BRIEF COMMUNICATIONS

Identification of WEHI-231 Phosphoproteins and Phosphorylation Sites Using IMAC and LC-MS/MS*

Shu H, Chen S, Bi Q, Mumby M, Brekken D

Alliance for Cellular Signaling Laboratories, University of Texas Southwestern Medical Center, Dallas, TX

Abstract: A goal of the Alliance for Cellular Signaling (AfCS) Protein Chemistry Laboratory is the identification of phosphoproteins in mouse B lymphocytes (B cells). In order to identify phosphoproteins on a proteome-wide basis, WEHI-231 cells were treated with calyculin A, a serine/threonine phosphatase inhibitor, or left untreated. Proteins were extracted from whole cell lysates and treated with trypsin to generate peptides. Phosphorylated peptides were enriched using immobilized metal affinity chromatography (IMAC) and identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS). A total of 107 proteins were identified using these methods. Forty-two of these proteins were previously reported to be phosphorylated, and 17 of them are on the AfCS protein list. The list also includes 11 proteins that appear to be completely novel. A total of 193 phosphorylation sites were identified within these proteins.

Introduction

An important goal of the AfCS Protein Chemistry Laboratory is the global analysis of ligand-induced changes in protein phosphorylation. Important steps in this process are the identification of phosphoproteins present in the AfCS model cell systems and the determination of their sites of phosphorylation. This information will be used to generate phospho-specific antibodies that will provide quantitative information on the effects of ligands on phosphorylation of specific sites. Recent improvements in phosphopeptide enrichment by immobilized metal affinity chromatography (IMAC) and in mass spectrometry make it possible to identify phosphoproteins on a proteome-wide basis (1). We used this approach to identify phosphorylation sites in whole cell lysates from the WEHI-231 B cell line. In this communication, we present the identification of phosphoproteins and their phosphorylation sites in control and calyculin A–treated WEHI cells.
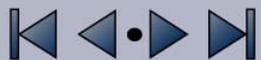
Methods

- WEHI-231 cells were either left untreated or treated with 100 nM calyculin A for 45 min according to AfCS protocol PS00000537.
- Cells were lysed and the protein was extracted using TriPure reagent according to protocol PP00000162, or the cells were lysed in a solution containing 0.1% sodium dodecyl sulfate (SDS; see Supplemental Methods).
- Immobilized metal affinity chromatography (IMAC) columns were prepared according to protocol PP00000163, or as described in Supplemental Methods, to compare different resins.
- Proteins were digested with trypsin and the phosphopeptides were enriched by IMAC according to protocol PP00000164.
- Proteins were identified by liquid chromatography and tandem mass spectrometry (LC-MS/MS) using a nanoscale C18 column coupled in-line with an ion trap mass spectrometer according to protocol PP00000164.
- The mass spectra and MS/MS data were used to search the nonredundant NCBI mouse protein database using SEQUEST software. Software parameters were set to detect a modification of 80 Da on Ser, Thr, or Tyr.
- The assignments of phosphopeptide sequences were then manually confirmed by comparing the acquired MS/MS spectra to the theoretical fragmentation pattern.

Flow chart of IMAC procedure for the enrichment of phosphorylated proteins from whole cell lysates

SDS Lysis

Construction of IMAC Micro-tip Columns

TriPure Lysis

Lyse cells in 0.1% SDS; 25 mM Tris-HCl, pH 7.5; 150 mM NaCl

Load 50% suspension of POROS 20 MC into a gel loader pipette tip to obtain 10 µl bed volume

Extract cell protein using TriPure reagent

Boil for 5 min

Charge resin with 100 mM GaCl$_3$

Solubilize protein in 6 M guanidine-HCl

Dilute 10-fold with 100 mM ammonium bicarbonate

Digest overnight with trypsin (37 °C)

Desalt using C18 cartridge

Load onto activated IMAC micro-tip column

Wash with 0.1% HAc; 50% ACN/0.1% HAc; 50% ACN/0.1% HAc/100 mM NaCl; and 0.1% HAc

Elute with 50 mM Na$_2$HPO$_4$

Load eluate onto the nanoscale C18 column

Desalt by washing with mobile phase A

Run LC-MS/MS in a data-dependent manner and use SEQUEST software to search database

