Mining the Tumor Phosphoproteome for Cancer Markers

Yoon Pin Lim

Abstract
Despite decades of cancer research, mortality rates remain high largely due to the failure of early detection, poor understanding of the epidemiology of rational drug targets, and molecular etiology of human cancers. The discovery of disease markers promises to deliver some solutions to these formidable challenges. Gene and protein expression profiling through DNA microarray and proteomics have already made a tremendous effect in this area. However, protein/gene expression does not necessarily reflect protein activity, which is often regulated via post-translation modifications, of which phosphorylation is one of the most prominent. This is an important consideration because the activity of protein is a more relevant phenotype than its expression during pathogenesis. Tyrosine kinases represent a very important class of enzymes that are critical regulators of mitogenic and angiogenic signaling, hence attractive targets for anticancer drugs as exemplified by BCR-ABL and ErbB2. More than 50% of them are overexpressed or mutated resulting in a gain of function in various human cancers. In this review, we discuss the potential effect of phosphoproteins as cancer markers in cancer diagnosis and therapeutics. Phosphoproteomics strategies that might pave the way to high-throughput analysis for routine clinical applications are also described.

The quest for cancer markers has led to the discovery of existing markers such as prostate-specific antigen, carcinoembryonic antigen, α-fetoprotein, and CA125 in the 20th century (1). However, current biomarkers often lack specificity and sensitivity, leading to false results and late detection of malignancies. With the completion of the human genome project and the advent of DNA microarray and proteomics technologies, expression profiling of multitudes of genes and proteins has begun to affect the field of cancer markers with the goal of achieving better markers (single or in combination). In the 21st century, we expect to see increasing research efforts on the global study of protein post-translational modifications like phosphorylation, degradation, and glycosylation and their relationship to cancer. Phosphorylation, in particular tyrosine phosphorylation, features prominently in the biology of cancer. Apart from being attractive drug targets, tyrosine kinases and their substrates may be useful molecular cancer markers that are invaluable for their potential discriminatory power in molecular classification of cancers (diagnosis), in predicting clinical outcome (prognosis) and response to drugs (therapy). Whereas the serine/threonine kinases represent another group of proteins important to cancer etiology, leading to false results and late detection of malignancies.

Growth factors, tyrosine kinases, and cancer biology. The process of tumor initiation, progression, and metastasis are multifactorial and highly complex. Growth factors are critical components in cancer biology. Epidermal growth factor (EGF) and platelet-derived growth factor are potent mitogens (4), whereas vascular endothelial growth factor and fibroblast growth factor have angiogenic properties (5). Growth factors stimulate a family of cell-surface receptors called receptor tyrosine kinases with intrinsic tyrosine kinase activity. Tyrosine phosphorylation is a key regulatory switch that controls biochemical cascades leading to cellular events such as growth and proliferation (6). It does so by controlling enzymatic activities, protein stability, protein-protein interactions, subcellular translocation, and transcriptional activities of proteins (3).

There are about 518 putative protein kinase genes identified in the human genome (7) and about 100 of these are tyrosine kinases (8). What is remarkable is that >50% of all tyrosine kinases have been implicated in human cancers as a result of aberrant activities arising from gain-of-function mutations such as deletion, translocation and viral insertion and/or overexpression. Examples include BCR-ABL tyrosine kinase in chronic myelogenous leukemia (CML), ErB2 receptor tyrosine kinase in breast cancers, platelet-derived growth factor receptor tyrosine kinase in gliomas, Src nonreceptor tyrosine kinase in colorectal cancers, c-Met receptor tyrosine kinase in gastric cancers, and focal adhesion kinases in invasive human tumors (3).

