Available online at www.sciencedirect.com
ScienceDirect
Methods 42 (2007) 289–297
METHODS
www.elsevier.com/locate/ymeth

Soluble nanopolymer-based phosphoproteomics for studying protein phosphatase

Minjie Guo, Jacob Galan, W. Andy Tao *

Departments of Biochemistry, Chemistry, and Medicinal Chemistry & Molecular Pharmacology, Purdue University, West Lafayette, IN 47907, USA

Accepted 24 February 2007

Abstract

Protein phosphorylation is a vital reversible post-translational modification that regulates protein–protein interactions, enzymatic activity, subcellular localization, complex formation and protein stability. The emerging field of mass spectrometry-based proteomics allows us to investigate phosphorylation and dephosphorylation on a global scale. In this review, we describe a new strategy based on soluble nanopolymers that have been used to selectively isolate phosphopeptides for mass spectrometric analysis. Functionalized soluble nanopolymers provide a homogeneous environment and linear reaction kinetics for chemical derivatization to isolate phosphopeptides with high specificity. Combined with phosphatase inhibitors and stable isotopic labeling, the approach has the capability of quantitatively measuring phosphorylation and dephosphorylation on individual sites. We provide experimental details for the approach and describe some other complementary techniques that can be used.

Keywords: Quantitative proteomics; Mass spectrometry; Nanopolymer; Phosphatase inhibitor; Database search; Isotopic labeling

1. Introduction

Reversible phosphorylation of proteins catalyzed by kinases and phosphatases plays a pivotal role in the regulation of important cellular functions such as growth, metabolism, division, and signaling. Under normal conditions, the coordinated and proper balance of kinase and phosphatase activity is tightly controlled. Any factor that obviates the balance and results in either an excessive or diminished substrate phosphorylation can cause a variety of human diseases such as cancer, immune diseases, and diabetes. It is estimated that there are hundreds of protein kinases/phosphatases differing in their substrate specificities, stoichiometry, cellular localization, and association with regulatory pathways. For instance, the complete genomic DNA sequence of *Saccharomyces cerevisiae* predicted 123 different protein kinases and 40 protein phosphatases that could be expressed. Thus, approximately 2% of the proteins expressed in yeast are involved in protein phosphorylation reactions, and presumably a much larger number of proteins are phosphorylated under specific physiological conditions [1]. The completion of genome-sequencing projects has also advanced the development of proteomics techniques, providing new tools for the global analysis of protein phosphorylation. A comprehensive study of protein phosphorylation involves: (i) the identification of phosphoproteins and sites of phosphorylation; (ii) the identification of the kinase(s) and phosphatase(s) responsible for reversible phosphorylation and dephosphorylation; (iii) the understanding of the biological consequence of the observed phosphorylation events.

Mass spectrometry (MS) has emerged as the method of choice for phosphoproteomics [2,3]. However, multiple factors can complicate the analysis of protein phosphorylation in complex mixtures by MS-based methods. First, the stoichiometry of phosphorylation is frequently low, and only fractions of expressed proteins may be phosphorylated at any given time. Second, a specific phosphoprotein may

* Corresponding author. Fax: +1 765 494 7897.
E-mail address: watao@purdue.edu (W.A. Tao).

1046-2023/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2007.02.019

exist in several differentially phosphorylated forms and the state of phosphorylation may be dynamically changing with changing states of the cell. In particular, many signaling molecules are present at extremely low abundance. Third, with the exception of tyrosine phosphate, the phosphate bonds range from labile to very labile. Therefore, specific precautions have to be taken to prevent the elimination of these phosphates during sample preparation. Specifically, *N*-phosphorylation and acyl-phosphorylation is extremely acid labile, while *O*-phosphorylation is relatively base labile. Forth and finally, chemical lability of the phosphate group on amino acid residues induced by collisions in the gas phase also has negative influence on database-searching results for protein/peptide identification.

While specific proteomic strategies for the investigation of protein phosphorylation keep evolving, there is a consensus on the general approach to study protein phosphorylation or dephosphorylation in complex samples. Fig. 1 schematically illustrates the proteomic strategy to study protein phosphatases. In the first step, cells are treated with phosphatase inhibitor before lysis. In the second step, phosphoproteins are enriched. This step is useful but not absolutely necessary unless proteins of very low abundance are being analyzed. In the third step, peptides are generated by chemical or enzymatic cleavage from the protein mixture. In the forth step, phosphopeptides are selectively isolated to identify sites of phosphorylation from complex samples. Finally, the phosphopeptides are analyzed by MS using specific data acquisition and database search protocols optimized for sequence and phosphorylation site determination. For a parallel comparison, the control will be processed similarly except that there is no treatment of phosphatase inhibitor in the first step. Differential labeling with stable isotopes can be performed on either the protein or peptide stage. The step to efficiently isolate phosphopeptides is critical and therefore it is the main focus in this review. We will briefly describe methods for other steps as well.

