JOURNAL OF PROTEOMICS 73 (2010) 1381–1390

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Phosphoproteomic analysis of primary human multiple myeloma cells

Feng Ge[a,1], *Chuan-Le Xiao*[a,1], *Xing-Feng Yin*[a], *Chun-Hua Lu*[a], *Hui-Lan Zeng*[b], *Qing-Yu He*[a,*]

[a] *Institute of Life and Health Engineering and National Engineering Research Center of Genetic Medicine, Jinan University, Guangzhou 510632, China*

[b] *Department of Hematology, The First Affiliated Hospital, Jinan University, Guangzhou 510632, China*

ARTICLE INFO

Article history:
Received 18 November 2009
Accepted 6 March 2010

Keywords:
Multiple myeloma
Phosphorylation
Phosphoproteome
Signaling pathway

ABSTRACT

Multiple myeloma (MM) is a malignant disorder of differentiated B cells. Clonal expansion of the tumor results in the excessive production of monoclonal immunoglobulin (Ig) which is a diagnostic feature of this disease. Previous investigations have demonstrated the alteration of the ERK, jun kinase, STAT, and AKT kinase signaling cascades in MM cells, suggesting that deregulated phosphorylation may contribute to MM pathogenesis. However, systematic analysis of the phosphoproteome in MM cells has not been reported. Here, we described a large-scale phosphorylation analysis of primary MM cells. Using a separation strategy involving immunomagnetic bead-positive selection of MM cells, preparative SDS-PAGE for prefractionation, in-gel digestion with trypsin, and titanium dioxide enrichment of phosphopeptides, followed by LC-MS/MS analysis employing a hybrid LTQ-Orbitrap mass spectrometer, we were able to catalog a substantial portion of the phosphoproteins present in primary MM cells. This analysis led to the identification of 530 phosphorylation sites from 325 unique phosphopeptides corresponding to 260 proteins at false positive rate (FPR) of 1.3%. This dataset provides an important resource for future studies on phosphorylation and carcinogenesis analysis of multiple myeloma.

1. Introduction

Essentially incurable with conventional or high-dose chemotherapy, multiple myeloma (MM) is a clonal B-cell hematological malignancy described as the latent accumulation in bone marrow (BM) of secretary plasma cells with a low proliferative index and a prolonged life span [1,2]. Enhanced proliferation and resistance to apoptosis account for the expansion of the MM malignant clone [3]. Previous investigations have demonstrated the alteration of the ERK [4], jun kinase [5], STAT [6], and AKT kinase [7] signaling cascades in MM cells and implicated the pathways in clonal expansion, suggesting that deregulated phosphorylation may contribute to MM pathogenesis.

Protein phosphorylation–dephosphorylation events play a primordial role in cell functions. Reversible phosphorylation of serine, threonine, and tyrosine regulates a variety of biological processes including intercellular communication, cell growth, proliferation, differentiation, and apoptosis. Characterization of phosphorylation status is critical to the elucidation of signal transduction pathways and to the understanding of the mechanisms of disease and drug actions [8,9]. It has been estimated that about 2–5% of the human genome codifies for kinases and phosphatases [10]. Protein

* *Corresponding author*. Institute of Life and Health Engineering, Jinan University, Guangzhou 510632, China. Tel./fax: +86 20 85227039.
E-mail address: tqyhe@jnu.edu.cn (Q.-Y. He).

[1] These authors contributed equally to this work.

1874-3919/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jprot.2010.03.004

phosphorylation is one of the most frequent posttranslational modifications (PTMs) in eukaryotic cells, where about one-third of all proteins in eukaryotic cells are phosphorylated at any given time [11].

Phosphoproteomics usually refers to a large-scale analysis of protein phosphorylation using mass spectrometry (MS)-based strategies [12–14]. Recent successes in this area owe much to the development of MS instrumentation such as the linear ion trap and the orbitrap, as well as novel phosphoprotein/peptide enrichment techniques including immobilized metal affinity chromatography (IMAC) [15,16], strong cation exchange chromatography (SCX) [17], or the two in combination [18].

Recently, titanium dioxide (TiO_2) chromatography has been adapted as an efficient method for enriching phosphopeptides from complex samples [19–21]. TiO_2 has a high selectivity for phosphorylated peptides and unspecific binding with non-phosphorylated peptides is reduced by including 2,5-dihydroxybenzoic acid (DHB), phthalic acid or glycolic acid and high concentration of trifluoroacetic acid (TFA) in the loading buffer [22]. In addition, TiO_2 chromatography of phosphorylated peptides is tolerant toward most buffers and salts used in biochemistry and cell biology laboratories [23].

Here, we described the phosphoproteome analysis of MM cells using TiO_2 enrichment directly coupled with the LC-MS/MS approach. After database search, stringent filtering, and manual validation of neutral loss in the MS/MS spectra, a total of 530 phosphorylation sites on 325 phosphopeptides were identified in the primary MM cells, including 417 on serine, 96 on threonine and 17 on tyrosine residues. To our knowledge, this is the first large scale phosphoproteomic analysis conducted on primary human MM cells to date.

2. Materials and methods

2.1. Experimental strategy

Large-scale phosphoproteomic studies have been hindered by the need for developing reliable methods to selectively enrich low-abundant phosphoproteins and phosphopeptides. Recent advances in phosphopeptide enrichment strategies made it possible for large-scale phosphorylation site analyses. The objective of this study is to investigate and optimize experimental conditions for large-scale proteomic profiling of peptide phosphorylation status in primary MM cells. The workflow in this study is outlined in Fig. 1 and the detailed experimental procedures were as follows.

2.2. Patient samples

Bone marrow (BM) aspirates were obtained from nine patients with newly diagnosed MM. Informed consent was obtained in accordance with the institutional policies and the Declaration of Helsinki protocol. Mononuclear cells were isolated from BM biopsies by purification over a Ficoll-Paque (Amersham, Piscataway, NJ, USA) gradient centrifugation. Briefly, BM aspirates (approximately 15 mL) were diluted to 1:1 with prewarmed (37 °C) PBS and overlaid onto 15 mL prewarmed Ficoll-Hypaque. After centrifugation at 2000 rpm for 20 min at room temperature, mononuclear cells were removed and

