19

PHOSPHOPROTEOMIC ANALYSIS OF LYMPHOCYTE SIGNALING

Lulu Cao, Kebing Yu, and Arthur R. Salomon

1. INTRODUCTION

The successful sequencing of the human genome is a monumental accomplishment. The application of this mountain of information to gain a better fundamental understanding of the molecular basis of disease represents a formidable challenge in the post-genomic era. For the first time, analysis of the whole complex system may be considered in parallel to traditional focused approaches. New global approaches must be developed to exploit this resource. One technique, the use of mRNA expression arrays, has gained new popularity in visualizing global patterns of transcription. Unfortunately there is a relatively weak correlation between transcript abundance and the abundance of individual proteins[1]. Also, critical signaling pathway components such as post-translational modifications, and protein-protein interactions are overlooked by the expression arrays. The emerging field of proteomics seeks to address some of these challenges by focusing on which proteins are expressed, their abundance, how they interact, and how they are modified in a global, unbiased fashion. Modern mass spectral proteomic tools leverage the availability of genomic sequence databases to match experimental spectra to genome-derived peptide and protein sequence. This report will focus on some recent developments in bioinformatics and proteomics methods based on mass spectrometry that provide new capabilities to examine the structure of signaling cascades through global phosphorylation site analysis from complex lysates.

Department of Molecular Biology, Cell biology and Biochemistry, Brown University, Box G-E335, Providence, RI 02912

2. METHODS FOR PHOSPHOPROTEOMICS

Phosphorylation sites from cell-derived signaling molecules are extremely difficult to detect due to their low abundance and low stoichiometry of phosphorylation in a cellular milieu of abundant unphosphorylated proteins. Although many approaches have been adopted to overcome this limitation, the limited specificity and efficiency of phosphorylation site isolation from highly multidimensional chromatographic separations such as MudPIT[2] or low pH, strong cation exchange[3] renders a unique technical challenge in revealing the phosphoproteome. Fundamentally, the poor quality of phosphopeptide tandem mass spectra arising from their low abundance, poor ionization, and abundant neutral loss of phosphate from phosphoserine and phosphothreonine residues necessitates the manual validation of every candidate MS/MS spectra. This time-consuming burden can be overcome by increasing the quality of the data through the use of accurate mass[4], better fragmentation methods such as ETD[5], more highly selective chromatography[6,7], and development of new statistical algorithms[8]. Limited dynamic range of mass spectrometers necessitates selective enrichment of phosphopeptides to obtain the highest numbers of phosphorylation sites from a complex mixture in the shortest amount of time. Furthermore, optimal MS/MS spectra are obtained in the positive ion mode where the additional negative charge of the phosphate group reduces the ionization efficiency of phosphopeptides relative to unphosphorylated peptides. Unfortunately, chemical removal of the phosphorylation site to improve the ionization of phosphopeptides such as via beta-elimination leads to ambiguity of the presence of the original phosphorylation site due to possible non-specific reactions[9]. If we want definitive proof of the modification from the mass spectrum, it is possible to overcome the dynamic range problem and poor ionization of phosphopeptides by enriching the sample as much as possible.

By far the most established method for phosphopeptide enrichment has been immobilized metal affinity chromatography (IMAC)[10]. A range of variations has been published, using different chromatographic materials, metal ions and manual or automated setups[10-13]. Often, it is much easier to isolate phosphopeptides from complex mixtures than to identify "the" phosphopeptide(s) from individual proteins. Background binding of peptides containing a high percentage of acidic amino acids to the IMAC column leads to significant contamination of the enriched phosphopeptide samples and difficulty applying IMAC alone to complex cellular lysates[6]. Increased IMAC selectivity may be obtained through methyl ester derivatization of acidic amino acids, resulting in higher yields of phosphopeptides[6,7,11,14].

3. PHOSPHOPROTEOMIC PLATFORM

Here I will describe an efficient phosphoproteomic method capable of discovery of hundreds of phosphorylation sites within a single day based on methyl esters IMAC[6]. This method has been improved significantly[6] from its original implementation[11,14] (Figure 1a). First, complex mixtures of proteins are obtained either from cells or tissue. Standard protein purification techniques such as immunoaffinity purification of phosphotyrosine-containing proteins are useful in focusing attention on smaller subpopulations of proteins such as tyrosine-phosphorylated proteins. After addition of a standard phosphopeptide(s), the cell or tissue derived proteins are most typically digested with trypsin and the