Results

Generation of phosphorylated proteins by calyculin A treatment

To maximize detection and identification of phosphoproteins, cells were treated with the serine/threonine phosphatase inhibitor calyculin A (CLA). A Western blot with an anti-P-Thr antibody shows the increase in threonine phosphorylation upon CLA treatment (Fig. 1).
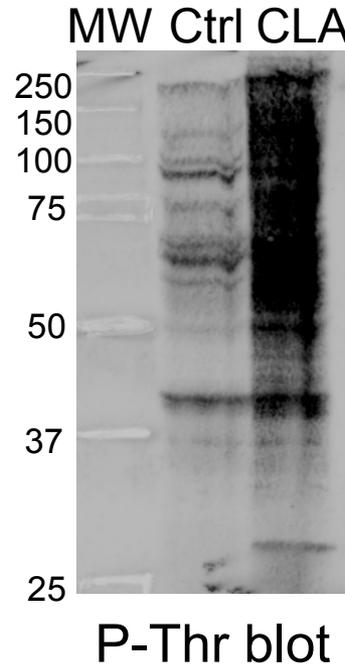
Fig. 1. Generation of phosphorylated proteins by calyculin A treatment. WEHI-231 cells were either left untreated (Ctrl) or treated with 20 nM CLA for 45 min (CLA) and lysed in SDS sample buffer. The proteins were resolved on a 10% SDS gel, transferred to a nitrocellulose membrane, and immunoblotted with anti-phosphothreonine polyclonal antibody P-Thr-polyclonal (Cell Signaling Technologies). The mass of each molecular weight marker (MW) is shown at the left.

Comparison of IMAC resins for the enrichment of phosphopeptides

To optimize enrichment of phosphopeptides, different metal chelating resins were charged with either iron or gallium and tested with a known mixture of phosphorylated and non-phosphorylated peptides. Of the resins tested, Ga^{3+}-charged POROS 20 MC and BRX-IDA resins resulted in the lowest background and highest recovery of each of the phosphopeptides (see Fig. 2). When compared to Fe^{3+}, the resins charged with Ga^{3+} resulted in a higher signal-to-noise ratio with our test sample.

Enrichment of phosphorylated peptides by IMAC

In order to test the IMAC procedure under more rigorous conditions, a test sample containing 1 pmol of a synthetic P-tyr phosphopeptide, 2 pmol alpha-casein, and 400 pmol of four non-phosphorylated proteins was used. Samples from each step of the procedure were monitored using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) (Fig. 3). The phosphopeptides were not detected in the total digest due to the abundance of non-phosphorylated peptides present. The total digest was loaded onto a Ga^{3+}-charged BRX-IDA column. None of the phosphopeptides were detected in either of the wash fractions. Analysis of the material eluted with 200 mM Na_2HPO_4 showed that all four phosphopeptides were recovered with very little contamination by non-phosphorylated peptides.