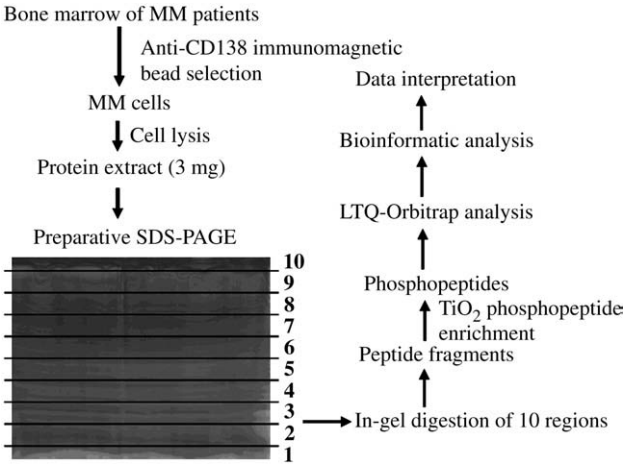

Fig. 1 – Workflow of the experiment to analyze the phosphoproteome in the primary human multiple myeloma cells.

washed with PBS. Plasma cell isolation from mononuclear cell fraction was performed by immunomagnetic bead selection with monoclonal mouse anti-human CD138 antibodies in association with LS separation columns separation system (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany). The purity of the isolated MM cells was monitored by BD-FACScalibur flow cytometry (BD Biosciences, San Jose, CA, USA) using CD138+/CD45- and CD38+/CD45-criteria [24]. The feasibility of this cell purification method has been tested in other reports [25]. MM cells represent an optimal model for phosphoproteomic research because they can be obtained as a pure cell population in high yield. As shown in Fig. 2, cell purity of more than 95% homogeneity was confirmed.

2.3. Sample preparation

Purified MM cells were washed three times with ice-cold washing buffer (10 μM Tris–HCl, 250 μM sucrose, pH 7.0) and transferred to a clean 1.5 mL Eppendorf tube. Cells were lysed with a buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 0.2 mg/mL PMSF, phosphatase inhibitors cocktail (Roche Diagnostics, Germany) and protease inhibitors (Complete tablets, Roche Diagnostics). Cellular debris was removed by centrifugation for 30 min at 13,200 g and at 4 °C. Protein concentrations were determined using Bradford assay. The protein sample (~3 mg) was resolved on a 10% SDS-PAGE gel (1.5 mm thick, 35 mm wide, 30 mm long), stained with Coomassie blue G250 and excised into 10 gel pieces for in-gel digestion. To prevent proteins from precipitation in the gel, the SDS concentration was adjusted from 0.1% to 0.2% in the gel and the running buffer. Gel lanes to be analyzed were excised from SDS-PAGE gels by razor blade and divided into 10–3 mm slices. Each slice was then further divided into ~1-mm^3 pieces. Each section was washed in water and completely destained using 100 mM ammonium bicarbonate in 50% acetonitrile. A reduction step was performed by addition of 100 μL of 10 mM DTT at 37 °C for 3 h. The proteins were alkylated by adding 100 μL of 50 mM iodoacetamide and allowed to react in the dark at 20 °C for 30 min. Gel sections were first washed in

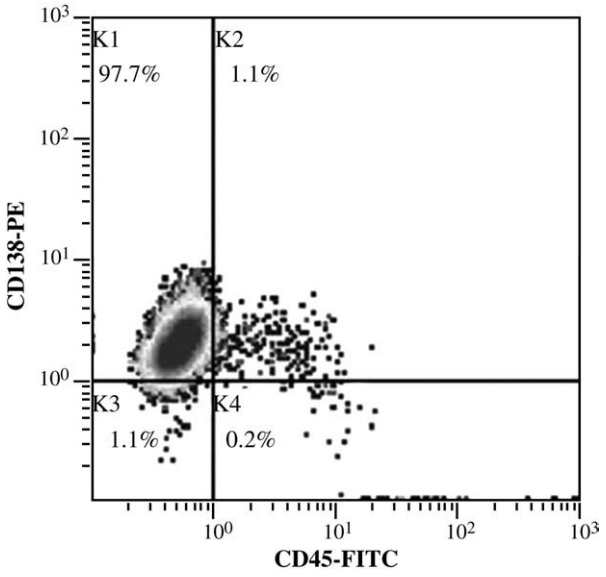

Fig. 2 – An example of flow cytometry reports of CD138-enriched bone marrow primary cells. Antibodies used in individual FACS analyses are indicated to the left and the bottom of each FACS plot. The percentage of cells in each gate window is indicated.

water, then acetonitrile, and finally dried by SpeedVac (Thermo Fisher Scientific, Waltham, MA) for 30 min. Digestion was carried out using 20 μg/mL sequencing grade modified trypsin (Promega, USA) in 50 mM ammonium bicarbonate. Sufficient trypsin solution was added to swell the gel pieces, which were kept at 4 °C for 45 min and then incubated at 37 °C overnight. The supernatants were transferred into a 200 μL microcentrifuge tube and the gels were extracted once with extraction buffer (67% acetonitrile containing 2.5% trifluoroacetic acid). The peptide extract and the supernatant of the gel slice were combined and then completely dried in a SpeedVac centrifuge.

2.4. *Phosphopeptide isolation using TiO_2 resin*

Phosphopeptides from digested peptides were enriched by using Phosphopeptide Enrichment TiO_2 kit (Calbiochem, San Diego, CA) according to the manufacturer's instruction with slight modifications. Briefly, the tryptic digest was dried, re-dissolved in 200 μL TiO_2 Phosphobind buffer containing 50 g/L 2,5-dihydroxybenzoic acid and then mixed with 50 μL TiO_2 Phosphobind Resin. After 30 minutes incubation, the supernatant was discarded, and TiO_2 was washed three times with the wash buffer. After that, the elution buffer was added two times to elute the phosphopeptides. The elutions were combined and dried using a Speed-Vac Concentrator and reconstituted in 2% ACN/1% TFA for LC-MS/MS analysis. All the buffers and the phosphoprotein purification resin were provided in the kit by the manufacturer.

2.5. *Peptide analysis by the LC-MS/MS approach*

Dried phosphopeptides were analyzed with a Finnigan Surveyor HPLC system coupled online with a LTQ-Orbitrap mass spectrometer (Thermo Electron, San Jose, CA) equipped with a nanospray source. Briefly, the peptide mixtures were loaded onto a C18 column (100 μm i.d., 10 cm long, 5 μm resin from Michrom Bioresources, Auburn, CA) using an autosampler. Peptides were eluted with a 0–35% gradient buffer solution (Buffer A, 0.1% formic acid, and 5% ACN; Buffer B, 0.1% formic acid and 95% ACN) over 120 min and then online detected in the LTQ-Orbitrap mass spectrometer using a data-dependent TOP5 method [26]. The general mass spectrometric conditions were: spray voltage, 1.85 kV; no sheath and auxiliary gas flow; ion transfer tube temperature, 200 °C; 35% normalized collision energy using for MS2. Ion selection thresholds were: 1000 counts for MS2. An activation $q=0.25$ and activation time of 30 ms were applied in MS2 acquisitions. The mass spectrometers were operated in positive ion mode with a data-dependent automatic switch between MS and MS/MS acquisition modes. For each cycle, one full MS scan in the Orbitrap at 1,000,000 AGC target was followed by five MS/MS (MS2) in the LTQ at 5000 AGC target on the five most intense ions. Selected ions were excluded from further selection for 90 s. Maximum ion accumulation times were 500 ms for full MS scans and 100 ms for MS2 scans. For the DDNL (data-dependent neutral loss) MS3 method, a MS3 was triggered if a neutral loss peak at −98.0, −49.0, −32.7 or −24.5 Da was observed in the MS2 and that peak was one of the five most intense ions of the MS2 spectra.

