

Review

## Post-translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches

Jawon Seo and Kong-Joo Lee\*

Center for Cell Signaling Research, Division of Molecular Life Sciences and College of Pharmacy,  
Ewha Womans University, Seoul 120-750, Korea

Received 26 December, 2003

Recently produced information on post-translational modifications makes it possible to interpret their biological regulation with new insights. Various protein modifications finely tune the cellular functions of each protein. Understanding the relationship between post-translational modifications and functional changes (“post-translatomics”) is another enormous project, not unlike the human genome project. Proteomics, combined with separation technology and mass spectrometry, makes it possible to dissect and characterize the individual parts of post-translational modifications and provide a systemic analysis. Systemic analysis of post-translational modifications in various signaling pathways has been applied to illustrate the kinetics of modifications. Availability will advance new technologies that improve sensitivity and peptide coverage. The progress of “post-translatomics”, novel analytical technologies that are rapidly emerging, offer a great potential for determining the details of the modification sites.

**Keywords:** Biological function, Post-translational modifications, Proteomics, Mass spectrometry, Systemic analysis

### Introduction

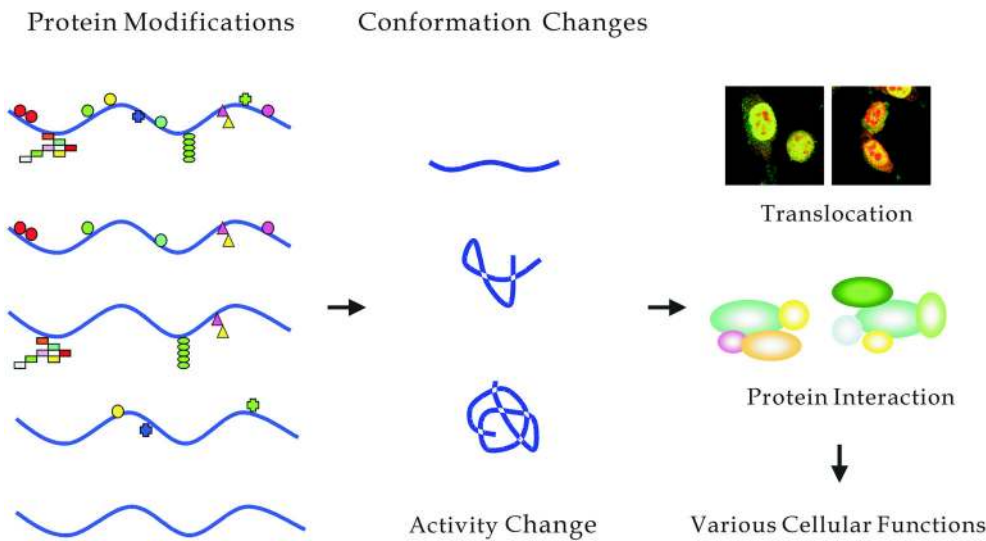
Primary structure of a protein obtained from human genome project is not sufficient to explain its various biological functions or their regulation mechanisms. Cellular homeostatic modifications of proteins have been shown to initiate various cellular processes. Post-translational modifications (PTMs) of protein call the covalent modifications of amino acids collectively. Protein has been

thought as a linear polymer decorated with simple modification, however, very complicated modifications in one protein are lately discovered in many processes. A variety of chemical modifications have been observed in a protein and these modifications alone or in various combinations occur time- and signal-dependent manner. PTMs of proteins determine their tertiary and quaternary structures and regulate their activities and functions. Some protein-protein interaction and localization for multiple functions of proteins are schematically presented in Fig. 1. Such information will help to understand biological phenomena and the disease states involving these proteins (Baenziger, 2003).

Identifications of PTMs resulting from various signals is lack of sensitive methodology and sufficient protein sample for analysis and characterization. Reviews in Pub Med reveals at least 200,000 PTMs, identified by studying the individual protein using radiolabeling, western analysis with antibody against specific modification, and mutagenesis of modification sites. This approach requires intensive and extensive labor, but does yield insufficient learning to the understanding of structure-function relationships.

Recent advances in proteomic methodology including mass spectrometry make it possible to identify proteins in complexes very rapidly (Mann and Jenson, 2003). While protein identification can be made by sequencing or to mapping only a few peptides, mapping of PTMs requires the complete coverage of peptides of a protein. Also the protein modifications *in vivo* occur only a small fraction of total proteins (less than 1%). Furthermore, the samples are a heterogeneous mixture of variously modified, and unmodified proteins. New methodology is required for the separation of homogeneous samples for detecting unknown protein modifications. Current proteomic technology is useful for detecting only simple modifications in large amounts of modified samples, and not for a through mapping of all cellular endogenous protein modifications. Since proteomic methodology has a tremendous potential for understanding the PTM, many efforts are progressing for enriching modified

\*To whom correspondence should be addressed.  
Tel: 82-2-3277-3038; Fax: 82-2-3277-3760  
E-mail: kjl@mm.ewha.ac.kr



**Fig. 1.** Schematic representation of protein modifications related to the regulation of biological processes.

samples and specific detection of modifications.

We will briefly review the protein modifications mainly in mammalian system, in the context of available methodologies for the detection of PTMs and the approaches for their systemic analysis.

### Biological Functions of Post-Translational Modifications

Post-translational modifications of proteins, which are not gene-template based, can regulate the protein functions, by causing changes in protein activity, their cellular locations and dynamic interactions with other proteins. A series of protein modifications are involved in the signaling pathway from membrane to nucleus in response to external stimuli. For example, receptor tyrosine kinases (RTKs) and ligand-mediated activation of RTKs result in autophosphorylation of both the receptor catalytic domain and noncatalytic regions of the cytoplasmic domain. Catalytic domain phosphorylation leads to activation and potentiation of receptor kinase activity. Noncatalytic domain phosphorylation creates docking sites for downstream cytoplasmic targets, which bind to specific receptor phosphotyrosine residues. Downstream signaling pathways are constructed in a modular fashion. In addition to SH2 and PTB (phosphotyrosine binding) domains, downstream signal proteins also contain domains that recognize other protein and phospholipid motifs. The arrangement and re-arrangement of various combinations of modular domains in different signaling proteins (combinatorial use) has allowed the creation of complex signaling networks and pathways. In addition to performing catalytic functions, signaling proteins modified by phosphorylation, myristoylation, farnesylation, cysteine oxidation, ubiquitination, acetylation, methylation,

nitrosylation, etc, serve as scaffolds for the assembly of multiprotein signaling complexes, as adaptors, as transcription factors and as signal pathway regulators, possibly (Table 1). This series of protein modifications in the cascade are transient and reversible between the switch-on and switch-off, and cross-talk between convergent pathways and different PTMs initiated independently, thereby providing further modulation of response. Because PTMs are central to all aspects of signaling cascades, understanding PTMs of proteins in signaling pathways will make it possible to find the common basis of biological regulation.