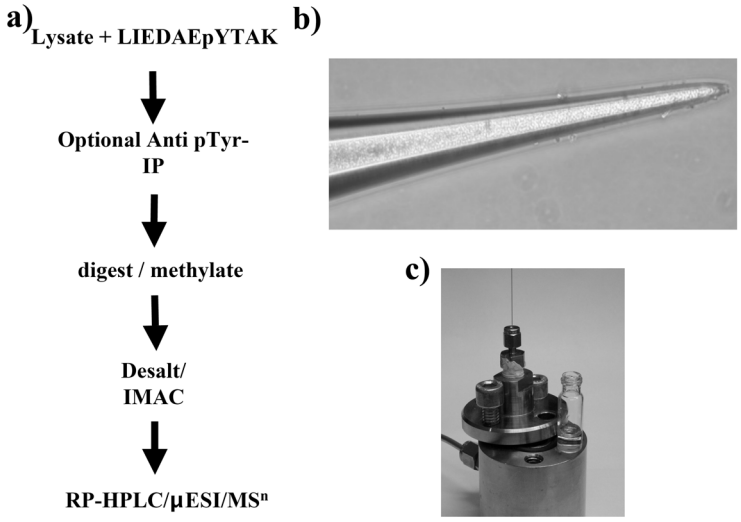

Figure 1. Proteomic platform for the analysis of phosphorylation sites from complex cell lysates as originally described by Brill et al[6] (Reprinted with permission; Copyright American Cancer Society). **(a)** After addition of a standard phosphopeptide used for external relative quantitation, proteins are optionally purified by phosphotyrosine immunoprecipitation prior to tryptic digestion of peptides. The resulting collection of peptides is methylated in methanolic HCl to increase the selectivity of subsequent separations and residual acid removed on a reversed-phase C18 column. Phosphorylated peptides are enriched on an immobilized metal affinity chromatography (IMAC) column and then eluted to a second reversed-phase C18 column. Peptides are then slowly eluted with a gradient of acetonitrile into the mass spectrometer through a **(b)** custom-made electrospray emitter tip at 20 nl/min. **(c)** All columns are hand prepared on a pressure bomb through the application of 600 psi of pressure to an emulsion of chromatographic resin through a 360 μm o.d. glass capillary fitted with a porous frit.

resulting peptides are converted to methyl esters with acidic methanol. Although global internal labeling such as with D3/D0 methyl esters or iTRAQ reagents provides the best relative quantitation[15], we have found addition of synthetic, non-genomic phosphopeptides to be a simple yet powerful method for relative quantitation of phosphorylation[14]. The addition of a reversed-phase C18 "desalting" column after peptide methylation and before the IMAC column efficiently removes the residual acidity from the methylation reaction and increases phosphopeptide retention on a high capacity IMAC column. Peptides are then slowly introduced to the high capacity IMAC column through gradient elution from the desalting column. Phosphorylated peptides are retained on the IMAC column while unphosphorylated proteins are removed with wash buffer. Finally peptides are slowly introduced to a 75 μm i.d. C18 "pre-column" with phosphate buffer and eluted into the mass spec through a 5 μm home-pulled electrospray emitter tip pulled on 75 μm i.d. capillary and packed with C18 resin (analytical column; Figure 1b). To maximize sensitivity, all columns in the system are prepared in glass capillaries with a "pressure bomb" (Figure 1c).

4. AUTOMATION OF THE PHOSPHOPROTEOMIC PLATFORM

Automation of proteomic methods is essential to increase both the reproducibility and the throughput of the analysis. The rapid and reproducible identification of phosphorylation sites from complex cell-derived mixtures is a crucial tool to accelerate signaling pathway discovery. It is critical to define not only the location of phosphorylation on proteins but to define each phosphorylation site's biological significance within the pathway. The analysis of site-directed mutants and protein disruptions of novel phosphorylation sites will provide insights into the placement of sites within pathways. As a complement to traditional approaches, the ability to define the perturbations in global patterns of phosphorylation resulting from disruption of a protein or novel phosphorylation site will provide vital clues about whether a new phosphorylation site is relevant to a given signaling pathway and its location within the pathway. A single time course phosphoproteomic experiment typically generates information about hundreds of sites of phosphorylation. A high-throughput system is therefore essential to perform both the initial timecourse phosphoproteomic experiments but also the timecourse data-inspired phosphoproteomic mutant analyses.

Here I describe a fully automated phosphoproteomic system for high-throughput phosphorylation site analysis controlled by custom-made automation software adapted from Ficarro et al.[7] (Figure 2). This system not only includes critical automation of the necessary phosphopeptide separations and introduction of peptides into the mass spectrometer as originally described by Ficarro et al[7],

