

Methods to profile protein kinase substrates in cells

Abstract

Protein kinases and protein phosphorylation play an important role in almost every process within a cell. Identifying specific substrates downstream of individual protein kinases is critical to understanding their function and role in human health and diseases. Given its importance, several techniques have been developed to determine direct substrates of protein kinases in cells. Here, I have summarized the advantages and limitations of some of the most widely used current techniques available to profile substrates of protein kinases.

Volume 4 Issue 4 - 2016

Shivaprasad H Venkatesha

Department of Microbiology and Immunology, University of Maryland School of Medicine, USA

Correspondence: Shivaprasad H Venkatesha, Department of Microbiology and Immunology, University of Maryland School of Medicine, HSF-1, Suite 380, 685 W. Baltimore Street, Baltimore, MD 21201, USA, Email svenkatesha@som.umaryland.edu**Received:** September 24, 2016 | **Published:** November 15, 2016

Introduction

Reversible protein phosphorylation plays an important role in essentially every aspect of life such as metabolism, cell cycle progression, proliferation, apoptosis, differentiation, and cell migration. They are known to regulate protein function, stability, ability to interact with other proteins, and localization.¹⁻³ The class of enzymes that mediates the covalent addition of phosphate group (from ATP or GTP) to amino acid side chains in proteins is known as protein kinases. The hydroxyl groups of serine, threonine, tyrosine or histidine amino acid side chains are the most common phosphoacceptor sites on target proteins. Interestingly, aberrant protein kinase activity is observed in several human diseases including cancer. Given the central role of protein phosphorylation in cellular function, there have been tremendous efforts to understand the regulation and downstream targets of protein kinases. Understanding the direct substrates of protein kinases and designing methods to disrupt specific kinase-substrate interaction can open new avenues for therapeutic intervention in diseases where kinase activity is deregulated.^{4,5} This review summarizes some of the current available techniques that have revolutionized the global analysis of protein phosphorylation and elucidation of protein kinase substrates.

Early methods for identifying protein kinase substrates

Early efforts to identify substrates of protein kinase relied on genetic screens and epitasis analysis in model organisms.⁶ A major issue with these methods was being labor intensive and low throughput. Despite these limitations these methods helped make crucial advances by placing substrates downstream of specific protein kinases, which led to identification of signaling pathways through which they act. A slightly different approach utilized yeast-two-hybrid screens to identify *in vivo* substrates of protein kinases. In this approach, the kinase of interest was used as bait.⁷ The caveat of this approach was that most kinase-substrate interactions are transient and would hence miss being detected using such an approach. This method also fails to distinguish between protein kinase binding partners and bona fide substrates thus leading to potentially high false positive rates.⁶ Solid-phase phosphorylation screening using phage expression libraries new Erk1 and Cdk substrates was successfully identified.^{8,9} In this method, cDNA libraries were cloned into phage expression vectors and expressed proteins were purified on plaques

on lawns of *E.coli*. These proteins contained bacterial and phage proteins in addition to proteins expressed from cDNA. These proteins were immobilized on nitrocellulose filters and probed for potential substrates by phosphorylation using purified Kinase and radio labeled ATP. Plaques containing putative substrate are identified by autoradiography. Although this method enabled easy identification of substrates from cDNA clones, high background phosphorylation of bacterial and phage proteins posed a significant challenge. Further, improper folding of proteins expressed from cDNA also resulted in false positives. The labor intensive and low throughput nature of this technique made it less preferable for large-scale screening. It also does not work well in situations that require additional interacting proteins for a kinase to phosphorylate a substrate. Due, to these limitations, despite its potential this technique was less preferred.

Computer-based methods for protein kinase substrate prediction

Neural network and pattern matching algorithms provide an easy way to relate protein kinases to substrates.¹⁰ Some of the commonly used protein kinase substrate prediction algorithms are: GPS 3.0, KinasePhos 2.0, PhosphoSite.^{11,12} The confidence of the algorithms to predict substrates of protein kinases and sites likely to be phosphorylated on the target protein depend on a number of factors such as how well the network is trained. Increase in the number of known substrates for a protein kinase increased the opportunity to train the algorithm and the confidence in its prediction. Although these prediction algorithms can serve as an excellent starting point, they have several drawbacks. *In silico* analysis often over simplifies the corresponding *in vivo* context and hence these analyses are at most times far from real. Computer algorithms lack considerations for three-dimensional structure of kinase and substrate, stoichiometry, substrate recruitment, protein localization, interacting partners, and different physiological conditions prevalent in the cell. As a result, despite being easy to use, these algorithms are far from flawless and often have an alarmingly high false-positive rate.