Identification of proteins enriched by IMAC

Proteins were identified using LC-MS/MS. The mass spectra and MS/MS data (Fig. 4) were used to search the nonredundant NCBI mouse protein database using SEQUEST software. Software parameters were set to detect a modification of 80 Da on Ser, Thr, or Tyr. The assignments of phosphopeptide sequences were then manually confirmed by comparing the acquired MS/MS spectra to the theoretical fragmentation pattern.
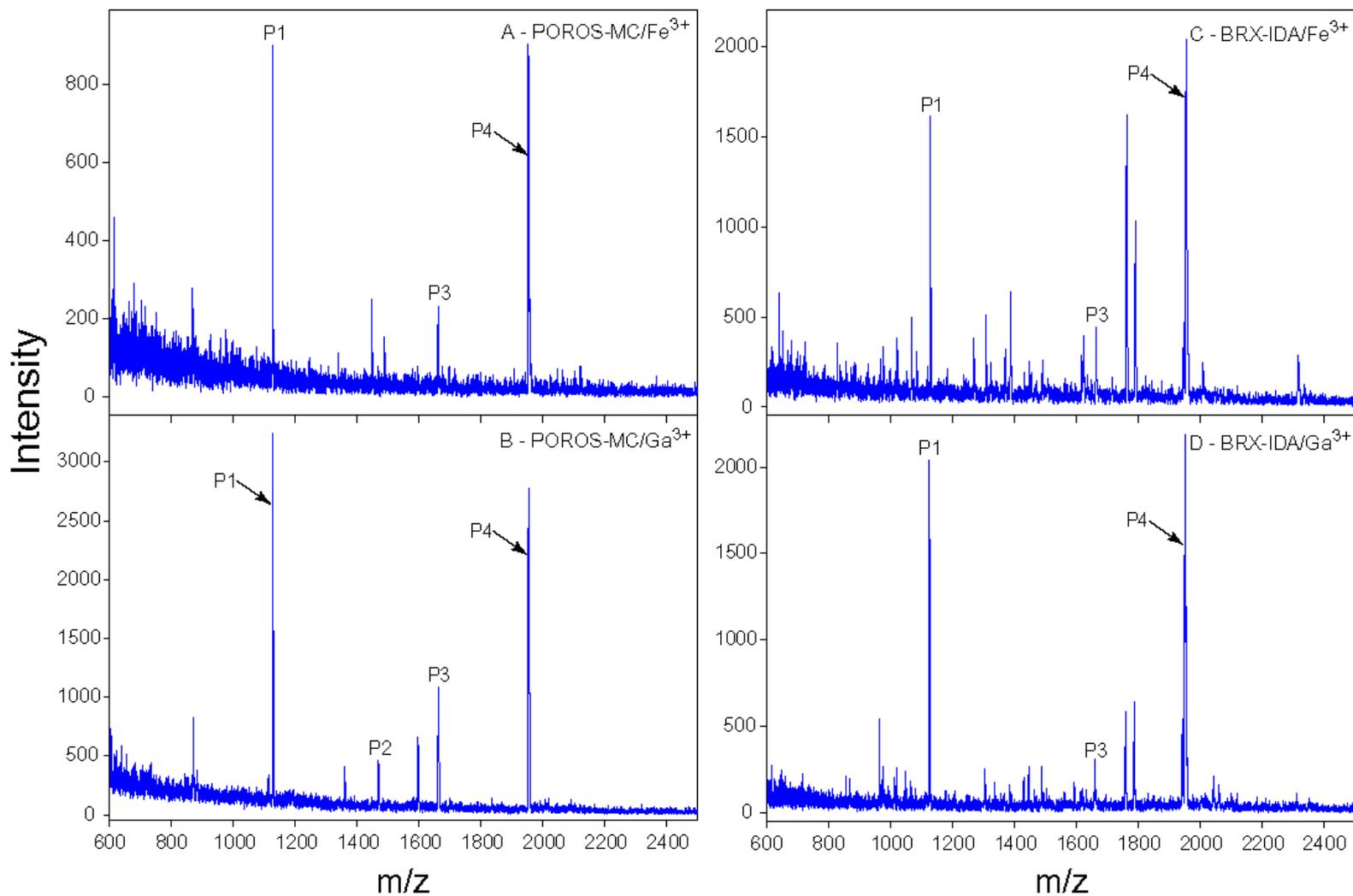
Fig. 2. (See Fig. legend on next page.)
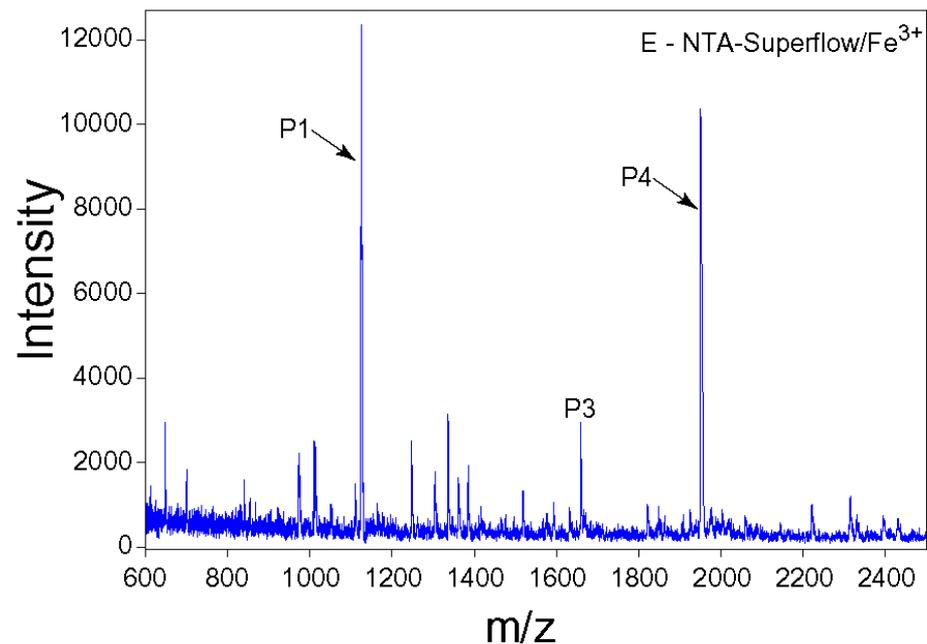
Fig. 2. Comparison of IMAC resins for the enrichment of phosphopeptides. The test sample was a standard mixture containing 1 pmol of a synthetic P-tyr phosphopeptide and a tryptic digest consisting of 1 pmol alpha-casein (containing three phosphopeptides) and 5 pmol each of four non-phosphorylated proteins (BSA, carbonic anhydrase, ubiquitin, and beta-lactoglobulin). The IMAC resins were (A) Fe^{3+}-charged POROS MC; (B) $Ga3^+$-charged POROS MC; (C) Fe^{3+}-charged BRX-IDA; (D) Ga^{3+}-charged BRX-IDA; and (E) Fe^{3+}-charged NTA-superflow. Samples were loaded onto IMAC columns prepared in micro-tips, washed, and eluted. The eluate was analyzed by MALDI-TOF mass spectrometry. The labeled peaks represent phosphorylated peptides; (P1) synthetic P-tyr phosphopeptide (m/z 1127, DRVpYIHPF); (P2) alpha-casein (m/z 1467, TVDMEpSTEVFTK); (P3) alpha-casein (m/z 1662, VPQLEIVPNpSAERR); and (P4) alpha-casein (m/z 1953, YKVPQLEIVPNpSAERR).