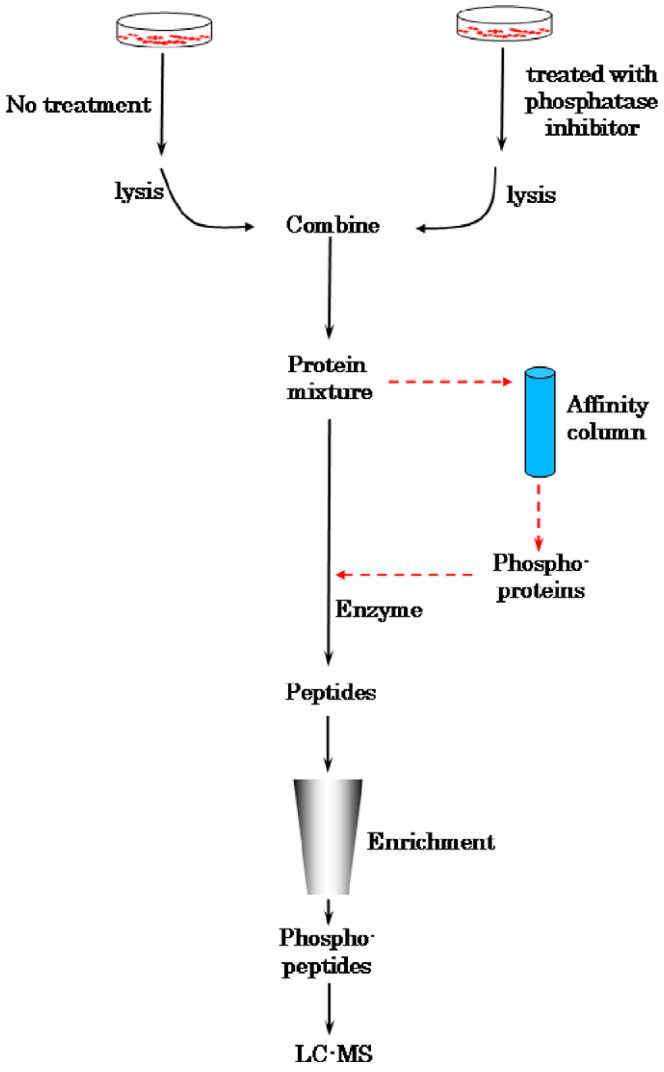

Fig. 1. A generic protocol for phosphoproteomic analysis. See text for details.

2. Isolation of phosphopeptides

Since phosphorylation is frequently a low stoichiometry event in a complex peptide mixture, nonphosphopeptides are the dominant species. In order to identify the sites of phosphorylation, it is essential to have an efficient strategy for the isolation of actual phosphopeptides. Several approaches have been explored for the selective isolation of phosphopeptides, the most notable of these being either affinity- or chemical derivatization-based. Antibodies have been used to enrich phosphopeptides but the efficiency of this affinity steps is relatively low [4,5]. Therefore, immobilized metal ion affinity chromatography (IMAC) has developed into the leading method to affinity enrich phosphopeptides [5–11]. Steady steps have been taken to improve its specificity [7,12]. However, the method appears to be highly dependent on the type of resin and pH condition for binding and elution, and prefers peptides with multiple phosphorylation sites [13]. More recent applications employed TiO_2 as the stationary phase and the results indicated increased specificity for phosphopeptides [14]. The negative charge carried by phosphate groups on phosphopeptides has also been adopted to enrich phosphopeptides using cation exchange chromatography [15]. Because a phosphate group maintains a negative charge at acidic pH values, the net charge state of a tryptic phosphopeptide is generally only 1+. Therefore on the ion exchange chromatography many phosphopeptides elute before non-phosphopeptides which are multiply charged. It is obvious that this strategy does not have high specificity and cannot recover all phosphopeptides either. However, the attraction of the method is that it can be fully automated and is particularly suitable for large scale experiments. Chemical modifications of phosphate groups are more favorable for the specificity but typically involve multiple derivatization steps, resulting in low yield for the analysis of phos-