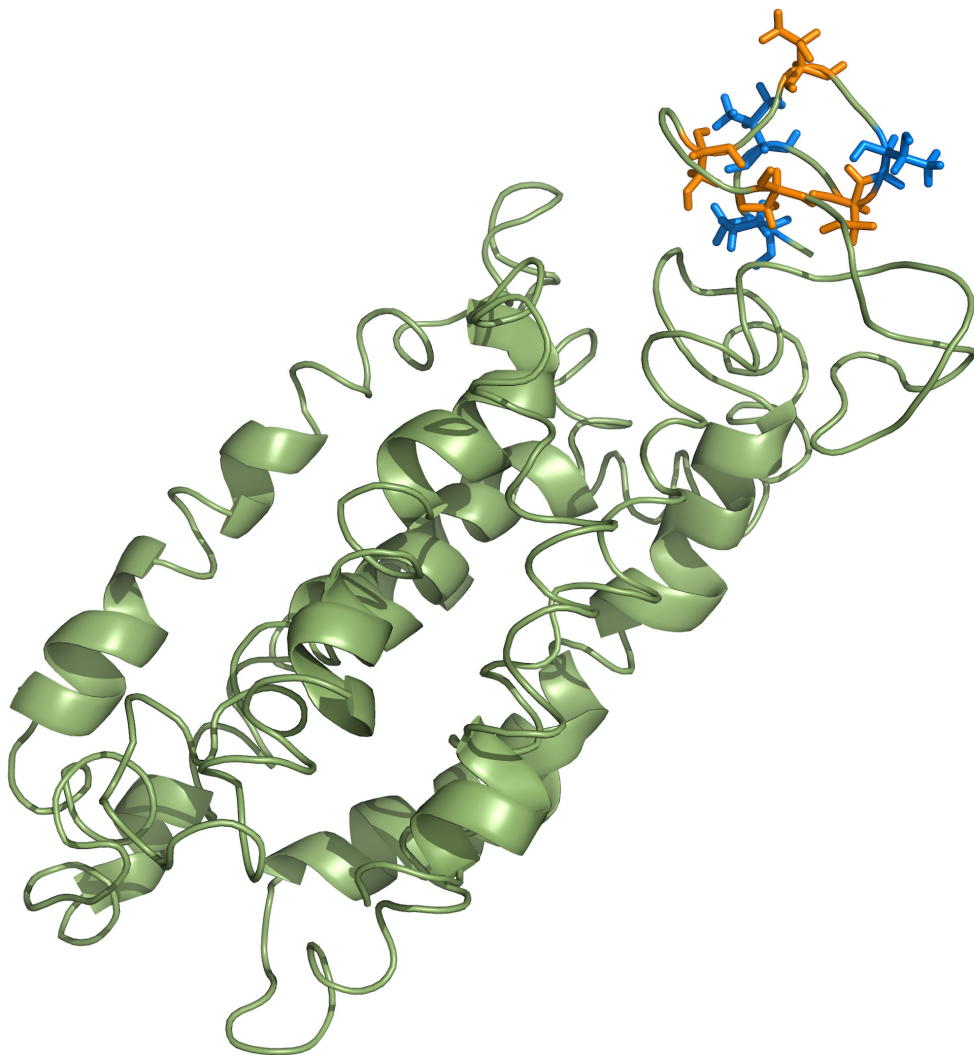


# PROTEIN PHOSPHORYLATION IN HEALTH AND DISEASE

EDITED BY: Allegra Via and Andreas Zanzoni  
PUBLISHED IN: Frontiers in Genetics





# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-900-6

DOI 10.3389/978-2-88919-900-6

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

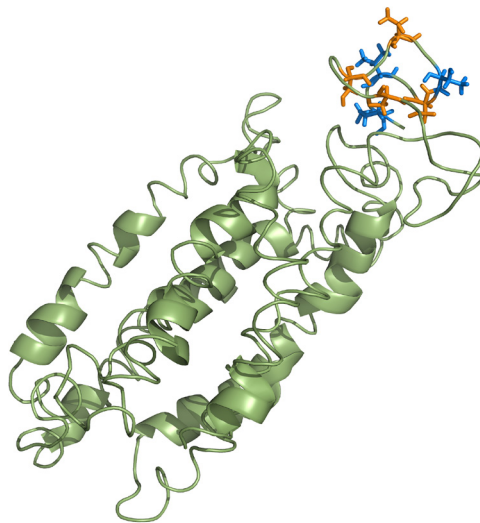
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# PROTEIN PHOSPHORYLATION IN HEALTH AND DISEASE

Topic Editors:

**Allegra Via**, CNR Institute of Biomembranes and Bioenergetics, Italy

**Andreas Zanzoni**, TAGC UMR S1090, Inserm, Aix-Marseille Université, France



Three-dimensional structure of rhodopsin [PDB:1JFP], a G-protein coupled receptor found in retina rods. The C-terminus of rhodopsin mediates the interaction with distinct protein partners, such as Tctex-1 and arrestin. The latter binds to rhodopsin C-terminus in a phosphorylation-dependent manner (phosphorylated residues are depicted in orange). Interestingly, this region also contains retinitis pigmentosa mutation sites (residues highlighted in blue) [MIM:613731], which are known to hamper the interaction ability of rhodopsin.

Figure by Allegra Via

Protein phosphorylation is one of the most abundant reversible post-translational modifications in eukaryotes. It is involved in virtually all cellular processes by regulating protein function, localization and stability and by mediating protein-protein interactions. Furthermore, aberrant protein phosphorylation is implicated in the onset and progression of human diseases such as cancer and neurodegenerative disorders.

In the last years, tens of thousands of *in vivo* phosphorylation events have been identified by large-scale quantitative phospho-proteomics experiment suggesting that a large fraction of the proteome might be regulated by phosphorylation. This data explosion is increasingly enabling the development of computational approaches, often combined with experimental validation, aiming at prioritizing phosphosites and assessing their functional relevance. Some computa-

tional approaches also address the inference of specificity determinants of protein kinases/phosphatases and the identification of phosphoresidue recognition domains. In this context, several challenging issues are still open regarding phosphorylation, including a better understanding of the interplay between phosphorylation and allosteric regulation, agents and mechanisms disrupting or promoting abnormal phosphorylation in diseases, the identification and modulation of novel phosphorylation inhibitors, and so forth. Furthermore, the determinants of kinase and phosphatase recognition and binding specificity are still unknown in several cases, as well as the impact of disease mutations on phosphorylation-mediated signaling.

The articles included in this Research Topic illustrate the very diverse aspects of phosphorylation, ranging from structural changes induced by phosphorylation to the peculiarities of phosphosite evolution. Some also provide a glimpse into the huge complexity of phosphorylation networks and pathways in health and disease, and underscore that a deeper knowledge of such processes is essential to identify disease biomarkers, on one hand, and design more effective therapeutic strategies, on the other.

**Citation:** Via, A., Zanzoni, A., eds. (2016). Protein Phosphorylation in Health and Disease. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-900-6



# Table of Contents

- 05    *A prismatic view of protein phosphorylation in health and disease***  
Allegra Via and Andreas Zanzoni
- 08    *Physicochemical mechanisms of protein regulation by phosphorylation***  
Hafumi Nishi, Alexey Shaytan and Anna R. Panchenko
- 18    *Heat shock protein 27 phosphorylation state is associated with cancer progression***  
Maria Katsogiannou, Claudia Andrieu and Palma Rocchi
- 23    *Phosphorylation of unique domains of Src family kinases***  
Irene Amata, Mariano Maffei and Miquel Pons
- 29    *Toward a systems-level view of dynamic phosphorylation networks***  
Robert H. Newman, Jin Zhang and Heng Zhu
- 51    *Exploiting holistic approaches to model specificity in protein phosphorylation***  
Antonio Palmeri, Fabrizio Ferrè and Manuela Helmer-Citterich
- 62    *Computational methods for analysis and inference of kinase/inhibitor relationships***  
Fabrizio Ferrè, Antonio Palmeri and Manuela Helmer-Citterich
- 67    *Defying c-Abl signaling circuits through small allosteric compounds***  
Stefania Gonfloni
- 73    *Modeling conformational transitions in kinases by molecular dynamics simulations: achievements, difficulties, and open challenges***  
Marco D'Abramo, Neva Besker, Giovanni Chillemi and Alessandro Grottesi
- 78    *Combining affinity proteomics and network context to identify new phosphatase substrates and adapters in growth pathways***  
Francesca Sacco, Karsten Boldt, Alberto Calderone, Simona Panni, Serena Paoluzi, Luisa Castagnoli, Marius Ueffing and Gianni Cesareni
- 93    *Phosphatases of  $\alpha$ -synuclein, LRRK2, and tau: important players in the phosphorylation-dependent pathology of Parkinsonism***  
Jean-Marc Taymans and Veerle Baekelandt
- 105    *Identification of developmentally-specific kinotypes and mechanisms of Varroa mite resistance through whole-organism, kinome analysis of honeybee***  
Albert J. Robertson, Brett Trost, Erin Scruten, Thomas Robertson, Mohammad Mostajeran, Wayne Connor, Anthony Kusalik, Philip Griebel and Scott Napper
- 117    *Turnover of protein phosphorylation evolving under stabilizing selection***  
Christian R. Landry, Luca Freschi, Taraneh Zarin and Alan M. Moses

# A prismatic view of protein phosphorylation in health and disease

Allegra Via<sup>1\*</sup> and Andreas Zanzoni<sup>2,3\*</sup>

<sup>1</sup> Department of Physics, Sapienza University of Rome, Rome, Italy, <sup>2</sup> Technological Advances for Genomics and Clinics (TAGC), UMR\_S1090, INSERM, Marseille, France, <sup>3</sup> Technological Advances for Genomics and Clinics (TAGC), UMR\_S1090, Aix Marseille Université, Marseille, France

**Keywords: protein phosphorylation, disease, evolution, cell signaling, systems biology, bioinformatics, phosphorylation networks**

The paramount relevance of protein phosphorylation in health and disease motivated us to gather several contributions from experts working in this area in order to portray the recent developments in the field. Our effort and the effort of 54 authors with their 12 contributions gave rise to this Research Topic, which represents a valuable forum where phosphorylation is discussed from different angles, including bioinformatics approaches and experimental methods that are currently used to decipher the complex mechanisms underlying this bewitching post-translational modification (PTM). The articles collected in this Research Topic illustrate very diverse aspects of phosphorylation, such as its biological effects and induced structural changes, the role of kinases and phosphatases as therapeutic targets, the use of phosphorylation profiles as biomarkers, how phosphorylation dys-regulation may cause disease, and more.

Nishi et al. (2014), in their extensive review of the representative studies on the biological effects of phosphorylation, show that a general mechanism of regulation by phosphorylation does not exist. Indeed, phosphorylation may serve as recognition/binding site or trigger allosteric effects inducing local structural changes, which may propagate into larger structure rearrangements. Some nice examples of the biological consequences of protein phosphorylation in physiological and disease states are described in two articles of this Research Topic.

The Hsp27 protein (coded by the HSPB1 gene) is a chaperone that is aberrantly expressed in many types of tumors and represents a promising drug target (Acunzo et al., 2014). Phosphorylation of serine residues affects the oligomerization state of Hsp27, favoring the recruitment of different client proteins involved in distinct cellular functions (Katsogiannou et al., 2014a). Katsogiannou et al. (2014b) suggest that better understanding Hsp27 phosphorylation dynamics in cancer may help improve existing and/or develop new therapies.

Amata et al. (2014) report the interesting case of the role of phosphorylation events in the Unique domain of Src family kinases (SFKs). This domain, an intrinsically disordered region little conserved across the family and linking the kinase to its membrane-anchoring domain, is stubbed with phosphosites involved in multilevel regulation of SFKs, including, among others, anchoring to the lipid membrane.

These examples hint at the high complexity of the cellular networks regulated by phosphorylation. Kinases and phosphatases can be also regulated by phosphorylation and most of signal transduction pathways involve cascades of phosphorylation and de-phosphorylation events. This scenario should give the reader a glimpse of the consequences on entire pathways of the abolition or stimulation of phosphorylation caused, for instance, by mutational events. In this context, the development of experimental and computational methods aimed at understanding the mechanisms of specificity and functioning of kinase and phosphatase repertoires is becoming increasingly relevant.

## OPEN ACCESS

### Edited and reviewed by:

Raina Robeva,  
Sweet Briar College, USA

### \*Correspondence:

Allegra Via and Andreas Zanzoni,  
allegra.via@uniroma1.it;  
zanzoni@tagc.univ-mrs.fr

### Specialty section:

This article was submitted to Systems  
Biology, a section of the journal  
Frontiers in Genetics

**Received:** 05 March 2015

**Accepted:** 18 March 2015

**Published:** 07 April 2015

### Citation:

Via A and Zanzoni A (2015) A  
prismatic view of protein  
phosphorylation in health and disease.  
Front. Genet. 6:131.  
doi: 10.3389/fgene.2015.00131

A comprehensive review (Newman et al., 2014) presents a broad picture on well established and emerging experimental methods that are providing new insights on the organization and regulation of phosphorylation networks. Thanks to these approaches, thousands of phosphorylation events have been identified in distinct cellular conditions. Importantly, these technological advances are stimulating the development of computational models to gain a systems-level understanding of phosphorylation maps (e.g., Linding et al., 2007; Song et al., 2012; Newman et al., 2013).

A full comprehension of such maps requires the elucidation of the molecular determinants of substrate recognition. The review of Palmeri and collaborators summarizes the current knowledge on kinase-substrate specificity (Palmeri et al., 2014), arguing that only the combination of different sources of contextual information will provide new insights into the molecular determinants of kinase specificity and improve the performance of kinase-substrate interaction prediction tools.

The understanding of kinase specificity is also particularly relevant for the identification, design and development of compounds modulating the kinase activity. In this context, reliable computational methods for kinase/inhibitor inference and analyses would play a crucial role. As pointed out by Ferrè et al. (2014), such methods are being successfully developed and applied to the whole kinome thanks to the increasing amount of homogeneous data provided by high-throughput profiling studies.

Most inhibition screenings focus on compounds able to bind the kinase catalytic site. However, on one hand, targeting the catalytic site lacks specificity and, on the other, the modulation of kinase activity can be achieved also through alternative routes. An illustrative example, is presented by Gonfloni in a perspective article reviewing the current knowledge of c-Abl regulation by allosteric compounds, such as GNF-2 (Gonfloni, 2014). The author proposes that the use of current and novel allosteric activators will elucidate the role of c-Abl regulation in cancer and neurodegenerative disorder signaling pathways.

This example suggests that rational drug design would tremendously benefit from a better comprehension of kinase conformational changes and flexibility. Whilst experimental techniques able to accurately describe such phenomena are still under development, computational methods for the study of kinase conformational transitions—especially Atomistic Molecular Dynamics (MD)—demonstrated to be of great help. In their minireview, D'Abramo et al. (2014) highlight that, since kinase conformational switches from inactive to active states occur on long time-scales, enhanced sampling techniques and brute-force approaches are being developed and successfully applied to rational drug design in a growing number of cases.

Despite phosphatases are key players in phosphorylation networks, they caught less attention than kinases as drug targets and the identification of phosphatase physiological substrates is still an open challenge (Brautigan, 2013).

In order to tackle this problem, Sacco and colleagues (Sacco et al., 2014), in their original research article, present computational an integrative approach combining functional

siRNA information (Sacco et al., 2012), interaction discovery experiments and network analyses to identify phosphatase substrates and potential scaffolds proteins that could mediate substrate recognition. Their strategy was able to recover known as well as novel phosphatase substrates/scaffolds, some of which were further validated.

The growing knowledge of phosphatases and their substrates is expanding the interest of the scientific community in these enzymes as potential therapeutic targets in order to modulate dys-regulated phosphorylation levels of disease proteins. One example is provided by Taymans and Baekelandt (2014) who review phosphorylation dys-regulation in the three main proteins linked to Parkinson Disease: alpha-synuclein, Leucine-rich repeat kinase type 2 (LRRK2), and microtubule associate protein tau (tau). They analyze the feasibility of targeting their phosphatases as a potential therapy for Parkinsonism.

Together with the study of kinases and phosphatases, the analysis of phosphorylation profiles is also important in developing therapies or biomarkers for phosphorylation-related diseases.

In an original research article, Robertson and colleagues (Robertson et al., 2014) describe an integrative approach to study the honeybee (*Apis mellifera*) kinome with the aim of detecting differences in phosphorylation profiles between bees with differential susceptibility to Varroa mite infestation at different developmental stages. They found that many peptides are differentially phosphorylated between the two phenotypes and bioinformatics analyses showed clear differences between resistant and susceptible phenotypes.

The evolution of PTMs and phosphorylation in particular, is another key aspect of the story. In a minireview, Landry et al. (2014) present recent evidences suggesting that clusters of sites rather than individual sites represent functional units evolving under stabilizing selection. Nevertheless, more experimental studies are needed to test whether the stabilizing selection model applies to phosphosite clusters.

With the multifaceted description of protein phosphorylation presented in this Research Topic, we hope we provided the readers with a taste of the huge complexity of phosphorylation networks and pathways and convinced them that a deeper knowledge of the interplay between kinases, phosphatases and their substrates is essential in the quest for new disease biomarkers and novel therapeutic targets.

Finally, we would like to thank all the authors for their pivotal contribution to this Research Topic, as well as the 25 peers that acted as review editors and the Frontiers in Genetics Editorial Office staff members for their support.

## Acknowledgments

AV acknowledges the King Abdullah University of Science and Technology (KAUST) Award n° KUK-I1-012-43 and AZ the French “Plan Cancer 2009–2013” program (Systems Biology call, A12171AS) and the A\*MIDEX project (n° ANR-11-IDEX-0001-02) for financial support.

## References

- Acunzo, J., Andrieu, C., Baylot, V., So, A., and Rocchi, P. (2014). Hsp27 as a therapeutic target in cancers. *Curr. Drug Targets* 15, 423–431. doi: 10.2174/13894501113146660230
- Amata, I., Maffei, M., and Pons, M. (2014). Phosphorylation of unique domains of Src family kinases. *Front. Genet.* 5:181. doi: 10.3389/fgene.2014.00181
- Brautigan, D. L. (2013). Protein Ser/Thr phosphatases—the ugly ducklings of cell signalling. *FEBS J.* 280, 324–345. doi: 10.1111/j.1742-4658.2012.08609.x
- D'Abramo, M., Besker, N., Chillemi, G., and Grottesi, A. (2014). Modeling conformational transitions in kinases by molecular dynamics simulations: achievements, difficulties, and open challenges. *Front. Genet.* 5:128. doi: 10.3389/fgene.2014.00128
- Ferrè, F., Palmeri, A., and Helmer-Citterich, M. (2014). Computational methods for analysis and inference of kinase/inhibitor relationships. *Front. Genet.* 5:196. doi: 10.3389/fgene.2014.00196
- Gonfloni, S. (2014). Defying c-Abl signaling circuits through small allosteric compounds. *Front. Genet.* 5:392. doi: 10.3389/fgene.2014.00392
- Katsogiannou, M., Andrieu, C., Baylot, V., Baudot, A., Duseti, N. J., Gayet, O., et al. (2014a). The functional landscape of Hsp27 reveals new cellular processes such as DNA repair and alternative splicing and proposes novel anticancer targets. *Mol. Cell. Proteomics* 13, 3585–3601. doi: 10.1074/mcp.M114.041228
- Katsogiannou, M., Andrieu, C., and Rocchi, P. (2014b). Heat shock protein 27 phosphorylation state is associated with cancer progression. *Front. Genet.* 5:346. doi: 10.3389/fgene.2014.00346
- Landry, C. R., Freschi, L., Zarin, T., and Moses, A. M. (2014). Turnover of protein phosphorylation evolving under stabilizing selection. *Front. Genet.* 5:245. doi: 10.3389/fgene.2014.00245
- Linding, R., Jensen, L. J., Ostheimer, G. J., van Vugt, M. A. T. M., Jørgensen, C., Miron, I. M., et al. (2007). Systematic discovery of *in vivo* phosphorylation networks. *Cell* 129, 1415–1426. doi: 10.1016/j.cell.2007.05.052
- Newman, R. H., Hu, J., Rho, H.-S., Xie, Z., Woodard, C., Neiswinger, J., et al. (2013). Construction of human activity-based phosphorylation networks. *Mol. Syst. Biol.* 9, 655. doi: 10.1038/msb.2013.12
- Newman, R. H., Zhang, J., and Zhu, H. (2014). Toward a systems-level view of dynamic phosphorylation networks. *Front. Genet.* 5:263. doi: 10.3389/fgene.2014.00263
- Nishi, H., Shaytan, A., and Panchenko, A. R. (2014). Physicochemical mechanisms of protein regulation by phosphorylation. *Front. Genet.* 5:270. doi: 10.3389/fgene.2014.00270
- Palmeri, A., Ferrè, F., and Helmer-Citterich, M. (2014). Exploiting holistic approaches to model specificity in protein phosphorylation. *Front. Genet.* 5:315. doi: 10.3389/fgene.2014.00315
- Robertson, A. J., Trost, B., Scruten, E., Robertson, T., Mostajeran, M., Connor, W., et al. (2014). Identification of developmentally-specific kinotypes and mechanisms of Varroa mite resistance through whole-organism, kinome analysis of honeybee. *Front. Genet.* 5:139. doi: 10.3389/fgene.2014.00139
- Sacco, F., Boldt, K., Calderone, A., Panni, S., Paoluzi, S., Castagnoli, L., et al. (2014). Combining affinity proteomics and network context to identify new phosphatase substrates and adapters in growth pathways. *Front. Genet.* 5:115. doi: 10.3389/fgene.2014.00115
- Sacco, F., Gherardini, P. F., Paoluzi, S., Saez-Rodriguez, J., Helmer-Citterich, M., Ragnini-Wilson, A., et al. (2012). Mapping the human phosphatome on growth pathways. *Mol. Syst. Biol.* 8, 603. doi: 10.1038/msb.2012.36
- Song, C., Ye, M., Liu, Z., Cheng, H., Jiang, X., Han, G., et al. (2012). Systematic analysis of protein phosphorylation networks from phosphoproteomic data. *Mol. Cell. Proteomics* 11, 1070–1083. doi: 10.1074/mcp.M111.012625
- Taymans, J.-M., and Baekelandt, V. (2014). Phosphatases of  $\alpha$ -synuclein, LRRK2, and tau: important players in the phosphorylation-dependent pathology of Parkinsonism. *Front. Genet.* 5:382. doi: 10.3389/fgene.2014.00382

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Via and Zanzoni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Physicochemical mechanisms of protein regulation by phosphorylation

Hafumi Nishi<sup>1\*</sup>, Alexey Shaytan<sup>2</sup> and Anna R. Panchenko<sup>2\*</sup>

<sup>1</sup> Graduate School of Medical Life Science, Yokohama City University, Yokohama, Japan

<sup>2</sup> National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

## Edited by:

Allegra Via, Sapienza University, Italy  
Andreas Zanzoni, Inserm TAGC,  
UMR1090, France

## Reviewed by:

Monika Heiner, Brandenburg  
University of Technology, Germany  
Fabrizio Ferrè, University of Rome Tor  
Vergata, Italy

## \*Correspondence:

Anna R. Panchenko, National Center  
for Biotechnology Information,  
National Library of Medicine, National  
Institutes of Health, Building 38A,  
8600 Rockville Pike, Bethesda,  
MD, USA

e-mail: panch@ncbi.nlm.nih.gov;  
Hafumi Nishi, Graduate School of  
Medical Life Science, Yokohama City  
University, 1-7-29 Suehiro-cho,  
Tsurumi-ku, Yokohama, Japan  
e-mail: hnishi@tsurumi.yokohama-cu.  
ac.jp

Phosphorylation offers a dynamic way to regulate protein activity and subcellular localization, which is achieved through its reversibility and fast kinetics. Adding or removing a dianionic phosphate group somewhere on a protein often changes the protein's structural properties, its stability and dynamics. Moreover, the majority of signaling pathways involve an extensive set of protein–protein interactions, and phosphorylation can be used to regulate and modulate protein–protein binding. Losses of phosphorylation sites, as a result of disease mutations, might disrupt protein binding and deregulate signal transduction. In this paper we focus on the effects of phosphorylation on protein stability, dynamics, and binding. We describe several physico-chemical mechanisms of protein regulation through phosphorylation and pay particular attention to phosphorylation in protein complexes and phosphorylation in the context of disorder–order and order–disorder transitions. Finally we assess the role of multiple phosphorylation sites in a protein molecule, their possible cooperativity and function.

**Keywords: protein phosphorylation, protein–protein interactions, allosteric regulation, protein disorder, multisite phosphorylation**

## INTRODUCTION

Cellular regulatory mechanisms provide a sensitive, specific, and robust response to external stimuli and posttranslational modifications (PTMs) play an important role in these mechanisms and control protein activity, subcellular localization, and stability (Olsen et al., 2006). Such dynamic regulation is achieved through reversibility and fast kinetics of PTMs. Recent phospho-proteomic analyses have revealed that the majority of proteins in a mammalian cell are phosphorylated (Olsen et al., 2010). In eukaryotes phosphoryl group can be attached to serine, threonine, and tyrosine residues and in prokaryotes the most commonly phosphorylated residues are histidine and aspartic acid. While majority of phospho complexes in human contain only few phosphorylation sites, some proteins have up to half of their serine, threonine, and tyrosine sites phosphorylated (Nishi et al., 2011). Overall, phosphorylated serines are the most abundant (86%), followed by threonine (12%), and tyrosine phosphorylations (2%; Olsen et al., 2006). The abundance and specificity of phosphorylation as regulatory mechanism is evident from the large number of genes (more than 500) encoding protein kinases which constitute almost 2% of human protein coding genes (Manning et al., 2002). The number of phosphatases is almost ten times smaller.

In this paper we summarize biological effects of phosphorylation which are explained through the lens of structural and dynamical changes. Below we will review a number of representative studies of computer simulations of the effects of phosphorylation on protein dynamics and stability together with

experimental techniques to reveal the details and underlying mechanisms of phosphorylation events at atomistic scale.

## EFFECT OF PHOSPHORYLATION ON STRUCTURE AND DYNAMICS

### STRUCTURAL CONSEQUENCES OF PHOSPHORYLATION

Phosphoryl group is dianionic at physiological pH and can form extensive hydrogen bond networks and salt bridges with neighboring residues of the same or different chains. One of the most dominant modes of interactions between phosphoryl and other residues is the interaction with the  $\alpha$ -helical dipole at the C-terminal main chain nitrogen to neutralize the combined effect of carbonyl dipoles (Johnson and Lewis, 2001). Another common mode of interaction is the formation of hydrogen bonds and salt bridges between the phosphate oxygens and arginine or lysine side chains. Arginine side chain usually makes stronger salt bridges with phosphorylated side chains compared to lysine whereas phosphoserine (pSer) hydrogen-bond acceptor forms more stable interactions than phosphoaspartate (pAsp) acceptor (Mandell et al., 2007). Although the strength of hydrogen bonds in general should depend on the phosphate protonation state, the latter effect was shown to be rather subtle (Mandell et al., 2007) with a more pronounced effect of protonation state on pAsp than on pSer. All things considered, adding or removing a dianionic phosphate group in a protein might considerably change its local physico-chemical properties and affect stability, kinetics, and dynamics (Johnson, 2009).



Analyses of phosphorylation in different proteins revealed the diversity and heterogeneity of its effects on protein structure (Zanzoni et al., 2011), phosphorylation can impact protein structure at local as well as global levels. A recent large-scale study compared the sets of phosphorylated and unphosphorylated protein structures and showed that phosphorylation produced local as well as global changes in structure (Xin and Radivojac, 2012). Structural changes produced by phosphorylation were the highest among other PTMs. However, according to this study, only 13% of proteins exhibited the root mean square deviation (RMSD) of 2 Å or higher between phosphorylated and unphosphorylated forms and it has been argued that phosphorylation in many cases might restrict the conformational flexibility of protein monomers (Xin and Radivojac, 2012; Li et al., 2013).

There were several attempts to predict structural rearrangements induced by phosphorylation or dephosphorylation events. Although accurate predictions could be made only for a few cases, such analyses allowed to pinpoint the underlying mechanisms which govern the transitions between phosphorylated and unphosphorylated states. For example, by *in silico* phosphorylating several proteins and evaluating their conformations by OPLS-AA (all atom Optimized Potentials for Liquid Simulations) force field and a Generalized Born implicit solvent model, it was shown that structures of phosphorylated regions and conformational changes induced by phosphorylation could be predicted in some cases with near-atomic accuracy compared to the actual phosphorylated conformations (Groban et al., 2006). In another study a coarse-grained model was applied to sample the conformations of nuclear factor of activated T cells (NFAT) which was phosphorylated at multiple sites. It was found that predicted changes produced by phosphorylation differed between cytoplasmic and nuclear forms of NFAT and were driven mostly by electrostatic and solvation energy contributions (Shen et al., 2005). Several cases of the effects of phosphorylation on structure and dynamics are reviewed in **Table 1**.

### COUPLING BETWEEN PHOSPHORYLATION AND PROLINE ISOMERIZATION

Interestingly, phosphorylation might not induce the structural change by itself but rather may serve as a recognition site for an enzyme which catalyzes the conformational switch. A classical example of such mechanism is proline-directed phosphorylation which occurs on serine or threonine residues preceding proline (Lu et al., 2002). This mechanism of regulation involves specific peptidyl-prolyl *cis/trans* isomerase Pin1 (Lu et al., 1996) which recognizes phosphorylated Ser/Thr-Pro motif and catalyzes *cis/trans* isomerization of phosphorylated Ser/Thr-Pro bonds. Phosphorylation dramatically slows down the uncatalyzed isomerization rate of Ser/Thr-Pro bonds, while rendering them inappropriate for the action of general peptidyl-prolyl *cis/trans* isomerases (Yaffe et al., 1997). This complex interplay of changes introduced by phosphorylation in relation to isomerization may affect dynamics and reaction kinetics of processes involved in timing and duration of cellular response. With the help of accelerated molecular dynamics (MD) simulations, it was elegantly demonstrated in molecular details how serine phosphorylation in Ser-Pro motifs may shift the equilibrium between *cis* and *trans* proline

isoforms and consequently slow down the rate of isomerization (Hamelberg et al., 2005). The authors found that isomerization of the omega-bond of proline is asymmetric and strongly depends on the psi-backbone angle of proline whereas phosphorylation might favor the  $\alpha$ -helical backbone conformation.

### ALLOSTERIC REGULATION BY PHOSPHORYLATION AND DISORDER

Phosphorylation may trigger the transitions between conformations with different activity and/or binding specificity leading to activation or deactivation of a protein (Dou et al., 2012; Kales et al., 2012). It can play a role of covalently attached allosteric effector which induces local changes at first which may propagate thereafter into larger tertiary or quaternary structure rearrangements (Nussinov et al., 2012). One of the classical examples of a protein with large conformational changes produced by phosphorylation is glycogen phosphorylase which exists as a homodimer in inactive T state and as a tetramer in an active R state. This transition is allosterically controlled by phosphorylation of only one residue Ser14 (Johnson, 1992). Catalytic sites of glycogen phosphorylase are buried and are not solvent accessible in inactive form. Phosphorylation of Ser14 leads to large conformational movements displacing protein N-terminal region by almost 50 Å so that some intrachain contacts of Ser14 can be replaced with the contacts between pSer14 and arginine of another identical chain in a homodimer. The change in a dimer binding mode causes reconfiguration of the catalytic site and a subsequent activation of glycogen phosphorylase. Recently the model was proposed which tried to explain the allosteric coupling between phosphorylation and allostery (Mitternacht and Berezhovsky, 2011). In the case of glycogen phosphorylase the authors found that the active sites had very high allosteric coupling via so-called binding leverage mechanism with those sites where the unphosphorylated N-terminal segment binds.

Another two examples illustrate the mechanism of allosteric regulation of protein activity through the coupled interplay between phosphorylation and disorder–order transitions. Activation of myosin in smooth muscle depends on the phosphorylation of regulatory light chain (RLC). In the unphosphorylated state myosin is auto-inhibited by interactions between the two catalytic domains, while phosphorylation of RLC at Ser19, which is rather distant from catalytic domain, disrupts these interactions and relieves the inhibition (Sellers, 1985). The complete mechanism by which phosphorylation of RLC activates myosin is still not known, but a series of combined experimental (EPR, TR-FRET) and MD studies were able to elucidate first steps in a cascade of conformational transitions (Nelson et al., 2005; Espinoza-Fonseca et al., 2007, 2008; Kast et al., 2010). In particular, upon phosphorylation little change in the direct vicinity of phosphorylation site is seen, while the  $\alpha$ -helical content in region Lys11–Ala17 increases dramatically. This finding revealed a disorder–order transition induced by phosphorylation, and corroborates published experimental data on site-directed spin labeling (Nelson et al., 2005). The thermodynamic and structural basis of this phosphorylation-induced disorder–order transitions were further studied (Espinoza-Fonseca et al., 2008) and revealed a delicate balance between the gain in enthalpy due to electrostatic interactions and loss in entropy due to constraining the

Table 1 | Selected examples demonstrating the mechanisms of phosphorylation on protein structure, stability, and dynamics.

Protein name	Method	Type of induced change upon phosphorylation	Specific outcome of phosphorylation	Reference
Short peptide with serine-proline motifs	AMD	Cis/trans isomerization of proline	Favors a helix formation	Hamelberg et al. (2005)
NFAT regulatory domain	BD, CG AMH	Conformational change	Blocks an access to the nuclear localization signaling sequence	Shen et al. (2005)
Na <sup>+</sup> /K <sup>+</sup> -ATPase	MD, electrophysiology experiments	Conformational change	Affects Na <sup>+</sup> binding site and modulates ion pumping	Poulsen et al. (2012)
H3 histone tail in complex with HP1	MD	Binding affinity change	Destabilizes the complex formation	Papamokos et al. (2012)
Smooth muscle myosin regulatory light chain	TR-FRET, MD	Conformational change	Shifts conformational equilibrium to an open state	Kast et al. (2010)
c-Src-kinase	US MD	Conformational change, activity change	Locks protein in an active conformation	Meng and Roux (2014)
Collagen domain protein Shc and Epidermal growth factor receptor (EGFR)	MD, MMPBSA	Conformational rearrangement of domains, binding affinity change	Causes significant rearrangement of protein domains.	Suenaga et al. (2009)
Myelin Basic Protein (MBP)	NMR, TFE-titration, MD	Conformational change, binding affinity change	Disfavors formation of amphipathic $\alpha$ -helix, inhibits MBP membrane interactions	Vassall et al. (2013)
H3 histone tail, checkpoint kinase 1 (Chk1), protein kinase C (PKC)	NMR, MD	Binding affinity change, phosphorylation cooperativity and crosstalk	Phosphorylation of H3 tail affects binding of Chk1 and inhibits phosphorylation of neighboring residues	Liokatis et al. (2012)
Extracellular signal-regulated kinases (ERK)	Mutagenesis, experimental catalytic activity measurement, MD, TMD	Influence of phosphomutations on autophosphorylation	Phosphomimic mutations affect activation through autophosphorylation and folding of homodimerization interface	Barr et al. (2011)

MD, molecular dynamics; BD, Brownian dynamics with implicit solvent; AMD, accelerated molecular dynamics; CG AMH, coarse-grain modeling with associative memory Hamiltonian; TR-FRET, time-resolved fluorescence energy transfer experiments; US MD, umbrella sampling molecular dynamics; TMD, targeted molecular dynamics; MM-PBSA, molecular mechanics Poisson–Boltzmann surface area method; NMR, nuclear magnetic resonance spectroscopy.



conformational dynamics of positively charged residues upon phosphorylation.

An interesting mechanism where phosphorylation inhibits disorder–order transition was reported for myelin basic protein (MBP), which includes a proline rich peptide containing two Thr-Pro motifs (-TPRTPPPS-) and an adjacent amphipathic  $\alpha$ -helix which can bind to membrane (Vassall et al., 2013). Both Thr-Pro motifs can be phosphorylated by mitogen-activated protein kinases. Using a combination of NMR spectroscopy, circular dichroism spectroscopy, trifluoroethanol-titration and MD simulations the authors investigated the structure of  $\alpha$ -helical and proline-rich regions and the effects of phosphorylation on their conformation. It was found that phosphorylation on one or both sites impedes the formation of the neighboring amphipathic  $\alpha$ -helix. This supports the hypothesis that structure of the membrane anchoring  $\alpha$ -helix is disrupted upon phosphorylation and thus regulates the association of MBP with the membrane. In addition, the proline-rich region may adopt PPII structure near the lipid interface when the MPB is anchored to the membrane via amphipathic helix.

## PHOSPHORYLATION IN PROTEIN–PROTEIN BINDING

### PHOSPHORYLATION ON INTERFACES MODULATES PROTEIN–PROTEIN BINDING

Many cellular control mechanisms operate at the level of protein–protein interactions, and main signaling pathways involve dense networks of protein–protein interactions and phosphorylation events. Phosphorylation may not only trigger the transitions between different conformation states of one protein but in some cases may modulate transitions between different conformations or oligomeric states in homooligomeric and heterooligomeric complexes and might represent an important mechanism for regulation of protein activity (Randez-Gil et al., 1998; Jia-Lin Ma and Stern, 2008; Hashimoto et al., 2011). Recently Nishi et al. performed a comprehensive analysis of phosphorylation sites on protein–protein binding interfaces (Nishi et al., 2011). They mapped experimentally identified phosphorylation sites onto crystal structures of human homo- and hetero-oligomers and showed that protein interfaces of transient homo- and hetero-oligomers are statistically enriched with phosphorylation sites compared to non-interfacial protein surface sites. The authors found that changes in binding affinity produced by substitutions at phosphorylation sites on binding interfaces of heterooligomers are larger compared to other sites on interfaces. In addition, consistent with the observation that phosphosites may frequently target binding hot spots, significant association between phosphosites and binding hotspots was observed (binding hot spots were defined if substitutions of residues in these sites into alanine considerably destabilizes the complex by more than 2 kcal/mol; Bogan and Thorn, 1998; Nishi et al., 2011).

Calculation of binding energy differences upon phosphorylation showed that the majority of phosphorylation events did not affect protein–protein binding (Nishi et al., 2011). It was consistent with several experimental studies pointing to the modest effect of phosphorylation on protein stability (Murray et al., 1998; Serber and Ferrell, 2007). Even if phosphorylation does not affect complex stability, it can provide diversity in recognition patterns

and offer recognition sites for binding of phosphoresidue binding domains thereby modulating binding selectivity. Phosphoresidue-binding domains are common functional modules distributed widely among cellular signaling proteins. Numerous studies have identified and investigated phosphoresidue binding domains in various proteins (Via et al., 2011; Reinhardt and Yaffe, 2013) such as SH2 and PTB domains for phosphotyrosine (Pawson et al., 2001), 14-3-3 domains for phosphoserine (Yaffe et al., 1997), and FHA domains for phosphothreonine (Durocher et al., 2000). Usually these domains contain arginine or lysine residues in their binding regions to form hydrogen bonds with phosphates, and may have neighboring residues (e.g., hydrophobic residues for phosphotyrosine) which help to recognize phosphorylated site, or any specific residues in the binding motifs (Liang and Van Doren, 2008; Johnson et al., 2010). Several cases of the effects of phosphorylation on protein–protein binding are reviewed in **Table 1**.

### REGULATION OF BINDING BY DISORDER–ORDER AND ORDER–DISORDER TRANSITIONS UPON PHOSPHORYLATION

Many proteins and protein regions are intrinsically disordered under native conditions, namely, they contain no or very little well-defined structure. Folding of disordered proteins into ordered structures may occur upon binding to their specific partners which in turn might provide high specificity even if binding affinity is low (Wright and Dyson, 1999; Sugase et al., 2007). On the other hand, a number of experimental studies on p53 (Scheinin et al., 1990), cystic fibrosis transmembrane-conductance regulator (CFTR; Bozoky et al., 2013), p27 (Yoon et al., 2012), and other proteins (Johnson, 2009) have shown that disordered regions often contain phosphorylation sites and (de)phosphorylation events can be coupled to disorder–order transitions. The first systematic study was performed on a large set of proteins trying to link disordered regions with the locations of experimental phosphorylation sites. This study found that intrinsically disordered regions were enriched in phosphorylation sites (Iakoucheva et al., 2004). Moreover, protein N- and C-terminal regions which are usually disordered often participate in binding to other proteins (Fong and Panchenko, 2010) and there are many cases where terminal regions contain multiple phosphorylation sites (Chacko et al., 2001). Functional diversity of disordered regions and their propensity for PTMs allow them to play a unique role in signaling networks where phosphorylation events might serve as switches and regulate binding events. In some cases, as was shown in the previous section, the regulation of binding might happen without invoking disordered regions while in others the regulatory mechanism might involve phosphorylation as well as disorder–order or order–disorder transitions.

Before describing specific cases of proteins involving the coupling between disorder, phosphorylation and binding, we would like to describe several studies which tried to generalize fundamental principles of such coupling in many different proteins. Mohan et al. analyzed relatively short (10–70 residues) segments called MoRFs (molecular recognition features) contained within longer disordered sequences that were structurally characterized in a complex with a larger protein (Mohan et al., 2006). It was assumed that MoRFs may undergo folding upon binding but would be disordered in their unbound state. The authors of this study applied DISorder-enhanced PHOSphorylation predictor

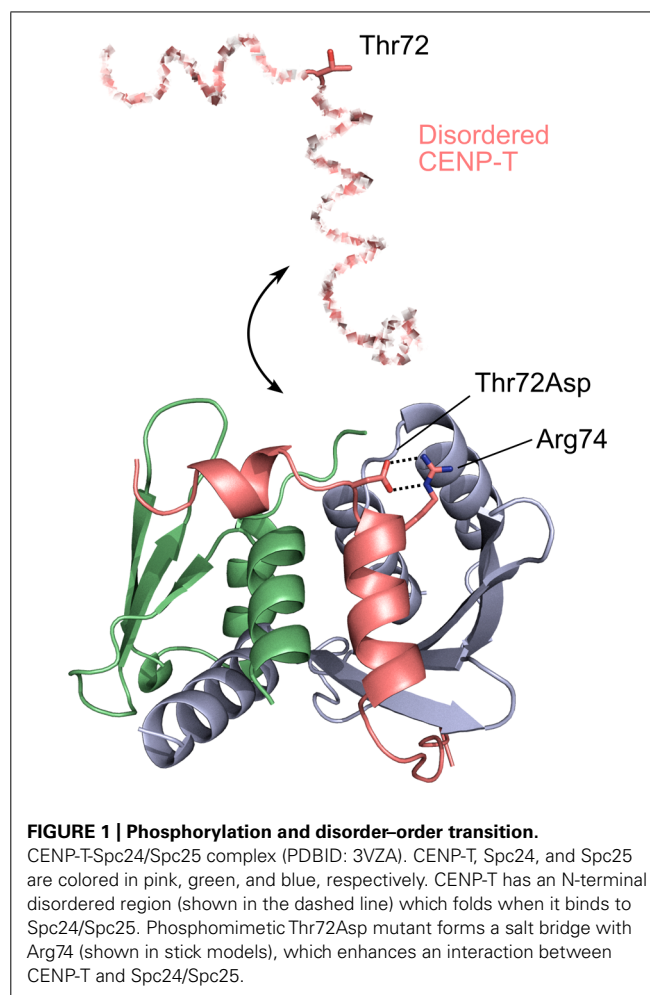
(DisPhos) to MoRF regions and found that in 305 MoRFs of more than 12 residues long, 159 of them had potential phosphorylation sites, suggesting that phosphorylation may be a common mechanism to modulate binding. Later, another group studied an association between phosphorylation and disordered binding regions in human protein complexes using experimentally identified phosphorylation sites and disorder prediction methods (Nishi et al., 2013). They showed that disordered interface residues (corresponding to sites disordered in unbound states and structured in the complex) had the highest fraction of phosphorylation sites (25%) compared to ordered interface (8%) or disordered non-interface (18%) residues, suggesting a strong association between phosphorylation and disordered interface residues. Disorder and interfacial location were significantly linked to phosphorylation of serine and to a lesser extent to phosphorylation of threonine. Tyrosine phosphorylation was not found to be directly associated with binding through disorder, and was often observed in ordered interface regions which were not predicted to be disordered in the unbound state. The fractions of phosphorylated Ser, Thr, and Tyr in disordered interfaces were 59, 26, and 15%, respectively, and were found to be quite different from those of structured interface (28, 22, and 50%; Nishi et al., 2013).

While in many disorder-involving transitions phosphorylated residues may directly regulate binding orthosterically, this is not necessarily the case. Centromere protein T (CENP-T) is an essential component of the inner kinetochore and consists of N-terminal disordered region and C-terminal histone-fold domain. The long disordered region is employed to bind to outer kinetochore complexes, namely, to Spc24/Spc25 subunits of Ndc80 complex (Gascoigne et al., 2011; Nishino et al., 2013). Nishino et al. (2013) revealed that phosphorylation on Thr72 of chicken CENP-T is crucial for its binding to Spc24/Spc25. The X-ray crystal structure of a complex between CENP and Spc24/Spc25 showed that binding segment comprising residues 63–93 contains two short helices, and Thr72 is located on a loop between these two helices. Interestingly, site-directed phosphomimicking mutagenesis experiments showed that Thr72Asp mutant forms a salt bridge with Arg74 on the second helix facilitating the orientation of hydrophobic residues on the second helix toward the hydrophobic patch on Spc25 partner. As a consequence, it enhances an interaction between CENP-T and Spc24/Spc25 (Figure 1). In addition, this phosphorylation site and salt bridge are conserved in many eukaryotic species which suggests that this mechanism is widespread for the CENP-T regulation. This example shows that phosphorylated residue can be critical for the complex formation through disorder even if it is located far away from the binding interface.

## MULTIPLE SITE PHOSPHORYLATION

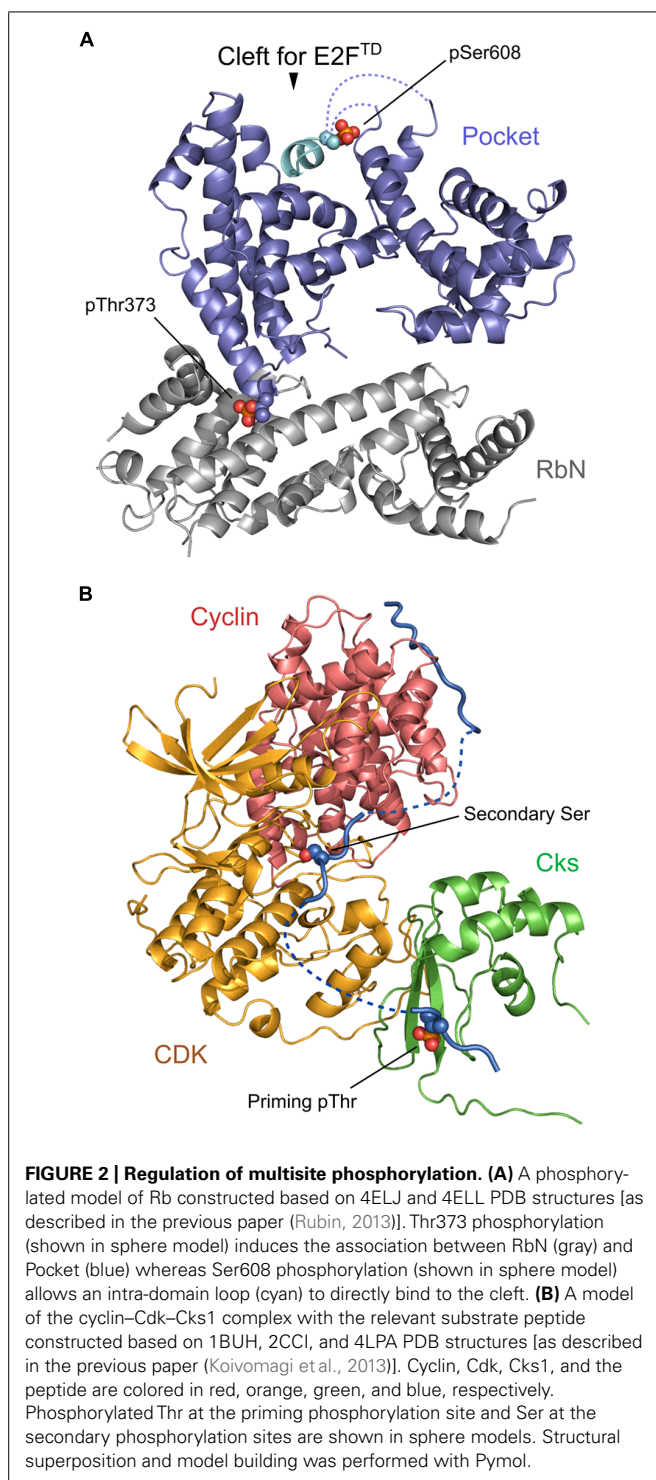
### DISTRIBUTION AND FUNCTION OF MULTIPLE PHOSPHORYLATION SITES IN PROTEINS

Single protein may contain multiple phosphorylation sites. Multisite phosphorylation can expand the patterns of regulation, give more accurate modulation of conformational change (Kumar et al., 2012) and cooperatively increase binding affinity to other proteins (Ferreon et al., 2009). Large scale analyses revealed that multiple phosphorylation sites are not distributed randomly, but



are often clustered on a protein (Li et al., 2009; Schweiger and Linial, 2010; Freschi et al., 2014). Namely 54% of all pSer/pThr sites are located within four residues of each other, while the tendency to form clusters is not very pronounced for pTyr sites. Clustered pSer/pThr sites are usually phosphorylated by the same kinase and clustered Ser/Thr prefer to be located in disordered regions compared to non-clustered Ser/Thr (Schweiger and Linial, 2010). Moreover, evolutionary clustered sites are 1.4 times more likely to be phosphorylated by the same kinases than expected by chance (Freschi et al., 2014).

Retinoblastoma protein (Rb) is one of the classical examples of a protein which has multiple phosphorylation sites and concerted phosphorylation patterns with very specific functional roles. Rb contains RbN, pocket, and RbC domains together with 13 different Ser/Thr phosphorylation sites that are phosphorylated by Cdk kinases. Phosphorylation sites are roughly grouped into eight clusters which mostly reside in flexible loop regions between structured regions or domains, and mediate domain-domain, domain-loop, and protein-protein interactions (Hassler et al., 2007; Burke et al., 2012). For example, Thr373 is located at the end of a flexible loop between RbN and pocket domains which do not interact if this residue is not phosphorylated. Phosphorylation of Thr373 induces large conformational changes



and, as a consequence, an interaction between RbN and pocket domains, which allosterically inhibits binding of transactivation domain of E2F transcription factor (E2F<sup>TD</sup>) to the pocket domain (Burke et al., 2012; **Figure 2A**). Meanwhile, phosphorylation on Ser608 and Ser612 directly and orthosterically inhibits binding of E2F<sup>TD</sup> to the pocket. This mechanism involves a competitive binding between E2F<sup>TD</sup> and Ser608/Ser612 containing loop,

namely, pSer608 stabilizes the association with the binding cleft, thereby mimicking and competing with E2F<sup>TD</sup> (Burke et al., 2010; **Figure 2A**). A recent study showed that phosphorylation of Ser788 and Ser795 may also cause the inhibition of E2F<sup>TD</sup> binding to the pocket by inducing the association between RbC and the pocket domain (Burke et al., 2014). This phosphorylation is additive with the effect of other preceding phosphorylations in inhibiting E2F<sup>TD</sup> binding, demonstrating separate regulatory mechanisms by different phosphorylation site clusters.

## MECHANISMS OF MULTIPLE PHOSPHOSITE PROCESSING

Processes of (de)phosphorylation on multiple sites can be classified by the order of (de)phosphorylation events, which can be sequential or random. In sequential phosphorylation sites are phosphorylated in a strict order of events where phosphorylation of one site depends on the phosphorylation state of another. Sequential phosphorylation has been observed for several kinases, especially Ser/Thr kinases (Salazar and Hofer, 2009). In contrast, random phosphorylation does not require the strict order of phosphorylation events. Kinetics of (de)phosphorylation can be distinguished by the number of binding events of kinases or phosphatases. A kinase may phosphorylate all sites while staying bound to the substrate (processive mechanism) or may bind and then dissociate after each phosphorylation (distributive mechanism; Patwardhan and Miller, 2007). For example, phosphorylation of Cdc25 by Cdk1 is most likely to be random and distributive, namely, mutations on single phosphorylation sites do not preclude other phosphorylation events, and intermediate levels of Cdk1 yield partially phosphorylated Cdc25 (Lu et al., 2012). On the other hand, some kinases require “priming” phosphorylation, which automatically determines the order of phosphorylation. Known priming recognition motifs include (S/T)XXX(pS/pT) motif for GSK3 (ter Haar et al., 2001) or (S/T)XX(E/D/pS/pT) for CK2 kinases (Meggio and Pinna, 2003). Interestingly, some recent studies on human CFTR protein showed that tyrosine residue in “SYDE” motif can act as both positive and negative regulator of phosphorylation of the first serine by CK2 kinase (Cesaro et al., 2013). While “SYDE” sequence matches the CK2 canonical phosphorylation motif (SXXE), this motif is not properly phosphorylated unless tyrosine is replaced or phosphorylated.

In sequential and processive phosphorylation, distances between phosphorylation sites can be critical to maintain the phosphorylation process. Koivomagi et al. (2013) revealed the molecular mechanism of semi-processive phosphorylation on Sic1 by cyclin-Cdk1-Cks1 complex (**Figure 2B**). Sic1 contains seven Ser/Thr sites in its N-terminal region, which are phosphorylated by Cdk1. Thr5 and Thr33 were previously identified as priming phosphorylation sites compared to downstream secondary phosphorylation sites (Thr45, Thr48, Ser69, Ser76, and Ser80). First, Sic1 binds to cyclin, which induces priming phosphorylation on Thr5 and Thr33 by Cdk1. This process is inhibited if priming Thr residues are placed closer to the cyclin docking sites, indicating the importance of maintaining the proper distance between cyclin-binding motif and the phosphorylation target residues of Sic1 (Koivomagi et al., 2013). Subsequently, the phosphorylated priming sites are docked to the Cks1 pocket which in turn allows



Cdk1 to access and phosphorylate downstream Ser/Thr sites which are located between the priming sites and the cyclin-binding sites. Additionally it has been shown that priming and secondary phosphorylation sites should be separated by at least 12 amino acids, otherwise the efficiency of secondary phosphorylation is greatly reduced. Overall, the authors of this paper proposed that the ability of Cdk1 to process multiple phosphorylation sites depends on spatial patterns of multiple phosphosite clusters and correct arrangements of cyclin and Cks binding elements (Koivomagi et al., 2013).

Sequential phosphorylation may also gradually increase the negative charge of a region and lead to the cooperative behavior between different phosphorylation sites. For example, phosphorylation of the *Neurospora* clock protein FREQUENCY (FRQ) is rate limiting for degradation and therefore crucial for circadian time keeping (Querfurth et al., 2011). This protein exists in closed and open states and in the course of the day the N-terminal domain of FRQ is sequentially phosphorylated at up to 46 sites, which increases its negative charge. As a result, the interaction with the negatively charged middle and C-terminal domains are destabilized which in turn shifts the equilibrium toward an open conformation. In an open conformation the signaling motif is exposed which targets protein for degradation.

Receptor tyrosine kinases (RTKs) represent another example. They transduce signals from the extracellular matrix to the cytoplasm of a cell and contain extracellular, transmembrane, and catalytic kinase domains and may include regulatory domains. In many cases binding of a ligand to the extracellular part induces dimerization or higher order oligomer formation and leads to the activation of intracellular kinase domain and its subsequent cross-phosphorylation. The interconversion between active and inactive states in kinases is highly regulated and kinases differ in their mechanisms of activation and inactivation (Hubbard and Miller, 2007). A key tyrosine in insulin receptor kinase domain protrudes into its active site, stabilizes inactive state and blocks access to ATP (Huse and Kuriyan, 2002) whereas tyrosines of activation loop in FGFR1 do not obstruct the ATP binding site but block the substrate binding site. Phosphorylated tyrosine can form an electrostatic contact with the basic residues, stabilize the active state of kinase and enable phosphorylation of other tyrosine residues on the C-terminal tail, which in turn mediate binding of SH2 and PTB domains of downstream signaling molecules. Phosphorylation of tyrosines happens in precise sequential order and autophosphorylation of Tyr653 in activation loop of FGFR1 increases kinase activity by 10–50-fold (Furdui et al., 2006) while subsequent phosphorylation of Tyr583, Tyr463, and Tyr585 boosts the catalytic activity up to 500-fold.

## PHOSPHORYLATION AND POST-TRANSLATIONAL MODIFICATION CROSSTALK

Post-translational modification crosstalk occurs in those cases where the presence of one modification influences the modification of another site. Phosphorylation can change the activity of proteins that regulate other types of PTMs and, as a consequence, can promote or inhibit the modification of other sites. Several studies attempted to identify crosstalk between concerted

phosphosites and other PTMs by looking at PTM sites within the sequence proximity from each other and by analyzing their evolutionary conservation and functional importance (Beltrao et al., 2012; Peng et al., 2014). Some of these studies have been recently reviewed (Lothrop et al., 2013; Gajadhar and White, 2014; Venne et al., 2014).

Phosphorylation in some cases can promote subsequent ubiquitylation and the crosstalk between phosphorylation and ubiquitylation is reciprocal, namely, phosphorylation can be regulated by ubiquitylation and vice versa (Swaney et al., 2013). Interplay between phosphorylation and protein GlcNAcylation was further examined and it was shown that an increased GlcNAcylation led to lower phosphorylation at 280 phosphosites while causing an increased phosphorylation at 148 sites (Wang et al., 2008). Different patterns of PTMs may govern the interactions with different proteins; these patterns are dynamic and may respond to changes in a cellular state. In particular, it was found that majority of proteins detected in response to stimulation with epidermal growth factor (EGF) were phosphorylated on multiple sites. Moreover, various phosphosites on one protein showed different kinetics pointing to the fact that they might play different functional roles (Olsen et al., 2006).

Several comprehensive statistical studies were recently performed trying to decipher the co-evolutionary links between different types of post-translationally modified sites. It was shown that phosphorylation associates with eleven other PTM types, followed by glycosylation and acetylation (Minguez et al., 2012; Beltrao et al., 2013). In addition it was reported that the coordination between different types of PTMs may occur at the level of one subunit in a protein complex since subunits highly modified by one PTM were also enriched by other PTM type (Woodsmith et al., 2013).

Crosstalk between phosphorylation and other PTMs can be illustrated by an example of histone tail modifications that are sometimes called the “histone code.” One important modification includes histone H3 Lys9 methylation that creates a binding site for chromodomain of heterochromatin protein 1 (HP1) which plays a key role in heterochromatin formation. Adjacent residue Ser10 is a known phosphorylation site. While the ultimate mechanism of Lys9 methylation and Ser10 phosphorylation crosstalk in H3 histone is still unknown, MD simulations showed that upon phosphorylation Ser10 forms a stable salt bridge with Arg8 rather than with the positively charged Lys9. It leads to a rearrangement of tail conformation and affects the binding of HP1 to methylated H3 Lys9 (Papamokos et al., 2012).

## CONCLUSION

In this paper we reviewed the present state of the structural and biophysical studies of protein phosphorylation. Physico-chemical consequences of phosphorylation are very diverse which makes it difficult to summarize and deduce general mechanisms of phosphorylation events. However, recent experimental and computational studies point to several major mechanisms for how phosphorylation may ultimately affect and modulate protein function. They include orthosteric and allosteric effects of phosphorylation on protein structure and protein–protein binding, disorder–order and order–disorder coupled transitions upon

phosphorylation and, finally, cooperativity and crosstalk between multiple phosphorylation sites or other PTMs. The structural and biophysical characterization of phosphorylation crosstalk is still in its infancy but in the future it will provide important clues about mechanisms of signal propagation, integration, and separation.

## ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the National Library of Medicine at the US National Institutes of Health. Alexey Shaytan was in part supported by US–Russia Collaboration in the Biomedical Sciences Fellowship Program. Hafumi Nishi was in part supported by JSPS Research Fellowships for Young Scientists.