2.6. *Database analysis and manual evaluation of mass spectra*

Peak lists for the database search were produced in the Mascot generic format using BioWorks 3.3.1 (Thermo Finnigan, San Jose, CA) and DTASuperCharge V 1.31 (SourceForge), and the derived peak lists were searched using the Mascot 2.2.04 search engine (Matrix Science, London, UK) against a real and false IPI human database (V3.56, including 153078 protein entries). The following search criteria were employed: full tryptic specificity was required; two missed cleavages were allowed; Carbamidomethylation was set as fixed modification, whereas Oxidation (M), Phospho (ST), and Phospho (Y) were considered as variable modifications. Precursor ion mass tolerances were 10 ppm for all MS acquired in the Orbitrap mass analyzer, fragment ion mass tolerance was 0.5 Da for all MS^2 spectra acquired in the LTQ. Mass spectra of identified phosphopeptides with peptide score >10 were further processed and validated with the MSQuant 1.5 software for post-translational modification (PTM) score analysis [27]. The following criteria were adopted for phosphopeptide identification: (1) Peptide score threshold was 17; (2) The total threshold of PTM score and peptide score was 36; (3) All p-Ser and p-Thr peptides were required to show a pronounced neutral loss of phosphoric acid from the precursor ion and/or fragment ions or trigger the neutral loss-dependant MS^3 scan [28,29]; (4) At least three consecutive b- and/or y-ion series were required, or two consecutive b- and/or y-ion series and extensive coverage of b- and/or y-ion series was required. The final estimated false positive rate based on the decoy database search was 1.3%. Annotated MS/MS spectra of all verified phosphopeptides are presented in Supplemental_Table_S1. (http://life-health.jnu.edu.cn/phospho/MMphosphodata/Supplemental_Table_S1.xls).

2.7. Bioinformatics analysis

The identified phosphoproteins were classified based on the PANTHER (Protein ANalysis THrough Evolutionary Relationships) system (http://www.pantherdb.org), a unique resource that classifies genes and proteins by their functions [30]. In the process, the PANTHER ontology, a highly controlled vocabulary (ontology terms) by biological process, molecular function and molecular pathway, was used to categorize proteins into families and subfamilies with shared functions. The cellular localization of identified phosphoproteins was assessed using Gene Ontology database (http://www.geneontology.org). The interaction network of phosphoproteins was generated with Pathway Studio version 5.0 software (Ariadne Genomis, Rockville, MD) and Resnet 5 database [31]. Common downstream targets or upstream regulators of multiple proteins were identified by using this software, which facilitated the process of selecting key factors and potential mechanisms from the large number of phosphoproteins. The dataset was compared to the public database of PhosphoSite (http://www.phosphosite.org/) to find out the novel phosphoproteins and phosphosites. Each confirmed phosphoprotein was searched with SCANSITE (http://scansite.mit.edu) [32] for potential kinase motifs with high, medium, and low stringency.

3. Results and discussion

3.1. Identification of phosphorylation proteins and sites

After LC-MS/MS analysis on the enriched phosphopeptides, all MS/MS spectra were searched, respectively, against the forward and reversed human protein sequence databases to estimate rates of false-positive matches. Search results were filtered based on peptide score of MASCOT and PTM score. A total of 1002 phosphopeptides (redundant) from the target database passed our criteria, allowing 13 decoy matches. The phosphopeptide false-positive rate was therefore estimated to be 1.3% (see Materials and methods section). Multiple filtering criteria were established to validate search results. For each of the phosphorylated peptides identified in this work, peptide sequences were manually confirmed. After validation we identified 530 unique phosphorylation sites from 325 unique phosphopeptides corresponding to 260 protein groups. This entire dataset is provided as Supplemental Table S1 in which hyperlinks were built up to view all the MS/MS spectra.

The phosphopeptides identified from human MM cells with precise sites were classified by site p value. The distributions of phosphorylated serine, threonine, and tyrosine sites including class I , class II and class III sites are 78.6% (417/530), 18.1% (96/530), and 3.2% (17/530), respectively. These results are in line with those from a previous study on a variety of cell types: the overall level of phosphotyrosine and phosphothreonine in proteins was low compared to the level of phosphoserine in human cells [33].

To further investigate the reliability of the results, PhosphoSite (http://www.Phosphosite.org) was used to distinguish known phosphorylation sites from novel phosphorylation sites. Among the identified 530 phosphorylated sites, 74.9% were also reported by others previously (for details, see Supplemental_Table_S2 (http://life-health.jnu.edu.cn/phospho/MMphosphodata/Supplemental_Table_S2.xls). In other words, many of the phosphorylation sites were also determined by researchers with other cancer cells, indicating that the phosphorylation sites detected by current strategy are of high confidence.

3.2. Functional analysis of identified phosphoproteins

In order to understand the biological relevance of phosphoproteins, PANTHER classification system was used to categorize the identified phosphoproteins according to their molecular functions in biological processes. The PANTHER classification analysis revealed that the phosphoproteins can be classified into 26 collections according to their biological processes and 26 groups according to their molecular functions (Figs 3A and B). Fig. 3 provides an overview of the human MM proteome based on the known or postulated functions or biological processes of the identified phosphoproteins.