Phosphorylation and dephosphorylation on S, T, Y and H residues are the best known modifications involved in reversible, activation and inactivation of enzyme activity and modulation of molecular interactions in signaling pathway (Pawson, 2002). However, the fraction of phosphorylated proteins is generally very small (<0.1%) *in vivo* and turnover of this modification is very fast. Therefore specific detection of phosphorylated peptides using phosphatase inhibitors are indispensable to identify phosphorylation.

Acetylation and deacetylation in N-terminal and K-residue are suggested as rival to phosphorylation (Kouzarides, 2000). Recent identification of enzymes that regulate histone acetylation suggests a broader use of acetylation in cellular regulation mechanism, than was suspected previously. Acetylases are now known to modify a variety of proteins, including transcription factors, nuclear import factors and alpha-tubulin. Acetylation regulates many diverse functions, including DNA recognition, protein-protein interaction and protein stability. As reported before, there is even a conserved structure, the bromodomain, that recognizes acetylated residues and serves as a signaling domain. These are features characteristic of kinases. So, the question arises: Is acetylation a modification analogous to phosphorylation? Recent work showed that the acetylation of hypoxia-inducible factor 1

**Table 1.** The list of post-translational modifications

PTM type	Average MH+	Modified amino acid residue	Position	Remarks	Reported in Pub Med (case)
Acetylation	42.04	S K	N-term anywhere	Reversible, protein stability, regulation of protein function	11,069
Phosphorylation	79.98	Y, S, T, H, D	anywhere	Reversible, regulation of protein activity, signaling	103,235
Cys oxidation			anywhere	Reversible, oxidative regulation of proteins	
disulfide bond	-2.0	C			23,538
glutathionylation	305.31	C			63
sulfenic acid	16.00	C			228
sulfinic acid	32.00	C			642
Acylation				Reversible, cellular localization to membrane	
farnesylation	204.36	C	anywhere		1,349
myristoylation	210.36	G	N-term		644
		K	anywhere		
palmitoylation	238.41	C (S, T, K)	anywhere		681
Glycosylation			anywhere	Reversible, cell-cell interaction and regulation of proteins	24,115
O-linked	>800	S, T			
(O-Glc-NAc)	203.20,				
N-linked	>800	N			
Deamidation	0.98	N, Q	anywhere	N to D, Q to E	711
Methylation			anywhere	Regulation of gene expression, protein stability	29,889
monomethylation	14.03	K			
dimethylation	28.05	K			
trimethylation	42.08	K			
Nitration	45.0	Y		Oxidative damage	62
S-Nitrosylation	29.00	C			399
Ubiquitination		K	anywhere	Reversible/irreversible	1951
Sumoylation		K	[ILFV]K.D		104
Hydroxyproline	16.00	P		Protein stability	11,424
Pyroglutamic acid	-17	Q	N-term		710

(HIF-1) by acetyltransferase ARD1 enhances interaction of HIF-1a with pVHL and HIF-1a ubiquitination, suggesting that the acetylation of HIF-1a by ARD1 is critical to proteasomal degradation and for providing a key regulatory mechanism underlying HIF-1a stability (Jeong *et al.*, 2002). This suggests a connection between acetylation and ubiquitination. A recent review (Freiman and Tjian, 2003) suggests the lists several cases of ubiquitination (Ub), sumoylation (SUMO), acetylation and methylation of lysine residues that link specific covalent modification of the transcriptional apparatus to their regulatory function. It proposes that potential cascades of modifications serve as molecular rheostats that fine-tune the control of transcription in diverse organisms.

Ubiquitination and deubiquitination, sumoylation and desumoylation alter the biological functions of many proteins through covalent attachment to polypeptide modifiers (Ub and SUMO etc.). Ubiquitin has a well-documented role in targeting proteins for degradation by the proteasome, and additional effects of protein ubiquitination are now being

uncovered (Schwartz, 2003). Furthermore, multiple polypeptides that are distinct from, but related to, ubiquitin are also enzymatically coupled to target macromolecules, and these ubiquitin-like proteins participate in diverse biological processes such as DNA repair, autophagy and signal transduction. Imbalances in these modifications seem to induce various diseases (Giasson and Lee, 2003; Schwartz and Hochstrasser, 2003). PTM by the ubiquitin-like SUMO protein is emerging as a defining feature of eukaryotic cells (Seeler and Dejean, 2003). Sumoylation has been reported to the crucial roles in the regulatory challenges in nucleate cells, including the control of nucleocytoplasmic signaling, and transport, faithful replication of a large and complex genome, as well as the regulation of gene expression. Since the regulatory relationship between Ub and SUMO is studying to unveil the mechanism, the proteomic methodology will be developed for this purpose.

Redox-dependent PTM of proteins is emerging as a key signaling system conserved through evolution, influences