## REFERENCES

- Barr, D., Oashi, T., Burkhard, K., Lucius, S., Samadani, R., Zhang, J., et al. (2011). Importance of domain closure for the autoactivation of ERK2. *Biochemistry* 50, 8038–8048. doi: 10.1021/bi200503a
- Beltrao, P., Albanese, V., Kenner, L. R., Swaney, D. L., Burlingame, A., Villen, J., et al. (2012). Systematic functional prioritization of protein posttranslational modifications. *Cell* 150, 413–425. doi: 10.1016/j.cell.2012.05.036
- Beltrao, P., Bork, P., Krogan, N. J., and Van Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. *Mol. Syst. Biol.* 9, 714. doi: 10.1002/msb.201304521
- Bogan, A. A., and Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* 280, 1–9. doi: 10.1006/jmbi.1998.1843
- Bozoky, Z., Krzeminski, M., Chong, P. A., and Forman-Kay, J. D. (2013). Structural changes of CFTR R region upon phosphorylation: a plastic platform for intramolecular and intermolecular interactions. *FEBS J.* 280, 4407–4416. doi: 10.1111/febs.12422
- Burke, J. R., Deshong, A. J., Pelton, J. G., and Rubin, S. M. (2010). Phosphorylation-induced conformational changes in the retinoblastoma protein inhibit E2F transactivation domain binding. *J. Biol. Chem.* 285, 16286–16293. doi: 10.1074/jbc.M110.108167
- Burke, J. R., Hura, G. L., and Rubin, S. M. (2012). Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. *Genes Dev.* 26, 1156–1166. doi: 10.1101/gad.189837.112
- Burke, J. R., Liban, T. J., Restrepo, T., Lee, H. W., and Rubin, S. M. (2014). Multiple mechanisms for E2F binding inhibition by phosphorylation of the retinoblastoma protein C-terminal domain. *J. Mol. Biol.* 426, 245–255. doi: 10.1016/j.jmb.2013.09.031
- Cesaro, L., Marin, O., Venerando, A., Donella-Deana, A., and Pinna, L. A. (2013). Phosphorylation of cystic fibrosis transmembrane conductance regulator (CFTR) serine-511 by the combined action of tyrosine kinases and CK2: the implication of tyrosine-512 and phenylalanine-508. *Amino Acids* 45, 1423–1429. doi: 10.1007/s00726-013-1613-y
- Chacko, B. M., Qin, B., Correia, J. J., Lam, S. S., De Caestecker, M. P., and Lin, K. (2001). The L3 loop and C-terminal phosphorylation jointly define Smad protein trimerization. *Nat. Struct. Biol.* 8, 248–253. doi: 10.1038/84995
- Dou, H., Buetow, L., Hock, A., Sibbet, G. J., Vousden, K. H., and Huang, D. T. (2012). Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl. *Nat. Struct. Mol. Biol.* 19, 184–192. doi: 10.1038/nsmb.2231
- Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., et al. (2000). The molecular basis of FHA domain: phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. *Mol. Cell* 6, 1169–1182. doi: 10.1016/S1097-2765(00)00114-3
- Espinoza-Fonseca, L. M., Kast, D., and Thomas, D. D. (2007). Molecular dynamics simulations reveal a disorder-to-order transition on phosphorylation of smooth muscle myosin. *Biophys. J.* 93, 2083–2090. doi: 10.1529/biophysj.106.095802
- Espinoza-Fonseca, L. M., Kast, D., and Thomas, D. D. (2008). Thermodynamic and structural basis of phosphorylation-induced disorder-to-order transition in the regulatory light chain of smooth muscle myosin. *J. Am. Chem. Soc.* 130, 12208–12209. doi: 10.1021/ja803143g
- Ferreon, J. C., Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009). Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6591–6596. doi: 10.1073/pnas.0811023106
- Fong, J. H., and Panchenko, A. R. (2010). Intrinsic disorder and protein multi-binding in domain, terminal, and linker regions. *Mol. Biosyst.* 6, 1821–1828. doi: 10.1039/c005144f
- Freschi, L., Osseni, M., and Landry, C. R. (2014). Functional divergence and evolutionary turnover in mammalian phosphoproteomes. *PLoS Genet.* 10:e1004062. doi: 10.1371/journal.pgen.1004062
- Furdui, C. M., Lew, E. D., Schlessinger, J., and Anderson, K. S. (2006). Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Mol. Cell* 21, 711–717. doi: 10.1016/j.molcel.2006.01.022
- Gajadhar, A. S., and White, F. M. (2014). System level dynamics of post-translational modifications. *Curr. Opin. Biotechnol.* 28C, 83–87. doi: 10.1016/j.copbio.2013.12.009
- Gascoigne, K. E., Takeuchi, K., Suzuki, A., Hori, T., Fukagawa, T., and Cheeseman, I. M. (2011). Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell* 145, 410–422. doi: 10.1016/j.cell.2011.03.031
- Groban, E. S., Narayanan, A., and Jacobson, M. P. (2006). Conformational changes in protein loops and helices induced by post-translational phosphorylation. *PLoS Comput. Biol.* 2:e32. doi: 10.1371/journal.pcbi.0020032
- Hamelberg, D., Shen, T., and Mccammon, J. A. (2005). Phosphorylation effects on cis/trans isomerization and the backbone conformation of serine-proline motifs: accelerated molecular dynamics analysis. *J. Am. Chem. Soc.* 127, 1969–1974. doi: 10.1021/ja0446707
- Hashimoto, K., Nishi, H., Bryant, S., and Panchenko, A. R. (2011). Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization. *Phys. Biol.* 8, 035007. doi: 10.1088/1478-3975/8/3/035007
- Hassler, M., Singh, S., Yue, W. W., Luczynski, M., Lakbir, R., Sanchez-Sanchez, F., et al. (2007). Crystal structure of the retinoblastoma protein N domain provides insight into tumor suppression, ligand interaction, and holoprotein architecture. *Mol. Cell* 28, 371–385. doi: 10.1016/j.molcel.2007.08.023
- Hubbard, S. R., and Miller, W. T. (2007). Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr. Opin. Cell Biol.* 19, 117–123. doi: 10.1016/j.ceb.2007.02.010
- Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell* 109, 275–282. doi: 10.1016/S0092-8674(02)00741-9
- Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., et al. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 32, 1037–1049. doi: 10.1093/nar/gkh253
- Jia-Lin Ma, N., and Stern, D. F. (2008). Regulation of the Rad53 protein kinase in signal amplification by oligomer assembly and disassembly. *Cell Cycle* 7, 808–817. doi: 10.4161/cc.7.6.5595
- Johnson, C., Crowther, S., Stafford, M. J., Campbell, D. G., Toth, R., and Mackintosh, C. (2010). Bioinformatic and experimental survey of 14-3-3-binding sites. *Biochem. J.* 427, 69–78. doi: 10.1042/BJ20091834
- Johnson, L. N. (1992). Glycogen phosphorylase: control by phosphorylation and allosteric effectors. *FASEB J.* 6, 2274–2282.
- Johnson, L. N. (2009). The regulation of protein phosphorylation. *Biochem. Soc. Trans.* 37, 627–641. doi: 10.1042/BST0370627
- Johnson, L. N., and Lewis, R. J. (2001). Structural basis for control by phosphorylation. *Chem. Rev.* 101, 2209–2242. doi: 10.1021/cr000225s
- Kales, S. C., Ryan, P. E., and Lipkowitz, S. (2012). Cbl exposes its RING finger. *Nat. Struct. Mol. Biol.* 19, 131–133. doi: 10.1038/nsmb.2241
- Kast, D., Espinoza-Fonseca, L. M., Yi, C., and Thomas, D. D. (2010). Phosphorylation-induced structural changes in smooth muscle myosin regulatory light chain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8207–8212. doi: 10.1073/pnas.1001941107
- Koivomagi, M., Ord, M., Iofik, A., Valk, E., Venta, R., Faustova, I., et al. (2013). Multisite phosphorylation networks as signal processors for Cdk1. *Nat. Struct. Mol. Biol.* 20, 1415–1424. doi: 10.1038/nsmb.2706
- Kumar, P., Chimenti, M. S., Pemble, H., Schonichen, A., Thompson, O., Jacobson, M. P., et al. (2012). Multisite phosphorylation disrupts arginine-glutamate salt bridge networks required for binding of cytoplasmic linker-associated protein 2 (CLASP2) to end-binding protein 1 (EB1). *J. Biol. Chem.* 287, 17050–17064. doi: 10.1074/jbc.M111.316661
- Li, H., Xing, X., Ding, G., Li, Q., Wang, C., Xie, L., et al. (2009). SysPTM: a systematic resource for proteomic research on post-translational modifications. *Mol. Cell. Proteomics* 8, 1839–1849. doi: 10.1074/mcp.M900030-MCP200

- Li, M., Shoemaker, B. A., Thangudu, R. R., Ferraris, J. D., Burg, M. B., and Panchenko, A. R. (2013). Mutations in DNA-binding loop of NFAT5 transcription factor produce unique outcomes on protein–DNA binding and dynamics. *J. Phys. Chem. B* 117, 13226–13234. doi: 10.1021/jp403310a
- Liang, X., and Van Doren, S. R. (2008). Mechanistic insights into phosphoprotein-binding FHA domains. *Acc. Chem. Res.* 41, 991–999. doi: 10.1021/ar700148u
- Liokatis, S., Stutzer, A., Elsasser, S. J., Theillet, F. X., Klingberg, R., Van Rossum, B., et al. (2012). Phosphorylation of histone H3 Ser10 establishes a hierarchy for subsequent intramolecular modification events. *Nat. Struct. Mol. Biol.* 19, 819–823. doi: 10.1038/nsmb.2310
- Lothrop, A. P., Torres, M. P., and Fuchs, S. M. (2013). Deciphering post-translational modification codes. *FEBS Lett.* 587, 1247–1257. doi: 10.1016/j.febslet.2013.01.047
- Lu, K. P., Hanes, S. D., and Hunter, T. (1996). A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* 380, 544–547. doi: 10.1038/380544a0
- Lu, K. P., Liou, Y. C., and Zhou, X. Z. (2002). Pinning down proline-directed phosphorylation signaling. *Trends Cell Biol.* 12, 164–172. doi: 10.1016/S0962-8924(02)02253-5
- Lu, L. X., Domingo-Sananes, M. R., Huzarska, M., Novak, B., and Gould, K. L. (2012). Multisite phosphoregulation of Cdc25 activity refines the mitotic entrance and exit switches. *Proc. Natl. Acad. Sci. U.S.A.* 109, 9899–9904. doi: 10.1073/pnas.1201366109
- Mandell, D. J., Chorny, L., Groban, E. S., Wong, S. E., Levine, E., Rapp, C. S., et al. (2007). Strengths of hydrogen bonds involving phosphorylated amino acid side chains. *J. Am. Chem. Soc.* 129, 820–827. doi: 10.1021/ja063019w
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298, 1912–1934. doi: 10.1126/science.1075762
- Meggio, F., and Pinna, L. A. (2003). One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* 17, 349–368. doi: 10.1096/fj.02-0473rev
- Meng, Y., and Roux, B. (2014). Locking the active conformation of c-Src kinase through the phosphorylation of the activation loop. *J. Mol. Biol.* 426, 423–435. doi: 10.1016/j.jmb.2013.10.001
- Minguez, P., Parca, L., Diella, F., Mende, D. R., Kumar, R., Helmer-Citterich, M., et al. (2012). Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.* 8:599. doi: 10.1038/msb.2012.31
- Mitternacht, S., and Berezovsky, I. N. (2011). Coherent conformational degrees of freedom as a structural basis for allosteric communication. *PLoS Comput. Biol.* 7:e1002301. doi: 10.1371/journal.pcbi.1002301
- Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., et al. (2006). Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.* 362, 1043–1059. doi: 10.1016/j.jmb.2006.07.087
- Murray, D., Hermida-Matsumoto, L., Buser, C. A., Tsang, J., Sigal, C. T., Ben-Tal, N., et al. (1998). Electrostatics and the membrane association of Src: theory and experiment. *Biochemistry* 37, 2145–2159. doi: 10.1021/bi972012b
- Nelson, W. D., Blakely, S. E., Nesmelov, Y. E., and Thomas, D. D. (2005). Site-directed spin labeling reveals a conformational switch in the phosphorylation domain of smooth muscle myosin. *Proc. Natl. Acad. Sci. U.S.A.* 102, 4000–4005. doi: 10.1073/pnas.0401664102
- Nishi, H., Fong, J. H., Chang, C., Teichmann, S. A., and Panchenko, A. R. (2013). Regulation of protein-protein binding by coupling between phosphorylation and intrinsic disorder: analysis of human protein complexes. *Mol. Biosyst.* 9, 1620–1626. doi: 10.1039/c3mb25514j
- Nishi, H., Hashimoto, K., and Panchenko, A. R. (2011). Phosphorylation in protein-protein binding: effect on stability and function. *Structure* 19, 1807–1815. doi: 10.1016/j.str.2011.09.021
- Nishino, T., Rago, F., Hori, T., Tomii, K., Cheeseman, I. M., and Fukagawa, T. (2013). CENP-T provides a structural platform for outer kinetochore assembly. *EMBO J.* 32, 424–436. doi: 10.1038/emboj.2012.348
- Nussinov, R., Tsai, C. J., Xin, F., and Radivojac, P. (2012). Allosteric post-translational modification codes. *Trends Biochem. Sci.* 37, 447–455. doi: 10.1016/j.tibs.2012.07.001
- Olsen, J. V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., et al. (2006). Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648. doi: 10.1016/j.cell.2006.09.026
- Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., et al. (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* 3:ra3. doi: 10.1126/scisignal.2000475
- Papamokos, G. V., Tziatzos, G., Papageorgiou, D. G., Georgatos, S. D., Politou, A. S., and Kaxiras, E. (2012). Structural role of RKS motifs in chromatin interactions: a molecular dynamics study of HP1 bound to a variably modified histone tail. *Biophys. J.* 102, 1926–1933. doi: 10.1016/j.bpj.2012.03.030
- Patwardhan, P., and Miller, W. T. (2007). Processive phosphorylation: mechanism and biological importance. *Cell. Signal.* 19, 2218–2226. doi: 10.1016/j.cellsig.2007.06.006
- Pawson, T., Gish, G. D., and Nash, P. (2001). SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol.* 11, 504–511. doi: 10.1016/S0962-8924(01)02154-7
- Peng, M., Scholten, A., Heck, A. J., and Van Breukelen, B. (2014). Identification of enriched PTM crosstalk motifs from large-scale experimental data sets. *J. Proteome Res.* 13, 249–259. doi: 10.1021/pr4005579
- Poulsen, H., Nissen, P., Mouritsen, O. G., and Khandelia, H. (2012). Protein kinase A (PKA) phosphorylation of Na<sup>+</sup>/K<sup>+</sup>-ATPase opens intracellular C-terminal water pathway leading to third Na<sup>+</sup>-binding site in molecular dynamics simulations. *J. Biol. Chem.* 287, 15959–15965. doi: 10.1074/jbc.M112.340406
- Querfurth, C., Diernfellner, A. C., Gin, E., Malzahn, E., Hofer, T., and Brunner, M. (2011). Circadian conformational change of the Neurospora clock protein FREQUENCY triggered by clustered hyperphosphorylation of a basic domain. *Mol. Cell* 43, 713–722. doi: 10.1016/j.molcel.2011.06.033
- Randez-Gil, F., Sanz, P., Entian, K. D., and Prieto, J. A. (1998). Carbon source-dependent phosphorylation of hexokinase PII and its role in the glucose-signaling response in yeast. *Mol. Cell. Biol.* 18, 2940–2948.
- Reinhardt, H. C., and Yaffe, M. B. (2013). Phospho-Ser/Thr-binding domains: navigating the cell cycle and DNA damage response. *Nat. Rev. Mol. Cell Biol.* 14, 563–580. doi: 10.1038/nrm3640
- Rubin, S. M. (2013). Deciphering the retinoblastoma protein phosphorylation code. *Trends Biochem. Sci.* 38, 12–19. doi: 10.1016/j.tibs.2012.10.007
- Salazar, C., and Hofer, T. (2009). Multisite protein phosphorylation – from molecular mechanisms to kinetic models. *FEBS J.* 276, 3177–3198. doi: 10.1111/j.1742-4658.2009.07027.x
- Scheinin, M., Koulou, M., Karhuvaara, S., and Zimmer, R. H. (1990). Evidence that the reversible MAO-A inhibitor moclobemide increases prolactin secretion by a serotonergic mechanism in healthy male volunteers. *Life Sci.* 47, 1491–1499. doi: 10.1016/0024-3205(90)90529-Z
- Schweiger, R., and Linial, M. (2010). Cooperativity within proximal phosphorylation sites is revealed from large-scale proteomics data. *Biol. Direct* 5:6. doi: 10.1186/1745-6150-5-6
- Sellers, J. R. (1985). Mechanism of the phosphorylation-dependent regulation of smooth muscle heavy meromyosin. *J. Biol. Chem.* 260, 15815–15819.
- Serber, Z., and Ferrell, J. E. Jr. (2007). Tuning bulk electrostatics to regulate protein function. *Cell* 128, 441–444. doi: 10.1016/j.cell.2007.01.018
- Shen, T., Zong, C., Hamelberg, D., Mccammon, J. A., and Wolynes, P. G. (2005). The folding energy landscape and phosphorylation: modeling the conformational switch of the NFAT regulatory domain. *FASEB J.* 19, 1389–1395. doi: 10.1096/fj.04-3590hyp
- Suenaga, A., Hatakeyama, M., Kiyatkin, A. B., Radhakrishnan, R., Taiji, M., and Kholodenko, B. N. (2009). Molecular dynamics simulations reveal that Tyr-317 phosphorylation reduces Shc binding affinity for phosphotyrosyl residues of epidermal growth factor receptor. *Biophys. J.* 96, 2278–2288. doi: 10.1016/j.bpj.2008.11.018
- Sugase, K., Dyson, H. J., and Wright, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021–1025. doi: 10.1038/nature05858
- Swaney, D. L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., et al. (2013). Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat. Methods* 10, 676–682. doi: 10.1038/nmeth.2519
- ter Haar, E., Coll, J. T., Austen, D. A., Hsiao, H. M., Swenson, L., and Jain, J. (2001). Structure of GSK3 $\beta$  reveals a primed phosphorylation mechanism. *Nat. Struct. Biol.* 8, 593–596. doi: 10.1038/89624
- Vassall, K. A., Bessonov, K., De Avila, M., Polverini, E., and Haraux, G. (2013). The effects of threonine phosphorylation on the stability and dynamics of the central molecular switch region of 18.5-kDa myelin basic protein. *PLoS ONE* 8:e68175. doi: 10.1371/journal.pone.0068175
- Venne, A. S., Kollipara, L., and Zahedi, R. P. (2014). The next level of complexity: crosstalk of posttranslational modifications. *Proteomics* 14, 513–524. doi: 10.1002/pmic.201300344

- Via, A., Diella, F., Gibson, T. J., and Helmer-Citterich, M. (2011). From sequence to structural analysis in protein phosphorylation motifs. *Front. Biosci. (Landmark Ed.)* 16:1261–1275.
- Wang, Z., Gucek, M., and Hart, G. W. (2008). Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13793–13798. doi: 10.1073/pnas.0806216105
- Woodsmith, J., Kamburov, A., and Stelzl, U. (2013). Dual coordination of post translational modifications in human protein networks. *PLoS Comput. Biol.* 9:e1002933. doi: 10.1371/journal.pcbi.1002933
- Wright, P. E., and Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293, 321–331. doi: 10.1006/jmbi.1999.3110
- Xin, F., and Radivojac, P. (2012). Post-translational modifications induce significant yet not extreme changes to protein structure. *Bioinformatics* 28, 2905–2913. doi: 10.1093/bioinformatics/bts541
- Yaffe, M. B., Rittinger, K., Volinia, S., Caron, P. R., Aitken, A., Leffers, H., et al. (1997). The structural basis for 14-3-3: phosphopeptide binding specificity. *Cell* 91, 961–971. doi: 10.1016/S0092-8674(00)80487-0
- Yoon, M. K., Mitrea, D. M., Ou, L., and Kriwacki, R. W. (2012). Cell cycle regulation by the intrinsically disordered proteins p21 and p27. *Biochem. Soc. Trans.* 40, 981–988. doi: 10.1042/BST20120092
- Zanzoni, A., Carbajo, D., Diella, F., Gherardini, P. F., Tramontano, A., Helmer-Citterich, M., et al. (2011). Phospho3D 2.0: an enhanced database of three-dimensional structures of phosphorylation sites. *Nucleic Acids Res.* 39, D268–D271. doi: 10.1093/nar/gkq936

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 April 2014; accepted: 22 July 2014; published online: 07 August 2014.

Citation: Nishi H, Shaytan A and Panchenko AR (2014) Physicochemical mechanisms of protein regulation by phosphorylation. *Front. Genet.* 5:270. doi: 10.3389/fgene.2014.00270

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Nishi, Shaytan and Panchenko. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Heat shock protein 27 phosphorylation state is associated with cancer progression

**Maria Katsogiannou<sup>1,2,3,4</sup>\*, Claudia Andrieu<sup>1,2,3,4</sup> and Palma Rocchi<sup>1,2,3,4</sup>\***

<sup>1</sup> Institut National de la Santé et de la Recherche Médicale, Unités Mixtes de Recherche 1068, Centre de Recherche en Cancérologie de Marseille, Marseille, France

<sup>2</sup> Institut Paoli-Calmettes, Marseille, France

<sup>3</sup> Centre de Recherche en Cancérologie de Marseille, Institut National de la Santé et de la Recherche Médicale Unités Mixtes de Recherche 1068, Aix-Marseille Université, Marseille, France

<sup>4</sup> Centre National de la Recherche Scientifique, Unités Mixtes de Recherche 7258, Centre de Recherche en Cancérologie de Marseille, Marseille, France

## Edited by:

Andreas Zanzoni, Technological Advances for Genomics and Clinics Laboratory INSERM UMR1090, France  
Allegra Via, Sapienza University, Italy

## Reviewed by:

Robert M. Tanguay, Université Laval, Canada  
Francesca Sacco, Max Plank Institute, Germany

## \*Correspondence:

Palma Rocchi and Maria Katsogiannou, Inserm, UMR1068, Centre de Recherche en Cancérologie de Marseille – Institut Paoli-Calmette – Aix-Marseille Université, 27 Boulevard Leï Roure BP30059, 13273 Marseille Cedex 9, France  
e-mail: palma.rocchi@inserm.fr; maria.katsogiannou@inserm.fr

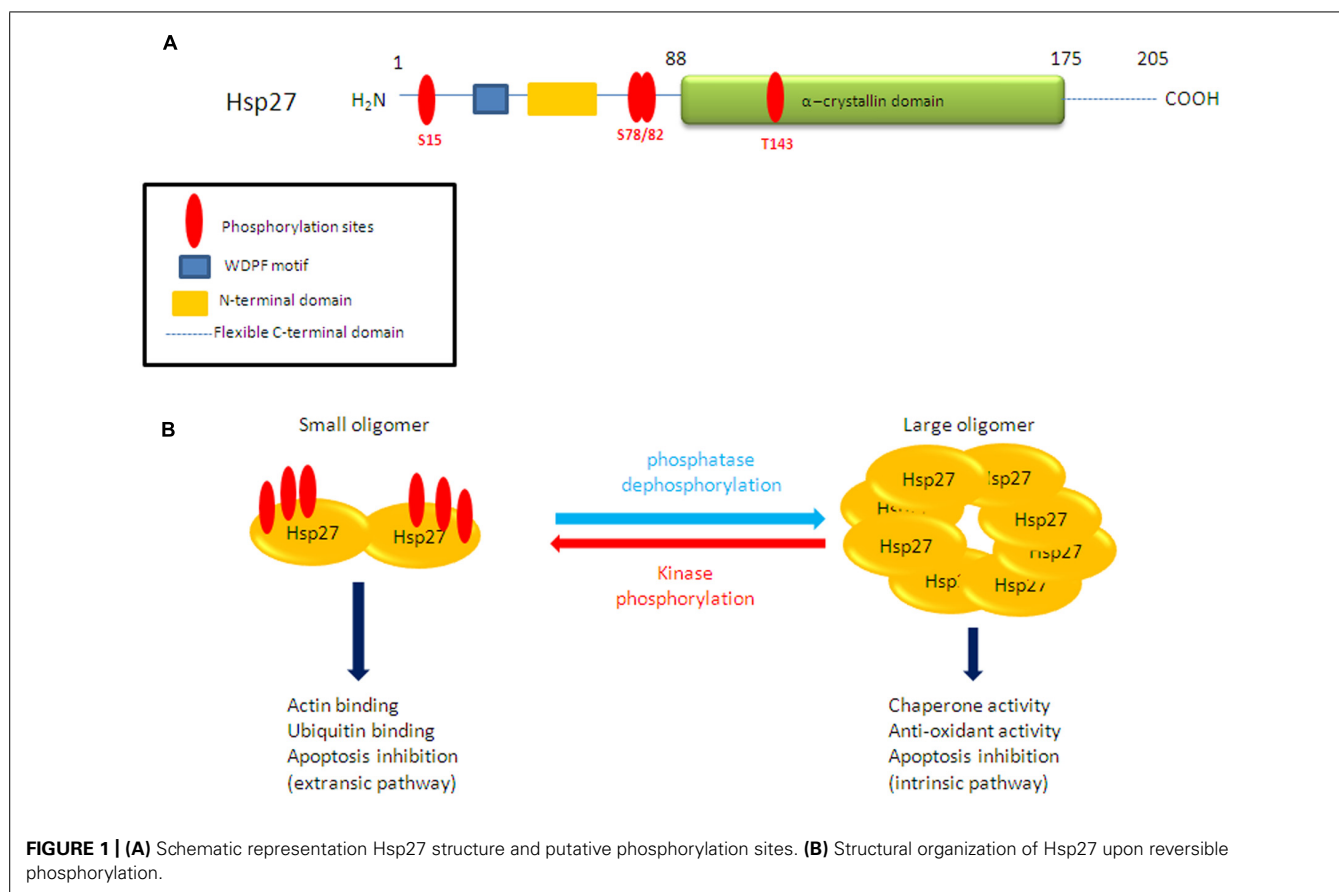
Understanding the mechanisms that control stress-induced survival is critical to explain how tumors frequently resist to treatment and to improve current anti-cancer therapies. Cancer cells are able to cope with stress and escape drug toxicity by regulating heat shock proteins (Hsps) expression and function. Hsp27 (HSPB1), a member of the small Hsp family, represents one of the key players of many signaling pathways contributing to tumorigenicity, treatment resistance, and apoptosis inhibition. Hsp27 is overexpressed in many types of cancer and its functions are regulated by post-translational modifications, such as phosphorylation. Protein phosphorylation is the most widespread signaling mechanism in eukaryotic cells, and it is involved in all fundamental cellular processes. Aberrant phosphorylation of Hsp27 has been associated with cancer but the molecular mechanisms by which it is implicated in cancer development and progression remain undefined. This mini-review focuses on the role of phosphorylation in Hsp27 functions in cancer cells and its potential usefulness as therapeutic target in cancer.

**Keywords: Hsp27, phosphorylation, stress-induced, cancer, apoptosis resistance**

## INTRODUCTION

Protein phosphorylation is the most widespread post-translational modification in eukaryotic cells, and it is involved in all fundamental cellular processes. Reversible phosphorylation based signaling networks are crucial to the cell's capacity to quickly respond to external and internal stimuli. An estimated 30% of cellular proteins are phosphorylated in an estimated total of 1000s of distinct phosphorylation sites (Cohen, 2000). This post-translational modification plays a crucial role in the cellular functions of proteins such as heat shock proteins (Hsps), particularly Hsp27 (HSPB1; Kostenko and Moens, 2009). Hsp27 is an ATP-independent molecular chaperone with well-described tumorigenic and metastatic roles, characterized by its dynamic phosphorylation leading to heterogeneous oligomerization under different conditions such as stress (Jakob et al., 1993; Martin et al., 1999; Garrido, 2002; Gusev et al., 2002; Kato et al., 2002; Koteiche and McHaourab, 2003; Webster, 2003; Taylor and Benjamin, 2005; Acunzo et al., 2012). Unphosphorylated Hsp27 is able to form multimers than can reach 800 kDa (Lentze and Narberhaus, 2004) while phosphorylation results in conformational changes leading to significantly decreased oligomeric size, complex dissociation, and subsequent loss of chaperone activity (Rogalla et al., 1999; Hayes et al., 2009). This supports the idea that Hsp27's reversible structural organization acts as a sensor allowing cells to adapt and eventually overcome lethal conditions by interacting with appropriate protein partners (Arrigo and Gibert, 2012).

It has been well documented that aberrations in protein phosphorylation are closely linked to major diseases such as cancer, diabetes, and rheumatoid arthritis (Radivojac et al., 2008; Watanabe and Osada, 2012; Hao et al., 2013; Nie et al., 2013; Streit et al., 2013). Moreover, Hsp27 overexpression contributes to the malignant progression of cancer cells including increased tumorigenicity, treatment resistance, and apoptosis inhibition (Hsu et al., 2011; Acunzo et al., 2012; Stope et al., 2012). While the aberrant expression of Hsp27 in human cancer have been and is still intensively studied and documented, its phosphorylation state in cancer cells compared to healthy cells are only starting to be examined (Arrigo et al., 2007; Calderwood and Ciocca, 2008; Arrigo and Gibert, 2012). Hsp27 is not the only chaperone whose functions in cancer cells are coordinated by phosphorylation regulation. A recent study identified C-terminal phosphorylation as a key mechanism for the dynamic regulation of Hsp90 and Hsp70 chaperone activity, and binding to co-chaperones to either fold or degrade client proteins (Muller et al., 2013). These co-chaperones are also regulated in a way that favors pro-folding environment in replicating tumor cells and degradation phenotype in non-proliferating cells (Muller et al., 2013). This mini-review briefly summarizes the regulation of Hsp27 by phosphorylation and its functional implications and focuses on the reports describing aberrant Hsp27 phosphorylation linked to cancer. The potential therapeutic strategies aiming at Hsp27 phosphorylation will also be discussed as future perspectives.



## Hsp27 PHOSPHORYLATION AND ASSOCIATED FUNCTIONS IN NORMAL CELLS: AN UP-TO-DATE OVERVIEW

Mapping the phosphorylation sites of Hsp27 showed the involvement of Serine (Ser)-15, Ser-82, Ser-78, and Threonine (Thr)-143 residues (**Figure 1A**). The contribution of single phosphorylation of Hsp27 at either of these sites, in biological processes has not yet been addressed. However, previous studies have shown that Hsp27 oligomerization is regulated by Ser-78 and/or Ser-82 phosphorylation while Ser-15 seems to induce small effect on oligomerization (Lambert et al., 1999; Gusev et al., 2002). Hsp27 phosphorylation/dephosphorylation equilibrium (**Figure 1B**) has been shown to be regulated by signals activating protein kinases and phosphatases. Numerous *in vitro* and *in vivo* studies in different cell types have described the roles of MAPK-activated protein kinase-2,-3,-5 (MK2, MK3, MK5), protein kinase (PK) A, B, C, and D in Hsp27 phosphorylation [for review (Kostenko and Moens, 2009)]. The choice of the kinase seems to depend on the cell type therefore kinase expression levels and the signaling pathway activated. Even though numerous kinases have been described to interact with and/or phosphorylate directly or indirectly Hsp27, controversy exists on the subject and the major kinases have been shown to be MK2, MK5, and PKD (Doppler et al., 2005). When induced upon stress, Hsp27 phosphorylation can be detected within a few minutes (Landry et al., 1992) and in a reversible manner which is controlled by phosphatases. Several studies have revealed the involvement of protein phosphatase 2A

(PP2A; Cairns et al., 1994; Tar et al., 2006) but the involvement of other protein phosphatases is not excluded.

In addition to controlling Hsp27 structural organization and oligomerization, phosphorylation seems to be a key mechanism which favors recognition of specific client proteins associating Hsp27 with specific functions (Arrigo and Gibert, 2012, 2013). Several functions are associated with Hsp27 phosphorylation in normal cells (**Figure 1B**). It has been well described that phosphorylated Hsp27 regulates actin filaments dynamics, in cytoskeleton organization during processes such as cell migration or cell stress (Lavoie et al., 1993; Clarke and Mearow, 2013). Various studies have also shown that Hsp27 overexpression and/or phosphorylation regulates cell cycle and therefore cell proliferation but this appears to be cell-specific. It has been demonstrated that phosphoHsp27 inhibited the MEK/ERK signaling pathway by a mechanism involving both c-Raf activity attenuation and stimulation of MAPK phosphatase-1 (MKP1) through p38 MAPK leading to significant reduction of cyclin D1 levels and subsequent cell cycle arrest (Matsushima-Nishiwaki et al., 2008). Moreover, Hsp27 is known to interact with p53, regulating its transcriptional activity (Venkatakrishnan et al., 2008), therefore having an effect in cell cycle regulation. Last but not least, phosphoHsp27 can prevent apoptosis by protecting cells against heat shock, apoptosis effectors, oxidative stress, and ischemia. Hsp27 can also inactivate Bax and block the release of Smac and cytochrome C (Garrido et al., 2006; Arrigo, 2007; Acunzo et al., 2012). It is important

to note at this point that in cells treated by apoptotic effectors that act on different pathways, Hsp27 has diverse localizations, oligomeric sizes, and phosphorylation states leading to negative regulation of apoptosis (Paul et al., 2010). More precisely, the two apoptotic effectors, etoposide, and Fas antibody, have the tendency to increase Hsp27 native sizes reflecting medium sized and large oligomers accumulation, while staurosporine and cytochalasin D induced Hsp27 in small oligomers (Paul et al., 2010). Hsp27 acts by regulating partner proteins involved in cell death pathways (Havasi et al., 2008; Acunzo et al., 2012; Sanchez-Nino et al., 2012).

## THE RATIONAL OF TARGETING Hsp27 ABERRANT PHOSPHORYLATION IN CANCER

Aberrant expression of Hsp27 in cancer cells has been intensively investigated and is known to be associated with aggressive tumor phenotype, increased therapy resistance, and poor prognosis for the patient. Targeting Hsp27 overexpression in different types of cancers has been shown promising (Rocchi et al., 2006; Cayado-Gutierrez et al., 2013; Lamoureux et al., 2014) but currently no clinical trial has passed phase III (Agensys, 2014). However, less focus has been given to the phosphorylation state of Hsp27 in cancer cells compared to healthy ones. Interestingly, a few recent studies demonstrated that phosphorylation levels of Hsp27 increased in advanced tumors and were correlated to treatment resistance (Taba et al., 2010; Wang et al., 2010; Sakai et al., 2012; Xu et al., 2014). A proteomics study identified phosphoHsp27 as part of the cancer-related phosphoprotein signature of prostate cancer (Chen et al., 2011). In the described study, phosphorylation of Hsp27 occurs upon androgen receptor (AR) activation by ligands, leading to Hsp90 displacement from the AR-complex and translocation of AR to the nucleus. Inhibition of Hsp27 phosphorylation shifted the association of AR with Hsp90 to the E3 ubiquitin ligase MDM2, increased AR degradation, decreased AR transcriptional activity and increased prostate cancer cell apoptotic rates (Chen et al., 2011). In pancreatic and prostate cancer cells, cytoprotection induced by Hsp27 is due, at least in part, to its interaction with eIF4E (eukaryotic translational initiation factor 4E) that increased when Hsp27 is phosphorylated. Hsp27 interaction protects eIF4E from its ubiquitin-proteasome-dependent degradation process, leading to apoptosis resistance induced by castration and chemotherapy (Andrieu et al., 2010; Baylot et al., 2011). A similar mechanism was previously described involving cooperative interactions between ligand-activated AR and Hsp27 phosphorylation that enhance AR stability, shuttling, and transcriptional activity, thereby increasing prostate cancer cell survival (Zoubeidi et al., 2007). Moreover, Hsp27 phosphorylation has been shown to regulate epithelial-mesenchymal transition process and NF- $\kappa$ B activity contributing to the maintenance of breast cancer stem cells (Wei et al., 2011). A comparative phosphoproteomic studies of HER-2/neu positive and -negative breast tumors revealed that Hsp27, one of the identified phosphoproteins, was highly phosphorylated on Ser78 in HER-2/neu positive tumors (Zhang et al., 2007). In MCF-7 cells, phosphoHsp27 plays different roles in regulating p53 pathway and cell survival (Xu et al., 2013). In ovarian and prostate cancers, p38 MAPK-MK2 dependent phosphorylation of Hsp27 was shown to be involved in remodeling of actin filaments required for pro-invasive and pro-metastatic

activities (Gurgis et al., 2014; Pavan et al., 2014). Interestingly, the authors propose targeting the MK2-Hsp27 axis in cancer cells as a strategy to reduce migration and metastasis in cancer cells. In a recent report, it was demonstrated that Hsp27 phosphorylation in liver cancer cells was associated with Hsp27 subcellular localization in the nucleus where it could perform specific functions such as mRNA processing (Bryantsev et al., 2007; Guo et al., 2012). Finally, Taba et al. (Taba et al., 2010) recently showed that phosphorylated Hsp27 played an important role in resistant to Gemcitabine in pancreatic cancer cells and propose phosphoHsp27 as a possible biomarker for predicting response of pancreatic cancer patients to Gemcitabine treatment.

In addition to targeting Hsp27 expression in cancer cells, it therefore appears of particular interest, to block the functions of phosphoHsp27. This approach may lead to new anti-cancer drug discovery. Less specific targeting strategies of the p38-MAPK signaling cascade have shown significant therapeutic potential in the treatment of endocrine resistant breast cancer through inhibition of downstream targets Hsp27 and MAPK (Antoon et al., 2012). Gilbert et al. (Gibert et al., 2011) developed peptide aptamers that specifically bind Hsp27, interfere with its structural organization (dimerization and oligomerization) and impair its anti-apoptotic and cytoprotective functions. We believe that interfering with specific phosphoHsp27-partner protein interactions in cancer cells may represent a promising therapeutic strategy with little or no side effects in normal cells. Targeting phosphoHsp27 in cancer cells constitutes a nascent field research that deserves more exploration in the future.

## CONCLUSION AND FUTURE DIRECTIONS

The future challenge lies in a deeper understanding of Hsp27 phosphorylation state in cancer cells in order to develop and/or improve therapies, specific to cancer cells. The role of Hsp27 phosphorylation in cancer progression has only started to be explored and the few studies published to date that are described in this mini-review suggest that phosphoHsp27 suppresses apoptosis, enhances invasion, and survival of cancer cells. Interestingly, some elements suggest that phosphoHsp27 could present modified subcellular localization which may account for specific roles in cancer cells. We believe that apart from a thorough understanding of Hsp27 phosphorylation state in cancer cells, subcellular localization, and protein partner interactions of phosphoHsp27 are aspects that require further exploration as they will certainly reveal new cancer-specific functions for Hsp27. Interfering with Hsp27's functions should be directed toward cancer cells considering the diversity of functions that Hsp27 exerts in normal cells. Several inhibitors against some Hsp27 kinases have been developed (Anderson et al., 2007; Schlapbach et al., 2008; Lopes et al., 2009) but clinical trials in patients with aberrant Hsp27 phosphorylation are lacking. Proteomics approaches are increasingly used to identify stress-induced chaperone phosphorylation in different pathophysiological conditions and may constitute useful tools for selecting patients who may respond to newly developed therapies.

## ACKNOWLEDGMENTS

The authors received financial support from: French Cancer Institute (InCa, PAIR prostate program #R10111AA), ITMO Cancer

(BioSys call, #A12171AS), Institut National de la Santé et de la Recherche Médicale (Inserm), Association pour la Recherche sur le Cancer (ARC), Agence Nationale pour la Recherche (ANR, Emergence Program #ANR-11-EMMA-0022), Aix-Marseille University, and competitiveness pole Eurobiomed.

## REFERENCES

- Acunzo, J., Katsogiannou, M., and Rocchi, P. (2012). Small heat shock proteins HSP27 (HspB1), alphaB-crystallin (HspB5) and HSP22 (HspB8) as regulators of cell death. *Int. J. Biochem. Cell Biol.* 44, 1622–1631. doi: 10.1016/j.biocel.2012.04.002
- Agensys. (2014). *Clinical Research and Development*. Available at: <http://clinicaltrials.gov/>
- Anderson, D. R., Meyers, M. J., Vernier, W. F., Mahoney, M. W., Kurumbail, R. G., Caspers, N., et al. (2007). Pyrrolopyridine inhibitors of mitogen-activated protein kinase-activated protein kinase 2 (MK-2). *J. Med. Chem.* 50, 2647–2654. doi: 10.1021/jm0611004
- Andrieu, C., Taieb, D., Baylot, V., Ettinger, S., Soubeyran, P., De-Thonel, A., et al. (2010). Heat shock protein 27 confers resistance to androgen ablation and chemotherapy in prostate cancer cells through eIF4E. *Oncogene* 29, 1883–1896. doi: 10.1038/ncr.2009.479
- Antoon, J. W., Bratton, M. R., Guillot, L. M., Wadsworth, S., Salvo, V. A., Elliott, S., et al. (2012). Pharmacology and anti-tumor activity of RWJ67657, a novel inhibitor of p38 mitogen activated protein kinase. *Am. J. Cancer Res.* 2, 446–458.
- Arrigo, A. P. (2007). The cellular “networking” of mammalian Hsp27 and its functions in the control of protein folding, redox state and apoptosis. *Adv. Exp. Med. Biol.* 594, 14–26. doi: 10.1007/978-0-387-39975-1\_2
- Arrigo, A. P., and Gibert, B. (2012). HspB1 dynamic phospho-oligomeric structure dependent interactome as cancer therapeutic target. *Curr. Mol. Med.* 12, 1151–1163. doi: 10.2174/156652412803306693
- Arrigo, A. P., and Gibert, B. (2013). Protein interactomes of three stress inducible small heat shock proteins: HspB1, HspB5 and HspB8. *Int. J. Hyperthermia* 29, 409–422. doi: 10.3109/02656736.2013.792956
- Arrigo, A. P., Simon, S., Gibert, B., Kretz-Remy, C., Nivon, M., Czekalla, A., et al. (2007). Hsp27 (HspB1) and alphaB-crystallin (HspB5) as therapeutic targets. *FEBS Lett.* 581, 3665–3674. doi: 10.1016/j.febslet.2007.04.033
- Baylot, V., Andrieu, C., Katsogiannou, M., Taieb, D., Garcia, S., Giusiano, S., et al. (2011). OGX-427 inhibits tumor progression and enhances gemcitabine chemotherapy in pancreatic cancer. *Cell Death Dis.* 2, e221. doi: 10.1038/cddis.2011.104
- Bryantsev, A. L., Kurchashova, S. Y., Golyshev, S. A., Polyakov, V. Y., Wunderink, H. F., Kanon, B., et al. (2007). Regulation of stress-induced intracellular sorting and chaperone function of Hsp27 (HspB1) in mammalian cells. *Biochem. J.* 407, 407–417. doi: 10.1042/BJ20070195
- Cairns, J., Qin, S., Philp, R., Tan, Y. H., and Guy, G. R. (1994). Dephosphorylation of the small heat shock protein Hsp27 in vivo by protein phosphatase 2A. *J. Biol. Chem.* 269, 9176–9183.
- Calderwood, S. K., and Ciocca, D. R. (2008). Heat shock proteins: stress proteins with Janus-like properties in cancer. *Int. J. Hyperthermia* 24, 31–39. doi: 10.1080/02656730701858305
- Cayado-Gutierrez, N., Moncalero, V. L., Rosales, E. M., Beron, W., Salvatierra, E. E., Alvarez-Olmedo, D., et al. (2013). Downregulation of Hsp27 (HSPB1) in MCF-7 human breast cancer cells induces upregulation of PTEN. *Cell Stress Chaperones* 18, 243–249. doi: 10.1007/s12192-012-0367-x
- Chen, L., Fang, B., Giorgianni, F., Gingrich, J. R., and Beranova-Giorgianni, S. (2011). Investigation of phosphoprotein signatures of archived prostate cancer tissue specimens via proteomic analysis. *Electrophoresis* 32, 1984–1991. doi: 10.1002/elps.201100101
- Clarke, J. P., and Mearow, K. M. (2013). Cell stress promotes the association of phosphorylated HspB1 with F-actin. *PLoS ONE* 8:e68978. doi: 10.1371/journal.pone.0068978
- Cohen, P. (2000). The regulation of protein function by multisite phosphorylation—a 25 year update. *Trends Biochem. Sci.* 25, 596–601. doi: 10.1016/S0968-0004(00)01712-6
- Doppler, H., Storz, P., Li, J., Comb, M. J., and Toker, A. (2005). A phosphorylation state-specific antibody recognizes Hsp27, a novel substrate of protein kinase D. *J. Biol. Chem.* 280, 15013–15019. doi: 10.1074/jbc.C400575200
- Garrido, C. (2002). Size matters: of the small HSP27 and its large oligomers. *Cell Death Differ.* 9, 483–485. doi: 10.1038/sj/cdd/4401005
- Garrido, C., Brunet, M., Didelot, C., Zermati, Y., Schmitt, E., and Kroemer, G. (2006). Heat shock proteins 27 and 70: anti-apoptotic proteins with tumorigenic properties. *Cell Cycle* 5, 2592–2601. doi: 10.4161/cc.5.22.3448
- Gibert, B., Hadchity, E., Czekalla, A., Aloy, M. T., Colas, P., Rodriguez-Lafrasse, C., et al. (2011). Inhibition of heat shock protein 27 (HspB1) tumorigenic functions by peptide aptamers. *Oncogene* 30, 3672–3681. doi: 10.1038/ncr.2011.73
- Guo, K., Gan, L., Zhang, S., Cui, F. J., Cun, W., Li, Y., et al. (2012). Translocation of HSP27 into liver cancer cell nucleus may be associated with phosphorylation and O-GlcNAc glycosylation. *Oncol. Rep.* 28, 494–500. doi: 10.3892/or.2012.1844
- Gurgis, F. M., Ziariaris, W., and Munoz, L. (2014). Mitogen-activated protein kinase-activated protein kinase 2 in neuroinflammation, heat shock protein 27 phosphorylation, and cell cycle: role and targeting. *Mol. Pharmacol.* 85, 345–356. doi: 10.1124/mol.113.090365
- Gusev, N. B., Bogatcheva, N. V., and Marston, S. B. (2002). Structure and properties of small heat shock proteins (sHsp) and their interaction with cytoskeleton proteins. *Biochemistry (Mosc)* 67, 511–519. doi: 10.1023/A:1015549725819
- Hao, J., Li, F., Liu, W., Liu, Q., Liu, S., Li, H., et al. (2013). Phosphorylation of PRAS40-Thr246 involves in renal lipid accumulation of diabetes. *J. Cell. Physiol.* 229, 1069–1077. doi: 10.1002/jcp.24533
- Havasi, A., Li, Z., Wang, Z., Martin, J. L., Botla, V., Ruchalski, K., et al. (2008). Hsp27 inhibits Bax activation and apoptosis via a phosphatidylinositol 3-kinase-dependent mechanism. *J. Biol. Chem.* 283, 12305–12313. doi: 10.1074/jbc.M801291200
- Hayes, D., Napoli, V., Mazurkie, A., Stafford, W. F., and Graceffa, P. (2009). Phosphorylation dependence of hsp27 multimeric size and molecular chaperone function. *J. Biol. Chem.* 284, 18801–18807. doi: 10.1074/jbc.M109.011353
- Hsu, H. S., Lin, J. H., Huang, W. C., Hsu, T. W., Su, K., Chiou, S. H., et al. (2011). Chemoresistance of lung cancer stemlike cells depends on activation of Hsp27. *Cancer* 117, 1516–1528. doi: 10.1002/cncr.25599
- Jakob, U., Gaestel, M., Engel, K., and Buchner, J. (1993). Small heat shock proteins are molecular chaperones. *J. Biol. Chem.* 268, 1517–1520.
- Kato, K., Ito, H., Kamei, K., Iwamoto, I., and Inaguma, Y. (2002). Innervation-dependent phosphorylation and accumulation of alphaB-crystallin and Hsp27 as insoluble complexes in disused muscle. *FASEB J.* 16, 1432–1434. doi: 10.1096/fj.02-0129fj
- Kostenko, S., and Moens, U. (2009). Heat shock protein 27 phosphorylation: kinases, phosphatases, functions and pathology. *Cell. Mol. Life Sci.* 66, 3289–3307. doi: 10.1007/s00018-009-0086-3
- Koteiche, H. A., and McHaourab, H. S. (2003). Mechanism of chaperone function in small heat-shock proteins. Phosphorylation-induced activation of two-mode binding in alphaB-crystallin. *J. Biol. Chem.* 278, 10361–10367. doi: 10.1074/jbc.M211851200
- Lambert, H., Charette, S. J., Bernier, A. F., Guimond, A., and Landry, J. (1999). HSP27 multimerization mediated by phosphorylation-sensitive intermolecular interactions at the amino terminus. *J. Biol. Chem.* 274, 9378–9385. doi: 10.1074/jbc.274.14.9378
- Lamoureux, F., Thomas, C., Yin, M. J., Fazli, L., Zoubeidi, A., and Gleave, M. E. (2014). Suppression of heat shock protein 27 using OGX-427 induces endoplasmic reticulum stress and potentiates heat shock protein 90 inhibitors to delay castrate-resistant prostate cancer. *Eur. Urol.* 66, 145–155. doi: 10.1016/j.eururo.2013.12.019
- Landry, J., Lambert, H., Zhou, M., Lavoie, J. N., Hickey, E., Weber, L. A., et al. (1992). Human HSP27 is phosphorylated at serines 78 and 82 by heat shock and mitogen-activated kinases that recognize the same amino acid motif as S6 kinase II. *J. Biol. Chem.* 267, 794–803.
- Lavoie, J. N., Hickey, E., Weber, L. A., and Landry, J. (1993). Modulation of actin microfilament dynamics and fluid phase pinocytosis by phosphorylation of heat shock protein 27. *J. Biol. Chem.* 268, 24210–24214.
- Lentze, N., and Narberhaus, F. (2004). Detection of oligomerisation and substrate recognition sites of small heat shock proteins by peptide arrays. *Biochem. Biophys. Res. Commun.* 325, 401–407. doi: 10.1016/j.bbrc.2004.10.043
- Lopes, L. B., Flynn, C., Komalavilas, P., Panitch, A., Brophy, C. M., and Seal, B. L. (2009). Inhibition of HSP27 phosphorylation by a cell-permeant



- MAPKAP Kinase 2 inhibitor. *Biochem. Biophys. Res. Commun.* 382, 535–539. doi: 10.1016/j.bbrc.2009.03.056
- Martin, J. L., Hickey, E., Weber, L. A., Dillmann, W. H., and Mestril, R. (1999). Influence of phosphorylation and oligomerization on the protective role of the small heat shock protein 27 in rat adult cardiomyocytes. *Gene Expr.* 7, 349–355.
- Matsushima-Nishiwaki, R., Takai, S., Adachi, S., Minamitani, C., Yasuda, E., Noda, T., et al. (2008). Phosphorylated heat shock protein 27 represses growth of hepatocellular carcinoma via inhibition of extracellular signal-regulated kinase. *J. Biol. Chem.* 283, 18852–18860. doi: 10.1074/jbc.M801301200
- Muller, P., Ruckova, E., Halada, P., Coates, P. J., Hrstka, R., Lane, D. P., et al. (2013). C-terminal phosphorylation of Hsp70 and Hsp90 regulates alternate binding to co-chaperones CHIP and HOP to determine cellular protein folding/degradation balances. *Oncogene* 32, 3101–3110. doi: 10.1038/onc.2012.314
- Nie, H., Zheng, Y., Li, R., Guo, T. B., He, D., Fang, L., et al. (2013). Phosphorylation of FOXp3 controls regulatory T cell function and is inhibited by TNF- $\alpha$  in rheumatoid arthritis. *Nat. Med.* 19, 322–328. doi: 10.1038/nm.3085
- Paul, C., Simon, S., Gibert, B., Viot, S., Manero, F., and Arrigo, A. P. (2010). Dynamic processes that reflect anti-apoptotic strategies set up by HspB1 (Hsp27). *Exp. Cell Res.* 316, 1535–1552. doi: 10.1016/j.yexcr.2010.03.006
- Pavan, S., Musiani, D., Torchiaro, E., Migliardi, G., Gai, M., Di Cunto, F., et al. (2014). HSP27 is required for invasion and metastasis triggered by hepatocyte growth factor. *Int. J. Cancer* 134, 1289–1299. doi: 10.1002/ijc.28464
- Radivojac, P., Baenziger, P. H., Kann, M. G., Mort, M. E., Hahn, M. W., and Mooney, S. D. (2008). Gain and loss of phosphorylation sites in human cancer. *Bioinformatics* 24, i241–i247. doi: 10.1093/bioinformatics/btn267
- Rocchi, P., Jugpal, P., So, A., Sinneman, S., Ettinger, S., Fazli, L., et al. (2006). Small interference RNA targeting heat-shock protein 27 inhibits the growth of prostatic cell lines and induces apoptosis via caspase-3 activation in vitro. *BJU Int.* 98, 1082–1089. doi: 10.1111/j.1464-410X.2006.06425.x
- Rogalla, T., Ehrnsperger, M., Preville, X., Kotlyarov, A., Lutsch, G., Ducasse, C., et al. (1999). Regulation of Hsp27 oligomerization, chaperone function, and protective activity against oxidative stress/tumor necrosis factor  $\alpha$  by phosphorylation. *J. Biol. Chem.* 274, 18947–18956. doi: 10.1074/jbc.274.27.18947
- Sakai, A., Otani, M., Miyamoto, A., Yoshida, H., Furuya, E., and Tanigawa, N. (2012). Identification of phosphorylated serine-15 and -82 residues of HSPB1 in 5-fluorouracil-resistant colorectal cancer cells by proteomics. *J. Proteomics* 75, 806–818. doi: 10.1016/j.jprot.2011.09.023
- Sanchez-Nino, M. D., Sanz, A. B., Sanchez-Lopez, E., Ruiz-Ortega, M., Benito-Martin, A., Saleem, M. A., et al. (2012). HSP27/HSPB1 as an adaptive podocyte antiapoptotic protein activated by high glucose and angiotensin II. *Lab. Invest.* 92, 32–45. doi: 10.1038/labinvest.2011.138
- Schlapbach, A., Feifel, R., Hawtin, S., Heng, R., Koch, G., Moebitz, H., et al. (2008). Pyrrolo-pyrimidones: a novel class of MK2 inhibitors with potent cellular activity. *Bioorg. Med. Chem. Lett.* 18, 6142–6146. doi: 10.1016/j.bmcl.2008.10.039
- Stope, M. B., Schubert, T., Staar, D., Ronnau, C., Streitberger, A., Kroeger, N., et al. (2012). Effect of the heat shock protein HSP27 on androgen receptor expression and function in prostate cancer cells. *World J. Urol.* 30, 327–331. doi: 10.1007/s00345-012-0843-z
- Streit, U., Reuter, H., Bloch, W., Wahlers, T., Schwinger, R. H., and Brixius, K. (2013). Phosphorylation of myocardial eNOS is altered in patients suffering from type 2 diabetes. *J. Appl. Physiol.* (1985) 114, 1366–1374. doi: 10.1152/japplphysiol.00011.2011
- Taba, K., Kuramitsu, Y., Ryozaawa, S., Yoshida, K., Tanaka, T., Maehara, S., et al. (2010). Heat-shock protein 27 is phosphorylated in gemcitabine-resistant pancreatic cancer cells. *Anticancer Res.* 30, 2539–2543.
- Tar, K., Csontos, C., Czokora, I., Olah, G., Ma, S. F., Wadgaonkar, R., et al. (2006). Role of protein phosphatase 2A in the regulation of endothelial cell cytoskeleton structure. *J. Cell. Biochem.* 98, 931–953. doi: 10.1002/jcb.20829
- Taylor, R. P., and Benjamin, I. J. (2005). Small heat shock proteins: a new classification scheme in mammals. *J. Mol. Cell. Cardiol.* 38, 433–444. doi: 10.1016/j.yjmcc.2004.12.014
- Venkatakrishnan, C. D., Dunsmore, K., Wong, H., Roy, S., Sen, C. K., Wani, A., et al. (2008). HSP27 regulates p53 transcriptional activity in doxorubicin-treated fibroblasts and cardiac H9c2 cells: p21 upregulation and G2/M phase cell cycle arrest. *Am. J. Physiol. Heart Circ. Physiol.* 294, H1736–H1744. doi: 10.1152/ajpheart.91507.2007
- Wang, H. Q., Yang, B., Xu, C. L., Wang, L. H., Zhang, Y. X., Xu, B., et al. (2010). Differential phosphoprotein levels and pathway analysis identify the transition mechanism of LNCaP cells into androgen-independent cells. *Prostate* 70, 508–517. doi: 10.1002/pros.21085
- Watanabe, N., and Osada, H. (2012). Phosphorylation-dependent protein-protein interaction modules as potential molecular targets for cancer therapy. *Curr. Drug Targets* 13, 1654–1658. doi: 10.2174/138945012803530035
- Webster, K. A. (2003). Serine phosphorylation and suppression of apoptosis by the small heat shock protein  $\alpha$ B-crystallin. *Circ. Res.* 92, 130–132. doi: 10.1161/01.RES.0000056967.51841.21
- Wei, L., Liu, T. T., Wang, H. H., Hong, H. M., Yu, A. L., Feng, H. P., et al. (2011). Hsp27 participates in the maintenance of breast cancer stem cells through regulation of epithelial-mesenchymal transition and nuclear factor- $\kappa$ B. *Breast Cancer Res.* 13, R101. doi: 10.1186/bcr3042
- Xu, F., Yang, T., Fang, D., Xu, Q., and Chen, Y. (2014). An investigation of heat shock protein 27 and P-glycoprotein mediated multi-drug resistance in breast cancer using liquid chromatography-tandem mass spectrometry-based targeted proteomics. *J. Proteomics* 108C, 188–197. doi: 10.1016/j.jprot.2014.05.016
- Xu, Y., Diao, Y., Qi, S., Pan, X., Wang, Q., Xin, Y., et al. (2013). Phosphorylated Hsp27 activates ATM-dependent p53 signaling and mediates the resistance of MCF-7 cells to doxorubicin-induced apoptosis. *Cell. Signal.* 25, 1176–1185. doi: 10.1016/j.cellsig.2013.01.017
- Zhang, D., Wong, L. L., and Koay, E. S. (2007). Phosphorylation of Ser78 of Hsp27 correlated with HER-2/neu status and lymph node positivity in breast cancer. *Mol. Cancer* 6, 52. doi: 10.1186/1476-4598-6-52
- Zoubeidi, A., Zardan, A., Beraldi, E., Fazli, L., Sowery, R., Rennie, P., et al. (2007). Cooperative interactions between androgen receptor (AR) and heat-shock protein 27 facilitate AR transcriptional activity. *Cancer Res.* 67, 10455–10465. doi: 10.1158/0008-5472.CAN-07-2057

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 April 2014; accepted: 16 September 2014; published online: 06 October 2014.

Citation: Katsogiannou M, Andrieu C and Rocchi P (2014) Heat shock protein 27 phosphorylation state is associated with cancer progression. *Front. Genet.* 5:346. doi: 10.3389/fgene.2014.00346

This article was submitted to *Systems Biology*, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Katsogiannou, Andrieu and Rocchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Phosphorylation of unique domains of Src family kinases

Irene Amata<sup>†</sup>, Mariano Maffei and Miquel Pons\*

Biomolecular NMR Laboratory, Department of Organic Chemistry, University of Barcelona, Barcelona, Spain

## Edited by:

Allegra Via, Sapienza University, Italy  
Andreas Zanzoni, Technological  
Advances for Genomics and Clinics,  
Inserm UMR1090, France

## Reviewed by:

Philipp Selenko, Leibniz Society,  
Germany  
Julie D. Forman-Kay, Hospital for Sick  
Children, Canada

## \*Correspondence:

Miquel Pons, Biomolecular NMR  
Laboratory, Department of Organic  
Chemistry, University of Barcelona,  
Baldri Reixac, 10-12, 08028  
Barcelona, Spain  
e-mail: mpons@ub.edu

## <sup>†</sup>Present address:

Irene Amata, Laboratoire de  
Cristallographie et RMN Biologiques –  
Université Paris Descartes, Paris,  
France

Members of the Src family of kinases (SFKs) are non-receptor tyrosine kinases involved in numerous signal transduction pathways. The catalytic, SH3 and SH2 domains are attached to the membrane-anchoring SH4 domain through the intrinsically disordered “Unique” domains, which exhibit strong sequence divergence among SFK members. In the last decade, structural and biochemical studies have begun to uncover the crucial role of the Unique domain in the regulation of SFK activity. This mini-review discusses what is known about the phosphorylation events taking place on the SFK Unique domains, and their biological relevance. The modulation by phosphorylation of biologically relevant inter- and intra- molecular interactions of Src, as well as the existence of complex phosphorylation/dephosphorylation patterns observed for the Unique domain of Src, reinforces the important functional role of the Unique domain in the regulation mechanisms of the Src kinases and, in a wider context, of intrinsically disordered regions in cellular processes.

**Keywords:** IDPs, IDRs, phosphorylation, unique domain, SFKs, Src

## INTRODUCTION

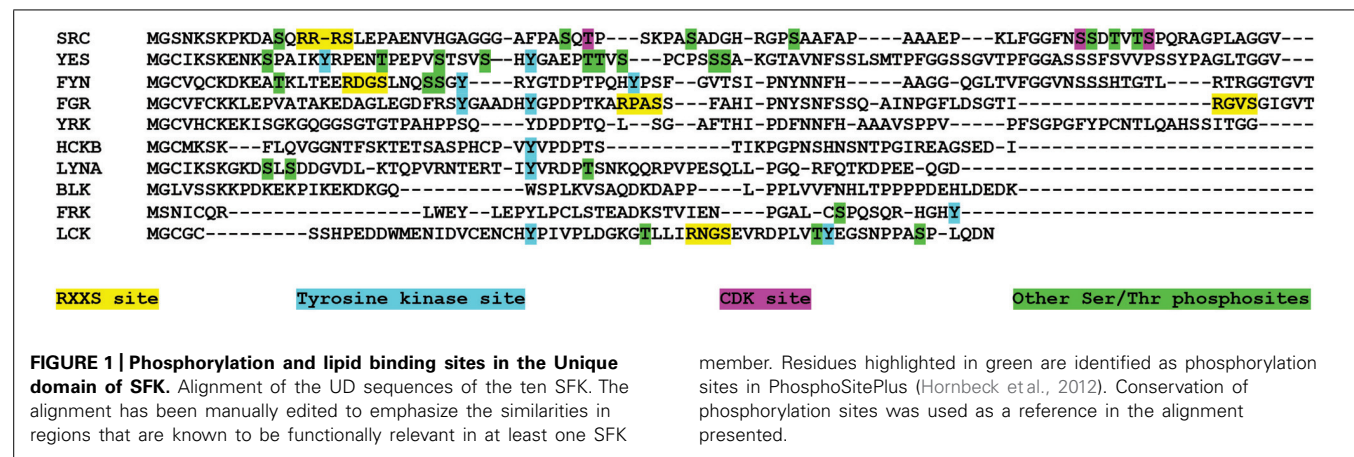
The Src kinase family is composed of 10 proteins: Src, Frk, Lck, Lyn, Blk, Hck, Fyn, Yrk, Fgr, and Yes. Src family kinases (SFKs) are membrane-associated, non-receptor tyrosine kinases that act as important signaling intermediaries regulating a variety of outputs, such as cell proliferation, differentiation, apoptosis, migration, and metabolism. All SFKs share the same domain arrangement: a large catalytic C-terminal domain is preceded by regulatory Src 2 and 3 homology domains (SH2 and SH3, respectively) and the membrane-anchoring SH4 N-terminal region, which contain myristoylation and palmitoylation sites as well as positively charged residues. The SH3 and SH4 domains are linked by an intrinsically disordered segment of 50–90 residues, called the Unique domain (UD) because of the lack of sequence similarity among the different SFKs. However, the UD of each individual SFK member is well conserved between different organisms suggesting a more specific role than that of a simple spacer. **Figure 1** shows the sequences of the UDs of the 10 SFKs, highlighting known phosphorylation sites.

Efforts to find specific UD functions during the last two decades have confirmed an active role of the UD in the regulation of SFK members. In the case of Lck, the UD mediates association with the cytoplasmic tail of CD4 and CD8 $\alpha$  through a zinc clasp structure (Kim et al., 2003). For most of the other SFK members the detailed mechanisms remain obscure. The Unique domains of Fyn and Lyn were observed to be cleaved during induction of apoptosis in intact hematopoietic cells, revealing a novel mechanism for the specific regulation of different SFKs, with important consequences

for their cellular localization and activity (Luciano et al., 2001). Several SFKs contain residues in the N-terminal region that are phosphorylated and dephosphorylated in cellular processes (Joung et al., 1995; Hansen et al., 1997; Johnson et al., 2000). Moreover, swapping the Unique domains of Src and Yes interchanges the functional specificity of the two SFKs (Hoey et al., 2000; Summy et al., 2003; Werdich and Penn, 2005). The versatility and relevance of the active role of the Unique domain in Src function was confirmed by the discovery of binding by the Unique domain to different targets, such as acidic lipids, the SH3 domain, and calmodulin (Pérez et al., 2013). The amino-acid sequence of the UD region responsible for some of these interactions is conserved in Src, Fyn, and Yes. The three proteins are co-expressed in many tissues and are partially redundant, in the sense that deficiency in one of them can be compensated by the action of the other two.

The involvement of SFK signaling in growth, motility and cell survival makes them important oncology targets. Activation of SFKs is tightly regulated in healthy cells (Boggon and Eck, 2004), while the kinase activation is often deregulated in cancer, giving rise to altered cellular shape, function, and growth (Vlahovic and Crawford, 2003). Among the SFKs, Src is the most studied and the most commonly discussed in the context of cancer (Yeatman, 2004). However, there has been growing interest in the other SFKs in both normal physiological and pathological states.

The UD is an intrinsically disordered region (IDR). Intrinsically disordered proteins (IDPs) are relatively more prevalent among signaling and cancer-related proteins (Iakoucheva et al., 2002). The connection between disorder and diseases such as



cancer, neurodegenerative conditions, amyloidosis, cardiovascular disease, and diabetes has been extensively explored in recent reviews (Uversky et al., 2008; Uros et al., 2009). Diseases arising from structural changes in proteins, loosely grouped as “conformational diseases” are caused not only by protein misfolding, but also by failures in post-translational modifications that result in aberrant interactions with physiological partners (Uversky et al., 2008). Remarkably, mutations in disordered regions can result in the loss of important post-translational modification sites, leading to disease (Li et al., 2010). The recognition of the crucial role of IDPs in a number of diseases has heralded a new era in the design of drugs (Chen et al., 2006).

While the catalytic and Src homology domains of SFKs have long been subjects of investigation, the study of the regulatory mechanisms involving the intrinsically disordered UD has only recently started gaining momentum. Phosphorylation events are often associated with the regulation of important functional regions. In this mini-review we discuss the current knowledge about the phosphorylation events taking place in the Unique domains of SFKs.

## PHOSPHORYLATION OF THE UNIQUE DOMAIN OF Lck

The Unique domain linking the SH4 and the SH3 domain of Lck is one of the smallest of the family, with only 60 amino acids. During T-cell activation, Ser59 in the Unique N-terminal region of Lck is phosphorylated (Watts et al., 1993; Winkler et al., 1993). The Unique domain of Lck contains a proline-rich region surrounding Ser59 (<sup>56</sup>PPASP<sup>60</sup>). Joung et al. (1995) found that modifications of Ser59 in the Unique N-terminal domain of tyrosine kinase Lck regulate specificity of its SH2 domain. Later on, Ser59 was identified as a site of *in vivo* mitotic phosphorylation in Lck (Kesavan et al., 2002). More recently, a new target was found for Lck: the protein Nck (Vázquez, 2007), consisting of one SH2 domain and three SH3 domains, which is known to link receptor tyrosine kinases to downstream proteins, and to be active in actin polymerization. It was also reported that Nck binds in T cells to the CD3 subunit of the T-cell antigen receptor (TCR) following TCR engagement. Interestingly, the inter-molecular interaction between Lck and Nck was observed to be disrupted by phosphorylation at Ser59 within the Unique domain of Lck. Wild-type Lck (wt-Lck) and

a Ser59Asp-Lck mutant transfected into Lck-deficient JCaM1.6 cells showed differences in the activation of proximal versus distal signaling events (Vázquez, 2007).

## PHOSPHORYLATION OF THE UNIQUE DOMAIN OF Hck

Hematopoietic cell kinase (Hck) is a potential drug target for cancer and HIV infections. High levels of Hck are associated with drug resistance in chronic myeloid leukemia and Hck activity has been connected with HIV-1 (Tintori et al., 2013). An important insight into the activation mechanism of this SFK member was the discovery that Hck is capable of performing autophosphorylation in its UD. Autophosphorylation of the activation loop in the kinase domain is a common process for SFK activation. Autophosphorylation of recombinant Hck leads to a 20-fold increase in its specific enzymatic activity (Johnson et al., 2000). Hck was found to autophosphorylate readily to a stoichiometry of 1.3 mol of phosphate per mol of enzyme, indicating that the kinase autophosphorylated at more than one site. In particular, Johnson and collaborators discovered — *in vitro* as well as *in vivo* — that Hck can undergo autophosphorylation at two different sites: Tyr388, which is located at the consensus autophosphorylation site in the well-characterized activation loop of similar kinases, and Tyr29, which is located in its intrinsically disordered Unique domain. By inspecting the activities and levels of phosphorylation of recombinant Hck mutants containing either the point mutation Tyr388Phe or Tyr29Phe, they demonstrated that phosphorylation at Tyr29 makes a crucial contribution to the activation of Hck through its autophosphorylation. Regulation of the catalytic activity by phosphorylation at Tyr29 in the Unique domain of Hck suggests that autophosphorylation within the N-terminal Unique region may also be an additional mechanism of regulation of other Src family tyrosine kinases.

## TYROSINE PHOSPHORYLATION IN OTHER UNIQUE DOMAINS: Lyn, Yes, Fgr, AND Frk

Interestingly, the tyrosine residue in position 32 of Lyn is located in an 8-residue-long sequence common to Lyn and Hck. EGFR phosphorylates the p56 isoform of Lyn, p56<sup>Lyn</sup>, at Tyr32, which then phosphorylates MCM7, a licensing factor critical for DNA replication. Phosphorylation at Tyr600 of MCM7 increases its association with other minichromosome maintenance complex



proteins, thereby promoting DNA synthesis complex assembly and cell proliferation. Both p56<sup>Lyn</sup> Tyr32 and MCM7 Tyr600 phosphorylation are enhanced in proliferating cells and correlated with poor survival of breast cancer patients (Huang et al., 2013).

All Unique domains contain aromatic residues, a rare feature in intrinsically disordered domains. A tyrosine residue is present between positions 25 and 34 in seven SFKs, and is known to be phosphorylated in five of them: Lyn, Hck, Lck (Hornbeck et al., 2012), Yes (Ariki et al., 1997), and Frg (Oppermann et al., 2009). Phosphorylation of Tyr34 of Fgr has been observed by mass spectrometry mainly in samples from leukemia patients.

Erk is the only SFK that is not myristoylated. Although it contains three tyrosine residues in its Unique domain, phosphorylation has only been reported for Tyr46, at the interface between the Unique and SH3 domains. The modification was observed by mass spectrometry mainly in samples from lung, liver, and gastric cancer (Stokes et al., 2012).

### PHOSPHORYLATION OF THE UNIQUE DOMAIN OF Fyn

Fyn is ubiquitously expressed together with Src and Yes, whereas other members of the Src family are expressed only in specific cell types. Fyn is primarily localized to the cytoplasmic leaflet of the plasma membrane, where it phosphorylates tyrosine residues on key targets involved in a variety of different signaling pathways. Fyn characterization has been mainly focused on its immune and neurological function. However, Fyn has also been recognized as an important mediator of mitogenic signaling and regulator of cell-cycle entry, growth, and proliferation, integrin-mediated interactions as well as cell adhesion and migration (Kawakami et al., 1988; Chen et al., 2001; Li et al., 2003; Kim et al., 2010). In particular, Fyn is over-expressed in several cancers, such as glioblastoma multiforme and melanoma. The role of Fyn over-expression in these systems, however, has not been well defined as yet.

Fyn presents multiple phosphorylation sites that can affect its kinase activity. Activation of Platelet-Derived Growth Factor (PDGF)  $\beta$ -receptor by binding to PDGF leads to activation of a member of the SFK (Alonso et al., 1995). The Unique domain of Fyn was found to be phosphorylated at Tyr28 (Hansen et al., 1997). The functional role of this phosphorylation was confirmed by the observation of significantly reduced activation following PDGF stimulation of a Fyn mutant in which Tyr28 was replaced by phenylalanine (Hansen et al., 1997). It was also proposed that the autophosphorylation of the N-terminal tyrosine residues plays a key role in the activation of Fyn, by complementing the PDGF receptor-induced phosphorylation of Tyr28.

Serine 21 within the UD of Fyn is part of a RxxS motif targeted by protein kinase A (PKA). Mutation of Ser21 to Alanine (Ser21Ala-Fyn) blocks PKA phosphorylation of Fyn and alters its tyrosine kinase activity (Yeo et al., 2011). In the same work, the authors showed that the over-expression of Ser21Ala-Fyn mutant in cells lacking Src/Yes/Fyn kinases (SYF cells) led to decreased tyrosine phosphorylation of focal adhesion kinase, resulting in reduced focal adhesion targeting, and slow lamellipodia dynamics and cell migration. These important changes in cell motility demonstrate a key role of UD phosphorylation at Ser21 in the Fyn

kinase activity that controls assembly and disassembly of focal adhesions in response to signals arising from cell-extracellular matrix interactions.

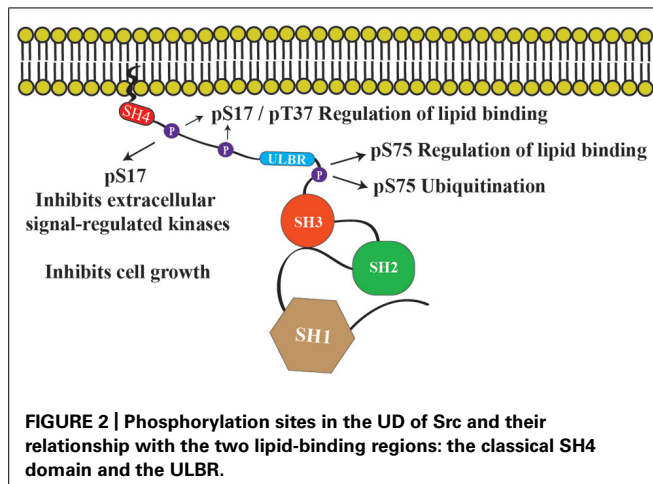
Protein kinase A is a crucial component of integrin-mediated signaling pathways. Interaction of cells with their substrate or adhesion to other cells leads to activation of PKA (Whittard and Akiyama, 2001), and induce the phosphorylation of a variety of protein substrates (Walsh and Van Patten, 1994). Four SFKs (Fyn, Src, Lck, and Fgr) contain a conserved RxxS motif suggesting they can be regulated by PKA.

### PHOSPHORYLATION OF THE UNIQUE DOMAIN OF Src

Src is a non-receptor protein tyrosine kinase with a key role in regulating cell-to-matrix adhesion, migration, and junctional stability (Frame, 2004). Thus, precise regulation of Src activity is critical for normal cell growth. The inactive state of Src is obtained by phosphorylated tyrosine near the C-terminus of Src (Tyr530 in mammalian Src; Tyr527 in chicken Src), which is recognized by its SH2 domain, while the SH3 domain interacts with a polyproline motif located in the linker region between the SH2 and kinase domains; these intramolecular interactions restrict access to the kinase domain (Xu et al., 1997). Dephosphorylation of Tyr530 is followed by autophosphorylation at Tyr419, leading to full activation of the kinase. Active Src may be deactivated by rephosphorylation of Tyr530 by C-terminal Src kinase (Csk; Piwnicka-Worms et al., 1987; Nada et al., 1991; Brown, 1996).

Although the phosphorylation of Src at Ser17 by PKA (cAMP-dependent protein kinase) is a well-characterized process, its biological significance remains unclear (Obara et al., 2004). Potential roles in protein-protein interactions or cellular localization have been postulated for this phosphorylation site. For instance, it has been observed that the treatment of 3T3 fibroblasts with PDGF results in the translocation of Src from the plasma membrane to the cytosol, concomitant with an increase in phosphorylation of Ser17 by PKA (Walker et al., 1993). This observation suggests that this phosphorylation could interfere with the electrostatic interactions that act to anchor Src to the lipid bilayer. PKA phosphorylation of Src at Ser17 is also required in cAMP activation of Rap1, inhibition of extracellular signal-regulated kinases, and inhibition of cell growth, although the mechanism by which this phosphorylation mediates these processes is not known (Obara et al., 2004).