The nucleic acid binding proteins represent the largest group in all the identified phosphoproteins. Nucleic acid binding proteins are very important constituents of proteomes of all species; they play crucial roles in transcription, replication, recombination, repairing, and other activities. Since gene expression is regulated during further processing of pre-mRNA into mRNA that is ready to be translated in the cytosol, proteins that bind to RNA during the processing may thus be involved in controlling gene expression. One such well-characterized group is the splicing regulatory proteins (SR proteins) [34]. The present study identified three SR proteins, namely, SFRS2, SFRS8 and SFRS9. It appears that the phosphorylation state of SR proteins changes as splicing proceeds [35]. Phosphorylation is essential for the SR proteins to function in constitutive and regulated pre-mRNA splicing and SR protein phosphorylation must be tightly regulated in the cell. Antitumor drugs, such as indolocarbazoles have been shown to act by inhibiting SR phosphorylation [36]. Given these findings, we postulate that these SR proteins may play significant roles in the pathogenesis of MM and indolocarbazoles could be used as the potential anti-MM drugs. In connection with this, we are investigating the effects of indolocarbazoles on MM cells.

Several translation factors, including three initiation and two elongation factors were identified in the present study. The process of translation is very complex, involving several translational factors that control three distinct steps of translation i.e., initiation, elongation and termination. Regulation of gene expression at the level of translation has attracted much attention, particularly in connection to cancer, with the observation that disruption by mutation or overexpression of translational factors can cause cellular transformation [37]. A number of reports have been published showing that alteration in expression of eIF2 and eIF4 translational factors is often associated with human carcinogenesis [38]. In spite of the evidence implicating the involvement of certain translation factors in neoplastic developments, the underlying mechanisms responsible for their oncogenic potential are less understood. Although no direct link of the identified factors with MM exists in the literature, there are a few of candidates that may be considered as relevant for pathogenesis of MM. For

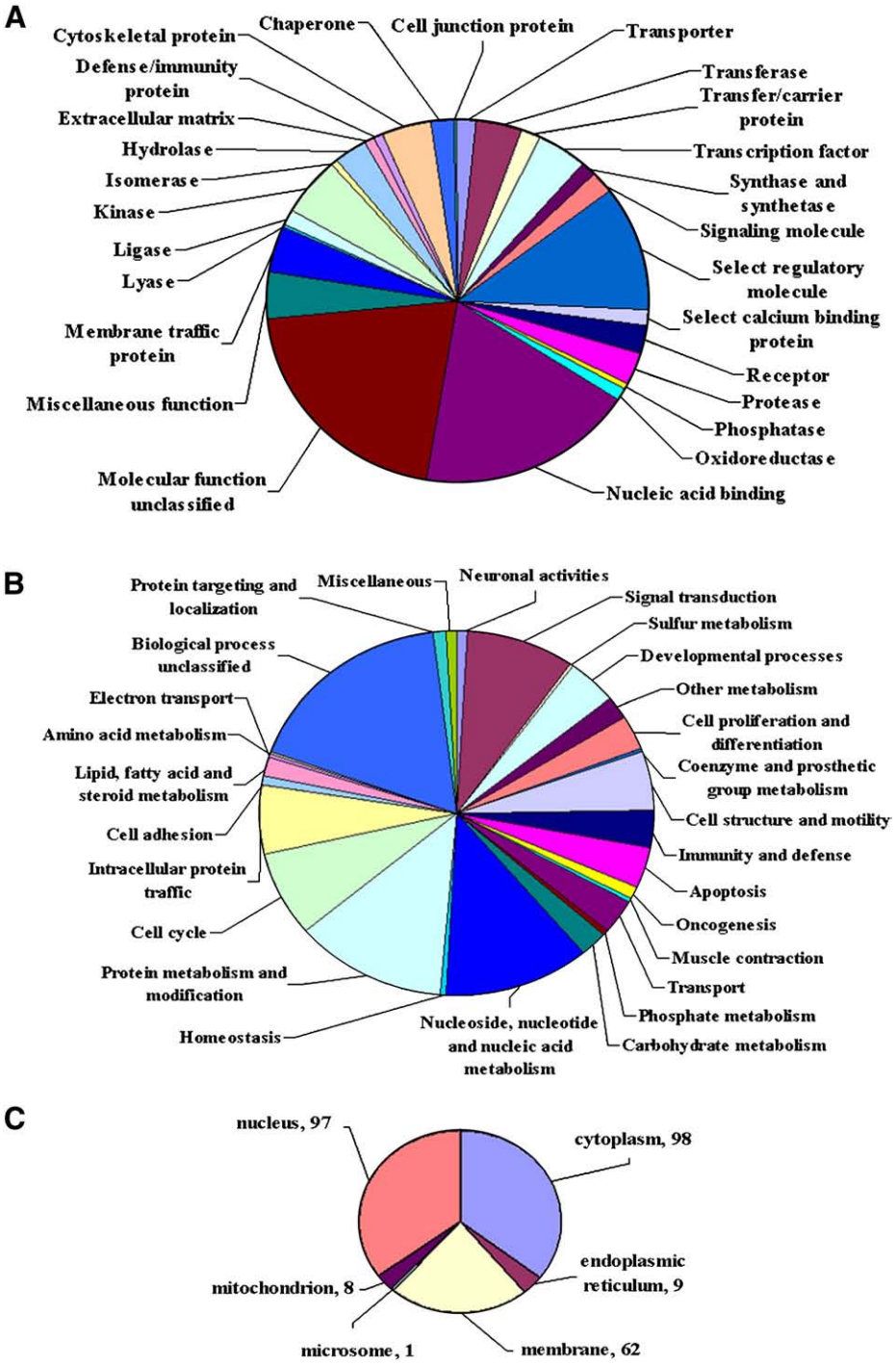

Fig. 3 – Pie chart representations of the distribution of identified phosphoproteins according to their (A) molecular functions (B) biological processes (C) cellular localization. Categorizations were based on information provided by the online resource PANTHER classification system and GO categorization system.

example, eEF2 can be phosphorylated by eEF2-kinase with activity increasing in several malignancies and non-specific inhibitors of this enzyme promoted cell death [39]. It can be therefore speculated that these translation factors could be the potential drug targets against MM.

Sixteen different kinases were identified in the current experiments. Among them, 13 proteins are protein kinases, including protein kinase C, 3-phosphoinositide dependent protein kinase-1(PDPK-1), p21(CDKN1A)-activated kinase 4 (PAK4), Cdc2-related kinase (CRKRS), B-Raf proto-oncogene serine/threonine-protein kinase (BRAF), etc. In human cells, most protein phosphorylation events are catalyzed by more than 500 members of the protein kinase superfamily of enzymes, also known as the human kinome [40]. Thus, protein kinases take the center stage in the complex signaling networks regulating a wide range of biological processes

such as cell proliferation, differentiation, and survival [41,42]. Many members of the protein kinase superfamily are positively or negatively regulated by either interacting with proteins or through specific phosphorylation events mediated by other protein kinases. As the phosphorylation states of protein kinases can provide a readout for the signaling activities within a cellular system, protein kinases identified in this study would provide informative insights into the cellular signal transduction networks of MM cells.