many aspects of cellular homeostasis. This system exemplify dynamic regulation of protein function by reversible modification of cysteine residues, which, in turn, regulate many cellular processes such as cell proliferation, differentiation and apoptosis. Cysteine oxidation including disulfide formation, sulfenic- and sulfinic acid formation, glutathionylation etc, is reversible and represents the cascades of proteins modifications with phosphorylation, acetylation, ubiquitination etc. One example of these cascades is interplay between phosphorylation- and redox-dependent signaling at the level of phosphotyrosine phosphatase-mediated regulation of receptor tyrosine kinases (RTKs) (Chiarugi and Cirri, 2003; van Montfort *et al.*, 2003). Signal transduction by oxygen species through reversible phosphotyrosine phosphatase inhibition, represents a widespread and conserved component of the biochemical machinery that is triggered by RTKs. This reversible reduction can be occurred by peroxiredoxin and thioredoxin family of proteins (Yano *et al.*, 2002; Woo *et al.*, 2003). The active-site cysteine of peroxiredoxins is selectively oxidized to cysteine sulfinic acid during catalysis, leading to inactivation of peroxidase activity. The sulfinic form of peroxiredoxin I, produced during the exposure of cells to H<sub>2</sub>O<sub>2</sub>, is rapidly reduced to the catalytically active thiol form. The mammalian cells' ability to reduce protein sulfinic acid might serve as a mechanism to repair oxidatively damaged proteins or represent a new type of cyclic modification by which the function of various proteins is regulated (Woo *et al.*, 2003). Thioredoxin, a ubiquitous 12 kDa protein with a catalytically active disulfide active site (Cys-Gly-Pro-Cys), plays a central role in controlling the redox status of disulfide bonds in proteins that regulate a range of processes. This family of proteins and its reaction mechanism can be identified by comprehensive analysis of proteins using mass spectrometry (Yano *et al.*, 2002). Redox regulation is important for cellular homeostasis (Georgiou, 2002). Tyrosine nitration is increasingly recognized as a prevalent, functionally significant PTM that serves as an indicator of nitric oxide (z.rad;NO)-mediated oxidative inflammatory reactions. Nitration of proteins modulates catalytic activity, cell signaling and cytoskeletal organization. Several reactions mediate protein nitration, and all of these predominantly depend on z.rad;NO- and nitrite-dependent formation of nitrogen dioxide, which is capable of nitrating aromatic amino acids, nucleotides and unsaturated fatty acids (Schopfer *et al.*, 2003).

Protein lipid-modifications including myristoylation, palmitoylation, farnesylation, and prenylation, known for a long time, have been shown to have a role in protein-membrane interactions, protein trafficking, and enzyme activity. Until recently, however, the molecular machinery that carries out reversible palmitoylation of proteins has been elusive (Linder and Deschenes, 2003). Protein palmitoyltransferases that carry out protein palmitoylation have been recently described. Two major classes of protein palmitoylthioesterases were also reported. Recent advances have been made in the understanding

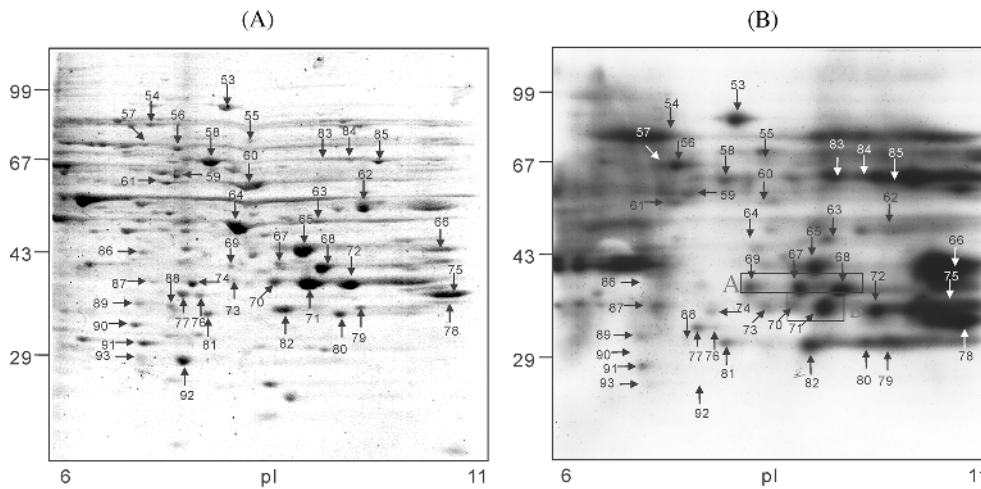
of mechanisms of addition of palmitate to proteins and removal of palmitate from proteins. Palmitoylation/depalmitoylation mechanism is another example of the reversible PTMs regulated by various stimuli.

Glycosylation is a regulatory PTM. O-linked beta-N-acetylglucosamine (O-GlcNAc) is both an abundant and dynamic posttranslational modification similar to phosphorylation that occurs on serine and threonine residues of cytosolic and nuclear proteins (Wells *et al.*, 2003; Whelan and Hart, 2003). Although O-GlcNAc was discovered more than 20 years ago, the elucidation of O-GlcNAc as a PTM has been slow, because of the lack of tools for study. Recently the enzymes catalyzing the addition and removal of GlcNAc were identified and proteomic site mapping of both O-GlcNAc and O-phosphate was developed. Recent technological innovations should advance our present understanding of O-GlcNAc and its role in diabetic and cardiovascular complications. There is considerable evidence that glycosylation plays a key role in immune regulation, e.g. in the development, survival and reactivity of T cells (Daniels *et al.*, 2002).

### **Proteomic Analysis of Post-translational Modifications**

Once the proteins are identified, the next question is whether or not the proteins are post-translationally modified and how. PTM analyses are more difficult than protein identification for the following reasons: Firstly, highly sensitive methods are required for detection due to a low stoichiometry. For example, since only 5-10% of a protein kinase substrate is phosphorylated, we need methods to detect the modified protein at very low levels (<5-10 fmol). Secondly, since the covalent bond between the PTM and amino acid side chain is typically labile, it is often difficult to maintain the peptide in its modified state during sample preparation and subsequent ionization in mass spectrometry. Thirdly, PTMs are frequently transient in dynamic homeostasis of nature. Therefore, information on the characteristics of the PTMs in protein is needed in order to find out the new PTMs identification and their position in the protein.

**Characteristics of post translationally modified proteins in gel** In general, the analysis of PTMs can be performed by a method combining 2-D gel electrophoresis (2 DE) and MS. In the case of phosphoproteins, the unique resolving power of 2 DE allows separation of different modified forms of the same protein into distinct spots on the gel (Gianazza *et al.*, 1989). Heterogeneous phosphorylations of a protein gives a series of spots with the same molecular weight but different pI values, i.e., spots located on a string horizontally over the gel (Fig. 2). In contrast, heterogeneous glycosylation of a protein causes a series of spots because they differ both in molecular weights and pI values due to the nature of glycan structure.