Interaction of Src with lipids is not restricted to the SH4 domain (Figure 2). The unique lipid binding region (ULBR) was discovered following NMR observations that revealed a partially structured region within the UD of Src (Pérez et al., 2009). While phosphorylation of Ser17 by PKA disturbed the interaction of the SH4 domain with lipids, phosphorylation of Thr37 and Ser75 by p25-Cdk5 decreased lipid binding by the ULBR (Pérez et al., 2009). Cross-effects were observed, suggesting a cooperative interaction of the two lipid binding regions with membranes. Conformational effects of phosphorylation at Ser17, Thr37, and Ser75 in the isolated UD are strictly local, indicating that electrostatic repulsion of the phosphorylated residues with the negatively charged lipids is the main mechanism by which lipid binding is disrupted.



The ULBR major regulatory role of the UD was shown experimentally in *Xenopus laevis* oocytes. Mutations abolishing lipid binding by the ULBR but not affecting the SH4 domain resulted in a conditional lethal phenotype after progesterone induced maturation (Pérez et al., 2013).

Previous results had identified a functional role for Thr37 and Ser75 (Thr34 and Ser72 in chicken Src) as well as Thr46 in chicken Src (with no correspondence in humans). These residues are phosphorylated by cyclin-dependent kinase 1 (Cdk1/Cdc2) during mitosis. These phosphorylations were found to activate Src by disrupting the interaction between the SH2 domain and Tyr527/Tyr530 (chicken/human) and facilitating the dephosphorylation by protein tyrosine phosphatases (Shenoy et al., 1992; Stover et al., 1994).

In addition (or possibly related to) its effect on lipid binding, phosphorylation of Ser75 is important in other aspects of Src regulation. Mitosis-independent phosphorylation of this site was observed in neurons and in certain cancer cell lines; this phosphorylation was shown to be due to Cdk5, a widely distributed proline-directed kinase with a substrate specificity similar to Cdk1 (Kato and Maeda, 1999). Active Src is reported to be irreversibly destroyed by Cullin-5-dependent ubiquitination and proteosomal degradation (Hakak and Martin, 1999; Laszlo and Cooper, 2009). More recently, it was shown that phosphorylation at Ser75 promotes the ubiquitin-dependent degradation of Src (Pan et al., 2011). Ser75-Src phosphorylation in epithelial cells was found to depend on the activation state of Src: only active Src was phosphorylated and eventually marked for ubiquitination. Thus, Cdk5-dependent phosphorylation of Ser75 within the UD of Src represents a mechanism to restrict the availability of active Src (Pan et al., 2011). Sequence alignment suggests that serine residues are present in homologous positions of Yes and possibly Frk.

NMR and mass-spectrometry studies of human Src UD added to *X. laevis* egg extracts showed the phosphorylation of Ser75 and Ser69, as well as Ser 17 (Pérez et al., 2013). Phosphorylation of Ser75 and Ser69 seem to be mutually interfering *in vitro*. Interestingly, *X. laevis* Src has a glycine residue at position 69 and phosphorylation of Ser69 in human Src had only been

previously detected by mass spectrometry in extracts of cancer lines HCT116 and MDA-MB-435S (Oppermann et al., 2009). It is tempting to speculate that pathological phosphorylation of Ser 69 in human cells could result in decreased phosphorylation of Ser75 and reduced degradation of active Src, leading to an oncogenic phenotype caused by Src overactivation.

The interplay between various phosphorylation sites within the UD emphasizes its role as a signaling integration hub. Further input signals, in addition to phosphorylation, include calcium-dependent interaction with calmodulin and the allosteric interaction with Src SH3 domain (Pérez et al., 2013; Figure 2). A further level of cross-talk between various phosphorylation events was observed when the time-dependent phosphorylation of the UD was studied by real-time NMR in cell extracts (Amata et al., 2013). In these experiments it was observed that the activity of a PKA-like kinase, which phosphorylates Ser17, also repressed the phosphatase(s) that catalyzed the dephosphorylation of Ser75. Thus, inhibition of PKA activity resulted in dephosphorylation of both sites. Similarly, inhibition of Cdk activity resulted in a reduction in the steady-state phosphorylation of both sites. On the other hand, addition of PKA caused a robust phosphorylation of Ser75 by preventing the action of the phosphatases that reverse the effect of Cdk. These results show that the phosphorylation state of the UD of Src represents a sensor of the kinase-phosphatase network active at any given moment in the cell. Experimental access to this information is possible by using real-time NMR techniques (Selenko et al., 2008; Theillet et al., 2012; Thongwichian and Selenko, 2012; Amata et al., 2013). In contrast to folded domains, intrinsically disordered domains, like the UD of Src, provide *in vivo* NMR resolution comparable to that obtainable *in vitro* (Kosol et al., 2013).

## CONCLUDING REMARKS

The study of the UD of SFKs, particularly that of Src, provides an example of multilevel regulation through phosphorylation of a membrane-bound intrinsically disordered domain. The occurrence of membrane tethering among IDRs is very common, although its relevance is not always recognized. The results arising from studies of UD should stimulate further research to uncover similar mechanisms beyond SFKs. In particular, switchable internal lipid binding sites (the ULBR in the case of Src) in disordered domains anchored to membranes by a second, more stable, binding site close to the protein termini (the SH4 domain in the case of SFKs) can modulate the position of the active domains (the kinase or regulatory domains in the case of SFKs) with respect to the membrane surface. Lipid binding by the internal site can switch off the access of the active domains to target sites of membrane-anchored substrates/partners located beyond a particular distance from the membrane surface or, conversely, facilitate the interaction with sites near the membrane surface. Phosphorylation or other modifications preventing the interaction of the internal site would result in a release of the active domains, which can then reach sites located further apart from the membrane surface or disfavor the interaction with membrane-proximal sites. This represents a new compartmentalization mechanism based on the relative position of interacting sites with respect to the membrane

surface. This mechanism enables modulating the functional interactions between two proteins anchored next to each other in the membrane.

## ACKNOWLEDGMENTS

This work was supported in part by funds from Fundació Marató TV3, the Spanish Government (BIO2010-15683), the Catalan Government (2009SGR1352), and the European Union 7FP (BioNMR contract 261863, and Marie Curie Action - COFUND to Irene Amata).

## REFERENCES

- Alonso, G., Koegl, M., Mazurenko, N., and Courtneidge, S. A. (1995). Sequence requirements for binding of src family tyrosine kinases to activated growth factor receptors. *J. Biol. Chem.* 270, 9840–9848. doi: 10.1074/jbc.270.17.9840
- Amata, I., Maffei, M., Igea, A., Gay, M., Vilaseca, M., Nebreda, A. R., et al. (2013). Multi-phosphorylation of the intrinsically disordered unique domain of c-Src studied by in-cell and real-time NMR spectroscopy. *Chembiochemistry* 14, 1820–1827. doi: 10.1002/cbic.201300139
- Ariki, M., Tanabe, O., Usui, H., Hayashi, H., Inoue, R., Nishito, Y., et al. (1997). Identification of autophosphorylation sites in c-Yes purified from rat liver plasma membranes. *J. Biochem.* 121, 104–111.
- Boggon, T. J., and Eck, M. J. (2004). Structure and regulation of Src family kinases. *Oncogene* 23, 7918–7927. doi: 10.1038/sj.onc.1208081
- Brown, E. (1996). A signaling pathway to translational control. *Cell* 86, 517–520. doi: 10.1016/S0092-8674(00)80125-7
- Chen, J. W., Romero, P., Uversky, V. N., and Dunker, A. K. (2006). Conservation of intrinsic disorder in protein domains and families: I. A database of conserved predicted disordered regions. *J. Proteome Res.* 5, 879–887. doi: 10.1021/pr060048x
- Chen, R., Kim, O., Yang, J., Sato, K., Eisenmann, K. M., McCarthy, J., et al. (2001). Regulation of Akt/PKB activation by tyrosine phosphorylation. *J. Biol. Chem.* 276, 31858–31862. doi: 10.1074/jbc.C100271200
- Frame, M. C. (2004). Newest findings on the oldest oncogene; how activated src does it. *J. Cell Sci.* 117(Pt 7), 989–998. doi: 10.1242/jcs.01111
- Hakak, Y., and Martin, S. G. (1999). Ubiquitin-dependent degradation of active Src. *Curr. Biol.* 9, 1039–1042. doi: 10.1016/S0960-9822(99)80453-9
- Hansen, K., Alonso, G., Courtneidge, S. A., Rönnstrand, L., and Heldin, C.-H. H. (1997). PDGF-induced phosphorylation of Tyr28 in the N-terminus of Fyn affects Fyn activation. *Biochem. Biophys. Res. Commun.* 241, 355–362. doi: 10.1006/bbrc.1997.7743
- Hoey, J. G., Summy, J., and Flynn, D. C. (2000). Chimeric constructs containing the SH4/Unique domains of cYes can restrict the ability of Src527F to upregulate heme oxygenase-1 expression efficiently. *Cell. Signal.* 12, 691–701. doi: 10.1016/S0898-6568(00)00116-9
- Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., et al. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270. doi: 10.1093/nar/gkr1122
- Huang, T. H., Huo, L., Wang, Y. N., Xia, W., Wei, Y., Chang, S. S., et al. (2013). Epidermal growth factor receptor potentiates MCM7-mediated DNA replication through tyrosine phosphorylation of Lyn kinase in human cancers. *Cancer Cell* 23, 796–810. doi: 10.1016/j.ccr.2013.04.027
- Iakouchcheva, L. M., Brown, C. J., Lawson, D. J., Obradović, Z., and Dunker, K. A. (2002). Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* 323, 573–584. doi: 10.1016/S0022-2836(02)00969-5
- Johnson, T. M., Williamson, N. A., Scholz, G., Jaworowski, A., Wettenhall, R. E., Dunn, A. R., et al. (2000). Modulation of the catalytic activity of the Src family tyrosine kinase Hck by autophosphorylation at a novel site in the unique domain. *J. Biol. Chem.* 275, 33353–33364. doi: 10.1074/jbc.M002022200
- Joung, I., Kim, T., Stolz, L. A., Payne, G., Winkler, D. G., Walsh, C. T., et al. (1995). Modification of Ser59 in the unique N-terminal region of tyrosine kinase p56lck regulates specificity of its Src homology 2 domain. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5778–5782. doi: 10.1073/pnas.92.13.5778
- Kato, G., and Maeda, S. (1999). Neuron-specific phosphorylation Cdk5 kinase is responsible of c-Src at Ser75 in human for mitosis-independent Y79 retinoblastoma. *J. Biochem.* 126, 957–961. doi: 10.1093/oxfordjournals.jbchem.a022540
- Kawakami, T., Kawakami, Y., Aaronson, S. A., and Robbins, K. C. (1988). Acquisition of transforming properties by FYN, a normal SRC-related human gene. *Proc. Natl. Acad. Sci. U.S.A.* 85, 3870–3874. doi: 10.1073/pnas.85.11.3870
- Kesavan, K. P., Isaacson, C. C., Ashendel, C. L., Geahlen, R. L., and Harrison, M. L. (2002). Characterization of the in vivo sites of serine phosphorylation on Lck identifying serine 59 as a site of mitotic phosphorylation. *J. Biol. Chem.* 277, 14666–14673. doi: 10.1074/jbc.M111911200
- Kim, H.-J., Warren, J. T., Kim, S.-Y., Chappel, J. C., DeSelm, C. J., Ross, F. P., et al. (2010). Fyn promotes proliferation, differentiation, survival and function of osteoclast lineage cells. *J. Cell. Biochem.* 111, 1107–1113. doi: 10.1002/jcb.22841
- Kim, P. W., Sun, Z.-Y. J., Blacklow, S. C., Wagner, G., and Eck, M. J. (2003). A zinc clasp structure tethers Lck to T cell coreceptors CD4 and CD8. *Science* 301, 1725–1728. doi: 10.1126/science.1085643
- Kosol, S., Contreras-Martos, S., Cedeño, C., and Tompa, P. (2013). Structural characterization of intrinsically disordered proteins by NMR spectroscopy. *Molecules* 18, 10802–10828. doi: 10.3390/molecules180910802
- Laszlo, G. S., and Cooper, J. A. (2009). Restriction of Src activity by Cullin-5. *Curr. Biol.* 19, 157–162. doi: 10.1016/j.cub.2008.12.007
- Li, S., Iakouchcheva, L. M., Mooney, S. D., and Radivojac, P. (2010). Loss of post-translational modification sites in disease. *Pac. Symp. Biocomput.* 337–347. doi: 10.1142/9789814295291\_0036
- Li, X., Yang, Y., Hu, Y., Dang, D., Regezi, J., Schmidt, B. L., et al. (2003). Alphavbeta6-Fyn signaling promotes oral cancer progression. *J. Biol. Chem.* 278, 41646–41653. doi: 10.1074/jbc.M306274200
- Luciano, E., Ricci, J.-E., and Auberger, P. (2001). Cleavage of Fyn and Lyn in their N-terminal unique regions during induction of apoptosis: a new mechanism for Src kinase regulation. *Oncogene* 20, 4935–4941. doi: 10.1038/sj.onc.1204661
- Nada, S., Okada, M., MacAuley, A., Cooper, J. A., and Nakagawa, H. (1991). Cloning of a complementary DNA for a protein-tyrosine kinase that specifically phosphorylates a negative regulatory site of p60c-src. *Nature* 351, 69–72. doi: 10.1038/351069a0
- Obara, Y., Labudda, K., Dillon, T. J., and Stork, P. J. S. (2004). PKA phosphorylation of Src mediates Rap1 activation in NGF and cAMP signaling in PC12 cells. *J. Cell Sci.* 117, 6085–6094. doi: 10.1242/jcs.01527
- Oppermann, F. S., Gnad, F., Olsen, J. V., Hornberger, R., Greff, Z., Kéri, G., et al. (2009). Large-scale proteomics analysis of the human kinome. *Mol. Cell. Proteomics* 8, 1751–1764. doi: 10.1074/mcp.M800588-MCP200
- Pan, Q., Qiao, F., Gao, C., Norman, B., Optican, L., and Zelenka, P. S. (2011). Cdk5 targets active Src for ubiquitin-dependent degradation by phosphorylating Src(S75). *Cell Mol. Life. Sci.* 68, 3425–3436. doi: 10.1007/s00018-011-0638-1
- Pérez, Y., Gairi, M., Pons, M., and Bernadó, P. (2009). Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* 391, 136–148. doi: 10.1016/j.jmb.2009.06.018
- Pérez, Y., Maffei, M., Igea, A., Amata, I., Gairi, M., Nebreda, A. R., et al. (2013). Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.* 3, 1295. doi: 10.1038/srep01295
- Piwnicka-Worms, H., Saunders, K. B., Roberts, T. M., Smith, A. E., and Cheng, S. H. (1987). Tyrosine phosphorylation regulates the biochemical and biological properties of pp60c-src. *Cell* 49, 75–82. doi: 10.1016/0092-8674(87)90757-4
- Selenko, P., Frueh, D. P., Elsaesser, S. J., Haas, W., Gygi, S. P., and Wagner, G. (2008). In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat. Struct. Mol. Biol.* 15, 321–329. doi: 10.1038/nsmb.1395
- Shenoy, S., Chackalaparampil, I., Bagrodia, S., Lin, P.-H., and Shalloway, D. (1992). Role of p34cdc2-mediated phosphorylations in two-step activation of pp60c-src during mitosis. *Proc. Natl. Acad. Sci. U.S.A.* 89, 7237–7241. doi: 10.1073/pnas.89.15.7237
- Stokes, M. P., Farnsworth, C. L., Moritz, A., Silva, J. C., Jia, X., Lee, K. A., et al. (2012). PTMScan direct: identification and quantification of peptides from critical signaling proteins by immunoaffinity enrichment coupled with LC-MS/MS. *Mol. Cell. Proteomics* 11, 187–201. doi: 10.1074/mcp.M111.015883
- Stover, R. D., Liebetanz, J., and Lydon, B. N. (1994). Cdc2-mediated Modulation of pp60c-src activity. *J. Biol. Chem.* 269, 26885–26889.
- Summy, J. M., Qian, Y., Jiang, B.-H., Guappone-Koay, A., Gatesman, A., Shi, X., et al. (2003). The SH4-Unique-SH3-SH2 domains dictate specificity in signaling

- that differentiate c-Yes from c-Src. *J. Cell Sci.* 116(Pt. 12), 2585–2598. doi: 10.1242/jcs.00466
- Theillet, F.-X., Smet-Nocca, C., Liokatis, S., Thongwichian, R., Kosten, J., Yoon, M.-K., et al. (2012). Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* 54, 217–236. doi: 10.1007/s10858-012-9674-x
- Thongwichian, R., and Selenko, P. (2012). In-cell NMR in *Xenopus laevis* oocytes. *Methods Mol. Biol.* 895, 33–41. doi: 10.1007/978-1-61779-927-3\_3
- Tintori, C., Laurenzana, I., La Rocca, F., Falchi, F., Carraro, F., Ruiz, A., et al. (2013). Identification of Hck inhibitors as hits for the development of antileukemia and anti-HIV agents. *Chem. Med. Chem.* 8, 1353–1360. doi: 10.1002/cmdc.2013.00204
- Uros, M., Chritopher, O. J., Keith, D. A., Zoran, O., and Uversky, V. N. (2009). Unfoldomics of human genetic diseases: illustrative examples of ordered and intrinsically disordered members of the human diseasome. *Protein Pept. Lett.* 16, 1533–1547. doi: 10.2174/092986609789839377
- Uversky, V. N., Oldfield, C. J., and Dunker, K. A. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* 37, 215–246. doi: 10.1146/annurev.biophys.37.032807.125924
- Vázquez, M. L. (2007). *Biological consequences of the Phosphorylation of Serine 59 on the Tyrosine Kinase Lck*. PhD Thesis, Purdue University, West Lafayette, IN. Available at: <http://search.proquest.com/docview/304827306?accountid=28979>
- Vlahovic, G., and Crawford, J. (2003). Activation of tyrosine kinases in cancer. *Oncologist* 8, 531–538. doi: 10.1634/theoncologist.8-6-531
- Walker, F., DeBlaquiere, J., and Burgess, A. W. (1993). Translocation of pp60C-src from the plasma membrane to the cytosol after stimulation by platelet-derived growth factor. *J. Biol. Chem.* 268, 19552–19558.
- Walsh, A. D., and Van Patten, M. S. (1994). Multiple pathway signal transduction by the cAMP-dependent protein kinase. *FASEB J.* 8, 1227–1236.
- Watts, J. D., Sanghera, J. S., Pelech, S. L., and Aebersold, R. (1993). Phosphorylation of serine 59 of p56lck in activated T cells. *J. Biol. Chem.* 268, 23275–23282.
- Werdich, X. Q., and Penn, J. S. (2005). Src, Fyn and Yes play differential roles in VEGF-mediated endothelial cell events. *Angiogenesis* 8, 315–326. doi: 10.1007/s10456-005-9021-x
- Whittard, J. D., and Akiyama, S. K. (2001). Positive regulation of cell-cell and cell-substrate adhesion by protein kinase A. *J. Cell Sci.* 114, 3265–3272.
- Winkler, D. G., Park, I., Kim, T., Payne, N. S., Walsht, C. T., Strominger, J. L., et al. (1993). Phosphorylation Ser-42 and Ser-59 in the N-terminal region of the tyrosine kinase p56lck. *Proc. Natl. Acad. Sci. U.S.A.* 90, 5176–5180. doi: 10.1073/pnas.90.11.5176
- Xu, W., Harrison, C. S., and Eck, J. M. (1997). Three-dimensional structure of the tyrosine kinase c-Src. *Nature* 385, 595–602. doi: 10.1038/385595a0
- Yeaman, T. J. (2004). A renaissance for SRC. *Nat. Rev. Cancer* 4, 470–480. doi: 10.1038/nrc1366
- Yeo, M. G., Oh, H. J., Cho, H.-S., Chun, J. S., Marcantonio, E. E., and Song, W. K. (2011). Phosphorylation of Ser 21 in Fyn regulates its kinase activity, focal adhesion targeting, and is required for cell migration. *J. Cell Physiol.* 226, 236–247. doi: 10.1002/jcp.22335

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 March 2014; accepted: 29 May 2014; published online: 30 June 2014.

Citation: Amata I, Maffei M and Pons M (2014) Phosphorylation of unique domains of Src family kinases. *Front. Genet.* 5:181. doi: 10.3389/fgene.2014.00181

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Amata, Maffei and Pons. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Toward a systems-level view of dynamic phosphorylation networks

Robert H. Newman<sup>1\*</sup>, Jin Zhang<sup>2,3,4,5\*</sup> and Heng Zhu<sup>2,6\*</sup>

<sup>1</sup> Department of Biology, North Carolina Agricultural and Technical State University, Greensboro, NC, USA

<sup>2</sup> Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>3</sup> The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>4</sup> Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>5</sup> Department of Chemical and Biomolecular Engineering, Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>6</sup> High-Throughput Biology Center, Institute for Basic Biomedical Sciences, Johns Hopkins University, Baltimore, MD, USA

## Edited by:

Andreas Zanzoni, Inserm TAGC  
UMR1090, France  
Allegra Via, Sapienza University, Italy

## Reviewed by:

Tamar Geiger, Tel-Aviv University,  
Israel  
Stephen Michnick, University of  
Montreal, Canada

## \*Correspondence:

Robert H. Newman, Department of  
Biology, North Carolina Agricultural  
and Technical State University,  
1601 E. Market St., Greensboro,  
NC 27411, USA  
e-mail: rhnewman@ncat.edu;  
Jin Zhang, Department of  
Pharmacology and Molecular  
Sciences, Johns Hopkins University  
School of Medicine, 725 N.  
Wolfe St., Baltimore, MD 21205,  
USA  
e-mail: jzhang32@jhmi.edu;  
Heng Zhu, Department of  
Pharmacology and Molecular  
Sciences, Johns Hopkins University  
School of Medicine, 733 N.  
Broadway, Baltimore, MD 21205,  
USA  
e-mail: hzhu4@jhmi.edu

To better understand how cells sense and respond to their environment, it is important to understand the organization and regulation of the phosphorylation networks that underlie most cellular signal transduction pathways. These networks, which are composed of protein kinases, protein phosphatases and their respective cellular targets, are highly dynamic. Importantly, to achieve signaling specificity, phosphorylation networks must be regulated at several levels, including at the level of protein expression, substrate recognition, and spatiotemporal modulation of enzymatic activity. Here, we briefly summarize some of the traditional methods used to study the phosphorylation status of cellular proteins before focusing our attention on several recent technological advances, such as protein microarrays, quantitative mass spectrometry, and genetically-targetable fluorescent biosensors, that are offering new insights into the organization and regulation of cellular phosphorylation networks. Together, these approaches promise to lead to a systems-level view of dynamic phosphorylation networks.

**Keywords: protein phosphorylation networks, protein microarrays, fluorescent biosensors, phosphoproteomics, systems biology, kinase-substrate relationship, cell signaling and regulation, quantitative mass spectrometry**

## INTRODUCTION

Protein phosphorylation, mediated by protein kinases and opposed by protein phosphatases, is one of the most widespread regulatory mechanisms in eukaryotes. Inside the cell, protein kinases, phosphatases, and their respective substrates are organized into complex phosphorylation networks that govern nearly all cellular processes. In order to achieve signaling specificity in response to a given environmental stimulus, these networks must be precisely coordinated in cellular space and time. This involves regulation at several different levels, including at the levels of protein expression (e.g., enzyme levels), substrate selection, and spatiotemporal regulation of enzymatic activity. Not surprisingly, disruption of these regulatory mechanisms has been implicated in many pervasive diseases, including cancer (Guha et al., 2008; Deschenes-Simard et al., 2014), diabetes (Guo, 2014; Mackenzie and Elliott, 2014; Ortsater et al., 2014), and heart disease (Kooij et al., 2014; Sciarretta et al., 2014). As a consequence, much effort

has been dedicated to understanding how these parameters are controlled inside the cell.

In this review, we will examine several methodologies used to study the organization and regulation of cellular phosphorylation networks. Our discussion will be organized around two central questions: (1) which cellular proteins are phosphorylated by a given kinase (or dephosphorylated by a given phosphatase) under physiological conditions and (2) how are specific kinases and phosphatases—and ultimately the signaling networks of which they are a part—regulated within the native cellular environment? We will begin by first discussing the properties of protein phosphorylation that make this post-translational modification an attractive chemical signal for cellular information transfer. We will then highlight some of the traditional methods that researchers have used to study the phosphorylation status of cellular proteins before turning our attention to several emerging technologies that are beginning to provide a dynamic and

systems-level view of phosphorylation networks. Together, these approaches promise to shed new light on the organization and regulation of cellular phosphorylation networks.

### CELLULAR AND MOLECULAR CONSEQUENCES OF PHOSPHORYLATION

Protein phosphorylation describes the covalent attachment of a negatively charged phosphate group to one of several amino acid residues. Though several reports suggest that phosphorylation of His residues (resulting in the formation of an acid labile phosphoramidate species) may account for up to 10% of protein phosphorylation in some eukaryotes (Klump and Kriegstein, 2002; Besant and Attwood, 2010), it is generally believed that the predominant sites of phosphorylation in mammalian proteins are on Ser (~86%), Thr (~12%), and Tyr (~2%) residues (Hunter, 1996). Phosphorylation of these hydroxyamino acids results in the formation of phosphate monoesters that are surpassingly stable under physiological conditions, exhibiting an estimated half-life of  $1.1 \times 10^{12}$  years in aqueous solution (Lad et al., 2003). Despite their extremely high stability, phosphate monoesters are nonetheless rapidly hydrolyzed through the catalytic action of protein phosphatases, thus making protein phosphorylation a highly dynamic and “regulatable” post-translational modification inside cells. Importantly, the chemical properties of phosphate monoesters confer unique characteristics to phosphorylated amino acids that are critically important for their role in cellular signal transduction (Dissmeyer and Schnittger, 2011). For instance, a phosphate monoester contains two ionizable oxygens which exhibit  $pK_a$ 's of ~2.2 and 5.8, respectively, (Hunter, 2012). Therefore, depending on its chemical context, the transferred phosphate group is expected to be either dianionic or partially dianionic at physiological pH. Such a high charge density makes phosphate groups well-suited to mediate intra- or intermolecular interactions, such as hydrogen bonds and salt bridges, which can impact protein function.

In most cases, phosphorylation is believed to induce conformational changes in the target protein that influence its behavior inside the cell. For example, phosphorylation-dependent conformational changes can alter the activity of enzymes either directly, by promoting the reorganization of the enzyme active site, or indirectly, through allosteric effects on regulatory domains or subunits (Hunter, 2012). These modes of regulation are particularly important in the coordination of kinase signaling cascades that underlie cellular phosphorylation networks. For instance, recent evidence suggests that phosphorylation of residues within the activation loop, a highly conserved region present in all eukaryotic protein kinases, triggers global reorganization of two critical hydrophobic “spines” within the kinase core that help to properly position key residues in the active site necessary for catalysis (Taylor and Kornev, 2011; Meharena et al., 2013). Of course, phosphorylation-dependent changes in activity are not restricted to enzymes. Indeed, many ion channels and cellular transporters are known to be regulated by phosphorylation. For instance, phosphorylation of ryanodine receptor (RyR)  $Ca^{2+}$  channels by either  $Ca^{2+}$ /calmodulin-dependent protein kinase II (CaMKII) or cAMP-dependent protein kinase (PKA) induces conformational changes that alter their open probability at the

sarcoplasmic reticulum, influencing both the intensity and the frequency of  $Ca^{2+}$  spikes in cardiomyocytes and striated muscle cells (Meissner, 2010; Niggli et al., 2013). While CaMKII-mediated phosphorylation of S2815 is believed to induce local conformational changes that directly affect the flow of  $Ca^{2+}$  ions through the channel, structural changes caused by phosphorylation of S2809 and/or S2030 by PKA appears to promote the dissociation of the stabilizing protein, calbabin 2, from a site far removed from the phosphorylation sites. These changes lead to “leaky” SR channels associated with aberrant contractile function (Meissner, 2010; Niggli et al., 2013).

In addition to modulating its activity, the phosphorylation status of a protein may also impact other cellular parameters important to its function, such as its subcellular location or its stability. For instance, hyper-phosphorylation of the regulatory domain of the transcription factor, nuclear factor of activated T cells 1 (NFAT1), promotes electrostatic interactions between negatively-charged phosphate groups and positively-charged Lys/Arg residues found within a conserved nuclear localization sequence (NLS) (Okamura et al., 2000). These interactions induce conformational changes in the regulatory region that effectively mask NFAT1's NLS, leading to a predominantly cytoplasmic distribution prior to dephosphorylation by the  $Ca^{2+}$ /calmodulin-dependent protein phosphatase, calcineurin (CaN).

In the previous example, the phosphorylation status of NFAT1 determines whether its NLS is sterically blocked by the regulatory region or exposed to the cellular environment. This is important because the NLS is, itself, a site of interaction with importin family members involved in nucleocytoplasmic shuttling (Marfori et al., 2011). In fact, many phosphorylation-dependent conformational changes are believed to regulate protein–protein interactions. In some instances, such as for the NFAT1-importin and RyR-calbabin 2 interactions described above, the phosphate group(s) is not directly involved in mediating interactions with partner proteins. In contrast, many phosphorylation-dependent interactions rely upon direct recognition of phosphorylated residues by a diverse set of phosphoamino acid binding domains (PAABDs). For instance, many members of the SCF family of ubiquitin E3 ligases associate with their substrates via either Trp-Asp 40 (WD40) or leucine-rich repeat (LRR) PAABDs that specifically bind pSer/pThr residues located in phosphodegrons (Ho et al., 2006; Reinhardt and Yaffe, 2013). Phosphorylation-dependent interaction of SCF family members leads to ubiquitylation and subsequent degradation of the substrate by the 26S proteasome. In this way, phosphorylation serves as a key upstream signal which influences protein stability.

Over the past 20 years, many PAABDs, each exhibiting distinct substrate specificities and binding affinities, have been characterized. These include PAABDs that exclusively recognize either pTyr residues or pSer/pThr residues, as well as others that exhibit dual specificity, interacting with the phosphate monoester of all three hydroxyamino acids (Liao et al., 1999; Jin and Pawson, 2012; Reinhardt and Yaffe, 2013). While most pTyr binding domains fall within two large families: (1) the Src homology 2 (SH2) domain family (~120 distinct members) and (2) the phosphotyrosine binding (PTB) domain family (~27 distinct members) (Yaffe, 2002; Liu et al., 2006; Jin and Pawson, 2012), the

diversity of pSer/pThr PAABD's is much greater. Indeed, as a testament to the prevalence of Ser/Thr phosphorylation in eukaryotes, 14 distinct families of PAABDs that recognize pSer/pThr residues have been described to date (Jin and Pawson, 2012). Such a large number of distinct PAABD families suggests that phosphorylation-dependent interactions are a widely employed means of regulating protein–protein interactions inside cells.

Together, phosphorylation-dependent changes in protein function, be it through modulation of enzymatic activity, changes in sub-cellular localization, alterations in protein stability, or regulation of protein–protein interactions (or a combination thereof), plays a major role in regulating cellular physiology. Therefore, it is important to understand what factors influence protein phosphorylation inside cells.

### FACTORS INFLUENCING PROTEIN PHOSPHORYLATION INSIDE CELLS

Inside the cell, the phosphorylation status of a particular phosphorylation site (phosphosite) is determined by the equilibrium between kinase-mediated phosphorylation, in the one direction, and phosphatase-mediated dephosphorylation in the other direction<sup>1</sup>. In order for a kinase to act on a cellular protein, several criteria must be met. For instance, a kinase and its substrate(s) must be expressed in the same cell and at the same time during development. This often involves several levels of control, including transcriptional regulation (e.g., control of gene expression), post-transcriptional regulation (e.g., regulation of mRNA stability) and modulation of protein levels (e.g., regulation of protein stability). Diverse and complementary experimental approaches have been developed to monitor these parameters, including reporter gene assays to track changes in transcriptional activation at a given promoter (Roura et al., 2013; van Rossum et al., 2013; Khan et al., 2014), quantitative real-time PCR (qPCR) and expression microarrays to measure the relative levels of mRNA transcripts (Skrzypski, 2008; Gorreta et al., 2012), and western blot analysis and fluorescent imaging techniques to measure changes in protein abundance over time (Wiechert et al., 2007; Chao et al., 2012). Though a comprehensive discussion of these techniques is beyond the scope of this review, the interested reader is referred to several recent reviews (Wilkins, 2009; van Rossum et al., 2013).

In addition to being co-expressed, a protein kinase and its substrate must physically interact with one another in order for phosphorylation to occur. Therefore, information about protein localization and protein–protein interactions is critical in determining whether a given kinase-substrate relationship (KSR) is likely to occur under physiological conditions. Traditionally, researchers have employed biochemical approaches to obtain this information. For instance, subcellular fractionation, generally achieved via differential centrifugation through a density gradient, has long been used to separate cellular organelles. When combined with western blot analysis, cellular fractionation is an effective means of determining the subcellular localization of a particular

protein species. More recently, fluorescence imaging techniques, such as immunofluorescence and live cell imaging, have offered insights into the subcellular localization of proteins without disrupting the cellular architecture. Importantly, in the case of live cell imaging, dynamic changes in the subcellular distribution of a fluorescently tagged protein can be visualized in real-time within single cells. The advent of real-time super-resolution microscopy techniques based on photoswitchable fluorescent proteins (FPs) has dramatically increased the resolution afforded by live cell fluorescence imaging experiments (reaching ~20 nm resolution in living cells compared to >200 nm using standard fluorescence imaging) (Dedecker et al., 2012; Agrawal et al., 2013; Persson et al., 2013). While, in theory, this resolution permits the visualization of several structures important for cellular signaling, such as membrane microdomains (Honigsmann et al., 2013) and large protein complexes (Biggs et al., 2011), it is still not sufficient to detect most protein–protein interactions directly. Instead, this information can be attained in several ways, as briefly summarized below.

Perhaps the most popular method for assaying protein–protein interactions is co-immunoprecipitation (co-IP). The immunoprecipitation step can be achieved using antibodies raised against either a specific protein-of-interest or an epitope tag, such as the hemagglutinin (HA)- or tandem affinity purification (TAP)-tags (Puig et al., 2001). Therefore, co-IPs can be used to detect interactions that occur in different cell types, ranging from naïve primary cells to transfected and/or genetically engineered cells. When coupled with quantitative tandem mass spectrometric (MS/MS) analysis (see Section “Top-Down” Methodologies: *In situ* Identification of Phosphosites), co-IP assays can yield a wealth of information about protein–protein interactions that occur under various cellular conditions. However, because many protein interactions occur in the context of multimeric protein complexes, in the absence of cross-linking agents, co-IPs do not provide definitive information about direct protein–protein interactions. Moreover, because the cells must be lysed prior to immunoprecipitation, weak and transient interactions are often missed using this approach.

In contrast, yeast two-hybrid analysis, which generally detects binary interactions, is able to detect both weak and strong interactions alike. However, it is important to note that, because both the bait and the prey proteins must be localized to the nucleus using this approach, interactions often occur within a cellular context much different from the one normally encountered by the proteins under study. Moreover, if the interaction is dependent upon post-translational modifications, such as phosphorylation or acetylation, the modification may not actually occur in yeast. In fact, this may be the case even if an ortholog of the appropriate modifying enzyme (e.g., the orthologous kinase or acetyl transferase) is present in yeast. Indeed, we and others have recently reported that many KSRs are not conserved between yeast and humans, despite the presence of orthologous kinase-substrate pairs (Mok et al., 2011; Hu et al., 2014).

Therefore, to directly visualize protein–protein interactions within the native cellular environment, several fluorescence imaging techniques have been developed (Ciruela, 2008; Shekhawat and Ghosh, 2011; Stynen et al., 2012). These include approaches

<sup>1</sup>For simplicity, unless otherwise noted, we will focus primarily on kinase-mediated phosphorylation of cellular substrates for the remainder of this discussion; however, it is important to note that the same principles apply for phosphatase-mediated dephosphorylation of cellular proteins.



based on fluorescence resonance energy transfer (FRET) between fluorescently tagged proteins (Padilla-Parra and Tramier, 2012; Zadran et al., 2012; Sun et al., 2013) as well as protein complementation assays (PCA), which rely on the interaction-dependent reassembly of N- and C-terminal fragments of FP color variants (Ciruela, 2008; Shekhawat and Ghosh, 2011). Not only do these techniques allow protein–protein interactions to be observed in many subcellular regions, but they also can be conducted in cell types (e.g., mammalian cells) that place the interacting proteins in the context of their endogenous regulatory networks.

Together, gene expression, subcellular localization and protein–protein interaction data can be used to construct extensive interaction networks that offer global information about the interactome under different cellular conditions (Pastrello et al., 2014). This information can be very useful in predicting KSRs that are likely to occur inside cells. For instance, protein–protein interactions, be they direct or indirect (e.g., those mediated by scaffold proteins), appear to be one of the strongest predictors of physiologically relevant KSRs (Newman et al., 2013). This is likely due to the key role that protein–protein interactions play in substrate selection, which is discussed below.

### SUBSTRATE SELECTION: IDENTIFYING THE CELLULAR TARGETS OF PROTEIN KINASES AND PHOSPHATASES

It is currently believed that ~40% of the proteins in the human proteome are phosphorylated at some point during their lifetime. By extension, since a given protein often contains multiple phosphorylation sites, the total number of phosphosites in the human proteome has been estimated to be ~100,000 sites (Zhang et al., 2002; Dephoure et al., 2013). Therefore, simply cataloging the complete complement of phosphosites in the human proteome, irrespective of their dynamic regulation or their functional consequences, is a seemingly daunting task. This task is made even more challenging by the fact that, in many cases, the phosphorylated form of a protein represents only a very small fraction of the total copies of that protein species inside the cell. That being said, while phosphosite identification is extremely important, simply knowing which sites are phosphorylated inside the cell only tells half of the story. Indeed, if we are to truly understand how dynamic phosphorylation networks are organized and regulated inside cells, it is important to know which kinases are actually mediating the phosphorylation event.

Therefore, to construct comprehensive maps of cellular phosphorylation networks, researchers have developed a series of experimental approaches that can be broadly grouped into two complementary categories, which we will refer to as “top-down” and “bottom-up” methodologies<sup>2</sup>. While “top-down” approaches begin at the cellular level and work their way down to the protein level, “bottom-up” approaches begin at the protein level and work their way up to the cellular level. Below, we describe each category and highlight how they can be used together to gain global insights into cellular phosphorylation networks.

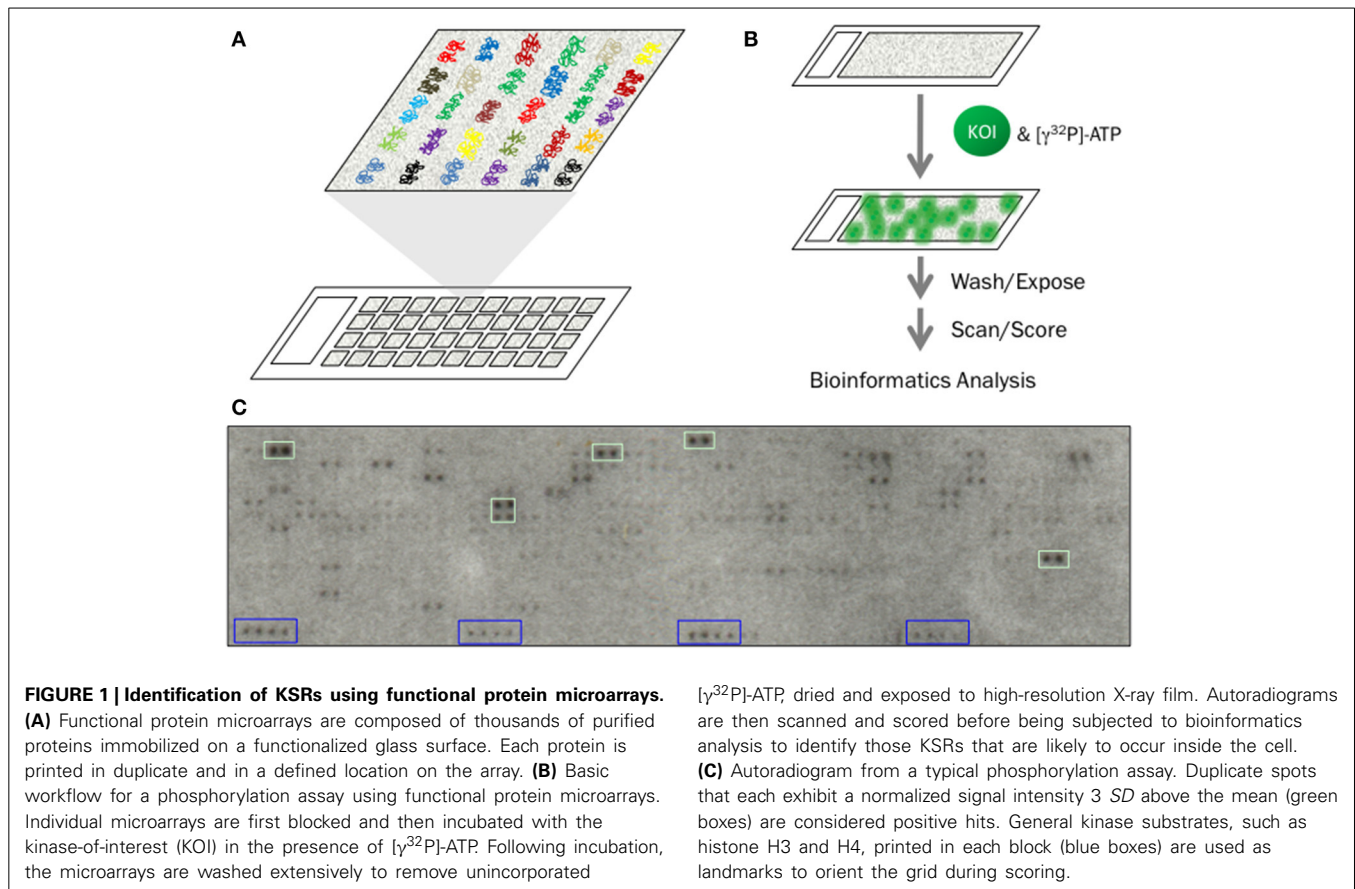
### “BOTTOM-UP” METHODOLOGIES: *IN VITRO* ANALYSIS OF KINASE-SUBSTRATE RELATIONSHIPS

In “bottom-up” approaches, biochemical analysis of individual kinase-substrate pairs is carried out *in vitro* using purified protein components. Traditionally, phosphorylation is detected using either autoradiography or scintillation counting following incubation of the kinase-of-interest with a single substrate and radiolabeled ATP (e.g., [ $\gamma$ -<sup>32</sup>P]-ATP or [ $\gamma$ -<sup>33</sup>P]-ATP). In order to identify the site(s) of phosphorylation on the substrate, the initial screen is often followed by mutational analysis of putative phosphorylation sites. Due to safety concerns associated with radioactive assays and the high cost of radioisotope disposal, several non-radioactive assays have been developed based on alternative detection methods (Glickman, 2012). In many cases, these methods employ coupled enzymatic reactions that measure activity-dependent depletion of ATP. On the other hand, the development of generalizable phospho-detection reagents, such as highly sensitive anti-pTyr antibodies or biotinylated Phos-Tag phosphochelators (see discussion in Section “Top-Down” Methodologies: *In situ* Identification of Phosphosites below), allows direct visualization of phosphorylation on protein substrates using fluorescent or chemiluminescent detection methods (Kinoshita et al., 2004, 2012a).

Regardless of the detection method employed, detailed biochemical analysis can offer a wealth of information about individual KSRs, including the kinetics of the phosphorylation reaction, the site(s) of phosphorylation and, in some cases, the functional consequences of phosphate addition. Each of these parameters is important for understanding how phosphorylation-dependent signaling is achieved inside cells. However, due to the time that it takes to fully characterize a single KSR, traditional biochemical approaches are not well suited to high-throughput analysis. Therefore, to increase the rate of KSR discovery, researchers have developed phosphorylation assays based on functional protein microarrays (Ptacek et al., 2005; Mok et al., 2009; Popescu et al., 2009; Newman et al., 2013).

Functional protein microarrays are composed of thousands of individually purified recombinant proteins immobilized in discrete spatial locations on a functionalized glass surface (Hu et al., 2011; Sutandy et al., 2013). While several surface chemistries have been used to immobilize the purified proteins [e.g., polyvinylidene fluoride (PVDF), nitrocellulose, streptavidin and acrylamide] (Hu et al., 2011), bi-functional cross-linking agents, such as functionalized aminosilane and carboxylic esters, are generally preferable for phosphorylation assays due to their relatively low background signal (Mok et al., 2009). On the microarray, each individual protein is printed multiple times (typically in duplicate or triplicate) to ensure “on-chip” reproducibility and to guard against signal artifacts that can arise due to incomplete washing or experimental variation (Figure 1A). Because each spot on the array typically contains only a few femtograms of protein, sensitive detection methods are necessary to visualize phosphorylation. Moreover, since each protein occupies a defined position on the array and many proteins can potentially be phosphorylated in the same experiment, direct visualization of phosphorylation using either radioactive ATP or specific phospho-detection reagents (e.g., pTyr antibodies) is essential.

<sup>2</sup>The “top-down” and “bottom-up” approaches described here should not be confused with similar terms used to describe MS-based proteomic analysis of intact proteins (“top-down”) or peptide fragments (“bottom-up”).



During a standard phosphorylation assay, each protein microarray is incubated with an active form of the kinase-of-interest in the presence of radioactive ATP (**Figure 1B**). Following incubation, the microarray is washed several times and dried by centrifugation before finally being exposed to high-resolution X-ray film for several hours or several days (**Figure 1C**). The resulting autoradiogram is then converted to a digital image and scored using image analysis software such as GenePix 6.0 (Axon, Inc.). General kinase substrates (e.g., histone H3 and H4) and/or kinases that undergo autophosphorylation are used as “landmarks” to align grids during scoring. In addition to the experimental arrays, it is important that each batch of experiments includes a negative control in order to identify those proteins that undergo autophosphorylation. Along with the landmarks, these species can also be useful for orienting the scoring grid during alignment. Using this protocol, a team of two researchers in our lab is able to conduct up to 144 phosphorylation assays per day, with each assay probing thousands of potential substrates. In this way, a large number of KSRs can be identified in a relatively short period of time. For instance, using high-density protein microarrays composed of 2158 unique proteins from *Arabidopsis thaliana*, Popescu and colleagues identified 570 substrates for 10 mitogen-activated protein kinase (MAPK) family members (Popescu et al., 2009). Likewise, we were able to identify over 24,000 *in vitro* substrates for 289 unique human kinases (Newman et al., 2013). While the latter studies used microarrays containing 4191 unique

human proteins (Hu et al., 2009), the recent development of a high-density human protein microarray composed of 16,368 unique human proteins (representing over 80% of the human proteome) creates exciting new opportunities for KSR discovery at a global scale (Jeong et al., 2012).

Due to the large number of phosphorylated species generated during a typical microarray phosphorylation assay, sophisticated statistical analyses are often required to identify those substrates that constitute positive “hits.” First, the signal intensity of each protein spot is measured and subjected to some form of background correction, be it background subtraction or ratiometric analysis (e.g., calculation of the foreground/background ratio) (Ptacek et al., 2005; Popescu et al., 2009; Newman et al., 2013). To eliminate spatial artifacts that can arise due to uneven washing or drying of the microarray, local background correction is often preferable to global correction (Hu et al., 2009). If the normalized signal intensity for each of the replicate spots is above a predefined threshold [e.g., >3 standard deviations (SD) above the mean intensity on the microarray], then the protein is considered a putative substrate. However, it is important to note that, because both the kinase and the substrate are removed from the cellular environment, a positive result *in vitro* does not guarantee that the KSR actually occurs under physiological conditions. Bioinformatics analysis can help curb this limitation by identifying “high confidence” KSRs that are likely to occur inside the cell. Here, accurate gene ontology (GO) data, such as

that obtained using the approaches outlined in Section Factors Influencing Protein Phosphorylation Inside Cells, is a valuable resource for determining whether a given cellular protein is likely to be phosphorylated by its cognate kinase *in vivo*.

Though necessary to eliminate false positives, the use of a stringent cutoff during KSR identification likely results in a relatively high false negative rate, as well (estimated to be ~95%, in some cases) (Newman et al., 2013). For instance, because the number of phosphosites for a given kinase varies from protein to protein, a substrate that contains only one phosphorylation site for a given kinase will exhibit a lower normalized signal intensity on the array than a protein that contains multiple sites for the same kinase. If the signal intensity of the first substrate falls below the threshold determined for the chip, then that substrate would not be considered a hit even though it is a *bona fide* substrate of the kinase (which likely would have been classified as such had the experiment been conducted using the purified kinase and substrate in isolation). Indeed, we have observed that known KSRs are sometimes absent from the final hit list. Visual inspection often reveals that this is not because the substrate is not actually phosphorylated on the microarray; rather it is because its normalized signal intensity does not exceed the stringent cut-off required to be considered a positive hit. Likewise, variations in protein abundance on the microarray or impairment of kinase-substrate interactions due to misfolding or truncation of the substrates during purification can bias KSR identification. Importantly, KSRs that are dependent on auxiliary factors, such as scaffolding proteins, co-factors, or post-translational modifications, may be missed all together using functional protein microarrays and other “bottom-up” approaches.

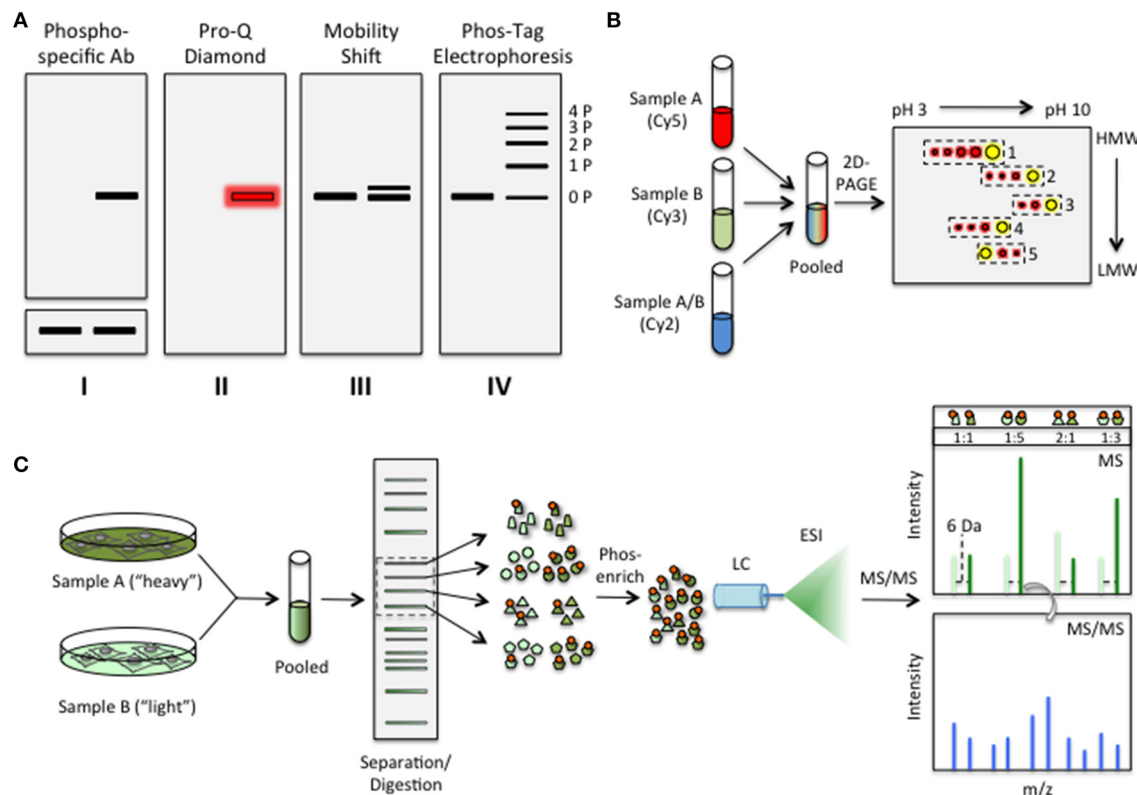
Therefore, as a complementary approach to the phosphorylation assays described above, protein microarray assays have been developed using cell lysates. For instance, Woodard et al. recently developed an assay platform that uses functional protein microarrays to measure phosphorylation profiles in whole-cell lysates obtained from cell or tissue samples (Woodard et al., 2013). This platform, which is similar to the phosphorylation assays described above except whole cell lysates are substituted for the purified kinase, revealed changes in the phosphorylation profiles of immortalized U373 cancer cells and glioblastoma tumors upon hepatocyte growth factor (HGF)-mediated activation of the c-Met signaling pathway. One of the primary advantages of this platform is that it more closely mirrors the cellular environment under which KSRs occur. For instance, because lysate assays preserve many of the elements of intact signaling networks, such as the endogenous levels of proteins, co-factors, second messengers, and inhibitors, while also maintaining protein complexes involved in signal transduction, they can offer insights into global changes in phosphorylation networks under different cellular conditions. Indeed, in addition to proteins known to be associated with the c-Met pathway, these analyses identified over 400 proteins that are likely influenced by HGF/c-Met signaling (Woodard et al., 2013). Moreover, the profiles obtained from these assays may be useful for biomarker discovery. Therefore, this approach has the potential to be a powerful diagnostic tool for rapid, cost-effective screening.

Similarly, reverse phase protein microarrays (RPPA) represent powerful tools for tracking changes in phosphorylation profiles across samples. Unlike the functional protein microarrays described above, RPPA fabrication entails the immobilization of cell lysates or tissue extracts on the functionalized glass surface. The arrays are then probed with an antibody specific for the epitope-of-interest (e.g., a particular phosphosite) to assess changes in its expression levels or phosphorylation status under different conditions. Because each spot contains only a few nanoliters of lysate, many arrays (each containing lysates derived from tens to thousands of unique samples) can be prepared using only a small amount of starting material (Pierobon et al., 2014). The microarrays can then be probed simultaneously, allowing a large number of analytes to be tested in parallel while reducing the effects of inter-assay variability. As a consequence, RPPA-based assays have been widely applied to the study of phosphorylation profiles in clinical samples. For instance, Ummanni et al. recently employed RPPA assays to probe the expression levels and phosphorylation status of 71 cancer-associated proteins in 84 non-small cell lung cancer cell lines (Ummanni et al., 2014). Interestingly, they found that the sensitivity toward the EGFR inhibitors, lapatinib and erlotinib, correlated more closely with EGFR/ERBB2 phosphorylation than with receptor expression levels. This was not the case for sensitivity toward the SRC/BCR-ABL inhibitor, dasatinib, which instead correlated with expression of proteins downstream of EGFR/ERBB2. Of course, a significant limitation of the RPPA approach is its dependence on the specificity and avidity of commercially produced antibodies. As the toolkit of reliable phosphospecific antibodies grows, so too will the utility of this approach.

It is important to note that, because many kinases are present in lysates simultaneously, the identity of the kinase actually mediating the phosphorylation of a given substrate cannot be definitively established using lysate-based approaches. Therefore, lysate-based assays should be considered orthogonal to assays using purified kinases. The same is true for other “top-down” approaches, as described below.

#### “TOP-DOWN” METHODOLOGIES: *IN SITU* IDENTIFICATION OF PHOSPHOSITES

“Top-down” approaches to substrate identification are designed to measure changes in the phosphorylation status of cellular proteins upon activation/inhibition of endogenous signaling pathways. Traditionally, this has been achieved using polyacrylamide gel electrophoresis (PAGE)-based methods (Figures 2A,B). Following cell lysis, the protein-of-interest is either immunoprecipitated or the whole cell lysate is loaded directly onto the gel. Separation is generally achieved via either one- or two-dimensional PAGE. In the case of one-dimensional PAGE, electrophoresis through a denaturing gel is often followed by western blot analysis to visualize the phosphorylation status and/or the migration pattern of the protein-of-interest (Figure 2A). If a phosphospecific antibody exists for the phosphosite under study, changes in phosphorylation levels at that site can be assessed directly by comparing the intensities of the bands before and after treatment (Figure 2AI). However, because phosphorylation can impact protein stability inside the cell,



**FIGURE 2 | “Top-down” approaches to assess changes in the phosphorylation status of cellular proteins.**

**(A)** Approaches based on 1D PAGE. **I** Western blot analysis using an antibody that specifically recognizes a particular phosphosite. To account for phosphorylation-dependent changes in protein stability, the signal must be normalized to that obtained using an antibody that recognizes the unphosphorylated form of the protein-of-interest (below); **II** Detection using a general phosphorylation detection reagent, such as Pro-Q Diamond or the Phos-Tag phospho-chelator; **III** Phosphorylation of some proteins can be assessed based on changes in their electrophoretic mobility through a standard SDS-PAGE gel; **IV** Phos-Tag electrophoresis allows detection of a wide range of phosphorylation events, often resulting in a ladder corresponding to multiply phosphorylated species (1P, 2P, 3P, etc.). **(B)** Approaches based on 2D-PAGE. The general workflow of a 2D-DIGE experiment is shown. Accordingly, lysates from treated (Sample A) or untreated (Sample B) cells are first labeled with size- and charge-matched fluorescent dyes, such as Cy3 and Cy5, before being pooled together. A third sample, composed of both lysates labeled with a third dye (e.g., Cy2), may also be included as an internal reference. The pooled samples are then resolved on a 2D-PAGE gel, which generally uses IEF in the first dimension to separate cellular proteins based on their pI's followed by SDS-PAGE in the second dimension to separate the proteins based on size (HMW, high MW; LMW, low MW). Composite spots [e.g., Cy3 (green) and Cy5 (red)] are shown in yellow.

Meanwhile, cellular proteins that are uniquely phosphorylated in the treated sample exhibit a spot train moving from right to left (e.g., boxes 1–4) while those proteins that are dephosphorylated under the experimental conditions are characterized by a spot train that moves in the opposite direction (e.g., box 5). The number of spots in the train corresponds to the number of phosphosites that are occupied in the protein (i.e., four spots represents four phosphorylation events). The intensity of each spot in the train can be used to gauge the relative levels of each phosphorylation state under each condition. **(C)** Approaches based on MS/MS. The basic workflow for a SILAC experiment is shown. To metabolically label proteins, cells are grown in the presence of either a “heavy” isotope of a particular amino acid (e.g.,  $^{13}\text{C}$ -Arg) or its “light” counterpart (e.g.,  $^{12}\text{C}$ -Arg). Cells are then pooled and lysed. Cellular proteins are resolved by 1D SDS-PAGE before being digested in the gel by an Arg/Lys-directed protease, such as trypsin. Peptide fragments in individual gel slices (represented by either parallelograms, circles, triangles, or pentagons) are then electro-eluted and phosphorylated species are enriched using one of several phospho-enrichment strategies outlined in the text. Phosphopeptides are then resolved by reverse-phase liquid chromatography (LC) and ionized by electrospray ionization (ESI) before being analyzed by in-line MS/MS. In the MS spectrum, fragments containing heavy isotopes are off-set by a known amount (e.g., 6 Da for  $^{13}\text{C}$ -Arg), allowing quantitation based on the relative intensity of each peak. The identity of individual peaks is determined based on the MS/MS spectrum.

the signal intensities obtained using phosphospecific antibodies must be normalized to the amount of target protein in each lane (e.g., by using an antibody raised against the unphosphorylated form of the protein). If a phosphospecific antibody does not exist (as is often the case) or the specific site of phosphorylation is not known, the phosphorylation status of proteins may still be assayed using general phospho-detection reagents, such as Pro-Q Diamond (Life Technologies, Inc.) or the Phos-Tag

phosphochelator, or based on changes in the electrophoretic mobility of the protein (**Figures 2AII,III**). In the case of the latter, it is believed that the high charge density of the phosphate monoester prevents the uniform binding of negatively charged SDS molecules in the vicinity of the phosphate group, causing some phosphorylated proteins to migrate more slowly than their unphosphorylated counterparts (Gaffken and Lampe, 2006). For instance, using a library of epitope-tagged proteins, the Matsuda



group profiled changes in the electrophoretic mobility of each of the proteins in the *S. pombe* proteome (Shirai et al., 2008). While this study did not focus specifically on phosphorylation, the authors noted that nearly 42% of the proteins in the *S. pombe* proteome migrated at a molecular weight that was significantly different from their calculated molecular weight (with ~28% of the proteins migrating at a higher-than-expected MW).

However, it is important to note that not all proteins exhibit an observable mobility shift upon phosphorylation. Therefore, to accentuate phosphorylation-induced changes in mobility in an unbiased manner, Kinoshita and colleagues recently developed several methodologies based on so-called Phos-Tag technologies (Kinoshita et al., 2004, 2012a,b; Tsunehiro et al., 2013). Phos-Tags, which are alkoxide-bridged dinuclear metal complexes, have been shown to selectively bind phosphate monoesters in a sequence-independent manner (Kinoshita et al., 2004). When conjugated to a gel matrix such as acrylamide, the Phos-Tag effectively slows the migration of any species containing a phosphate monoester, leading to a discernable shift in its mobility (Kinoshita et al., 2012b) (**Figure 2AIV**). Consistent with the notion that many proteins are multiply phosphorylated inside the cell, a ladder of clearly distinguishable bands is often observed using Phos-Tag electrophoresis. Though several studies have used Phos-Tag technologies either to profile the phosphorylation status of select substrates (Kinoshita et al., 2009; Aguilar et al., 2011; Kinoshita-Kikuta et al., 2012) or to validate substrates identified by other methods (Mukai et al., 2008; Deswal et al., 2010; Mok et al., 2010; Huang et al., 2014; Yip et al., 2014), to date Phos-Tag electrophoresis has not been applied to large-scale proteomic analyses. It will be interesting to compare the results of such studies with those obtained using standard 1D SDS-PAGE—such as the one conducted by Matsuda et al. in *S. pombe* (Shirai et al., 2008)—as well as with those that utilize 2D-PAGE, as described below.

In addition to one-dimensional PAGE, 2D-PAGE is can also be used to assess the phosphorylation status of cellular proteins. In a 2D-PAGE experiment, proteins are typically separated via isoelectric focusing (IEF) in the first dimension followed by SDS-PAGE in the second dimension (though other separation strategies have also been successfully combined) (Camacho-Carvajal et al., 2004; Kinoshita et al., 2009) (**Figure 2B**). Because the negative charge of the phosphate monoester causes a reduction in the pI of the protein, the migration of the phosphorylated species is altered during IEF. Separation in the second dimension therefore leads to clearly discernable spot trains along the vertical or horizontal axes, corresponding to differentially phosphorylated species. When coupled with western blot analysis, 2D-PAGE can offer insights into either the phosphorylation status of a specific cellular protein or global changes in phosphorylation across the entire proteome (e.g., changes in Tyr phosphorylation). Alternatively, proteins can first be stained with a contrast agent, such as Coomassie dye or silver stain, then excised from the gel, and sequenced using either Edman sequencing or quantitative MS/MS.

One of the greatest strengths of 2D-PAGE is that it simultaneously offers insights into the relative expression levels and post-translational modifications of intact cellular proteins. At the same time, this also highlights one of its limitations with

respect to phosphosite identification. Indeed, because phosphorylated proteins are often present at relatively low levels inside cells, many phosphoproteins can be missed using this approach. In some cases, this limitation can be overcome through the use of sensitive phosphodetection reagents, such as radioactive ATP or phosphospecific antibodies, or through phosphoprotein enrichment procedures, such as immobilized metal affinity chromatography (IMAC) using trivalent metal ions such as  $\text{Fe}^{3+}$  and  $\text{Ga}^{3+}$  (Posewitz and Tempst, 1999; Kosako and Nagano, 2011). Likewise, subcellular fractionation can be used to examine the phosphorylation profiles of proteins within specific subcellular compartments (Stasyk et al., 2007).

To gain insights into differential phosphorylation of cellular proteins using 2D-PAGE, two sets of lysates must be compared (e.g., cell extracts from healthy vs. diseased cells). However, due to inherent technical obstacles stemming from gel-to-gel variability and differential staining of proteins, cross-gel comparisons are often difficult using this approach. To address these limitations, researchers have developed 2D-differential gel electrophoresis (2D-DIGE) methodologies (Minden, 2012) (**Figure 2B**). In a 2D-DIGE experiment, spectrally distinct, charge-matched fluorescent dyes, such as Cy3 and Cy5, are used to label the proteins in individual lysates. The two lysates, each labeled with a distinct dye, are then mixed and examined by 2D-PAGE. Because 2D-DIGE permits the analysis of multiple lysates on a single gel, direct comparison between different experimental conditions is relatively straightforward. Moreover, due to their high intrinsic brightness, the fluorescent dyes used for labeling are well-suited for the detection of low abundance protein species (Mujumdar et al., 1993; Kosako and Nagano, 2011). Importantly, quantitation of the fluorescence intensity of each spot also allows relative levels of phosphorylation to be measured directly. Using this approach, Pellegrin et al. conducted a proteome-wide analysis of primary erythrocytes undergoing apoptosis following erythropoietin withdrawal (Pellegrin et al., 2012). These studies revealed changes in both protein abundance and the phosphorylation profiles of many important cellular proteins, including several heat shock protein 90 (Hsp90) isoforms.

Despite the improved sensitivity and analytical power afforded by 2D-DIGE, methods based on 2D-PAGE are still only able to identify changes in a few hundred proteins at a time (Choudhary and Mann, 2010). This makes it difficult to gain a truly systems-level view of phosphorylation networks, where thousands of different cellular proteins may be phosphorylated at any given time. In this respect, the development of shotgun proteomics approaches based on MS/MS has been truly revolutionary. Indeed, studies employing shotgun MS/MS have recently led to an explosion in the number of annotated cellular phosphorylation sites, opening new avenues of research in systems biology.

Like the other “top-down” approaches discussed above, MS-based detection of phosphosites requires cell lysis prior to analysis. Following lysis, proteins are generally resolved by a separation technique suitable to the sample volume to be analyzed (e.g., SDS-PAGE or column chromatography) and digested using an Arg/Lys-directed protease, such as trypsin, chymotrypsin or lysyl endopeptidase (**Figure 2C**). Phosphopeptides are then enriched using one of several phosphoenrichment procedures discussed



below. The enriched phosphopeptides are separated by reverse-phase liquid chromatography (LC) before being ionized by electrospray ionization as they elute from the column. Finally, the ionized peptides are loaded onto an in-line mass analyzer for analysis. In the case of MS/MS, two mass analyzers, for example an ion trap instrument and an orbitrap instrument, are often used in tandem to increase the resolution of detection<sup>3</sup>. In this scheme, the first mass analyzer is used to obtain a mass spectrum of all peptides eluting from the column at a given time. Once the first mass spectrum is obtained, peptides with a specific mass-to-charge ratio ( $m/z$ ) are isolated, further fragmented through high-energy interactions, and analyzed by the second mass analyzer (modern mass spectrometers can isolate individual peptides in a matter of milliseconds) (Choudhary and Mann, 2010). The tandem mass spectra obtained from these experiments are then compared against protein sequence databases or *in silico* peptide fragmentation spectra in order to match the peptide fragments to a corresponding cellular protein (Gafken and Lampe, 2006). A 79.97 Da increase in the mass spectrum of a peptide fragment is indicative of phosphorylation; however, it is important to note that, if multiple Ser/Thr/Tyr residues are present in close proximity to one another in a peptide, it is often difficult to unambiguously identify the site of phosphorylation.

Though MS analysis has been used for decades to analyze relatively simple mixtures consisting of a few hundred proteins, recent advances have now made possible the identification of thousands of phosphopeptides in a single study. As a consequence, MS-based studies have contributed to an unprecedented increase in the number of known phosphoproteins and phosphosites, which currently stands at just under 118,500 unique phosphosites in ~16,400 non-redundant proteins across human, mouse, and several other species [PhosphositePlus database (as of 4/13/2014)]. As we discuss below, this veritable explosion in phosphosite data has been made possible by several key technological advances in MS-based analysis of the phosphoproteome, including (1) the development of phosphopeptide enrichment methods, (2) improvements in the resolution and accuracy of mass analyzers, (3) the emergence of alternative fragmentation methods, and (4) the development of computational algorithms that allow high-confidence peptide identification and phosphosite localization. Moreover, the introduction of quantitation strategies based on isotopic labeling allows researchers to track dynamic changes in global phosphorylation profiles in response to a particular cellular stimulus. Due to the existence of a number of excellent reviews on these topics, (Choudhary and Mann, 2010; Kosako and Nagano, 2011; Johnson and White, 2012; Kinoshita-Kikuta et al., 2012; Nilsson, 2012; Ong, 2012; Rigbolt and Blagoev, 2012) here we will only highlight these innovations and briefly discuss how they have been used to gain a systems-level view of phosphorylation networks.