The cytoskeletal proteins represent the fourth largest group in the identified phosphoproteins. The cytoskeleton is a highly complex and dynamic system comprising structural proteins forming polymers (actin, tubulin and intermediate filaments) and several associated proteins with regulatory functions. In addition to their well-known structural function, cytoskeletal proteins play important roles in cell mobility and migration, immunological synapse formation and apoptosis [43]. Tumor-associated changes in the cytoskeleton are well documented and even utilized in cancer diagnostics [44]. The present study identified 12 different cytoskeletal proteins, which can be divided into three groups, namely, actin binding cytoskeletal proteins, intermediate filament and microtubule family cytoskeletal proteins. Six different actin-binding protein members were also identified, including cortactin, myosin, L-plastin, vasodilator-stimulated phosphoprotein (VASP), etc. The actin-binding protein family represents a large number of cytoskeletal proteins with a crucial role in the regulation of microfilaments implicating in many pathologies [45]. For example, Vasodilator-stimulated phosphoprotein (VASP) is an actin regulatory protein that links signaling pathways to remodeling of the cytoskeleton. VASP effects on actin turnover are regulated by phosphorylation, and in vitro studies have identified three serine/threonine phosphorylation sites (Ser-157, Ser-239, and Thr-278) in the protein [46]. Both in vitro phosphorylated VASP and VASP mutants mimicking defined phosphorylation states suggested that phospho-VASP interferes with actin filament assembly, turnover, and branching. Mechanistically, VASP phosphorylation seems to modulate F-actin binding, to reduce actin-polymerization-promoting activity [47], and to interfere with anticapping activity [48]. Originally, VASP phosphorylation was described in platelets, where both cAMP- and cGMP-dependent protein kinases (PKA and PKG, respectively) were shown to phosphorylate residues Ser-157 and Ser-239 [49]. Subsequently, the importance of phospho-Ser-157 (pSer-157) and phospho-Ser-239 (pSer-239) in cyclic-nucleotide dependent-kinase signaling cascades was established in other cardiovascular cells including aortic smooth muscle cells [50] and endothelial cells [51], in vessels of animal models, and in humans [52]. Very recently, it was reported that VASP functions are precisely regulated by differential phosphorylation and differential VASP phosphorylation controls remodeling of the actin cytoskeleton [53]. The functional implication of the VASP phosphorylation in MM is not clear and thus deserves further investigation.

The identified chaperone class comprises six members, including three heat shock proteins (HSPs) and nucleophosmin. HSPs are the products of several distinct gene families that are required for cell survival during stress. Different classes of HSPs play diverse roles in governing proper protein assembly, folding, and translocation. Regulation of HSP synthesis creates a unique defense system to maintain cellular protein homeostasis and to ensure cell survival [54]. For example, HSP90 is an emerging therapeutic target that may be of interest for the treatment of MM. Its role in protein homeostasis and the selective chaperoning of key signaling proteins in cancer survival and proliferation pathways has made HSP90 an attractive target of small molecule therapeutic intervention [55]. Phosphorylation of HSP90 on Ser and/or Thr residues has clearly been shown to be involved in its function. Hyperphosphorylation at these sites has been shown to result in the inhibition of the chaperoning function of HSP90 [56,57]. Thus, phosphorylation/dephosphorylation of HSPs represents a key regulatory mechanism for their functions in MM.

The identified proteins categorized into other functional groups include isomerase (0.7%), miscellaneous function (4.3%), transferase (4.0%), transfer/carrier protein (1.7%), transporter (1.7%), synthase and synthetase (1.3%), extracellular matrix (0.7%), phosphatase (0.7%), defense/immunity protein (1.0%), transcription factor (4.0%), hydrolase (2.7%), ligase (1.3%), receptor (2.7%), select regulatory molecule (11.0%), membrane traffic protein (4.0%), select calcium binding protein (1.3%), lyase (0.3%), oxidoreductase (1.0%), protease (2.7%) and signaling molecule (2.0%), cell junction protein (0.3%). Other identified proteins with their subfamilies not being classified were categorized as “molecular function unclassified” (20.7%).

As shown in Fig. 3C, a total of 181 phosphoproteins were annotated as being associated with cellular localization. Of these, 97 occur in the nucleus and 98 occur in the cytoplasm, with the remainder occurring in the membrane, mitochondrion, microsome and endoplasmic reticulum. These observations agree with previous findings that nuclear and cytoplasm proteins are subjected to a greater occurrence of phosphorylation events [27]. Although membrane proteins are often of low abundance, there were about 62 phosphoproteins identified as membrane proteins in this study. Most of the proteins identified belong to several GO cellular categories, and for this reason, the sum of the proteins listed in the different compartments is higher than 260 (Fig. 3C).

3.3. *Analysis of the novel identified phosphorylation sites*

The dataset (all 260 phosphoproteins and the 530 sites) in this work was compared to the public database of PhosphoSite (http://www.phosphosite.org/). Approximately 12.3% (29/260) of the phosphoproteins and 24.7% (131/530) of the phosphorylation sites identified appeared to be novel (Supplemental Table S2). Some of these novel phosphoproteins and phosphorylated sites may play important roles in the organism. For example, heterogeneous nuclear ribonucleoprotein K (hnRNP K) is a 65-kDa nuclear phosphoprotein that acts as a docking platform to integrate signals from a variety of kinase-mediated signal transduction pathways controlling nucleic acid-directed processes [58]. Cytoplasmic accumulation of hnRNP K via ERK mediated phosphorylation of hnRNP K S284 and S353 has been reported in cervical carcinoma HeLa cells [59] and chronic myelogenous leukemia cells [60]. Cytoplasmic translocation of hnRNP K is critical to its ability to regulate translation and promote migration of fibrosarcoma

cells [61]. As reported, S284 phosphorylation plays an important role in hnRNP K localization to the nucleus, possibly through altered conformation of the hnRNP K nuclearshuttling (KNS) domain which confers bi-directional transport across the nuclear membrane [59]. hnRNP K phosphorylation may also affect other members of the hnRNP family, including E1, E2, I and L, which may contribute to the diverse functions in which K protein has been implicated [62,63]. Importantly, hnRNP K can upregulate c-myc expression through the c-myc internal ribosome entry segment (IRES) in MM cells [64], suggesting the special function of hnRNPK in MM pathogenesis. In this study, we found two novel phosphorylation sites, S417 and S420, for hnRNPK. The functional implication of above mentioned novel phosphorylation sites