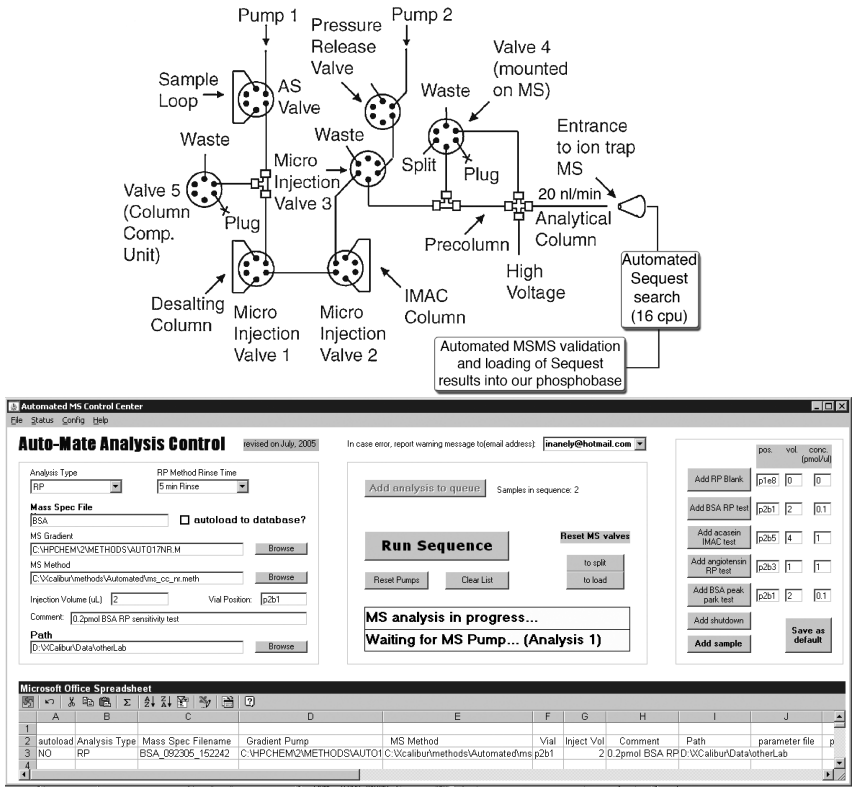

Figure 2. Fully automated system for the analysis of phosphorylated peptides adapted from Ficarro et al[7] (Copyright John Wiley and Sons Ltd; Reproduced with permission.

but also the bioinformatic analysis of the resulting data. A typical time course experiment generates tens of thousands of MS/MS spectra in a single day that need to be compared to theoretical spectra derived from genomic sequence databases with algorithms such as SEQUEST. Also software must be written to provide the storage, organization, visualization and statistical analysis of the data. We have directly coupled our automated mass spectrometer to a 16 cpu sequest cluster and a custom-made phosphoproteomic relational database. After completion of data acquisition, the automation software controls the transfer of data to the SEQUEST cluster, searching of the data, and deposition of the searched data into our custom-made proteomic relational database on a 3

terrabyte data server. After loading an autosampler with methylated peptides from complex mixtures, the system provides fully automated, unattended multi-dimensional chromatography with IMAC, nanospray with peak parking, SEQUEST searching on a cluster of computers, and bioinformatic analysis of the SEQUEST results in our database. To maintain 24/7 operational status, the automation software alerts the operator if any problems arise via email through its diagnostic capabilities.

Currently, there are no commercially available relational databases for the analysis of large proteomic data sets. We have designed a proteomic relational database that contains functionality to accelerate both the manual validation of spectra and the bioinformatic analysis of proteomic data. The database has the ability to display proteomics comparative analyses in a user-friendly, intuitive interface allowing the investigator to quickly discern temporal and spatial patterns of phosphorylation sites. One key feature of the database is an automated match of newly acquired spectra to previously manually validated spectra (2732 spectra so far), obviating redundant manual validation. Another critical feature of the database is an interactive MS/MS spectral editor allowing the electronic storage of manual annotations added directly to spectra during manual spectral validation. We have integrated other newly developed databases, such as the Human Reference Protein Database (HPRD)[16], directly within our database to provide automated lookup of phosphorylation sites and protein-protein interactions from the literature and hyperlinking of the relevant journal articles directly to the proteomic data. This feature allows the user to quickly investigate literature connections to the novel phosphorylation sites revealed in phosphoproteomic experiments, minimizing manual PubMed searching. Ambiguous protein naming in the literature is clarified through querying HPRD by peptide sequence (not currently available through the HPRD website). While the data is loaded into the database with the automation software, every peptide is BLAST searched against the genomic database. These archived BLAST searches allow the user to rapidly analyze the many names associated with a peptide assignment and permanently reassign the correct name to each peptide. Together, this bioinformatic infrastructure provides for the rapid analysis of phosphoproteomic data.