As alluded to above, the occupancy of many phosphosites is sub-stoichiometric. As a consequence, at any given time, only a small fraction of the total copies of a protein species are expected

to be phosphorylated at a particular site. When coupled with the fact that many phosphoproteins involved in signal transduction are already expressed at relatively low levels, detection of phosphorylated species represents a major challenge to phosphosite identification. To overcome these challenges, researchers have recently developed a variety of phosphoenrichment techniques designed to selectively bind phosphorylated peptides. In addition to the IMAC- (Andersson and Porath, 1986; Posewitz and Tempst, 1999) and Phos-Tag-based (Nabetani et al., 2009) approaches discussed above, phosphoenrichment strategies using chromatographic separation by strong cation exchange (SCX), hydrophilic interactions, or immobilized metal oxides, such as titanium dioxide (TiO<sub>2</sub>) and zirconium dioxide (ZrO<sub>2</sub>), have also been successfully combined with MS/MS for phosphosite identification (Ruprecht and Lemeer, 2014; Yang et al., 2014). Likewise, immunopurification using phosphospecific antibodies, such as anti-pTyr antibodies or antibodies raised against the phosphorylated form of the consensus phosphorylation motif of a kinase-of-interest (Zhang et al., 2002), provide information about phosphorylation on a particular type of residue (e.g., Tyr) or in a particular sequence context (e.g., a kinase consensus motif), respectively. Due to their high degree of enrichment and ease of use, IMAC- and metal oxide-mediated enrichment strategies are currently the most popular methods for phosphoproteomic analysis. However, these methods are also prone to non-specific interactions with acidic peptides containing a disproportionate number of Asp and Glu residues. Therefore, it is sometimes necessary to combine one or more enrichment methods prior to analysis by LC-MS/MS.

In addition to the development of phosphoenrichment methods, improvements in the resolution and mass accuracy of modern mass analyzers have greatly improved our ability to detect phosphosites in complex mixtures. For instance, whereas standard ion trap mass spectrometers can typically resolve a few hundred to a thousand individual peptides as they emerge from the reverse phase column, modern time-of-flight (TOF) and orbitrap instruments exhibit a mass resolution of 20,000 and 100,000, respectively. Moreover, these instruments exhibit mass accuracies in the low parts-per-million, dramatically improving the coverage that can be achieved in a single experiment. Together, the increased resolution and mass accuracy afforded by modern mass spectrometers allows much more complex mixtures to be analyzed at once, markedly improving both the throughput of a given experiment and the confidence in the resulting peptide assignments.

One of the primary challenges in phosphosite detection is preservation of phosphorylation during peptide fragmentation. This is because collision-induced dissociation (CID) methods, which accelerate the ionized peptides in a collision cell and then bombard them with inert gases such as helium, nitrogen, or argon, tend to cause the cleavage of the phosphate monoester linkage on hydroxyamino acid residues (Kosako and Nagano, 2011). The development of “phosphate-friendly” fragmentation methods, such as higher energy collisional dissociation (HCD) (Olsen et al., 2007), electron capture dissociation (ECD) (Zubarev et al., 2000), and electron transfer dissociation (ETD) (Syka et al., 2004), permits efficient fragmentation

<sup>3</sup>It should be noted that MS/MS can also be achieved “in time” using a single mass analyzer with multiple separation steps being instituted at different times.

without cleavage of the phosphoester bond. As a consequence, more phosphosites will be preserved and available for detection. Importantly, because the fragmentation mechanisms used by ETD and CID are complementary, these methods can be used together to improve identification of phosphopeptides (Na and Paek, 2014).

Once mass spectra are generated, peptide sequences are typically identified by comparing the spectra with databases specific for the species under study. A number of currently available database search algorithms, such as MASCOT, SEQUEST, X!Tandem, and the Open Mass Spectrometry Search Algorithm (OMSSA), (Eng et al., 1994; Perkins et al., 1999; Craig and Beavis, 2003; Geer et al., 2004) are amenable to the identification of phosphorylated peptides. In general, these algorithms compare the experimentally determined mass spectra to theoretical fragmentation spectra generated *in silico*. Various scoring metrics are then used to determine the likelihood that a given match represents a *bona fide* phosphorylation event.

In addition to database searches, *de novo* sequencing approaches have also been developed. Unlike database searches, which are restricted by available sequence data, modern *de novo* approaches, such as DeNovoPTM, (He et al., 2013) infer the amino acid sequence (and any modifications thereto) directly from the MS/MS spectrum. Consequently, they are able to detect modifications in mutated proteins that may otherwise be missed. However, because *de novo* approaches are highly dependent on the quality of the spectral data, they are often prone to errors caused by incomplete or random fragmentation patterns (Na and Paek, 2014). To overcome these limitations, hybrid methods have been developed that conduct the database search using short sequence tags composed of two to four amino acids (determined by *de novo* sequencing). The use of sequence tags, which are less sensitive to sequencing errors than complete *de novo* sequencing, can dramatically reduce the database search time by reducing the search space to only a few candidate peptides that contain the sequence tag. By comparing the experimentally determined mass spectra with the theoretical spectra associated with the sequence tag, hybrid approaches have the potential to rapidly identify phosphorylation events on peptides (as well as many other modifications).

However, as alluded to above, in many cases a phosphorylated peptide may contain multiple phosphorylatable residues. In these cases, the unambiguous assignment of phosphosites is a non-trivial task. Therefore, to facilitate phosphosite identification, programs such as MASCOT and MSQuant include algorithms that assign a confidence score to each Ser, Thr, and Tyr residue present in a peptide sequence. Similar algorithms, such as the Ascore algorithm, can also be applied to MS spectra to facilitate phosphosite identification (Beausoleil et al., 2006; Olsen et al., 2006; Savitski et al., 2011).

Finally, due to the complex nature of biological samples and the run-to-run variations that can occur at several steps in the fractionation/detection protocol, traditionally it has been difficult to monitor changes in the phosphorylation status of cellular proteins using MS/MS-based approaches. However, the development of quantitative MS methods, such as stable isotope labeling of amino acids in cell culture (SILAC) (Ong, 2012) and isobaric

tags for relative and absolute quantitation (iTRAQ) (Evans et al., 2012), have made it possible to directly compare phosphorylation profiles of multiple samples in a single experiment. These approaches, which rely on isotopic labeling of protein and peptide fragments, respectively, have quickly become mainstays in systems biology research.

In SILAC, two sets of cells are grown in culture media containing either a stable isotope of a particular “heavy” amino acid(s) (usually  $^{13}\text{C}$ -Arg and/or  $^{13}\text{C}$ -Lys) or its normal, “light” counterpart (Figure 2C). As a consequence, cellular proteins are metabolically labeled with either the heavy or the light isotope. Following lysis, the differentially labeled lysates are mixed together and subjected to identical digestion, separation, enrichment, and detection procedures. Meanwhile, iTRAQ and its close cousin, tandem mass tag (TMT) (Dayon and Sanchez, 2012; Jia et al., 2012), incorporate an isotopically labeled tag on peptide fragments *in vitro* following protease digestion. Because peptide fragments are labeled after lysis using iTRAQ and TMT, these approaches are amenable to analysis of primary cells as well as cultured cells. However, as with any *in vitro* reaction, care must be taken to minimize side reactions that can obscure analysis.

The power of isotopic labeling techniques stems from the fact that isotopic isomers are chemically identical—and therefore behave identically during LC-MS/MS—but can be distinguished from one another during detection due to a measurable offset between their  $m/z$  ratios (Figure 2C). As a consequence, relative changes in the phosphorylation status at a particular site can be measured by calculating the ratio of intensities between the heavy and light isotopes for each fragment. These quantitative MS approaches have proven to be powerful means of tracking system-wide changes in phosphorylation networks in a variety of cellular contexts and, as we will see in Section Computational Models of Phosphorylation Networks, provide critical information that facilitates systems-level modeling of dynamic phosphorylation networks.

Together, the advances in MS technology described above have led to the identification and, in many cases quantitative profiling, of a very large number of phosphosites that occur within the cell. Indeed, the identification of thousands, and even tens of thousands, of phosphosites in a single study is often achieved (Choudhary and Mann, 2010). For instance, in a SILAC-based experiment, Olsen et al. employed SCX/TiO<sub>2</sub> enrichment and ion trap-orbitrap MS/MS to identify over 20,400 unique phosphorylation sites on ~6000 cellular proteins isolated from HeLa cells undergoing mitosis (Olsen et al., 2010). In addition to recovering a large percentage of known phosphosites involved in the mitotic transition, this study also uncovered nearly 10,000 previously unknown sites. More recently, Mertins et al. used an i-TRAQ-based protocol to profile ischemia-induced changes in the phosphoproteomes of human ovarian tumor and breast cancer xenograft tissues (Mertins et al., 2014). Among the >25,000 phosphosites identified in these studies, approximately one quarter of them (24%) were either up- or down-regulated in response to short periods of ischemia (<60 min). The altered phosphosites were predominantly associated with stress-response pathways, such as EGFR and MAPK pathways, suggesting that these

pathways are activated at very early stages during the ischemic response.

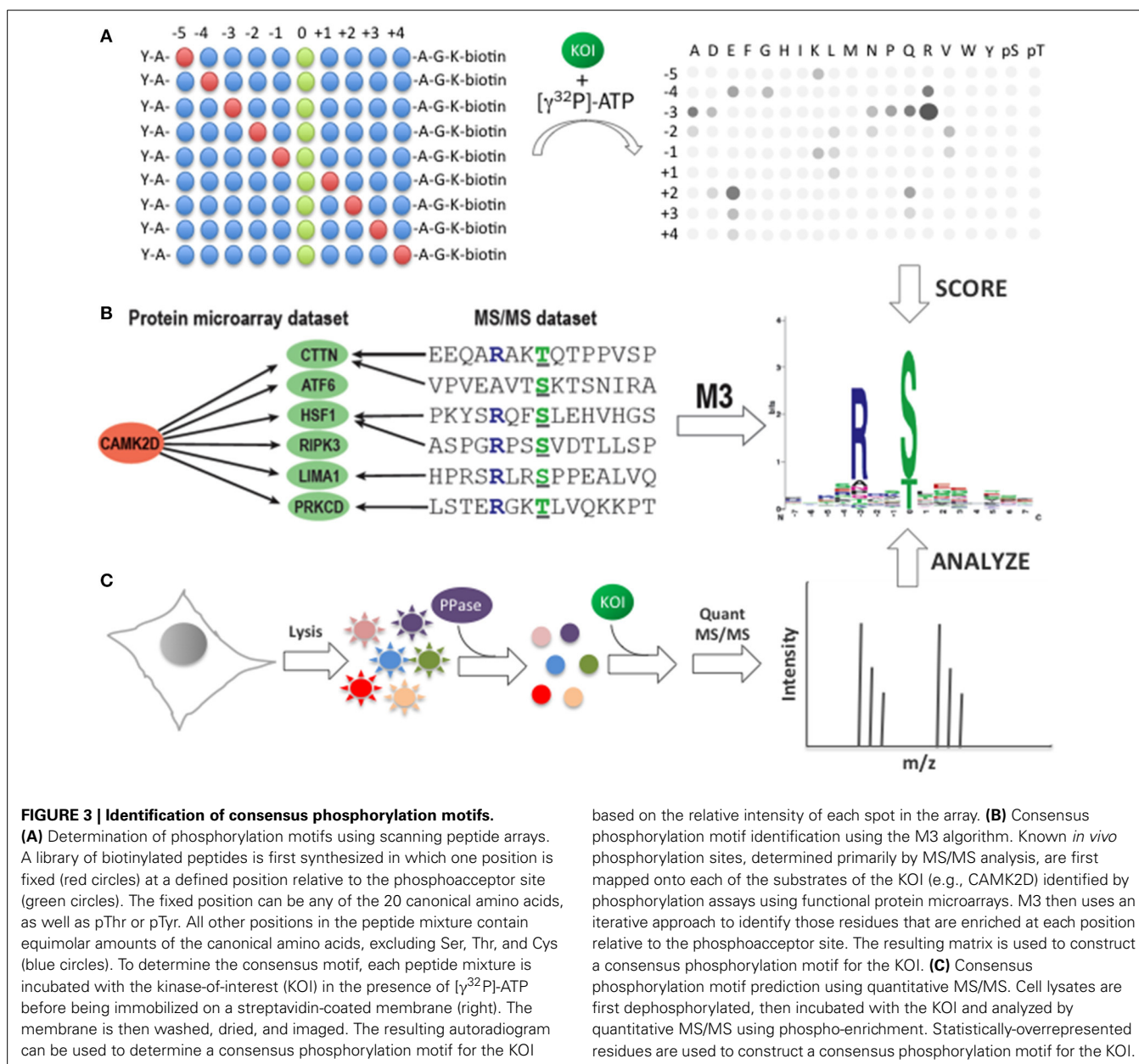
Together, large-scale MS-based studies have greatly expanded the number of annotated phosphoproteins and phosphosites in the phosphoproteome. These data have facilitated global analysis of phosphorylation networks and provided important insights into the cellular targets of protein kinases and phosphatases. However, as alluded to above, MS/MS and other “top-down” approaches are unable to definitively match a given phosphorylation event with the upstream kinase or phosphatase. This notion is underscored by several recent studies that directly examined the interconnectedness of components within phosphorylation-dependent signaling networks. For instance, Bodenmiller et al. employed targeted gene disruption and select analog sensitive kinase (as-kinase) variants to conduct a systems-wide analysis of yeast phosphorylation networks (Bodenmiller et al., 2010). Using quantitative MS/MS, they identified over 8800 phosphorylation events that are regulated by 97 kinases and 27 phosphatases. Not surprisingly, in addition to direct KSRs, they also observed phosphorylation events that were modulated indirectly by disruption of the target kinase/phosphatase activity. Perhaps more interestingly, these analyses revealed that the number of indirect interactions greatly outnumbered the number of direct interactions, accounting for ~2/3 of the total regulated phosphorylation events in the case of kinase gene disruption. Similarly, Breitkreutz et al. generated an extensive kinase-phosphatase interaction (KPI) network in yeast and used it to examine the relationships between phosphorylation-dependent signaling modules. These studies revealed locally dense regions within the network centered around several key signaling molecules, such as the Ser/Thr phosphatase, cell division cycle 14 (Cdc14), and the Ser/Thr kinases, target of rapamycin complex 1 and 2 (TORC1 and TORC2) (Breitkreutz et al., 2010). Within this network, they also uncovered a significantly enriched kinase–kinase (K–K) interaction network that was extremely robust to fragmentation by hub deletion. A similar K–K network is highly conserved between yeast and humans (Newman et al., 2013). Together, these studies suggest that components within phosphorylation-dependent signaling networks are highly interconnected, exhibiting far less modularity than would be expected from the simple linear pathways typically used to depict cellular signaling pathways. As a consequence, modulation of a single network component (e.g., an upstream kinase) is likely to impact the phosphorylation status not only of its immediate downstream substrates, but also that of a large number of substrates that are not directly connected to the target within the network. This makes it extremely difficult to identify direct KSRs for the majority of kinases *in vivo*. Therefore, in addition to the identification of additional phosphoproteins and phosphosites, a major challenge in systems biology research has been to match annotated phosphosites to their upstream kinases and/or phosphatases. This requires information about phosphosites that occur under various cellular conditions, derived from “top-down” approaches, as well as information about direct KSRs, obtained by “bottom-up” approaches. To this end, several methods have been developed to integrate data obtained *in vitro* and *in vivo* to create high-resolution maps of phosphorylation networks.

## COMBINED APPROACHES TO CONNECT PHOSPHOSITES TO THEIR COGNATE KINASES

Methods designed to connect phosphosites to their immediate upstream kinase(s) generally require information about consensus phosphorylation motifs (Figure 3). Traditionally, consensus phosphorylation sites have been derived from positional scanning peptide arrays (Figure 3A). For instance, positional scanning peptide arrays have been designed to identify phosphorylation motifs for Ser/Thr kinases (Chen and Turk, 2010). These arrays are composed of a collection of synthetic peptides containing a central phosphoacceptor site surrounded by degenerate positions containing equimolar amounts of 17 amino acids (Cys, Ser, and Thr are generally excluded). In each peptide, one position (encoding either one of the 20 common amino acids or pThr or pTyr) remains fixed while all other positions are varied. Each peptide mixture is then incubated with the kinase-of-interest in the presence of radioactive ATP, immobilized on a functionalized glass surface, and washed extensively. In a related approach using so-called peptide SPOT microarrays, the peptides are immobilized on a functionalized glass surface prior to incubation with the kinase-of-interest (Leung et al., 2009). After washing, the intensity of each spot is used to identify those residues that are preferred by the kinase-of-interest at a particular position relative to the phosphoacceptor site. Using this approach, consensus phosphorylation motifs have been determined for a number of important protein kinases, including PKA, Src, Akt, PIM1, and PKC.

One of the primary advantages of peptide arrays for consensus phosphorylation motif discovery is that the peptides in the matrix are relatively well defined (at least to the extent that the phosphoacceptor site and the fixed position are constant and all other positions contain equimolar amounts of the other amino acids). However, because the length of the derived motif is limited by the number of residues in the peptides and post-translational modifications other than pTyr or pThr are not present in the synthetic peptides, this approach may lack important information regarding the molecular determinants of kinase specificity. Moreover, since the peptides are not in the context of the full-length protein substrate, other factors involved in substrate recognition, such as docking sites or tertiary structure, are absent. Therefore, as a complementary approach, we developed an algorithm, termed Motif discovery based on protein Microarrays and MS/MS (M3), that combines data about KSRs determined *in vitro* using functional protein microarrays with information about phosphorylation sites identified *in vivo* by MS/MS analyses (Newman et al., 2013) (Figure 3B). M3 first maps *in vivo* phosphorylation sites onto the each of the substrates identified on the microarray for a given kinase. It then uses an iterative approach to identify those residues that are enriched at various positions around the phosphoacceptor site. Using this approach, we predicted consensus motifs for each of the 289 unique kinases in our human kinase collection (including distinct Tyr and Ser/Thr motifs for several dual specificity kinases). When motifs generated by M3 were compared to those generated by positional scanning peptide arrays, a high degree of similarity was observed between them.

One of the greatest strengths of M3 is its flexibility. For instance, though we chose to examine 15-mers in our study in



order to facilitate comparison with motifs determined using scanning peptide arrays, in theory, motifs of various lengths can be generated in a relatively straightforward manner. Likewise, because M3 includes information about the entire amino acid sequence surrounding the putative phosphoacceptor site, aside from information about which residues are preferred at a particular position, it also provides insights into which residues are *disfavored* at select positions. This information may be particularly useful in understanding the substrate selectivity of kinases that have very similar motifs (e.g., closely-related members of the same kinase family). That being said, because it uses an iterative approach for motif discovery, M3 requires a relatively large number of substrates/phosphosites for accurate motif prediction. Likewise, the accuracy of M3 is limited by the number of known KSRs, on the one hand, and the number of annotated

*in vivo* phosphosites, on the other. As these data sets continue to expand, it is likely that the accuracy of M3 with continue to improve, as well.

In addition to the microarray-based strategies described above, several groups have recently developed MS/MS-based approaches for motif discovery (Huang et al., 2007; Chou et al., 2012; Kettenbach et al., 2012; Knight et al., 2012) (Figure 3C). In general, these methods rely on the phosphorylation of eukaryotic cell lysates by an active, recombinant kinase. To identify consensus motifs, the phosphorylation reaction is carried out in lysates that have previously been dephosphorylated, followed by trypsin digestion, phosphopeptide enrichment, and MS/MS analysis. Like M3, this approach benefits from the use of full-length protein substrates. However, because it is often difficult to fully dephosphorylate the lysates, the background signal may be high



using this approach. Recently, Chou and colleagues described an interesting solution to this problem (Chou et al., 2012). Their approach, termed Proteomic Peptide Library (ProPeL), relies on exogenous expression of eukaryotic kinases in *E. coli*. Because *E. coli* express only two Ser/Thr kinases, ProPeL is expected to exhibit a very low background signal when examining eukaryotic Ser/Thr kinases. As a consequence, Ser/Thr phosphorylation events mediated by the kinase-of-interest should be more readily detectable. Using this approach, which uses the bacterial proteome as surrogate substrates for the kinase-of-interest, the authors accurately identified motifs for two Ser/Thr kinases, PKA and Casein Kinase II, suggesting that the use of bacterial proteins does not alter the substrate preferences of these kinases. Though it is well known that the activity of many eukaryotic kinases is poor in bacteria, the use of bacterial lysates with recombinant eukaryotic kinases could facilitate motif discovery using MS/MS-based approaches in the future.

Another attractive solution to this problem is the use of the as-kinases pioneered by Kevan Shokat's group (Bishop et al., 2001; Koch and Hauf, 2010). Analog sensitive kinases are generated by incorporating a functionally silent active site mutation into the kinase-of-interest, allowing the mutant enzyme to utilize an ATP analog that contains a bulky substituent at the N6 position (e.g., N6-benzyl-, N6-cyclopentyl-, or N6-phenylethyl-ATP). This is usually achieved by converting a bulky Met, Leu, Phe, or Thr residue in the ATP binding pocket to a smaller residue, such as Gly or Ala. Because the bulky substituent on the ATP analog cannot be used by either the wild-type version of the kinase or by the majority of other cellular kinases, incubation with the ATP analog containing radioactive phosphate at the  $\gamma$ -position is expected to label only direct substrates of the as-kinase. Subsequent analysis by 2D-PAGE and MS permits the identification of direct substrates. Variations on this approach, for example through the incorporation of a thiol group into the  $\gamma$  phosphate of the ATP analog (e.g., via  $[\gamma\text{S}]\text{-ATP}^4$ ), also allow direct as-kinase substrates to be identified through thiol-dependent enrichment and MS/MS-based analysis. Based on the substrates and phosphosites identified through these analyses, a consensus phosphorylation motif for the as-kinase can be derived. It should be noted that, because the ATP analogs are not cell permeable, this approach is generally restricted to cell lysates. However, as-kinase-specific inhibitors, such as a series of 4-amino-1-tert-butyl-3-phenylpyrazolo[3,4-d]pyrimidine derivatives, can be used to specifically inhibit as-kinases in cells. In this way, indirect identification of as-kinase substrates can be achieved (though, as demonstrated by Bodenmiller et al., this approach may also identify substrates whose phosphorylation status is not directly modulated by the as-kinase Bodenmiller et al., 2010). Recently, Zhang and colleagues used a structure-guided approach to develop a series of potent inhibitors that target a variety of as-kinases (Zhang et al., 2013). Though, to date, as-kinase variants and corresponding ATP analogs have been generated for only a subset of the kinases in the human kinome, the majority of kinases in the human kinome contain a bulky "gatekeeper" residue in the ATP

binding pocket. Therefore, in theory, this approach is generally applicable to a large number of kinases.

Once a consensus motif is identified, it is possible to make predictions about which kinases are responsible for the phosphorylation event. Indeed, by combining information about consensus phosphorylation motifs with phosphosites identified by MS/MS analysis, several groups have developed computational methods to predict which kinase(s) is likely to phosphorylate a given site *in vivo* (Linding et al., 2007, 2008; Song et al., 2012; Newman et al., 2013). Though many of the early KSR prediction methods utilized sophisticated search algorithms, such as position-specific scoring matrices, neural networks, or support vector machines, to scan the primary amino acid sequence of a given protein for the motif-of-interest (Yaffe et al., 2001; Brinkworth et al., 2003; Blom et al., 2004; Hjerrild et al., 2004; Kim et al., 2004; Xue et al., 2005, 2006; Wong et al., 2007), these algorithms were generally more successful at identifying putative *in vitro* KSRs than they were at accurately predicting KSRs that actually occur inside the cell. This is due, in large part, to lack of contextual information about the kinase-substrate pair (Linding et al., 2008; Song et al., 2012). Therefore, in addition to consensus phosphorylation motifs, more recent platforms also integrate GO data, such as information about protein expression, subcellular localization, and protein-protein interactions, to improve the predictive power of the algorithms (Linding et al., 2007, 2008; Song et al., 2012; Newman et al., 2013). In this way, extensive phosphorylation networks have been constructed that connect cellular substrates and/or phosphosites to their immediate upstream kinase(s). Though still in their infancy, the utility of these methods has grown considerably in recent years. For instance, using current data about validated *in vivo* KSRs, the true positive rate of the NetworKIN algorithm, which pioneered the use of consensus motifs (determined by scanning peptide arrays) with contextual information about kinase-substrate pairs in cells, is 0.76% (48 true positives out of a total 6338 phosphosite identifications) (Linding et al., 2007; Newman et al., 2013). Meanwhile, more recent approaches, such as the *in vivo* group-based prediction system (iGPS) (Song et al., 2012) and the CEASAR approach (which Connects Enzymes And Substrates at Amino acid Resolution) perform markedly better, with the latter exhibiting a true positive rate of 17.2% (758/4417) (Newman et al., 2013). Though the predictive power of these approaches is currently limited by several parameters, including the availability of GO data, *in vivo* phosphosites, and information about consensus phosphorylation motifs, the high resolution maps that they generate have the potential to provide a wealth of information about the organization of cellular phosphorylation networks. This information will be invaluable for identifying those KSRs that are likely to occur inside the cell. Importantly, they also provide a framework for identifying points of signal integration between distinct pathways.

## REGULATION OF PHOSPHORYLATION-DEPENDENT SIGNALING PATHWAYS

In addition to information about the sites on a substrate that *can* be phosphorylated by a given kinase (or dephosphorylated by given phosphatase), it is important to know *when* and *where*

<sup>4</sup> $[\gamma\text{S}]\text{-ATP}$  can be utilized by most protein kinases, but not by most other cellular ATPases (Koch and Hauf, 2010).



phosphorylation is likely to occur inside the cell. Such information can provide key insights into the physiological consequences of phosphorylation. Indeed, the tight spatial and temporal regulation of cellular phosphorylation networks is believed to underlie the specificity of cellular signal transduction. Therefore, in order to gain a more comprehensive understanding of dynamic phosphorylation networks, the static maps of phosphorylation networks described in the previous section will need to be supplemented with information about the spatiotemporal regulation of the protein kinases and phosphatases within these networks.

Traditionally, this has been achieved by examining changes in the phosphorylation state of representative cellular substrates before and after stimulation. Changes in the phosphorylation status of these proteins are typically determined using two complementary approaches. In the first approach, cellular proteins are isolated by either subcellular fractionation or immunoprecipitation so that their phosphorylation status can be probed using one of the phosphodetection methods discussed above (e.g., 2D-PAGE, MS/MS or western blotting using phosphospecific antibodies). While careful experimental design and the use of general phosphatase and/or kinase inhibitors allows for decent temporal resolution using this approach, its spatial resolution is often poor because cellular architecture is disrupted during the lysis procedure. Moreover, since a large number of cells are required to obtain enough protein for analysis, this approach lacks single cell resolution, potentially masking subtle, yet important, differences between individual cells. The other approach, immunofluorescence using phosphospecific antibodies, offers excellent spatial resolution and can provide insights into cell-to-cell differences within a given population. However, because the cells must be fixed prior to analysis, this approach suffers from poor temporal resolution. Importantly, neither approach is able to examine changes in phosphorylation in living cells. Therefore, these methods offer only a “snapshot” of the dynamic changes in kinase and phosphatase activities that characterize most phosphorylation-dependent signal networks under physiological conditions.

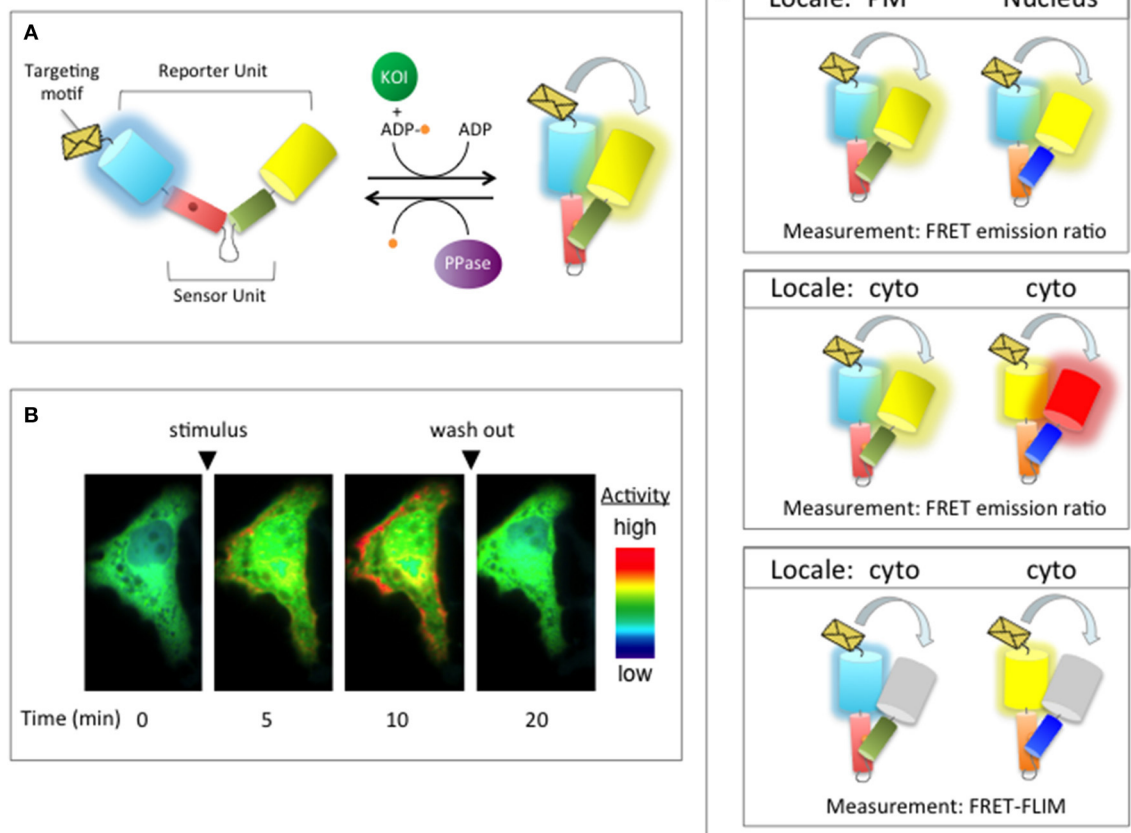
#### GENETICALLY-TARGETABLE BIOSENSORS TO TRACK KINASE ACTIVITY PROFILES IN LIVING CELLS

In order to study the activity profiles of cellular kinases and phosphatases in single, living cells with high spatiotemporal resolution, several groups have developed genetically-targetable FRET-based biosensors (Zhang and Allen, 2007; Herbst et al., 2009; Newman et al., 2011; Sipiet et al., 2013) (**Figures 4A,B**). These biosensors, which can be directed to specific subcellular regions through the incorporation of a targeting motif (e.g., a NLS) or a component of a signaling complex (e.g., a scaffold protein) (Zhang et al., 2001; Kunkel and Newton, 2014), are able to monitor real-time changes in the activity profiles of specific pools of a given kinase or phosphatase in living cells (Kunkel and Newton, 2009; Gao and Zhang, 2010). Importantly, as outlined in Section Computational Models of Phosphorylation Networks, the high spatiotemporal resolution afforded by these sensors provides kinetic data that can be integrated into computational models that provide insights into the behaviors of entire signaling networks (Saucerman et al., 2006; Violin et al., 2008; Ni et al.,

2011; Greenwald et al., 2014). Moreover, with the development of bright, spectrally distinct FPs that span the visible spectrum (Day and Davidson, 2009; Sample et al., 2009; Newman et al., 2011; Ai et al., 2014) and sophisticated imaging/deconvolution protocols (Grant et al., 2008; Woehler, 2013), it is now possible to measure changes in the activity of multiple cellular signaling enzymes (e.g., two kinases or a kinase and a phosphatase) simultaneously in the same cell (Carlson and Campbell, 2009; Woehler, 2013) (**Figure 4C**). Such information will be particularly valuable for understanding crosstalk between distinct signaling pathways within phosphorylation networks.

FRET-based biosensors contain two basic components: (1) a “sensor unit,” which undergoes a conformational change in response to a given cellular stimulus (e.g., phosphorylation or enzyme activation) and (2) a “reporter unit,” which converts the induced conformational change into a change in FRET (Frommer et al., 2009; Newman et al., 2011) (**Figure 4A**). While the reporter units of most FRET-based kinase and phosphatase reporters utilize a FP FRET pair consisting of cyan FP (CFP) and yellow FP (YFP) color variants, other FP combinations, such as green/red FP (GFP/RFP), yellow/orange FP (YFP/OF), YFP/RFP and CFP/RFP, have also been used successfully (Sample et al., 2009; Newman et al., 2011; Day and Davidson, 2012; Lam et al., 2012; Ai et al., 2014). The continued refinement of the photophysical properties of FPs for FRET-based applications will be critical for the development of future biosensors. As alluded to above, this will also facilitate the development of orthogonal activity reporters that can be used to monitor more than one signaling enzyme in the same cell (**Figure 4C**). In this respect, the development of FPs that exhibit a large Stokes shift, such as mAmetrine (Ai et al., 2008) and T-Sapphire (Zapata-Hommer and Griesbeck, 2003), hold great promise for multicomponent FRET imaging (Carlson and Campbell, 2009).

While the “reporter unit” relies upon the FRET efficiency between the FP FRET pair, the molecular switch utilized by the “sensor unit” can take several different forms. The primary consideration when developing a molecular switch is that it promotes a conformational change in response to the cellular parameter under study. For instance, some FRET-based biosensors are designed to monitor the conformational state of the signaling enzyme itself. Indeed, kinase activation sensors, which directly monitor the activation of a kinase-of-interest, have been used to study the spatiotemporal regulation of several important kinases, including extracellular-regulated kinase 2 (ERK2) (Fujioka et al., 2006), MAP kinase-activated protein kinase 2 (MK2) (Neininger et al., 2001), Akt (Yoshizaki et al., 2006; Ananthanarayanan et al., 2007; Calleja et al., 2007), focal adhesion kinase (FAK) (Ritt et al., 2013), CaMKII (Erickson et al., 2011), and phosphoinositide-dependent kinase 1 (PDK1) (Calleja et al., 2007; Gao et al., 2011). For instance, this approach was recently used to monitor the activation of PDK1 in different membrane microdomains. To this end, Gao et al. sandwiched full-length PDK1 between enhanced CFP (ECFP) and the YFP variant, mCitrine (Gao et al., 2011). The resulting PDK1 activation reporter (PARE) was localized to either raft or non-raft regions of the plasma membrane using targeting motifs derived from Lyn and K-Ras, respectively, and used to track changes in PDK1 activation following



**FIGURE 4 | Genetically-targetable FRET-based biosensors. (A)** Design of a FRET-based kinase activity reporter based on an engineered molecular switch. In this design, the sensor unit is composed of a PAABD (red cylinder) tethered to a substrate domain (green cylinder) that is specifically phosphorylated by the kinase-of-interest (KOI; green circle). The sensor unit is sandwiched between the reporter unit, which is comprised of two FPs that are able to undergo FRET [e.g., CFP (cyan cylinder) and YFP (yellow cylinder)]. A targeting motif (orange envelope) is used to direct the reporter to distinct subcellular regions. In the unphosphorylated state, the FPs are far removed from one another and, therefore, do not undergo FRET. An increase in the activity of the KOI leads to a phosphorylation-dependent conformational change that alters the distance and/or orientation of the FPs, increasing FRET between them. Phosphatase (PPase; purple oval)-mediated dephosphorylation of the reporter switches it back to the open conformation, reducing FRET. **(B)** Pseudocolor images of a live cell imaging experiment. In this experiment, a time-dependent increase in the FRET emission ratio (YFP

FRET/CFP) is observed following stimulation with a pharmacological activator (stimulus). Upon removal of the stimulus (wash out), the emission ratio returns to basal levels. Warmer colors indicate high activity while cooler colors indicate low activity. **(C)** Several ways in which genetically-targetable biosensors can be used to monitor real-time changes in the activity profiles of two or more signaling enzymes in the same cell. Top panel: Activity reporters that utilize the same FP FRET pair can be monitored simultaneously in the same cell provided that each biosensor is targeted to a distinct subcellular locale, such as the plasma membrane (PM) and the nucleus. Middle and bottom panels: To track the activities of two or more signaling enzymes in the same subcellular region, activity reporters that utilize spectrally distinct FP FRET pairs (e.g., CFP/YFP and YFP/RFP) can be used (middle panel) or alternative fluorescence imaging techniques, such as FRET-fluorescence lifetime imaging (FRET-FLIM), which only measures changes in emission of the donor fluorophore, can be employed (bottom panel). Cyto, cytoplasm; PM, plasma membrane.

stimulation with the growth factor, platelet-derived growth factor (PDGF). Interestingly, these studies demonstrated that, following PDGF stimulation, PDK1 is activated in raft, but not in non-raft, regions. The differential regulation of PDK1 appears to be dependent on the presence of the lipid phosphatase PTEN, which dephosphorylates phosphatidylinositol (3,4,5) triphosphate (PIP<sub>3</sub>) required for PDK1 activation. Because PDK1 is a master regulator of many AGC family members, including Akt, PKA, SGK and several PKC isoforms, these results have important implications for understanding the regulation of PDK1-dependent signaling networks (Arencibia et al., 2013).

In addition to biosensors designed to monitor kinase *activation*, many biosensors have also been developed to monitor changes in kinase *activity*. Unlike the activation reporters discussed above, which exhibit a linear relationship between the signal response and the activation “event,” kinase activity reporters benefit from enzymatic amplification of the signal. This is because multiple reporter species can be phosphorylated by a single activated kinase. This feature stems from the design of the sensor unit. Indeed, the sensor units of all currently available kinase activity reporters utilize an engineered molecular switch based upon a modular design (Zhang and Allen, 2007; Mehta and

Zhang, 2011; Newman et al., 2011) (**Figure 4A**). Accordingly, a consensus phosphorylation site specific for the kinase-of-interest serves as the “substrate domain” while a PAABD specific for the phosphorylated form of the substrate functions as the “switching segment.” The substrate domain and the switching segment are typically joined together by a flexible linker and sandwiched between a FP FRET pair, such as CFP and YFP. Whereas the length of the linker and the choice of FRET pairs influence the dynamic range of the reporter (Komatsu et al., 2011; Belal et al., 2014), the choice of the substrate domain and PAABD contribute to its specificity and reversibility, respectively. The reversibility of the reporter is critical for monitoring kinase attenuation, which provides a clearer picture of kinase regulation within intact signaling networks. For instance, although the first generation A-kinase activity reporter, AKAR1, enabled kinetic analysis of subcellular PKA activity induced by the  $\beta$ -AR agonist, isoproterenol, this sensor was unable to report the attenuation of PKA activity following receptor desensitization because its response was essentially irreversible inside cells (Zhang et al., 2001). This is likely due to the fact that the 14-3-3 $\tau$  PAABD utilized by AKAR1 binds the phosphorylated form of the substrate domain very tightly, preventing cellular phosphatases from gaining access to the substrate region as PKA activity declines. Therefore, in order to visualize both increases and decreases in PKA activity, a second generation PKA reporter, AKAR2, was constructed by replacing 14-3-3 $\tau$  with the weaker-binding PAABD, forkhead-associated 1 (FHA1) (Zhang et al., 2005a). While the activation profile of AKAR2 is similar to that observed for AKAR1, the former is also readily reversible following removal of agonist or treatment with the PKA inhibitor, H89.

The basic modular design described above has been used to create activity reporters for a number of protein kinases, including PKA, protein kinase C (PKC), ataxia telangiectasia mutated (ATM), Akt, Abl, Src, aurora kinase B, ERK, c-Jun N-terminal kinase (JNK), cyclin-dependent kinase 1 (CDK1), AMP-activated protein kinase (AMPK), and the epidermal growth factor receptor (EGFR) (Zhang and Allen, 2007; Newman et al., 2011; Tsou et al., 2011; Belal et al., 2014). Using these reporters, researchers have uncovered important details about both the kinetics and the spatial distribution of endogenous kinase action in a variety of cellular contexts (Zhang et al., 2005a; Zhang and Allen, 2007; Erickson et al., 2011; Gao et al., 2011; Komatsu et al., 2011; Mehta and Zhang, 2011; Newman et al., 2011; Tsou et al., 2011; Arencibia et al., 2013; Ritt et al., 2013; Belal et al., 2014). However, despite their unique ability to track kinase activity in real time and at single cell resolution, to date, kinase activity reporters are available for <3% of the 518 human kinases in the human kinome. This is due, in large part, to limited information about the substrate specificity of most human kinases. Indeed, this is one of the primary obstacles to the large-scale development of kinase activity reporters. The recent identification of consensus phosphorylation motifs for 289 unique kinases (representing ~55% of the human kinome) (Newman et al., 2013), as well as the continued characterization of PAABD family members (Jin and Pawson, 2012; Reinhardt and Yaffe, 2013), will facilitate the development of a large number of novel kinase activity sensors that promise to offer important insights into

the spatial and temporal regulation of cellular phosphorylation networks.

### GENETICALLY-TARGETABLE BIOSENSORS TO TRACK PHOSPHATASE ACTIVITY PROFILES IN LIVING CELLS

In addition to information about kinase regulation, a comprehensive map of phosphorylation networks will also require an understanding of the dynamic regulation of protein phosphatases. However, until recently, a general design for phosphatase activity reporters had not been described. To address this issue, we engineered a phosphatase activity sensor designed to measure the activity of the  $\text{Ca}^{2+}$ /calmodulin-regulated Ser/Thr protein phosphatase, CaN (Newman and Zhang, 2008; Mehta and Zhang, 2014). This reporter, termed CaN activity reporter 1 (CaNAR1), utilizes an intrinsic molecular switch based upon dephosphorylation-induced conformational changes within the regulatory region of NFAT1 (Okamura et al., 2000). Upon dephosphorylation, CaNAR1 exhibits an increase in its emission ratio that is dependent on CaN activity. Importantly, because the regulatory region of NFAT1 is hyperphosphorylated in resting cells by cellular kinases such as p38 and the constitutively-active kinase, casein kinase 1  $\alpha$  (CK1 $\alpha$ ), CaNAR1 does not require activation of additional kinases to put it into a “dephosphorylation competent” state. This feature ensures that the cellular environment remains relatively unperturbed prior to  $\text{Ca}^{2+}$  stimulation. This and other design features utilized by CaNAR1 should be generally applicable to other protein phosphatases as specific molecular switches are identified or engineered. Thus, as a prototype phosphatase activity sensor, CaNAR1 lays a foundation for studying the targeting and compartmentalization of protein phosphatases within the cellular environment.

### COMPUTATIONAL MODELS OF PHOSPHORYLATION NETWORKS

The development of novel biosensors will enable the activity profiles of a large number of protein kinases and phosphatases to be measured within the cellular environment, providing the experimental information necessary to build detailed computational models of phosphorylation networks. Such approaches will be important in order to gain a systems-level understanding of the complex regulation of phosphorylation networks in response to various cellular stimuli. To this end, quantitative FRET-based imaging data has been incorporated into mathematical models of phosphorylation networks based on ordinary differential equations (Saucerman et al., 2006; Violin et al., 2008; Ni et al., 2011; Song et al., 2012). One of the primary goals of this approach is to determine the relative contributions of multiple feedback and feed-forward loops in producing the tight spatiotemporal control exhibited by many signal transduction pathways. Not only does this help to uncover details about the behavior of individual components within a given signaling network, but it can also provide crucial insights into how information is propagated throughout the entire system. Because these types of questions are difficult (if not impossible) to answer using experimental approaches alone, computational methods can provide a more comprehensive view of signaling networks that will ultimately promote a better understanding of cellular signaling at the systems level. For instance, computational models have been developed to study the

impact of signaling enzymes in diverse cellular processes, including PKA-mediated phosphorylation gradients in cardiomyocytes (Saucerman et al., 2006), hippocampal neurons (Neves et al., 2008), and pancreatic  $\beta$ -cells (Ni et al., 2011). Similar models have also been constructed to examine the mechanisms by which the activities of  $\text{Ca}^{2+}$ /calmodulin-dependent signaling enzymes, such as CaMKII and CaN, are differentially regulated in response to dynamic changes in intracellular  $\text{Ca}^{2+}$  concentrations (Song et al., 2008).

In addition to computational models based on fluorescence imaging data, the advent of quantitative MS/MS methods, such as SILAC and iTRAQ, has also permitted the development of computational models that focus on global changes in phosphorylation profiles (Kozuka-Hata et al., 2012). While models based on data obtained using fluorescent biosensors can best be described as kinetic and/or stochastic because they are built upon detailed kinetic information about the behavior of each component in the system (e.g., individual kinase and phosphatase activities), models based on quantitative MS/MS data are best characterized as discrete models because they describe changes in the general profile of an entire population or system (Wu et al., 2009). For instance, computational methods based on self-organizing maps (Zhang et al., 2005b), partial least squares regression analysis (Wolf-Yadlin et al., 2006; Kumar et al., 2007), Bayesian network modeling (Bose et al., 2006; Guha et al., 2008), and numerical modeling (Tasaki et al., 2006) have been developed. More recently, Tian and Song described a general computational framework that can be used to develop mathematical models derived from multiple quantitative phosphoproteomic data sets (Tian and Song, 2012). This model, which was used to model MAPK signaling pathways, will be particularly useful for modeling highly time-resolved phosphoproteomic data sets, such as the one reported by Oyama et al. (2009). In this study, the authors used SILAC and pTyr-enrichment to construct a detailed kinetic analysis of the Tyr phosphoproteome following EGFR activation in human epithelial A431 cells. These studies, which measured relative levels of Tyr phosphorylation at 0.5, 2, 5, 10, 15, 20, 25, and 30 min following EGFR activation, identified both transient and sustained changes in pTyr levels among a collection of 77 different cellular proteins.

Importantly, with continued advances in biosensor development and quantitative MS/MS methodologies, the quality and depth of information that can be incorporated into models of phosphorylation networks will continue to increase, promoting a more quantitative understanding of the molecular mechanisms that influence signaling dynamics inside the cell. Such quantitative information will be critical if we are to truly understand the functional interactions that occur between individual signaling enzymes to drive complex cellular behaviors.

## CONCLUSION AND PERSPECTIVE

Phosphorylation-dependent signaling networks underlie diverse cellular processes, including metabolism, cell cycle progression, the immune response, and cell migration. Recently, the emergence of several novel technologies has provided a truly systems-level view of dynamic phosphorylation networks. Together, these technologies are beginning to provide us with a clearer picture about (1) which cellular proteins are phosphorylated in

response to a given stimulus and/or in a particular cellular context, (2) where on the protein the phosphorylation event occurs, (3) which kinases and phosphatases are mediating phosphorylation/dephosphorylation of the phosphosite, and (4) how these modifying enzymes are regulated inside the cell. In the future, it will be important to expand this knowledge-base through (1) systems-wide profiling of phosphoproteomes under different cellular conditions and disease states using quantitative MS/MS and 2D-DIGE; (2) the identification of additional KSRs [and their relatively under-studied counterparts, phosphatase-substrate relationships (PSRs)] using “whole-proteome” microarrays; (3) the development of novel fluorescent biosensors and orthogonal imaging modalities; and (4) the integration of these data into predictive models of phosphorylation networks. Though much work remains to be done, these goals appear to be within reach. The next challenge will be to gain a detailed understanding of the functional consequences of specific phosphorylation events and to integrate this information with other resources, such as protein–protein interaction data sets, global expression profiles, and metabolomics data (Derouiche et al., 2012; Harrold et al., 2013; Medina-Cleghorn and Nomura, 2013; Bordbar et al., 2014). Likewise, in order to understand crosstalk between different cellular signals (e.g., O-glycosylation and phosphorylation), it will be necessary to adapt the approaches developed to study phosphoproteomics to other post-translational modifications (Choudhary and Mann, 2010; Zhu et al., 2012; D’hondt et al., 2013; Sutandy et al., 2013). Together, this information will provide a comprehensive view of the organization and regulation of cellular phosphorylation networks and beyond.

## ACKNOWLEDGMENTS

This work was supported by NSF HRD-1038160 (to Robert H. Newman), NIH R01 DK073368 (to Jin Zhang) and NIH GM076102, RR020839, CA160036, and HG 006434 (to Heng Zhu).

## REFERENCES

- Agrawal, U., Reilly, D. T., and Schroeder, C. M. (2013). Zooming in on biological processes with fluorescence nanoscopy. *Curr. Opin. Biotechnol.* 24, 646–653. doi: 10.1016/j.copbio.2013.02.016
- Aguilar, H. N., Tracey, C. N., Tsang, S. C., McGinnis, J. M., and Mitchell, B. F. (2011). Phos-tag-based analysis of myosin regulatory light chain phosphorylation in human uterine myocytes. *PLoS ONE* 6:e20903. doi: 10.1371/journal.pone.0020903
- Ai, H. W., Baird, M. A., Shen, Y., Davidson, M. W., and Campbell, R. E. (2014). Engineering and characterizing monomeric fluorescent proteins for live-cell imaging applications. *Nat. Protoc.* 9, 910–928. doi: 10.1038/nprot.2014.054
- Ai, H. W., Hazelwood, K. L., Davidson, M. W., and Campbell, R. E. (2008). Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors. *Nat. Methods* 5, 401–403. doi: 10.1038/nmeth.1207
- Ananthanarayanan, B., Fosbrink, M., Rahdar, M., and Zhang, J. (2007). Live-cell molecular analysis of Akt activation reveals roles for activation loop phosphorylation. *J. Biol. Chem.* 282, 36634–36641. doi: 10.1074/jbc.M706227200
- Andersson, L., and Porath, J. (1986). Isolation of phosphoproteins by immobilized metal ( $\text{Fe}^{3+}$ ) affinity chromatography. *Anal. Biochem.* 154, 250–254. doi: 10.1016/0003-2697(86)90523-3
- Arencibia, J. M., Pastor-Flores, D., Bauer, A. F., Schulze, J. O., and Biondi, R. M. (2013). AGC protein kinases: from structural mechanism of regulation to allosteric drug development for the treatment of human diseases. *Biochim. Biophys. Acta* 1834, 1302–1321. doi: 10.1016/j.bbapap.2013.03.010



- Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., and Gygi, S. P. (2006). A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* 24, 1285–1292. doi: 10.1038/nbt1240
- Belal, A. S., Sell, B. R., Hoi, H., Davidson, M. W., and Campbell, R. E. (2014). Optimization of a genetically encoded biosensor for cyclin B1-cyclin dependent kinase 1. *Mol. Biosyst.* 10, 191–195. doi: 10.1039/c3mb70402e
- Besant, P. G., and Attwood, P. V. (2010). Histidine phosphorylation in histones and in other mammalian proteins. *Methods Enzymol.* 471, 403–426. doi: 10.1016/S0076-6879(10)71021-1
- Biggs, M. J., Milone, M. C., Santos, L. C., Gondarenko, A., and Wind, S. J. (2011). High-resolution imaging of the immunological synapse and T-cell receptor microclustering through microfabricated substrates. *J. R. Soc. Interface* 8, 1462–1471. doi: 10.1098/rsif.2011.0025
- Bishop, A. C., Buzko, O., and Shokat, K. M. (2001). Magic bullets for protein kinases. *Trends Cell Biol.* 11, 167–172. doi: 10.1016/S0962-8924(01)01928-6
- Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633–1649. doi: 10.1002/pmic.200300771
- Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P. G., et al. (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci. Signal.* 3, rs4. doi: 10.1126/scisignal.2001182
- Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* 15, 107–120. doi: 10.1038/nrg3643
- Bose, R., Molina, H., Patterson, A. S., Bitok, J. K., Periaswamy, B., Bader, J. S., et al. (2006). Phosphoproteomic analysis of Her2/neu signaling and inhibition. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9773–9778. doi: 10.1073/pnas.0603948103
- Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., et al. (2010). A global protein kinase and phosphatase interaction network in yeast. *Science* 328, 1043–1046. doi: 10.1126/science.1176495
- Brinkworth, R. I., Breinl, R. A., and Kobe, B. (2003). Structural basis and prediction of substrate specificity in protein serine/threonine kinases. *Proc. Natl. Acad. Sci. U.S.A.* 100, 74–79. doi: 10.1073/pnas.0134224100
- Calleja, V., Alcor, D., Laguerre, M., Park, J., Vojnovic, B., Hemmings, B. A., et al. (2007). Intramolecular and intermolecular interactions of protein kinase B define its activation *in vivo*. *PLoS Biol.* 5:e95. doi: 10.1371/journal.pbio.0050095
- Camacho-Carvajal, M. M., Wollscheid, B., Aebersold, R., Steimle, V., and Schamel, W. W. (2004). Two-dimensional Blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. *Mol. Cell Proteomics* 3, 176–182. doi: 10.1074/mcp.T300010-MCP200
- Carlson, H. J., and Campbell, R. E. (2009). Genetically encoded FRET-based biosensors for multiparameter fluorescence imaging. *Curr. Opin. Biotechnol.* 20, 19–27. doi: 10.1016/j.copbio.2009.01.003
- Chao, J. A., Yoon, Y. J., and Singer, R. H. (2012). Imaging translation in single cells using fluorescent microscopy. *Cold Spring Harb. Perspect. Biol.* 4:a012310. doi: 10.1101/cshperspect.a012310
- Chen, C., and Turk, B. E. (2010). Analysis of serine-threonine kinase specificity using arrayed positional scanning peptide libraries. *Curr. Protoc. Mol. Biol.* Chapter 18:Unit 18.14. doi: 10.1002/0471142727.mb1814s91
- Chou, M. F., Pristic, S., Lubner, J. M., Church, G. M., Husson, R. N., and Schwartz, D. (2012). Using bacteria to determine protein kinase specificity and predict target substrates. *PLoS ONE* 7:e52747. doi: 10.1371/journal.pone.0052747
- Choudhary, C., and Mann, M. (2010). Decoding signalling networks by mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* 11, 427–439. doi: 10.1038/nrm2900
- Ciruela, F. (2008). Fluorescence-based methods in the study of protein-protein interactions in living cells. *Curr. Opin. Biotechnol.* 19, 338–343. doi: 10.1016/j.copbio.2008.06.003
- Craig, R., and Beavis, R. C. (2003). A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun. Mass Spectrom.* 17, 2310–2316. doi: 10.1002/rcm.1198
- D'hondt, C., Iyyathurai, J., Vinken, M., Rogiers, V., Leybaert, L., Himpens, B., et al. (2013). Regulation of connexin- and pannexin-based channels by post-translational modifications. *Biol. Cell* 105, 373–398. doi: 10.1111/boc.201200096
- Day, R. N., and Davidson, M. W. (2009). The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* 38, 2887–2921. doi: 10.1039/b901966a
- Day, R. N., and Davidson, M. W. (2012). Fluorescent proteins for FRET microscopy: monitoring protein interactions in living cells. *Bioessays* 34, 341–350. doi: 10.1002/bies.201100098
- Dayon, L., and Sanchez, J. C. (2012). Relative protein quantification by MS/MS using the tandem mass tag technology. *Methods Mol. Biol.* 893, 115–127. doi: 10.1007/978-1-61779-885-6\_9
- Dedecker, P., Mo, G. C., Dertinger, T., and Zhang, J. (2012). Widely accessible method for superresolution fluorescence imaging of living systems. *Proc. Natl. Acad. Sci. U.S.A.* 109, 10909–10914. doi: 10.1073/pnas.1204917109
- Dephoure, N., Gould, K. L., Gygi, S. P., and Kellogg, D. R. (2013). Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. *Mol. Biol. Cell* 24, 535–542. doi: 10.1091/mbc.E12-09-0677
- Derouiche, A., Cousin, C., and Mijakovic, I. (2012). Protein phosphorylation from the perspective of systems biology. *Curr. Opin. Biotechnol.* 23, 585–590. doi: 10.1016/j.copbio.2011.11.008
- Deschenes-Simard, X., Kottakis, F., Meloche, S., and Ferbeyre, G. (2014). ERKs in cancer: friends or foes? *Cancer Res.* 74, 412–419. doi: 10.1158/0008-5472.CAN-13-2381
- Deswal, S., Beck-Garcia, K., Blumenthal, B., Dopfer, E. P., and Schamel, W. W. (2010). Detection of phosphorylated T and B cell antigen receptor species by Phos-tag SDS- and Blue Native-PAGE. *Immunol. Lett.* 130, 51–56. doi: 10.1016/j.imlet.2009.12.012
- Dissmeyer, N., and Schnittger, A. (2011). The age of protein kinases. *Methods Mol. Biol.* 779, 7–52. doi: 10.1007/978-1-61779-264-9\_2
- Eng, J. K., McCormack, A. L., and Yates, J. R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 5, 976–989. doi: 10.1016/1044-0305(94)80016-2
- Erickson, J. R., Patel, R., Ferguson, A., Bossuyt, J., and Bers, D. M. (2011). Fluorescence resonance energy transfer-based sensor Camui provides new insight into mechanisms of calcium/calmodulin-dependent protein kinase II activation in intact cardiomyocytes. *Circ. Res.* 109, 729–738. doi: 10.1161/CIRCRESAHA.111.247148
- Evans, C., Noirel, J., Ow, S. Y., Salim, M., Pereira-Medrano, A. G., Couto, N., et al. (2012). An insight into iTRAQ: where do we stand now? *Anal. Bioanal. Chem.* 404, 1011–1027. doi: 10.1007/s00216-012-5918-6
- Frommer, W. B., Davidson, M. W., and Campbell, R. E. (2009). Genetically encoded biosensors based on engineered fluorescent proteins. *Chem. Soc. Rev.* 38, 2833–2841. doi: 10.1039/b907749a
- Fujioka, A., Terai, K., Itoh, R. E., Aoki, K., Nakamura, T., Kuroda, S., et al. (2006). Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *J. Biol. Chem.* 281, 8917–8926. doi: 10.1074/jbc.M509344200
- Galken, P. R., and Lampe, P. D. (2006). Methodologies for characterizing phosphoproteins by mass spectrometry. *Cell Commun. Adhes.* 13, 249–262. doi: 10.1080/15419060601077917
- Gao, X., Lowry, P. R., Zhou, X., Depry, C., Wei, Z., Wong, G. W., et al. (2011). PI3K/Akt signaling requires spatial compartmentalization in plasma membrane microdomains. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14509–14514. doi: 10.1073/pnas.1019386108
- Gao, X., and Zhang, J. (2010). FRET-based activity biosensors to probe compartmentalized signaling. *Chembiochem* 11, 147–151. doi: 10.1002/cbic.200900594
- Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., et al. (2004). Open mass spectrometry search algorithm. *J. Proteome Res.* 3, 958–964. doi: 10.1021/pr0499491
- Glickman, J. F. (2012). “Assay development for protein kinases” in *Source Assay Guidance Manual [Internet]*, eds G. S. Sittampalam, N. Gal-Edd, M. Arkin, D. Auld, C. Austin, B. Bejcek, et al. (Bethesda, MD: Eli Lilly & Company and the National Center for Advancing Translational Sciences). (Updated 2012 October 01).
- Gorreta, E., Carbone, W., and Barzaghi, D. (2012). Genomic profiling: cDNA arrays and oligoarrays. *Methods Mol. Biol.* 823, 89–105. doi: 10.1007/978-1-60327-216-2\_7
- Grant, D. M., Zhang, W., McGhee, E. J., Bunney, T. D., Talbot, C. B., Kumar, S., et al. (2008). Multiplexed FRET to image multiple signaling events in live cells. *Biophys. J.* 95, L69–L71. doi: 10.1529/biophysj.108.139204



- Greenwald, E. C., Polanowska-Grabowska, R. K., and Saucerman, J. J. (2014). Integrating fluorescent biosensor data using computational models. *Methods Mol. Biol.* 1071, 227–248. doi: 10.1007/978-1-62703-622-1\_18
- Guha, U., Chaerkady, R., Marimuthu, A., Patterson, A. S., Kashyap, M. K., Harsha, H. C., et al. (2008). Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14112–14117. doi: 10.1073/pnas.0806158105
- Guo, S. (2014). Insulin signaling, resistance, and the metabolic syndrome: insights from mouse models into disease mechanisms. *J. Endocrinol.* 220, T1–T23. doi: 10.1530/JOE-13-0327
- Harrold, J. M., Ramanathan, M., and Mager, D. E. (2013). Network-based approaches in drug discovery and early development. *Clin. Pharmacol. Ther.* 94, 651–658. doi: 10.1038/clpt.2013.176
- He, L., Han, X., and Ma, B. (2013). De novo sequencing with limited number of post-translational modifications per peptide. *J. Bioinform. Comput. Biol.* 11, 1350007. doi: 10.1142/S0219720013500078
- Herbst, K. J., Ni, Q., and Zhang, J. (2009). Dynamic visualization of signal transduction in living cells: from second messengers to kinases. *IUBMB Life* 61, 902–908. doi: 10.1002/iub.232
- Hjerrild, M., Stensballe, A., Rasmussen, T. E., Kofoed, C. B., Blom, N., Sicheritz-Ponten, T., et al. (2004). Identification of phosphorylation sites in protein kinase A substrates using artificial neural networks and mass spectrometry. *J. Proteome Res.* 3, 426–433. doi: 10.1021/pr0341033
- Ho, M. S., Tsai, P. I., and Chien, C. T. (2006). F-box proteins: the key to protein degradation. *J. Biomed. Sci.* 13, 181–191. doi: 10.1007/s11373-005-9058-2
- Honigsmann, A., Mueller, V., Hell, S. W., and Eggeling, C. (2013). STED microscopy detects and quantifies liquid phase separation in lipid membranes using a new far-red emitting fluorescent phosphoglycerolipid analogue. *Faraday Discuss.* 161, 77–89. doi: 10.1039/c2fd20107k
- Hu, J., Rho, H. S., Newman, R. H., Hwang, W., Neiswinger, J., Zhu, H., et al. (2014). Global analysis of phosphorylation networks in humans. *Biochim. Biophys. Acta* 1844, 224–231. doi: 10.1016/j.bbapap.2013.03.009
- Hu, S., Xie, Z., Onishi, A., Yu, X., Jiang, L., Lin, J., et al. (2009). Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell* 139, 610–622. doi: 10.1016/j.cell.2009.08.037
- Hu, S., Xie, Z., Qian, J., Blackshaw, S., and Zhu, H. (2011). Functional protein microarray technology. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 3, 255–268. doi: 10.1002/wsbm.118
- Huang, S. Y., Tsai, M. L., Chen, G. Y., Wu, C. J., and Chen, S. H. (2007). A systematic MS-based approach for identifying in vitro substrates of PKA and PKG in rat uteri. *J. Proteome Res.* 6, 2674–2684. doi: 10.1021/pr070134c
- Huang, Z. X., Zhao, P., Zeng, G. S., Wang, Y. M., Sudbery, I., and Wang, Y. (2014). Phosphoregulation of Nap1 plays a role in septin ring dynamics and morphogenesis in *Candida albicans*. *MBio* 5:e00915-13. doi: 10.1128/mBio.00915-13
- Hunter, T. (1996). Tyrosine phosphorylation: past, present and future. *Biochem. Soc. Trans.* 24, 307–327.
- Hunter, T. (2012). Why nature chose phosphate to modify proteins. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 367, 2513–2516. doi: 10.1098/rstb.2012.0013
- Jeong, J. S., Jiang, L., Albino, E., Marrero, J., Rho, H. S., Hu, J., et al. (2012). Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. *Mol. Cell Proteomics* 11:O111. doi: 10.1074/mcp.O111.016253
- Jia, W., Andaya, A., and Leary, J. A. (2012). Novel mass spectrometric method for phosphorylation quantification using cerium oxide nanoparticles and tandem mass tags. *Anal. Chem.* 84, 2466–2473. doi: 10.1021/ac203248s
- Jin, J., and Pawson, T. (2012). Modular evolution of phosphorylation-based signalling systems. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 367, 2540–2555. doi: 10.1098/rstb.2012.0106
- Johnson, H., and White, F. M. (2012). Toward quantitative phosphotyrosine profiling in vivo. *Semin. Cell Dev. Biol.* 23, 854–862. doi: 10.1016/j.semcdb.2012.05.008
- Kettenbach, A. N., Wang, T., Faherty, B. K., Madden, D. R., Knapp, S., Bailey-Kellogg, C., et al. (2012). Rapid determination of multiple linear kinase substrate motifs by mass spectrometry. *Chem. Biol.* 19, 608–618. doi: 10.1016/j.chembiol.2012.04.011
- Khan, P., Idrees, D., Moxley, M. A., Corbett, J. A., Ahmad, F., von Figura, G., et al. (2014). Luminol-based chemiluminescent signals: clinical and non-clinical application and future uses. *Appl. Biochem. Biotechnol.* 173, 333–355. doi: 10.1007/s12010-014-0850-1
- Kim, J. H., Lee, J., Oh, B., Kimm, K., and Koh, I. (2004). Prediction of phosphorylation sites using SVMs. *Bioinformatics* 20, 3179–3184. doi: 10.1093/bioinformatics/bth382
- Kinoshita, E., Kinoshita-Kikuta, E., and Koike, T. (2012b). Phos-tag SDS-PAGE systems for phosphorylation profiling of proteins with a wide range of molecular masses under neutral pH conditions. *Proteomics* 12, 192–202. doi: 10.1002/pmic.201100524
- Kinoshita, E., Kinoshita-Kikuta, E., Matsubara, M., Aoki, Y., Ohie, S., Mouri, Y., et al. (2009). Two-dimensional phosphate-affinity gel electrophoresis for the analysis of phosphoprotein isotypes. *Electrophoresis* 30, 550–559. doi: 10.1002/elps.200800386
- Kinoshita, E., Kinoshita-Kikuta, E., Sugiyama, Y., Fukada, Y., Ozeki, T., and Koike, T. (2012a). Highly sensitive detection of protein phosphorylation by using improved Phos-tag Biotin. *Proteomics* 12, 932–937. doi: 10.1002/pmic.201100639
- Kinoshita, E., Takahashi, M., Takeda, H., Shiro, M., and Koike, T. (2004). Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc(II) complex. *Dalton Trans.* 1189–1193. doi: 10.1039/b400269e
- Kinoshita-Kikuta, E., Kinoshita, E., and Koike, T. (2012). Separation and identification of four distinct serine-phosphorylation states of ovalbumin by Phos-tag affinity electrophoresis. *Electrophoresis* 33, 849–855. doi: 10.1002/elps.201100518
- Klumpp, S., and Krieglstein, J. (2002). Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur. J. Biochem.* 269, 1067–1071. doi: 10.1046/j.1432-1033.2002.02755.x
- Knight, J. D., Tian, R., Lee, R. E., Wang, F., Beauvais, A., Zou, H., et al. (2012). A novel whole-cell lysate kinase assay identifies substrates of the p38 MAPK in differentiating myoblasts. *Skelet. Muscle* 2:5. doi: 10.1186/2044-5040-2-5
- Koch, A., and Hauf, S. (2010). Strategies for the identification of kinase substrates using analog-sensitive kinases. *Eur. J. Cell Biol.* 89, 184–193. doi: 10.1016/j.ejcb.2009.11.024
- Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y., et al. (2011). Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* 22, 4647–4656. doi: 10.1091/mbc.E11-01-0072
- Kooij, V., Venkatraman, V., Tra, J., Kirk, J., Rowell, J., Blice-Baum, A., et al. (2014). Sizing up models of heart failure: Proteomics from flies to humans. *Proteomics Clin. Appl.* doi: 10.1002/prca.201300123. [Epub ahead of print].
- Kosako, H., and Nagano, K. (2011). Quantitative phosphoproteomics strategies for understanding protein kinase-mediated signal transduction pathways. *Expert. Rev. Proteomics* 8, 81–94. doi: 10.1586/epr.10.104
- Kozuka-Hata, H., Nasu-Nishimura, Y., Koyama-Nasu, R., Ao-Kondo, H., Tsumoto, K., Akiyama, T., et al. (2012). Phosphoproteome of human glioblastoma initiating cells reveals novel signaling regulators encoded by the transcriptome. *PLoS ONE* 7:e43398. doi: 10.1371/journal.pone.0043398
- Kumar, N., Wolf-Yadlin, A., White, F. M., and Lauffenburger, D. A. (2007). Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput. Biol.* 3:e4. doi: 10.1371/journal.pcbi.0030004
- Kunkel, M. T., and Newton, A. C. (2009). Spatiotemporal dynamics of kinase signaling visualized by targeted reporters. *Curr. Protoc. Chem. Biol.* 1, 17–18. doi: 10.1002/9780470559277.ch090106
- Kunkel, M. T., and Newton, A. C. (2014). Imaging kinase activity at protein scaffolds. *Methods Mol. Biol.* 1071, 129–137. doi: 10.1007/978-1-62703-622-1\_10
- Lad, C., Williams, N. H., and Wolfenden, R. (2003). The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5607–5610. doi: 10.1073/pnas.0631607100
- Lam, A. J., St. Pierre, F., Gong, Y., Marshall, J. D., Cranfill, P. J., Baird, M. A., et al. (2012). Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods* 9, 1005–1012. doi: 10.1038/nmeth.2171
- Leung, G. C., Murphy, J. M., Briant, D., and Sicheri, F. (2009). Characterization of kinase target phosphorylation consensus motifs using peptide SPOT arrays. *Methods Mol. Biol.* 570, 187–195. doi: 10.1007/978-1-60327-394-7\_7
- Liao, H., Byeon, I. J., and Tsai, M. D. (1999). Structure and function of a new phosphopeptide-binding domain containing the FHA2 of Rad53. *J. Mol. Biol.* 294, 1041–1049. doi: 10.1006/jmbi.1999.3313
- Linding, R., Jensen, L. J., Ostheimer, G. J., van Vugt, M. A., Jorgensen, C., Miron, I. M., et al. (2007). Systematic discovery of in vivo phosphorylation networks. *Cell* 129, 1415–1426. doi: 10.1016/j.cell.2007.05.052