**Fig. 2.** Two dimensional gel images visualized by silver staining and immunostaining. (A, B) Silver stained 2D gel in pH range 6-11 of mouse fibrosarcoma cells (A), and western analysis of 2D gel with anti-phosphotyrosine antibody (B). Box A indicates the modifications of MAP kinase kinase (spot 67, 68, 69) and Box B presents the GAPDH (spot 70, 71). Adapted and reprinted with permission from Kim *et al.*, 2002.

**Detection of modified proteins by autoradiography and western analysis** This method can be used to detect different forms of phosphorylation (S, T, Y, H) at low level by incorporation of a radiolabeled phosphate ( $^{32}\text{P}$ ) followed by visualization by autoradiography. Western analysis with antibodies against specific types of phosphorylation site is an alternative way to allow the distinction between different types of phosphorylation (Soskic *et al.*, 1999). Likewise, acetylated proteins can be detected with antibody against anti-acetyl lysine. In this case, the quality of antibody including specificity and sensitivity is the matter of detection. Detection of glycosylated proteins on a gel or a western blot can be accomplished by utilizing any of four different methods (Packer *et al.*, 1999): periodic acid/schiff (PAS) staining, digoxigenin (DIG)/anti-DIG alkaline phosphatase labeling, affinity blotting and monosaccharide composition analysis of the liberated glycan using high-pressure anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD). Glycosylphosphatidylinositol (GPI) anchored proteins have recently been characterized by 2 DE after GPI-specific phospholipase C partitioning in Triton X114 (Sherrier and Prime, 1999).

**Analysis of PTMs in proteins by mass spectrometry** PTMs in proteins can be characterized by MS with the following strategies (Andersen *et al.*, 1996); determination of the intact molecular weight, mass fingerprinting of peptide mixtures derived by proteolytic digestion, chemical or enzymatic removal of the modification ( $\beta$ -elimination, phosphatase) and sequencing of the modified peptide by tandem MS. With these strategies, more than 200 PTMs of amino acids have been reported, all associated with either an increase or a decrease in molecular mass. Therefore, MS is an ideal tool for identification and characterization of PTMs with

proteomic studies. Information on the nature of PTMs will facilitate understanding of protein function. A list of reported PTMs and the corresponding site of amino acids is shown in Table 1.

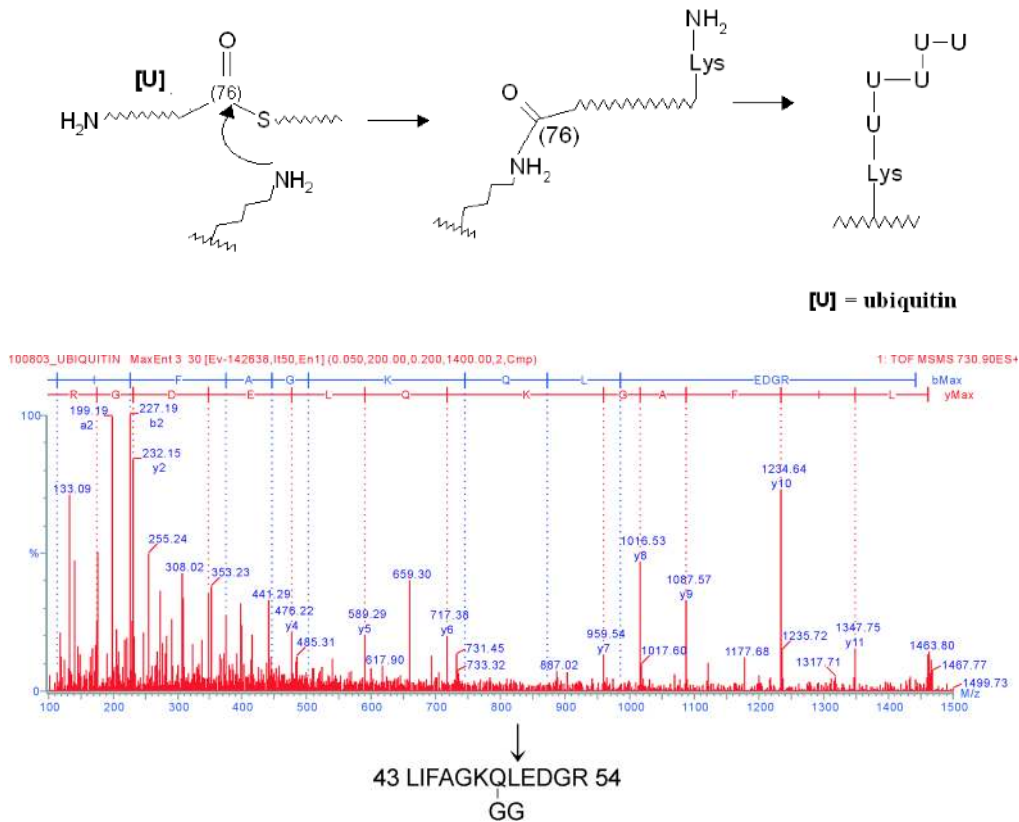
**General problems associated with the characterization of post-translational modifications** Several of criteria must be fulfilled before full characterization of the protein can be accomplished. The amount of protein available must be sufficient to perform the analysis. In general, a spot being clearly visible with coomassie blue staining, is often sufficient for localization and verifying the modification. In some cases, depending on the type of modification, even smaller amounts of starting material is sufficient to characterize the modifications in proteins from gels. However, silver stained spots must be concentrated to perform the PTM analysis. Presence of contaminants, such as alkali metal ions, often causes reduced intensity or suppression of the peptide signals, resulting in decreased sequence coverage. Simultaneous removal of contaminants and sample concentration can be achieved by purification of the peptide mixture on micro-columns prior to mass analysis (Gobom *et al.*, 1999). The sequence coverage can frequently be further improved by dividing the sample into several fractions using stepwise elution from the column with increasing organic solvent concentration. Co-migration of proteins in the gel as well as contamination with human keratins and other proteins also represents a major problem because proteolytic digestion results in a mixed population of peptides derived from different proteins. Due to ion suppression effects, the sequence coverage will be lower for each of the identified proteins in the mixture. Use of narrow pI gels, which have better resolution, may solve these problems. Also certain modifications reduce the ion yield of the corresponding

peptide resulting in suppression of its signal. This is especially the case for acidic modifications, e.g., phosphorylation and acetylation. Modified proteins are often separated from the non-modified proteins in the 2-D gel. However, a given spot may contain several molecular species caused by modification of one of several possible positions in a protein. This means that all molecules in a spot may contain the same modification which may or may not necessarily be at the same position. In such cases only a small amount of modified peptides may be present, in the rest being the nonmodified peptides. This may result in failure to observe the modified peptide (Mortz *et al.*, 1996; Spangfort *et al.*, 1996). Since protein phosphorylation results in a mass increase of 80 Da per phosphate group, MS is the obvious method for characterizing this type of modification.

