Journal of Pharmaceutical and Biomedical Analysis xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Large-scale profiling of protein kinases for cellular signaling studies by mass spectrometry and other techniques

Naoyuki Sugiyama, Yasushi Ishihama*

Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida-shimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form 26 May 2016
Accepted 27 May 2016
Available online xxx

Keyword:
Proteomics
Protein kinase
Mass spectrometry
Signal transduction

ABSTRACT

Protein phosphorylation mediated by protein kinases plays key roles in cellular signal transduction in both eukaryotic and prokaryotic organisms. Many researches have been conducted to investigate their functions and regulatory mechanisms, but these were mostly focused on particular kinases related to well-known signaling pathways, drug targets, diseases, and so on. Therefore our overall understanding of the whole set of kinases (the kinome), including their structural basis for substrate recognition, is still limited. In this review, we describe technologies for determination of sequence preference and physiological substrates of individual target kinases and for kinome-scale profiling, focusing on peptide array technology or proteomic approaches based on liquid chromatography-mass spectrometry (LC–MS).

1. Introduction

Reversible phosphorylation of cellular proteins, mediated by protein kinases and phosphatases, is universally utilized in eukaryotes, as well as prokaryotes, and indeed, is one of the most intensively studied post-translational modifications [1]. Protein kinases catalyze transfer of γ-phosphate from ATP to serine, threonine and tyrosine residues of substrate proteins in eukaryotes, and to histidine, aspartic acid, lysine and arginine in addition to serine/threonine/tyrosine in prokaryotes [2]. Phosphorylation causes a change of higher-order structure of proteins, resulting in functional change of the targeted protein, such as switching on/off of enzymatic activity and molecular recognition ability. The functions of protein kinases themselves are also regulated via phosphorylation by other kinases or/and auto-phosphorylation [3–5]. Therefore, phosphorylation networks between kinases and substrate proteins play key roles in cellular signal transduction pathways that regulate a wide range of cellular functions, including cell growth, division, apoptosis, and mitosis [6,7].

Abnormal signal transduction resulting from dysregulation of protein phosphorylation is associated with various diseases, including cancer [6], diabetes [8], CNS disorders [9] and so on. Consequently, many drugs that inhibit a specific kinase or kinases have been developed for molecular-targeting therapy. To date, more than 19,000 kinase inhibitors targeting about 260 protein kinases have been reported [10], and about 30 small-molecular kinase inhibitors have been approved for clinical use by the US Food and Drug Administration (FDA) [11]. Although these kinase inhibitors are therapeutically useful, they sometimes exhibit side-effects, or it turns out that some patients are unresponsive to them [12–14]. One reason for this may be that the overall effects of these kinase inhibitors on global cellular signaling pathways still remain unclear, even though various new technologies such as chemical proteomic approaches have been employed to characterize them [15,16].

To obtain an overview of the whole signaling network in a cell is important for elucidation of the global regulation mechanisms of cellular events, as well as for drug development, therapy and diagnosis. Human genome sequencing analysis has revealed that at least 518 genes encode protein kinases [17], and recent large-scale analyses of protein phosphorylation events (phosphoproteomics) indicated that more than 75% of all human proteins are phosphorylated by these kinases [18,19]. Thus, each protein kinase phosphorylates hundreds of substrates on average, and the resultant phosphorylation networks must be extremely complex.

Since the early 2000's, large-scale identification of phosphorylated proteins based on mass spectrometry, namely phosphoproteomics, has been utilized in many studies aimed at large-scale analysis of cellular signaling. The comprehensiveness of phosphoproteome analysis has been dramatically improved in the past decade due to enormous technical advances in mass spectrometry and selective phosphopeptide enrichment; examples include immobilized metal ion affinity chromatography (IMAC) [20,21], metal oxide chromatography [22–26], immunoprecipita-

* Corresponding author.
E-mail address: yishiham@pharm.kyoto-u.ac.jp (Y. Ishihama).

http://dx.doi.org/10.1016/j.jpba.2016.05.046
0731-7085/

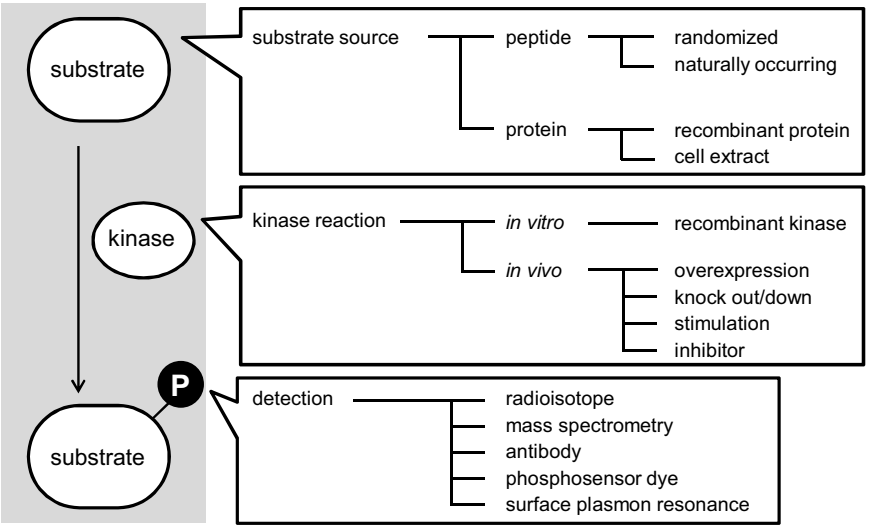

Fig. 1. Methods for profiling of kinases.
In a kinase profiling experiment, substrate source was reacted by a target kinase, and then phosphorylated substrates were detected (and sequenced) by an analytical technique. The typical methods and materials were summarized in this figure.