8
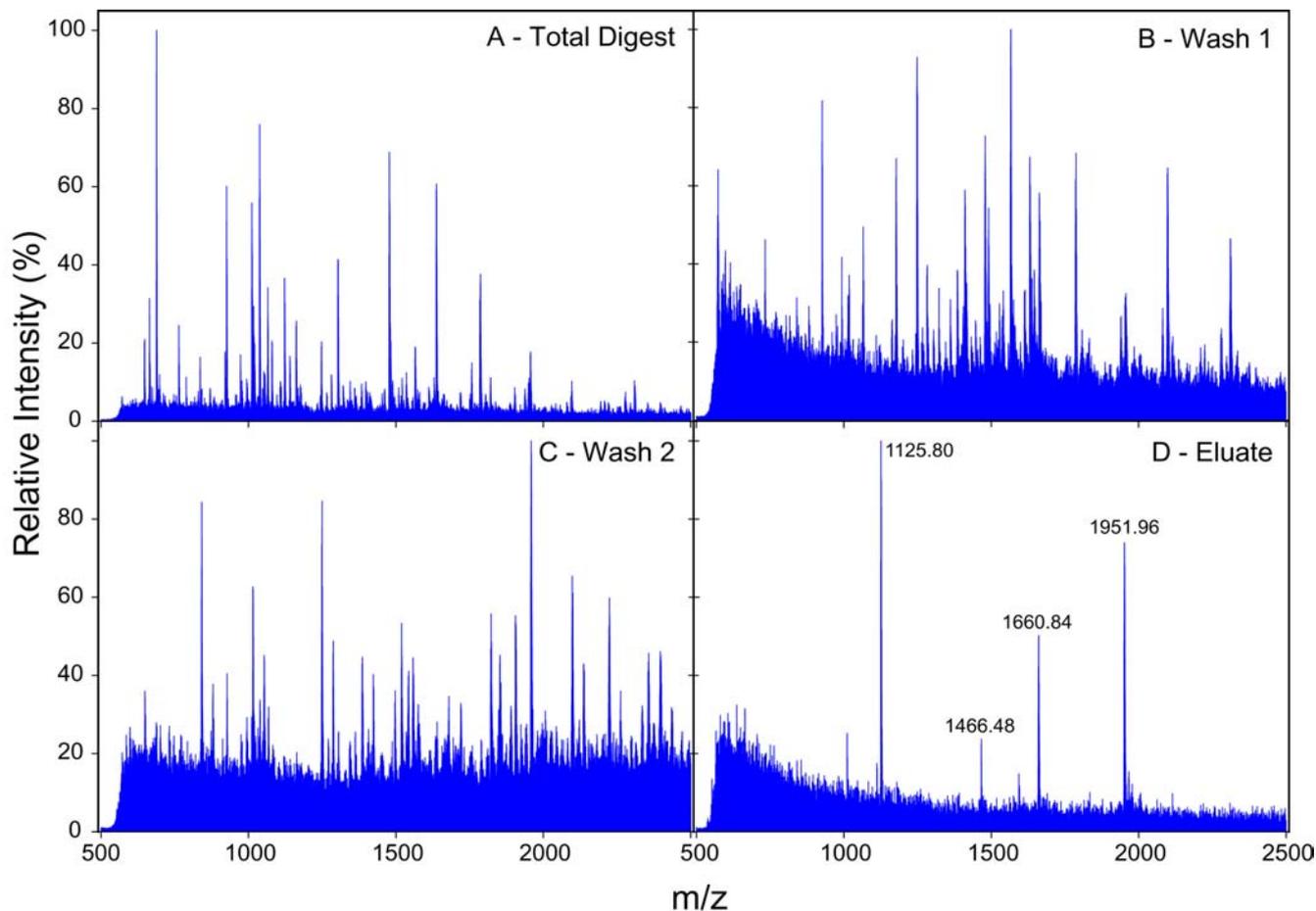
Fig. 3. Enrichment of phosphorylated peptides by IMAC. Peptides present in each fraction from the IMAC procedure were monitored by MALDI-TOF mass spectrometry. A standard mixture containing 1 pmol of a synthetic P-tyr phosphopeptide and a tryptic digest of 2 pmol alpha-casein and 400 pmol each of four non-phosphorylated proteins (BSA, carbonic anhydrase, ubiquitin, and beta-lactoglobulin) was precleaned with a C18 cartridge. The total digest (A) was loaded onto an IMAC micro-tip containing Ga^{3+}-charged BRX-IDA resin, washed with 50% ACN/0.1% HAc (B) and 50% ACN/0.1% HAc/100 mM NaCl (C), and eluted with 200 mM Na_2HPO_4 (D).
