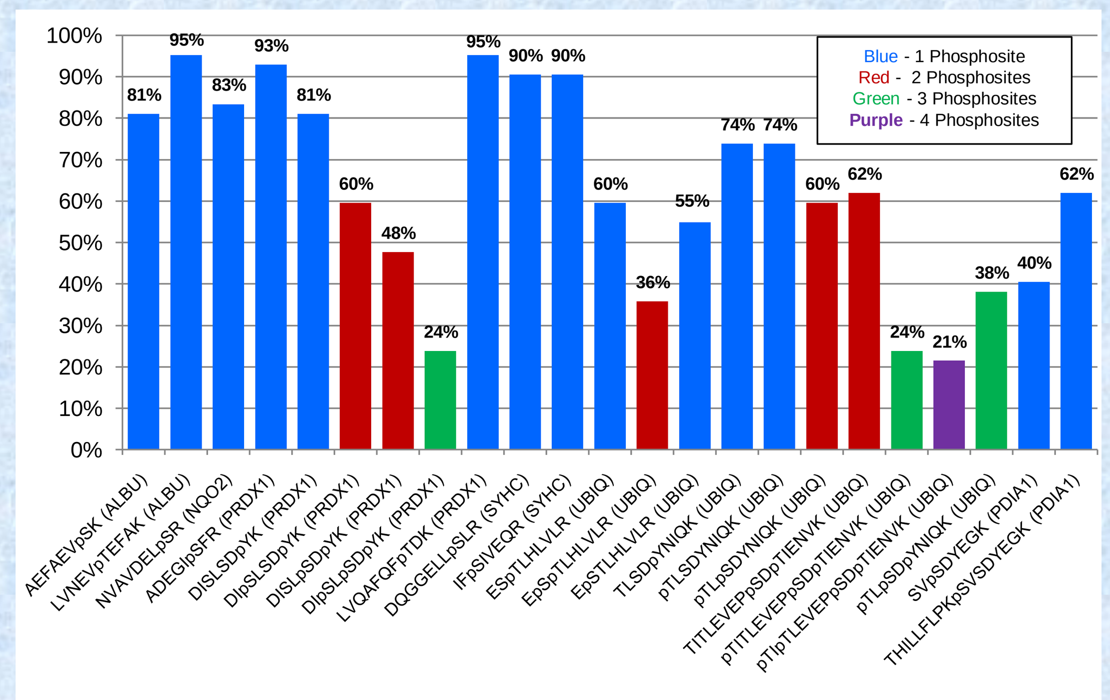

Figure 2. Phosphopeptide Identification Rate

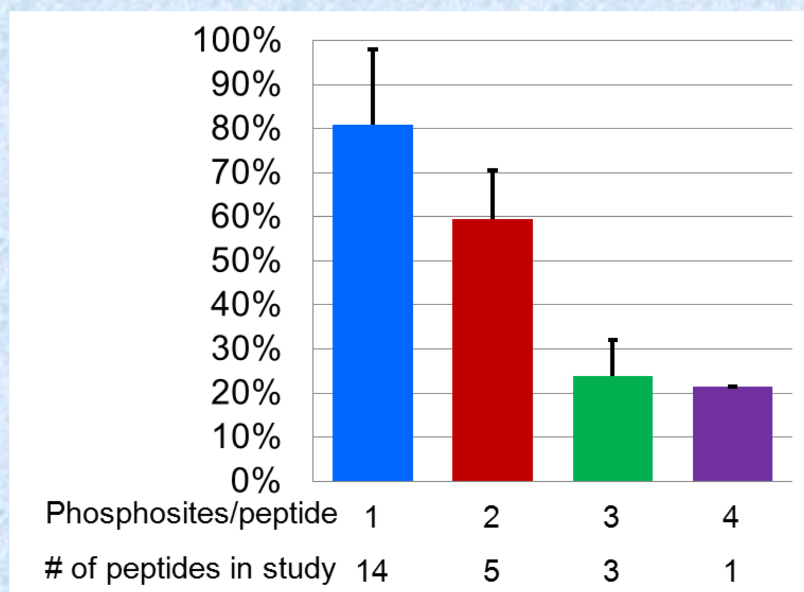

Figure 3. Success of Detection by Number of Phosphosites per Peptide

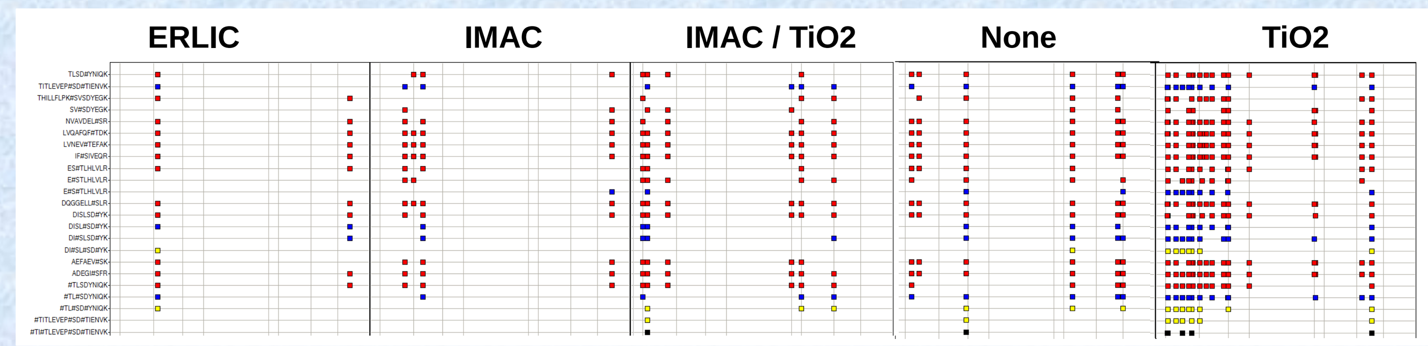

Figure 4. Phosphopeptides Detected by Enrichment Method. Sorted by the number of Phosphosites per Peptide

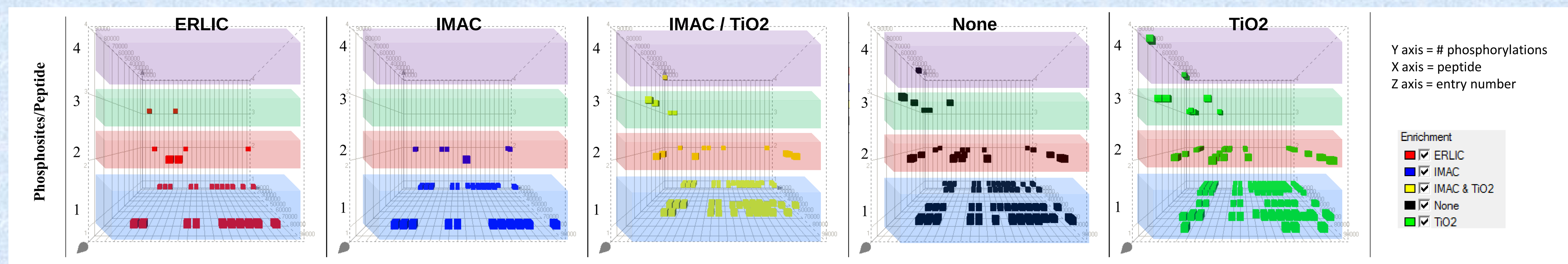

Figure 5. Phosphopeptides Detected by Enrichment and Number of Phosphosites per Peptide.

Results

Forty three result datasets were submitted anonymously through an on-line survey and included experimental, instrumental and other details. The datasets were from 18 ABRF members, 23 non-members, and 2 vendors. Detailed demographics and research experience is outlined in Figure 1. A summary of the survey results submitted by the participants is shown in Table 1. Included are analyses from laboratories of sPRG members (indicated by an “s” after the ID number); these analyses were conducted “blindly” in that the laboratory staff member who performed the analysis was provided with the same information as the other participants. Additionally, vendors were assigned a “v” after the ID number. None of the responding laboratories reported assignment/identification of all of the phosphopeptides that were added to the study sample. The results are suggested that the methods of phosphopeptide enrichment have different capabilities depending on the number of phosphosites in a peptide. Among the limited number of responses received TiO2 was the only one that afforded detection of the quadruply phosphorylated peptide. Interestingly, one laboratory that did not use enrichement prior to LC-MS was able to detect the quadruply phosphorylated peptide. It is important to keep in mind that the phosphpeptides were at equimolar concentation with the corresponding tryptic digest. As such, different results might be obtained if phospeptides are present at sub-stoichoimetric levels.

Conclusions

These results indicate that identification of multiply-phosphorylated peptides remains a challenge, despite advances in sample preparation and instrumentation. It is important to note that there was an insufficient number of responses using any of the individual technologies to afford a statistically significant measure of the ability of any method to “get the correct answer.”

Acknowledgements

Mathias Madalinski of the Mechtler Lab (ICP Austria) for peptide synthesis, purification and analysis.

Jim Makusky (NIH) as the anonymizer

Sigma for the purified proteins

Please visit www.abrf.org/sPRG for more in-depth examination of the methods used and the experimental results obtained.