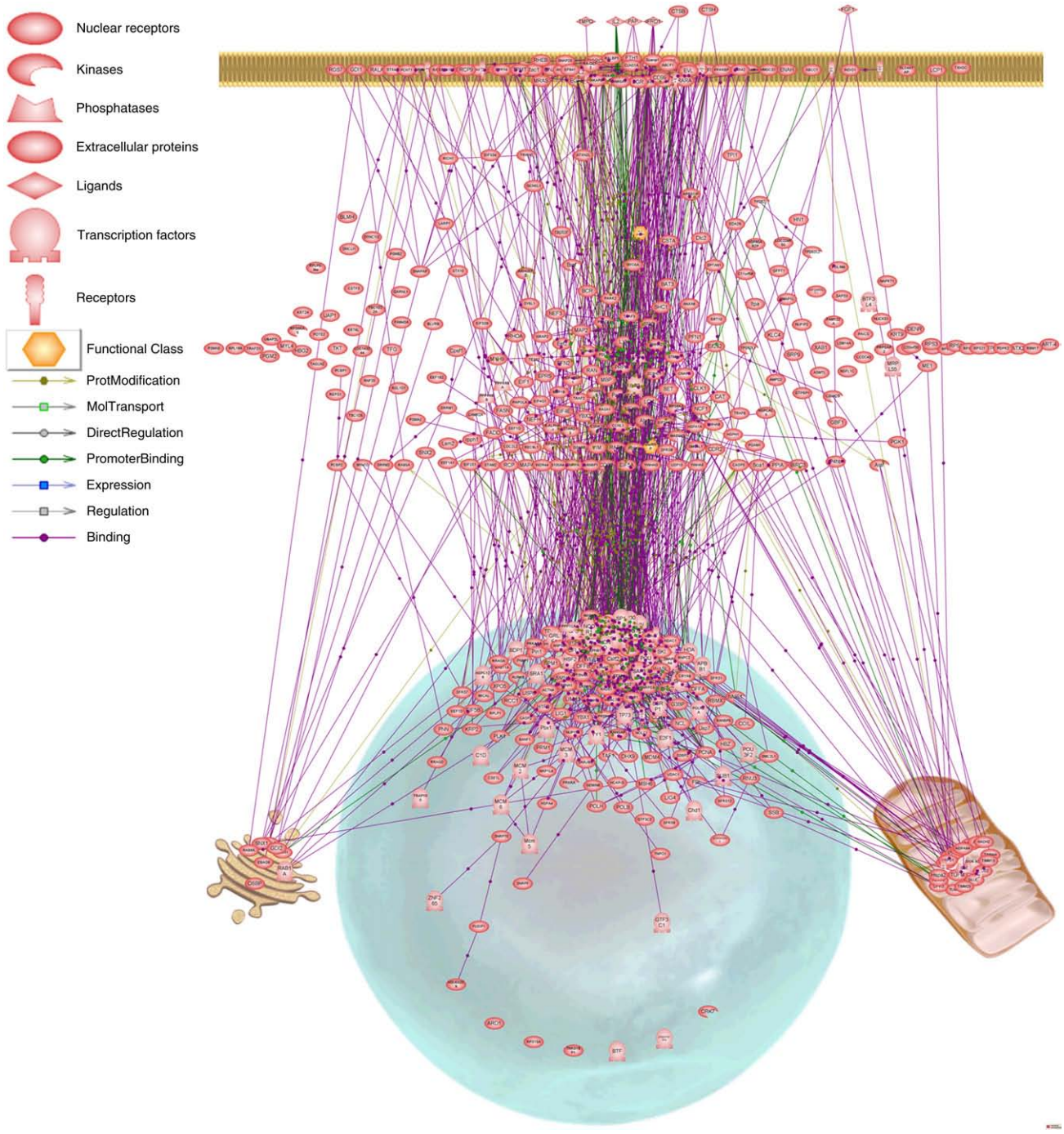

Fig. 4 – Pathway mapping of proteins with altered expression using Pathway Studio. Identified phosphoproteins listed inSupplemental Table S1 were imported into PathwayAssist, and an interaction map was generated. Shown are proteins that bound to another identified protein either directly or indirectly. Each node represents either a protein entity or a control mechanism of the interaction. The biological network can be accessed via internet page (http://life-health.jnu.edu.cn/MMphosphopathway.html).

is not clear. This is an issue which would deserve further investigation.

3.4. *Biological interaction networking (BIN) of identified proteins*

Supplemental Table S1 only shows a list of phosphoproteins in MM cells. It lacks the biochemical context. To create significance out of otherwise static proteomic data, we constructed a biological interaction networking (BIN) of the phosphoproteins identified in MM cells (Fig. 4). The phosphoproteins that could be networked were linked by various relationships such as protein interactions, modifications and regulation of expression. These relationships are color coded, and the legends are provided next to the map. However, this does not mean that all the interactions took place within a single spatial and temporal situation, but it enables the identification of those biochemical pathways that may be altered in MM. We have created a webpage that hosts the BIN created in this study (http://life-health.jnu.edu.cn/MMphosphopathway.html). The webpage contains hyperlinks through which the information on the phosphoproteins and the nature of their interactions with each other are provided. Where available, the references including description of the experimental findings from which the BIN was based on were also included. This should be useful for formulating testable hypotheses to understand the function of the identified phosphoproteins in MM cells. The network strongly suggested that the majority of the proteins identified in this study were integral parts of protein complexes. Two proteins with the most numbers of connectivity were MAPK3 and PDPK-1, suggesting that these proteins represent central nodes in the MM phosphoproteome. Several proteins remained

Table 1 – SCANSITE prediction at high stringency (0.2%), medium stringency (1.0%), and low stringency (5.0%) within identified phosphorylation sites for kinase phosphorylation and binding motifs. Phosphorylation sites in the MM phosphoproteome dataset were grouped according to kinase motifs. The detailed information of the phosphorylated sites and kinases is provided as Supplemental_Table_S3. (http://life-health.jnu.edu.cn/phospho/MMphosphodata/Supplemental_Table_S3.xls).