- Linding, R., Jensen, L. J., Pasculescu, A., Olhovskiy, M., Colwill, K., Bork, P., et al. (2008). NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* 36, D695–D699. doi: 10.1093/nar/gkm902
- Liu, B. A., Jablonowski, K., Raina, M., Arce, M., Pawson, T., and Nash, P. D. (2006). The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. *Mol. Cell* 22, 851–868. doi: 10.1016/j.molcel.2006.06.001
- Mackenzie, R. W., and Elliott, B. T. (2014). Akt/PKB activation and insulin signaling: a novel insulin signaling pathway in the treatment of type 2 diabetes. *Diabetes Metab. Syndr. Obes.* 7, 55–64. doi: 10.2147/DMSO.S48260
- Marfori, M., Mynott, A., Ellis, J. J., Mehdi, A. M., Saunders, N. F., Curmi, P. M., et al. (2011). Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim. Biophys. Acta* 1813, 1562–1577. doi: 10.1016/j.bbamcr.2010.10.013
- Medina-Cleghorn, D., and Nomura, D. K. (2013). Chemical approaches to study metabolic networks. *Pflugers Arch.* 465, 427–440. doi: 10.1007/s00424-012-1201-0
- Meharena, H. S., Chang, P., Keshwani, M. M., Oruganty, K., Nene, A. K., Kannan, N., et al. (2013). Deciphering the structural basis of eukaryotic protein kinase regulation. *PLoS Biol.* 11:e1001680. doi: 10.1371/journal.pbio.1001680
- Mehta, S., and Zhang, J. (2011). Reporting from the field: genetically encoded fluorescent reporters uncover signaling dynamics in living biological systems. *Annu. Rev. Biochem.* 80, 375–401. doi: 10.1146/annurev-biochem-060409-093259
- Mehta, S., and Zhang, J. (2014). Using a genetically encoded FRET-based reporter to visualize calcineurin phosphatase activity in living cells. *Methods Mol. Biol.* 1071, 139–149. doi: 10.1007/978-1-62703-622-1\_11
- Meissner, G. (2010). Regulation of ryanodine receptor ion channels through post-translational modifications. *Curr. Top. Membr.* 66, 91–113. doi: 10.1016/S1063-5823(10)66005-X
- Mertins, P., Yang, F., Liu, T., Mani, D. R., Petyuk, V. A., Gillette, M. A., et al. (2014). Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Mol. Cell Proteomics* 13, 1690–1704. doi: 10.1074/mcp.M113.036392
- Minden, J. S. (2012). DIGE: past and future. *Methods Mol. Biol.* 854, 3–8. doi: 10.1007/978-1-61779-573-2\_1
- Mok, J., Im, H., and Snyder, M. (2009). Global identification of protein kinase substrates by protein microarray analysis. *Nat. Protoc.* 4, 1820–1827. doi: 10.1038/nprot.2009.194
- Mok, J., Kim, P. M., Lam, H. Y., Piccirillo, S., Zhou, X., Jeschke, G. R., et al. (2010). Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci. Signal.* 3, ra12. doi: 10.1126/scisignal.2000482
- Mok, J., Zhu, X., and Snyder, M. (2011). Dissecting phosphorylation networks: lessons learned from yeast. *Expert. Rev. Proteomics* 8, 775–786. doi: 10.1586/epr.11.64
- Mujumdar, R. B., Ernst, L. A., Mujumdar, S. R., Lewis, C. J., and Waggoner, A. S. (1993). Cyanine dye labeling reagents: sulfoindocyanine succinimidyl esters. *Bioconjug. Chem.* 4, 105–111. doi: 10.1021/bc00020a001
- Mukai, N., Nakanishi, T., Shimizu, A., Takubo, T., and Ikeda, T. (2008). Identification of phosphotyrosyl proteins in vitreous humours of patients with vitreoretinal diseases by sodium dodecyl sulphate-polyacrylamide gel electrophoresis/Western blotting/matrix-assisted laser desorption time-of-flight mass spectrometry. *Ann. Clin. Biochem.* 45, 307–312. doi: 10.1258/acb.2007.007151
- Na, S., and Paek, E. (2014). Software eyes for protein post-translational modifications. *Mass Spectrom. Rev.* doi: 10.1002/mas.21425. [Epub ahead of print].
- Nabetani, T., Kim, Y. J., Watanabe, M., Ohashi, Y., Kamiguchi, H., and Hirabayashi, Y. (2009). Improved method of phosphopeptides enrichment using biphasic phosphate-binding tag/C18 tip for versatile analysis of phosphorylation dynamics. *Proteomics* 9, 5525–5533. doi: 10.1002/pmic.200900341
- Neininger, A., Thielemann, H., and Gaestel, M. (2001). FRET-based detection of different conformations of MK2. *EMBO Rep.* 2, 703–708. doi: 10.1093/embo-reports/kve157
- Neves, S. R., Tsokas, P., Sarkar, A., Grace, E. A., Rangamani, P., Taubenfeld, S. M., et al. (2008). Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks. *Cell* 133, 666–680. doi: 10.1016/j.cell.2008.04.025
- Newman, R. H., Fosbrink, M. D., and Zhang, J. (2011). Genetically encodable fluorescent biosensors for tracking signaling dynamics in living cells. *Chem. Rev.* 111, 3614–3666. doi: 10.1021/cr100002u
- Newman, R. H., Hu, J., Rho, H. S., Xie, Z., Woodard, C., Neiswinger, J., et al. (2013). Construction of human activity-based phosphorylation networks. *Mol. Syst. Biol.* 9, 655. doi: 10.1038/msb.2013.12
- Newman, R. H., and Zhang, J. (2008). Visualization of phosphatase activity in living cells with a FRET-based calcineurin activity sensor. *Mol. Biosyst.* 4, 496–501. doi: 10.1039/b720034j
- Ni, Q., Ganesan, A., Aye-Han, N. N., Gao, X., Allen, M. D., Levchenko, A., et al. (2011). Signaling diversity of PKA achieved via a Ca<sup>2+</sup>-cAMP-PKA oscillatory circuit. *Nat. Chem. Biol.* 7, 34–40. doi: 10.1038/nchembio.478
- Niggli, E., Ullrich, N. D., Gutierrez, D., Kyrychenko, S., Polakova, E., and Shirokova, N. (2013). Posttranslational modifications of cardiac ryanodine receptors: Ca(2+) signaling and EC-coupling. *Biochim. Biophys. Acta* 1833, 866–875. doi: 10.1016/j.bbamcr.2012.08.016
- Nilsson, C. L. (2012). Advances in quantitative phosphoproteomics. *Anal. Chem.* 84, 735–746. doi: 10.1021/ac202877y
- Okamura, H., Aramburu, J., Garcia-Rodriguez, C., Viola, J. P., Raghavan, A., Tahliliani, M., et al. (2000). Concerted dephosphorylation of the transcription factor NFAT1 induces a conformational switch that regulates transcriptional activity. *Mol. Cell* 6, 539–550. doi: 10.1016/S1097-2765(00)00053-8
- Olsen, J. V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., et al. (2006). Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648. doi: 10.1016/j.cell.2006.09.026
- Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* 4, 709–712. doi: 10.1038/nmeth1060
- Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., et al. (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* 3, ra3. doi: 10.1126/scisignal.2000475
- Ong, S. E. (2012). The expanding field of SILAC. *Anal. Bioanal. Chem.* 404, 967–976. doi: 10.1007/s00216-012-5998-3
- Ortsater, H., Grankvist, N., Honkanen, R. E., and Sjöholm, A. (2014). Protein phosphatases in pancreatic islets. *J. Endocrinol.* 221, R121–R144. doi: 10.1530/JOE-14-0002
- Oyama, M., Kozuka-Hata, H., Tasaki, S., Semba, K., Hattori, S., Sugano, S., et al. (2009). Temporal perturbation of tyrosine phosphoproteome dynamics reveals the system-wide regulatory networks. *Mol. Cell Proteomics* 8, 226–231. doi: 10.1074/mcp.M800186-MCP200
- Padilla-Parra, S., and Tramier, M. (2012). FRET microscopy in the living cell: different approaches, strengths and weaknesses. *Bioessays* 34, 369–376. doi: 10.1002/bies.201100086
- Pastrello, C., Pasini, E., Kotlyar, M., Otasek, D., Wong, S., Sangrar, W., et al. (2014). Integration, visualization and analysis of human interactome. *Biochem. Biophys. Res. Commun.* 445, 757–773. doi: 10.1016/j.bbrc.2014.01.151
- Pellegrin, S., Heesom, K. J., Satchwell, T. J., Hawley, B. R., Daniels, G., van den Akker, E., et al. (2012). Differential proteomic analysis of human erythroblasts undergoing apoptosis induced by epo-withdrawal. *PLoS ONE* 7:e38356. doi: 10.1371/journal.pone.0038356
- Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567. doi: 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2
- Persson, F., Barkefors, I., and Elf, J. (2013). Single molecule methods with applications in living cells. *Curr. Opin. Biotechnol.* 24, 737–744. doi: 10.1016/j.copbio.2013.03.013
- Pierobon, M., Wulfkühle, J., Liotta, L., and Petricoin, E. (2014). Application of molecular technologies for phosphoproteomic analysis of clinical samples. *Oncogene*. doi: 10.1038/onc.2014.16. [Epub ahead of print].
- Popescu, S. C., Popescu, G. V., Bachan, S., Zhang, Z., Gerstein, M., Snyder, M., et al. (2009). MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Dev.* 23, 80–92. doi: 10.1101/gad.174009
- Posewitz, M. C., and Tempst, P. (1999). Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* 71, 2883–2892. doi: 10.1021/ac981409y

- Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., et al. (2005). Global analysis of protein phosphorylation in yeast. *Nature* 438, 679–684. doi: 10.1038/nature04187
- Puig, O., Caspari, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., et al. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24, 218–229. doi: 10.1006/meth.2001.1183
- Reinhardt, H. C., and Yaffe, M. B. (2013). Phospho-Ser/Thr-binding domains: navigating the cell cycle and DNA damage response. *Nat. Rev. Mol. Cell Biol.* 14, 563–580. doi: 10.1038/nrm3640
- Rigbolt, K. T., and Blagoev, B. (2012). Quantitative phosphoproteomics to characterize signaling networks. *Semin. Cell Dev. Biol.* 23, 863–871. doi: 10.1016/j.semcdb.2012.05.006
- Ritt, M., Guan, J. L., and Sivaramakrishnan, S. (2013). Visualizing and manipulating focal adhesion kinase regulation in live cells. *J. Biol. Chem.* 288, 8875–8886. doi: 10.1074/jbc.M112.421164
- Roura, S., Galvez-Monton, C., and Bayes-Genis, A. (2013). Bioluminescence imaging: a shining future for cardiac regeneration. *J. Cell Mol. Med.* 17, 693–703. doi: 10.1111/jcmm.12018
- Ruprecht, B., and Lemeer, S. (2014). Proteomic analysis of phosphorylation in cancer. *Expert Rev. Proteomics* 11, 259–267. doi: 10.1586/14789450.2014.901156
- Sample, V., Newman, R. H., and Zhang, J. (2009). The structure and function of fluorescent proteins. *Chem. Soc. Rev.* 38, 2852–2864. doi: 10.1039/b913033k
- Saucerman, J. J., Zhang, J., Martin, J. C., Peng, L. X., Stenbit, A. E., Tsien, R. Y., et al. (2006). Systems analysis of PKA-mediated phosphorylation gradients in live cardiac myocytes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12923–12928. doi: 10.1073/pnas.0600137103
- Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., et al. (2011). Confident phosphorylation site localization using the Mascot Delta Score. *Mol. Cell Proteomics* 10:M110. doi: 10.1074/mcp.M110.003830
- Sciarretta, S., Volpe, M., and Sadoshima, J. (2014). Mammalian target of rapamycin signaling in cardiac physiology and disease. *Circ. Res.* 114, 549–564. doi: 10.1161/CIRCRESAHA.114.302022
- Shekhawat, S. S., and Ghosh, I. (2011). Split-protein systems: beyond binary protein-protein interactions. *Curr. Opin. Chem. Biol.* 15, 789–797. doi: 10.1016/j.cbpa.2011.10.014
- Shirai, A., Matsuyama, A., Yashiroda, Y., Hashimoto, A., Kawamura, Y., Arai, R., et al. (2008). Global analysis of gel mobility of proteins and its use in target identification. *J. Biol. Chem.* 283, 10745–10752. doi: 10.1074/jbc.M709211200
- Sipiet, F., Vandame, P., Spriet, C., Leray, A., Vincent, P., Trinell, D., et al. (2013). From FRET imaging to practical methodology for kinase activity sensing in living cells. *Prog. Mol. Biol. Transl. Sci.* 113, 145–216. doi: 10.1016/B978-0-12-386932-6.00005-3
- Skrzypski, M. (2008). Quantitative reverse transcriptase real-time polymerase chain reaction (qRT-PCR) in translational oncology: lung cancer perspective. *Lung Cancer* 59, 147–154. doi: 10.1016/j.lungcan.2007.11.008
- Song, C., Ye, M., Liu, Z., Cheng, H., Jiang, X., Han, G., et al. (2012). Systematic analysis of protein phosphorylation networks from phosphoproteomic data. *Mol. Cell Proteomics* 11, 1070–1083. doi: 10.1074/mcp.M111.012625
- Song, Q., Saucerman, J. J., Bossuyt, J., and Bers, D. M. (2008). Differential integration of Ca<sup>2+</sup>-calmodulin signal in intact ventricular myocytes at low and high affinity Ca<sup>2+</sup>-calmodulin targets. *J. Biol. Chem.* 283, 31531–31540. doi: 10.1074/jbc.M804902200
- Stasyk, T., Schiefermeier, N., Skvortsov, S., Zwierzina, H., Peranen, J., Bonn, G. K., et al. (2007). Identification of endosomal epidermal growth factor receptor signaling targets by functional organelle proteomics. *Mol. Cell Proteomics* 6, 908–922. doi: 10.1074/mcp.M600463-MCP200
- Stynen, B., Tournu, H., Tavernier, J., and Van Dijck, P. (2012). Diversity in genetic *in vivo* methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. *Microbiol. Mol. Biol. Rev.* 76, 331–382. doi: 10.1128/MMBR.05021-11
- Sun, Y., Rombola, C., Jyothikumar, V., and Periasamy, A. (2013). Forster resonance energy transfer microscopy and spectroscopy for localizing protein-protein interactions in living cells. *Cytometry A* 83, 780–793. doi: 10.1002/cyto.a.22321
- Sutandy, F. X., Qian, J., Chen, C. S., and Zhu, H. (2013). Overview of protein microarrays. *Curr. Protoc. Protein Sci.* Chapter 27:Unit 27.1. doi: 10.1002/0471140864.ps2701s72
- Syka, J. E., Marto, J. A., Bai, D. L., Horning, S., Senko, M. W., Schwartz, J. C., et al. (2004). Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.* 3, 621–626. doi: 10.1021/pr0499794
- Tasaki, S., Nagasaki, M., Oyama, M., Hata, H., Ueno, K., Yoshida, R., et al. (2006). Modeling and estimation of dynamic EGFR pathway by data assimilation approach using time series proteomic data. *Genome Inform.* 17, 226–238.
- Taylor, S. S., and Kornev, A. P. (2011). Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* 36, 65–77. doi: 10.1016/j.tibs.2010.09.006
- Tian, T., and Song, J. (2012). Mathematical modelling of the MAP kinase pathway using proteomic datasets. *PLoS ONE* 7:e42230. doi: 10.1371/journal.pone.0042230
- Tsou, P., Zheng, B., Hsu, C. H., Sasaki, A. T., and Cantley, L. C. (2011). A fluorescent reporter of AMPK activity and cellular energy stress. *Cell Metab.* 13, 476–486. doi: 10.1016/j.cmet.2011.03.006
- Tsunehiro, M., Meki, Y., Matsuoka, K., Kinoshita-Kikuta, E., Kinoshita, E., and Koike, T. (2013). A Phos-tag-based magnetic-bead method for rapid and selective separation of phosphorylated biomolecules. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 925, 86–94. doi: 10.1016/j.jchromb.2013.02.039
- Ummanni, R., Mannsperger, H. A., Sonntag, J., Oswald, M., Sharma, A. K., König, R., et al. (2014). Evaluation of reverse phase protein array (RPPA)-based pathway-activation profiling in 84 non-small cell lung cancer (NSCLC) cell lines as platform for cancer proteomics and biomarker discovery. *Biochim. Biophys. Acta* 1844, 950–959. doi: 10.1016/j.bbapap.2013.11.017
- van Rossum, T., Kengen, S. W., and van der Oost, J. (2013). Reporter-based screening and selection of enzymes. *FEBS J.* 280, 2979–2996. doi: 10.1111/febs.12281
- Violin, J. D., DiPilato, L. M., Yildirim, N., Elston, T. C., Zhang, J., and Lefkowitz, R. J. (2008). beta2-adrenergic receptor signaling and desensitization elucidated by quantitative modeling of real time cAMP dynamics. *J. Biol. Chem.* 283, 2949–2961. doi: 10.1074/jbc.M707009200
- Wiechert, W., Schweissgut, O., Takanaga, H., and Frommer, W. B. (2007). Fluxomics: mass spectrometry versus quantitative imaging. *Curr. Opin. Plant Biol.* 10, 323–330. doi: 10.1016/j.pbi.2007.04.015
- Wilkins, M. R. (2009). Hares and tortoises: the high- versus low-throughput proteomic race. *Electrophoresis* 30(Suppl. 1), S150–S155. doi: 10.1002/elps.200900175
- Woehler, A. (2013). Simultaneous quantitative live cell imaging of multiple FRET-based biosensors. *PLoS ONE* 8:e61096. doi: 10.1371/journal.pone.0061096
- Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S., Zaman, M., Kim, H. D., et al. (2006). Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Mol. Syst. Biol.* 2, 54. doi: 10.1038/msb4100094
- Wong, Y. H., Lee, T. Y., Liang, H. K., Huang, C. M., Wang, T. Y., Yang, Y. H., et al. (2007). KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Res.* 35, W588–W594. doi: 10.1093/nar/gkm322
- Woodard, C. L., Goodwin, C. R., Wan, J., Xia, S., Newman, R., Hu, J., et al. (2013). Profiling the dynamics of a human phosphoproteome reveals new components in HGF/c-Met signaling. *PLoS ONE* 8:e72671. doi: 10.1371/journal.pone.0072671
- Wu, M., Yang, X., and Chan, C. (2009). A dynamic analysis of IRS-PKR signaling in liver cells: a discrete modeling approach. *PLoS ONE* 4:e8040. doi: 10.1371/journal.pone.0008040
- Xue, Y., Li, A., Wang, L., Feng, H., and Yao, X. (2006). PPSP: prediction of PK-specific phosphorylation site with Bayesian decision theory. *BMC Bioinformatics* 7:163. doi: 10.1186/1471-2105-7-163
- Xue, Y., Zhou, F., Zhu, M., Ahmed, K., Chen, G., and Yao, X. (2005). GPS: a comprehensive web server for phosphorylation sites prediction. *Nucleic Acids Res.* 33, W184–W187. doi: 10.1093/nar/gki393
- Yaffe, M. B. (2002). Phosphotyrosine-binding domains in signal transduction. *Nat. Rev. Mol. Cell Biol.* 3, 177–186. doi: 10.1038/nrm759
- Yaffe, M. B., Leparc, G. G., Lai, J., Obata, T., Volinia, S., and Cantley, L. C. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nat. Biotechnol.* 19, 348–353. doi: 10.1038/86737
- Yang, C., Zhong, X., and Li, L. (2014). Recent advances in enrichment and separation strategies for mass spectrometry-based phosphoproteomics. *Electrophoresis*. doi: 10.1002/elps.201400017. [Epub ahead of print].

- Yip, Y. Y., Yeap, Y. Y., Bogoyevitch, M. A., and Ng, D. C. (2014). cAMP-dependent protein kinase and c-Jun N-terminal kinase mediate stathmin phosphorylation for the maintenance of interphase microtubules during osmotic stress. *J. Biol. Chem.* 289, 2157–2169. doi: 10.1074/jbc.M113.470682
- Yoshizaki, H., Aoki, K., Nakamura, T., and Matsuda, M. (2006). Regulation of RalA GTPase by phosphatidylinositol 3-kinase as visualized by FRET probes. *Biochem. Soc. Trans.* 34, 851–854. doi: 10.1042/BST0340851
- Zadran, S., Standley, S., Wong, K., Otiniano, E., Amighi, A., and Baudry, M. (2012). Fluorescence resonance energy transfer (FRET)-based biosensors: visualizing cellular dynamics and bioenergetics. *Appl. Microbiol. Biotechnol.* 96, 895–902. doi: 10.1007/s00253-012-4449-6
- Zapata-Hommer, O., and Griesbeck, O. (2003). Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP. *BMC Biotechnol.* 3:5. doi: 10.1186/1472-6750-3-5
- Zhang, C., Lopez, M. S., Dar, A. C., Ladow, E., Finkbeiner, S., Yun, C. H., et al. (2013). Structure-guided inhibitor design expands the scope of analog-sensitive kinase technology. *ACS Chem. Biol.* 8, 1931–1938. doi: 10.1021/cb400376p
- Zhang, H., Zha, X., Tan, Y., Hornbeck, P. V., Mastrangelo, A. J., Alessi, D. R., et al. (2002). Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. *J. Biol. Chem.* 277, 39379–39387. doi: 10.1074/jbc.M206399200
- Zhang, J., and Allen, M. D. (2007). FRET-based biosensors for protein kinases: illuminating the kinome. *Mol. Biosyst.* 3, 759–765. doi: 10.1039/b706628g
- Zhang, J., Hupfeld, C. J., Taylor, S. S., Olefsky, J. M., and Tsien, R. Y. (2005a). Insulin disrupts beta-adrenergic signalling to protein kinase A in adipocytes. *Nature* 437, 569–573. doi: 10.1038/nature04140
- Zhang, J., Ma, Y., Taylor, S. S., and Tsien, R. Y. (2001). Genetically encoded reporters of protein kinase A activity reveal impact of substrate tethering. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14997–15002. doi: 10.1073/pnas.211566798
- Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J., Rush, J., Lauffenburger, D. A., et al. (2005b). Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell Proteomics* 4, 1240–1250. doi: 10.1074/mcp.M500089-MCP200
- Zhu, H., Cox, E., and Qian, J. (2012). Functional protein microarray as molecular decathlete: a versatile player in clinical proteomics. *Proteomics Clin. Appl.* 6, 548–562. doi: 10.1002/prca.201200041
- Zubarev, R. A., Horn, D. M., Fridriksson, E. K., Kelleher, N. L., Kruger, N. A., Lewis, M. A., et al. (2000). Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* 72, 563–573. doi: 10.1021/ac990811p

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 May 2014; accepted: 16 July 2014; published online: 15 August 2014.

Citation: Newman RH, Zhang J and Zhu H (2014) Toward a systems-level view of dynamic phosphorylation networks. *Front. Genet.* 5:263. doi: 10.3389/fgene.2014.00263

This article was submitted to *Systems Biology*, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Newman, Zhang and Zhu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Exploiting holistic approaches to model specificity in protein phosphorylation

Antonio Palmeri\*, Fabrizio Ferrè and Manuela Helmer-Citterich

Department of Biology, Centre for Molecular Bioinformatics, University of Rome Tor Vergata, Rome, Italy

## Edited by:

Andreas Zanzoni, Inserm TAGC  
UMR1090, France  
Allegra Via, Sapienza University, Italy

## Reviewed by:

Caroline Evans, University of  
Sheffield, UK  
Florian Gnad, Genentech, USA

## \*Correspondence:

Antonio Palmeri, Department of  
Biology, Centre for Molecular  
Bioinformatics, University of Rome  
Tor Vergata, Via della Ricerca  
Scientifica snc, 00133 Rome, Italy  
e-mail: antonio.palmeri@uniroma2.it

Phosphate plays a chemically unique role in shaping cellular signaling of all current living systems, especially eukaryotes. Protein phosphorylation has been studied at several levels, from the near-site context, both in sequence and structure, to the crowded cellular environment, and ultimately to the systems-level perspective. Despite the tremendous advances in mass spectrometry and efforts dedicated to the development of *ad hoc* highly sophisticated methods, phosphorylation site inference and associated kinase identification are still unresolved problems in kinome biology. The sequence and structure of the substrate near-site context are not sufficient alone to model the *in vivo* phosphorylation rules, and they should be integrated with orthogonal information in all possible applications. Here we provide an overview of the different contexts that contribute to protein phosphorylation, discussing their potential impact in phosphorylation site annotation and in predicting kinase-substrate specificity.

**Keywords: kinase-substrate specificity, phosphorylation context, phosphorylation prediction, cellular signaling, kinase-peptide specificity, substrate recruitment, signaling networks**

## INTRODUCTION

Phosphorylation, the enzymatic reaction resulting in the addition of a phosphate group to several types of residues, which in eukaryotes are mainly serines, threonines, or tyrosines, generates *de facto* a new side chain whose physico-chemical properties are different from those of the unmodified residues. This mechanism of Post-Translational Modification (PTM) is strikingly common throughout evolution and in particular for eukaryotes where it is involved in a myriad of cellular processes (Manning et al., 2002a,b, 2008, 2011; Caenepeel et al., 2004; Bradham et al., 2006).

The chemical properties of phosphate make this group a perfect candidate for protein modification, and allow its broad use as a molecular switch within the cell (Hunter, 2012). Indeed the hydrolytic stability of phosphate esters (for instance phosphoserine, phosphotyrosine, phosphothreonine, etc.) in aqueous solutions at pH7 allows the cell to minimize the noise in signal transduction due to non-enzymatically catalyzed hydrolyses. In addition, phosphate monoesters act as sensors, as their electric charge can be influenced by the chemical environment. Lastly, phosphate is a largely available molecule, as it is abundant on Earth and particularly within the cell, where it is included in a fundamental energy storage molecule, i.e., ATP. Differently from other types of PTMs, only one group can be enzymatically added to one residue, underlining the peculiar binary nature of this protein modification. The modified residue can undergo inter- or intra-molecular interactions, causing changes to the protein structure or interfering with its function, probably the most famous and complex example being the allosteric regulation of glycogen phosphorylase (Barford et al., 1991). Additional mechanisms for phosphorylation-mediated modulation have also been reported, such as for instance the inhibition of a binding site (Hurley et al.,

1990). A beautiful electrostatic-based tuning of protein function mediated by phosphorylation has been described in yeast cell-cycle regulation, where the membrane localization of the MAPKs scaffold protein Ste5 is disrupted by phosphorylation of a cluster of sites flanking a basic membrane binding motif (Strickfaden et al., 2007).

However, the reason for the success of this type of PTM during evolution, at least in eukaryotes, has to be found largely in its ability to be edited and recognized selectively by specific protein domains, thus providing an efficient tool for transient molecular recognition in the context of signal transduction networks (Lim and Pawson, 2010).

With PTM-based proteomics, phosphorylation sites, as well as other PTMs, are identified and stored in large-scale datasets (Olsen and Mann, 2013). As a consequence of this explosion of data, there is great demand for functional annotation studies that largely exceeds what current technology offers. Furthermore, some observations question the functionality of a substantial fraction of these sites (Landry et al., 2009; Moses and Landry, 2010; Levy et al., 2012; Tan and Bader, 2012).

Given the difficulties in the experimental annotation of the kinase responsible for the phosphorylation, many attempts have been made to computationally model cellular signaling events. Some of the published reviews examine the field of kinase specificity from a more biological perspective, discussing the protein kinase specificity rules in sequence and in structure, while some others compare the different tools, and the techniques used to model kinase-substrate interaction and in general those used to build phosphorylation site predictors (Zhu et al., 2005; Ubersax and Ferrell, 2007; Miller and Blom, 2009; Xue et al., 2010; Trost and Kusalik, 2011; Via et al., 2011). Here we will focus on kinase-substrate interaction at the kinase domain and



the substrate-peptide level, and then we will summarize the contextual information that could help to better understand the molecular determinants of kinase specificity, contributing also to boost the performances of phosphorylation site predictors.

## INFERRING KINASES RESPONSIBLE FOR PHOSPHORYLATIONS *IN SILICO*

While recent advancements in phosphoproteomics allow the identification of phosphosites from entire proteomes with ever increasing reliability and higher coverage, no high-throughput method is able to pinpoint which kinases are responsible for phosphorylating which protein substrates. Therefore, in a high-throughput context, only *in silico* approaches can effectively help in reconstructing molecular signaling circuits.

All the methods can be grouped according to different criteria, but arguably the main differences are between motif- or PSSM-based and machine learning-based methods and in the use of evolutionary information. We select seven major aspects, as exemplars of different methodologies that have been developed, namely: motif-based identification of phosphorylation sites, structural information integration, integration of phosphorylation site structural context, phospho-clusters modeling, integration of Protein-Protein Interaction Network (PPIN) information and multi-organisms prediction. For a complete list of currently available methods, see **Table 1**.

The first method to predict the specific kinases that are responsible for the phosphorylations is Scansite (Yaffe et al., 2001), developed by Yaffe and colleagues, using Position Specific Scoring Matrices (PSSMs) for 62 different kinase phosphorylation motifs. Following an extensive analysis of the PKA motifs, PkaPS (Neuberger et al., 2007) was developed, exclusively suited for the prediction of protein kinase A-specific phosphorylation sites. Taking advantage of the structural information, Kobe and his collaborators developed PrediKin (Brinkworth et al., 2003; Ellis and Kobe, 2011), which is based on the analysis of the contact positions between kinases and substrates in proteins of known structure. The authors were able to associate the identification of specific kinase residues with a corresponding preference in the sequence of the substrate. PrediKin outperformed other three predictors in the DREAM4 challenge, whose goal was to predict peptide recognition domain specificity in protein kinases. In another work the information about the 3d-context of phosphorylation sites has been directly integrated in kinase-specific predictions, defining 3d-signatures motifs, even if the improvement with respect to sequence information is small (Durek et al., 2009). Conservation-based methods for predicting kinase-substrates usually assume that phosphorylation sites should be positionally conserved in Multiple Sequence Alignments (MSA) of orthologs (Budovskaya et al., 2005; Gnad et al., 2011). However, it was observed that phospho-motifs may also be found in different positions of the same local regions of orthologous proteins (Moses et al., 2007). In these cases only the local density of phosphorylation sites, but not their exact position, is conserved across orthologs. Lai et al. designed a method, ConDens, which computes the probability of observing a number of matches to a kinase motif in a MSA, under a null evolutionary model (Lai et al., 2012).

The most complete and updated collection of kinase classifiers is NetPhorest (Miller et al., 2008), currently covering 222 kinases and other fundamental signaling domains (Horn et al., 2014). Another milestone in the classification of the kinases responsible for the phosphorylations is NetworKIN (Linding et al., 2007, 2008; Horn et al., 2014), which combines the NetPhorest score with a score that considers the network context of kinases and phosphoproteins, derived from STRING (Franceschini et al., 2013) and based on genomic context, primary experimental evidence, manually curated pathway databases, and automatic literature mining.

Thanks to recent genome sequencing initiatives and phosphoproteomic efforts in several eukaryotes, organism-specific predictors have been developed (Ingrell et al., 2007; Miller et al., 2009; Gao et al., 2010). These methods aim at increasing the prediction accuracy by training on phosphopeptides derived from single organisms. The rationale for these organism-based approaches is that phosphopeptides observed in mass spectrometry experiments performed in these organisms should better represent kinome-specific phosphorylation motifs preferences (Palmeri et al., 2011).

The choice of the predictor is dramatically dependent on user needs, in terms of sensitivity and tolerance to false positives. Some predictors offer to set specific thresholds for specificity and sensitivity (Gao et al., 2010; Xue et al., 2010). Motif-based methods, depending on the motif length and distribution in the proteome, are likely to produce false positives, which can be pruned out by adding more contextual or evolutionary information. Currently, to the best of our knowledge, there is no method that takes advantage of all the aspects here reviewed. Performances will greatly vary from kinase family to family. From Src family to CDK, there are several families whose members share the same motif, and only by deploying contextual information it is possible to distinguish between those members.

Different predictors are often benchmarked using different datasets, at different redundancy levels, with different criteria, and reporting different performance measures (**Table 1**). Therefore, it is quite impractical to rank all the available predictors precisely, only considering the reported accuracies, and establishing the state-of-the-art is unfeasible. Initiatives, like predictors competitions such as DREAM, could be valuable opportunities to set the standards and offer more reliable evaluations.

## KINASE-PEPTIDE SPECIFICITY: THE KINASE SIDE

Given the relatively high frequency of Ser, Thr and Tyr residues in proteomes (in human 8.5, 5.1, and 2.5% respectively), biological systems have evolved efficient strategies to increase the signal to noise ratio and more importantly to minimize those off target phosphorylations leading to detrimental consequences.

The mechanisms of kinase-substrate specificity can be explored at several levels. A major separation is usually operated between peptide specificity and recruitment. Peptide specificity arises from the interactions between the catalytic kinase domain and the substrate peptide, while recruitment is based on interactions between kinase and substrate that do not involve surfaces localized at the catalytic center. During the phosphorylation reaction, the substrate is located together with the ATP in

**Table 1 | Computational methods for kinase-specific phosphorylation site prediction.**

Method name	Approach	Training data	References	Website	Benchmark	Number of kinase families	Homology reduction	Accuracy	Cross validation
Scansite	PSSM	OPL experimental data	Yaffe et al., 2001	scansite.mit.edu	nr*	62	nr	nr	nr
Scansite 2.0	PSSM	OPL experimental data	Obenauer et al., 2003	scansite3.mit.edu	nr	62	nr	nr	nr
Predikin 1.0	Structural	PhosphoBase	Brinkworth et al., 2003		nr	All kinases with structural homologs	nr	Up to 70% of known PhosphoBase sites in top 5%	nr
NetphosK	Artificial Neural Network	PhosphoBase	Blom et al., 2004	cbs.dtu.dk/services/NetPhosK/	YES	6	nr	0.22–0.61 MCC depending on the kinase	3-fold
GPS	BLOSUM62 similarity + Markov Clustering	phosphoELM	Zhou et al., 2004	http://gps.biocuckoo.org/	YES	52	nr	91.8 Sn, 85.0 Sp (PKA) 94.4 Sn, 97.1 Sp (Aurora-B)	nr
KinasePhos 1.0	HMM	PhosphoBase + SwissProt	Huang et al., 2005	kinasephos.mbc.nctu.edu.tw/	Partial	18	nr	0.87	k-fold
PPSP	Bayesian	phosphoELM	Xue et al., 2006	bioinformatics.lcd-ustc.org/PPSP	YES	68	40–90%	90.1Sn 91.7Sp (PKA) 83.2 Sn 90.0Sp (CK2) 93.0Sn 94.1 Sp (ATM) 92.8 Sn 98.0 Sp (S6K)	Leave-one-out
pkaPS	Structural	phosphoELM	Neuberger et al., 2007	mendel.imp.univie.ac.at/sat/pkaPS	YES	1	30%	95.8 Sn 93.5 Sp (PKA)	Neighbor-jackknife
	Cluster-based	phosphoELM + MitoCheck	Moses et al., 2007		YES	1		nr	nr
KinasePhos 2.0	SVM	phosphoELM + SwissProt	Wong et al., 2007	http://KinasePhos2.mbc.nctu.edu.tw/		4	YES	91	Leave-one-out
NetworkKIN	Artificial Neural Network/PSSM + PPIN context	phosphoELM	Linding et al., 2007	networkkin.info	YES	222 kinases	YES	nr	YES
NetPhorest	Artificial Neural Network/PSSM	phosphoELM	Miller et al., 2008	netphorest.info	YES	222 kinases	YES	0.58–1 AUC	YES
PhoScan	PSSM	phosphoELM	Li et al., 2008	http://bioinfo.au.tsinghua.edu.cn/phoscan/		3		At high-stringency cutoffs, ~50% Sn up to 99%. Sp At low-stringency cutoffs, ~90% Sn and Sp for CDK, PKA, CK2	nr
MetaPredPS	MetaPredictor	phosphoELM + PhosphoSite + SwissProt	Wan et al., 2008	http://MetaPred.umn.edu/MetaPredPS/	YES	15	YES	0.86 (CDK) 0.89 (CK2) 0.85 (PKA) 0.78 (PKC)	YES

(Continued)

**Table 1 | Continued**

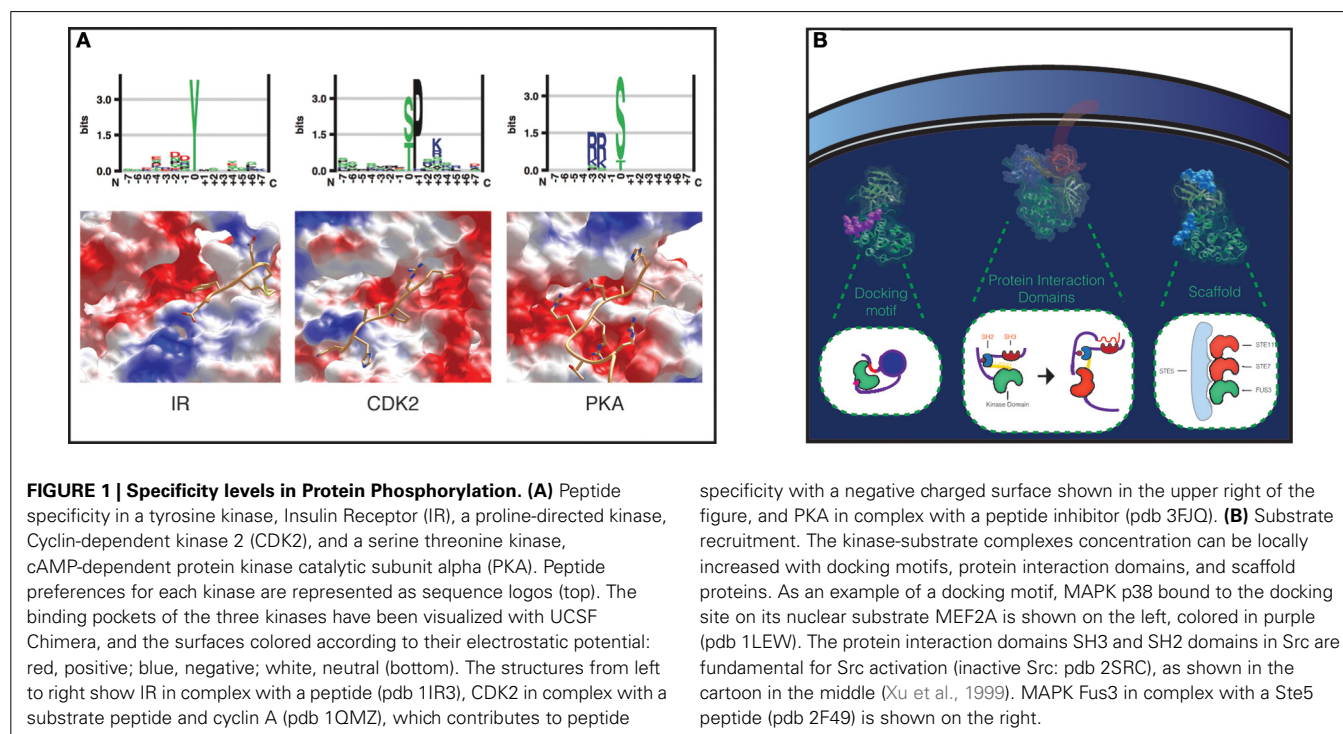
Method name	Approach	Training data	References	Website	Benchmark	Number of kinase families	Homology reduction	Accuracy	Cross validation
Predikin 2.0	Structural with HMM-based kinase identification	phosphoELM + SwissProt	Saunders et al., 2008	<a href="http://predikin.biosci.uq.edu.au/">http://predikin.biosci.uq.edu.au/</a>	YES	All kinases with structural homologs	nr	0.86–0.88 AUC (S/T) 0.66–0.76 AUC (Y)	10-fold
GPS 2.0	Optimized BLOSUM62 similarity	phosphoELM	Xue et al., 2008	<a href="http://gps.biocuckoo.org/">http://gps.biocuckoo.org/</a>	YES	70	YES	83.1 Sn 95.0 Sp (PKA) 100.0 Sn 94.0 Sp (ATM) 78.0 Sn 95.2 Sp (CDC2) 54.0 Sn 95.3 Sp (Src)	Leave-one-out
CRPhos	Conditional Random Fields	phosphoELM	Dang et al., 2008	<a href="http://www.ptools.ua.ac.be/CRPhos">http://www.ptools.ua.ac.be/CRPhos</a>	YES	4	YES	0.93 AUC (CK2) 0.96 AUC (PKA) 0.97 AUC (CDK) 0.91 AUC (PKC)	k-fold
Phos3D	SVM using structural information	phosphoELM	Durek et al., 2009	<a href="http://phos3d.mpimp-golm.mpg.de/">http://phos3d.mpimp-golm.mpg.de/</a>	YES	6	YES	0.73 (Ser kinases) 0.69 (Thr kinases) 0.67 (Tyr kinases)	YES
Musite	SVM	phosphoELM + SwissProt + PhosphoPep + PhosphAt	Gao et al., 2010	<a href="http://musite.sourceforge.net/">http://musite.sourceforge.net/</a>	YES	13 in 6 organisms	50%	1.7–85.5 Sn, 99.9–90.0 Sp (PKA) 6.1–83.6 Sn, 99.9–90.0 Sp (CK2) 0.9–81.9 Sn, 99.9–90.0 Sp (MAPK)	k-fold
GPS 2.1	Optimized BLOSUM62 similarity	phodphoELM	Xue et al., 2011	<a href="http://gps.biocuckoo.org/">http://gps.biocuckoo.org/</a>	Partial	70	nr	70.73 as average on a set of kinase families	nr
PrediKin	Structural with HMM-based kinase identification	phosphoELM + SwissProt	Ellis and Kobe, 2011	<a href="http://predikin.biosci.uq.edu.au/">http://predikin.biosci.uq.edu.au/</a>	YES	All kinases with structural homologs	nr	0.86–0.88 AUC (S/T) 0.66–0.76 AUC (Y) but lower Frobenius distance for PWMs	nr
RegPhos	SVM + PPI + subcellular localization	dbPTM	Lee et al., 2011	<a href="http://regphos.mbc.nctu.edu.tw/">http://regphos.mbc.nctu.edu.tw/</a>	Partial	ca. 100	YES	87.7 (PKC) 92.1 (PIKK) 92.8 (CDK) 91.9 (INSR)	5-fold
ConDens	Conservation of Local Motif Density	No need for training dataset	Lai et al., 2012	<a href="http://www.moseslab.csb.utoronto.ca/andy/">http://www.moseslab.csb.utoronto.ca/andy/</a>	YES	All kinases with known motifs	Does not apply	0.79 AUC	Does not apply
PKIS	SVM	phosphoELM	Zou et al., 2013	<a href="http://bioinformatics.ustc.edu.cn/pkis/">http://bioinformatics.ustc.edu.cn/pkis/</a>	YES	56	YES	13.9 Sn 97.6 Sp (Erk2) 13.5 Sn 97.3 Sp (p38a) 60.7 Sn (0.1 Sp (CK2 $\alpha$ ) 37.5 Sn 93.3 Sp (CDC2) 37.3 Sn 99.8 Sp (PKC $\alpha$ ) 45.0 Sn 93.0 Sp (SYK) 40.0 Sn 97.4 Sp (LCK) 23.5 Sn 94.6 Sp (FYN)	Leave-one-out

(Continued)

**Table 1 | Continued**

Method name	Approach	Training data	References	Website	Benchmark	Number of kinase families	Homology reduction	Accuracy	Cross validation
RegPhos 2.0	SVM + PPI + subcellular localization	dbPTM	Huang et al., 2014	<a href="http://csb.cse.yzu.edu.tw/RegPhos2/">http://csb.cse.yzu.edu.tw/RegPhos2/</a>		122	nr	nr	nr
	SVM	phosphoELM	Xu et al., 2014		YES	7	70%	nr	10-fold
phos_pred	Random Forest	phosphoELM	Fan et al., 2014	<a href="http://bioinformatics.yescustc.edu.cn/phos_pred/">http://bioinformatics.yescustc.edu.cn/phos_pred/</a>		54	YES	68.1–83.4 Sn, 99.1–97.2 Sp (CK2 $\alpha$ ) 68.9–82.2 Sn, 100–95.6 Sp (GSK3B) 71.5–83.5 Sn, 99.1–98.3 Sp (MAPK1) 83.8–88.0 Sn, 99.1–98.3 Sp (MAPK3) etc.	nr

The methods that are currently available for predicting kinase-specific phosphorylation. nr, not reported. \*Impossible to benchmark, because no other methods were available at that time.



the structural region between the two kinase domain lobes, so that the gamma phosphate of the ATP can be transferred to the substrate site. The binding site differs between Ser/Thr and Tyr kinases, allowing the enzymes to discriminate between the three residues. In general each kinase shows a preference for one of these residues. Not only the site, but also its surrounding sequence provides information that is used by kinases to recognize their

target sites (**Figure 1A**). The geometrical and electrostatic properties of the substrate binding sites across the kinases have a substantial impact on substrate specificity. Also different kinases show different electrostatic distributions over their entire surfaces that can influence substrate binding.

Usually, screenings for kinase peptide specificity are performed with Oriented Peptide Libraries (OPL) (Hutti et al., 2004). This

approach revolutionized the determination of kinase specificity, using a mix of solution-phase and solid-support strategies, making kinase specificity screenings both scalable and accurate (Yaffe, 2004). It consists in the quantification of the phosphorylation frequency in degenerate peptide libraries, composed of peptides with a fixed central phosphoacceptor residue and a fixed amino acid in any one of the positions flanking the phosphoacceptor, while the remaining positions are usually drawn from a uniform amino acids distribution. The phosphorylation reaction is performed incubating the kinase with radio-labeled ATP in solution-phase, and after, thanks to a C-terminal biotin tag that is present in all libraries, the peptides are fixed to avidin-coated membranes (Songyang et al., 1994; Hutti et al., 2004; Turk, 2008). The kinase preferences for certain amino acids in fixed peptide positions can then be encoded in *consensus* sequences, in Position Specific Scoring Matrices or more complex classifiers (Miller et al., 2008). From these data, it emerges that peptide specificities of distinct protein kinases are highly variable (Ubersax and Ferrell, 2007; Turk, 2008).

It is generally assumed that the specificity between kinases and substrates is mostly driven by the substrate-binding pocket residues (Ellis and Kobe, 2011), even if also residues localized far from the kinase binding cleft may contribute to shape the peptide specificity.

### THE SUBSTRATE SIDE: PEPTIDE SEQUENCE vs. 3d MOTIFS

Durek and colleagues attempted to characterize 3d-signature phosphorylation site motifs and evaluated their contribution to phosphorylation site prediction performance (Durek et al., 2009). They studied the spatial distribution of amino acids from 2 to 10 Angstrom around each phosphosite, and defined family-specific 3d-profiles. They reported a modest improvement in predicting the kinase families that phosphorylate serine phosphosites, due to the inclusion of structural information. Despite the small discriminatory power of 3d motifs, structural information, like disorder and secondary structure predictions can more efficiently be deployed to improve phosphorylation site predictors performances (Iakoucheva et al., 2004; Durek et al., 2009).

### THE SUBSTRATE SIDE: PEPTIDE INTERPOSITIONAL DEPENDENCE

In 2012 Joughin et al. explored the inter-positional dependence on substrates of ATM/ATR, Cdk1/Cyclin B and CK2 kinases (Joughin et al., 2012). They found only a few significant substrate sequence position pairs that show deviations from position-wise independence. They also tested the ability of first and second order models to correctly separate between the true kinase substrates and mock substrates. Firstly they just used shuffled negative controls (i.e., they shuffled the substrate peptide positions, drawing from the distribution of the true substrates in each position), and they uncovered that mock substrates were similar in quality to the true ones. Then, by using proteomically derived mock substrates, they uncovered that second order were either equal to first order models, or due to over-fitting in training, even worse. Therefore, they concluded that higher-order interdependencies in peptides do not seem to give a significant contribution to predictive performances.

This work has interesting implications for the evolution of signaling networks. There are several examples in the literature showing that the molecular recognition of substrates and phosphopeptide-binding domains is subjected to inter-positional dependences. If also other kinases not included in this study turn out not to show marked second or higher-order preferences on substrate sequences, this would mean that there is a fundamental difference in the way these two components behave in the evolution of signaling networks. As the authors point out, the fitness landscape might look smooth for kinase peptide specificity, and the fitness of the kinase substrate could be boosted after sequential mutations on the peptide, while the fitness landscape for phosphopeptide-binding domain substrate may contain energetic barriers. From a phosphosite predictor perspective, this means that greater efforts should be centered around the development of context-dependent methodologies.

### PLACING KINASE-PEPTIDE SPECIFICITY IN CONTEXT

Although modeling kinase-peptide specificity is fundamental for understanding kinases preferences for their substrates, to study signal propagation in biological systems all phosphorylation events need also to be placed in time and space. As Alexander and colleagues clearly demonstrated in a paper published in 2011, the *in vivo* specificity of mitotic kinases arises from both sub-cellular localization and preferences for phosphorylation motifs (Alexander et al., 2011). For the first time they described an evolutionary conserved mechanism based on a combination of negative and positive phosphorylation motifs selection and spatial localization, to secure proper signal propagation during mitosis. Thus, even if two kinases can share a phosphorylation motif or can localize in the same place, none of the mitotic kinases shares similar preferences in phosphorylation motifs and is also co-localized with any other mitotic kinase.

### SUBSTRATE RECRUITMENT

Protein kinases are highly flexible molecules, and this intrinsic flexibility has likely favored the engineering of complex regulatory and specificity mechanisms throughout eukaryotic evolution. Several mechanisms of recruitment are peculiar to some kinase families, and they can generally be grouped into: scaffold interactions, docking sites, and domain-domain interactions (Reményi et al., 2006; Miller et al., 2008) (Figure 1B).

Scaffold proteins can contribute to specificity increasing the local concentration of the kinase and the substrate, thus enhancing phosphorylation. Probably the best known are MAPK and PKA scaffolds (Wong and Scott, 2004; Strickfaden et al., 2007).

Docking motifs are distant from the phosphosite and facilitate the kinase-substrate recognition (Biondi and Nebreda, 2003). They can be discovered using experimental screening of focused or randomized peptide libraries (Reményi et al., 2006). In the case of Tyr kinases, the motif is usually found in domains that are different from the ones that catalyze the phosphorylation reaction. The motif can also be induced, as in the case of conditional docking sites, where the kinase is recruited only after a phosphorylation event takes place in the motif (Elia et al., 2003). This could also be a way used by the cell to implement logic gates and



keep the timing of phosphorylation. Protein interaction domains, like SH2, SH3, PTB, 14-3-3 can also promote the association between kinases and substrates. Src activation, for instance, is mediated by its SH2 domain (Xu et al., 1999) (see **Figure 1B**). Domain-peptide interactions can be studied experimentally with peptide binding assays, while domain-domain interactions can be modeled using data collected in several databases (Luo et al., 2011; Yellaboina et al., 2011; Kim et al., 2012; Mosca et al., 2014), from high-throughput experiments, like yeast two hybrid, or extracted from the literature (see **Figure 2** for an overview of the major experimental techniques used to identify phosphorylation sites and kinase-substrates interactions).

In 2011 Won and colleagues described the contribution of recruitment interactions to the kinase specificity of Ste7, a MAPKK involved in mating signal flow in *S. cerevisiae* (Won et al., 2011). Ste7 has an interaction with the scaffold protein Ste5 and two docking interactions that allow it to bind to the MAPK Fus3. They uncovered that two out of the three other MAPKK encoded in *S. cerevisiae* genome can functionally replace the MAPKK Ste7, grafting recruitment interactions in their kinase domain. Notably, grafting only the scaffold, or only the docking interactions is not enough to restore the mating signal. This underlines the critical importance of recruitment mechanisms acting concertedly. Scaffold proteins mediating interactions may in theory be discovered using a yeast three hybrid approach, where the kinase and the substrate are fused as bait and prey, and their indirect interaction could be tested expressing the scaffold.

The cellular context has a fundamental role in the determination of the substrate specificity. For instance, kinase localization is important for proper CDKs function. A number of cyclins activate and localize CDKs to different compartments. Overexpression of cyclin B1 causes chromosome condensation, reorganization of the microtubules, and disassembly of the nuclear *lamina* and of the Golgi *apparatus*, while overexpression of cyclin B2 only causes the disassembly of Golgi *apparatus*. Changing the localization motifs, and swapping the two cyclins localizations reverses their phenotypes (Draviam et al., 2001).

Several mechanisms of specificity have recently been explored also for protein phosphatases (Tiganis and Bennett, 2007; Roy and Cyert, 2009). Despite these enzymes lack strong preferences for substrate sequences, higher specificity is obtained with recruitment *via* domains and short linear motifs-mediated interactions and subcellular localization (Sacco et al., 2012).

## PROPERTIES AND EVOLUTION OF POST-TRANSLATIONAL REGULATORY NETWORKS

Currently tens of thousands of phosphorylation sites can be identified by MS-based proteomics in a single experiment (Olsen and Mann, 2013). These large-scale datasets challenged the view of PTMs gained from low-throughput experiments, where a few highly important sites are studied, questioning the functionality of all these PTMs. Several evolutionary studies on phosphorylation sites have confirmed that sites known to be associated to a function are significantly more conserved than non-phosphorylated residues (Gnad et al., 2007; Malik et al., 2008; Landry et al., 2009; Tan et al., 2009a; Moses and Landry, 2010).

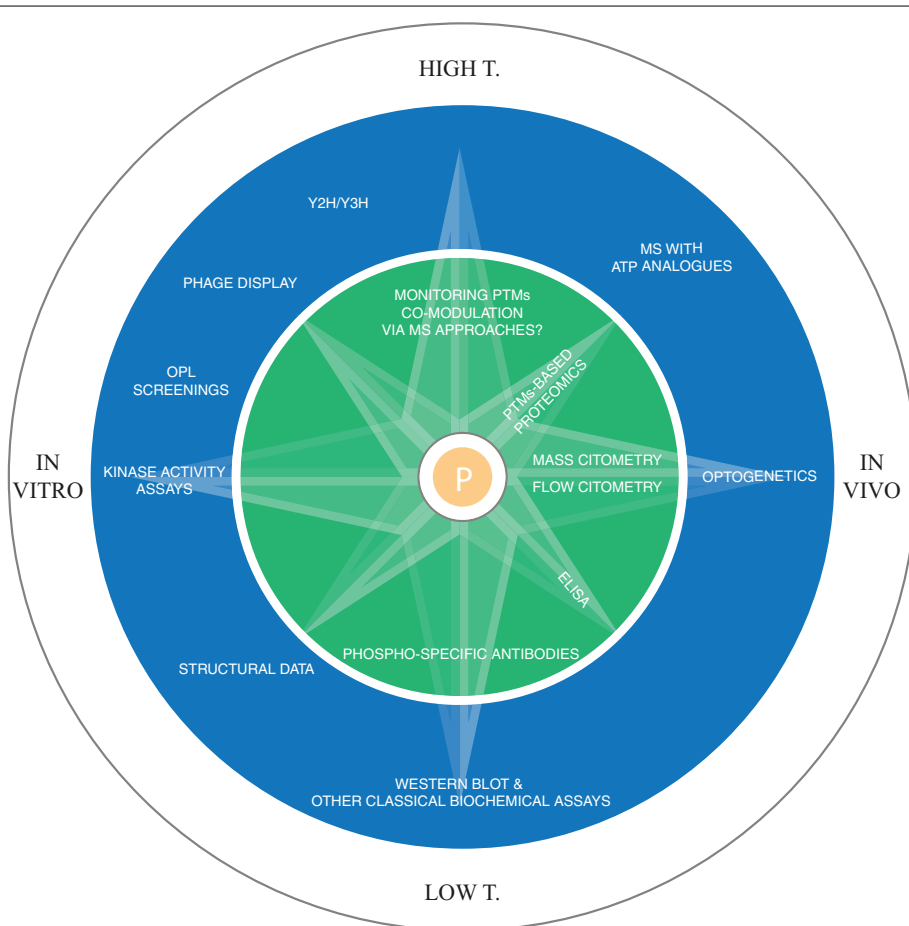
However, a large fraction of phosphorylation sites identified in high-throughput experiments does not show strong evolutionary conservation.

The high evolutionary turnover of phosphoproteomes may be due either to non-functional phosphorylation sites or to species-specific regulation. Recently, a model has been proposed that could explain observations about phosphorylation enrichment in abundant proteins combined with the low stoichiometry of these phosphorylation sites (Levy et al., 2012; Tan and Bader, 2012). According to this model, random encounters between a kinase and proteins in the same subcellular location could end up in a-specific phosphorylations, that will more likely affect highly abundant proteins. This model implies also that only a minimal fraction of an abundant protein population should host these off-target sites. Not all *unintended* phosphorylations are necessarily damaging the cell, otherwise they would have been removed during evolution. Therefore, a fraction of all phosphorylation events could be neutral from an evolutionary perspective. In this *scenario*, the upper bound to the accumulation in the proteome of such sites during evolution is the signaling networks tolerance to noise levels. A nice example of noise minimization has been observed in metazoan lineage evolution, where it has been hypothesized that the signaling networks may have eliminated detrimental phosphorylation sites and limited the noise in the system as tyrosine kinases expanded, by tyrosine-removing mutations (Tan et al., 2009b). In shorter evolutionary distances, the signaling networks properties will more tightly be coupled with the mutational properties of the codons encoding the different phosphorylatable residues. Amongst all residues, serine is considered a mutational hub, as it is very close in mutational space to most residues (Creixell et al., 2012). Indeed it is the only amino acid whose codons are distributed in two groups that are at least two mutations away from each other.

Another explanation for the low conservation of some phosphorylation sites could reside in compensatory mechanisms. In a pioneering work Bodenmiller and colleagues performed single deletions of all yeast kinases and phosphatases, surprisingly observing only a small amount of regulation in the phosphoproteome (Bodenmiller et al., 2010). Even more strikingly, the indirect effects are predominant on effects on the direct targets of the deleted kinases, without strong phenotype alterations, in agreement with the view of signaling networks as systems that are robust to perturbations. Crucially this highlights that a similar cellular state can be the result of different systems regulations. From an evolutionary perspective, different signaling solutions, independently evolved, could be analogous implementations of the same function.

A consistent fraction of eukaryotic phosphoproteomes may represent an evolutionary reservoir that the different organisms could exploit to evolve specific regulation. Estimating more precisely the magnitude of these *non-functional* phosphorylations will contribute in the near future to improve our understanding of post-translational regulatory networks and their properties.

In the case of phosphosites involved in modulating protein-protein interactions, the site may not necessarily be positionally



**FIGURE 2 | Methodologies for the experimental identification of phosphorylation sites.** In the inner circle the techniques for the detection of phosphorylation sites are reported, while the outer circle displays the major techniques for dissecting kinase-substrates interactions, both at the level of direct determination of kinase-substrate interaction (kinase activity assays, OPL, MS with ATP analogs, structural data, western blot, optogenetics) and at the contextual information generation level, i.e., the methodologies that allow the identification of interacting domain

preferences, domain-peptide interactions, etc. (Y2H, Y3H, phage display, structural data, western blot, optogenetics). In the field of PTM identification, future advancements in MS will allow the monitoring of multiple PTMs co-modulation, while for kinase-substrate interactions, the use of ATP analogs coupled with MS/MS is currently the most promising high-throughput technique to link kinases to their substrates *in vivo*. Y2H/Y3H: yeast 2/3 hybrid system (Y3H could be deployed for the study of scaffold proteins-mediated interactions).

conserved (Tan et al., 2010). Phosphosites in different organisms at the same interface tend also to be phosphorylated by kinases of similar specificity. Therefore, the same protein interface may be modulated by functionally redundant sites that are weakly conserved in sequence (Tan et al., 2010; Palmeri et al., 2014).

Many domains in the human proteome seem to have peculiar preferences for being targets of phosphorylation. Some domains, like the kinase domain, tend to be significantly enriched, while other ones tend to be depleted in phosphorylation. Within the same domain, phospho-hot spots can also be identified, i.e., regions that are highly enriched in phosphorylation, suggesting modulation of the domain function via these segments, as in the kinase activation loop, or the C-terminus of HSP90 domain (Beltrao et al., 2012). The domain context of a phosphosite can then be used to functionally characterize the site, and also to improve phosphosite predictor performances (Palmeri et al., 2014).

### PTMs CROSS-TALK

Several low-throughput experiments offer nice examples of how the cell uses PTMs combinations to reach highly sophisticated levels of control (Lo et al., 2001; Choudhary et al., 2009; Wang et al., 2011; Zheng et al., 2011). Difficulties in high-throughput determination of co-modulation between different PTMs currently limit the scale at which analysis of cross-regulation can be conducted. However, in a recent work, Swaney and colleagues studied the cross-talk between phosphorylation and ubiquitylation after proteasome inhibition, thus identifying potential phosphodegrons, analyzing the pairs of phosphorylation sites and ubiquitylation sites that increased in abundance after proteasome inhibition (Swaney et al., 2013). Computational works have already started exploring this relatively novel territory. For instance, Woodsmith and collaborators suggested that PTMs clusters may represent signal integration platforms (Woodsmith et al., 2013). In another work, Minguez and colleagues from

Bork's group, used the concept of correlated evolution to discover new types of co-regulation within different PTMs (Mínguez et al., 2012). Greater efforts in the developments of experimental methods for large scale monitoring of co-modulated PTMs will enormously help in understanding how the signaling networks respond to and integrate different inputs from a systems level perspective. But also new computational models will have to be developed and may benefit from coordinated modeling of the different PTMs.

## CONCLUSION: KINASE-SUBSTRATE SPECIFICITY MODELING

Modeling kinase specificity for substrates is one of the most challenging bioinformatics contributions to cellular signaling. Mass Spectrometry is able to generate large amount of data, but there is currently no high-throughput experimental way to identify the candidate responsible for the phosphorylation.

The two main challenges in developing computational approaches to kinase-substrate specificity are: modeling kinase-peptide specificity and substrate recruitment preferences. Bioinformatics solutions have extensively explored the motif-context, and as pointed out by Joughin et al. the construction of higher order mathematical models might have limited, if any, advantage, at the high cost of overfitted models (Joughin et al., 2012). *In silico* modeling efforts should be centered around a more effective integration of different levels of contextual information, placing the kinases and the substrates in correct space and time, but also considering the interactions outside the kinase domain (domain-domain, scaffolds and docking sites interactions), that can increase locally the concentration of kinases and substrates. Data is obviously critical to all this, therefore more OPL screenings on kinases with unknown specificity, combined with more comprehensive studies to explore scaffold-mediated kinase-substrate interactions and also more efforts dedicated to assess the impact of mutations in domain-domain interactions involved in signaling could contribute to the development of more refined models of kinase-substrate specificity. Probably the most difficult challenge is to convert results from *in vitro* studies to approaches that make reliable *in vivo* predictions. Different kinases vary in their dependence on contextual information. Holistic, i.e., highly integrative, approaches, that allow the modeling of many contexts at the same time, will be able in the future to dissect for each kinase (or kinase group) the different contributions that shape the logic of the signaling system, like for instance in the remarkable case of mitotic kinases, studied by Yaffe's group (Alexander et al., 2011). High-throughput identification of the kinases responsible for the phosphorylation events is critical to achieve this. MS coupled with ATP analogs is currently the most promising approach in this field (Lopez et al., 2013).

Function-dependent classifiers, like those considering the identity of the domain where the phosphorylation is located, can be an alternative way to boost performances in phosphorylation prediction and they might be considered also for biomedical applications. Lastly, from the computational integration of different PTMs models (Creixell and Linding, 2012; Mínguez et al., 2012), it may be possible, in the future, to infer and monitor system-level regulatory centers, whose function might be impaired in complex diseases.

## ACKNOWLEDGMENTS

This work was supported by the EPIGEN flagship project MIUR-CNR to Manuela Helmer-Citterich.