Kinase substrate tracking and elucidation (KRESTEL)

Kinase Substrate Tracking and Elucidation (KRESTEL) is an *in vitro* global protein kinase substrate identification tool developed by Phillip Cohen and colleagues.¹³ In KESTREL, cell lysate containing

potential substrates of a particular protein kinase is incubated with the purified catalytically active kinase and radioactive ATP. The phosphorylated proteins are either enriched or directly visualized by autoradiography following SDS-PAGE. To reduce problems with ATPase activity in cell lysate and facilitate protein identification by mass spectrometry, the cell lysate is often pre-fractionated.¹³ This helps to reduce complexity of the proteome under analysis and separates potential kinase-substrate pair in the proteome. To increase signal to noise ratio, a higher concentration of purified protein kinase whose substrate is to be identified is used in the *in vitro* reaction and the reaction is usually carried out for short duration. Although this technique was a significant breakthrough and enabled identification of many novel substrates for a number of protein kinases, KRESTEL was clearly far from perfect. The central issue facing KRESTEL analysis was the high false positive rate and background phosphorylation arising from phosphorylation of proteins in the lysate by endogenous protein kinases. This concern was partly alleviated by heat-inactivating the cell lysate prior to addition of purified kinase and radioactive ATP 13. However, heat-inactivation raised the issue of denaturing proteins in the lysate and exposing residues that are usually non-accessible to protein kinases. In addition to these, KRESTEL carried the criticism common to most *in vitro* techniques such as, altered stoichiometry, absence of cellular context and interacting proteins important in imparting specificity, risk of making substrates usually inaccessible in cells readily available for phospho-transfer etc. KRESTEL was also a low throughput method.

Degenerate peptide library screen-based method

Large scale identification of protein kinase substrates have been accomplished using degenerate peptide library screens.^{6,14} In this approach a mixture of peptides of random sequences, oriented around an invariant serine, threonine or tyrosine residue is phosphorylated *in vitro* using a purified kinase. A ferric column is used to isolate and enrich the small percentage of phosphorylated peptides from a mixture of billions of peptides. The isolated phosphorylated peptide is then subjected to Edman degradation to identify the sequence and identity of the target peptide. To identify potential substrates having this target peptide motif, the obtained data is formatted into a selectivity matrix and searched against protein databases for candidate kinase-substrates using web-based programs such as Scansite.⁶ This method relies on *in vitro* kinase assay to determine and verify protein kinase substrates and hence carries caveats associated with the assay. In addition, this method also relies on data collected from short (~20 amino acid long) peptides, which lack the structural complexity observed in proteins. Another potential limitation of using the degenerate peptide library is that since the peptide synthesis is not directed, this approach cannot consciously exhaust all possible permutations of peptides possible neither does will this analysis provide valuable information regarding stoichiometry of phosphorylation of a particular site by a given protein kinase. Despite all these limitations and potential pitfalls, this tool is highly powerful and can be used to narrow the search of putative substrates of a protein kinase and is suited for use in combination with other kinase-substrate discovery tools such as immunoblotting with anti-phospho-antibodies.⁶

Chemical-genetic method to identify protein kinase substrate

Structural conservation of the ATP binding pocket of all protein kinases permits the design of mutations that would bind unnatural