tion using antibodies such as anti-phosphotyrosine (pY) antibody [27], phosphopeptide-specific detection with anti-pY antibody or phostag [28], and phosphopeptide-specific identification using neutral loss-triggered MS/MS/MS [29]. These techniques have made it possible to utilize phosphoproteomics to elucidate major signal transduction mechanisms, including the epidermal growth factor receptor pathway [30], regulation of the cell cycle [30,31], embryonic stem cell differentiation [32], mammalian target of rapamycin (mTOR) signaling pathway [33,34] and tissue-specific phosphorylation events [35].

Although MS analysis can now identify up to about 10,000 sites in a single experiment and numerous recent studies have led to extensive catalogs of phosphorylated substrates, the identities of the kinases that phosphorylate these sites remain largely unknown. In addition, even at the protein level, only 40% of phosphoproteins can be mapped to all pathways in the KEGG database [36], and in most cases the functional effect of phosphorylation also remains unknown.

Generally each kinase has some preference for a particular substrate sequence, i.e., a phosphorylation motif, and this has been used to predict kinase-substrate relationships. Kreegipuu et al. investigated the sequence specificity of 58 protein kinases using 1008 phosphorylation sites in 406 substrate proteins, which were obtained from the literature and public databases [37]. They found that sequence specificity was not stringent, but was determined by the residues located within the −4 to +4 positions relative to the phosphosite, and most phosphosites were located on the surface of substrate proteins. Based on limited *in vitro* kinase—synthetic peptide datasets combined with computational approaches, tools to predict kinases from substrates or substrates from kinases, such as Scansite [38,39] Predikin [40,41] and NetPhorest [42], have been developed. Protein-protein interaction (PPI) databases such as STRING [43] have also been utilized to predict signaling networks for large-scale phosphoproteome analysis [31,35,44]. NetworKIN can predict a responsible kinase from a queried phosphosite dataset by using a combination of phosphorylation motifs, PPI information and a pathway database [45], and the predicted kinase-substrate information has been used to estimate regulatory kinases for observed phosphorylation events [18,46]. Although many computational approaches have been reported to predict kinase-substrate relationships [47], a common bottleneck is the lack of an adequate experimental kinase-substrate dataset with extracted phosphorylation motifs.

Here, we focus on approaches for determination of sequence preference, *in vitro* substrates and physiological substrates of a target kinase using peptide libraries, peptide arrays, mass spectrometry and combinations of these methods (Fig. 1). We also describe kinome-scale profiling using peptide arrays or phosphoproteomics based on LC–MS.

2. Peptide libraries

To explore kinase substrates, inhibition of a specific kinase *in vivo* using knockout/knockdown or a kinase inhibitor with known specificity has been utilized. However, it is difficult to distinguish true substrates that are directly phosphorylated by a kinase of interest from downstream molecules or phosphorylated by another kinase(s) [48]. Inhibition of a target kinase may cause serial down-regulation of downstream kinases, so that phosphorylation of their substrates is also decreased as a result.

Although most protein kinases require high-order structure for substrate recognition under physiological conditions, and substrate selection may be influenced by cellular localization and the existence of scaffold protein(s) [49], phosphoacceptor preferences are relatively well defined, as mentioned above. Although sequence preference for structure adjacent to the phosphorylated residue is insufficient to fully account for the substrate recognition mechanisms of a kinase, these characteristics are useful to predict substrate proteins or responsible kinases in combination with other experimental data and/or computational tools. *In vitro* kinase profiling using peptide libraries allows high-throughput profiling of sequence preference to be conducted more easily and widely than can be done with *in vivo* studies.

2.1. Random peptide libraries

Phosphorylation motifs of cyclic AMP-dependent kinase (PKA) were investigated using a randomized synthetic peptide library [50,51]. Wu et al. screened substrates of PKA using synthetic peptide libraries containing degenerate penta- and heptapeptides that were individually expressed on solid-phase beads and phosphorylated *in vitro* by the kinase in the presence of [γ-^{32}P]ATP [50]. Finally, the beads expressing phosphorylated peptides were

Please cite this article in press as: N. Sugiyama, Y. Ishihama, Large-scale profiling of protein kinases for cellular signaling studies by mass spectrometry and other techniques, J. Pharm. Biomed. Anal. (2016), http://dx.doi.org/10.1016/j.jpba.2016.05.046