## REFERENCES

- Alexander, J., Lim, D., Joughin, B. A., Hegemann, B., Hutchins, J. R. A., Ehrenberger, T., et al. (2011). Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling. *Sci. Signal.* 4:ra42. doi: 10.1126/scisignal.2001796
- Barford, D., Hu, S. H., and Johnson, L. N. (1991). Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP. *J. Mol. Biol.* 218, 233–260. doi: 10.1016/0022-2836(91)90887-C
- Biondi, R. M., and Nebreda, A. R. (2003). Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem. J.* 372, 1–13. doi: 10.1042/BJ20021641
- Beltrao, P., Albanese, V., Kenner, L. R., Swaney, D. L., Burlingame, A., Villén, J., et al. (2012). Systematic functional prioritization of protein posttranslational modifications. *Cell* 150, 413–425. doi: 10.1016/j.cell.2012.05.036
- Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633–1649. doi: 10.1002/pmic.200300771
- Bodenmiller, B., Wanka, S., Kraft, C., Urban, J., Campbell, D., Pedrioli, P. G., et al. (2010). Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Sci. Signal.* 3:rs4. doi: 10.1126/scisignal.2001182
- Bradham, C. A., Foltz, K. R., Beane, W. S., Arnone, M. I., Rizzo, F., Coffman, J. A., et al. (2006). The sea urchin kinase: a first look. *Dev. Biol.* 300, 180–193. doi: 10.1016/j.ydbio.2006.08.074
- Brinkworth, R. I., Breinl, R. A., and Kobe, B. (2003). Structural basis and prediction of substrate specificity in protein serine/threonine kinases. *Proc. Natl. Acad. Sci. U.S.A.* 100, 74–79. doi: 10.1073/pnas.0134224100
- Budovskaya, Y. V., Stephan, J. S., Deminoff, S. J., and Herman, P. K. (2005). An evolutionary proteomics approach identifies substrates of the cAMP-dependent protein kinase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13933–13938. doi: 10.1073/pnas.0501046102
- Caenepeel, S., Charyczak, G., Sudarsanam, S., Hunter, T., and Manning, G. (2004). The mouse kinase: discovery and comparative genomics of all mouse protein kinases. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11707–11712. doi: 10.1073/pnas.0306880101
- Choudhary, C., Kumar, C., Gnäd, F., Nielsen, M. L., Rehman, M., Walther, T. C., et al. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834–840. doi: 10.1126/science.1175371
- Creixell, P., and Linding, R. (2012). Cells, shared memory and breaking the PTM code. *Mol. Syst. Biol.* 8, 598. doi: 10.1038/msb.2012.33
- Creixell, P., Schoof, E. M., Tan, C. S. H., and Linding, R. (2012). Mutational properties of amino acid residues: implications for evolvability of phosphorylatable residues. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2584–2593. doi: 10.1098/rstb.2012.0076
- Dang, T. H., Van Leemput, K., Verschoren, A., and Laukens, K. (2008). Prediction of kinase-specific phosphorylation sites using conditional random fields. *Bioinformatics* 24, 2857–2864. doi: 10.1093/bioinformatics/btn546
- Draviam, V. M., Orrechia, S., Lowe, M., Pardi, R., and Pines, J. (2001). The localization of human cyclins B1 and B2 determines Cdk1 substrate specificity and neither enzyme requires MEK to disassemble the Golgi apparatus. *J. Cell Biol.* 152, 945–958. doi: 10.1083/jcb.152.5.945
- Durek, P., Schudoma, C., Weckwerth, W., Selbig, J., and Walther, D. (2009). Detection and characterization of 3D-signature phosphorylation site motifs and their contribution towards improved phosphorylation site prediction in proteins. *BMC Bioinformatics* 10:117. doi: 10.1186/1471-2105-10-117
- Elia, A. E. H., Cantley, L. C., and Yaffe, M. B. (2003). Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates. *Science* 299, 1228–1231. doi: 10.1126/science.1079079
- Ellis, J. J., and Kobe, B. (2011). Predicting protein kinase specificity: predikin update and performance in the DREAM4 challenge. *PLoS ONE* 6:e21169. doi: 10.1371/journal.pone.0021169
- Fan, W., Xu, X., Shen, Y., Feng, H., Li, A., and Wang, M. (2014). Prediction of protein kinase-specific phosphorylation sites in hierarchical structure using

- functional information and random forest. *Amino Acids* 46, 1069–1078. doi: 10.1007/s00726-014-1669-3
- Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., et al. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808–D815. doi: 10.1093/nar/gks1094
- Gao, J., Thelen, J. J., Dunker, A. K., and Xu, D. (2010). Musite, a tool for global prediction of general and kinase-specific phosphorylation sites. *Mol. Cell. Proteomics* 9, 2586–2600. doi: 10.1074/mcp.M110.001388
- Gnad, F., Gunawardena, J., and Mann, M. (2011). PHOSIDA 2011: the post-translational modification database. *Nucleic Acids Res.* 39, D253–D260. doi: 10.1093/nar/gkq1159
- Gnad, F., Ren, S., Cox, J., Olsen, J. V., Macek, B., Oroschi, M., et al. (2007). PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol.* 8:R250. doi: 10.1186/gb-2007-8-11-r250
- Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., et al. (2014). KinomeXplorer: an integrated platform for kinome biology studies. *Nat. Methods* 11, 603–604. doi: 10.1038/nmeth.2968
- Huang, H.-D., Lee, T.-Y., Tzeng, S.-W., and Horng, J.-T. (2005). KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. *Nucleic Acids Res.* 33, W226–W229. doi: 10.1093/nar/gki471
- Huang, K.-Y., Wu, H.-Y., Chen, Y.-J., Lu, C.-T., Su, M.-G., Hsieh, Y.-C., et al. (2014). RegPhos 2.0: an updated resource to explore protein kinase-substrate phosphorylation networks in mammals. *Database* 2014:bau034. doi: 10.1093/database/bau034
- Hunter, T. (2012). Why nature chose phosphate to modify proteins. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 2513–2516. doi: 10.1098/rstb.2012.0013
- Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E., and Stroud, R. M. (1990). Regulation of an enzyme by phosphorylation at the active site. *Science* 249, 1012–1016. doi: 10.1126/science.2204109
- Hutti, J. E., Jarrell, E. T., Chang, J. D., Abbott, D. W., Storz, P., Tokar, A., et al. (2004). A rapid method for determining protein kinase phosphorylation specificity. *Nat. Methods* 1, 27–29. doi: 10.1038/nmeth708
- Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 32, 1037–1049. doi: 10.1093/nar/gkh253
- Ingrell, C. R., Miller, M. L., Jensen, O. N., and Blom, N. (2007). NetPhosYeast: prediction of protein phosphorylation sites in yeast. *Bioinformatics* 23, 895–897. doi: 10.1093/bioinformatics/btm020
- Joughin, B. A., Liu, C., Lauffenburger, D. A., Hogue, C. W. V., and Yaffe, M. B. (2012). Protein kinases display minimal interpositional dependence on substrate sequence: potential implications for the evolution of signalling networks. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 2574–2583. doi: 10.1098/rstb.2012.0010
- Kim, Y., Min, B., and Yi, G.-S. (2012). IDDI: integrated domain-domain interaction and protein interaction analysis system. *Proteome Sci.* 10(Suppl. 1):S9. doi: 10.1186/1477-5956-10-S1-S9
- Lai, A. C. W., Nguyen Ba, A. N., and Moses, A. M. (2012). Predicting kinase substrates using conservation of local motif density. *Bioinformatics* 28, 962–969. doi: 10.1093/bioinformatics/bts060
- Landry, C. R., Levy, E. D., and Michnick, S. W. (2009). Weak functional constraints on phosphoproteomes. *Trends Genet.* 25, 193–197. doi: 10.1016/j.tig.2009.03.003
- Lee, T.-Y., Bo-Kai Hsu, J., Chang, W.-C., and Huang, H.-D. (2011). RegPhos: a system to explore the protein kinase-substrate phosphorylation network in humans. *Nucleic Acids Res.* 39, D777–D787. doi: 10.1093/nar/gkq970
- Levy, E. D., Michnick, S. W., and Landry, C. R. (2012). Protein abundance is key to distinguish promiscuous from functional phosphorylation based on evolutionary information. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 2594–2606. doi: 10.1098/rstb.2012.0078
- Li, T., Li, F., and Zhang, X. (2008). Prediction of kinase-specific phosphorylation sites with sequence features by a log-odds ratio approach. *Proteins* 70, 404–414. doi: 10.1002/prot.21563
- Lim, W. A., and Pawson, T. (2010). Phosphotyrosine signaling: evolving a new cellular communication system. *Cell* 142, 661–667. doi: 10.1016/j.cell.2010.08.023
- Linding, R., Jensen, L. J., Ostheimer, G. J., van Vugt, M. A., Jørgensen, C., Miron, I. M., et al. (2007). Systematic discovery of *in vivo* phosphorylation networks. *Cell* 129, 1415–1426. doi: 10.1016/j.cell.2007.05.052
- Linding, R., Jensen, L. J., Pasculescu, A., Olhovskiy, M., Colwill, K., Bork, P., et al. (2008). NetworkKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* 36, D695–D699. doi: 10.1093/nar/gkm902
- Lo, W. S., Duggan, L., Emre, N. C., Belotserkovskaya, R., Lane, W. S., Shiekhata, R., et al. (2001). Snf1—a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* 293, 1142–1146. doi: 10.1126/science.1062322
- Lopez, M. S., Choy, J. W., Peters, U., Sos, M. L., Morgan, D. O., and Shokat, K. M. (2013). Staurosporine-derived inhibitors broaden the scope of analog-sensitive kinase technology. *J. Am. Chem. Soc.* 135, 18153–18159. doi: 10.1021/ja408704u
- Luo, Q., Pagel, P., Vilne, B., and Frishman, D. (2011). DIMA 3.0: Domain interaction map. *Nucleic Acids Res.* 39, D724–D729. doi: 10.1093/nar/gkq1200
- Malik, R., Nigg, E. A., and Körner, R. (2008). Comparative conservation analysis of the human mitotic phosphoproteome. *Bioinformatics* 24, 1426–1432. doi: 10.1093/bioinformatics/btn197
- Manning, G., Plowman, G. D., Hunter, T., and Sudarsanam, S. (2002a). Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* 27, 514–520. doi: 10.1016/S0968-0004(02)02179-5
- Manning, G., Reiner, D. S., Lauwaet, T., Dacre, M., Smith, A., Zhai, Y., et al. (2011). The minimal kinome of *Giardia lamblia* illuminates early kinase evolution and unique parasite biology. *Genome Biol.* 12:R66. doi: 10.1186/gb-2011-12-7-r66
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002b). The protein kinase complement of the human genome. *Science* 298, 1912. doi: 10.1126/science.1075762
- Manning, G., Young, S. L., Miller, W. T., and Zhai, Y. (2008). The protist, *monosiga brevicollis*, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9674–9679. doi: 10.1073/pnas.0801314105
- Miller, M., Jensen, L., Diella, F., Jørgensen, C., Tinti, M., Li, L., et al. (2008). Linear motif atlas for phosphorylation-dependent signaling. *Sci. Signal.* 1:ra2. doi: 10.1126/scisignal.1159433
- Miller, M. L., and Blom, N. (2009). Kinase-specific prediction of protein phosphorylation sites. *Methods Mol. Biol.* 527, 299–310. doi: 10.1007/978-1-60327-834-8\_22
- Miller, M. L., Soufi, B., Jers, C., Blom, N., Macek, B., and Mijakovic, I. (2009). NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics* 9, 116–125. doi: 10.1002/pmic.200800285
- Minguez, P., Parca, L., Diella, F., Mende, D. R., Kumar, R., Helmer-Citterich, M., et al. (2012). Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.* 8, 599. doi: 10.1038/msb.2012.31
- Mosca, R., Céol, A., Stein, A., Olivella, R., and Aloy, P. (2014). 3did: a catalog of domain-based interactions of known three-dimensional structure. *Nucleic Acids Res.* 42, D374–D379. doi: 10.1093/nar/gkt887
- Moses, A. M., Hériché, J.-K., and Durbin, R. (2007). Clustering of phosphorylation site recognition motifs can be exploited to predict the targets of cyclin-dependent kinase. *Genome Biol.* 8:R23. doi: 10.1186/gb-2007-8-2-r23
- Moses, A. M., and Landry, C. R. (2010). Moving from transcriptional to phospho-evolution: generalizing regulatory evolution? *Trends Genet.* 26, 462–467. doi: 10.1016/j.tig.2010.08.002
- Neuberger, G., Schneider, G., and Eisenhaber, F. (2007). pkaPS: prediction of protein kinase phosphorylation sites with the simplified kinase-substrate binding model. *Biol. Direct* 2:1. doi: 10.1186/1745-6150-2-1
- Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003). Scansite 2.0: proteome-wide prediction of cell signalling interactions using short sequence motifs. *Nucleic Acids Res.* 31, 3635–3641. doi: 10.1093/nar/gkg584
- Olsen, J. V., and Mann, M. (2013). Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteomics* 12, 3444–3452. doi: 10.1074/mcp.O113.034181
- Palmeri, A., Ausiello, G., Ferre, F., Helmer-Citterich, M., and Gherardini, P. F. (2014). A proteome-wide domain-centric perspective on protein phosphorylation. *Mol. Cell. Proteomics* 13, 2198–2212. doi: 10.1074/mcp.M114.039990
- Palmeri, A., Gherardini, P. F., Tsigankov, P., Ausiello, G., Späth, G. F., Zilberstein, D., et al. (2011). PhosTryp: a phosphorylation site predictor specific for parasitic protozoa of the family trypanosomatidae. *BMC Genomics* 12:614. doi: 10.1186/1471-2164-12-614

- Reményi, A., Good, M. C., and Lim, W. A. (2006). Docking interactions in protein kinase and phosphatase networks. *Curr. Opin. Struct. Biol.* 16, 676–685. doi: 10.1016/j.sbi.2006.10.008
- Roy, J., and Cyert, M. S. (2009). Cracking the phosphatase code: docking interactions determine substrate specificity. *Sci. Signal.* 2:re9. doi: 10.1126/scisignal.2100re9
- Sacco, F., Perfetto, L., Castagnoli, L., and Cesareni, G. (2012). The human phosphatase interactome: an intricate family portrait. *FEBS Lett.* 586, 2732–2739. doi: 10.1016/j.febslet.2012.05.008
- Saunders, N. F. W., Brinkworth, R. I., Huber, T., Kemp, B. E., and Kobe, B. (2008). Predikin and PredikinDB: a computational framework for the prediction of protein kinase peptide specificity and an associated database of phosphorylation sites. *BMC Bioinformatics* 9:245. doi: 10.1186/1471-2105-9-245
- Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Pwnica-Worms, H., and Cantley, L. C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr. Biol.* 4, 973–982. doi: 10.1016/S0960-9822(00)00221-9
- Strickfaden, S. C., Winters, M. J., Ben-Ari, G., Lamson, R. E., Tyers, M., and Pryciak, P. M. (2007). A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* 128, 519–531. doi: 10.1016/j.cell.2006.12.032
- Swaney, D. L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., et al. (2013). Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nat. Methods* 10, 676–682. doi: 10.1038/nmeth.2519
- Tan, C. S. H., and Bader, G. D. (2012). Phosphorylation sites of higher stoichiometry are more conserved. *Nat. Methods* 9, 317; author reply: 318. doi: 10.1038/nmeth.1941
- Tan, C. S. H., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M. O., Jorgensen, C., et al. (2009a). Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci. Signal.* 2:ra39. doi: 10.1126/scisignal.2000316
- Tan, C. S. H., Jorgensen, C., and Linding, R. (2010). Roles of “junk phosphorylation” in modulating biomolecular association of phosphorylated proteins? *Cell cycle Georg. Tex* 9, 1276–1280. doi: 10.4161/cc.9.7.11066
- Tan, C. S. H., Pasculescu, A., Lim, W. A., Pawson, T., Bader, G. D., and Linding, R. (2009b). Positive selection of tyrosine loss in metazoan evolution. *Science* 325, 1686–1688. doi: 10.1126/science.1174301
- Tiganis, T., and Bennett, A. M. (2007). Protein tyrosine phosphatase function: the substrate perspective. *Biochem. J.* 402, 1–15. doi: 10.1042/BJ20061548
- Trost, B., and Kusalik, A. (2011). Computational prediction of eukaryotic phosphorylation sites. *Bioinformatics* 27, 2927–2935. doi: 10.1093/bioinformatics/btr525
- Turk, B. E. (2008). Understanding and exploiting substrate recognition by protein kinases. *Curr. Opin. Chem. Biol.* 12, 4–10. doi: 10.1016/j.cbpa.2008.01.018
- Ubersax, J. A., and Ferrell, J. E. (2007). Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* 8, 530–541. doi: 10.1038/nrm2203
- Via, A., Diella, F., Gibson, T. J., and Helmer-Citterich, M. (2011). From sequence to structural analysis in protein phosphorylation motifs. *Front. Biosci.* 16, 1261–1275. doi: 10.2741/3787
- Wan, J., Kang, S., Tang, C., Yan, J., Ren, Y., Liu, J., et al. (2008). Meta-prediction of phosphorylation sites with weighted voting and restricted grid search parameter selection. *Nucleic Acids Res.* 36, e22. doi: 10.1093/nar/gkm848
- Wang, Z. A., Singh, D., van der Wel, H., and West, C. M. (2011). Prolyl hydroxylation- and glycosylation-dependent functions of Skp1 in O2-regulated development of Dictyostelium. *Dev. Biol.* 349, 283–295. doi: 10.1016/j.ydbio.2010.10.013
- Won, A. P., Garbarino, J. E., and Lim, W. A. (2011). Recruitment interactions can override catalytic interactions in determining the functional identity of a protein kinase. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9809–9814. doi: 10.1073/pnas.1016337108
- Wong, W., and Scott, J. D. (2004). AKAP signalling complexes: focal points in space and time. *Nat. Rev. Mol. Cell Biol.* 5, 959–970. doi: 10.1038/nrm1527
- Wong, Y.-H., Lee, T.-Y., Liang, H.-K., Huang, C.-M., Wang, T.-Y., Yang, Y.-H., et al. (2007). KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Res.* 35, W588–W594. doi: 10.1093/nar/gkm322
- Woodsmith, J., Kamburov, A., and Stelzl, U. (2013). Dual coordination of post translational modifications in human protein networks. *PLoS Comput. Biol.* 9:e1002933. doi: 10.1371/journal.pcbi.1002933
- Xu, W., Doshi, A., Lei, M., Eck, M. J., and Harrison, S. C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* 3, 629–638. doi: 10.1016/S1097-2765(00)80356-1
- Xu, X., Li, A., Zou, L., Shen, Y., Fan, W., and Wang, M. (2014). Improving the performance of protein kinase identification via high dimensional protein-protein interactions and substrate structure data. *Mol. Biosyst.* 10, 694–702. doi: 10.1039/c3mb70462a
- Xue, Y., Gao, X., Cao, J., Liu, Z., Jin, C., Wen, L., et al. (2010). A summary of computational resources for protein phosphorylation. *Curr. Protein Pept. Sci.* 11, 485–96. doi: 10.2174/138920310791824138
- Xue, Y., Li, A., Wang, L., Feng, H., and Yao, X. (2006). PPSP: prediction of PK-specific phosphorylation site with Bayesian decision theory. *BMC Bioinformatics* 7:163. doi: 10.1186/1471-2105-7-163
- Xue, Y., Liu, Z., Cao, J., Ma, Q., Gao, X., Wang, Q., et al. (2011). GPS 2.1: enhanced prediction of kinase-specific phosphorylation sites with an algorithm of motif length selection. *Protein Eng. Des. Sel.* 24, 255–260. doi: 10.1093/protein/gzq094
- Xue, Y., Ren, J., Gao, X., Jin, C., Wen, L., and Yao, X. (2008). GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy. *Mol. Cell. Proteomics* 7, 1598–1608. doi: 10.1074/mcp.M700574-MCP200
- Yaffe, M. B. (2004). Novel at the library. *Nat. Methods* 1, 13–14. doi: 10.1038/nmeth1004-13
- Yaffe, M. B., Leparc, G. G., Lai, J., Obata, T., Volinia, S., and Cantley, L. C. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nat. Biotechnol.* 19, 348–353. doi: 10.1038/86737
- Yellaboina, S., Tasneem, A., Zaykin, D. V., Raghavachari, B., and Jothi, R. (2011). DOMINE: a comprehensive collection of known and predicted domain-domain interactions. *Nucleic Acids Res.* 39, D730–D735. doi: 10.1093/nar/gkq1229
- Zheng, B., Han, M., Shu, Y.-N., Li, Y.-J., Miao, S.-B., Zhang, X.-H., et al. (2011). HDAC2 phosphorylation-dependent Klf5 deacetylation and RAR $\alpha$  acetylation induced by RAR agonist switch the transcription regulatory programs of p21 in VSMCs. *Cell Res.* 21, 1487–1508. doi: 10.1038/cr.2011.34
- Zhou, F.-F., Xue, Y., Chen, G.-L., and Yao, X. (2004). GPS: a novel group-based phosphorylation predicting and scoring method. *Biochem. Biophys. Res. Commun.* 325, 1443–1448. doi: 10.1016/j.bbrc.2004.11.001
- Zhu, G., Liu, Y., and Shaw, S. (2005). Protein kinase specificity: a strategic collaboration between kinase peptide specificity and substrate recruitment. *Cell Cycle* 4, 52–56. doi: 10.4161/cc.4.1.1353
- Zou, L., Wang, M., Shen, Y., Liao, J., Li, A., and Wang, M. (2013). PKIS: computational identification of protein kinases for experimentally discovered protein phosphorylation sites. *BMC Bioinformatics* 14:247. doi: 10.1186/1471-2105-14-247

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 June 2014; accepted: 21 August 2014; published online: 30 September 2014.

Citation: Palmeri A, Ferrè F and Helmer-Citterich M (2014) Exploiting holistic approaches to model specificity in protein phosphorylation. *Front. Genet.* 5:315. doi: 10.3389/fgene.2014.00315

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Palmeri, Ferrè and Helmer-Citterich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Computational methods for analysis and inference of kinase/inhibitor relationships

Fabrizio Ferrè\*, Antonio Palmeri and Manuela Helmer-Citterich

Centre for Molecular Bioinformatics, Department of Biology, University of Rome Tor Vergata, Rome, Italy

## Edited by:

Allegra Via, Sapienza University of Rome, Italy

## Reviewed by:

Jongrae Kim, University of Glasgow, UK

Cao Dongsheng, Central South University, China

## \*Correspondence:

Fabrizio Ferrè, Centre for Molecular Bioinformatics, Department of Biology, University of Rome Tor Vergata, Via della Ricerca Scientifica s.n.c., Rome, Italy  
e-mail: fabrizio.ferre@uniroma2.it

The central role of kinases in virtually all signal transduction networks is the driving motivation for the development of compounds modulating their activity. ATP-mimetic inhibitors are essential tools for elucidating signaling pathways and are emerging as promising therapeutic agents. However, off-target ligand binding and complex and sometimes unexpected kinase/inhibitor relationships can occur for seemingly unrelated kinases, stressing that computational approaches are needed for learning the interaction determinants and for the inference of the effect of small compounds on a given kinase. Recently published high-throughput profiling studies assessed the effects of thousands of small compound inhibitors, covering a substantial portion of the kinome. This wealth of data paved the road for computational resources and methods that can offer a major contribution in understanding the reasons of the inhibition, helping in the rational design of more specific molecules, in the *in silico* prediction of inhibition for those neglected kinases for which no systematic analysis has been carried yet, in the selection of novel inhibitors with desired selectivity, and offering novel avenues of personalized therapies.

**Keywords: kinase inhibitors, kinase activity modulation, kinase/inhibitor inference, drug design and development, chemogenomics**

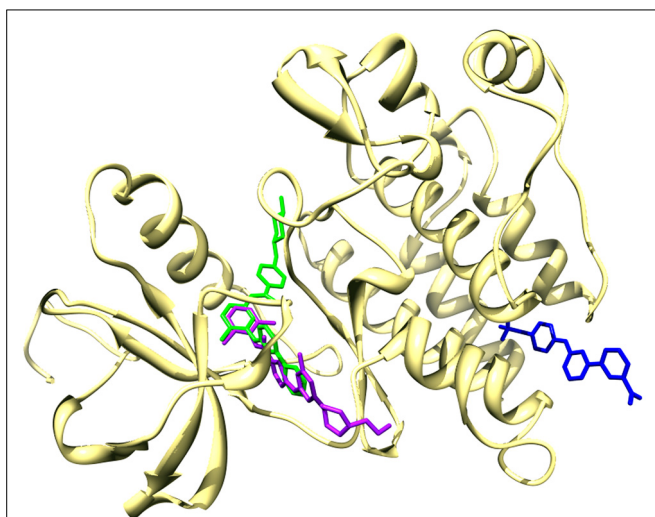
## INTRODUCTION

The kinome plays a predominant role in signal transduction networks and cellular responses; its involvement in a large number of pathologies is a major impulse for the identification and development of compounds modulating the activity of individual kinases or kinase families. Currently, eleven kinase inhibitors are FDA-approved for cancer treatment, and 149 inhibitors and 42 distinct kinase targets are being tested in clinical trials (Fedorov et al., 2010; Chahrour et al., 2012; see <http://www.brimr.org/PKI/PKIs.htm> for an updated list). In addition to their promises as therapeutic agents, kinase inhibitors are commonly used as research tools to disclose the biological consequences of the inactivation of their targets. Generally, kinase inhibitors are ATP-mimetic compounds. The majority of known inhibitors belong to the so-called type I class, and they occupy directly the ATP binding site, located in a hydrophobic cleft between the two lobes of the kinase domain, while type II inhibitors target the ATP binding site as well, but extend also to an allosteric pocket adjacent to the ATP binding site; additional non-ATP-mimetic inhibitor classes (type III, IV, and V), of which a limited number of examples is currently known, seem very promising therapeutic agents given their generally high specificity (Liu and Gray, 2006; Garuti et al., 2010; Chahrour et al., 2012; Gavrin and Saiah, 2013). An example of type I, II, and IV inhibitors is provided in **Figure 1**. For type I and II inhibitors, the evolutionary structural conservation of the kinase ATP-binding site can lead to off-target binding, and while similar kinases tend to show similar inhibition profiles by sharing recurring sequence and structural patterns (Chiu et al., 2013), often complex kinase/inhibitor relationships occur, where kinase bioactivity profiles cannot be reconciled to

their phylogenetic relationships (Paricharak et al., 2013). While absolute specificity toward an individual kinase is not always necessary for a compound to achieve a therapeutic effect (Mencher and Wang, 2005), a detailed knowledge of target selectivity for kinase inhibitors is crucial for predicting and interpreting the effects of inhibitors, and for designing drugs with a desired selectivity. However, kinase inhibitor selectivity is generally not inclusively known for the majority of the tested compounds, as kinase research has been principally focused on a small subset of the kinome.

Traditional kinase inhibitor analysis is a low-throughput process in which the capability of small compounds to decrease the phosphorylation activity (usually reported as the IC<sub>50</sub> or as the remaining or residual activity of the kinase) or their binding affinity (as its dissociation constant) is measured, but are generally not extended to the characterization of the inhibitory abilities of a given compound against the entire kinome. Such data are mined from the literature and collected in general-purpose databases such as ChEMBL (Gaulton et al., 2012) and STITCH (Kuhn et al., 2014), or in kinase-dedicated public resources such as the ChEMBL Kinase SARfari, or the commercially available Kinase Knowledgebase (KKB) by Eidogen-Sertanty (Oceanside, CA, USA) and the kinase inhibitor database provided by GVK Biosciences (Hyderabad, India). While largely populated, such databases tend to be highly heterogeneous by including evidences obtained by diverse means.

However, in recent years the results of medium- and high-throughput profiling studies became available, tackling inhibition of the phosphorylation activity for panels of widely used research compounds and clinical agents against large subsets



**FIGURE 1 | Binding of the ABL kinase with *dasatinib* (type I inhibitor, shown in purple), *imatinib* (type II inhibitor, shown in green), and *GNF-2* (type IV inhibitor, shown in blue).** The human ABL kinase co-crystallized with *imatinib* (PDB code 1IEP) was used as reference for the structural superposition of the human ABL co-crystallized with *dasatinib* (PDB code 2GQG) and of the mouse ABL in complex with the allosteric inhibitor *GNF-2* (PDB code 3K5V). Only the ribbon representation of the human ABL kinase domain from 1IEP (chain A) is shown.

of the human kinome (Table 1). These studies were able to identify novel inhibitor chemotypes for specific kinase targets and to reveal the target specificities of a large set of kinase inhibitors. Importantly, these panels also provide negative results, i.e., inhibitors having little or no effect on tested

kinases, which are instrumental for computational learning techniques and are generally absent or scarce in low-throughput settings.

Additionally, a large and growing number of known three-dimensional (3D) structures of whole kinases or kinase domains are available in the Protein Data Bank (PDB, Berman et al., 2013), and, in few cases, the kinase was also co-crystallized with an inhibitor. These structures provide a rich background for a detailed analysis of kinase binding pockets and for a better identification of binding determinants.

Computational methods for kinase/inhibitor relationships analysis and inference were successfully attempted in the past (e.g., Manallack et al., 2002; Vieth et al., 2004; Xia et al., 2004; Chuaqui et al., 2005), but were limited by the incomplete and heterogeneous data available at the time. In this review we focused on recent computational methods and resources that employ the latest kinase inhibition profiling data but go beyond standard quantitative structure-activity relationship (QSAR) modeling approaches, which are generally specific for a single target, being instead purposely tailored toward kinase inhibition analysis and applied to the whole kinome, taking advantage from the overall kinase domain conservation and from shared binding patterns and characteristics and providing multidimensional structure-activity relationships concerning tens or hundreds of targets at the same time (Goldstein et al., 2008).

## METHODS FOR KINASE/INHIBITOR INFERENCE

Procedures that use inhibition data from panels of proteins tested against panels of compounds are generally based on numerical descriptions of physicochemical, structural and/or geometrical properties of both ligands and targets, and seek possibly non-linear relationships that explain the binding profiles. Machine learning methods are therefore particularly suited, either for classification (binds/does not bind) or regression on the measured inhibition values (e.g.,  $IC_{50}$  or  $K_d$ ). Since all information available for any kinase target and/or inhibitor is used for learning, these studies can be considered a multi-target approach. Additionally, they can be used to infer novel kinase/inhibitor relationships, also for kinases and compounds not included in the training set.

A number of recent papers explored this kind of approach, differing in the employed training dataset, in the way compounds and proteins are described and in the learning algorithm, but following similar pipelines. For example, Nijijima et al. (2012) and Cao et al. (2013) both started from data extracted from Kinase SARfari [in Nijijima et al. (2012) the Metz dataset was additionally used for external validation], and propose a similar kinase/inhibitor deconvolution approach, in which the whole kinase sequences, or only the kinase ATP-binding pockets, are deconstructed into residues (either described simply by amino acid type or by physicochemical characteristics) and compounds into chemical fragments or in topological Daylight fingerprints. Yabuuchi et al. (2011) developed a method, called CGBVS (chemical genomics-based virtual screening), in which compounds were represented by a large set of substructure descriptors and physicochemical properties, and protein descriptors were computed from the protein sequence dipeptide composition using a

**Table 1 | Kinase/inhibitor profiling panels.**

Dataset	Kinases	Compounds	Readout
Fabian et al. (2005)	119	20	$K_d$
Bain et al. (2007)	70–80*	65	% Activity; $IC_{50}$
Fedorov et al. (2007)	60	156	$\Delta T_m$
Bamborough et al. (2008)	203	577	% Control
Karaman et al. (2008)	317	38	$K_d$
Posy et al. (2011)	317–402*	21,851	% Control
Davis et al. (2011)	442	72	$K_d$
Miduturu et al. (2011)	353	118	% Control; $K_d$
Metz et al. (2011)	172	3,858	$K_i$
Anastassiadis et al. (2011)	300	178	% Activity
Cao et al. (2013)	234	158	% Activity; $IC_{50}$
Sutherland et al. (2013)	100	2,871	$IC_{50}$

For each dataset, the number of tested kinases and compounds is reported, together with the type or provided readout:  $K_d$  (dissociation constant);  $K_i$  (inhibition constant); % Activity (remaining catalytic activity); % Control (percentage of kinase bound to the inhibitor compared to a control);  $IC_{50}$  (half maximal inhibitory concentration);  $\Delta T_m$  (thermal stability shift upon inhibitor binding); \*: not all kinase/inhibitor combinations were tested.

string kernel. Originally developed for G-protein-coupled receptor inhibitors, the method was also applied to kinases, using a panel of 143 kinases and 8830 inhibitors, for a total of more than 15,000 tested interactions extracted from the commercial GVK Biosciences kinase inhibitor database. In Lapins and Wikberg (2010), starting from the Karaman dataset, compounds were described by physicochemical and geometrical characteristics, while kinases were described with either alignment-independent or alignment-based methods, by building a multiple alignment of the kinase domains, excluding gap-rich positions, describing columns of the alignment with physicochemical properties, and applying principal component analysis (PCA) and partial least squares discriminant analysis to summarize descriptors. Schürer and Muskál (2013) employed the Eidogen-Sertanty KKB Q4 2009 release, including more than 430,000 tested kinase/compound pairs extracted from literature and patents. Given the heterogeneous nature of the dataset, data were subject to filtering, standardization, and clustering procedures. For each kinase in the dataset, active and inactive compounds were described using extended connectivity fingerprints, and negative instances for training were either known as inactive on a given kinase, or taken as the entire set of molecules not tested on that kinase.

Then, in these works, machine learning algorithms were trained on kinases and compounds converted into numerical descriptors, to learn associations between kinase residues and compound fragments, and for inference. Variants of a naïve Bayesian (NB) classifier or of a support vector machine (SVM) were used in Nijima et al. (2012), a random forest (RF) in Cao et al. (2013), SVM, decision trees, k-nearest neighbors, and partial least squares projections in Lapins and Wikberg (2010), an SVM in Yabuuchi et al. (2011), Laplacian-corrected NB classifiers, k-nearest neighbors, and partial least squares regression in Schürer and Muskál (2013). All these studies achieved good prediction performances: from 0.67 to 0.73 correlation coefficient in Lapins and Wikberg (2010); accuracy between 74 and 81% and Matthews correlation coefficient (MCC) between 0.3 and 0.48 in different tested datasets and with different encodings and learning methods in Nijima et al. (2012); 94% accuracy and 0.98 area under the ROC curve (auROC) in Cao et al. (2013). In Schürer and Muskál (2013), the auROC for individual kinase models vary from around 0.93 to 1, and the prediction accuracy showed a positive correlation with the number of known inhibitors available for training. In Yabuuchi et al. (2011), some predicted novel inhibitors for the epidermal growth factor receptor kinase and the cyclin-dependent kinase 2 were experimentally confirmed, sometimes showing scaffold hopping (i.e., having radically different characteristics than known inhibitors).

Another class of methods includes those taking advantage of kinase 3D structures, used to obtain a more accurate representation of kinase binding sites. A reasonable assumption is that the affinity that a kinase, or a set of kinases, show toward a compound can be ascribed to set of residues that either allow or hinder the binding, and that, once identified in the 3D structures, can be looked for in other kinases to infer their binding ability, even for those kinases for which the 3D structure is unknown, by taking advantage of the kinase domain sequence conservation. Such sets

of residues can additionally be converted in numerical descriptors for machine learning.

A subset of kinase/inhibitor pairs extracted from the Fabian and Karaman datasets was used in Caffrey et al. (2008). For these inhibitors the structure of the kinase/compound complex is known, and the specificity determinants can be rationalized. An algorithm was developed to predict specificity determinants given a kinase multiple sequence alignment and structural information, which was able to reproduce the known determinants and to highlight non-trivial additional factors, and can be used as basis for the design of drugs with a desired specificity.

X-React<sup>KIN</sup> (Brylinski and Skolnick, 2010) is a machine learning method for assessment of cross-reactivity in which each human kinase domain structure was obtained through homology modeling, and binding sites residues were predicted using computational methods. Similarity between kinases was computed by different metrics using sequence, structure, and ligand binding profiles. The system employed data from the Fabian and Karaman panels for training and validation of a NB classifier, obtaining sensitivity higher than 0.5 for around 70% of the tested compounds, and the Bamborough dataset was used for further validation, finding significant correspondence (0.53 average Pearson correlation) between predicted and experimental activity profiles. The computed cross-reactivity profiles are freely available for download.

In Huang et al. (2010), all kinase 3D structures available in the PDB at the time were superposed to obtain a fine description of a series of features known in the literature to be related to inhibitor specificity, e.g., the size of the gatekeeper residue, that affects the pocket accessibility, the hydrogen bonding and covalent bonding ability at specific positions, the flexibility of the hinge loop connecting the kinase domain small and large lobes, and others. These features were extended to kinases for which the structure is unknown via multiple alignments, converted into numerical vectors and used to estimate a similarity between each pair of kinases. Using these distances, a network of kinase binding sites was constructed, which recapitulated well a network based on the similarity between the inhibitor profiles in the Karaman dataset. Integration of the binding site similarity network with the inhibition profile network led to inference of off-target interactions, some of which were validated experimentally.

On the same lines, in Anderson et al. (2012), starting from the Karaman dataset, first kinases were clustered by similarity in binding affinity profiles for the inhibitors tested in the dataset. Kinases within the same cluster were shown to have more similar binding sites, as detected by the comparison of the binding site 3D structures extracted from the PDB. *In silico* docking procedures then highlighted cluster-specific residues acting as interaction hot spots, which were converted into a series of descriptors, used for RF training, achieving 76% of prediction accuracy. The RF was then used for the prediction of novel kinase/inhibitor relationships, some of which were experimentally tested, obtaining a good agreement with the predicted  $K_i$  values in 70% of the cases.

The Karaman dataset, crossed with kinase 3D structures available in the PDB, were also the starting point for the work presented in Bryant et al. (2013); the structure of a kinase bound to a known



type II kinase inhibitor, *imatinib*, was used as template to identify contact residues, mapped to all other considered kinases using the Pfam (Punta et al., 2012) kinase family multiple alignment. A combinatorial clustering was used to find subsets of binding site residues that better correlate with the binding affinities reported in the Karaman dataset. An SVM was then trained on these data, and the prediction performance was estimated individually for each inhibitor as the auROC, which ranges from 0.5 to 1 (mean 0.8). Finally, the trained SVM was used to infer the binding ability of unlabeled kinases.

## INTEGRATIVE APPROACHES

The wealth of kinase inhibition profiling data presents great opportunities for being analyzed as a whole, by integrating data from different resources in order to provide a unified view on kinome inhibition. The whole kinase/inhibitor data can therefore be represented as a network, where binding can be treated as a binary on–off relation or weighted by the binding affinity or by the strength of the inhibitory effect. This kind of network can aid in the identification and rationalization of drugs secondary effects and facilitate drug repositioning.

KIDFamMap (Chiu et al., 2013) and K-Map (Kim et al., 2013) are free web-databases in which kinase/inhibitor relationships, retrieved from different sources, are connected and integrated with other annotations to facilitate the at-a-glance investigation of the kinome inhibition. In KIDFamMap, the Karaman, Anastassiadis and Davis profiling panels, Kinase SARfari, the PDB, and others resources, for a total of more than 186,000 kinase/compound pairs, are investigated by decomposing each interaction into a series of binding pocket sub-regions and compound fragments preferences (Chen et al., 2010), and then extending the identified rules to the whole kinome (introducing the concept of kinase/inhibitor families) and associated to known pathologies involving kinases. Queries can start from a kinase, a compound or a disease, retrieving a detailed overview of the kinase/inhibitor interaction, all the other interactions belonging to the same family, and a description of associated diseases and how allelic variants might affect the compound binding. In K-Map the Anastassiadis and Davis datasets were analyzed by building connectivity maps based on the Kolmogorov–Smirnov statistic to find correlations between inhibitors and lists of kinases. K-Map allows querying these datasets by kinase, kinase family, custom lists, or kinase-related GO terms, obtaining lists of associated inhibitors ranked by correlation significance. Similarly, the user can start from lists of inhibitors. The intent of K-Map is to provide insights for drug development and repositioning.

Caveats of integrative approaches are that to convert data into an on–off relation would require setting thresholds that might not be easy to optimize, and that data from different sources might not be directly comparable, so they must be opportunely processed. In Sutherland et al. (2013) the Anastassiadis, Metz and Davis datasets were compared to each other and to an additional profiling panel (the Sutherland dataset in Table 1), by converting each readout in an estimated IC<sub>50</sub>, testing the concordance between IC<sub>50</sub> in different panels, and for promiscuity and selectivity measures. They found that the all panels have good agreement in assessing whether a compound is active or inactive on a given kinase, but the exact

inhibition values show instead low levels of concordance, as well as measures of how much selective is a compound.

In Tang et al. (2014) the Metz, Davis, and Anastassiadis datasets were compared and integrated with data from ChEMBL and STITCH. Since these panels employed different assays and different readouts (K<sub>d</sub>, K<sub>i</sub> and percentage of remaining activity for the Davis, Metz and Anastassiadis datasets, respectively), a new method called KIBA (kinase inhibitor bioactivity) is introduced to obtain a single comparable activity score for each kinase/compound pair. The three panels have a relatively small number of common tested kinase/inhibitor pairs; in such cases, the Metz and Davis datasets show good degree of correlation between readouts, which is smaller when both are compared with the Anastassiadis panel. The project resulted in a kinase/inhibitor bioactivity map comprising 467 kinases and more than 50,000 compounds, which is freely available.

## CONCLUSION

While different in methods and scope, the approaches presented here highlight the need for original and effective computational methods to unravel the rich and complex kinase/inhibitor relationships systematically measured in inhibition profiling panels, which can have significant implications in understanding the reasons of the inhibition, helping in the rational design of bioactive molecules, and can be used for the *in silico* prediction of inhibition for those neglected kinases for which no systematic analysis has been carried yet, and for the selection of inhibitors with desired promiscuity. Additionally, a better understanding of the kinase determinants of inhibition can help in apprehending the different response of individual patients to treatment, such as inhibitor resistance due to specific mutations, moving toward a more personalized treatment.

## ACKNOWLEDGMENT

This work was supported by Programmi di Ricerca di rilevante Interesse Nazionale (PRIN) 2010 (prot. 20108XYHJS\_006 to Manuela Helmer-Citterich).

## REFERENCES

- Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H., and Peterson, J. R. (2011). Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1039–1045. doi: 10.1038/nbt.2017
- Anderson, P. C., De Sapio, V., Turner, K. B., Elmer, S. P., Roe, D. C., and Schoeniger, J. S. (2012). Identification of binding specificity-determining features in protein families. *J. Med. Chem.* 55, 1926–1939. doi: 10.1021/jm200979x
- Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C. J., McLauchlan, H., et al. (2007). The selectivity of protein kinase inhibitors: a further update. *Biochem. J.* 408, 297–315. doi: 10.1042/BJ20070797
- Bamborough, P., Drewry, D., Harper, G., Smith, G. K., and Schneider, K. (2008). Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *J. Med. Chem.* 51, 7898–7914. doi: 10.1021/jm8011036
- Berman, H. M., Kleywegt, G. J., Nakamura, H., and Markley, J. L. (2013). How community has shaped the Protein Data Bank. *Structure* 21, 1485–1491. doi: 10.1016/j.str.2013.07.010
- Bryant, D. H., Moll, M., Finn, P. W., and Kavraki, L. E. (2013). Combinatorial clustering of residue position subsets predicts inhibitor affinity across the human kinome. *PLoS Comput. Biol.* 9:e1003087. doi: 10.1371/journal.pcbi.1003087

- Brylinski, M., and Skolnick, J. (2010). Cross-reactivity virtual profiling of the human kinome by X-react(KIN): a chemical systems biology approach. *Mol. Pharm.* 7, 2324–2333. doi: 10.1021/mp1002976
- Caffrey, D. R., Lunney, E. A., and Moshinsky, D. J. (2008). Prediction of specificity-determining residues for small-molecule kinase inhibitors. *BMC Bioinformatics* 9:491. doi: 10.1186/1471-2105-9-491
- Cao, D. S., Zhou, G. H., Liu, S., Zhang, L. X., Xu, Q. S., He, M., et al. (2013). Large-scale prediction of human kinase/inhibitor interactions using protein sequences and molecular topological structures. *Anal. Chim. Acta* 792, 10–18. doi: 10.1016/j.aca.2013.07.003
- Chahrouh, O., Cairns, D., and Omran, Z. (2012). Small molecule kinase inhibitors as anti-cancer therapeutics. *Mini Rev. Med. Chem.* 12, 399–411. doi: 10.2174/138955712800493915
- Chen, Y. F., Hsu, K. C., Lin, S. R., Wang, W. C., Huang, Y. C., and Yang, J. M. (2010). SIMMap: a web server for inferring site-moiety map to recognize interaction preferences between protein pockets and compound moieties. *Nucleic Acids Res.* 38, W424–W430. doi: 10.1093/nar/gkq480
- Chiu, Y. Y., Lin, C. T., Huang, J. W., Hsu, K. C., Tseng, J. H., You, S. R., et al. (2013). KIDFamMap: a database of kinase/inhibitor-disease family maps for kinase inhibitor selectivity and binding mechanisms. *Nucleic Acids Res.* 41, D430–D440. doi: 10.1093/nar/gks1218
- Chuaqui, C., Deng, Z., and Singh, J. (2005). Interaction profiles of protein kinase/inhibitor complexes and their application to virtual screening. *J. Med. Chem.* 48, 121–133. doi: 10.1021/jm049312t
- Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., et al. (2011). Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1046–1051. doi: 10.1038/nbt.1990
- Fabian, M. A., Biggs W. H. III, Treiber, D. K., Atteridge, C. E., Azimioara, M. D., Benedetti, M. G., et al. (2005). A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* 23, 329–336. doi: 10.1038/nbt1068
- Fedorov, O., Marsden, B., Pogacic, V., Rellos, P., Müller, S., Bullock, A. N., et al. (2007). A systematic interaction map of validated kinase inhibitors with Ser/Thr kinases. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20523–20528. doi: 10.1073/pnas.0708800104
- Fedorov, O., Muller, S., and Knapp, S. (2010). The (un)targeted cancer kinome. *Nat. Chem. Biol.* 6, 166–169. doi: 10.1038/nchembio.297
- Garuti, L., Roberti, M., and Bottegioni, G. (2010). Non-ATP competitive protein kinase inhibitors. *Curr. Med. Chem.* 17, 2804–2821. doi: 10.2174/092986710791859333
- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., et al. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 40, D1100–D1107. doi: 10.1093/nar/gkr777
- Gavrin, L. K., and Saiah, E. (2013). Approaches to discover non-ATP site kinase inhibitors. *Med. Chem. Commun.* 4, 41–51. doi: 10.1039/c2md20180a
- Goldstein, D. M., Gray, N. S., and Zarrinkar, P. P. (2008). High-throughput kinase profiling as a platform for drug discovery. *Nat. Rev. Drug Discov.* 7, 391–397. doi: 10.1038/nrd2541
- Huang, D., Zhou, T., Lafleur, K., Nevado, C., and Caflisch, A. (2010). Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* 26, 198–204. doi: 10.1093/bioinformatics/btp650
- Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., et al. (2008). A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 26, 127–132. doi: 10.1038/nbt1358
- Kim, J., Yoo, M., Kang, J., and Tan, A. C. (2013). K-Map: connecting kinases with therapeutics for drug repurposing and development. *Hum. Genomics* 7, 20. doi: 10.1186/1479-7364-7-20
- Kuhn, M., Szklarczyk, D., Pletscher-Frankild, S., Blicher, T. H., von Mering, C., Jensen, L. J., et al. (2014). STITCH 4: integration of protein-chemical interactions with user data. *Nucleic Acids Res.* 42, D401. doi: 10.1093/nar/gkt1207
- Lapins, M., and Wikberg, J. E. (2010). Kinome-wide interaction modelling using alignment-based and alignment-independent approaches for kinase description and linear and non-linear data analysis techniques. *BMC Bioinformatics* 11:339. doi: 10.1186/1471-2105-11-339
- Liu, Y., and Gray, N. S. (2006). Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* 2, 358–364. doi: 10.1038/nchembio799
- Manallack, D. T., Pitt, W. R., Gancia, E., Montana, J. G., Livingstone, D. J., Ford, M. G., et al. (2002). Selecting screening candidates for kinase and G protein-coupled receptor targets using neural networks. *J. Chem. Inf. Comput. Sci.* 42, 1256–1262. doi: 10.1021/ci020267c
- Mencher, S. K., and Wang, L. G. (2005). Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clin. Pharmacol.* 5:3. doi: 10.1186/1472-6904-5-3
- Metz, J. T., Johnson, E. F., Soni, N. B., Merta, P. J., Kifle, L., and Hajduk, P. J. (2011). Navigating the kinome. *Nat. Chem. Biol.* 7, 200–202. doi: 10.1038/nchembio.530
- Miduturu, C. V., Deng, X., Kwiatkowski, N., Yang, W., Brault, L., Filippakopoulos, P., et al. (2011). High-throughput kinase profiling: a more efficient approach toward the discovery of new kinase inhibitors. *Chem. Biol.* 18, 868–879. doi: 10.1016/j.chembiol.2011.05.010
- Nijima, S., Shiraishi, A., and Okuno, Y. (2012). Dissecting kinase profiling data to predict activity and understand cross-reactivity of kinase inhibitors. *J. Chem. Inf. Model.* 52, 901–912. doi: 10.1021/ci200607f
- Paricharak, S., Klenka, T., Augustin, M., Patel, U. A., and Bender, A. (2013). Are phylogenetic trees suitable for chemogenomics analyses of bioactivity data sets: the importance of shared active compounds and choosing a suitable data embedding method, as exemplified on Kinases. *J. Cheminform.* 5, 49. doi: 10.1186/1758-2946-5-49
- Posy, S. L., Hermsmeier, M. A., Vaccaro, W., Ott, K. H., Todderud, G., Lippy, J. S., et al. (2011). Trends in kinase selectivity: insights for target class-focused library screening. *J. Med. Chem.* 54, 54–66. doi: 10.1021/jm101195a
- Punta, M., Coghill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., et al. (2012). The Pfam protein families database. *Nucleic Acids Res.* 40, D290–D301. doi: 10.1093/nar/gkr1065
- Schürer, S. C., and Muskal, S. M. (2013). Kinome-wide activity modeling from diverse public high-quality data sets. *J. Chem. Inf. Model.* 53, 27–38. doi: 10.1021/ci300403k
- Sutherland, J. J., Gao, C., Cahya, S., and Vieth, M. (2013). What general conclusions can we draw from kinase profiling data sets? *Biochim. Biophys. Acta* 1834, 1425–1433. doi: 10.1016/j.bbapap.2012.12.023
- Tang, J., Sz wajda, A., Shakyawar, S., Xu, T., Hintsanen, P., Wennerberg, K., et al. (2014). Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. *J. Chem. Inf. Model.* 54, 735–743. doi: 10.1021/ci400709d
- Vieth, M., Higgs, R. E., Robertson, D. H., Shapiro, M., Gragg, E. A., and Hemmerle, H. (2004). Kinomics-structural biology and chemogenomics of kinase inhibitors and targets. *Biochim. Biophys. Acta* 1697, 243–257. doi: 10.1016/j.bbapap.2003.11.028
- Xia, X., Maliski, E. G., Gallant, P., and Rogers, D. (2004). Classification of kinase inhibitors using a Bayesian model. *J. Med. Chem.* 47, 4463–4470. doi: 10.1021/jm0303195
- Yabuuchi, H., Nijima, S., Takematsu, H., Ida, T., Hirokawa, T., Hara, T., et al. (2011). Analysis of multiple compound-protein interactions reveals novel bioactive molecules. *Mol. Syst. Biol.* 7, 472. doi: 10.1038/msb.2011.5

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 May 2014; paper pending published: 30 May 2014; accepted: 13 June 2014; published online: 30 June 2014.

Citation: Ferrè F, Palmeri A and Helmer-Citterich M (2014) Computational methods for analysis and inference of kinase/inhibitor relationships. *Front. Genet.* 5:196. doi: 10.3389/fgene.2014.00196

This article was submitted to *Systems Biology*, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Ferrè, Palmeri and Helmer-Citterich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Defying c-Abl signaling circuits through small allosteric compounds

Stefania Gonfloni \*

Department of Biology, University of Rome Tor Vergata, Rome, Italy

## Edited by:

Andreas Zanzoni, Inserm TAGC  
UMR1090, France  
Allegra Via, Sapienza University,  
Italy

## Reviewed by:

Oliver Hantschel, École  
Polytechnique Fédérale de  
Lausanne, Switzerland  
Stephan Grzesiek, University of  
Basel, Switzerland

## \*Correspondence:

Stefania Gonfloni, Department of  
Biology, University of Rome Tor  
Vergata, Via della Ricerca  
Scientifica, 00133 Rome, Italy  
email: stefania.gonfloni@  
uniroma2.it

Many extracellular and intracellular signals promote the c-Abl tyrosine kinase activity. c-Abl in turn triggers a multitude of changes either in protein phosphorylation or in gene expression in the cell. Yet, c-Abl takes part in diverse signaling routes because of several domains linked to its catalytic core. Complex conformational changes turn on and off its kinase activity. These changes affect surface features of the c-Abl kinase and likely its capability to bind actin and/or DNA. Two specific inhibitors (ATP-competitive or allosteric compounds) regulate the c-Abl kinase through different mechanisms. NMR studies show that a c-Abl fragment (SH3-SH2-linker-SH1) adopts different conformational states upon binding to each inhibitor. This supports an unconventional use for allosteric compounds to unraveling physiological c-Abl signaling circuits.

**Keywords: c-Abl signaling motifs, stress responses, allosteric compounds**

## INTRODUCTION

Proteomics has revealed a rather complex picture underlying cellular signaling circuits. Recent studies have indicated an extensive overlap between diverse cellular responses. DNA damage-induced stress response overlaps with pathogen infection response and with heat stress in intact animals (*Caenorhabditis elegans*; Ermolaeva et al., 2013). Signaling connections between DNA damage, stress response, and aging remain elusive in other intact organisms (Gartner and Akay, 2013). Yet, evidence from Ermolaeva et al. (2013) supports a model of an integrated signaling network at the interface of various cell stress routes.

We have discussed about the aberrant c-Abl signaling in the molecular events at the interface of oxidative stress – metabolic regulation, protein aggregation, and DNA damage in neurons (Gonfloni et al., 2012). We have proposed that various stress responses seem to rely on a small set of recurring c-Abl-mediated regulatory circuits (Gonfloni et al., 2012). An emerging theme in neuronal diseases is the aberrant interplay between c-Abl phosphorylation of transcription factors, adaptors, modifiers/enzymes, and ubiquitin-mediated signaling responses (Gonfloni et al., 2012; Ciccone et al., 2013).

In this perspective, I will discuss how modulation of c-Abl, through small molecule allosteric inhibitors/ligands could be exploited to tackle the interface of c-Abl signaling circuits.

## SURFING THE CELL SIGNALING CIRCUITS

Cell metabolism and homeostasis rely on signaling networks of interacting proteins. Posttranslational modifications (PTMs) are crucial for the network. Colocalization of the binding partners converts protein interactions into functional consequences

(Kuriyan and Eisenberg, 2007). Crosstalk and interplay between different PTMs give rise to a versatile, rich, and dynamic framework of signaling circuits. Negative (or positive) feedback loops control the amplitude of signaling pathway and the sustained activation in time, conveying signals for irreversible decisions of the cell. A crucial issue for understanding cell signaling is to define how PTMs control changes in metabolism and homeostasis. A simple way is to consider protein domains as basic units of cell signaling (Kuriyan and Cowburn, 1997). Cells may use modular binding motifs like a broad device dedicated to the selective recognition of PTMs (Seet et al., 2006). However, complex biochemical responses can be only achieved in the context of multidomain proteins or multiprotein complexes. Of note, PTMs can also induce a new conformational state that in turn promotes an allosteric regulation of the targeted protein. An example of such an allosteric modulation is the phosphorylation of the activation loop of Src tyrosine kinases. This phosphorylation promotes the productive configuration of the catalytic site. By this mechanism, Src activity (i.e., a productive conformation of the active site in the catalytic domain) also controls the availability of the regulatory domains (SH3-linker; Gonfloni et al., 2000). This effect in turn promotes selective binding/recruitment in multiprotein complexes.

Posttranslational modifications are dynamic and necessary for assembling a temporary platform of local signaling circuits. PTMs work as an allosteric device both for the modified protein and for the assembly/(or reshaping?) of multiprotein complexes. Enzymes are often “switchable,” with their activities controlled by many targets/effectors. The tyrosine kinases are themselves regulated by phosphorylation through various allosteric mechanisms.

Non-receptor tyrosine kinases (RTKs) have a conserved catalytic domain (kinase domain = SH1), auto-inhibited by the binding of regulatory domains (SH2, SH3). Such an allosteric regulation links the enzymatic activity of the kinase with the colocalization of its substrate. Allosteric auto-regulation seems to be a recurring feature in cell signaling (Liu and Nussinov, 2013). Protein domains with enzymatic activity (acting as modifiers/writers) are often in tandem with the binding motifs devoted to recognition (acting as readers) of the same modification. This concept is also well substantiated by ubiquitin-mediated signaling. Ubiquitin represents a transferable interaction domain recognized by specialized binding motifs (ubiquitin binding domains, UBDs). Monoubiquitinated proteins often contain a UBD required for their auto-regulation (Seet et al., 2006).

### c-Abl IS AN ALLOSTERIC SIGNALING SWITCH FOR VARIOUS CELL RESPONSES

Non-receptor Abl tyrosine kinases regulate a diverse range of cellular signaling paths. Recent reviews have discussed both the biological functions of the mammalian c-Abl tyrosine kinase (Colicelli, 2010; Wang, 2014) and the role of Abl family kinases in cancer (Greuber et al., 2013). It is beyond the scope of this review discussing these aspects. Interested readers are directed to several excellent reviews on this topic (Sirvent et al., 2008; Hossain et al., 2012; Greuber et al., 2013; Wang, 2014).

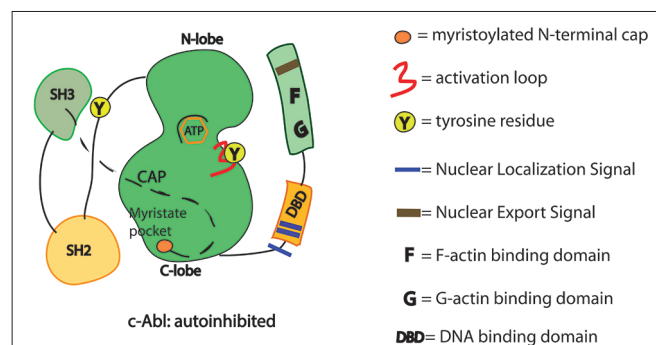
Here, I will recall the mechanisms of c-Abl auto-inhibition. The purpose is to highlight the effects of small molecule inhibitors on the conformation of c-Abl.

The Abl kinase family comprises two related proteins Abl1 (c-Abl) and Abl2 (Arg). Both kinases have redundant and unique roles due to their conserved sequence/domain structures. c-Abl and Arg have two different variants (1a and 1b). Both variants are ubiquitously expressed. Abl kinases share a conserved assembly of amino-terminal regulatory and catalytic domain (SH3–SH2–linker–SH1 domain). At the carboxyl terminus region, Abl kinases contain a filamentous (F)-actin-binding domain (ABD; Van Etten et al., 1994). c-Abl and Arg are less conserved in the middle region. So, Arg lacks of the three nuclear localization signal (NLS) motifs and localizes in the cytoplasm and in cell periphery (Miller et al., 2004). By contrast, c-Abl is present in the cytosol but also in organelles, such as the endoplasmic reticulum (ER) the mitochondria (Ito et al., 2001), or the nucleus (Wen et al., 1996). The diverse localization of c-Abl is modulated by PTMs (Yoshida et al., 2005). The formation of distinct multiprotein complexes is likely regulated by a dynamic spatial recruitment. Spatial distribution of c-Abl is linked to the catalytic competence of the kinase. The latter is in constant equilibrium between low (fully inhibited) and high (fully activated) levels of activity (Hantschel and Superti-Furga, 2004). c-Abl signaling is not only dependent from the outcomes derived from its enzymatic kinase activity. But, it depends from dynamic recruitment of c-Abl into different protein complexes (and subcellular compartments). In vertebrates, its C-terminal F-ABD mediates actin binding, bundling and microtubule crosslinking (Bradley and Koleske, 2009). This has important consequences for cell adhesion, migration (Woodring et al., 2003), intracellular trafficking (Rotty et al., 2013), endocytosis (Lonskaya et al., 2013), autophagy (Yogalingam and Pendegast,

2008; Hebron et al., 2013b; Lonskaya et al., 2013). In *Drosophila*, D-abl signaling is linked to actin dynamics and cell adhesion (reviewed by Lanier and Gertler, 2000; Hernandez et al., 2004). Recent evidence indicates that the D-abl kinase signaling regulates the Golgi complex architecture in neurons (Kannan et al., 2014). These data suggest that some of the effects of c-Abl signaling may arise from alterations of protein trafficking and secretion (Kannan et al., 2014). Kinase-independent functions of c-Abl have been already described (Henkemeyer et al., 1990; Chen et al., 2006). Evidence supports a kinase-independent function of *Drosophila* Abl for axonal guidance outcomes (O'Donnell and Bashaw, 2013). These results are consistent with a model for stepwise scaffolding and kinase functions of Abl in cell motility (Lapetina et al., 2009). Likely in a stepwise manner c-Abl promotes (or prevents) the formation of diverse signaling platforms within the cell. Specific outcomes rely on the full catalytic competence of the Abl kinase. The latter is due to local enrichment and/or a concomitant allosteric binding/removal of activators/adaptors/co-inhibitors (as it occurs in the nucleus during apoptosis). Both local enrichment and expression/localization of binding partners (adaptors/co-inhibitors) depend from cellular context.

### ALLOSTERIC REGULATION OF c-Abl

The auto-inhibited conformation of c-Abl is controlled through SH3–SH2–linker unit as in the Src family tyrosine kinases. In c-Src, the SH2 domain interacts with the C-terminal tail phosphotyrosine residue (Y527). By contrast, in c-Abl, the SH2 domain interacts more intimately with the large C-terminal lobe of the kinase domain (SH1). Interestingly, the tight interactions of the SH2–SH1 domain are induced by the binding of the myristoylated residues of the N-terminal region into a hydrophobic pocket of the kinase (Figure 1; Nagar et al., 2003). c-Abl requires the N-terminal myristoyl group (only present in Abl1b variant) to help the proper SH3–SH2-linker docking and inhibition (Iacob et al., 2011; Corbi-Verge et al., 2013; de Oliveira et al., 2013). Allosteric inhibitory interactions for the Abl1a variant are still poorly understood. Such interactions likely involve the binding of other inhibitors/adaptors. Small molecule compounds (GNF-2 and GNF-5) targeting the myristate pocket in the C-lobe of the



**FIGURE 1 | Schematics of the functional domains in c-Abl.** SH3, Src homology 3 domain; SH2, Src homology 2 domain. Y245 is in the linker region connecting SH2 domain with the kinase domain. Y412 is in the activation loop of the kinase domain. Phosphorylation in these two sites may prevent the auto-inhibited conformation of the c-Abl.

kinase domain do act as allosteric c-Abl inhibitors (Adrian et al., 2006, Fabbro et al., 2010, Zhang et al., 2010). The relevance of the myristoyl-binding pocket is further reinforced by the recent discovery of small-molecule c-Abl activators that dock into the same site (Yang et al., 2011, Hong et al., 2014). Upon c-Abl activation and removal of the allosteric interactions, the SH2 domain interacts with the N-terminal lobe of the kinase domain by using different surfaces of the SH2 domain (Hantschel, 2012). Compelling evidence indicates that the SH2 domain acts as a positive allosteric activator via the formation of an internal interface with the N-terminal lobe of kinase domain (Hantschel, 2012). Of note, the positioning of the SH2 domain facilitates multisite phosphorylation of substrates by c-Abl (Filippakopoulos et al., 2008, Grebien et al., 2011). However, alternative active states of c-Abl that do not require the SH2/kinase interface to function may occur when local clustering of c-Abl kinase core is sufficient for triggering transphosphorylation of the activation loop. In these circumstances the SH2 domain displacement from the back of the kinase domain is dispensable (Panjarian et al., 2013a,b). In short, the multidomain kinases like c-Abl can assume various conformational states and take more than one path to activation.

## EMERGING CONCEPTS FROM THE SOLUTION CONFORMATIONS OF c-Abl

Recent structural studies using NMR in combination with small angle X-ray scattering (SAXS) of a c-Abl fragment (SH3–SH2–linker–SH1 domains) provide the first structural information of apo form of c-Abl in the absence of small molecule inhibitors (Skora et al., 2013). The apo form of c-Abl adopts the “closed” conformation with the SH3–SH2 regulatory unit engaged with the kinase domain. However, addition of Imatinib (an ATP-competitive inhibitor) induces both a large structural rearrangement of the kinase domain and the detachment of the SH3–SH2 regulatory unit from the kinase domain leading to the formation of an “open” inactive state, where the ATP binding site is not accessible. In contrast to Imatinib, addition of the myristoyl pocket ligand GNF-5 to apo c-Abl induces only limited local changes around the myristoyl-binding pocket and keeps the SH3–SH2 regulatory unit in the “closed” state. Addition of GNF-5 to the “open” inactive state (c-Abl in complex with Imatinib) restores the “closed” inactive conformation (Skora et al., 2013). Under physiological conditions the “open” and “closed” conformations of c-Abl may be in equilibrium, which can be altered by the presence of specific inhibitors (ATP-competitive and/or allosteric ones).

It has been proposed that the ABD may stabilize the auto-inhibited conformation of the kinase by binding to F-actin (Woodring et al., 2003). Interestingly, the inhibitory effect of F-actin requires the SH2-kinase domain interaction to maintain the auto-inhibited conformation (Woodring et al., 2003). Small molecule inhibitors may induce a structural remodeling of the auto-inhibited conformation. This in turn may perturb domain interactions and consequently impinge on adaptor/effector/substrate binding, modulating c-Abl signaling dynamics. Cells treated with Imatinib show a profound change in the shape and a more rapid migration when plated on collagen-coated substrates (Chen et al., 2013). GNF-2 promotes a

translocation of c-Abl to the endoplasmic reticulum (Choi et al., 2009). It is tempting to consider that GNF-2 promotes a dynamic recruitment of c-Abl into different subcellular compartments and protein contexts. A timely relocation of c-Abl/GNF-2 complex in a specific subcellular compartment may induce *per se* a signaling circuitry.

An unproductive conformation (“open” inactive state) of the catalytic site induced by Imatinib may promote the availability of the regulatory domains (SH3–SH2–linker–ABD) with profound effects on c-Abl-interactome.

Molecular switches like c-Abl have modular domains required for their assembly into multiprotein complexes. Yet, c-Abl has also a modifier/kinase domain to regulate scaffold/signaling dynamics. The challenge now is to understand how such a complex signaling assembly is regulated in time and space (Hossain et al., 2012). Of note, both local enrichment and expression/localization of binding partners (adaptors/co-inhibitors) are dictated from the cell type and signaling context.

## ABERRANT c-Abl SIGNALING

The c-Abl kinase was early discovered as the oncogene in the Abelson murine leukemia virus (Goff et al., 1980) and then associated with human leukemias (Ben-Neriah et al., 1986). Several reports have indicated that c-Abl is a substrate and an activator of RTK. This bidirectional activation contributes to robust and persistent RTK signaling (Bromann et al., 2004). Cancer cells, expressing high levels of c-Abl, become dependent from its catalytic activity for growth and viability (reviewed in Greuber et al., 2013). On the other hand, in neurons aberrant c-Abl activation causes hyperphosphorylation, misfolding, and protein aggregation of tau protein, or alpha-synuclein. Such effects are considered hallmarks of neurodegenerative diseases (Ciccone et al., 2013, Tenreiro et al., 2014).

## c-Abl SIGNALING MEETS UBIQUITIN-MEDIATED RESPONSE

Cell signaling relies on PTMs for its regulation. The interplay and the crosstalk between phosphorylation and ubiquitination represent a recurrent theme in cell signaling (Hunter, 2007). We have discussed about some connections occurring between c-Abl phosphorylation and ubiquitin-mediated signaling in DNA damage response (Maiani et al., 2011). Kinase phosphorylation/activation often triggers ubiquitination. Activated forms of c-Abl are more unstable than wild-type (Echarri and Pendergast, 2001).

Compelling evidence indicates that c-Abl modulates the degradation of two proteins implicated in the pathogenesis of Parkinson's disease (Mahul-Mellier et al., 2014). A specific inhibitor of c-Abl like Nilotinib, used for leukemia treatment, promotes autophagic degradation of  $\alpha$ -synuclein while protecting neurons (Hebron et al., 2013a, Lonskaya et al., 2014). Interestingly, Nilotinib-induced autophagic changes increase endogenous Parkin level and ubiquitination, favoring amyloid clearance (Lonskaya et al., 2014). Taken together the data indicate that small molecule c-Abl inhibitors (ATP-competitive) may modulate the interplay between c-Abl and ubiquitin-mediated signaling. Convincing evidence indicates that c-Abl-mediated phosphorylation directly regulates the activity of some substrate E3 ligases (Zuckerman et al., 2009, Chan et al., 2013). In addition, negative regulation of E3 ligase activity by c-Abl can require a specific



recruitment of c-Abl into a complex with adaptor molecules, necessary for proper localization (Skouloudaki and Walz, 2012). Of note, tyrosine phosphorylation of Parkin on Y143 inhibits its E3 ligase activity leading to an accumulation of Parkin substrates (Ko et al., 2010). On the contrary, Imatinib treatment restores the E3 ligase activity of Parkin and its protective function. Interestingly, administration of Nilotinib has reduced the c-Abl activation and the levels of the Parkin substrate (PARIS) without preventing tyrosine phosphorylation of Parkin and accumulation of the Parkin substrate AIMP2. This suggests that the protective effect of Nilotinib may be in part Parkin-independent or related to the pharmacodynamics properties of Nilotinib (Karuppagounder et al., 2014). Of note, Nilotinib belongs to a second generation of inhibitor, ATP-competitive like Imatinib. It is more potent (>20-fold), exhibiting activity toward the majority of Imatinib-resistant mutations. The majority of Nilotinib–c-Abl interactions overlap with those described in the Imatinib–c-Abl complex. However, Nilotinib binds many of mutants resistant to Imatinib and is less sensitive to mutations of the C-lobe of the kinase (Reddy and Aggarwal, 2012). This may reflect that Nilotinib induces a dynamic rearrangement in the catalytic domain, attenuating the effects of mutations on the auto-inhibited conformation.

## CONCLUSION

Compelling evidence indicates the physiological relevance of the interface between c-Abl signaling and stress response, metabolic regulation mediated by transcription factors (Gonfloni et al., 2012). A small molecule that binds to the myristate binding pocket in the C-lobe of the Abl kinase was shown to inhibit Bcr-Abl (Zhang et al., 2010). This result indicates a functional connection between the myristate pocket and the kinase active site. Data from hydrogen exchange mass spectrometry indicate that binding of GNF-2/GNF-5 induces a dynamic conformation of residues near Thr315 that allows ATP-competitive inhibitors to tolerate the isoleucine at this position (Zhang et al., 2010). However, it remains still elusive how changes in the myristate pocket are communicated to the ATP binding site of the kinase. Molecular dynamic (MD) simulations (Fallacara et al., 2014) and emerging evidence from recent structural studies using NMR indicate that c-Abl may assume different conformational states in presence with different small molecule inhibitors (Skora et al., 2013). Together these data indicate that GNF-2 binding induces a more compact conformation of SH2-kinase domain interface of c-Abl. Therefore allosteric ligands for myristoyl pocket may be valuable tools for tackling the interface of c-Abl signaling. They could represent a way to attenuate the enzymatic activity while impinging on critical SH2 domain interactions. This in turn may rewire downstream signaling circuits and/or pathways. Recent evidence on the use of GNF-2 *in vivo* supports such a model (Maiani et al., 2012). c-Abl interacts with a large number of proteins (up to now more than 100) most of them are also c-Abl substrates (85%; Colicelli, 2010), in line with the idea that substrates can work as allosteric activators of the kinase. A better understanding of the spatiotemporal regulation of c-Abl signaling may allow us to modulate c-Abl signaling into specific subcellular compartments, with important consequences for cell homeostasis. These studies will take advantage from small allosteric compounds/activators as

tools to investigate the biological functions of c-Abl (Hong et al., 2014). Surely, a detailed understanding of c-Abl functions will help to develop combined targeted therapies in order to rewire the physiological regulatory circuits in cancer cells and in aged neurons.