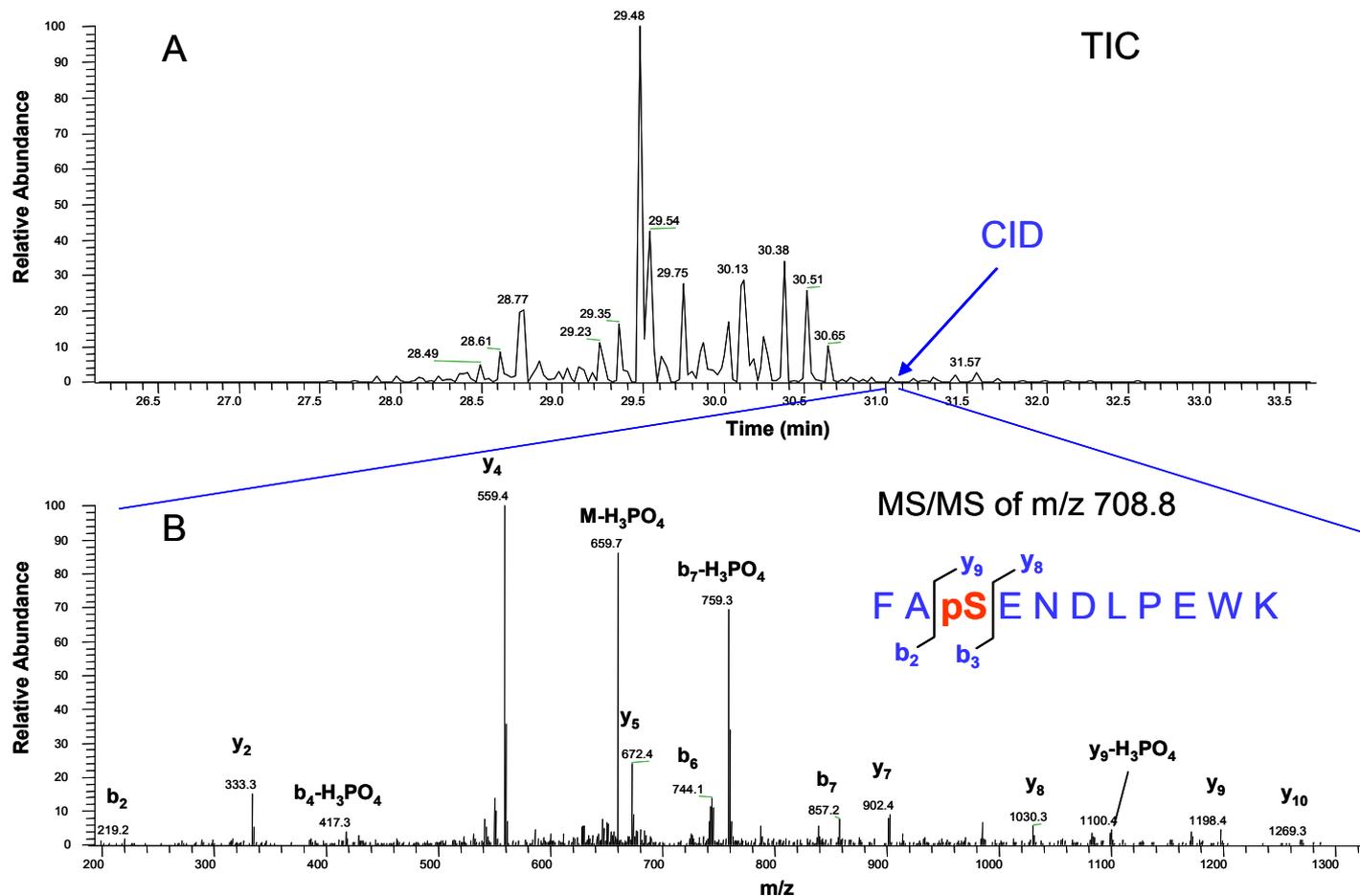
Fig. 4. LC-MS/MS analysis of peptides enriched by IMAC. The sample eluted from IMAC resin was loaded onto a capillary HPLC C18 column, washed to desalt, and coupled online with an ion trap mass spectrometer. The peptides were resolved by gradient elution and identified by tandem mass spectrometry. Panel A shows the total ion chromatogram (TIC) of the sample, and panel B shows the tandem mass spectrum (CID) of an ion with mass-to-charge ratio (m/z) of 708.8, eluting at 31 min. The peptide sequence FApSENDLPEWK from RAN binding protein 1 was derived from the fragment ion spectrum.

