ELSEVIER



# Phosphoproteomics in drug discovery

**Melody K. Morris[1], An Chi[1], Ioannis N. Melas[2,3] and Leonidas G. Alexopoulos[2,3]**

[1] Merck & Co., Boston, MA, USA
[2] ProtATonce Ltd, Athens, Greece
[3] Department of Mechanical Engineering, National Technical University of Athens, Athens, Greece

**Several important aspects of the drug discovery process, including target identification, mechanism of action determination and biomarker identification as well as drug repositioning, require complete understanding of the effects of drugs on protein phosphorylation in relevant biological systems. Novel high-throughput phosphoproteomic technologies can be employed to measure these phosphorylation events. In this review, we describe the advantages and limitations of state-of-the-art phosphoproteomic approaches such as mass spectrometry and antibody-based technologies in terms of sample and data throughput as well as data quality. We then discuss how datasets from each technology can be analyzed and how the results can be and have been applied to advance different aspects of the drug discovery process.**

## Introduction

The pharmaceutical industry is tasked with delivering drugs of high efficacy and low toxicity. The drug discovery pipeline needs to become fast and effective in response to the industry-wide challenge regarding a low first-in-human to registration rate and pressure from strict regulatory requirements, budget cuts in the healthcare system, vigilant patient foundations and time constrains caused by patent expiry. Acquiring a better understanding of how drugs target cells in the human body is of the utmost importance for increasing drug development efficiency and decreasing high attrition rates because success rates from first-in-human studies to registration are only 11% [1].

Although not yet proven, high-throughput phosphoproteomic technologies together with well-established pharmacogenomic and pharmacogenetic measurements hold promise for improving the drug discovery process because phosphorylation events are proximal to many disease-causing signaling mechanisms. For example, dysregulated phosphosignaling caused by mutations is a known driving mechanism of several types of cancer, and inhibitors of kinases such as Raf (e.g. sorafenib), anaplastic lymphoma kinase (ALK; e.g. crizotinib) and epidermal growth factor receptor (EGFR; e.g. erlotinib) have proven efficacy in the treatment of a variety of cancers. Additionally, kinase inhibition has proved to be an effective method for inhibiting activation of immune cells important in autoimmune disease [e.g. Janus kinase (JAK) inhibition for the treatment of rheumatoid arthritis]. By understanding the phosphosignaling networks underlying aberrant growth in cancer or immune cell activation in autoimmunity, drug developers can choose better targets and better understand how therapeutics will alter cellular processes.

Several phosphoproteomic technologies have emerged in recent years that have become invaluable tools for drug discovery and biomarker development. In this review, we discuss how modern phosphoproteomic technologies can be used to aid target identification, understand drug mechanism of action (MOA), construct signaling pathways, predict toxicity and/or efficacy, as well as their involvement in drug repositioning when coupled with appropriate computational analyses. For each technology we describe its distinct advantages and inherent limitations. We then provide examples of how these technologies have been applied to enhance the drug discovery process.

*Corresponding author.* Alexopoulos, L.G. (leo@protATonce.com)

1359-6446/06/$ - see front matter  http://dx.doi.org/10.1016/j.drudis.2013.10.010

## Phosphoproteomic technologies

Protein measurements can be divided into two distinct categories: those that make no a priori assumption about proteins to be measured [i.e. 2D-PAGE and mass spectrometry (MS) technology]; and those that are based on a pre-determined set of measured proteins (i.e. antibody- or aptamer-based approaches). In the first category, MS approaches promise an unbiased (i.e. hypothesis free) screening of thousands of phosphoprotein targets. In the second category, antibody-based technologies exist in several different formats and varieties that promise a throughput of thousands of samples per day and improved quantification on pre-determined phosphosignals. Despite the tremendous improvements in both technologies, inherent limitations exist: MS-based technologies offer great coverage of the phosphoproteome but require significant sample preparation and data post-processing time, leading to limited sample throughput; by contrast, antibody-based formats are restricted by limited antibody quality and target availability. All commonly used technologies are presented in Fig. 1 and described below.

### *Forward- and reverse-phase protein microarrays for phosphoprotein measurements*

A protein microarray consists of a solid surface, typically a glass slide or membrane, on top of which antibodies, aptamers, purified proteins or cell lysates are spotted and then probed with molecules interacting with the spots [2–5]. Depending on the immobilized and probing molecules, two types of phosphoprotein microarray are the most prominent: (i) forward-phase protein microarrays (FPPAs – also known as antibody microarrays) with immobilized phospho-antibodies and (ii) reverse-phase protein microarrays (RPPAs) with immobilized samples.