The strategy used in our laboratory for site specific determination of phosphorylation is differential peptide mapping before and after treatment with alkaline phosphatase, or removal of phosphate with alkali buffer, or analysis of the phosphorylated peptides purified by immobilized metal affinity chromatography (IMAC). The first approach will be illustrated by identification of the phosphorylation site in an *in vivo* phosphorylated human protein. Phosphorylated proteins were detected in the gel by autoradiography after labeling with  $^{33}\text{P}$ . The spot of interest was excised in the autoradiogram and submitted to in-gel tryptic digestion. A characteristic metastable loss of  $\text{H}_3\text{PO}_4$  and  $\text{HPO}_3$  was associated with peaks having slightly poorer resolution (Kusmann *et al.*, 1997). The mass spectrum of phosphorylated peptide shows a mass decrease of 80 Da, confirming the loss of a phosphate group. Because of the suppressed signals originating from phosphorylated peptides, selective isolation of the phosphopeptide(s) using immobilized metal affinity chromatography (IMAC) (Annan and Carr, 1996; Figeys *et al.*, 1998) prior to MS analysis may be advantageous. This procedure allows identification of the phosphorylation sites in b-casein from as little as 500 fmol protein applied on a SDS-gel (Jensen *et al.*, 2000). The phosphopeptides are eluted with a high pH buffer on a micro IMAC column. Upon on-target dephosphorylation with alkaline phosphatase a mass decrease of 80 Da is observed. An alternative way to identify phosphorylation sites in proteins is by parent ion scanning (Carr *et al.*, 1996; Kim *et al.*, 2002). Different modifications yield different diagnostic low-mass ions during collision induced dissociation (CID). Thus, phosphorylated peptides yield fragment ions at  $m/z$  63 and 79 corresponding to  $\text{PO}_2$  and  $\text{PO}_3$ , respectively. This technique is mainly used in combination with LC-ESI-q-TOF tandem MS, but has also been successfully applied to the identification of phosphorylation sites by direct mass spectrometric analysis of the peptide mixture (Neubauer and Mann, 1999).

Use of MS in glycoprotein analysis has been reviewed (Burlingame, 1996; Kuster *et al.*, 1997; Harvey *et al.*, 1998). The general strategy for mass spectrometric characterization of purified glycoproteins is to measure the molecular weight

of the intact proteins by mass spectrometry, followed by identification of the sites of glycosylation. Frequently, differential peptide mapping before and after treatment of the peptide mixture with appropriate endoglycosidases is performed. Structural characterization of the glycan structure can be performed after chemical or enzymatic liberation of the glycan, and monitored by high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) or by mass spectrometry. Alternatively, direct mass spectrometric analysis of the glycopeptide often in combination with sequential exoglycosidase treatment can be performed (Kuster *et al.*, 1997). The methods used for glycan structure determination cannot be directly used in proteomic analysis because the amount of protein in a gel generally is too low. Glycopeptide signals are often suppressed in the presence of the non-glycosylated peptides especially if the glycans are terminated with a negatively charged sialic acid residue. Presence of sialic acid causes a pronounced metastable state and fragmentation in MALDI-TOF MS, both of which complicate interpretation of the spectra. In spite of these difficulties, a few reports have described accomplishing characterization of N-linked glycans in glycoproteins separated by gel electrophoresis (Kuster and Mann, 1999). The intact molecular weights of the protein in the different bands as determined by MALDI-TOF MS, indicated the presence of the non-glycosylated, mono- and di-glycosylated species in the three bands, respectively. By direct MALDI-TOF MS analysis, the peptide mixture generated by in-gel digestion of the glycosylated peptides was observed. Subsequent sequential treatment of the peptide mixture with different exoglycosidases and endoglycosidases with monitoring by MS, allowed determination of the glycan structure in this site. After microbore HPLC separation of the mixture, a fraction containing typical glycopeptide peak pattern was identified by MALDI-TOF MS. Using the same strategy it was possible to demonstrate a second glycosylated peptide and to identify the glycan structures on this site. For N-linked glycans, recent work demonstrated that it is possible to perform in-gel endoglycosidase treatment followed by extraction of the liberated glycan and mass spectrometric characterization of the glycans (Kuster and Mann, 1999). By performing the treatment in a buffer containing  $\text{H}_2^{16}\text{O}$  and  $\text{H}_2^{18}\text{O}$  in a 1 : 1 proportion followed by in-gel digestion and mass spectrometric analysis of the resulting peptide mixture, it was possible to determine which Asn residues were glycosylated, because the Asp residues obtained by conversion of asparagine would contain equal amounts of  $^{16}\text{O}/^{18}\text{O}$  (Packer *et al.*, 1998). This method, however, does not allow site specific assignment of the different glycan structures. Packer and coworkers have characterized human a2-HS glycoprotein and human a1-protease inhibitor (Yan *et al.*, 1999). Different forms of the glycoproteins were separated on narrow pI range 2 DE gels (pH 4.5-5.5), transferred to a PVDF membrane following enzymatic (for N-linked glycans) or chemical (for O-linked glycans) release of the glycans. The



**Fig. 3.** Determination of a ubiquitination site of multi-ubiquitinated protein by LC-ESI-Q-TOF tandem MS (Song *et al.*, unpublished data).

glycans were extracted, separated by HPAEC-PAD, and the isolated glycans were analyzed by ESIMSMS. No labeling procedures have been found that allow identification of N-terminal processing. Instead, peptide mass mapping after in-gel digestion with two or more different enzymes was used to putatively locate the processing position.

Using the strategy described above, ubiquitinated proteins have been identified in yeast proteome analysis. An example is shown in Fig. 3. Ubiquitination results in a 76 amino acid polypeptide attached to a lysine residue through an isopeptide bond. When a ubiquitinated protein is digested with trypsin, Gly-Gly dipeptide from the original ubiquitin polypeptide still covalently attached to the affected lysine of the original protein is left (Peng *et al.*, 2003). By interrogating sequence database with and without the net additional mass of 114 Da, the precise site of ubiquitination often is inferred. This resulted in identification of a peptide that corresponded to ubiquitination on the lysine residue (1460.8432 m/z Da).