ATP-analogs in addition to ATP 15. These engineered mutations usually involve replacing a bulky amino acid in the kinase's ATP binding pocket with either alanine or glycine. This creates a structural 'hole' in ATP binding pocket. Such a structural alteration would then permit the engineered kinase to recognize an ATP-analog containing a large hydrophobic moiety. These ATP analogs carrying large hydrophobic moiety do not usually bind to the ATP-binding pockets of endogenous protein kinases due to structural constraints.¹⁵ Cell lysates can be incubated with the recombinant engineered protein kinase and radio labeled ATP-analog. Since the ATP analog cannot be utilized by the endogenous kinases in the cell lysate, this would result in specific phosphorylation of the substrates by recombinant kinase specifically designed to accommodate the ATP-analog, thereby allowing identification of its direct substrates in the cell lysate.¹⁵ This method can be potentially extended to *in vivo* detection of protein kinase substrates by expressing the modified kinase in cells and treating the cells with ATP analogs.⁶ This would help overcome potential false-positive identification as a result of using cell lysate in a *in vitro* setting. The 'chemical-genetic method' can be used in conjunction with modified inhibitors that are recognized by only the ATP-binding pocket of the mutated kinase to verify *in vivo* substrates. This method has been successfully used to identify novel substrates of many protein kinases such as JNK, ERK1.⁶ The primary concern with this method has been the difficulty to deliver the bulky ATP-analogs into cells to enable *in vivo* substrate identification. In addition there are also reports suggesting recognition of the ATP-analog molecules by some endogenous kinases, thereby not offering the predicted specificity. If used to identify substrates *in vitro*, then the chemical-genetic method raises most of the generic concerns associated with most *in vitro* kinase-substrate detection technologies. Overall, the 'chemical-genetic methods' offers significantly improved signal-to-noise ratio and remains a valuable tool in detecting novel direct substrates of protein kinases.

Protein chip-based method

This method developed by Michael Snyder's involved using a protein chip containing specific proteins immobilized on a chip. These immobilized substrate array is phosphorylated by the protein kinase of choice using radio labeled ATP. Since the proteins in the array are pre-defined and their location is known, the candidate *in vitro* substrates of a particular kinase can be easily identified.¹⁶ Although high throughput, this strategy had several drawbacks. This method uses a biased approach because the immobilized substrates are pre-defined. This can be overcome by immobilizing all the proteins in the proteome on the protein chip. However this is technically challenging and labor intensive. Another concern is mis-folding of the proteins upon being immobilized on the chip. This could expose surfaces on proteins that are usually shielded and contribute to identification of false positives. Again, similar to KRESTEL and the chemical-genetic approach, this method does not consider the complexities of an *in vivo* system.

Reverse in-gel kinase assay (RIKA)

Reverse In-gel kinase assay is a novel protein kinase-substrate identification technology that enables identification of direct substrates of a protein kinase from an unbiased pool of substrates.¹⁷⁻²¹ The protein kinase whose substrates are to be identified is first immobilized in a polyacrylamide gel. The cell lysate from which target substrates are to be identified are then resolved on this polyacrylamide gel containing the immobilized kinase. The immobilized kinase and

resolved substrates are subsequently refolded in the gel by a series of buffer-exchange processes. This enables the proteins to regain their structural conformation and the kinase to regain its catalytic activity. The gel containing the immobilized catalytically active kinase and the refolded resolved substrates are then incubated along with γ -³²P ATP in a buffer favoring phospho-transfer reaction thus permitting the *in-situ* phosphorylation of the substrates by the kinase in the gel.¹⁷ The position of the phosphorylated substrate is identified by autoradiography after washing away the excess unincorporated radio labeled ATP. The identity of the phosphorylated substrate and the site of phosphorylation can be deduced by mass spectrometry. Li X et al., have demonstrated RIKA to be far more sensitive and to offer better signal-to-noise ratio compared to KRESTEL.¹⁷ The false-positive identification rate by RIKA was also demonstrated to be very extremely low compared to other contemporary methods.¹⁷ Also, unlike protein kinase chips, RIKA permits the use of an unbiased substrate pool. Since the proteins in the gel are refolded before phospho-transfer, it alleviates the issue of exposing usually inaccessible regions of a protein to the kinase. However, whether the process achieves proper refolding of all proteins is debatable. Although RIKA scores over other *in vitro* methods, a valid concern in using RIKA has been that RIKA permits phosphorylation of substrates that are normally inaccessible to the kinase in the cells. This has been the biggest issue with all *in vitro* technologies to identify kinase-substrate relationship.