## ACKNOWLEDGMENTS

This work was supported by AIRC grant (IG grant 2011 no. 11344) to Stefania Gonfloni. I thank all my collaborators (past and present) for stimulating discussions.

## REFERENCES

- Adrian, F. J., Ding, Q., Sim, T., Velentza, A., Sloan, C., Liu, Y., et al. (2006). Allosteric inhibitors of Bcr-abl-dependent cell proliferation. *Nat. Chem. Biol.* 2, 95–102. doi: 10.1038/nchembio760
- Ben-Neriah, Y., Daley, G. Q., Mes-Masson, A. M., Witte, O. N., and Baltimore, D. (1986). The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. *Science* 233, 212–214. doi: 10.1126/science.3460176
- Bradley, W. D., and Koleske, A. J. (2009). Regulation of cell migration and morphogenesis by Abl-family kinases: emerging mechanisms and physiological contexts. *J. Cell Sci.* 122, 3441–3454. doi: 10.1242/jcs.039859
- Bromann, P. A., Korkaya, H., and Courtneidge, S. A. (2004). The interplay between Src family kinases and receptor tyrosine kinases. *Oncogene* 23, 7957–7968. doi: 10.1038/sj.onc.1208079
- Chan, A. L., Grossman, T., Zuckerman, V., Campigli Di Giammartino, D., Moshel, O., Scheffner, M., et al. (2013). c-Abl phosphorylates E6AP and regulates its E3 ubiquitin ligase activity. *Biochemistry* 52, 3119–3129. doi: 10.1021/bi301710c
- Chen, X., Zhang, J., Lee, J., Lin, P. S., Ford, J. M., Zheng, N., et al. (2006). A kinase-independent function of c-Abl in promoting proteolytic destruction of damaged DNA binding proteins. *Mol. Cell* 22, 489–499. doi: 10.1016/j.molcel.2006.04.021
- Chen, Z., Lessey, E., Berginski, M. E., Cao, L., Li, J., Trepatt, X., et al. (2013). Gleevec, an Abl family inhibitor, produces a profound change in cell shape and migration. *PLoS ONE* 8:e52233. doi: 10.1371/journal.pone.0052233
- Choi, Y., Seeliger, M. A., Panjarian, S. B., Kim, H., Deng, X., Sim, T., et al. (2009). N-myristoylated c-Abl tyrosine kinase localizes to the endoplasmic reticulum upon binding to an allosteric inhibitor. *J. Biol. Chem.* 284, 29005–29014. doi: 10.1074/jbc.M109.026633
- Ciccone, S., Maiani, E., Bellusci, G., Diederich, M., and Gonfloni, S. (2013). Parkinson's disease: a complex interplay of mitochondrial DNA alterations and oxidative stress. *Int. J. Mol. Sci.* 14, 2388–2409. doi: 10.3390/ijms14022388
- Colicelli, J. (2010). ABL tyrosine kinases: evolution of function, regulation, and specificity. *Sci. Signal.* 3, re6. doi: 10.1126/scisignal.3139re6
- Corbi-Verge, C., Marinelli, E., Zafra-Ruano, A., Ruiz-Sanz, J., Luque, I., and Faraldo-Gomez, J. D. (2013). Two-state dynamics of the SH3–SH2 tandem of Abl kinase and the allosteric role of the N-cap. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3372–E3380. doi: 10.1073/pnas.1303966110
- de Oliveira, G. A., Pereira, E. G., Ferretti, G. D., Valente, A. P., Cordeiro, Y., and Silva, J. L. (2013). Intramolecular dynamics within the N-Cap–SH3–SH2 regulatory unit of the c-Abl tyrosine kinase reveal targeting to the cellular membrane. *J. Biol. Chem.* 288, 28331–28345. doi: 10.1074/jbc.M113.500926
- Echarri, A., and Pendergast, A. M. (2001). Activated c-Abl is degraded by the ubiquitin-dependent proteasome pathway. *Curr. Biol.* 11, 1759–1765. doi: 10.1016/S0960-9822(01)00538-3
- Ermolaeva, M. A., Segref, A., Dakhovnik, A., Ou, H. L., Schneider, J. I., Utermohlen, O., et al. (2013). DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature* 501, 416–420. doi: 10.1038/nature12452
- Fabbro, D., Manley, P. W., Jahnke, W., Liebetanz, J., Szyttenholm, A., Fendrich, G., et al. (2010). Inhibitors of the Abl kinase directed at either the ATP- or myristate-binding site. *Biochim. Biophys. Acta* 1804, 454–462. doi: 10.1016/j.bbapap.2009.12.009
- Fallacara, A. L., Tintori, C., Radi, M., Schenone, S., and Botta, M. (2014). Insight into the allosteric inhibition of Abl kinase. *J. Chem. Inf. Model.* 54, 1325–1338. doi: 10.1021/ci500060k
- Filippakopoulos, P., Kofler, M., Hantschel, O., Gish, G. D., Grebien, F., Salah, E., et al. (2008). Structural coupling of SH2-kinase domains links Fes and

- Abl substrate recognition and kinase activation. *Cell* 134, 793–803. doi: 10.1016/j.cell.2008.07.047
- Gartner, A., and Akay, A. (2013). Stress response: anything that doesn't kill you makes you stronger. *Curr. Biol.* 23, R1012–R1014. doi: 10.1016/j.cub.2013.09.036
- Goff, S. P., Gilboa, E., Witte, O. N., and Baltimore, D. (1980). Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. *Cell* 22, 777–785. doi: 10.1016/0092-8674(80)90554-1
- Gonfloni, S., Maiani, E., Di Bartolomeo, C., Diederich, M., and Cesareni, G. (2012). Oxidative stress, DNA damage, and c-Abl signaling: at the crossroad in neurodegenerative diseases? *Int. J. Cell Biol.* 2012, 683097. doi: 10.1155/2012/683097
- Gonfloni, S., Weijland, A., Kretzschmar, J., and Superti-Furga, G. (2000). Crosstalk between the catalytic and regulatory domains allows bidirectional regulation of Src. *Nat. Struct. Biol.* 7, 281–286. doi: 10.1038/74041
- Grebien, F., Hantschel, O., Wojcik, J., Kaup, I., Kovacic, B., Wyrzucki, A. M., et al. (2011). Targeting the SH2-kinase interface in Bcr-Abl inhibits leukemogenesis. *Cell* 147, 306–319. doi: 10.1016/j.cell.2011.08.046
- Greuber, E. K., Smith-Pearson, P., Wang, J., and Pendergast, A. M. (2013). Role of ABL family kinases in cancer: from leukaemia to solid tumours. *Nat. Rev. Cancer* 13, 559–571. doi: 10.1038/nrc3563
- Hantschel, O. (2012). Structure, regulation, signaling, and targeting of abl kinases in cancer. *Genes Cancer* 3, 436–446. doi: 10.1177/1947601912458584
- Hantschel, O., and Superti-Furga, G. (2004). Regulation of the c-Abl and Bcr-Abl tyrosine kinases. *Nat. Rev. Mol. Cell Biol.* 5, 33–44. doi: 10.1038/nrm1280
- Hebron, M. L., Lonskaya, I., and Moussa, C. E. (2013a). Nilotinib reverses loss of dopamine neurons and improves motor behavior via autophagic degradation of alpha-synuclein in Parkinson's disease models. *Hum. Mol. Genet.* 22, 3315–3328. doi: 10.1093/hmg/ddt192
- Hebron, M. L., Lonskaya, I., and Moussa, C. E. (2013b). Tyrosine kinase inhibition facilitates autophagic SNCA/alpha-synuclein clearance. *Autophagy* 9, 1249–1250. doi: 10.4161/auto.25368
- Henkemeyer, M., West, S. R., Gertler, F. B., and Hoffmann, F. M. (1990). A novel tyrosine kinase-independent function of Drosophila abl correlates with proper subcellular localization. *Cell* 63, 949–960. doi: 10.1016/0092-8674(90)90498-4
- Hernandez, S. E., Krishnaswami, M., Miller, A. L., and Koleske, A. J. (2004). How do Abl family kinases regulate cell shape and movement? *Trends Cell Biol.* 14, 36–44. doi: 10.1016/j.tcb.2003.11.003
- Hong, X., Cao, P., Washio, Y., Simpson, G., Campobasso, N., Yang, J., et al. (2014). Structure-guided optimization of small molecule c-Abl activators. *J. Comput. Aided Mol. Des.* 28, 75–87. doi: 10.1007/s10822-014-9731-5
- Hossain, S., Dubielecka, P. M., Sikorski, A. F., Birge, R. B., and Kotula, L. (2012). Crk and ABI1: binary molecular switches that regulate abl tyrosine kinase and signaling to the cytoskeleton. *Genes Cancer* 3, 402–413. doi: 10.1177/1947601912460051
- Hunter, T. (2007). The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* 28, 730–738. doi: 10.1016/j.molcel.2007.11.019
- Jacob, R. E., Zhang, J., Gray, N. S., and Engen, J. R. (2011). Allosteric interactions between the myristate- and ATP-site of the Abl kinase. *PLoS ONE* 6:e15929. doi: 10.1371/journal.pone.0015929
- Ito, Y., Pandey, P., Mishra, N., Kumar, S., Narula, N., Kharbanda, S., et al. (2001). Targeting of the c-Abl tyrosine kinase to mitochondria in endoplasmic reticulum stress-induced apoptosis. *Mol. Cell Biol.* 21, 6233–6242. doi: 10.1128/MCB.21.18.6233-6242.2001
- Kannan, R., Kuzina, I., Wincovitch, S., Nowotarski, S. H., and Giniger, E. (2014). The Abl/enabled signaling pathway regulates Golgi architecture in *Drosophila* photoreceptor neurons. *Mol. Biol. Cell* 25, 2993–3005. doi: 10.1091/mbc.E14-02-0729
- Karuppagounder, S. S., Brahmachari, S., Lee, Y., Dawson, V. L., Dawson, T. M., and Ko, H. S. (2014). The c-Abl inhibitor, nilotinib, protects dopaminergic neurons in a preclinical animal model of Parkinson's disease. *Sci. Rep.* 4, 4874. doi: 10.1038/srep04874
- Ko, H. S., Lee, Y., Shin, J. H., Karuppagounder, S. S., Gadad, B. S., Koleske, A. J., et al. (2010). Phosphorylation by the c-Abl protein tyrosine kinase inhibits Parkinson's ubiquitination and protective function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16691–16696. doi: 10.1073/pnas.1006083107
- Kuriyan, J., and Cowburn, D. (1997). Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* 26, 259–288. doi: 10.1146/annurev.biophys.26.1.259
- Kuriyan, J., and Eisenberg, D. (2007). The origin of protein interactions and allostery in colocalization. *Nature* 450, 983–990. doi: 10.1038/nature06524
- Lanier, L. M., and Gertler, F. B. (2000). From Abl to actin: Abl tyrosine kinase and associated proteins in growth cone motility. *Curr. Opin. Neurobiol.* 10, 80–87. doi: 10.1016/S0959-4388(99)00058-6
- Lapetina, S., Mader, C. C., Machida, K., Mayer, B. J., and Koleske, A. J. (2009). Arg interacts with cortactin to promote adhesion-dependent cell edge protrusion. *J. Cell Biol.* 185, 503–519. doi: 10.1083/jcb.200809085
- Liu, J., and Nussinov, R. (2013). The role of allostery in the ubiquitin-proteasome system. *Crit. Rev. Biochem. Mol. Biol.* 48, 89–97. doi: 10.3109/10409238.2012.742856
- Lonskaya, I., Hebron, M. L., Desforges, N. M., Franjie, A., and Moussa, C. E. (2013). Tyrosine kinase inhibition increases functional parkin-Beclin-1 interaction and enhances amyloid clearance and cognitive performance. *EMBO Mol. Med.* 5, 1247–1262. doi: 10.1002/emmm.201302771
- Lonskaya, I., Hebron, M. L., Desforges, N. M., Schachter, J. B., and Moussa, C. E. (2014). Nilotinib-induced autophagic changes increase endogenous parkin level and ubiquitination, leading to amyloid clearance. *J. Mol. Med. (Berl.)* 92, 373–386. doi: 10.1007/s00109-013-1112-3
- Mahul-Mellier, A. L., Fauvet, B., Gysbers, A., Dikiy, I., Oueslati, A., Georgeon, S., et al. (2014). c-Abl phosphorylates alpha-synuclein and regulates its degradation: implication for alpha-synuclein clearance and contribution to the pathogenesis of Parkinson's disease. *Hum. Mol. Genet.* 23, 2858–2879. doi: 10.1093/hmg/ddt674
- Maiani, E., Di Bartolomeo, C., Klinger, F. G., Cannata, S. M., Bernardini, S., Chateauvieux, S., et al. (2012). Reply to: cisplatin-induced primordial follicle oocyte killing and loss of fertility are not prevented by imatinib. *Nat. Med.* 18, 1172–1174. doi: 10.1038/nm.2852
- Maiani, E., Diederich, M., and Gonfloni, S. (2011). DNA damage response: the emerging role of c-Abl as a regulatory switch? *Biochem. Pharmacol.* 82, 1269–1276. doi: 10.1016/j.bcp.2011.07.001
- Miller, A. L., Wang, Y., Mooseker, M. S., and Koleske, A. J. (2004). The Abl-related gene (Arg) requires its F-actin-microtubule cross-linking activity to regulate lamellipodial dynamics during fibroblast adhesion. *J. Cell Biol.* 165, 407–419. doi: 10.1083/jcb.200308055
- Nagar, B., Hantschel, O., Young, M. A., Scheffzek, K., Veach, D., Bornmann, W., et al. (2003). Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* 112, 859–871. doi: 10.1016/S0092-8674(03)00194-6
- O'Donnell, M. P., and Bashaw, G. J. (2013). Distinct functional domains of the Abelson tyrosine kinase control axon guidance responses to Netrin and Slit to regulate the assembly of neural circuits. *Development* 140, 2724–2733. doi: 10.1242/dev.093831
- Panjarian, S., Jacob, R. E., Chen, S., Engen, J. R., and Smithgall, T. E. (2013a). Structure and dynamic regulation of Abl kinases. *J. Biol. Chem.* 288, 5443–5450. doi: 10.1074/jbc.R112.438382
- Panjarian, S., Jacob, R. E., Chen, S., Wales, T. E., Engen, J. R., and Smithgall, T. E. (2013b). Enhanced SH3/linker interaction overcomes Abl kinase activation by gatekeeper and myristic acid binding pocket mutations and increases sensitivity to small molecule inhibitors. *J. Biol. Chem.* 288, 6116–6129. doi: 10.1074/jbc.M112.431312
- Reddy, E. P., and Aggarwal, A. K. (2012). The ins and outs of bcr-abl inhibition. *Genes Cancer* 3, 447–454. doi: 10.1177/1947601912462126
- Rotty, J. D., Wu, C., and Bear, J. E. (2013). New insights into the regulation and cellular functions of the ARP2/3 complex. *Nat. Rev. Mol. Cell Biol.* 14, 7–12. doi: 10.1038/nrm3492
- Seet, B. T., Dikic, I., Zhou, M. M., and Pawson, T. (2006). Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* 7, 473–483. doi: 10.1038/nrm1960
- Sirvent, A., Benistant, C., and Roche, S. (2008). Cytoplasmic signalling by the c-Abl tyrosine kinase in normal and cancer cells. *Biol. Cell* 100, 617–631. doi: 10.1042/BC20080020
- Skora, L., Mestan, J., Fabbro, D., Jahnke, W., and Grzesiek, S. (2013). NMR reveals the allosteric opening and closing of Abelson tyrosine kinase by ATP-site and myristoyl pocket inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 110, E4437–E4445. doi: 10.1073/pnas.1314712110
- Skouloudaki, K., and Walz, G. (2012). YAP1 recruits c-Abl to protect angiotensin-like 1 from Nedd4-mediated degradation. *PLoS ONE* 7:e35735. doi: 10.1371/journal.pone.0035735



- Tenreiro, S., Eckermann, K., and Outeiro, T. F. (2014). Protein phosphorylation in neurodegeneration: friend or foe? *Front. Mol. Neurosci.* 7:42. doi: 10.3389/fnmol.2014.00042
- Van Etten, R. A., Jackson, P. K., Baltimore, D., Sanders, M. C., Matsudaira, P. T., and Janney, P. A. (1994). The COOH terminus of the c-Abl tyrosine kinase contains distinct F- and G-actin binding domains with bundling activity. *J. Cell Biol.* 124, 325–340. doi: 10.1083/jcb.124.3.325
- Wang, J. Y. (2014). The capable ABL: what is its biological function? *Mol. Cell. Biol.* 34, 1188–1197. doi: 10.1128/MCB.01454-13
- Wen, S. T., Jackson, P. K., and Van Etten, R. A. (1996). The cytostatic function of c-Abl is controlled by multiple nuclear localization signals and requires the p53 and Rb tumor suppressor gene products. *EMBO J.* 15, 1583–1595.
- Woodring, P. J., Hunter, T., and Wang, J. Y. (2003). Regulation of F-actin-dependent processes by the Abl family of tyrosine kinases. *J. Cell Sci.* 116, 2613–2626. doi: 10.1242/jcs.00622
- Yang, J., Campobasso, N., Biju, M. P., Fisher, K., Pan, X. Q., Cottom, J., et al. (2011). Discovery and characterization of a cell-permeable, small-molecule c-Abl kinase activator that binds to the myristoyl binding site. *Chem. Biol.* 18, 177–186. doi: 10.1016/j.chembiol.2010.12.013
- Yogalingam, G., and Pendergast, A. M. (2008). Abl kinases regulate autophagy by promoting the trafficking and function of lysosomal components. *J. Biol. Chem.* 283, 35941–35953. doi: 10.1074/jbc.M804543200
- Yoshida, K., Yamaguchi, T., Natsume, T., Kufe, D., and Miki, Y. (2005). JNK phosphorylation of 14-3-3 proteins regulates nuclear targeting of c-Abl in the apoptotic response to DNA damage. *Nat. Cell Biol.* 7, 278–285. doi: 10.1038/ncb1228
- Zhang, J., Adrian, F. J., Jahnke, W., Cowan-Jacob, S. W., Li, A. G., Iacob, R. E., et al. (2010). Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors. *Nature* 463, 501–506. doi: 10.1038/nature08675
- Zuckerman, V., Lenos, K., Popowicz, G. M., Silberman, I., Grossman, T., Marine, J. C., et al. (2009). c-Abl phosphorylates Hdmx and regulates its interaction with p53. *J. Biol. Chem.* 284, 4031–4039. doi: 10.1074/jbc.M809211200

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 June 2014; accepted: 25 October 2014; published online: 12 November 2014.

Citation: Gonfloni S (2014) Defying c-Abl signaling circuits through small allosteric compounds. *Front. Genet.* 5:392. doi: 10.3389/fgene.2014.00392

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Gonfloni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Modeling conformational transitions in kinases by molecular dynamics simulations: achievements, difficulties, and open challenges

Marco D'Abramo<sup>1,2\*</sup>, Neva Besker<sup>3</sup>, Giovanni Chillemi<sup>1</sup> and Alessandro Grottesi<sup>1</sup>

<sup>1</sup> CINECA, Rome, Italy

<sup>2</sup> Dipartimento di Chimica, Sapienza University of Rome, Rome, Italy

<sup>3</sup> Dipartimento di Scienze e Tecnologie Chimiche, Università di Roma "Tor Vergata," Rome, Italy

## Edited by:

Andreas Zanzoni, Inserm TAGC  
UMR1090, France  
Allegria Via, Sapienza University, Italy

## Reviewed by:

Miquel Pons, University of  
Barcelona, Spain  
Marco Mazzorana, Diamond Light  
Source Ltd., UK

## \*Correspondence:

Marco D'Abramo, Dipartimento di  
Chimica, Sapienza University of  
Rome, P.le A. Moro, 5, 00182  
Rome, Italy  
e-mail: marco.dabramo@uniroma1.it

Protein kinases work because their flexibility allows to continuously switch from inactive to active form. Despite the large number of structures experimentally determined in such states, the mechanism of their conformational transitions as well as the transition pathways are not easily to capture. In this regard, computational methods can help to shed light on such an issue. However, due to the intrinsic sampling limitations, much efforts have been done to model in a realistic way the conformational changes occurring in protein kinases. In this review we will address the principal biological achievements and structural aspects in studying kinases conformational transitions and will focus on the main challenges related to computational approaches such as molecular modeling and MD simulations.

**Keywords:** molecular dynamics, conformational transitions, kinases, computational modeling, conformational pathways

## INTRODUCTION

Protein kinases, one of the largest gene families in eukaryotes, are a class of enzymes which regulates several key cellular processes, such as cell growth and differentiation.

They all achieve their function by favoring particular structural arrangements, able to catalyze protein phosphorylation. Despite the high degree of the conservation of the overall fold, each kinase is able to respond to specific signals tuning its activity.

Therefore, these molecules represent a unique example of the nature's ability to introduce diversification through relatively small variations to a highly conserved motif.

Such variations, which rarely modify the optimized structure on the protein free-energy minimum, can severely affect protein flexibility which in turn might modify the transition patterns designed to be optimal for their function.

Protein kinases probably represent one of the most remarkable examples of the importance of the protein flexibility: in fact, they regulate—continuously switching from active to inactive state—multiple biological processes by posttranslational phosphorylation of serine, threonine and tyrosine residues. In this context, the knowledge of their structures, although mandatory, it is not sufficient for a complete understanding of their function, thus, the description of their conformational transition pathways is of paramount importance.

However, the experimental techniques able to accurately describe the transient/intermediate states populated between the conformational changes are still under development (Ban et al., 2011; Waldrop, 2014), thus, computational methods can greatly help in this case.

Atomistic Molecular Dynamics (MD) simulations—based on the classical mechanics laws and using well-refined empirical

force fields—are probably the most accurate computational technique for the study of protein flexibility, being able to describe the dynamical evolution of molecular systems at atomistic level of detail. Unfortunately, the time-scale reached by standard MD simulations is still limited by the computational power to the microsecond, thus excluding processes occurring on longer times, such as the conformational transitions of almost all kinases. Therefore, several theoretical-computational techniques able to enhance the conformational sampling were applied to model such a process.

Here, after presenting the structural features of protein kinases, we discuss some recent results obtained in modeling kinase conformational transitions by enhancing sampling techniques and brute-force approaches, highlighting their advantages and limits. Finally, due to the involvement of this protein family in very important diseases such as cancer, we briefly introduce the results reached in drug design using computational techniques able to provide a complete description of the accessible conformational space. We discuss their still full unexploited potential to further enhance the kinase inhibitor performance.

## KINASE STRUCTURE AND CONFORMATIONAL VARIABILITY

To get insights into the structure/dynamics/function relationship we need to describe their molecular structure in detail. Although the number of human protein kinase family members is large, the existing X-ray crystallographic studies showed that the three-dimensional (3D) structures of their catalytic domains are similar. Studies based on X-ray structures of PKA revealed that catalytic subunits of protein kinases share a conserved core consisting of two lobes. The N-terminal small lobe (N-lobe) and C-terminal large lobe (C-lobe). These two lobes host a deep pocket that

accommodates a molecule of ATP bound to one or two divalent cations: magnesium or manganese that compensate for the strong negative charge of the ATP phosphates (Kornev and Taylor, 2010). This special structural arrangement is required because the phosphoryl transfer reactions are highly dependent on the bond distances and charge distribution in the reaction transition state. In particular, structural studies performed with transition state mimics showed that the presence of Mg<sup>2+</sup> ions and positively charged groups in its vicinity would imply an associative mechanism for phosphoryl transfer, thus optimizing the overall reaction (Madhusudan et al., 2002). The N-lobe usually includes five  $\beta$ -strands and an  $\alpha$ -helix (hereafter called  $\alpha$ C-helix). Despite the fact that the  $\beta$ -strands form a relatively rigid antiparallel  $\beta$  sheet, the N-lobe is very dynamic and flexible. The C-lobe is mainly  $\alpha$ -helical (Figure 1) and includes the activation segment, a 20–35 residues stretch located between a conserved DFG motif, and the APE motif (Huse and Kuriyan, 2002; Nolen et al., 2004), that is highly conformationally variable and its conformation can influence both substrate binding and catalytic efficiency. The C-lobe serves as a docking site for substrate peptides/proteins. The N-terminal part of the peptide lies in a groove between the  $\alpha$ D and  $\alpha$ F-helices on one side and the  $\alpha$ G-helix on the other side.

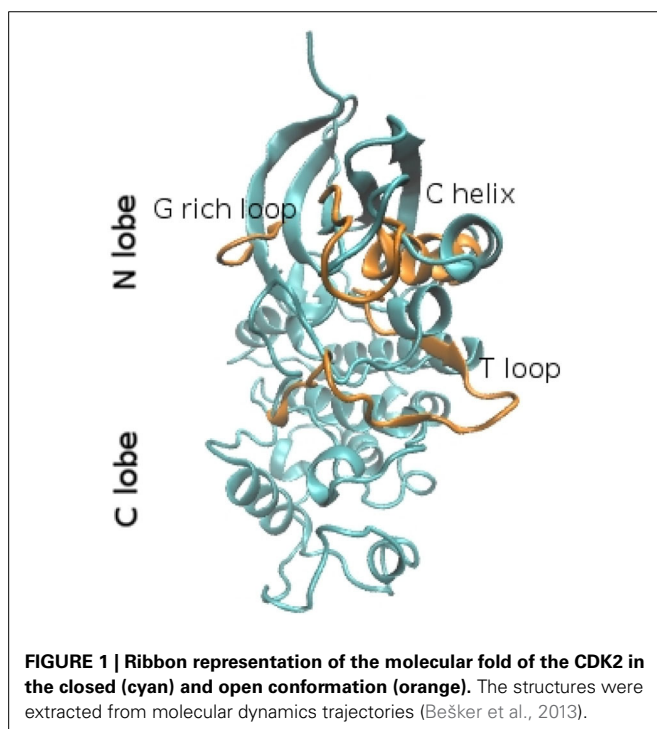
Bioinformatics approaches (Kornev et al., 2006) have shown that key hydrophobic residues in kinases have a strict organization around cores. In particular, the overall fold is joined by two hydrophobic non-contiguous motifs referred as “spines” (Kornev et al., 2008). The spines connect all critically important elements of the protein kinase molecule to a single hydrophobic helix ( $\alpha$ F) that provides their positioning in space. These motifs regulate the protein kinase activity. The regulatory spine is typically assembled after phosphorylation loop and the catalytic one is

completed by the ATP adenine ring. All these are anchored to the buried hydrophobic  $\alpha$ F helix of the C lobe and provide the basic architecture of the core.

The “activation loop” is a highly flexible loop, which adopts an extended conformation in the active kinase domain to function as a binding platform for the peptide substrate. The experimental structures of kinases show that  $\alpha$ C-helix (in the small lobe) and the activation loop (in the large lobe) adopt distinct conformations associated with the active or different inactive forms. This extended conformation is stabilized by phosphorylation, as observed in the active and phosphorylated kinase domain of the Src-family kinase Lck (Hantschel and Superti-Furga, 2004). In contrast to the active kinase, the inactive kinases are structurally diverse (Noble et al., 2004). This diversity arises because no catalytic requirements constrain the fold when it is inactive, with the presence of different inactive conformations that, nevertheless, share a number of common structural themes.

In many kinases  $\alpha$ C-helix can rotate and change its position, thereby altering the orientation of key catalytic residues. In the active state of EGFR kinase, the activation loop maintains a  $\beta$ 9 strand and an overall conformation compatible with substrate binding while the  $\alpha$ C-helix is adjacent to the ATP-binding site (the “ $\alpha$ C-in” conformation), and the catalytic important Asp831 residue of the conserved “DFG” motif is found within that site (“DFG-in”). That is, in the active state, the phenylalanine (Phe) side-chain occupies the ATP-binding pocket, and the aspartate (Asp) side-chain is located in the outside of the pocket. When the so-called “DFG-flip” occurs, the Asp and Phe residues swap: the Asp side-chain rotates into the ATP-binding pocket, and the Phe side-chain rotates out of the ATP-binding pocket (DFG-out conformation), leading the kinases to the inactive state (Shan et al., 2009). Some human kinases were shown to be able to adopt the DFG-out conformation (Debondt et al., 1993; Siccheri et al., 1997; Xu et al., 1997) and it was suggested that the DFG-in and DFG-out conformations are in dynamic equilibrium and can interconvert to determine the kinase function (Levinson et al., 2006).

The protein kinase fold may switch between an active or inactive conformation by extra domains or separate subunits. How the active and inactive states are stabilized and how states interconvert are key questions in understanding kinase regulation. Typically, the activation segment contains a serine/threonine or tyrosine residue that can be phosphorylated. This site is referred to as first “phosphorylation site”; some kinases can have up to three phosphorylation site in the activation loop and perform function specific for these kinases. Other kinases, called non-RD kinases, lack the conserved arginine residue in kinase subdomain and it has been observed that some of them do not autophosphorylate the activation loop and are either constitutively active or regulated through alternative mechanisms (Dardick and Ronald, 2006). Most protein kinases are comprised of more than just the kinase domain. This can be flanked by other domains which tend to be quite dynamic or are part of a multi-subunit complex such as the Cyclin-dependent Kinases (CDK) which require an activating cyclin or phosphorylase kinase. Others such as the receptor tyrosine kinases are anchored to membranes and often have long segments that tether them to the membrane as well



as long C-terminal tails. Certain protein kinases can be activated or inhibited by specific polypeptide cofactors. In CDK2 the activation depends on association with a cyclin subunit via the interaction of the  $\alpha$ C-helix with specific hydrophobic residues on cyclin that enable the active site to be accessible for phosphorylation (Lowe et al., 1997; Lolli, 2010). Many protein kinases dimerize as part of their activation mechanism, and dimerization can be regarded as a special case of kinase activation by accessory proteins or domains. In such cases, either both partners are activated by reciprocal phosphorylation or one partner (the activator kinase) activates the other (the receiver kinase) through an allosteric mechanism (Endicott et al., 2012).

## COMPUTATIONAL STUDIES: AN OVERVIEW

Computational works focusing on kinase conformational transitions can be divided in two groups: those based on brute-force approach, where the system was simulated for microseconds and those where the dynamical evolution of the system was altered to speed up the sampling. Using the former approach and a special-purpose supercomputer, the group of D.E. Shaw described spontaneous transitions in the microsecond time-scale for the kinase domain of the EFGR from the active to the so-called “Src-like inactive” conformation by way of two sets of intermediate conformations, significantly different from the existing crystal structures. Interestingly, in one of them the helical arrangement of the activation loop (A-loop) leaves the well-known phosphorylation site exposed (Shan et al., 2013). Similarly, Yang et al. used several independent MD trajectories starting from the putative intermediate conformations to reconstruct the transitions pathways between active and inactive states for the Src kinase and identified the concerted conformational rearrangements in terms of structural regions of the protein, such as the A-loop and the  $\alpha$ C-helix (Yang et al., 2009). They also found that transient structures in not fully active conformations allow exposure of Tyr416 to the bulk solvent—the residue whose phosphorylation locks the Src in the active state—a feature typical of the fully active states. In an extensive application of the Replica Exchange MD—an enhancing sampling technique favoring the overcome of potential energy barriers—Huang et al. found that the  $\alpha$ C-helix acts as an energy switch between the active and inactive conformations and describe the sequence of events along the optimal inactivation path as the fold of the A-loop into the active site followed by the  $\alpha$ C-helix movement (Huang et al., 2012).

Due to the large computational requirements, brute-force approaches are still limited to few research groups and inevitably to few systems. Therefore, a class of theoretical-computational methods able to sample the conformational rearrangements at limited computational costs were applied to protein kinases. Perturbative approaches, which are likely to be the most powerful techniques to model the free energy of a chemical reaction, are not particularly suited in this case, mainly because it is very difficult to find a proper reaction coordinate in a high-dimensional space able to fully discriminate the two conformational states. Three of such methods, the Targeted, Steered and Biased MD were applied to model the conformational transitions between the active/inactive forms of the catalytic domain of LYN kinase, a member of the Src family of protein tyrosine kinases: the results

showed that although the transition pathways described by these approaches are similar, the path is strongly dependent on the choice of progress variable (Huang et al., 2009).

One of the most interesting approaches, however, is based on the possibility to iteratively select the conformations sampled by unbiased MD which are most likely to be productive, in order to limit the time spent in sampling the conformational space regions within the free-energy minimum basin. In this regard, Beckstein et al. applied the RMSD as a measure to apply their Dynamic IMportance Sampling (DIMS) algorithm to describe the zipping and unzipping of adenylate kinase (Beckstein et al., 2009). From their results, a cooperative salt bridge zipper is hypothesized to be the rate-limiting step of the apo-AdK conformational transition.

Following a similar approach, our group applied the Essential Dynamics algorithm to the description of the conformational transitions in the CDK2 (Bešker et al., 2013). The power of such a technique is due to the fact the direction in the high-dimensional space is described by the native protein movements in its stable conformational states, e.g., open and closed. In this case, the choice of the progress variable(s) is natural and self-defined by the unbiased MD simulations of the starting and target state. In the case of CDK2, we found that both the opening and closure follow common transition paths in the essential subspace, involving the same structural determinants (i.e., the  $\alpha$ C-helix and the activation loop) but in opposite order, indicating that the time-sequence of such rearrangements determines the specific protein conformational transition. It is worth noting that due to their high computational efficiency, the last two methods can be used to fill the gap between single-system and system-wide studies, thus providing an invaluable source of information at the protein family level.

All the above methods, with their advantages and shortcomings, represent an intriguing field of research and, in the case of the kinase conformational transitions, they have contributed to a better description of the transition pathways between the active and inactive forms.

Nevertheless, several important aspects from a biochemical point of view are still not well-characterized in detail, including for example the possibility to bridge the gap between the atomic description of the transitions and the protein activity and/or the role played by possible biological partners. In the next few years, with the support of new or refined enhancing sampling algorithm and the fast growth of the computational power, we envision that a better modeling of large conformational rearrangements of biomolecules, as those occurring in kinases, will be within range of the scientific community.

## INHIBITORS

The central biological role of protein kinases leads to their frequent alteration in several pathologies, first of all cancer. It is not surprising, therefore, that they are major targets for therapy. The great number of family members and their high degree of conservation, however, easily lead to multi target inhibition and therefore important side effects. First-generation CDK inhibitors, for instance, targeted the non-specific ATP-binding site and were discarded in the pre-clinical phase because of their toxic effects (Lapenna and Giordano, 2009).



The importance of kinase structural flexibility on the rational drug design of their inhibitor can be appreciate looking at the experience of imatinib, the first Bcr-Abl inhibitor approved by the US Food and Drug Administration that revolutionized treatment of Chronic Myelogenous Leukaemia (CML). At the time of this drug discovery, the structural knowledge on the kinase family was not deep enough to rationalize its mode of action, but several subsequent studies demonstrated that imatinib exploits the change of conformation of the DFG-motif in the unphosphorylated inactive form (Lambert et al., 2013). These second generation inhibitors, targeting a variable region in the inactive form, are more capable of discerning among kinases but at the same time have the Achilles' heel of a greatest rate of drug-resistant mutations (Noble et al., 2004).

First generation inhibitors, in fact, targeting the active conformation, seldom developed drug-resistant mutations because they would have compromised the kinase activity. The war against cancer, therefore, moved another step, going against imatinib resistant cancers.

A growing number of experimentally solved kinase structures (more than 5000 in PDB database), both in active and inactive forms, and several MD simulations at microsecond-timescale have changed our knowledge on this protein family. These deep roots made possible the successful application of *in silico* rational drug design in a growing number of cases: (1) the so-called hybrid-design was carried out, in which the same kinase inhibitor targets different protein regions, thus merging the characteristic of potency (typical of first generation inhibitors) and selectivity (obtained by second generation ones) (Albaugh et al., 2012); (2) imatinib was also reengineered to reduce its cardiotoxicity, by making use of a different water propensity of two residues in the principal and secondary target kinases (Fernandez et al., 2007). Several MD simulations were carried out at this scope, that were able to discriminate the dynamic and structural characteristics of Bcr-Abl and C-Abl kinases, this last responsible for the cardiotoxic effects; (3) systematic rigid body docking of potential inhibitors against 84 unique protein kinases identified three derivatives of indirubin (Zahler et al., 2007), one of which has been recently indicated as particularly active against cancer metastasis (Braig et al., 2013).

The research on better kinase inhibitors by computational methods has focused not only on the target but on the ligand as well, with the so-called Pharmacophore Modeling approach, i.e., “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response” (Wermuth et al., 1998). Modern drug design combines more computational approach, as in the case of the identification of novel inhibitors against thymidine monophosphate kinase of *M. tuberculosis* with potent antitubercular activity, obtained through the combination of pharmacophore modeling, docking simulation, and structure interaction fingerprint analysis (Kumar et al., 2009).

It is easy to predict that these coming years will still be rich of new kinase inhibitors with improved selectivity and mutant resistance. In fact, new targets, coming from the computational and experimental characterization of kinase intermediates

are combined with improved computer-aided drug design methods.

## CONCLUSIONS AND PERSPECTIVES

As briefly highlighted in this review, the intrinsic flexibility of the kinases represents a very intense field of research. The recent developments in theoretical-computational modeling are contributing to elucidate the actual dynamical behavior of the kinase domains, thus leading to the possibility of enhancing the design of new and selective inhibitors using a proper atomistic description of their conformational basins and the pathways along which they continuously move to tune their activity. Finally, the growth of computational power as well as the development of new theoretical treatments able to model slow and/or rare events at limited computational costs are making it possible to switch from single-system modeling of the kinase dynamical behavior to a system-wide modeling.

## ACKNOWLEDGMENTS

This work was supported by MIUR-FIRB grant (n. RBF12BGHO) and “Rita Levi Montalcini” research program.

## REFERENCES

- Albaugh, P., Fan, Y., Mi, Y., Sun, F., Adrian, F., Li, N., et al. (2012). Discovery of GNF-5837, a Selective TRK Inhibitor with efficacy in rodent cancer tumor models. *ACS Med. Chem. Lett.* 3, 140–145. doi: 10.1021/ml200261d
- Ban, D., Funk, M., Gulich, R., Egger, D., Sabo, T. M., Walter, K. F. A., et al. (2011). Kinetics of conformational sampling in ubiquitin. *Angew. Chem. Int. Ed.* 50, 11437–11440. doi: 10.1002/anie.201105086
- Beckstein, O., Denning, E. J., Perilla, J. R., and Woolf, T. B. (2009). Zipping and unzipping of adenylate kinase: atomistic insights into the ensemble of open <-> closed transitions. *J. Mol. Biol.* 394, 160–176. doi: 10.1016/j.jmb.2009.09.009
- Bešker, N., Amadei, A., and D'Abramo, M. (2013). Molecular mechanisms of activation in CDK2. *J. Biomol. Struct. Dyn.* doi: 10.1080/07391102.2013.844080. [Epub ahead of print].
- Braig, S., Kressirer, C. A., Liebl, J., Bischoff, F., Zahler, S., Meijer, L., et al. (2013). Indirubin derivative 6BIO suppresses metastasis. *Cancer Res.* 73, 6004–6012. doi: 10.1158/0008-5472.CAN-12-4358
- Dardick, C., and Ronald, P. (2006). Plant and animal pathogen recognition receptors signal through non-RD kinases. *PLoPathog.* 2:e2. doi: 10.1371/journal.ppat.0020002
- Debondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O., and Kim, S. H. (1993). Crystal-structure of cyclin-dependent kinase-2. *Nature* 363, 595–602. doi: 10.1038/363595a0
- Endicott, J. A., Noble, M. E. M., and Johnson, L. N. (2012). The structural basis for control of eukaryotic protein kinases. *Annu. Rev. Biochem.* 81, 587–613. doi: 10.1146/annurev-biochem-052410-090317
- Fernandez, A., Sanguino, A., Peng, Z., Ozturk, E., Chen, J., Crespo, A., et al. (2007). An anticancer C-Kit kinase inhibitor is reengineered to make it more active and less cardiotoxic. *J. Clin. Invest.* 117, 4044–4054. doi: 10.1172/JCI32373
- Hantschel, O., and Superti-Furga, G. (2004). Regulation of the C-Abl and Bcr-Abl tyrosine kinases. *Nat. Rev. Mol. Cell Biol.* 5, 33–44. doi: 10.1038/nrm1280
- Huang, H., Ozkirimli, E., and Post, C. B. (2009). Comparison of three perturbation molecular dynamics methods for modeling conformational transitions. *J. Chem. Theory Comput.* 5, 1304–1314. doi: 10.1021/ct9000153
- Huang, H., Zhao, R., Dickson, B. M., Skeel, R. D., and Post, C. B. (2012). alpha C Helix as a Switch in the conformational transition of Src/CDK-like kinase domains. *J. Phys. Chem. B* 116, 4465–4475. doi: 10.1021/jp301628r
- Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell* 109, 275–282. doi: 10.1016/S0092-8674(02)00741-9
- Kornev, A. P., Haste, N. M., Taylor, S. S., and Ten Eyck, L. F. (2006). Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17783–17788. doi: 10.1073/pnas.0607656103



- Kornev, A. P., and Taylor, S. S. (2010). Defining the conserved internal architecture of a protein kinase. *Biochim. Biophys. Acta* 1804, 440–444. doi: 10.1016/j.bbapap.2009.10.017
- Kornev, A. P., Taylor, S. S., and Ten Eyck, L. F. (2008). A helix scaffold for the assembly of active protein kinases. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14377–14382. doi: 10.1073/pnas.0807988105
- Kumar, A., Chaturvedi, V., Bhatnagar, S., Sinha, S., and Siddiqi, M. I. (2009). Knowledge based identification of potent antitubercular compounds using structure based virtual screening and structure interaction fingerprints. *J. Chem. Inf. Model.* 49, 35–42. doi: 10.1021/ci8003607
- Lambert, G. K., Duhme-Klair, A. K., Morgan, T., and Ramjee, M. K. (2013). The background, discovery and clinical development of BCR-ABL inhibitors. *Drug Discov. Today* 18, 992–1000. doi: 10.1016/j.drudis.2013.06.001
- Lapenna, S., and Giordano, A. (2009). Cell cycle kinases as therapeutic targets for cancer. *Nat. Rev. Drug Discov.* 8, 547–566. doi: 10.1038/nrd2907
- Levinson, N. M., Kuchment, O., Shen, K., Young, M. A., Koldobskiy, M., Karplus, M., et al. (2006). A Src-like inactive conformation in the Abl tyrosine kinase domain. *PLoS Biol.* 4:e144. doi: 10.1371/journal.pbio.0040144
- Lolli, G. (2010). Structural dissection of cyclin dependent kinases regulation and protein recognition properties. *Cell Cycle* 9, 1551–1561. doi: 10.4161/cc.9.8.11195
- Lowe, E. D., Noble, M. E. M., Skamnaki, V. T., Oikonomakos, N. G., Owen, D. J., and Johnson, L. N. (1997). The crystal structure of a phosphorylase kinase peptide substrate complex: kinase substrate recognition. *Embo. J.* 16, 6646–6658. doi: 10.1093/emboj/16.22.6646
- Madhusudan, P. A., Xuong, N. H., and Taylor, S. S. (2002). Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nat. Struct. Biol.* 9, 273–277. doi: 10.1038/nsb780
- Noble, M. E. M., Endicott, J. A., and Johnson, L. N. (2004). Protein kinase inhibitors: Insights into drug design from structure. *Science* 303, 1800–1805. doi: 10.1126/science.1095920
- Nolen, B., Taylor, S., and Ghosh, G. (2004). Regulation of protein kinases: controlling activity through activation segment conformation. *Mol. Cell* 15, 661–675. doi: 10.1016/j.molcel.2004.08.024
- Shan, Y., Arkhipov, A., Kim, E. T., Pan, A. C., and Shaw, D. E. (2013). Transitions to catalytically inactive conformations in EGFR kinase. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7270–7275. doi: 10.1073/pnas.1220843110
- Shan, Y., Seeliger, M. A., Eastwood, M. P., Frank, F., Xu, H., Jensen, M. O., et al. (2009). A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc. Natl. Acad. Sci. U.S.A.* 106, 139–144. doi: 10.1073/pnas.0811223106
- Sicheri, F., Moarefi, I., and Kuriyan, J. (1997). Crystal structure of the Src family tyrosine kinase Hck. *Nature* 385, 602–609. doi: 10.1038/385602a0
- Waldrop, M. M. (2014). The big guns. *Nature* 505, 604–606. doi: 10.1038/505604a
- Wermuth, G., Ganellin, C. R., Lindberg, P., and Mitscher, L. A. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* 70, 1129–1143. doi: 10.1351/pac199870051129
- Xu, W. Q., Harrison, S. C., and Eck, M. J. (1997). Three-dimensional structure of the tyrosine kinase c-Src. *Nature* 385, 595–602. doi: 10.1038/385595a0
- Yang, S., Banavali, N. K., and Roux, B. (2009). Mapping the conformational transition in Src activation by cumulating the information from multiple molecular dynamics trajectories. *Proc. Natl. Acad. Sci. U.S.A.* 106, 3776–3781. doi: 10.1073/pnas.0808261106
- Zahler, S., Tietze, S., Totzke, F., Kubbutat, M., Meijer, L., Vollmar, A. M., et al. (2007). Inverse *in silico* screening for identification of kinase inhibitor targets. *Chem. Biol.* 14, 1207–1214. doi: 10.1016/j.chembiol.2007.10.010

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 March 2014; accepted: 22 April 2014; published online: 13 May 2014.

Citation: D'Abramo M, Besker N, Chillemi G and Grottesi A (2014) Modeling conformational transitions in kinases by molecular dynamics simulations: achievements, difficulties, and open challenges. *Front. Genet.* 5:128. doi: 10.3389/fgene.2014.00128  
This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 D'Abramo, Besker, Chillemi and Grottesi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Combining affinity proteomics and network context to identify new phosphatase substrates and adapters in growth pathways

Francesca Sacco<sup>1\*</sup>, Karsten Boldt<sup>2</sup>, Alberto Calderone<sup>1</sup>, Simona Panni<sup>3</sup>, Serena Paoluzi<sup>1</sup>, Luisa Castagnoli<sup>1</sup>, Marius Ueffing<sup>2,4</sup> and Gianni Cesareni<sup>1,5\*</sup>

<sup>1</sup> Department of Biology, University of Rome Tor Vergata, Rome, Italy

<sup>2</sup> Division of Experimental Ophthalmology, Centre for Ophthalmology, Institute for Ophthalmic Research, University of Tuebingen, Tuebingen, Germany

<sup>3</sup> Department DiBEST, University of Calabria, Rende, Italy

<sup>4</sup> Research Unit for Protein Science, Helmholtz Zentrum München, Neuherberg, Germany

<sup>5</sup> Istituto Ricovero e Cura a Carattere Scientifico, Fondazione Santa Lucia, Rome, Italy

## Edited by:

Andreas Zanzoni, Inserm TAGC  
UMR1090, France  
Allegra Via, Sapienza University, Italy

## Reviewed by:

Antonio Feliciello, University  
Federico II, Italy  
Roberto Sacco, CeMM-Research  
Center for Molecular Medicine of  
the Austrian Academy of Sciences,  
Austria  
Montserrat Soler-Lopez, Institute for  
Research in Biomedicine, Spain

## \*Correspondence:

Francesca Sacco and Gianni  
Cesareni, Department of Biology,  
University of Rome Tor Vergata, Via  
della ricerca scientifica, 00133  
Rome, Italy  
e-mail: francesca.sacco@  
uniroma2.it;  
gianni.cesareni@uniroma2.it

Protein phosphorylation homeostasis is tightly controlled and pathological conditions are caused by subtle alterations of the cell phosphorylation profile. Altered levels of kinase activities have already been associated to specific diseases. Less is known about the impact of phosphatases, the enzymes that down-regulate phosphorylation by removing the phosphate groups. This is partly due to our poor understanding of the phosphatase-substrate network. Much of phosphatase substrate specificity is not based on intrinsic enzyme specificity with the catalytic pocket recognizing the sequence/structure context of the phosphorylated residue. In addition many phosphatase catalytic subunits do not form a stable complex with their substrates. This makes the inference and validation of phosphatase substrates a non-trivial task. Here, we present a novel approach that builds on the observation that much of phosphatase substrate selection is based on the network of physical interactions linking the phosphatase to the substrate. We first used affinity proteomics coupled to quantitative mass spectrometry to saturate the interactome of eight phosphatases whose down regulations was shown to affect the activation of the RAS-PI3K pathway. By integrating information from functional siRNA with protein interaction information, we develop a strategy that aims at inferring phosphatase physiological substrates. Graph analysis is used to identify protein scaffolds that may link the catalytic subunits to their substrates. By this approach we rediscover several previously described phosphatase substrate interactions and characterize two new protein scaffolds that promote the dephosphorylation of PTPN11 and ERK by DUSP18 and DUSP26, respectively.

**Keywords: phosphatase, signal transduction, systems biology, cell biology, protein protein interaction**

## INTRODUCTION

Protein phosphorylation is a common post-translational modification governing signal propagation (Mann and Jensen, 2003). The concerted activity of kinases and phosphatases modulate protein phosphorylation levels and control key physiological processes, such as migration, proliferation, inflammation, and apoptosis (Graves and Krebs, 1999; Manning et al., 2002a,b). Till recently protein phosphatases have been considered uninteresting housekeeping enzymes and have received less attention compared to kinases (Bardelli and Velculescu, 2005). However, evidence accumulated over the past decades have indicated that this enzyme class plays an important regulatory role and that the deregulation of the concentration or activity of specific phosphatases correlate with a variety of human disorders (Wera and Hemmings, 1995; Tonks, 2006). Notably, approximately 40% of protein phosphatases are implicated in tumor development, highlighting the central role of this enzyme group in growth regulation

and identifying some members of this enzyme class as promising therapeutic targets (Julien et al., 2011; Liberti et al., 2013). One of the problems in the characterization, on a large scale, of the functional role of members of the phosphatase family is the lack of a simple, robust, method to identify physiologically relevant substrates. Many phosphatases have low intrinsic enzymatic specificity and are able to de-phosphorylate many substrates non-specifically *in vitro* (Tremblay, 2009). Alternative methods such as the use of trapping mutants (Blanchetot et al., 2005) are often used, but the identification of direct phosphatase substrates still remains a challenge.

In order to characterize new modulators of some key cancer associated pathways and to identify their direct targets, we have recently proposed a novel strategy based on a phosphatase high content siRNA screening combined with modeling and simulation. This approach enabled the identification of 62 phosphatase catalytic or regulatory subunits whose down-regulation affects

one or more of five readouts linked to cell proliferation: ERK, p38, and NFkB activation, rpS6 phosphorylation and autophagy (Sacco et al., 2012a). However, this approach was not designed to identify the direct phosphatase substrates, responsible for the phenotypic effect.

Here we delineate a strategy to identify protein scaffolds that may contribute to substrate recognition specificity by bridging the phosphatases to their targets. To develop this strategy we focused on eight phosphatase subunits whose down-regulation was found to affect ERK and/or RPS6 phosphorylation and are therefore modulators of the RAS-PI3K pathway (Figure 1). To identify new phosphatase substrates involved in the control of the RAS-PI3K pathway we first built a protein interaction network (PPI) by combining information extracted from protein interaction databases and results from new affinity purification experiments. This analysis confirmed that the identified phosphatase interactors often act as molecular bridges linking enzymes to substrates. In addition we independently validated a subset of these predictions

## RESULTS

### THE PHOSPHATASE INTERACTOME

We have used the results of the siRNA screening (Sacco et al., 2012a) to select eight phosphatases that modulate the activity of the RAS-PI3K pathway. The phosphatase catalytic or regulatory subunits were cloned in frame C-terminal to an SF-TAP cassette and transiently transfected in HeLa cells. These constructs direct the synthesis of four tyrosine phosphatases (PTPN21, PTPN3, DUSP18, and DUSP26), three components of the PP2A holoenzyme (the PPP2R3C regulatory subunit, the PPP2R1A scaffold subunit and the PPP2CA catalytic subunit) and the PPP3CA (calcineurin) serine/threonine phosphatase. As negative control, HeLa cells were transiently transfected with the empty vector SF-TAP. Since in our siRNA screening phosphatases controlling the activity of the RAS-PI3K pathway were identified in HeLa cells stimulated with TNF $\alpha$  for 10 min, we decided to perform the affinity purification experiments in the same experimental condition. Thus, phosphatase transfected cells were stimulated with TNF $\alpha$  for 10 min or left untreated. While the control cells were grown in a medium containing natural amino acids, phosphatase transfected cells with or without TNF $\alpha$  were, respectively, grown in media containing isotopically labeled lysine and arginine amino acids (SILAC) (Ong et al., 2002). After lysis, phosphatases, and regulatory subunits were affinity purified and analyzed by mass spectrometry (Table S1), as described in Materials and Methods (Figure 2). Contaminants binding to these baits were identified by their equal abundance in both (Table S3), the affinity-purified phosphatase sample and the negative control (Meixner et al., 2011), whereas true co-purified interactors to a given phosphatase were identified by selective enrichment of their peptides (Table S2). Only those that were significantly enriched in our samples were considered for further analysis, as described in Materials and Methods. As shown in Figure 3, this strategy resulted in a highly connected interaction network. Approximately 10% of the identified interactions have already been reported in the literature. Indeed we were able to recapitulate most of the interactions occurring between the catalytic

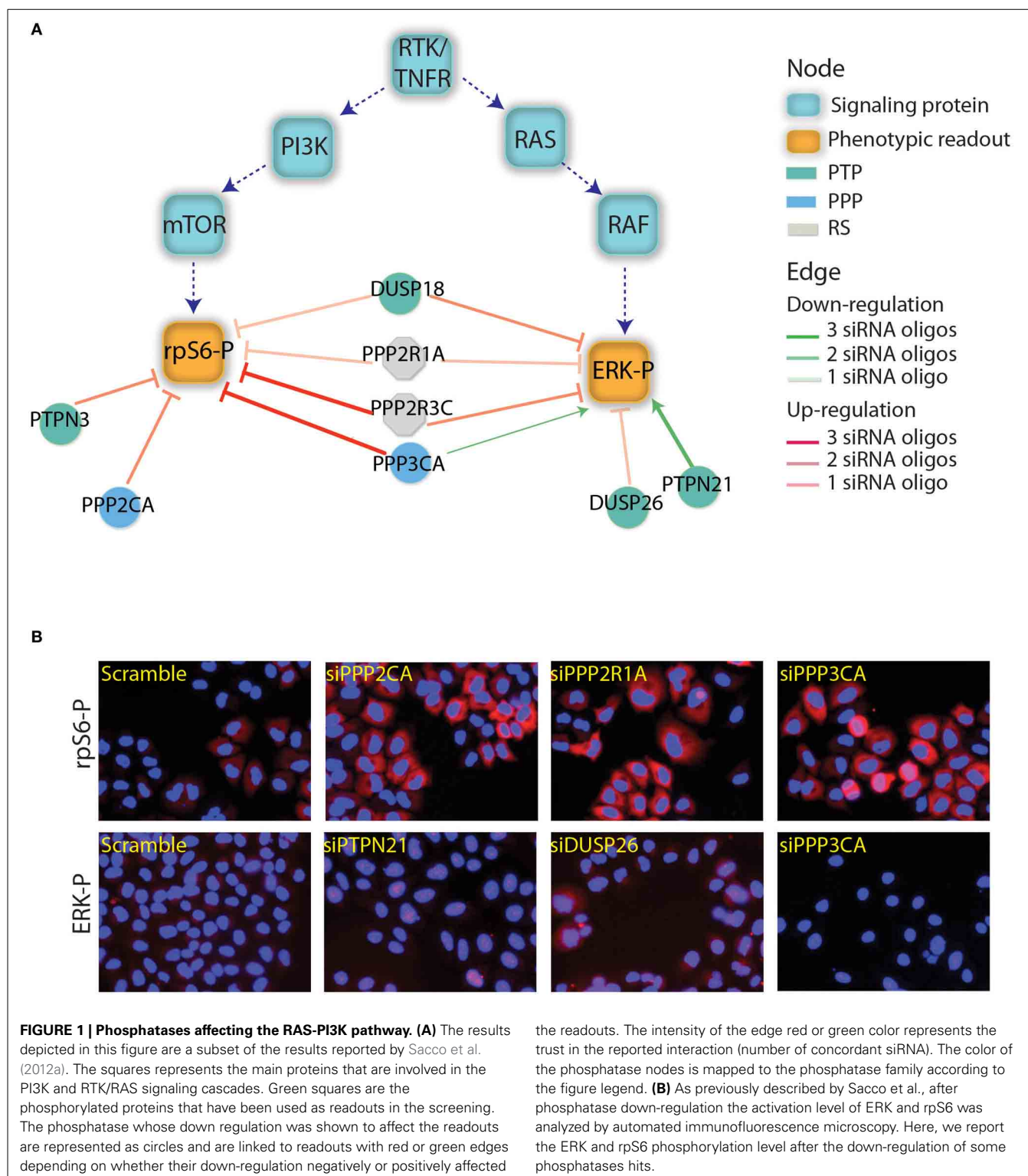
and the scaffold subunits of the PP2 holoenzyme, which, as expected, share a significant number of common interactors, many of which are regulatory subunits. These observations taken together with the validation by coimmunoprecipitation assays of some of the newly identified phosphatase interactions (Figure S1) confirm the reliability of our experimental approach. In addition in Figure S1, we demonstrated that our affinity purification experiments enabled the identification of new phosphatase interactors (the dynein protein, DLC1 and the serine threonine kinase, ATM) already involved in the regulation of the autophagy process. With the exception of the PPP2CA-DLC1 binding, both DUSP26-ATM and PTPN21-DLC1 associations are decreased upon an autophagic stimulus (starvation), suggesting that these interactions may have a regulatory role in the autophagy process.

Since the affinity purification was also carried out with or without incubation with TNF $\alpha$  we can also provide dynamically regulated interactions in response to TNF $\alpha$  treatment. As shown in Figure 3, a few interactions are positively (green edge) or negatively (red edge) regulated by TNF $\alpha$  incubation. The vast majority of the co-purified ligands, however, are TNF $\alpha$  independent (black edges).

### GUILT BY ASSOCIATION

Next we used the phosphatase interactome derived from the *in vivo* pull down experiment to ask whether the phosphatase interaction network could provide hints toward specific pathways that are affected by phosphatase activity. To this aim, we performed a KEGG- pathway enrichment analysis of the ligands of each of the phosphatases, by using the Functional Annotation Tool, David (Huang da et al., 2009). The two phosphatases PTPN21 and PPP2CA and the PP2 scaffold subunit PPP2R1A were significantly associated to RTK signaling (Figure 3), in agreement with their established involvement in the modulation of EGF signaling by controlling the SRC and S6K kinases, respectively (Cardone et al., 2004; Carlucci et al., 2010; Hahn et al., 2010). In addition our enrichment analysis reveals a statistically significant association of PPP3CA with cell differentiation signaling. This result is consistent with the report by Kao et al. that the differentiation of Schwann cells requires the activity of the PPP3CA phosphatase (Kao et al., 2009). Similarly, we found that DUSP26 is significantly correlated with the DNA damage response. This conclusion is in accordance with the observations that DUSP26 inhibits the p53 tumor suppressor function, by suppressing doxorubicin-induced apoptosis in human neuroblastoma cells (Shang et al., 2010). On the other hand, our experimental strategy led to the identification of new biological processes that are controlled by these protein phosphatases (e.g., vesicular trafficking and cell metabolism). As shown in Figure 3, the interactors of both DUSP18 and PTPN3 were not significantly associated to any specific pathway.

Next we looked for evidence that the proteins that copurified with each phosphatase may form complexes. To this end, we queried the *mentha* database (Calderone et al., 2013), and looked for evidence of interactions between ligands of each bait phosphatase. The interactors of five of the eight phosphatases are linked by direct interactions. As illustrated in Figure 4A, we found

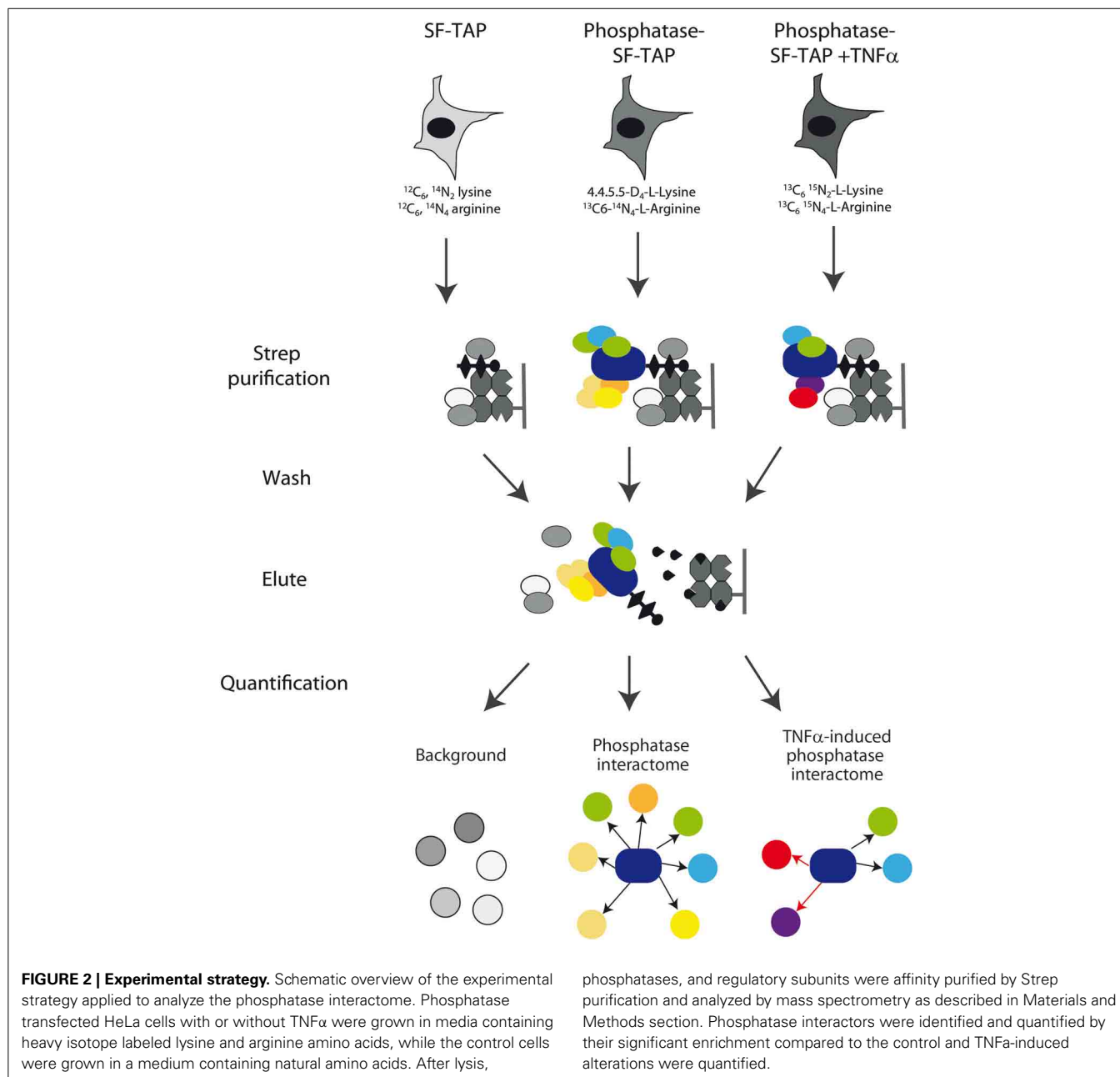


that DUSP26 copurifies with the serine/threonine kinase ATM, which, in response to genotoxic stress, phosphorylates the two Fanconi proteins FANCI and FANCD2, triggering the S-phase checkpoint activation (Taniguchi et al., 2002). A third DUSP26 ligand, TELO2, which is a member, together with TTI1 and

TTI2, of the TTT complex (Hurov et al., 2010) also interacts with ATM.

PTPN21, on the other hand, interacts with the scaffold protein GRB2, which associates with DNAJB11, DYNLL1, and UBR, suggesting that some of the identified interactors may copurify by





indirect interactions (Figure 4C). Similarly we found that PTPN3 interact with the mitochondrial ribosomal subunit ICT1 that binds GADD45GIP1 and POLRMT1 proteins (Figure 4B).

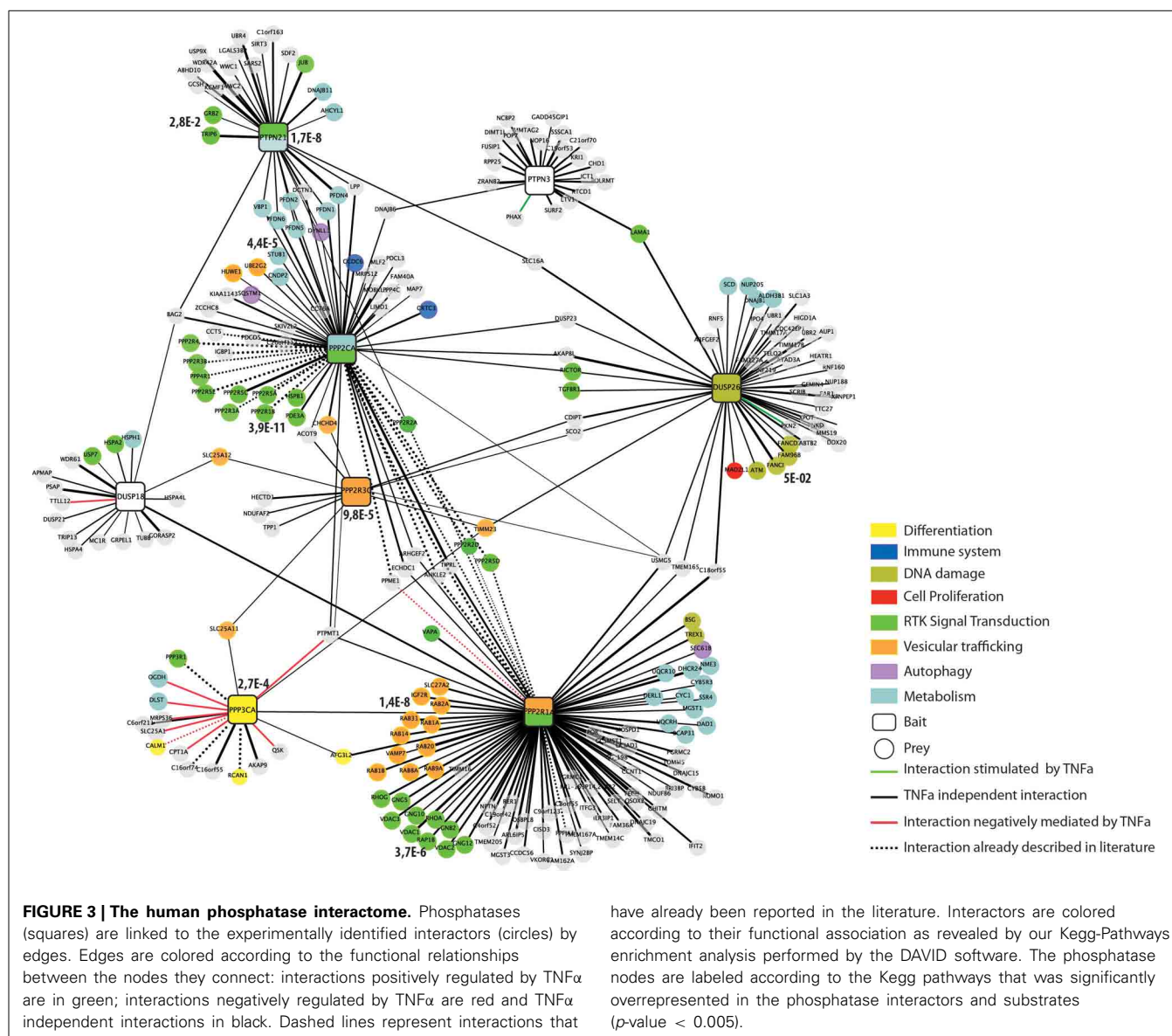
As expected, the catalytic and scaffold subunit of PP2A share many interactors, confirming that such heterodimer forms different protein complexes that act on distinct substrates, by recruiting multiple regulatory subunits (Figure 4D).