New PTMs can be identified by using MS in combination with gel electrophoresis and observation of unexpected peptide mass shifts during mass analysis. Thus, characterization of lysine methylation has recently been reported by the identification of the modified residues by high mass accuracy MALDI-TOF MS. As shown in above, the combining the separation with 2 DE and nano-LC, and the identification with peptide mass fingerprinting and sequencing

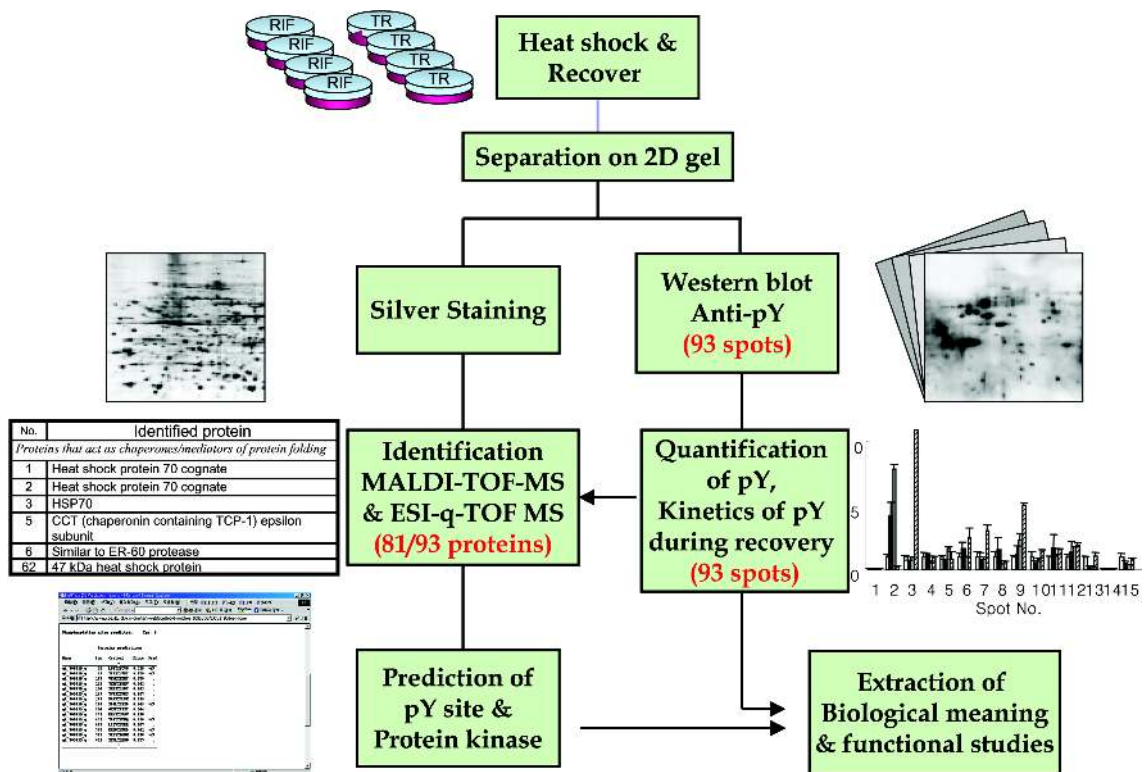
with MS, and biological detection with radiolabel and western analysis will give us the thorough understanding of PTMs.

### Systemic Approach for Protein Modifications

Proteomic analysis is allowing the elucidation of expression differences in control and disease states or in stimulated biological condition. However, reversible and dynamic cellular processes can be regulated by protein modifications. Although the protein modifications of individual protein occurring in response to the stimuli are identified and how they work together as systems, is not clear.

To comprehensively understand the protein modifications occurring in response to various signaling steps, several approaches were tried to identify the modified proteins simultaneously, using proteomics, despite technical difficulties in detection. This methodology is not possible with high throughput detection using LC/MS/MS because of lack of methods for the detection of modified proteins. Several modified populations of one protein should be separated on 2D-gel, detected quantitatively with high sensitivity and identified with mass spectrometry. Protein glutathionylation (Fratelli *et al.*, 2002) and nitrosylation (Kuncewicz *et al.*, 2003), and tyrosine phosphorylation (Kim *et al.*, 2002; Kim *et al.*, 2003; Salomon *et al.*, 2003) occurring in response to





**Fig. 4.** Schematic diagram of systemic analysis of protein tyrosine phosphorylation and its kinetics in response to heat shock.

several stimuli were investigated. Forty-two glutathionylated proteins formed in response to oxidative stress (hydrogen peroxide and diamide) were identified in human T lymphocytes by labeling with  $^{35}\text{S}$  labeling of the intracellular GSH pool. Enzymatic activities of some of the proteins were regulated by glutathionylation, suggesting that protein glutathionylation might be a possible mechanism for global regulation of protein functions (Fratelli *et al.*, 2002). Similarly, induction of S-nitrosylation was identified in 34 proteins by treating mesangial cells with IL-1 $\beta$  broadening our knowledge of potential signal transduction pathways and other cellular functions mediated by NO S-nitrosylation (Kuncewicz *et al.*, 2003). Many proteins modified in this process were identified, but, it is difficult to understand the biological rules from such massive identification of modified proteins, without studying the detail of each PTM individual using low throughput traditional way.