More recently RIKA was adapted to address this concern and to determine the stoichiometry of specific *in vivo* phosphorylation events.^{19,20,22} Achieving this relies on complete dephosphorylation of proteins in a sample by hydrogen fluoride treatment. Chemical dephosphorylation by hydrogen fluoride has been shown to non-specifically remove post-translational modifications from proteins without affecting its integrity.²³ To determine *in vivo* phosphorylation status of a specific protein using RIKA, the protein is purified from the cell lysate using a specific tag or antibody raised against that protein. The eluate containing purified protein is divided into two. One half is completely dephosphorylated by hydrogen fluoride (HF) treatment.²³ The HF-treated and control eluates are subjected to RIKA using the kinase whose *in vivo* phosphorylation site is being analyzed. Since during RIKA (and other *in vitro* methods) only sites that are previously non-phosphorylated are available for phosphorylation, the intensity of the RIKA signal corresponds to the amount of substrates available for phosphorylation by the kinase in the gel. Thus, the signal from HF treated samples represents phosphorylation signal from total substrate (completely dephosphorylated) and the signal from control treated samples represents the pool of substrate in the cells that was not phosphorylated at the site recognized by that particular kinase. If the substrate is phosphorylated *in vivo* at the site recognized by the particular kinase, then the RIKA signal in the control-treated sample will be less than the HF-treated sample. The signals from the two samples can be normalized to the corresponding western blots (protein level) and the phosphorylation index of the substrate at the site phosphorylated by the specific kinase (or % of substrate phosphorylated *in vivo*) can be calculated. This method can be further extended to further confirm the *in vivo* status of phosphorylation of this site by performing the same assay with cells treated with inhibitor of a specific kinase and comparing the phosphorylation index with and without kinase inhibition.^{19,22} A *bona fide in vivo* phosphorylation event would show decreased phosphorylation index upon inhibitor treatment. The applicability of this method however heavily relies on the ability of a kinase to refold and regain its activity in the gel. This could be challenging especially for protein kinases with large domains. Overcoming this hurdle would predicate the

large-scale applicability of this method. Also, in its present state, the method is quite labor intensive compared to other high-throughput techniques. Combining RIKA with the chemical-genetic approach would eliminate background phosphorylation by endogenous kinases present in the cell lysate and significantly improve the signal-noise ratio of this method. Nevertheless, this method has lot of advantages and if improved to permit high-throughput analysis can serve as a valuable tool.

SILAC-based approaches

Recently many high throughput proteomic approaches have been developed to identify *in vivo* phosphorylation sites on proteins. Most of them exploit stable isotope labeling with amino acids in cell culture (SILAC), which relies on metabolic incorporation of the 'light' or 'heavy' form of a given amino acid.²⁴ The standard procedure involves treating cells grown in media containing different isotopically labeled amino acids with either inhibitor of a particular kinase or vehicle control. This enables generation of peptides from these samples, which can be subsequently quantitated relative to each other.²⁴ The cell lysate from the inhibitor treated (heavy isotope labeled) cells and mock treated (light isotope labeled), are mixed, digested using trypsin (or other enzymes) and the peptides are first separated by chromatography and then the phosphopeptides are enriched using IMAC resin. Analysis of the phosphopeptides using LC-MS/MS yields information on the sites on these peptides that are phosphorylated and by comparing relative abundance of particular phosphopeptides containing heavy and light isotopes reveals the stoichiometry with which these sites are phosphorylated *in vivo*.²⁴ Techniques based on this strategy are high throughput and provides precise *in vivo* phosphorylation status of proteins. Although this technique can identify global phosphorylation change in response to activation or inhibition of a particular protein kinase it does not conclusively identify a particular protein as a substrate of a specific kinase. Many of the sites identified through this approach could reflect indirect changes as a result of inhibition or activation of a particular protein kinase and associated signaling pathways. Despite this limitation, SILAC-based approaches have been extensively used and have provided valuable information on previously unknown kinase-substrate relationships.

Conclusion

Recent advances in proteomic technologies have paved way for the development of several new methods for identifying substrates of protein kinases. These techniques have greatly contributed to our understanding of protein kinase signaling pathways and biology. To realize the complete potential of these techniques and to determine the method best suited to address a specific problem, it is important to understand the strengths and limitations of each method. This review highlights these aspects for some of the popular and widely used methods to identify protein kinase substrates.