5. PHOSPHOPROTEOMICS APPLIED TO LYMPHOCYTE SIGNALING

To determine whether this IMAC technology platform is sensitive enough to view phosphorylation sites from human T cells, pervanadate treated Jurkat T cells were analyzed as described in Brill et al[6]. This analysis of phosphotyrosine immunoprecipitated and IMAC enriched phosphopeptides from 20 minute pervanadate stimulated Jurkat cells revealed 182 unique phosphorylation sites

Table 1. Phosphopeptides detected in pervanadate treated Jurkat cells as reported by Brill at al[6]

Protein	Peptide	Protein	Peptide
26S proteasome 4	KQEGTPEGLpYL	MAP4	TDpYIPLLDVDEK
Ataxin-2-like	EIESSPQpYR	NICE-4	RYPSSISSpSPQKDLTQAK
Ataxin-2-like	GPHHLDNSpSPGPGSEAR	NICE-4	NPSDSAVHpSPFTK
Ataxin-2-like	LQPSSpSENSLDPFPPR	NICE-4	RYPSSIpSSpSPQK
Ataxin-2-like	GPPQpSPVFEGVYNNSR	dynamitin	pYADLPGIAR
AF6	ADHRpSSPNVANQPPpSPGGK	P53bp	VLLpSPSIPSVGQDQTLpSPGSK
AF6	EpYFTFPASK	PAR3	FSPDSQpYIDNR
AF6	ADHRSpSPNVANQPPpSPGGK	elfin	VTPPEGpYEVVTVFPK
AF6	SQDADSPGpSpSGAPENLTFK	Cbp/PAG	AEFAEpYASVDR
AHCP	VGSAAPSRSPSEpTGR	PLCγ1	pYQQPFEDFR
Apg3p	ALEVAEpYLTPVLK	PLCγ1	IGTAEPDpYGALYEGR
ARMET	MWApTQGLAVALALpSVLPGSRALR	PLCγ2	DINSLpYDVSR
BICD2	RpSPILLPK	PLCγ2	EFSVNENQLQLpYQEK
BICD2	pSPILLPK	PIP5K	SApYSSFVNLFR
HEF1	TGHGYVpYEYPSR	Plakophilin 4	TYpYSPVYR
CBL	IKPSSpSANAIpYSLAAR	Plakophilin 4	VGpSPLTLTDAQTR
CD28	HYQPpYAPPR	Plakophilin 4	SAVpSPDLHITPIpYEGR
CD2 binding	pSIEVENDFLPVEK	Pre-mRNA Factor I	TPAILpYTYSGLR
CD3δ	NDQVpYQPLR	PTP12A	SGSpYSYLEER
CD3ε	DLpYSGLNQR	PTP12A	pYETSSTSAGDRpYDSLLGR
CD3ζ	SADAPApYQQGQNQLpYNELNLGR	PTPB	YKIQILpTVpSGGLFSK
CD3ζ	NPQEGLpYNELQK	R3HD	ASpSFSGISVLTR
CD3ζ	KNPQEGLpYNELQK	RAMA1	pSPQLSDFGLER
CD3ζ	GHDGLpYQGLSTATK	RCAS	KLpSGDQITLPTTVDpYSSVPK
CD3ζ	GKGHDGLpYQGLSTATK	RCAS	KLSGDQIpTLPTTVDpYSSVPK
CD3ζ	pSADAPAYQQGQNQLpYNELNLGR	Rho6	p(SSS)LSAANTSQTNPQGAVSSTVSGLQRQpSK
CD3ζ	SADAPAYQQGQNQLpYNELNLGR	Rho7	p(ST)AALEEDAQILK
CD5	pSHAENPTASHVDNEpYSQPPR	Rho7	SGTLKpSPPKRFDTTAINK
c-Mpl binding	EQYVPPRpSPK	RhoGAP12	ATpTPPNQGRPDpSPVpYANLQELK
CQT7	GEKGpTAGLKGK	RIK1	LQDEANYHLpYGSR
DOCK180	RSpSVVFADEK	RIK1	LQDEANYHLYGpSR
DOCK180	VEQEEPISPGp(ST)LPEVK	RIK1	LQDEANpYHLYGSR
p56dok	GQEGEpYAVPFDAVAR	RNA Pol II	YSPTpSPTpYSPTSPK
DYRK1B	IYQpYIQSR	RNPL	p(YY)DSRPGGYGYGYGR