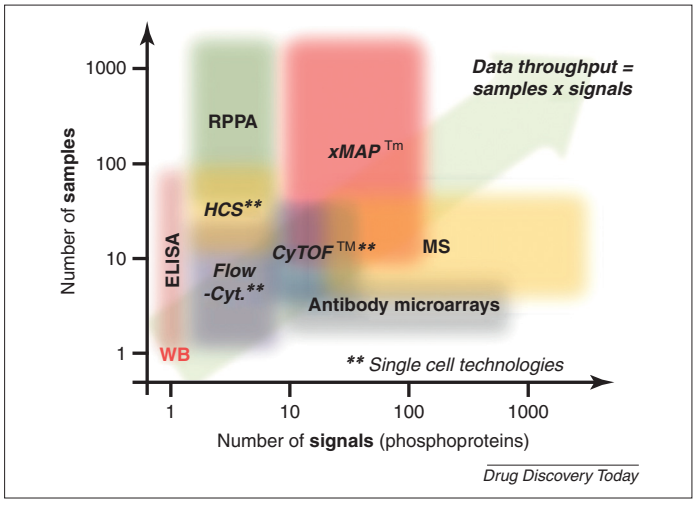


**FIGURE 1**

State-of-the-art phosphoproteomic platforms for high-throughput measurements of phosphosignaling. We illustrate the difference in sample and signal throughput of each technology by plotting number of samples readily assayed by number of phosphorylation sites (phosphosignals) readily measured. Reverse-phase protein arrays (RPPA) utilize a single-antibody format to probe thousands of lysate spots. Protein arrays follow a regular sandwich ELISA format where lysates are incubated with dozens of antibodies. High-content screening (HCS) is a single-cell technology that usually employs an automated fluorescent microscope capable of visualizing 96- or 384-well plates. Fluorescence-based flow cytometry (Flow-Cyt.) can detect fewer than 12 phosphoproteins at a single-cell level. Mass cytometry (CyTOF) offers greater multiplexability for single-cell analysis via transition metal ion tags. Mass spectrometry (MS) utilizes affinity phosphorylaton enrichment approaches with high-performance MS instrumentations capable of identifying thousands phosphorylation sites. xMAP® is a suspension ELISA microarray on microbeads that utilizes a dual-antibody format where thousands of samples can be probed in a single day with dozens of antibodies. Western blot (WB) and ELISAs are the standard low-throughput assays that have been mapped for comparison reasons.

Antibody microarrays are ELISA-type assays where the 'sandwich' antibody format enables high fidelity measurements of several intracellular signals on a limited number of samples. Several vendors offer ready-to-use arrays. For example, R&D Systems offers an array for detecting 43 phosphoproteins, PathScan® from Cell Signaling can measure 18 phosphosignals from a single lysate and the RayBioTech Phosphorylation Array (RayBiotech Inc., Norcross GA, USA) can detect the activation of 71 tyrosine kinases. In all cases, a handful of samples can be measured by a simple incubation of samples with the arrays. During incubation, phosphoproteins in the lysate are pulled down by the capture antibodies on the spots, and a secondary antibody that carries a detection motif (fluorophore, Horseradish peroxidase, etc.) binds to a different epitope of the phosphoprotein and produces a quantifiable signal (i.e. fluorescent intensity). In most cases, an image analysis algorithm is required to measure the intensity of the spot and thus quantify the phosphoprotein. The spotting procedure together with the imaging step can introduce some artifacts on the signal quantification (i.e. spot evaporation, donut shape spot, uneven spotting, 'growing' spot size, etc.). Major bottlenecks for establishing high-throughput assays are sample preparation and handling and image analysis procedures that usually require a manual curation step. Recent developments have established protein arrays in a 96-well plate format (i.e. Meso Scale Discovery, Rockville, MD USA) but at the expense of multiplexability (usually less than ~6 signals are measured per sample).

In general, phosphoprotein antibody microarrays are a well-established technique with several products available to buy off the shelf. Owing to the 'sandwich' antibody format, antibody microarrays are ideal for high-quality measurements of several phosphoproteins. However, their total data throughput (samples × signals per time for assay run) is facing strong competition from bead-suspension systems that are based on the same protein detection principle, are compatible with 96-well plate automation and offer similar multiplexability.

RPPAs bypass the need for antibody pairs by direct printing (spotting) of the lysate onto a functionalized glass or membrane. The single-antibody detection format enables a large variety of phopshoproteins to be measured, and multiple slides can be spotted with the same lysate and probed with different antibodies to enable analysis of multiple phosphoproteins in the same samples. Several hundred samples can be spotted on a single slide but a very limited number of proteins can be measured on each slide. Compared with ready-to-use protein arrays, RPPAs require significant technical expertise for protein spotting and protein detection that can be established in-house or out-sourced. In addition, the single-antibody detection format is prone to low data quality as a result of higher background noise and antibody specificity. By contrast, the sample throughput of RPPAs is unmatched compared to any other known technology. When thousands of samples need to be measured, RPPAs offer by far the lowest cost per data point, but at the expense of data quality and multiplexability.