Kinase	Kinase group type	Hits (0.2%)	Hits (1%)	Hits (5%)
GSK3-improved	Acidophilic serine/threonine kinase	0	11	2
Casein kinase 1	Acidophilic serine/threonine kinase	1	3	22
ATM kinase	Acidophilic serine/threonine kinase	1	4	17
GSK3 kinase	Acidophilic serine/threonine kinase	6	9	30
Casein kinase 2	Acidophilic serine/threonine kinase	37	25	13
PKC delta	Basophilic serine/threonine kinase group	0	1	16
PKC alpha/beta/gamma	Basophilic serine/threonine kinase group	0	1	16
PKC mu	Basophilic serine/threonine kinase group	0	2	14
AMP-activated protein kinase	Basophilic serine/threonine kinase group	0	4	3
PKC zeta	Basophilic serine/threonine kinase group	0	5	3
Calmodulin dependent kinase 2	Basophilic serine/threonine kinase group	1	4	18
PKC epsilon	Basophilic serine/threonine kinase group	2	5	16
Clk2 kinase	Basophilic serine/threonine kinase group	4	2	22
Protein kinase A	Basophilic serine/threonine kinase group	9	11	30
Akt kinase	Basophilic serine/threonine kinase group	11	8	19
DNA PK	DNA damage kinase group	2	8	36
PDK1 binding	Kinase binding site group	3	0	8
p38 MAPK	Proline-dependent serine/threonine kinase group	1	1	9
Cdc2 kinase	Proline-dependent serine/threonine kinase group	6	26	52
Cdk5 kinase	Proline-dependent serine/threonine kinase group	6	30	56
Erk1 kinase	Proline-dependent serine/threonine kinase group	11	13	40
14-3-3 Mode 1	Phosphoserine/threonine binding group	17	14	28
Shc PTB	hosphotyrosine binding group	0	0	1
Crk SH2	Src homology 2 group	0	0	1
Fyn SH2	Src homology 2 group	0	0	1
Itk SH2	Src homology 2 group	0	0	1
Nck SH2	Src homology 2 group	0	0	1
PLCg N-terminal SH2	Src homology 2 group	0	0	1
Abl Kinase	Src homology 2 group	0	0	2
Lck SH2	Src homology 2 group	0	0	2
PLCg C-terminal SH2	Src homology 2 group	0	0	2
p85 SH2	Src homology 10 group	0	0	2
Fgr SH2	Src homology 7 group	0	0	3
Abl SH2	Src homology 4 group	1	0	1
Grb2 SH2	Src homology 8 group	2	0	0
Fgr kinase	Tyrosine kinase group	0	0	2
Itk kinase	Tyrosine kinase group	0	0	2
EGFR kinase	Tyrosine kinase group	0	2	0
PDGFR kinase	Tyrosine kinase group	2	0	0

orphans because there is insufficient information in the database to link them to other proteins in the BIN.

3.5. *SCANSITE analysis*

To predict the kinase substrate relationships from the dataset, the computer algorithm SCANSITE was used (http://scansite.mit.edu). SCANSITE makes use of soluble peptide library phosphorylation data to predict substrates recognized by specific kinases. Table 1 shows the results of phosphopeptides that were identified in this study and were predicted to be associated with a kinase binding motif by SCANSITE at the highest stringency (0.2%), medium stringency (1.0%), and low stringency (5.0%) levels. Totally, 123 unique phosphorylated sites were matched with the predicted ones in highest stringency level. It was found that most of the phosphorylation sites determined in this study were phosphorylated by acidophilic serine/threonine kinase and basophilic serine/threonine kinase group.

Notably, we found seventeen sites that were phosphorylated by the kinase 14-3-3 Mode 1. The 14-3-3 proteins are a family of multifunctional phosphoserine/phosphotheronine binding molecules that can serve as effectors of survival signaling [65]. They are involved in a variety of important cellular processes including cell cycle progression, growth, differentiation as well as apoptosis. 14-3-3 proteins participate in phosphorylation-dependent protein–protein interactions that control progression through the cell cycle, initiation and maintenance of DNA damage checkpoints, activation of MAP kinase, prevention of apoptosis, and coordination of integrin signaling and cytoskeletal dynamics [66,67]. The oncogenic role of 14-3-3 proteins has been proposed in recent studies [68]. Our SCANSITE analysis suggests that 14-3-3 proteins may play an important role in maintaining MM cell survival and thus implicates an attractive opportunity for therapeutic development by inhibiting 14-3-3 proteins to sensitize myeloma cells to apoptosis.

4. Conclusions

The rich and complex regulatory nature of protein phosphorylation offers an exciting and challenging opportunity for future proteomic studies. Here, using emerging technologies for the enrichment of phosphopeptides, we presented the first large-scale phosphoproteomic analysis of primary human MM cells. The feasibility of extracting and recovering phosphopeptides from primary human MM cells was demonstrated. A variety of cellular functions was covered from the identified phosphoproteins, including nucleic acid binding, kinase, signal transduction, cytoskeleton maintenance, chaperone, etc. Many phosphosites identified in this study appear to be novel, providing a valuable resource and starting point for future functional studies of human MM. The newly identified phosphoproteins may contribute to the identification of novel regulatory mechanisms of MM development. However, it must be kept in mind that phosphorylation in signal transduction pathways is a dynamic process, in which successive phosphorylation and dephosphorylation events at multiple sites are responsible for biological activity in the highly regulated systems. Further work is ongoing that combines this methodology with quantitative methods to detect the dynamics of the protein phosphorylation, comparing human MM cells with normal plasma cells to analyze the difference of kinase or signal pathway in biological processes in the molecular level. This analysis will provide further insights into the regulation of variable phosphorylations in MM cells.

Acknowledgements

This work was partially supported by the 2007 Chang-Jiang Scholars Program, “211” Projects, Talents Start-up Foundation of Jinan University (Grant 51207040) and Open Research Fund Program of the State Key Laboratory of Virology of China (Grant 2009003).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jprot.2010.03.004.