Emerin	DSAYQpSITHpYRPVSASR	RNPL	YSGGNYRDNpYDN
endofin	NEIIQpSPISQVPSVEK	RN-tre	IEVLPVDTGAGGpYSGNSGpSPK
endofin	NEIIQpSPISQVPpSVEK	RN-tre	LIIPPVDpYLPDNR
ERBIN	SKSApTLLpYDQPLQVFTGSSSSSDLISGTK	ROA3	p(SSGS)PYGGGYGSGGGSGSRGGY
ERBIN	AQIPEGDpYLSYR	hnRNPH1	HpTGPNSPDTANDGFVR
Formin-binding 17	TVpSDNSLSNSR	U1 snRNP p70	YDERPGPpSPLPHR
FOP	GPTTGEGALDLSDVHpSPPKpSPEGK	SCA3	QpYATLDVpYNPFETR
FXR2	TDGpSIpSGDRQPVTVADYISR	SCA3	NpYGSpYSTQASAAAATAELLK
FYB	TTAVEIDpYDSLK	SCA3	NpYGSYSTQASAAAATAELLK
G3BP	p(SSS)PAPADIAQTVQEDLR	SCA3	QpYATLDVYNPFETR
GADS	AELGSQEGpYVPK	sim. to BMI gene	TVDLPKpSPK
GADS	RHpTDPVQLQAAGR	sim. PTP1c	TApSFGGITVLTR
GIT2	QATTNVYQVQpTGSEYTDTpSNHSSLK	sim. cdc42GAP	SpYAFETQANPGK
GOA5	NTDpYTELHQQNTDLIpYQTGPK	sim. cdc42GAP	NGGSLpSFDAAVALAR
H4	LDQPVpSAPPpSPR	Sim. cell surface ant.	EHLRSpSVIDRK
HIP-55	LRpSPFLQK	sim. R-spondin	QFpYLYKGK
hyp. FLJ20686	STSLpSALVR	Sim. Ser/Arg matrix 1	HRPpSPPApTPPPK
hyp. H41	LQLDNQpYAVLENQK	SIT	pSGESVEEVPLpYGNLHYLQTGR
hyp. LOC144100	KVTpSPLQSPpTK	SIT	SGESVEEVPLpYGNLHYLQTGR
hyp. PRO 0971	SQpYEVFR	SIT	IPGPGpTPVKpYSEVVLDSEPK
hyp. XP_211316	VTPTLpSLKGK	SIT	pYSEVVLDSEPK
hyp. XP_298586	QSpTRSpSAGR	SIT	IPGPGTPVKpYpSEVVLDSEPK
ICAM-5	ARGAVLpTApTVLAR	SIT	ARApSFPDQApYANSQPAAS
ITK	FVLDDQpYTSSTGTK	SIT	IPGPGTPVKpYSEVVLDSEPK
intersectin 2	EEPEALpYAAVNK	SLP-76	NHpSPLPPPQTNHEEPSR
intersectin 2	REEPEALpYAAVNK	SRm300	p(SS)TPPGESpYFGVSSLQLK
JNK/SAPK-assoc	SASQpSpSLDKLDQELK	SRPK2	TVpSASSpTGDLPK
KIAA0147	NpSLESISSIDR	STAM2	LVNEAPVYSVpYSK
KIAA0157	PQAVGSSNpYASTSAGLK	TRAP150	IDIpSPSTFR
KIAA0853	GPRpTPpSPPPPIPEDIALGK	TBC1 domain prot.	RTpp(SST)LDSEGTFNSYR
KIAA0853	pTPpSPPPPIPEDIALGK	tubulin chaperone B	LGEpYEDVSR
KIAA1046	VQRPKEEpSpSEDENEVSNILR	Ubiquitin-associated 2	LPVDpYpYGIPFAAPTALASR
KIAA1389	FSFYGNLpSPR	UBPO	VYDQTNPpYTDVR
KIRREL	TPYEApYDPIGK	UBPO	VSDQNpSPVLPK
KIS1	DLPNpYNWNSFGLRFGKR	ZAP-70	RIDTLNpSDGpYTPEPAR
LAT	LPGSYDSTpSSDSLYPR	ZAP-70	IDTLNSDGpYTPEPAR
HS1	GFGGQpYGIQK	ZAP-70	RIDTLNSDGpYTPEPAR
Leupaxin	STLQDpSDEpYSNPAPLPLDQHSR	ZAP-70	ALGADDSYpYTAR
Limatin	TSSESIpYSRPGSSIPGSPGHTIpYAK	ZAP-70	IDTLNpSDGpYTPEPAR
LIMD1	VpSPGLPpSPNLENGAPAVGPVQPR	ZDH5	p(SSS)LKpSAQGTGFELGQLQSIR
LIM lipoma	YpYEGYYAAGPGYGGR	ZO-2	HDDIpYAVPIK
LIM lipoma	YYEGYpYAAGPGYGGR		