Further efforts were made to determine the biological functions of PTMs from the large set of tyrosine phosphorylated proteins caused by heat shock and VEGF (Kim *et al.*, 2002; Kim *et al.*, 2003). Control and experimental proteins were compared with kinetic study and the proteomic information obtained from experimental process was analyzed with bioinformatic tools. For example, tyrosine phosphorylation by heat shock was compared in control and thermotolerant cells, and simultaneously the changes occurring during recovery after heat shock treatment were analyzed by proteomics, combined with the separation on 2D-

gel, detection with western analysis using anti-phosphotyrosine antibodies, quantitative analysis and identification of modified proteins with mass spectrometry (Fig. 4). Eighty one of tyrosine phosphorylated proteins were identified, exhibiting various cellular functions; as chaperones, ion channels, signaling molecules, in transcription and translation processes, in amino acid biosynthesis, oxidoreduction, energy metabolism, and cell motility or structure, suggesting that heat shock turns on the various signaling pathways by activating protein-tyrosine kinases. Of these, 20 proteins were previously identified phosphorylated proteins and 64 were new from kinetic studies, the identified proteins can be grouped into three families; 1) proteins highly phosphorylated in thermotolerance cells at basal level and phosphorylated more significantly by heat shock in control than thermotolerant cells; 2) proteins highly phosphorylated in control cells at basal level and phosphorylated more easily by heat shock in thermotolerant cells; 3) proteins with similar basal phosphorylation level in both control and thermotolerant cells, and responding to heat shock similarly in both cells. Most of the phosphorylated proteins are possibly involved in heat shock signaling in different ways, with the first and second families of proteins influencing thermotolerances. The possible tyrosine phosphorylation sites, the possible protein tyrosine kinases phosphorylating these proteins, and the proteins binding to these phosphorylated sites were predicted employing bioinformatic tools. This study is the first example of use of proteomics to study protein modifications in order to



understand the dynamic cellular phenomena during heat shock response. In further work to determine the protein modifications involved in the differential signaling pathways, tyrosine phosphorylations in response to VEGF treatment were compared with heat shock response (Kim *et al.*, 2002). This comparison gave information on the proteins phosphorylated commonly and differentially in both signaling pathways. In more recent work the sites of tyrosine phosphorylation on different proteins from T cells in response to Gleevec have examined using multidimensional liquid chromatography and tandem mass spectrometry (Salomon *et al.*, 2003). This type of experimental designs should lead to new insights into signaling pathways as they occur in states of health and disease.

## Conclusions

Many protein modifications in different cellular signaling pathways regulate the various biological responses by changing the protein activity, localization and binding partners, called as post-translatomics. Understanding these relationships between PTMs and physiology will give us the common theory: how complicated biological processes can be regulated with fine tune having low error rate and how the network can be working from the combination of various signaling pathways. Post-translatomics, another huge challenge in biology for the future, include the individual biological studies of each PTM, powerful proteomic technology development for routine PTM analysis and systematic approach for understanding the biological network using systemic biology and bioinformatics. This will contribute to our understanding of biological processes and disease states in molecular level.

**Acknowledgments** This work was supported by KOSEF through the Center for Cell Signaling Research (CCSR) at Ewha Womans University and by KISTEP (IMT 2000-CS-2, MOST FPR03B3-04-110). J. Seo was supported by the Brain Korea 21 Project.

## References

- Andersen, J. S., Svensson, B. and Roepstorff, P. (1996) Electrospray ionization and matrix assisted laser desorption/ionization mass spectrometry: powerful analytical tools in recombinant protein chemistry. *Nat. Biotechnol.* **14**, 449-457.
- Annan, R. S. and Carr, S. A. (1996) Phosphopeptide analysis by matrix-assisted laser desorption time-of-flight mass spectrometry. *Anal. Chem.* **68**, 3413-3421.
- Baenziger, J. U. (2003) A major step on the road to understanding a unique posttranslational modification and its role in a genetic disease. *Cell* **113**, 421-422.
- Burlingame, A. L. (1996) Characterization of protein glycosylation by mass spectrometry. *Curr. Opin. Biotechnol.* **7**, 4-10.
- Carr, S. A., Huddleston, M. J. and Annan, R. S. (1996) Selective detection and sequencing of phosphopeptides at the femtomole level by mass spectrometry. *Anal. Biochem.* **239**, 180-192.
- Chiarugi, P. and Cirri, P. (2003) Redox regulation of protein tyrosine phosphatases during receptor tyrosine kinase signal transduction. *Trends Biochem. Sci.* **28**, 509-514.
- Daniels, M. A., Hogquist, K. A. and Jameson, S. C. (2002) Sweet 'n' sour: the impact of differential glycosylation on T cell responses. *Nat. Immunol.* **3**, 903-910.
- Figeys, D., Gygi, S. P., Zhang, Y., Watts, J., Gu, M. and Aebersold, R. (1998) Electrophoresis combined with novel mass spectrometry techniques: powerful tools for the analysis of proteins and proteomes. *Electrophoresis* **19**, 1811-1818.
- Fratelli, M., Demol, H., Puype, M., Casagrande, S., Eberini, I., Salmona, M., Bonetto, V., Mengozzi, M., Duffieux, F., Miclet, E., Bachi, A., Vandekerckhove, J., Gianazza, E. and Ghezzi, P. (2002) Identification by redox proteomics of glutathionylated proteins in oxidatively stressed human T lymphocytes. *Proc. Natl. Acad. Sci. USA* **99**, 3505-3510.
- Freiman, R. N. and Tjian, R. (2003) Regulating the regulators: lysine modifications make their mark. *Cell* **112**, 11-17.
- Georgiou, G. (2002) How to flip the (redox) switch. *Cell* **111**, 607-610.
- Gianazza, E., Celentano, F., Magenes, S., Etori, C. and Righetti, P. G. (1989) Formulations for immobilized pH gradients including pH extremes. *Electrophoresis* **10**, 806-808.
- Giasson, B. I. and Lee, V. M. (2003) Are ubiquitination pathways central to Parkinson's disease? *Cell* **114**, 1-8.
- Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R. and Roepstorff, P. (1999) Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **34**, 105-116.
- Harvey, D. J., Kuster, B. and Naven, T. J. (1998) Perspectives in the glycosciences--matrix-assisted laser desorption/ionization (MALDI) mass spectrometry of carbohydrates. *Glycoconj. J.* **15**, 333-338.
- Jensen, O. N., Stensballe, A. and Andersen, S. (2000) Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* **1**, 207-222.
- Jeong, J. W., Bae, M. K., Ahn, M. Y., Kim, S. H., Sohn, T. K., Bae, M. H., Yoo, M. A., Song, E. J., Lee, K. J. and Kim, K. W. (2002) Regulation and destabilization of HIF-1 $\alpha$  by ARD1-mediated acetylation. *Cell* **111**, 709-720.
- Kim, H. J., Song, E. J. and Lee, K. J. (2002) Proteomic analysis of protein phosphorylations in heat shock response and thermotolerance. *J. Biol. Chem.* **277**, 23193-23207.
- Kim, J. Y., Kim, K. W., Kwon, H. J., Lee, D. W. and Yoo, J. S. (2002) Probing lysine acetylation with a modification-specific marker ion using high-performance liquid chromatography/electrospray-mass spectrometry with collision-induced dissociation. *Anal. Chem.* **74**, 5443-5449.
- Kim, Y. M., Song, E. J., Kim, H. J. and Lee, K. J. (2003) Proteomic analysis of tyrosine phosphorylations in vascular endothelial growth factor- and reactive oxygen species-mediated signaling pathway of endothelial cells. *J. Biol. Chem.* Submitted.
- Kouzarides, T. (2000) Acetylation: a regulatory modification to