REFERENCES

[1] Kyle RA, Rajkumar SV. N Engl J Med 2004;351:1860–73.
[2] Sirohi B, Powles R. Lancet 2004;363:875–87.
[3] Lichtenstein A, Tu Y, Fady C, Vescio R, Berenson J. Cell Immunol 1995;162:248–55.
[4] Ogata A, Chauhan D, Teoh G, Treon SP, et al. J Immunol 1997;159:2212–21.
[5] Chauhan D, Kharbanda S, Ogata A, Urashima M, et al. Blood 1997;89:227–34.
[6] Catlett-Falcone R, Landowski TH, Oshiro MM, Turkson J, et al. Immunity 1999;10:105–15.
[7] Tu Y, Gardner A, Lichtenstein A. Cancer Res 2000;60:6763–70.
[8] de Graauw M, Hensbergen P, van de Water B. Electrophoresis 2006;27:2676–86.
[9] Schmelzle K, White FM. Curr Opin Biotechnol 2006;17:406–14.
[10] Gronborg M, Kristiansen TZ, Stensballe A, Andersen JS, et al. Mol Cell Proteomics 2002;1:517–27.
[11] Boekhorst J, van Breukelen B, Heck Jr A, Snel B. Genome Biol 2008;9:R144.
[12] Hoffert JD, Knepper MA. Anal Biochem 2008;375:1–10.
[13] Paradela A, Albar JP. J Proteome Res 2008;7:1809–18.
[14] Johnson SA, Hunter T. Nat Biotechnol 2004;22:1093–4.
[15] Thingholm TE, Jensen ON, Robinson PJ, Larsen MR. Mol Cell Proteomics 2008;7:661–71.
[16] Nuhse TS, Stensballe A, Jensen ON, Peck SC. Mol Cell Proteomics 2003;2:1234–43.
[17] Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, et al. Proc Natl Acad Sci U S A 2004;101:12130–5.
[18] McNulty DE, Annan RS. Methods Mol Biol 2009;527:93–105 x.
[19] Mazanek M, Mituloviae G, Herzog F, Stingl C, et al. Nat Protoc 2007;2:1059–69.
[20] Thingholm TE, Jensen ON, Larsen MR. Proteomics 2009;9:1451–68.
[21] Kyono Y, Sugiyama N, Imami K, Tomita M, Ishihama Y. J Proteome Res 2008;7:4585–93.
[22] Thingholm TE, Jorgensen TJ, Jensen ON, Larsen MR. Nat Protoc 2006;1:1929–35.
[23] Belu AM, Graham DJ, Castner DG. Biomaterials 2003;24:3635–53.

[24] Jennings CD, Foon KA. Blood 1997;90:2863–92.
[25] Zhan F, Tian E, Bumm K, Smith R, et al. Blood 2003;101: 1128–40.
[26] Haas W, Faherty BK, Gerber SA, Elias JE, et al. Mol Cell Proteomics 2006;5:1326–37.
[27] Olsen JV, Blagoev B, Gnad F, Macek B, et al. Cell 2006;127: 635–48.
[28] Macek B, Gnad F, Soufi B, Kumar C, et al. Mol Cell Proteomics 2008;7:299–307.
[29] Macek B, Mijakovic I, Olsen JV, Gnad F, et al. Mol Cell Proteomics 2007;6:697–707.
[30] Mi H, Guo N, Kejariwal A, Thomas PD. Nucleic Acids Res 2007;35:D247–52.
[31] Nikitin A, Egorov S, Daraselia N, Mazo I. Bioinformatics 2003;19:2155–7.
[32] Obenauer JC, Cantley LC, Yaffe MB. Nucleic Acids Res 2003;31: 3635–41.
[33] Sui S, Wang J, Yang B, Song L, et al. Proteomics 2008;8:2024–34.
[34] Manley JL, Tacke R. Genes Dev 1996;10:1569–79.
[35] Kanopka A, Muhlemann O, Petersen-Mahrt S, Estmer C, et al. Nature 1998;393:185–7.
[36] Krecic AM, Swanson MS. Curr Opin Cell Biol 1999;11:363–71.
[37] Kerekatte V, Smiley K, Hu B, Smith A, et al. Int J Cancer 1995;64:27–31.
[38] Joseph P, O'Kernick CM, Othumpangat S, Lei YX, et al. Mol Carcinog 2004;40:171–9.
[39] Arora S, Yang JM, Kinzy TG, Utsumi R, et al. Cancer Res 2003;63:6894–9.
[40] Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. Science 2002;298:1912–34.
[41] Blume-Jensen P, Hunter T. Nature 2001;411:355–65.
[42] Chang L, Karin M. Nature 2001;410:37–40.
[43] Richter-Landsberg C. J Mol Neurosci 2008;35:55–63.
[44] Ditzel HJ, Strik MC, Larsen MK, Willis AC, et al. J Biol Chem 2002;277:21712–22.
[45] Khurana S, George SP. FEBS Lett 2008;582:2128–39.
[46] Butt E, Abel K, Krieger M, Palm D, et al. J Biol Chem 1994;269: 14509–17.
[47] Harbeck B, Huttelmaier S, Schluter K, Jockusch BM, Illenberger S. J Biol Chem 2000;275:30817–25.
[48] Barzik M, Kotova TI, Higgs HN, Hazelwood L, et al. J Biol Chem 2005;280:28653–62.
[49] Halbrugge M, Friedrich C, Eigenthaler M, Schanzenbacher P, Walter U. J Biol Chem 1990;265:3088–93.
[50] Zhuang S, Nguyen GT, Chen Y, Gudi T, et al. J Biol Chem 2004;279:10397–407.
[51] Smolenski A, Poller W, Walter U, Lohmann SM. J Biol Chem 2000;275:25723–32.
[52] Munzel T, Feil R, Mulsch A, Lohmann SM, et al. Circulation 2003;108:2172–83.
[53] Benz PM, Blume C, Seifert S, Wilhelm S, et al. J Cell Sci 2009;122:3954–65.
[54] Calderwood SK, Khaleque MA, Sawyer DB, Ciocca DR. Trends Biochem Sci 2006;31:164–72.
[55] Mitsiades CS, Mitsiades NS, McMullan CJ, Poulaki V, et al. Blood 2006;107:1092–100.
[56] Ogiso H, Kagi N, Matsumoto E, Nishimoto M, et al. Biochemistry 2004;43:15510–9.
[57] Zhao YG, Gilmore R, Leone G, Coffey MC, et al. J Biol Chem 2001;276:32822–7.
[58] Bomsztyk K, Denisenko O, Ostrowski J. Bioessays 2004;26: 629–38.
[59] Habelhah H, Shah K, Huang L, Ostareck-Lederer A, et al. Nat Cell Biol 2001;3:325–30.
[60] Notari M, Neviani P, Santhanam R, Blaser BW, et al. Blood 2006;107:2507–16.
[61] Inoue A, Sawata SY, Taira K, Wadhwa R. Proc Natl Acad Sci U S A 2007;104:8983–8.
[62] Kim JH, Hahm B, Kim YK, Choi M, Jang SK. J Mol Biol 2000;298: 395–405.
[63] Ostareck-Lederer A, Ostareck DH, Hentze MW. Trends Biochem Sci 1998;23:409–11.
[64] Evans JR, Mitchell SA, Spriggs KA, Ostrowski J, et al. Oncogene 2003;22:8012–20.
[65] Fu H, Subramanian RR, Masters SC. Annu Rev Pharmacol Toxicol 2000;40:617–47.
[66] Aitken A. Semin Cancer Biol 2006;16:162–72.
[67] Jin J, Smith FD, Stark C, Wells CD, et al. Curr Biol 2004;14: 1436–50.
[68] Li Z, Zhao J, Du Y, Park HR, et al. Proc Natl Acad Sci U S A 2008;105:162–7.