Adapted with permission from L.M. Brill et al, *Anal. Chem.* 76(10), 2763-2772 (2004).

Table 2: Phosphopeptides from anti-CD3/CD4 stimulated Jurkat cells as reported in Salomon et al[14].

Protein	NCBI GI #	Peptide Sequence	PO_4 Site	Known Site	0 min	0.5 min	1 min	2 min	5 min
Jurkat									
CAS-L	5453680	TGHGYVpYEYPSR	Y166		x				
CBL	115855	IKPSSpSANAIpYSLAAR	S669,Y674	Y674[17]				x	
CD3δ	4502669	NDQVpYQPLR	Y149				x	x	
CD3δ	4502669	DRDDAQpYSHLGGNWAR	Y160		x	x	x	x	x
CD3ε	4502671	DLpYSGLNQR	Y199	Y199[18]	x	x	x	x	x
CD3ζ	115997	NPQEGLpYNELQK	Y110	Y110[19]	x				
CD3ζ	115997	RKNPQEGLpYNELQK	Y110	Y110[19]			x		x
CD3ζ	115997	MAEApYSEIGMK	Y122	Y122[19]			x	x	
CD3ζ	115997	GHDGLpYQGLSTATK	Y141	Y141[19]	x	x	x	x	x
CD3ζ	115997	REEpYDVLDKR	Y83	Y83[19]				x	
CD5	7656965	pSHAENPTASHVDNEpYSQ PPRNpSR	S439,Y453, S460	Y453[20]				x	
GADS	4758476	RHpTDPVQLQAAGR	T262		x		x	x	
HS1	4885405	GFGGQpYGIQK	Y198				x	x	
LIM lipoma	5031887	YYEGYpYAAGPGYGGR	Y301						x
PYK2	4758976	YIEDEDpYpYKASVTR	Y579,Y580	Y579, Y580[21]			x		
ZAP-70	1177033	RIDTLNSDGpYTPEPAR	Y292	Y292[22]			x	x	x
ZAP-70	1177044	PMPMDTSVpYESPpYSDPEE LKDK	Y315,Y319	Y315, Y319[23]			x	x	
ZAP-70	1177044	ALGADDSpYpYTAR	Y492,Y493	Y492, Y493[22]			x		
ZAP-70	1177044	ALGADDSYpYTAR	Y493	Y493[22]			x	x	x
Jurkat Lck depleted (J.CaM1.6)									
PYK2	4758976	YIEDEDpYpYKASVTR	Y579,Y580	Y579, Y580[21]	x				

Times are after anti-mouse IgG treatment (TCR crosslinking). Shaded areas indicate time points not analyzed in the experiments.
Adapted with permission from A.R. Salomon et al, *Proc. Natl. Acad. Sci. USA*, 100(2), 443-448 (2003).

residing on 151 different peptides. A total of 80 tyrosine, 83 serine, and 19threonine phosphorylation sites were observed (Table 1)[6]. Most importantly,

this experiment established that all of these phosphorylation sites are associated with a single cell line from a single experiment and that our methodology is sensitive enough to see them. In contrast with traditional approaches, these phosphorylation sites were not inferred from site-directed mutagenesis or in vitro kinase reactions but were obtained directly from MS/MS sequencing of peptides derived from whole cell lysates. The fact that many of these phosphorylation sites are known to play critical roles in T cell signaling suggests that pervanadate stimulation is a useful technique that provides many biologically relevant phosphorylation sites.

In an initial approach to understand the signaling pathways involved in T cell activation, phosphorylation sites were comprehensively determined after T cell receptor stimulation as described in Salomon et al[14]. This experiment utilized an earlier version of the phosphoproteomic technology platform lacking the desalting column and using the less sensitive LCQ mass spectrometer. Phosphorylated peptides were isolated and sites of phosphorylation characterized by MS/MS sequencing after receptor activation (0, 0.5, 1, 2, and 5 minute time points)[14]. Up-regulation of ZAP-70 catalytic activity results from phosphorylation of Tyr^{493} by Lck and autophosphorylation of Tyr^{319} following tandem interactions of the ZAP-70 SH2 domains with immunoreceptor tyrosine-based activation motifs (ITAMs) on CD3ζ[23, 24]. Although we found fewer phosphorylation sites due to the older methodology used in this experiment, this time course experiment was able to show the sequential phosphorylation of CD3ζ ITAM phosphorylation sites prior to ZAP-70 phosphorylation with careful relative quantitation (as illustrated in Salomon et al[14]). All of the known *in vivo* tyrosine phosphorylation sites of ZAP-70 were detected (Tyr^{292}, Tyr^{315}, Tyr^{319}, Tyr^{492}, and Tyr^{493})[22,23]. Furthermore, tyrosine phosphorylation was effectively absent in the Lck-depleted mutant cell line J.CaM1.6 on CD3/CD4 stimulation (Table 2). Interestingly, despite the large volume of published work on T cell signaling, these experiments also identified unreported sites of tyrosine phosphorylation such as Tyr^{166} of Cas-L, which contains a strong consensus site for binding to the Lck SH2 domain. Cas-L is known to bind to Lck kinase in CD3-stimulated T cells, but the site of interaction has not been defined[25]; phosphorylation of a Lck SH2 site on Cas-L suggests that Lck could bind this site. Another novel phosphorylation site was observed at Tyr^{198} of HS1. HS1 is known to be tyrosine phosphorylated as a consequence of anti-CD3 stimulation, possibly through the action of Syk[26,27], but again the site of phosphorylation was not previously identified. Induced phosphorylation of CD3ζ, CD3δ, and CD3ε was consistent with the expected phosphorylation of these ITAM regions[19].