10

Identification of proteins enriched by IMAC

Table 1 (see link below) lists the proteins identified in five different experiments. A total of 107 proteins were identified, including 17 proteins on the AfCS Protein List. Forty-two of the identified proteins are previously characterized phosphoproteins (PubMed links to relevant publications are included in the table). Eleven of the proteins are novel proteins that have only been predicted from genomic or cDNA sequences.

One hundred and ninety-three distinct phosphorylation sites were identified. Only one of the sites identified was a phosphorylated tyrosine residue (dipeptidylpeptidase 8). All of the others were serine or threonine phosphorylation sites. This result is similar to those reported by others using IMAC methods (1, 2) and probably reflects the very low abundance of tyrosine phosphorylation relative to serine and threonine phosphorylation in most cells. One third of the total sites (69/193) were proline directed phosphorylation sites (pSer-Pro or pThr-Pro), suggesting that many of these proteins are phosphorylated by members of the cyclin-dependent kinase or MAP kinase families. Of the 147 distinct phosphopeptides identified, 62 contained more than one phosphorylated amino acid and 85 had a single phosphorylated residue. Based on the relative numbers of acidic (Asp, Glu) and basic (Lys, Arg) residues present in the peptides, 75 were acidic, 49 were basic, and 23 were neutral in the non-phosphorylated form.
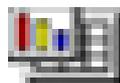
Table 1. *Phosphoproteins and phosphorylation sites identified in WEHI-231 cells by IMAC and LC-MS/MS*.