Tyrosine kinases as rational drug targets. Given the critical roles of tyrosine kinases in cancers, researchers and drug companies have for a long time envisaged the arrival of target specific drugs or magical bullets against these signaling molecules. In recent years, Herceptin (trastuzumab) against ErB2 receptor tyrosine kinase for the treatment of metastatic breast cancer and Gleevec (imatinib mesylate) against BCR-ABL for the therapy of patients with CML and gastrointestinal...
stroma tumors emerged as early examples of protein-based cancer drugs against signal transduction molecules (9). More recently, Iressa (gefitinib, AstraZeneca) has been approved by for chemotherapy-recalcitrant non–small cell lung cancer (NSCLC) by the Food and Drug Administration (10). These agents not only prolong and improve the quality of lives they also provided clinical validation of rational drug design.

Gleevec, in particular, is a star performer. Its impressive results during phase II trials with 532 patients with CML, achieving a 95% response, all of whom had previously failed standard IFN therapy, has led to its fast-track approval by Food and Drug Administration (11). A new wave of "smart drugs" are currently making its way to the market and they include Avastin (Genentech, South San Francisco, CA) and vascular endothelial growth factor trap (Regeneron, Tarrytown, NY), both with antiangiogenic properties against vascular endothelial growth factor for colorectal cancer and non-Hodgkins lymphoma, respectively; ABX-EGF (Abgenix, Fremont, CA) and Tarceva (Genentech) against EGF receptor (EGFR) for NSCLC (12). The mode of action of these drugs includes inhibition of ATP binding to the kinase domain, ligand-mediated receptor modulation (endocytosis and/or degradation) and trapping of receptor or ligand (9).

Although some difficulties have been encountered by some of these drugs (e.g., the relapse of patients seen in late-stage CML as a result of mutations in BCR-ABL resulting in drug resistance; refs. 13–18), the stage is set for more demonstrations of the power of target-directed therapy development, as is obvious from the emergence of drugs against Gleevec-resistant CML as a result of mutations in BCR-ABL resulting in drug resistance, and the recent reports of target-specific drugs against insulin-like growth factor-I receptor tyrosine kinase as having significant antitumor activity (19–22).
Phosphorylation Isoforms in Fingerprinting of Molecular Pathways

Phosphoproteomics potentially provides two additional layers of information compared with conventional expression profiling studies. First, it provides clues on what protein or pathway might be activated because a change in the phosphorylation status of proteins almost always reflects a change in protein activity. Second, it indicates what proteins might be potential drug targets because phosphoproteomics focuses on proteins that have kinase activities or are substrates of kinases. Thus, there is increasing interests in phosphoproteomics as reflected by a considerable number of works describing various strategies including the use of different isotopic technologies (e.g., isotope-coded affinity tag, stable isotope labeling by amino acids in cell culture), protein separation techniques (e.g., gel-based versus liquid chromatography–based methods), and chemical modifications (e.g., β-elimination and esterification) to identify phosphoproteins and/or phosphorylation sites in a global fashion (58–67). Recently, our group developed a method that combines the capability of anti-phosphotyrosine 4G10 antibodies (68) to enrich for tyrosine-phosphorylated proteins, the high-resolution power of two-dimensional PAGE to separate proteins and the high sensitivity of mass spectrometry to identify proteins of interest. The tool enabled the concurrent visualization of >50 protein spots in a single phosphoproteome.
map belonging to about 20 proteins which can be broadly categorized into five functional classes of signaling proteins activated by EGF in the A431 human cancer cell line (69). They were the receptor tyrosine kinases (e.g., EGFR), SH2-containing adaptor proteins [e.g., growth factor receptor binding protein 2, p85 PI3-K, Src homology and collagen protein], SH2-containing signaling enzymes (e.g., PLC-γ), heat shock proteins (e.g., Hsc71), and cytoskeletal proteins (e.g., tubulin and γ-catenin), all of which have pro-mitogenic or anti-apoptotic properties. The potential application of this valuable tool is tremendous, ranging from the fingerprinting of growth factor signaling pathways to drug screening and to comparative studies between normal and tumor tissues. However, the current sensitivity of this gel-based tool limits its success to only in vitro studies and we are attempting to achieve greater sensitivity for tissue analysis by adapting liquid chromatography–based mass spectrometry into this approach.