6. CONCLUDING REMARKS

Through fusion of innovations in high-throughput chromatographic separations of phosphopeptides, detection by mass spectrometry, and bioinformatic analysis, we have assembled a formidable tool to complement the traditional methodologies typically used to study signaling pathways. Although discovery of a large and dynamic set of cell-derived phosphorylation sites in a single proteomics experiment is an impressive accomplishment, understanding which of these sites participate in a given signaling pathway and the nature of this participation is the fundamental challenge confronting the proteomics researcher. It is simply not feasible to make hundreds of site-directed mutants and the associated knockout mice with current methodologies within a single lab in a reasonable amount of time. Therefore, complimentary high-throughput follow-up strategies must be developed to ascertain whether newly discovered phosphorylation sites participate in pathways and the nature of this participation. Our new approach streamlines the usual signaling pathway analysis paradigm by allowing for the production of mutants of phosphorylation sites and signaling proteins shown to exist within cells as opposed to the usual motif-driven site-directed mutagenesis approach. Also, phosphoproteomic characterization of perturbations of global phosphorylation patterns in mutant cells will likely provide a useful perspective on signaling pathways and ultimately a more rapid understanding of the molecular basis of a wide array of diseases.

7. REFERENCES

1. S. P. Gygi, Y. Rochon, B. R. Franza, and R. Aebersold, Correlation between protein and mRNA abundance in yeast, *Mol Cell Biol* **19**(3), 1720-1730 (1999).
2. M. P. Washburn, D. Wolters, and J. R. Yates, 3rd, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat Biotechnol* **19**(3), 242-247 (2001).
3. S. A. Beausoleil, M. Jedrychowski, D. Schwartz, J. E. Elias, J. Villen, J. Li, M. A. Cohn, L. C. Cantley, and S. P. Gygi, Large-scale characterization of HeLa cell nuclear phosphoproteins, *Proc Natl Acad Sci U S A* **101**(33), 12130-12135 (2004).
4. M. Mann, and O. N. Jensen, Proteomic analysis of post-translational modifications, *Nat Biotechnol* **21**(3), 255-261 (2003).
5. J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, and D. F. Hunt, Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proc Natl Acad Sci U S A* **101**(26), 9528-9533 (2004).
6. L. M. Brill, A. R. Salomon, S. B. Ficarro, M. Mukherji, M. Stettler-Gill, and E. C. Peters, Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry, *Anal Chem* **76**(10), 2763-2772 (2004).
7. S. B. Ficarro, A. R. Salomon, L. M. Brill, D. E. Mason, M. Stettler-Gill, A. Brock, and E. C. Peters, Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites, *Rapid Commun Mass Spectrom* **19**(1), 57-71 (2005).