Discussion

The major goal of the work reported here was the identification of phosphoproteins and their sites of phosphorylation in WEHI-231 cells. Identification of these proteins is part of the larger AfCS effort to map signaling pathways. This information will be used to better define the signaling networks present in these cells and provide new probes (primarily phosphospecific antibodies) for quantitative measurements of signaling molecule activation during ligand-induced signaling.

IMAC coupled to LC-MS/MS has proven to be a powerful method to identify phosphoproteins present in complex mixtures of non-phosphorylated proteins. Of the resins tested, Ga^{3+}-charged POROS 20 MC and BRX-IDA resins resulted in the greatest recovery of the phosphopeptides with the lowest background. An advantage of the IMAC method is the simultaneous identification of phosphoproteins and their phosphorylation sites. Other groups have utilized the conversion of carboxylic acid groups to methyl esters to reduce nonspecific binding of acidic peptides to IMAC resins (1, 2, 3). We and others (4) have not found that esterification causes a marked enhancement of recovery of phosphopeptides from digests of total cell protein. However, we have found that esterification does enhance the detection and identification of phosphotyrosine peptides from samples previously enriched by immunoaffinity chromatography with anti-phosphotyrosine antibodies.

12

Discussion (continued)

A total of 107 proteins and 193 phosphorylation sites were identified using the IMAC procedure. The proteins included known phosphorylated proteins and proteins important in B cell signaling. However, Table 1 is not a complete list of phosphorylated proteins, and a number of known phosphoproteins were not identified in these experiments. While forty-two of the proteins were previously characterized phosphoproteins, sixty-five new phosphoproteins were identified in these experiments. The new phosphoproteins include 54 previously characterized proteins and 11 completely novel proteins that have only been inferred from cDNA or genomic sequencing.

This information will be used to expand the signaling "parts list" identified in these cells and allow the production of phosphospecific antibodies that will be used to monitor ligand-dependent changes in protein phosphorylation. The identification of these novel phosphoproteins should provide new avenues for investigating signaling pathways in these cells. Important issues will be the identification of the protein kinases and phosphatases that act on these sites and identification of factors that lead to changes in their levels of phosphorylation.

References

1. Ficarro SB, McCleland ML, Stukenberg PT, et al. (2002) Nat. Biotechnol. 20(3), 301-305. [PubMed]

2. Ficarro S, Chertihin O, Westbrook VA, et al. (2003) J. Biol. Chem. 278(13), 11579-11589. [PubMed]

3. Salomon AR, Ficarro SB, Brill LM, et al. (2003) Proc. Natl. Acad. Sci. U. S .A. 100(2), 443-448. [PubMed]

4. Nuhse TS, Stensballe A, Jensen ON, and Peck SC (2003) Mol. Cell Proteomics 2(11), 1234-1243. [PubMed]

Authors*	Editors
Concept, Design, & Supervision Deirdre Brekken[†‡] Hongjun Shu[†‡] Marc Mumby[‡] *Technical & Data Collection* Qun Bi[‡] She Chen[‡]	Ashley K. Butler *Duke University, Durham, NC* Gilberto R. Sambrano[‖] *University of California San Francisco, San Francisco, CA*
Acknowledgements	**Reviewers**
Robert Cox[‡] Farah El Mazouni[‡] Christine Horvath[§] Kathleen Lyons[‡] Deepa Sethuraman[‡]	Lewis Cantley *Harvard Medical School, Boston, MA*

* Please refer to the AfCS policy on authorship.
† To whom scientific correspondence should be addressed.
‡ AfCS Protein Chemistry Lab and § AfCS Cell Preparation and Analysis Lab
 University of Texas Southwestern Medical Center, Dallas, TX.
‖ To whom questions or comments about the *AfCS Communications* should be addressed.