Whereas general profiling of phosphorylation status in proteins offers useful information, it is possible that proteins are tyrosine phosphorylated in dissimilar fashions in terms of the sites and degree of phosphorylation in different states of disease and between subjects. Our recent study reveals that of the 20 or so proteins mapped on two-dimensional gels, about nine or 45% of them have varying numbers of tyrosine-phosphorylated isoforms within the same signaling system (EGF) and between two different signaling systems (EGF versus platelet-derived growth factor; Table 1; ref. 69). These differences reflected the predominant functional effects of the individual signaling system. Platelet-derived growth factor induced a more potent PI3-K/actin–mediated chemotactic pathway relevant to tumor metastasis, whereas EGF induced a more prominent focal adhesion pathway involving γ-catenin. Similarly, phosphorylation isoforms exist for proteins that are phosphorylated at serine and threonine (70).

Will phosphorylation isoforms have important bearing on the causative nature of oncoproteins? The EGFR receptor is known to be phosphorylated at multiple sites upon activation. Each phosphorylated tyrosine residue on the receptor is a potential docking site for downstream signaling proteins, thereby triggering different effector pathways. It has been shown that c-Src phosphorylates Y845 in the activation loop of the EGFR and maintains the enzyme in an active state (56). Phosphorylated Y992 and Y1068 of EGFR are direct binding sites for the SH2 domain of PLC-γ and growth factor receptor binding protein 2, respectively, and these interactions stimulate the mitogen-activated protein kinase cascade critical for growth and proliferation in the cells (71). On the other hand, phosphorylation of Y1045 creates a major docking site for c-Cbl, the binding of which leads to receptor ubiquitinylation and degradation (71). Examples of the interaction of other cytosolic signaling proteins with various receptor tyrosine kinases have been reviewed elsewhere (72). Some questions
that beg to be answered are whether all the sites are phosphorylated in an all-or-none fashion or are these phosphorylations hierarchical and temporal in nature and whether subpopulations of receptor with distinct phosphorylation sites exist. For the non-receptor tyrosine kinase, c-Src is reciprocally phosphorylated at Y527 and Y416 when it is activated and inactivated, respectively (73). In the case of receptor tyrosine kinases where multiple phosphorylation sites are involved, the temporal/spatial stoichiometry of phosphorylation events may be more complicated. Nevertheless, it is obvious from the above examples that site-specific phosphorylation has important functional implication in terms of the molecular pathways activated and needs to be more thoroughly understood with respect to tumor biology and cancer therapeutics.

Technical Challenges in Clinical Applications of Phosphoproteomics

The prerequisite features for the translation of protein-based techniques into routine clinical use include high throughput and high sensitivity because critical decisions have to be made within the shortest possible time from minute quantities of patients’ samples. Whereas liquid chromatography–based mass spectrometry is instrumental in mapping of the phosphorylation sites on different cancer-causing tyrosine kinases and/or their substrates, it is tedious and is not amenable for high-throughput analysis of tissues. A viable strategy towards the fine scale mapping of the phosphorylation status of tyrosine sites in biomarkers is to exploit the experimental data from liquid chromatography–based mass spectrometry analysis to generate phosphorylation site-specific antibodies, which could be used as reagents in the development of antibody array/chip or for immunohistochemistry on tissue microarrays (TMA) in molecular pathology. The establishment of a methodology for raising phosphorylation site-specific antibodies (74) and the availability of a vast arsenal of phosphorylation site-specific antibodies (70) have kick-started pilot-scale studies employing