8. M. J. MacCoss, C. C. Wu, and J. R. Yates, 3rd, Probability-based validation of protein identifications using a modified SEQUEST algorithm, *Anal Chem* **74**(21), 5593-5599 (2002).
9. Y. Oda, T. Nagasu, and B. Chait, Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome, *Nature Biotechnology* **19**, 379-382 (2001).
10. M. C. Posewitz, and P. Tempst, Immobilized gallium(III) affinity chromatography of phosphopeptides, *Anal Chem* **71**(14), 2883-2892. (1999).
11. S. B. Ficarro, M. L. McCleland, P. T. Stukenberg, D. J. Burke, M. M. Ross, J. Shabanowitz, D. F. Hunt, and F. M. White, Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae, *Nat Biotechnol* **20**(3), 301-305 (2002).
12. D. C. Neville, C. R. Rozanas, E. M. Price, D. B. Gruis, A. S. Verkman, and R. R. Townsend, Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry, *Protein Sci* **6**(11), 2436-2445 (1997).
13. A. Stensballe, S. Andersen, and O. N. Jensen, Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis, *Proteomics* **1**(2), 207-222 (2001).
14. A. R. Salomon, S. B. Ficarro, L. M. Brill, A. Brinker, Q. T. Phung, C. Ericson, K. Sauer, A. Brock, D. M. Horn, P. G. Schultz, and E. C. Peters, Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry, *Proc Natl Acad Sci U S A* **100**(2), 443-448 (2003).
15. Y. Zhang, A. Wolf-Yadlin, P. L. Ross, D. J. Pappin, J. Rush, D. A. Lauffenburger, and F. M. White, Time-resolved Mass Spectrometry of Tyrosine Phosphorylation Sites in the Epidermal Growth Factor Receptor Signaling Network Reveals Dynamic Modules, *Mol Cell Proteomics* **4**(9), 1240-1250 (2005).
16. S. Peri, J. D. Navarro, R. Amanchy, T. Z. Kristiansen, C. K. Jonnalagadda, V. Surendranath, V. Niranjan, B. Muthusamy, T. K. Gandhi, M. Gronborg, N. Ibarrola, N. Deshpande, K. Shanker, H. N. Shivashankar, B. P. Rashmi, M. A. Ramya, Z. Zhao, K. N. Chandrika, N. Padma, H. C. Harsha, A. J. Yatish, M. P. Kavitha, M. Menezes, D. R. Choudhury, S. Suresh, N. Ghosh, R. Saravana, S. Chandran, S. Krishna, M. Joy, S. K. Anand, V. Madavan, A. Joseph, G. W. Wong, W. P. Schiemann, S. N. Constantinescu, L. Huang, R. Khosravi-Far, H. Steen, M. Tewari, S. Ghaffari, G. C. Blobe, C. V. Dang, J. G. Garcia, J. Pevsner, O. N. Jensen, P. Roepstorff, K. S. Deshpande, A. M. Chinnaiyan, A. Hamosh, A. Chakravarti, and A. Pandey, Development of human protein reference database as an initial platform for approaching systems biology in humans, *Genome Res* **13**(10), 2363-2371 (2003).
17. H. Steen, B. Kuster, M. Fernandez, A. Pandey, and M. Mann, Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway, *J Biol Chem* **277**(2), 1031-1039. (2002).
18. I. de Aos, M. H. Metzger, M. Exley, C. E. Dahl, S. Misra, D. Zheng, L. Varticovski, C. Terhorst, and J. Sancho, Tyrosine phosphorylation of the CD3-epsilon subunit of the T cell antigen receptor mediates enhanced association with phosphatidylinositol 3-kinase in Jurkat T cells, *J Biol Chem* **272**(40), 25310-25318. (1997).
19. E. N. Kersh, A. S. Shaw, and P. M. Allen, Fidelity of T cell activation through multistep T cell receptor zeta phosphorylation, *Science* **281**(5376), 572-575. (1998).
20. J. M. Vila, I. Gimferrer, O. Padilla, M. Arman, L. Places, M. Simarro, J. Vives, and F. Lozano, Residues Y429 and Y463 of the human CD5 are targeted by protein tyrosine kinases, *Eur J Immunol* **31**(4), 1191-1198. (2001).
21. D. D. Schlaepfer, C. R. Hauck, and D. J. Sieg, Signaling through focal adhesion kinase, *Prog Biophys Mol Biol* **71**(3-4), 435-478 (1999).
22. J. D. Watts, M. Affolter, D. L. Krebs, R. L. Wange, L. E. Samelson, and R. Aebersold, Identification by electrospray ionization mass spectrometry of the sites of tyrosine phosphorylation induced in activated Jurkat T cells on the protein tyrosine kinase ZAP-70, *J Biol Chem* **269**(47), 29520-29529. (1994).

23. V. Di Bartolo, D. Mege, V. Germain, M. Pelosi, E. Dufour, F. Michel, G. Magistrelli, A. Isacchi, and O. Acuto, Tyrosine 319, a newly identified phosphorylation site of ZAP-70, plays a critical role in T cell antigen receptor signaling, *J Biol Chem* **274**(10), 6285-6294. (1999).
24. A. C. Chan, M. Dalton, R. Johnson, G. H. Kong, T. Wang, R. Thoma, and T. Kurosaki, Activation of ZAP-70 kinase activity by phosphorylation of tyrosine 493 is required for lymphocyte antigen receptor function, *Embo J* **14**(11), 2499-2508. (1995).
25. H. Kanda, T. Mimura, K. Hamasaki, K. Yamamoto, Y. Yazaki, H. Hirai, and Y. Nojima, Fyn and Lck tyrosine kinases regulate tyrosine phosphorylation of p105CasL, a member of the p130Cas docking protein family, in T-cell receptor-mediated signalling, *Immunology* **97**(1), 56-61. (1999).
26. J. E. Hutchcroft, J. M. Slavik, H. Lin, T. Watanabe, and B. E. Bierer, Uncoupling activation-dependent HS1 phosphorylation from nuclear factor of activated T cells transcriptional activation in Jurkat T cells: differential signaling through CD3 and the costimulatory receptors CD2 and CD28, *J Immunol* **161**(9), 4506-4512. (1998).
27. M. Ruzzene, A. M. Brunati, O. Marin, A. Donella-Deana, and L. A. Pinna, SH2 domains mediate the sequential phosphorylation of HS1 protein by p72syk and Src-related protein tyrosine kinases, *Biochemistry* **35**(16), 5327-5332. (1996).