Acknowledgements

None.

Conflict of interest

The author declares no conflict of interest.

References

1. Padmanabhan A, Rao V, De Marzo AM, et al. Regulating NKX3.1 stability and function: Post-translational modifications and structural determinants. *Prostate*. 2016;76(6):523–533.

2. Padmanabhan A, Gosc EB, Bieberich CJ. Stabilization of the prostate-specific tumor suppressor NKX3.1 by the oncogenic protein kinase Pim-1 in prostate cancer cells. *J Cell Biochem*. 2013;114(5):1050–1057.
3. Chisholm C, Padmanabhan A, Guan B, et al. Protein Kinase CK2 Controls the Stability of Prostate Derived ETS Factor. *The Open Cancer Journal*. 2010;3:109–115.
4. Padmanabhan A. Targeting Protein Kinase Substrate Docking in Cancers. *Journal of Glycobiology*. 2014;3:e1105.
5. Padmanabhan A. MYC, Metabolism and Malignancy: Opportunities for Therapeutic Intervention. *Journal of Carcinogenesis & Mutagenesis*. 2013;4:146.
6. Manning BD, Cantley LC. Hitting the target: emerging technologies in the search for kinase substrates. *Sci STKE*. 2002;2002(162):pe49.
7. Yang X, Hubbard EJ, Carlson M. A protein kinase substrate identified by the two-hybrid system. *Science*. 1992;257:680–682.
8. Fukunaga R, Hunter T. MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO J*. 1997;16(8):1921–1933.
9. Jiang W, Jimenez G, Wells NJ, et al. PRC1: a human mitotic spindle-associated CDK substrate protein required for cytokinesis. *Mol Cell*. 1998;2(6):877–885.
10. Xue Y, Gao X, Cao J, et al. A summary of computational resources for protein phosphorylation. *Curr Protein Pept Sci*. 2010;11(6):485–596.
11. Xue Y, Ren J, Gao X, et al. GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarc. *Mol Cell Proteomics*. 2008;7(9):1598–608.
12. Huang HD, Lee TY, Tzeng SW, et al. KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. *Nucleic Acids Res*. 2005;33:W226–W229.
13. Cohen P, Knebel A. KESTREL: a powerful method for identifying the physiological substrates of protein kinases. *Biochem J*. 2006;393:1–6.
14. Songyang Z, Blechner S, Hoagland N, et al. Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol*. 1994;4:973–982.
15. Shah K, Shokat KM. A chemical genetic approach for the identification of direct substrates of protein kinases. *Methods Mol Biol*. 2003;233:253–271.
16. Ptacek J, Devgan G, Michaud G, et al. Global analysis of protein phosphorylation in yeast. *Nature*. 2005;438(7060):679–684.
17. Li X, Guan B, Srivastava MK, et al. The reverse in-gel kinase assay to profile physiological kinase substrates. *Nat Methods*. 2007;4(11):957–962.
18. Toughiri R, Padmanabhan A, Li X, et al. Discovering a new physiological substrate of Aurora-A kinase. *Cancer Research*. 2011;71:1619–1619.
19. Padmanabhan A, Li X, Bieberich CJ. Protein kinase A regulates MYC protein through transcriptional and post-translational mechanisms in a catalytic subunit isoform-specific manner. *J Biol Chem*. 2013;288:14158–14169.
20. Padmanabhan A. *In Biological Sciences 241*. Baltimore County: University Of Maryland; 2012.
21. Padmanabhan A. The Search of New Cancer Biomarkers: Moving Forward. *Journal of Clinical Trials*. 2013;3.
22. Li X, Rao V, Jin J, et al. Identification and validation of inhibitor-responsive kinase substrates using a new paradigm to measure kinase-specific protein phosphorylation index. *J Proteome Res*. 2012;11(7):3637–3649.
23. Kuyama H, Toda C, Watanabe M, et al. An efficient chemical method for dephosphorylation of phosphopeptides. *Rapid Commun Mass Spectrom*. 2003;17(12):1493–1496.
24. Pimienta G, Chaerkady R, Pandey A. SILAC for global phosphoproteomic analysis. *Methods Mol Biol*. 2009;527:107–116.