![Fig. 3 Phosphoproteomic identification of proteins involved in clinical resistance to anticancer drugs. A, activation of PKC, MAPK, and AKT pathways by EGF upon binding of EGF. B, inhibition of EGF-induced activation of various signaling pathways by Iressa. C, potential genetic alterations during cancer formation that might confer resistance to Iressa are color coded, indicating the frequency of overexpression or mutation in tumors or cancer cell lines derived from primary tumors. Yellow, high frequency and includes PLC-γ (83–85), Grb2 (86, 87), and ras (88, 89). Green, sporadic frequency and includes GAB2 (90), P13-kinase (91), AKT (91), mitogen-activated protein kinase (MAPK; ref. 92), and PKC (93). Blue, unknown occurrence. Scenarios A and B will generate specific phosphoproteome 1 and 2, respectively. Depending on the type/s of genetic changes in scenario C, different phosphoproteomes will be produced. Comparative analysis of these phosphoproteomes would allow drug resistance markers to be identified. Due to space constraint, activated EGFR is depicted as a monomer instead of a dimer.](www.aacrjournals.org)
these reagents for tumor analysis (75–78). One very recent and interesting application of this phosphorylation site–specific antibodies was in single cell profiling of signaling networks using multivariable flow cytometry (79). However, current literature reporting the use of phosphorylation site–specific antibodies in tissues is restricted to the concurrent investigation of at most a limited number of target proteins such as ErbB2, p38, mitogen-activated protein kinase, and signal transducers and activators of transcriptions. It is becoming increasing evident that multiple markers is of better utility than single marker in cancer prognosis and diagnosis (80). Hence, arrays of phosphospecific site–specific antibodies against multiple targets are likely to emerge in greater visibility in the next decade.

It is obvious that one of the key bottlenecks of such a challenge lies in the cost-effective production of large quantities of a wide range of highly purified and phosphorylation site–specific antibodies. High-density antibody array remains difficult to construct and its current density of about $10^3/cm^2$ is considerably lesser than that of $10^7/cm^2$ in DNA microarray. This means that a larger amount of samples would be needed for protein analysis than for DNA analysis. Another major hurdle is the concoction of a condition whereby all antibody-antigen interactions could take place optimally. This is apparently impossible and is a major drawback of antibody array. Thus, a considerable amount of technical inertia has to be overcome before a high-throughput antibody/protein chip could be clinically applied in a routine basis. Whereas the Atpamer technology is making inroads to provide some solutions to this problem (81), semi-high throughput TMAs may be the first platform through which routine clinical application of phosphoantibodies could be feasibly achieved. Tissue microarrays consisting of up to 1,000 tissues can now be constructed and immunohistochemistry on TMAs is now routinely done (82). Application of the phosphorylation site–specific antibodies on TMA and correlation of the phosphorylation status of proteins at specific tyrosine residues with clinical data is likely to lead to the discovery of potential biomarkers. Current efforts are being made to construct TMAs from frozen tissue blocks, which will be especially useful for protein analysis because better preservation of protein antigens is expected than from paraffin blocks.

**Conclusion**

Whereas technology development is a critical component, what is equally important, if not more important, is the definition of the molecular signature of various cancers through biomarker discovery. In this respect, much effort remains necessary to study aberrant signaling pathways in cancer cells from tumors of different origins in a global fashion. The discovery of cancer markers from the tumor phosphoproteomes will undoubtedly contribute to a paradigm shift towards individualized medicine. Along with this is the spinning off of basic research programs to understand the role of phosphorylation in regulating the function of potential phosphoprotein markers in cancer biology. Ironically, whereas phosphoproteomics may yield hundreds or even thousands of candidate genes/proteins, the final cancer-specific signature may comprise of only some 10 to 20 targets that are of significant discriminatory value. This bodes well for the feasibility of creating phosphoprotein-specific antibody chip, which is one of the ultimate goals for routine application of phosphoproteomics in the clinical setting.

**References**


**Clin Cancer Res 2005;11(9) May 1, 2005 3168**

www.aacrjournals.org
41. Finlay GJ. Genetics, molecular biology and colorectal cancer. Mutat Res 1993;290:3–12.
Mining the Tumor Phosphoproteome for Cancer Markers

Yoon Pin Lim


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/11/9/3163

Cited articles
This article cites 89 articles, 27 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/11/9/3163.full#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/11/9/3163.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.