- rival phosphorylation? *EMBO J.* **19**, 1176-1179.
- Kuncewicz, T., Sheta, E. A., Goldknopf, I. L. and Kone, B. C. (2003) Proteomic analysis of s-nitrosylated proteins in mesangial cells. *Mol. Cell. Proteomics* **2**, 156-163.
- Kusmann, M., Lassing, U., Sturmer, C. A., Przybylski, M. and Roepstorff, P. (1997) Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of the neural cell adhesion protein neurolin purified by sodium dodecyl sulfate polyacrylamide gel electrophoresis or acidic precipitation. *J. Mass. Spectrom.* **32**, 483-493.
- Kuster, B. and Mann, M. (1999) <sup>18</sup>O-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. *Anal. Chem.* **71**, 1431-1440.
- Kuster, B., Wheeler, S. F., Hunter, A. P., Dwek, R. A. and Harvey, D. J. (1997) Sequencing of N-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography. *Anal. Biochem.* **250**, 82-101.
- Linder, M. E. and Deschenes, R. J. (2003) New insights into the mechanisms of protein palmitoylation. *Biochemistry* **42**, 4311-4320.
- Mann, M. and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**, 255-261.
- Mortz, E., Sarenava, T., Haebel, S., Julkunen, I. and Roepstorff, P. (1996) Mass spectrometric characterization of glycosylated interferon-gamma variants separated by gel electrophoresis. *Electrophoresis* **17**, 925-931.
- Neubauer, G. and Mann, M. (1999) Mapping of phosphorylation sites of gel-isolated proteins by nanoelectrospray tandem mass spectrometry: potentials and limitations. *Anal. Chem.* **71**, 235-242.
- Packer, N. H., Ball, M. S. and Devine, P. L. (1999) Glycoprotein detection of 2-D separated proteins. *Methods Mol. Biol.* **112**, 341-352.
- Packer, N. H., Lawson, M. A., Jardine, D. R., Sanchez, J. C. and Gooley, A. A. (1998) Analyzing glycoproteins separated by two-dimensional gel electrophoresis. *Electrophoresis* **19**, 981-988.
- Pawson, T. (2002) Regulation and targets of receptor tyrosine kinases. *Eur. J. Cancer Suppl* **5**, S3-10.
- Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D. and Gygi, S. P. (2003) A proteomics approach to understanding protein ubiquitination. *Nat. Biotechnol.* **21**, 921-926.
- Salomon, A. R., Ficarro, S. B., Brill, L. M., Brinker, A., Phung, Q. T., Ericson, C., Sauer, K., Brock, A., Horn, D. M., Schultz, P. G. and Peters, E. C. (2003) Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc. Natl. Acad. Sci. USA* **100**, 443-448.
- Schopfer, F. J., Baker, P. R. and Freeman, B. A. (2003) NO-dependent protein nitration: a cell signaling event or an oxidative inflammatory response? *Trends Biochem. Sci.* **28**, 646-654.
- Schwartz, D. C. and Hochstrasser, M. (2003) A superfamily of protein tags: ubiquitin, SUMO and related modifiers. *Trends Biochem. Sci.* **28**, 321-328.
- Schwartz, J. H. (2003) Ubiquitination, protein turnover, and long-term synaptic plasticity. *Sci. STKE*. **190**, pe26.
- Seeler, J. S. and Dejean, A. (2003) Nuclear and unclear functions of SUMO. *Nat. Rev. Mol. Cell. Biol.* **4**, 690-699.
- Sherrier, D. J. and Prime, T. A. (1999) Glycosylphosphatidylinositol-anchored cell-surface proteins from Arabidopsis. *Electrophoresis* **10**, 2027-2035.
- Soskic, V., Gorlach, M., Poznanovic, S., Boehmer, F. D. and Godovac-Zimmermann, J. (1999) Functional proteomics analysis of signal transduction pathways of the platelet-derived growth factor beta receptor. *Biochemistry* **38**, 1757-1764.
- Spangfort, M. D., Ipsen, H., Sparholt, S. H., Aasmul-Olsen, S., Larsen, M. R., Mortz, E., Roepstorff, P. and Larsen, J. N. (1996) Characterization of purified recombinant Bet v 1 with authentic N-terminus, cloned in fusion with maltose-binding protein. *Protein Expr. Purif.* **8**, 365-373.
- van Montfort, R. L., Congreve, M., Tisi, D., Carr, R. and Jhoti, H. (2003) Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B. *Nature* **423**, 773-777.
- Wells, L., Whalen, S. A. and Hart, G. W. (2003) O-GlcNAc: a regulatory post-translational modification. *Biochem. Biophys. Res. Commun.* **302**, 435-441.
- Whelan, S. A. and Hart, G. W. (2003) Proteomic approaches to analyze the dynamic relationships between nucleocytoplasmic protein glycosylation and phosphorylation. *Circ. Res.* **93**, 1047-1058.
- Woo, H. A., Chae, H. Z., Hwang, S. C., Yang, K. S., Kang, S. W., Kim, K. and Rhee, S. G. (2003) Reversing the inactivation of peroxiredoxins caused by cysteine sulfinic acid formation. *Science* **300**, 653-656.
- Yan, J. X., Sanchez, J. C., Binz, P. A., Williams, K. L. and Hochstrasser, D. F. (1999) Method for identification and quantitative analysis of protein lysine methylation using matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry and amino acid analysis. *Electrophoresis* **20**, 749-754.
- Yano, H., Kuroda, S. and Buchanan, B. B. (2002) Disulfide proteome in the analysis of protein function and structure. *Proteomics* **2**, 1090-1096.