#### PTPN21 ASSOCIATES WITH THE SH3 DOMAIN OF GRB2

Among all phosphatase-interaction partners, we focused on the newly discovered interaction between the scaffold protein GRB2 and the tyrosine phosphatase PTPN21, both partners mapping to the RAS-PI3K signaling pathway. Cardone et al. reported that

PTPN21 is recruited to mitochondria by binding the scaffold protein AKAP121 and that this interaction is essential for the phosphatase dependent dephosphorylation of the inhibitory tyrosine 527 of the SRC kinase (Cardone et al., 2004; Carlucci et al., 2010).

GRB2 is an essential adapter protein consisting of two SH3 domains flanking one central SH2 domain. The affinity purification assay results suggest that the GRB2-PTPN21 interaction is not likely to occur in a phosphorylation dependent manner, since it is not modulated by TNF $\alpha$  (Figure 3). However, phosphoproteomics of both cancer and embryonic stem cells revealed that PTPN21 contains multiple tyrosine phosphorylated residues (Rikova et al., 2007; Guo et al., 2008; Brill et al., 2009). In order to

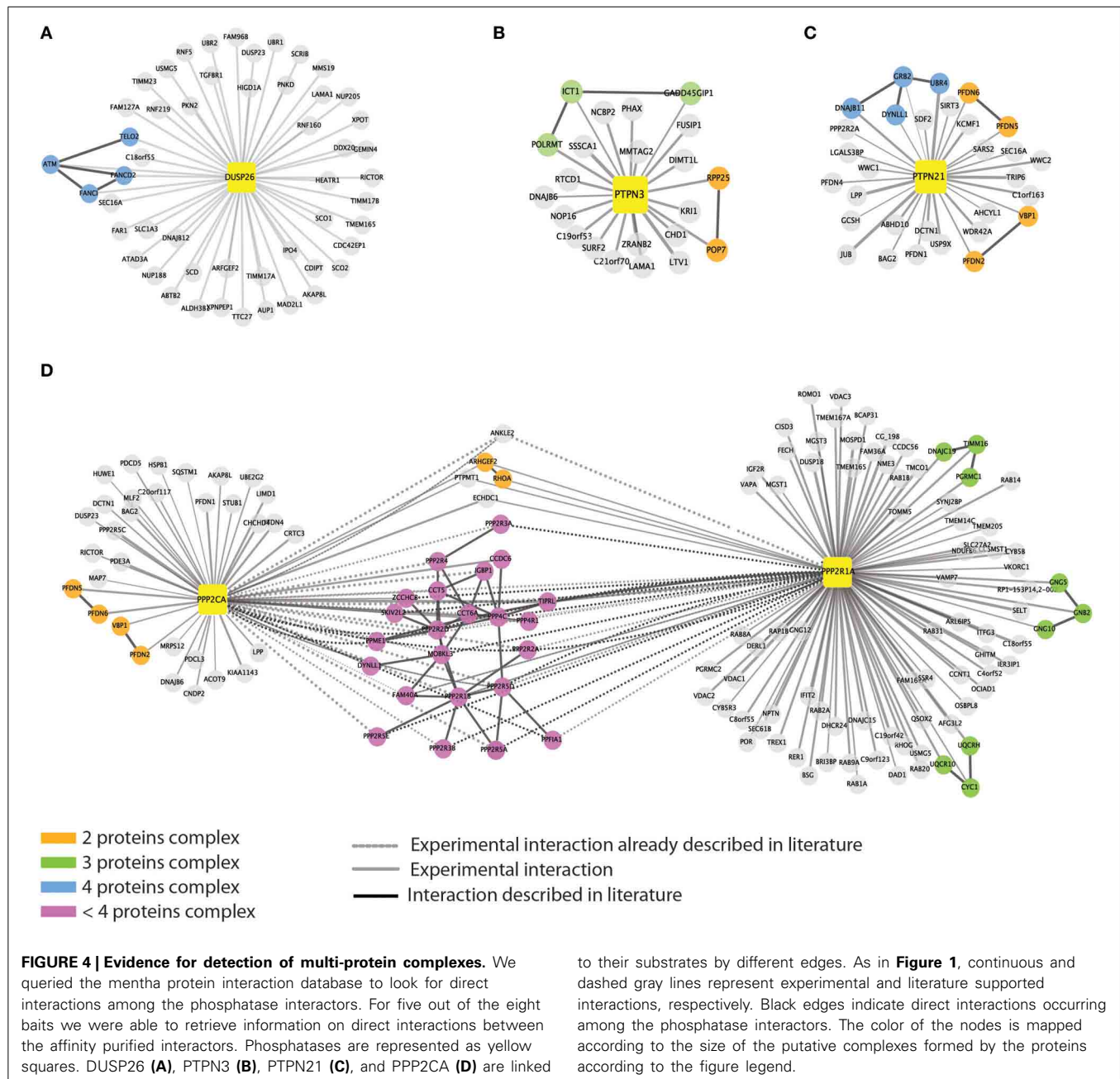


map the GRB2-PTPN21 interaction to a specific GRB2 domain and assess whether such binding occurs in a phosphorylation dependent manner, the SH2 domain of GRB2 as well as its two SH3 domains were purified as GST fusion proteins and incubated with whole protein extracts co-transfected with Flag-PTPN21 in presence or in absence of a constitutively active SRC kinase mutant (Y527F). As shown in **Figure 5A**, PTPN21 strongly binds the C-terminal SH3 domain of GRB2 and to a lesser extent the N-term SH3 domain, independently from SRC. On the other hand, the GRB2 SH2 domain does not interact with PTPN21. The analysis of PTPN21 protein sequence reveals that it contains a SH3 binding motif (<sup>564</sup>RP PPPYPPRP<sup>574</sup>), whose sequence matches the GRB2 binding specificity described by Carducci et al. (2012). These results support the existence of a PTPN21-GRB2 complex that is phosphorylation independent and likely occurs between the carboxy-terminal SH3 domain of GRB2 and PTPN21 (**Figure 5B**).

Next we asked whether the formation of the PTPN21 GRB2 complex promotes the dephosphorylation of GRB2. For this purpose, GRB2 tyrosine phosphorylation was induced by transfecting HeLa cells with the constitutively active SRC kinase mutant (Y527F) in presence or in absence of Flag-PTPN21. As shown in **Figure 5C**, after cell lysis and immunoprecipitation with anti-GRB2, PTPN21 was found to associate with GRB2. However, when PTPN21 is overexpressed, GRB2 phosphorylation, if anything, seems to be slightly increased, as revealed by probing the GRB2 protein with an anti-phospho tyrosine antibody (**Figure 5C**). Thus, GRB2 is not a substrate of PTPN21 but may play a role in targeting PTPN21 to different substrates.

#### A STRATEGY TO IDENTIFY NEW PHOSPHATASE SUBSTRATES IN GROWTH PATHWAYS

Having obtained a high coverage interactome of the eight phosphatases that affect the RAS-PI3K pathway we used it to develop

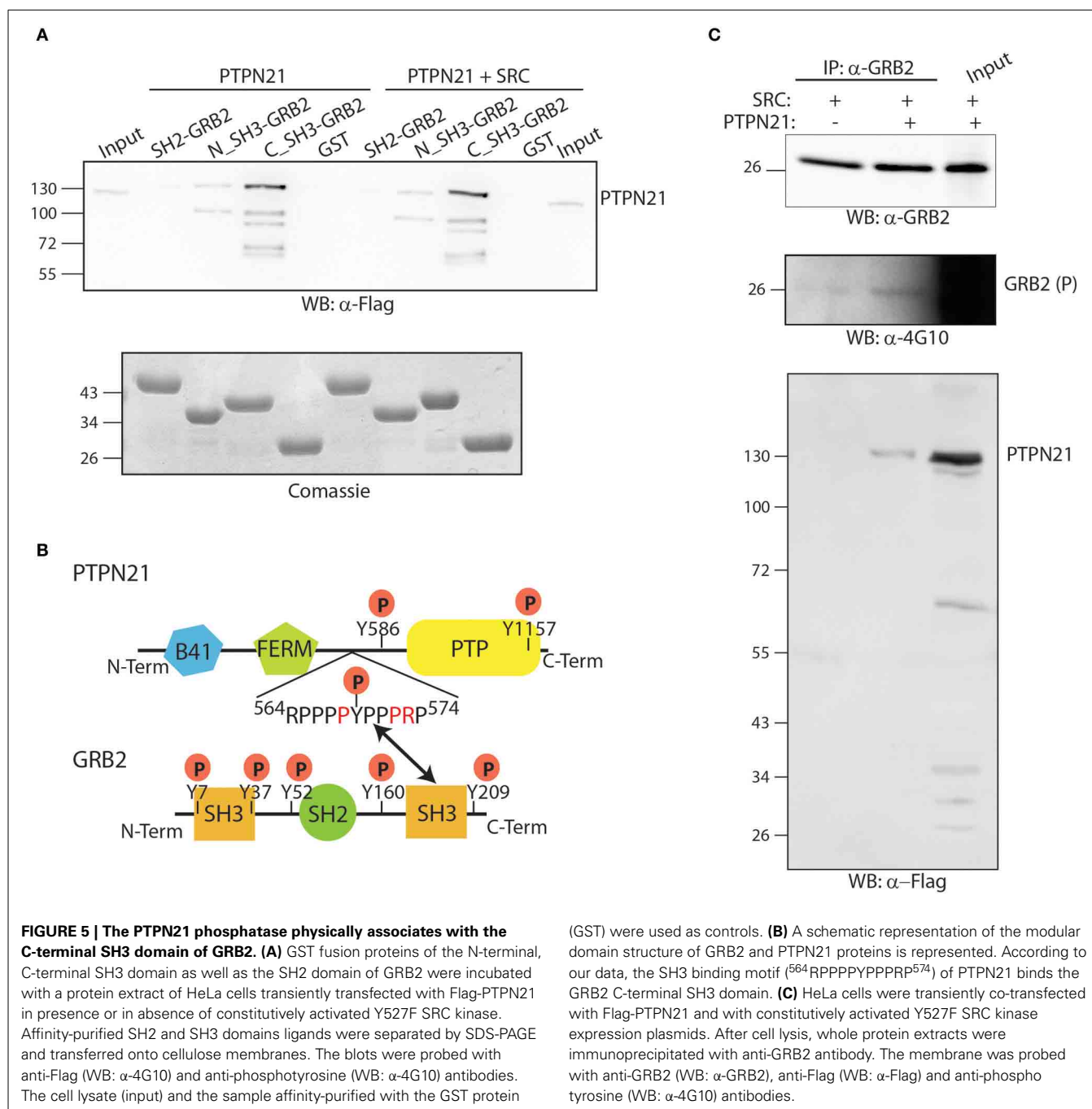


a general strategy that could infer the direct target of these phosphatases. Phosphatase-substrate interaction is weak and transient, thus it is unlikely that substrates can be identified by co-immunoprecipitation. In fact none of the interactors identified in the affinity purification experiments are among the validated substrates annotated in the HUPHO and DEPOD databases (Li et al., 2013; Liberti et al., 2013). It has been reported that much of phosphatase substrate specificity, localization and activity is modulated by the interaction with scaffold/regulatory proteins that target them to specific locations (Roy and Cyert, 2009; Sacco et al., 2012b). We hypothesized that some of the interactors identified by our approach act as molecular bridges linking phosphatase to substrates participating in the RAS-PI3K pathway. For this reason,

we made use of the PPI network downloaded from the *mentha* database (Calderone et al., 2013) to link phosphatase interactors to putative substrates in the RAS-PI3K pathway (Figure 6A).

The strategy that we used is based on the following steps:

- (1) Draw a literature derived directed network of the RAS-PI3K pathway and identify the participating proteins as putative targets of the “modulator phosphatases” (Supplementary Material, Table S4).
- (2) Combine the phosphatase interactors identified in the affinity purification MS experiment with the ones already described in the literature and reported in the *mentha* database (red and black edges, respectively, in Figure 6B).



- (3) Define paths in the protein interaction graph that connect each phosphatase to the proteins participating in a given pathway (here RAS-PI3K signaling).

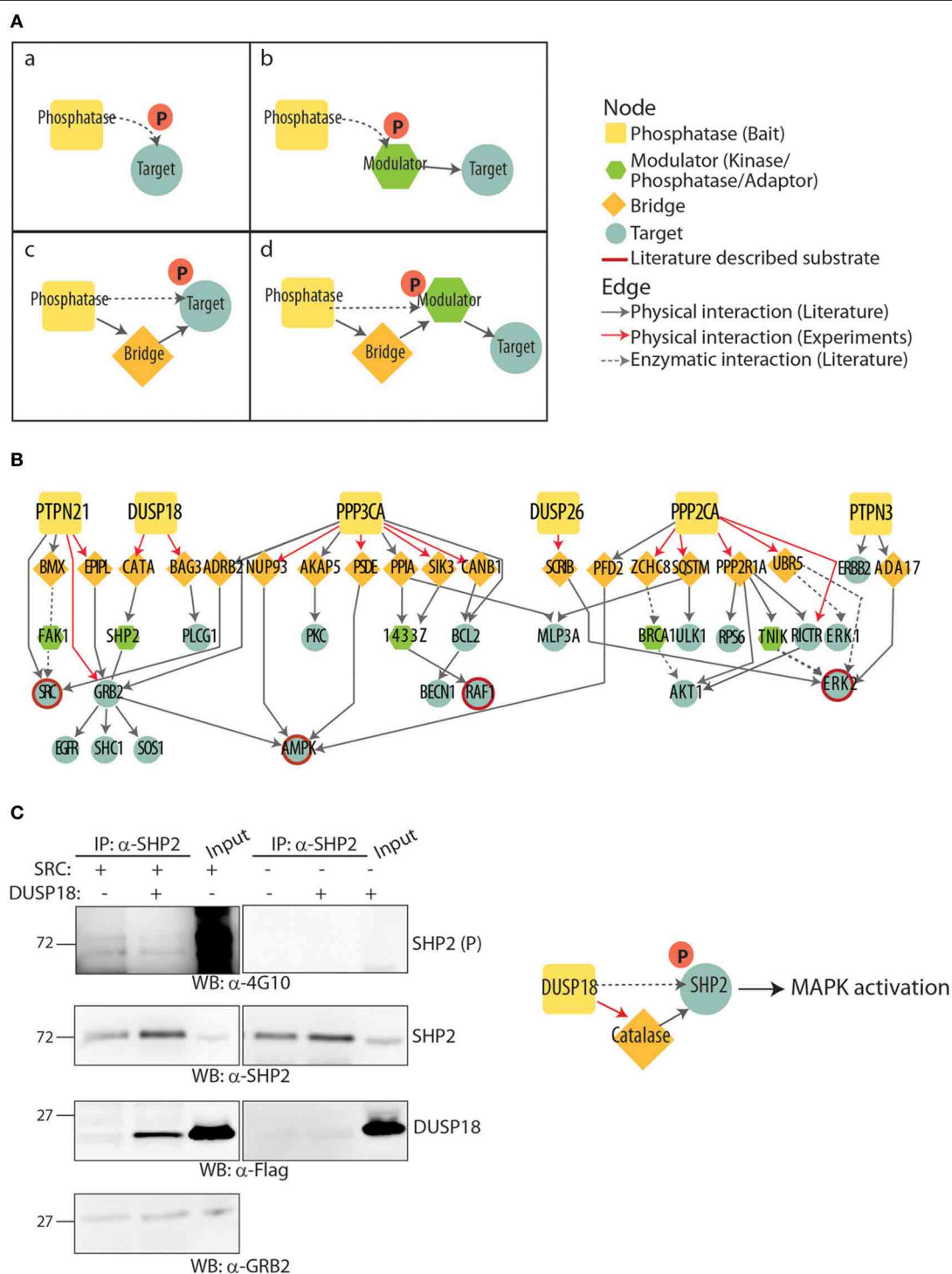
By this strategy, each interactor was linked to RAS-PI3K signaling proteins and a by a large number of possible paths. The resulting complex graph was filtered according to the following rules (illustrated in **Figure 6A**):

- (1) Longer paths are filtered out. Only paths connecting tyrosine phosphatases to protein members of the growth network

with up to two “binding steps” are considered. For phosphatases subunits that form holoenzymes with regulatory subunits such as PP2A and PPP3CA we allowed three binding steps.

- (2) The inferred substrate in the RAS-PI3K pathway has to be a kinase, a phosphatase or a scaffold protein whose dephosphorylation controls either enzyme activity or the molecular association with other regulatory proteins.
- (3) The inferred substrate has to contain phosphorylation sites with defined functional roles as annotated in the Phosphosite database (Hornbeck et al., 2012).





**FIGURE 6 | Inferring new phosphatase substrates. (A)** Schematic representation of the multiple paths going from phosphatase to substrates. **(B)** The multiple paths going from phosphatases to substrates are represented as a graph. Nodes have different shapes according to their functional role: phosphatases are indicated as squares, bridge proteins as diamonds, modulators as hexagons and inferred substrate as circles. The red border outlines phosphatase substrates that have been already reported in literature. Solid black and red lines indicate physical interaction literature and experimentally supported, respectively, while black dashed line represent

enzymatic interaction already described in literature. **(C)** HeLa cells were transiently co-transfected with expression plasmids expressing Flag-SHP2 and a constitutively active mutant of the SRC kinase (Y527F) expression plasmids. After cell lysis, whole protein extracts were immunoprecipitated with anti-SHP2 antibody. The beads were washed with lysis buffer, and the immunoprecipitation (IP) was revealed with anti-SHP2 (WB: α-SHP2), anti-GRB2 (WB: α-GRB2), anti-Flag (WB: α-Flag) and anti-phospho tyrosine (WB: α-4G10) antibodies. GRB2 which is an established ligand of SHP2 was used as a positive control.

- (4) The phosphorylation sites of the inferred substrates have to be compatible with the nature of the phosphatase (Tyrosine phosphatase can only dephosphorylate tyrosine residues, etc).
- (5) The effect of the phosphatase induced dephosphorylation of the inferred substrate has to explain the phenotypic effect observed upon phosphatase down-regulation in the functional screenings (Sacco et al., 2012a).

The result of this approach (**Table S5**) is illustrated in the filtered graph in **Figure 6B**. Remarkably, our strategy was validated by the recovery of phosphatase substrates already reported in the literature. For instance, Duan and Cobb already demonstrated that PPP3CA induces the MAPK activation by dephosphorylating Thr401 in RAF1 (Duan and Cobb, 2010). In addition the inhibitory effect of PTPN3 on ERK phosphorylation was already reported by Han et al. (2000). Interestingly, both PPP3CA and PPP2CA phosphatases have been already described to be negative modulators of autophagy (Magnaudeix et al., 2013; He et al., 2014). Our approach enabled the identification of a new potential molecular mechanism that these two phosphatases may control to modulate autophagy. SIK3 and SQSTM proteins have been identified by our affinity purification experiment as two novel interactors of PPP3CA and PPP2CA, respectively. In our approach, we propose that SIK3 and SQSTM proteins act as bridge to connect PPP3CA and PPP2CA phosphatases to the autophagy marker MLP3A (LC3A). This observation suggests that our approach can be used to propose new potential molecular mechanisms linking a phosphatase to an established biological process. This graph links phosphatase to putative adapter and to putative substrates. In principle depending on the available information one can use it (1) to infer new substrates starting from a consolidated PPI or (2) to validate molecular bridges that target a phosphatase to an established substrate. In the two following paragraphs we will demonstrate these strategies in two specific cases.

#### SHP2 CAN BE DEPHOSPHORYLATED BY DUSP18

DUSP18 was shown by our screening to negatively regulate the RAS pathway. The graph in **Figure 6B** indicates that the regulatory protein that is closest to DUSP18 in the RAS pathway is SHP2 and that DUSP18 and SHP2 are connected by catalase. Indeed it has been shown that the SH2 domains of SHP2 bind tyrosine phosphorylated catalase (Yano et al., 2004), and catalase was recovered as a DUSP18 interactor in our approach. We can therefore picture catalase acting as a bridge linking the phosphatase to its putative target. To test this hypothesis, HeLa cells were transiently co-transfected with Y527F SRC kinase, to enhance phosphorylation, and Flag-DUSP18. As shown in **Figure 6C**, after cell lysis and endogenous immunoprecipitation with anti-SHP2, DUSP18 was found to associate with SHP2 only in SRC transfected cells. This data is compatible with the model whereby the SH2 domains of SHP2 bind tyrosine phosphorylated catalase which in turn binds to DUSP18. In addition the over-expression of DUSP18 induces SHP2 dephosphorylation, without affecting its association with GRB2. Since it has been shown that the C-terminal tyrosine residues of SHP2 bind GRB2, this result suggests that DUSP18 likely dephosphorylates the Tyr62 and

Tyr63 residues. Although the biological relevance of the inferred dephosphorylation needs to be proven in more physiological conditions, this result shows that DUSP18 has the potential to dephosphorylate SHP2 as inferred by our approach.

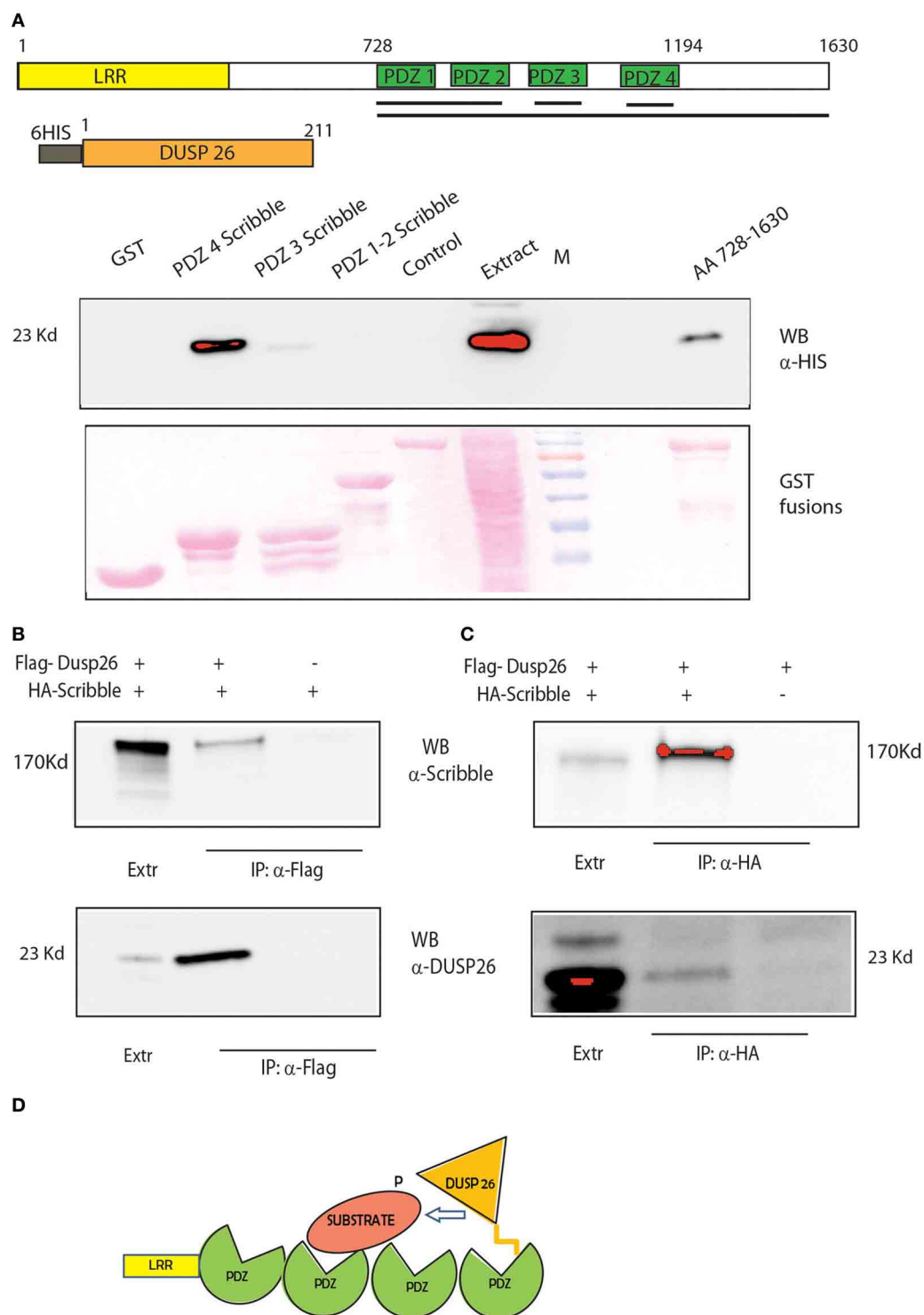
#### SCRIB ACTS AS A BRIDGE TO TARGET DUSP26 TO ERK

Knock down of DUSP26 by siRNA negatively affects the activation of ERK (Sacco et al., 2012a). This is in agreement with the ability of DUSP26 to inhibit cell proliferation in epithelial cell lines (Hu and Mivechi, 2006; Patterson et al., 2010). Consistent with a role as tumor suppressor, DUSP26 is down-regulated, in several human cancer cell lines, as well as in some primary tumors (Tanuma et al., 2009; Patterson et al., 2010). However, DUSP26 is not able to directly bind ERK to dephosphorylate it (Hu and Mivechi, 2006; Patterson et al., 2010) suggesting the existence of a molecular bridge. The heat shock transcription factor Hsf4b, a substrate of ERK, was proposed as a possible bridge to link DUSP26 to ERK (Hu and Mivechi, 2006). Similarly, more recently, the adenylate kinase 2 was proposed to be a bridge that directs DUSP26 to dephosphorylate FADD (Kim et al., 2014).

Our approach identified SCRIB as a potential bridge that would modulate the de-phosphorylation of ERK by DUSP26. SCRIB is an adapter protein that was recently suggested to down-regulate ERK by binding and activating the phosphatase PP1 gamma (Nagasaka et al., 2013). We propose here that SCRIB may also promote the de-phosphorylation of ERK by DUSP26. SCRIB directly binds to ERK through two KIM motifs and regulates its activation and nuclear translocation (Nagasaka et al., 2010). The protein contains four PDZ domains (**Figure 6A**). The C-terminus region of DUSP26 contains an atypical motif for PDZ binding L-D/E- $\Phi$ , where  $\Phi$  is a hydrophobic residue (Tonikian et al., 2008). Thus, we asked whether the binding of SCRIB to DUSP26, as identified in our affinity purification experiment, could be mediated by any of the SCRIB PDZ domains. To this end we performed a GST pull down experiments by affinity purifying extracts of *E. coli* cells expressing HIS-tagged DUSP26 with GST fusion of SCRIB PDZ domains (**Figure 7A**). Only the fourth PDZ domain of SCRIB was able to bind DUSP26. The binding was confirmed by co-immunoprecipitation assay, after cotransfecting HA-SCRIB and Flag-DUSP26 in H1299 cells. As shown in **Figure 7B**, co-immunoprecipitated SCRIB was detected by western blotting with an anti-SCRIB antibody. Similarly SCRIB was immunoprecipitated with anti-HA antibody and the presence of DUSP26 was revealed by western blotting with anti-DUSP26 antibody **Figure 7C**. These data suggest that SCRIB could direct the phosphatase activity of DUSP26 toward ERK as suggested in the cartoon in **Figure 7D**.

#### DISCUSSION

Although protein phosphorylation has been considered as a key post-translational mechanism controlling a variety of physiological processes and a number of reports have contributed to describe the phosphatase interaction network, a comprehensive characterization of phosphatase substrates is still missing (Goudreault et al., 2009; Breitreutz et al., 2010; Skarra et al., 2011; Couzens et al., 2013). Recently we have reported an unbiased siRNA screening aimed at identifying phosphatases



**FIGURE 7 | DUSP 26 interacts with the adapter protein SCRIB. (A)** GST Pull Down of full-length DUSP26 by GST constructs fused to PDZ 1-2, PDZ3, PDZ4, or all the four PDZ domains (aa 728-1630) of SCRIB fused to GST. Black Lines under the schematic representation of the domain structure of SCRIB indicate the protein regions that were fused to GST. Two controls were added: GST alone and a PDZ containing protein of similar length

(PDZK1, lane “control”). **(B,C)** Co-immunoprecipitation of SCRIB and DUSP26. H1299 cells were transfected with HA-SCRIB, Flag-DUSP26 or both plasmids as indicated. Lysates were immunoprecipitated with anti-Flag **(B)** or with anti-HA **(C)** and detection was performed with anti-DUSP26 or anti-SCRIB antibodies. **(D)** Cartoon picturing the proposed role of SCRIB as adapter protein to mediate DUSP26 dephosphorylation specificity.

controlling key growth pathways in cancer cells (Sacco et al., 2012a). Combining the siRNA screening results with modeling techniques, we were able to map phosphatases on specific nodes of the growth signaling model. However, our approach only identified phosphatases modulating the growth pathway but did not enable us to link phosphatases to specific substrates.

For this purpose, we set up to develop an experimental strategy that combines the functional information obtained with the siRNA screening and PPI network context information. We first enriched the literature derived interactome of six phosphatases and two phosphatase accessory subunits by affinity purification experiments of phosphatase complex followed by quantitative mass spectrometry based proteomics in cancer cells stimulated with TNF $\alpha$ . By this approach we were able to recapitulate most of the interactions occurring between the catalytic, scaffold, and regulatory subunit of the PP2A holoenzyme, confirming the robustness of our approach. The resulting interactome is completely connected, since each phosphatase shares at least one ligand with one of the remaining phosphatases. For instance the tyrosine phosphatase PTPN21 and the catalytic subunit of the serine threonine phosphatase PP2A share a common group of interactors, mainly involved in controlling cell metabolism. We observed that the phosphatase interactome is largely insensitive to stimulation with TNF $\alpha$ , suggesting that these interactions may be either constitutive or triggered by other types of stimuli. For instance, while the DUSP26-ATM interaction is not modulated by TNF $\alpha$ , we show that nutrients and amino acids deprivation increases the binding (Supplementary Material, **Figure S1A**), suggesting that these proteins may play a role in controlling the autophagy process. Indeed, we have previously shown that the siRNA interference of DUSP26 results in a decrease of the autophagy marker LC3, while much evidence suggest that the ATM kinase promotes the autophagy induced by ionizing radiation and ROS (Liang et al., 2013; Tripathi et al., 2013). In addition, as shown in **Figure 3**, about 50% of the PPP3CA interactions are negatively modulated by TNF $\alpha$ , including the binding to its activator subunit calmodulin. This result suggests that the TNF $\alpha$  stimulation may have an inhibitory role on PPP3CA activity. However, Fernandez et al. have recently demonstrated that in reactive, but not in quiescent astrocytes, PPP3CA dephosphorylates the transcription factor Foxo3 in response to TNF $\alpha$ , suggesting that depending on the cells type, this phosphatase may have opposite functions (Fernandez et al., 2012).

Interestingly our experimental approach enabled us to identify a novel interaction between the scaffold protein GRB2 and the tyrosine phosphatase PTPN21. Here, we report that PTPN21 binds the C-terminal SH3 domain of GRB2 *in vitro*, but it does not dephosphorylate its phosphotyrosine residues. Indeed our affinity purification experiment failed to identify known phosphatase substrates that had already been described in the literature. This observation is not surprising if we consider that phosphatases rapidly dephosphorylate the substrate and the phosphatase-substrate interaction is so transient and weak that coimmunoprecipitation-based approaches likely fail to identify phosphatase substrates. In addition, while most protein kinases recognize a specific amino acid motifs on their targets, phosphatase substrates specificity is weaker

and mainly based on the interaction with regulatory subunits (Roy and Cyert, 2009).

To infer new phosphatase substrates, we have here outlined a combined experimental-bioinformatic strategy based on the integration of the phosphatase interactome with network context information, extracted from the *mentha* PPI database (Calderone et al., 2013). Although this approach lead us to recover some of the phosphatase-substrate relationships already described in literature, we are aware of some relations that are missed by our approach [e.g., the RAF1 dephosphorylation by PPP2CA (Dent et al., 1995)]. These failures can be explained by several factors: (1) some interactions may be cell type dependent or rely on specific stimulations; (2) some phosphatase partners may have very low level of expression that remains undetected in our affinity purification experiments and (3) some PPI relations may have not been reported yet or may have not been annotated in *mentha*. In addition we want to stress that we used rather stringent filtering criteria to reduce the total number of inferred phosphatase-scaffold-substrate complexes. This might increase the chance of missing already validated enzyme-substrate relationships or of identifying new interesting regulation mechanisms. If desirable, these criteria can be relaxed at the cost of increasing the noise of false positives.

In essence our method combines functional information with the interactome and analyses the resulting graph to identify paths between phosphatases and putative substrates. By this approach new substrates may be inferred or alternatively proteins that form molecular bridges between the phosphatase and the substrates can be identified. To assess the robustness and reliability of our strategy, two specific cases were analyzed. Firstly we demonstrated that DUSP18 induces SHP2 dephosphorylation. Our siRNA screening revealed that DUSP18 negatively controls ERK phosphorylation (Sacco et al., 2012a). This is consistent with SHP2 being a positive modulator of the MAPK signaling (Cai et al., 2002). Here we infer that catalase acts as a bridge to enable the DUSP18 mediated de-phosphorylation of SHP2. Although our approach does not identify the specific SHP2 tyrosine residues dephosphorylated by DUSP18, we demonstrated that the C-terminal residues involved in the GRB2 interaction are not targeted by the phosphatase (**Figure 5**). DUSP18 may negatively controls the MAPK signaling by dephosphorylating and inactivating SHP2. Finally, we demonstrated that SCRIB acts as a bridge to mediate the dephosphorylation of ERK by DUSP26 (**Figure 7B**). DUSP26 is a poorly characterized dual specificity phosphatase whose negative regulation of the MAPK signaling has been already reported.

Taken together these observations show that the combination of the topological information contained in the phosphatase interactome with functional information obtained by siRNA screening can be valid tool to infer new phosphatase substrates and modes of targeting.

## MATERIALS AND METHODS

### ANTIBODIES AND REAGENTS

Anti-HA, anti-FLAG and anti-Flag M1 agarose beads and anti DUSP26 were from Sigma; anti-SHP2 and anti SCRIB were from Santa Cruz Biotechnology; anti-GRB2 and anti-4G10 was



from Upstate Biotechnology, Inc. Peroxidase-conjugated anti-rabbit, anti-mouse and anti-goat secondary antibodies were from Jackson ImmunoResearch. PPP2CA, PTPN3, PTPN21, DUSP26, PPP2R3C encoding plasmids were purchased from OpenBiosystem. DUSP18, PPP2R1A, and PPP3CA constructs were kindly provided by Marc Vidal. Phosphatase cDNAs were cloned in pDNOR vector (Invitrogen) and cloned in the SF-TAP plasmid by using the Gateway Recombinant Cloning Technology from Invitrogen. The cDNA of DUSP26 was also cloned in Pet28 and PC-DNA plasmids. HA-DLC1 was kindly provided by Prof. Cecconi. The cDNA encoding SRC Y527F was cloned in pSGT (Gonfloni et al., 1997). HA-SCRIB, PDZ3-GST, and PDZ4 -GST were a generous gift of L. Banks. Construct containing human SCRIB PDZ1-2 and 1-4 (aa 728-1630) were cloned in pGex2TK.

## CELL CULTURE

Cells were maintained in a humidified atmosphere at 37°C and 5% CO<sub>2</sub> in Dulbecco's modified Eagle's medium (Invitrogen), supplemented with 10% fetal bovine serum (Sigma) and 0.1% penicillin/streptomycin (Invitrogen). For SILAC experiments, SILAC DMEM (PAA, Pasching, Austria) deficient of L-Lysine and L-Arginine, supplemented with 10% (v/v) dialyzed fetal bovine serum (FBS; PAA, Pasching, Austria), 50 units/ml Penicillin, 0.05 mg/ml Streptomycin and 0.55 mM lysine, 0.4 mM arginine was used. Light labeled medium was supplemented with <sup>12</sup>C<sub>6</sub>, <sup>14</sup>N<sub>2</sub> lysine and <sup>12</sup>C<sub>6</sub>, <sup>14</sup>N<sub>4</sub> arginine, medium labeled medium with 4.4.5.5-D<sub>4</sub>-L-Lysine and <sup>13</sup>C<sub>6</sub>-<sup>14</sup>N<sub>4</sub>-L-Arginine and heavy labeled medium with <sup>13</sup>C<sub>6</sub> <sup>15</sup>N<sub>2</sub>-L-Lysine and <sup>13</sup>C<sub>6</sub> <sup>15</sup>N<sub>4</sub>-L-Arginine. Proline was added to a final concentration of 0.5 mM to prevent arginine to proline conversion (Bendall et al., 2008), which could impair the quantification. All amino acids were purchased from Silantes. Human epithelial carcinoma (HeLa) cells were purchased from the ATCC. HeLa cells were transfected with Lipofectamine 2000 (Invitrogen) according to manufacturer's protocol.

## AFFINITY PURIFICATION OF PROTEIN COMPLEXES

For one step Strep purifications, SF-TAP tagged proteins and associated protein complexes were purified essentially as described earlier (Gloeckner et al., 2007; Boldt et al., 2011). HeLa cells, transiently expressing the SF-TAP tagged constructs or SF-TAP alone as control were either stimulated with 50 ng/ml TNFα or mock treated. They were next lysed in lysis buffer (containing 150 mM NaCl, 50 mM Tris-HCl, 1% Nonidet P-40, and 0.25% sodium deoxycholate, protease inhibitor cocktail (Roche) and phosphatase inhibitor cocktails II and III (Sigma-Aldrich), for 20 min at 4°C. After sedimentation of nuclei at 10,000 × g for 10 min, the protein concentration of the lysates were determined by a Bradford assay before equal amounts of the cleared lysates were transferred to Strep-Tactin-Superflow beads (IBA) and incubated for 1 h before the resin was washed three times with wash buffer (TBS containing 0.1% NP-40, phosphatase inhibitor cocktail I and II). The protein complexes were eluted by incubation for 10 min in Strep-elution buffer (IBA). Following elution, the corresponding samples were combined. The combined samples were concentrated using 10 kDa cut-off VivaSpin 500 centrifugal devices (Sartorius Stedim Biotech) and pre-fractionated

using SDS-Page and in-gel tryptic cleavage as described elsewhere (Gloeckner et al., 2009).

## MASS SPECTROMETRY AND DATA ANALYSIS

LC-MS/MS analysis was performed on an Ultimate3000 nano RSLC system (Thermo Fisher Scientific) coupled to a LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) by a nano spray ion source. Tryptic peptide mixtures were automatically injected and separated by a linear gradient from 5 to 40% of buffer B in buffer A (2% acetonitrile, 0.1% formic acid in HPLC grade water) in buffer A (0.1% formic acid in HPLC grade water) at a flow rate of 300 nl/min over 90 min. Remaining peptides were eluted by a short gradient from 40 to 100% buffer B in 5 min. The eluted peptides were analyzed by the LTQ Orbitrap Velos mass spectrometer. From the high resolution MS pre-scan with a mass range of 300–1500, the 10 most intense peptide ions were selected for fragment analysis in the linear ion trap if they exceeded an intensity of at least 500 counts and if they were at least doubly charged. The normalized collision energy for CID was set to a value of 35 and the resulting fragments were detected with normal resolution in the linear ion trap. The lock mass option was activated, the background signal with a mass of 445.12002 was used as lock mass (Olsen et al., 2005). Every ion selected for fragmentation, was excluded for 20 s by dynamic exclusion. For SILAC experiments, all acquired spectra were processed and analyzed using the MaxQuant software (Cox and Mann, 2008) (version 1.0.13.13) and the human specific IPI database version 3.52 (<http://www.maxquant.org/>) in combination with Mascot (Matrix Science, version 2.2). Cysteine carbamidomethylation was selected as fixed modification, methionine oxidation and protein acetylation were allowed as variable modifications. The peptide and protein false discovery rates were set to 1%. Contaminants like keratins were removed. Proteins, identified and quantified by at least two unique peptides were considered for further analysis. The significance values were determined by Perseus tool using significance B. Those proteins whose ratio was greater than 1.9 and significance B was lesser than 0.1 were considered significantly enriched.

## PULL-DOWN ASSAY

After 24 h of transfection, confluent HeLa cells were washed with ice-cold PBS and lysed in RIPA buffer (150 mM NaCl, 50 mM Tris-HCl, 1% Nonidet P-40, 0.25% sodium deoxycholate) supplemented with 1 mM pervanadate, 1 mM NaF, protease inhibitor mixture 200× (Sigma), inhibitor phosphatase mixture I and II 100× (Sigma). The samples were kept on ice for 30 min and centrifuged at 15,000 rpm at 4°C for 30 min. The supernatant was collected, and the total amount of protein was determined by Bradford colorimetric assay (Bio-Rad). The whole cell lysates were incubated with 50 µg of the indicated GST fusion protein at 4°C for 1 h. Thus, glutathione-Sepharose 4B beads were blocked by incubating with 3% bovine serum albumin with rocking at 4°C for 1 h, and then after centrifugation for 3 min at 4000 × g, at 4°C, the dry beads were bound to lysates mixed with GST fusion proteins at 4°C for 1 h. The supernatant was discarded by centrifugation, and the beads were washed six times with lysis buffer for 3 min at 4000 × g, at 4°C, and then the dry beads

were resuspended in SDS sample buffer, boiled and analyzed by SDS-PAGE and Western blotting on nitrocellulose membrane.

### IMMUNOPRECIPITATION AND IMMUNOBLOT ANALYSIS

HeLa cells were lysed as described previously. The whole cell lysates were incubated with anti-Flag antibody conjugated to Sepharose beads over-night at 4°C. The beads were washed with lysis buffer, and the immunoprecipitated proteins were separated by SDS-PAGE, transferred onto a nitrocellulose membrane, and immunoblotted with antibodies. The immunoreactions were visualized using ECL detection system (Amersham Biosciences).

### ACKNOWLEDGMENTS

This work was supported by the European Community's Seventh Framework Programme FP7 under grant agreement no. 241955; SYSCILIA (to Marius Ueffing), FP7 grant agreement no. 278568, PRIMES (to Marius Ueffing and Karsten Boldt), FP7 grant agreement no. 241481, AFFINOMICS (to Gianni Cesareni and Marius Ueffing) and by the Telethon Italy grant GGP09243 and the FIRB Oncodiet project to Gianni Cesareni. We thank Marc Vidal for providing DUSP18, PPP2R1A and PPP3CA encoding plasmids. We thank Lawrence Banks for providing HA-SCRIB and PDZ3 and 4-GST the SCRIB PDZ constructs.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fgene.2014.00115/abstract>

**Figure S1 | Validation of some of the newly identified phosphatase interactions (A) HeLa cells were transiently transfected with Flag-DUSP26 expression plasmid.** Twenty-four hours post transfection, cells were serum and amino acids starved for 1 h or left untreated and then lysed. Whole protein extracts were immunoprecipitated with anti-Flag antibody to purify the DUSP26 phosphatase. The membranes were probed with anti-ATM (WB:  $\alpha$ -ATM) and anti-Flag (WB:  $\alpha$ -Flag) antibodies. **(B)** HeLa cells were transiently co-transfected with Flag-PTPN21 and with HA-DLC1 expression plasmids. Twenty-four hours post transfection, cells were serum and amino acids starved for 1 h or left untreated and then lysed. Whole protein extracts were immunoprecipitated with anti-Flag antibody to purify PTPN21 phosphatase. The membranes were probed with anti-HA (WB:  $\alpha$ -HA), anti-Flag (WB:  $\alpha$ -Flag) and anti-GRB2 antibodies. **(C)** HeLa cells were transiently co-transfected with Flag-PPP2CA and with HA-DLC1 expression plasmids. Twenty-four hours post transfection, cells were serum and amino acids starved for 1 h or left untreated and then lysed. Whole protein extracts were immunoprecipitated with anti-Flag antibody to purify PTPN21 phosphatase. The membranes were probed with anti-HA (WB:  $\alpha$ -HA) and anti-Flag (WB:  $\alpha$ -Flag) antibodies.

**Table S1 | Protein groups identified by mass spectrometry based proteomics of phosphatase pull-down are reported with protein quantification, number of peptides and intensities.**

**Table S2 | After statistical analysis, for each phosphatase the corresponding interactor is reported.** In the "Phosph-SF" column, the intensity value of each interactor in phosphatase transfected cells was divided by its intensity in not transfected cells (Control). In the "PhosphTNF-Phosph" column, the intensity value of each interactor in cells over-expressing the phosphatase and stimulated with TNF $\alpha$  was

divided by its intensity in transfected unstimulated cells. Finally in "PhosphTNF-SF" column, the intensity value of each interactor in cells over-expressing the phosphatase and stimulated with TNF $\alpha$  was divided by its intensity in not transfected cells. For each ratio, the corresponding significance B is reported.

**Table S3 | List of common contaminants was collected from the literature.**

**Table S4 | Experimental data describing the functional relationships between signaling proteins in the pathways of interest were collected from the literature (PMID column).** Each enzyme-substrate relationship is described as activating (1) or inhibitory (-1). For each protein, Uniprot ID and gene name have been reported.

**Table S5 | Experimental and literature extracted binary interactions, describing the paths from a phosphatase to its target in the RAS-PI3K pathway.**

### REFERENCES

- Bardelli, A., and Velculescu, V. E. (2005). Mutational analysis of gene families in human cancer. *Curr. Opin. Genet. Dev.* 15, 5–12. doi: 10.1016/j.gde.2004.12.009
- Bendall, S. C., Hughes, C., Stewart, M. H., Doble, B., Bhatia, M., and Lajoie, G. A. (2008). Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Mol. Cell. Proteomics* 7, 1587–1597. doi: 10.1074/mcp.M800113-MCP200
- Blanchetot, C., Chagnon, M., Dube, N., Halle, M., and Tremblay, M. L. (2005). Substrate-trapping techniques in the identification of cellular PTP targets. *Methods* 35, 44–53. doi: 10.1016/j.ymeth.2004.07.007
- Boldt, K., Mans, D. A., Won, J., Van Reeuwijk, J., Vogt, A., Kinkl, N., et al. (2011). Disruption of intraflagellar protein transport in photoreceptor cilia causes Leber congenital amaurosis in humans and mice. *J. Clin. Invest.* 121, 2169–2180. doi: 10.1172/JCI45627
- Breitkreutz, A., Choi, H., Sharom, J. R., Boucher, L., Neduva, V., Larsen, B., et al. (2010). A global protein kinase and phosphatase interaction network in yeast. *Science* 328, 1043–1046. doi: 10.1126/science.1176495
- Brill, L. M., Xiong, W., Lee, K. B., Ficarro, S. B., Crain, A., Xu, Y., et al. (2009). Phosphoproteomic analysis of human embryonic stem cells. *Cell Stem Cell* 5, 204–213. doi: 10.1016/j.stem.2009.06.002
- Cai, T., Nishida, K., Hirano, T., and Khavari, P. A. (2002). Gab1 and SHP-2 promote Ras/MAPK regulation of epidermal growth and differentiation. *J. Cell Biol.* 159, 103–112. doi: 10.1083/jcb.200205017
- Calderone, A., Castagnoli, L., and Cesareni, G. (2013). mentha: a resource for browsing integrated protein-interaction networks. *Nat. Methods* 10, 690–691. doi: 10.1038/nmeth.2561
- Cardone, L., Carlucci, A., Affaitati, A., Livigni, A., Decristofaro, T., Garbi, C., et al. (2004). Mitochondrial AKAP121 binds and targets protein tyrosine phosphatase D1, a novel positive regulator of src signaling. *Mol. Cell Biol.* 24, 4613–4626. doi: 10.1128/MCB.24.11.4613-4626.2004
- Carducci, M., Perfetto, L., Briganti, L., Paoluzi, S., Costa, S., Zerweck, J., et al. (2012). The protein interaction network mediated by human SH3 domains. *Biotechnol. Adv.* 30, 4–15. doi: 10.1016/j.biotechadv.2011.06.012
- Carlucci, A., Porpora, M., Garbi, C., Galgani, M., Santoriello, M., Mascolo, M., et al. (2010). PTPD1 supports receptor stability and mitogenic signaling in bladder cancer cells. *J. Biol. Chem.* 285, 39260–39270. doi: 10.1074/jbc.M110.174706
- Couzens, A. L., Knight, J. D., Kean, M. J., Teo, G., Weiss, A., Dunham, W. H., et al. (2013). Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci. Signal.* 6, rs15. doi: 10.1126/scisignal.2004712
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. doi: 10.1038/nbt.1511
- Dent, P., Jelinek, T., Morrison, D. K., Weber, M. J., and Sturgill, T. W. (1995). Reversal of Raf-1 activation by purified and membrane-associated protein phosphatases. *Science* 268, 1902–1906. doi: 10.1126/science.7604263
- Duan, L., and Cobb, M. H. (2010). Calcineurin increases glucose activation of ERK1/2 by reversing negative feedback. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22314–22319. doi: 10.1073/pnas.1016630108

- Fernandez, A. M., Jimenez, S., Mecha, M., Davila, D., Guaza, C., Vitorica, J., et al. (2012). Regulation of the phosphatase calcineurin by insulin-like growth factor I unveils a key role of astrocytes in Alzheimer's pathology. *Mol. Psychiatry* 17, 705–718. doi: 10.1038/mp.2011.128
- Gonfloni, S., Williams, J. C., Hattula, K., Weijland, A., Wierenga, R. K., and Superti-Furga, G. (1997). The role of the linker between the SH2 domain and catalytic domain in the regulation and function of Src. *EMBO J.* 16, 7261–7271. doi: 10.1093/emboj/16.24.7261
- Gloeckner, C. J., Boldt, K., Schumacher, A., Roepman, R., and Ueffing, M. (2007). A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics* 7, 4228–4234. doi: 10.1002/pmic.200700038
- Gloeckner, C. J., Boldt, K., and Ueffing, M. (2009). Strep/FLAG tandem affinity purification (SF-TAP) to study protein interactions. *Curr. Protoc. Protein Sci.* Chapter 19:Unit19.20. doi: 10.1002/0471140864.ps1920s57
- Goudreaux, M., D'Ambrosio, L. M., Kean, M. J., Mullin, M. J., Larsen, B. G., Sanchez, A., et al. (2009). A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol. Cell. Proteomics* 8, 157–171. doi: 10.1074/mcp.M800266-MCP200
- Graves, J. D., and Krebs, E. G. (1999). Protein phosphorylation and signal transduction. *Pharmacol. Ther.* 82, 111–121. doi: 10.1016/S0163-7258(98)00056-4
- Guo, A., Villen, J., Kornhauser, J., Lee, K. A., Stokes, M. P., Rikova, K., et al. (2008). Signaling networks assembled by oncogenic EGFR and c-Met. *Proc. Natl. Acad. Sci. U.S.A.* 105, 692–697. doi: 10.1073/pnas.0707270105
- Hahn, K., Miranda, M., Francis, V. A., Vendrell, J., Zorzano, A., and Teleman, A. A. (2010). PP2A regulatory subunit PP2A-B' counteracts S6K phosphorylation. *Cell Metab.* 11, 438–444. doi: 10.1016/j.cmet.2010.03.015
- Han, S., Williams, S., and Mustelin, T. (2000). Cytoskeletal protein tyrosine phosphatase PTPH1 reduces T cell antigen receptor signaling. *Eur. J. Immunol.* 30, 1318–1325. doi: 10.1002/(SICI)1521-4141(200005)30:5%3C1318::AID-IMMU1318%3E3.0.CO;2-G
- He, H., Liu, X., Lv, L., Liang, H., Leng, B., Zhao, D., et al. (2014). Calcineurin suppresses AMPK-dependent cytoprotective autophagy in cardiomyocytes under oxidative stress. *Cell Death Dis.* 5, e997. doi: 10.1038/cddis.2013.533
- Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., et al. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270. doi: 10.1093/nar/gkr1122
- Hu, Y., and Mivechi, N. F. (2006). Association and regulation of heat shock transcription factor 4b with both extracellular signal-regulated kinase mitogen-activated protein kinase and dual-specificity tyrosine phosphatase DUSP26. *Mol. Cell. Biol.* 26, 3282–3294. doi: 10.1128/MCB.26.8.3282-3294.2006
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Hurov, K. E., Cotta-Ramusino, C., and Elledge, S. J. (2010). A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes Dev.* 24, 1939–1950. doi: 10.1101/gad.1934210
- Julien, S. G., Dube, N., Hardy, S., and Tremblay, M. L. (2011). Inside the human cancer tyrosine phosphatome. *Nat. Rev. Cancer* 11, 35–49. doi: 10.1038/nrc2980
- Kao, S. C., Wu, H., Xie, J., Chang, C. P., Ranish, J. A., Graef, I. A., et al. (2009). Calcineurin/NFAT signaling is required for neuregulin-regulated Schwann cell differentiation. *Science* 323, 651–654. doi: 10.1126/science.1166562
- Kim, H., Lee, H. J., Oh, Y., Choi, S. G., Hong, S. H., Kim, H. J., et al. (2014). The DUSP26 phosphatase activator adenylate kinase 2 regulates FADD phosphorylation and cell growth. *Nat. Commun.* 5, 3351. doi: 10.1038/ncomms4351
- Li, X., Wilmanns, M., Thornton, J., and Kohn, M. (2013). Elucidating human phosphatase-substrate networks. *Sci. Signal.* 6, rs10. doi: 10.1126/scisignal.2003203
- Liang, N., Jia, L., Liu, Y., Liang, B., Kong, D., Yan, M., et al. (2013). ATM pathway is essential for ionizing radiation-induced autophagy. *Cell. Signal.* 25, 2530–2539. doi: 10.1016/j.cellsig.2013.08.010
- Liberti, S., Sacco, F., Calderone, A., Perfetto, L., Iannuccelli, M., Panni, S., et al. (2013). HuPho: the human phosphatase portal. *FEBS J.* 280, 379–387. doi: 10.1111/j.1742-4658.2012.08712.x
- Magnaudeix, A., Wilson, C. M., Page, G., Bauvy, C., Codogno, P., Leveque, P., et al. (2013). PP2A blockade inhibits autophagy and causes intraneuronal accumulation of ubiquitinated proteins. *Neurobiol. Aging* 34, 770–790. doi: 10.1016/j.neurobiolaging.2012.06.026
- Mann, M., and Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21, 255–261. doi: 10.1038/nbt0303-255
- Manning, G., Plowman, G. D., Hunter, T., and Sudarsanam, S. (2002a). Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* 27, 514–520. doi: 10.1016/S0968-0004(02)02179-5
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002b). The protein kinase complement of the human genome. *Science* 298, 1912–1934. doi: 10.1126/science.1075762
- Meixner, A., Boldt, K., Van Troys, M., Askenazi, M., Gloeckner, C. J., Bauer, M., et al. (2011). A QUICK screen for Lrrk2 interaction partners—leucine-rich repeat kinase 2 is involved in actin cytoskeleton dynamics. *Mol. Cell. Proteomics* 10, M110 001172. doi: 10.1074/mcp.M110.001172
- Nagasaka, K., Pim, D., Massimi, P., Thomas, M., Tomaic, V., Subbiah, V. K., et al. (2010). The cell polarity regulator hScrib controls ERK activation through a KIM site-dependent interaction. *Oncogene* 29, 5311–5321. doi: 10.1038/nc.2010.265
- Nagasaka, K., Seiki, T., Yamashita, A., Massimi, P., Subbiah, V. K., Thomas, M., et al. (2013). A novel interaction between hScrib and PP1gamma downregulates ERK signaling and suppresses oncogene-induced cell transformation. *PLoS ONE* 8:e53752. doi: 10.1371/journal.pone.0053752
- Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., et al. (2005). Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* 4, 2010–2021. doi: 10.1074/mcp.T500030-MCP200
- Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., et al. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376–386. doi: 10.1074/mcp.M200025-MCP200
- Patterson, K. I., Brummer, T., Daly, R. J., and O'Brien, P. M. (2010). DUSP26 negatively affects the proliferation of epithelial cells, an effect not mediated by dephosphorylation of MAPKs. *Biochim. Biophys. Acta* 1803, 1003–1012. doi: 10.1016/j.bbamer.2010.03.014
- Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., et al. (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131, 1190–1203. doi: 10.1016/j.cell.2007.11.025
- Roy, J., and Cyert, M. S. (2009). Cracking the phosphatase code: docking interactions determine substrate specificity. *Sci. Signal.* 2, re9. doi: 10.1126/scisignal.21000re9
- Sacco, F., Gherardini, P. F., Paoluzi, S., Saez-Rodriguez, J., Helmer-Citterich, M., Ragnini-Wilson, A., et al. (2012a). Mapping the human phosphatome on growth pathways. *Mol. Syst. Biol.* 8, 603. doi: 10.1038/msb.2012.36
- Sacco, F., Perfetto, L., Castagnoli, L., and Cesareni, G. (2012b). The human phosphatase interactome: an intricate family portrait. *FEBS Lett.* 586, 2732–2739. doi: 10.1016/j.febslet.2012.05.008
- Shang, X., Vasudevan, S. A., Yu, Y., Ge, N., Ludwig, A. D., Wesson, C. L., et al. (2010). Dual-specificity phosphatase 26 is a novel p53 phosphatase and inhibits p53 tumor suppressor functions in human neuroblastoma. *Oncogene* 29, 4938–4946. doi: 10.1038/nc.2010.244
- Skarra, D. V., Goudreaux, M., Choi, H., Mullin, M., Nesvizhskii, A. I., Gingras, A. C., et al. (2011). Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5. *Proteomics* 11, 1508–1516. doi: 10.1002/pmic.201000770
- Taniguchi, T., Garcia-Higuera, I., Xu, B., Andreassen, P. R., Gregory, R. C., Kim, S. T., et al. (2002). Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. *Cell* 109, 459–472. doi: 10.1016/S0092-8674(02)00747-X
- Tanuma, N., Nomura, M., Ikeda, M., Kasugai, I., Tsubaki, Y., Takagaki, K., et al. (2009). Protein phosphatase Dusp26 associates with KIF3 motor and promotes N-cadherin-mediated cell-cell adhesion. *Oncogene* 28, 752–761. doi: 10.1038/nc.2008.431
- Tonikian, R., Zhang, Y., Sazinsky, S. L., Currell, B., Yeh, J. H., Reva, B., et al. (2008). A specificity map for the PDZ domain family. *PLoS Biol.* 6:e239. doi: 10.1371/journal.pbio.0060239
- Tonks, N. K. (2006). Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* 7, 833–846. doi: 10.1038/nrm2039
- Tremblay, M. L. (2009). The PTP family photo album. *Cell* 136, 213–214. doi: 10.1016/j.cell.2009.01.006

- Tripathi, D. N., Chowdhury, R., Trudel, L. J., Tee, A. R., Slack, R. S., Walker, C. L., et al. (2013). Reactive nitrogen species regulate autophagy through ATM-AMPK-TSC2-mediated suppression of mTORC1. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2950–E2957. doi: 10.1073/pnas.1307736110
- Wera, S., and Hemmings, B. A. (1995). Serine/threonine protein phosphatases. *Biochem. J.* 311(pt 1), 17–29.
- Yano, S., Arroyo, N., and Yano, N. (2004). SHP2 binds catalase and acquires a hydrogen peroxide-resistant phosphatase activity via integrin-signaling. *FEBS Lett.* 577, 327–332. doi: 10.1016/j.febslet.2004.10.011

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 February 2014; accepted: 16 April 2014; published online: 07 May 2014.  
Citation: Sacco F, Boldt K, Calderone A, Panni S, Paoluzi S, Castagnoli L, Ueffing M and Cesareni G (2014) Combining affinity proteomics and network context to identify new phosphatase substrates and adapters in growth pathways. *Front. Genet.* 5:115. doi: 10.3389/fgene.2014.00115

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Sacco, Boldt, Calderone, Panni, Paoluzi, Castagnoli, Ueffing and Cesareni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Phosphatases of $\alpha$ -synuclein, LRRK2, and tau: important players in the phosphorylation-dependent pathology of Parkinsonism

Jean-Marc Taymans\* and Veerle Baekelandt

Department of Neurosciences, Laboratory for Neurobiology and Gene Therapy, KU Leuven, Leuven, Belgium

## Edited by:

Allegra Via, Sapienza University, Italy  
Andreas Zanzoni, Inserm  
Technological Advances for Genomics  
and Clinics, UMR1090, France

## Reviewed by:

Eric Yang, Johnson and Johnson, USA  
Christian Johannes Gloeckner,  
Helmholtz Zentrum München,  
Germany

## \*Correspondence:

Jean-Marc Taymans, Jean-Pierre  
Aubert Research Center, UMR837,  
Rue Polonovski – 1 Place de Verdun,  
59045 Lille, France  
e-mail: jean-marc.taymans@inserm.fr

An important challenge in the field of Parkinson's disease (PD) is to develop disease modifying therapies capable of stalling or even halting disease progression. Coupled to this challenge is the need to identify disease biomarkers, in order to identify pre-symptomatic hallmarks of disease and monitor disease progression. The answer to these challenges lies in the elucidation of the molecular causes underlying PD, for which important leads are disease genes identified in studies investigating the underlying genetic causes of PD. LRRK2 and  $\alpha$ -syn have been both linked to familial forms of PD as well as associated to sporadic PD. Another gene, microtubule associated protein tau (MAPT), has been genetically linked to a dominant form of frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17) and genome-wide association studies report a strong association between MAPT and sporadic PD. Interestingly, LRRK2,  $\alpha$ -syn, and tau are all phosphorylated proteins, and their phosphorylation patterns are linked to disease. In this review, we provide an overview of the evidence linking LRRK2,  $\alpha$ -syn, and tau phosphorylation to PD pathology and focus on studies which have identified phosphatases responsible for dephosphorylation of pathology-related phosphorylations. We also discuss how the LRRK2,  $\alpha$ -syn, and tau phosphatases may point to separate or cross-talking pathological pathways in PD. Finally, we will discuss how the study of phosphatases of dominant Parkinsonism proteins opens perspectives for targeting pathological phosphorylation events.

**Keywords: PP1, PP2A, phosphorylation, phosphatase, Parkinson disease, LRRK2, alpha-synuclein, tauopathies, tau proteins**

## INTRODUCTION

Parkinson's disease (PD) is an incurable disease of aging characterized by the progressive death of dopaminergic cells in the midbrain as well as by  $\alpha$ -synuclein-rich intracytoplasmic depositions called Lewy bodies (LBs). Treatments which alleviate disease symptoms have been available for several decades; however, these do not halt disease progression. The development of disease-modifying therapies to replace the symptomatic treatments is therefore a major priority in the biomedical research field. Genetic studies of families with a history of PD (genetic linkage studies) as well as of PD patient groups compared to matched groups of healthy individuals (genetic association studies) have identified genes and genomic variants which contribute to the development of PD (Gasser, 2009; Nalls et al., 2011). These studies have revealed at least 20 PD genes, many of which are currently the subject of studies aiming to understand their biology and disease mechanisms. For instance, several PD genes, such as parkin, DJ-1, Pten induced kinase 1 (PINK1), or ATP13A2, contribute to early onset autosomal recessive forms of Parkinsonism. Other genes are linked to Parkinsonism in an autosomal dominant fashion and are responsible for early onset forms of PD ( $\alpha$ -synuclein duplications or triplications, some families with mutated  $\alpha$ -synuclein) as well as the more common late onset forms of Parkinsonism ( $\alpha$ -synuclein mutants, tau, LRRK2, VPS35, or EIF4G1; Houlden and Singleton, 2012). In this review, we focus

on the dominant proteins  $\alpha$ -synuclein, tau, and LRRK2 in light of the importance of their phosphorylation for their biological functioning.

Mutations in  $\alpha$ -synuclein (SNCA, PARK1/4) and mutations in leucine-rich repeat kinase type 2 (LRRK2, PARK8) are linked to autosomal-dominant forms of PD (Gasser, 2009). Also, although protein deposition of microtubule associated protein tau (MAPT) is a feature of Alzheimer's disease (AD), MAPT gene mutations cause frontotemporal dementia (FTD) with Parkinsonism. Interestingly, these three dominant genes in Parkinsonism (MAPT, SNCA, and LRRK2) have also been identified as risk factors for sporadic PD in genome-wide association studies (GWAS; Taymans and Cookson, 2010; Sharma et al., 2012). The dominant mode of disease transmission through these genes also suggests a gain of toxic function mechanism pointing to an inhibition of toxic function as potential therapeutic strategies.

LRRK2,  $\alpha$ -syn, and tau are all phosphorylated proteins, and their phosphorylation patterns are linked to disease (Lobbestael et al., 2012; Tenreiro et al., 2014). Early work showed that hyperphosphorylation of tau is correlated to pathology of tauopathies and phosphorylation of  $\alpha$ -syn at serine129 is correlated to synucleinopathies (for reviews, see references Martin et al., 2011; Tenreiro et al., 2014); therefore much work has focused on identifying and characterizing kinases of these proteins (for reviews, see references

Vancraenenbroeck et al., 2011; Martin et al., 2013b; Tenreiro et al., 2014). The characterization of LRRK2 phosphorylation and the link to disease is still underway although some evidence suggests that a site-dependent mixed phosphorylation state is indicative of disease. Tau and synuclein kinases have been considered as potential therapeutic targets for synucleinopathies and tauopathies and several compounds have been developed for these kinases and tested in preclinical models (for reviews on these topics, see references Vancraenenbroeck et al., 2011; Kramer et al., 2012; Tell and Hilgeroth, 2013). In this review, we will discuss the second main component in the regulation of protein phosphorylation of LRRK2,  $\alpha$ -syn, and tau, namely phosphatases. We will briefly introduce the three proteins and discuss what is known about their dephosphorylation and which phosphatases and phosphatase regulators are involved. We will also discuss the relationships between the three proteins with regards to their cognate phosphatases and discuss targeting of phosphatase holoenzymes of LRRK2,  $\alpha$ -syn, and tau as a potential phosphomodulatory therapeutic approach.

## ALPHA-SYNUCLEIN

The involvement of  $\alpha$ -syn in PD was initially identified through genetic linkage studies in a small number of families (Polymeropoulos et al., 1997), including mutations as well as gene duplications (Chartier-Harlin et al., 2004) and triplications (Singleton et al., 2003). Recently, strong association was shown between  $\alpha$ -syn and sporadic PD in GWAS (Satake et al., 2009; Simon-Sanchez et al., 2009). Also,  $\alpha$ -syn is a major component of LBs (Spillantini et al., 1997). These arguments illustrate that  $\alpha$ -syn is a central player in the pathogenesis of PD.

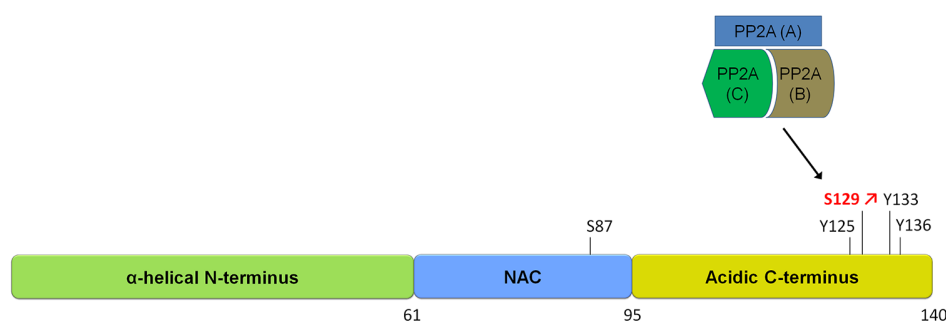
Studies investigating the phosphorylation of  $\alpha$ -syn in diseased and aged brains have shown that  $\alpha$ -syn can be phosphorylated at serines (S87, S129) as well as at several tyrosines including Y125, Y133, and Y136 (Figure 1). The pY125 modification has been reported to be inversely correlated with PD-related pathology. Indeed, pY125 appears to protect brains against  $\alpha$ -syn mediated toxicity, as this modification is reduced in aged human brain tissue and absent in brain tissue affected by Lewy body dementia (Chen and Feany, 2005; Chen et al., 2009). The pS129 modification on the other hand is most often correlated with PD

pathology. This notion is primarily supported by the finding that the majority of  $\alpha$ -syn in LBs in postmortem PD brains is phosphorylated at S129 (pS129; Fujiwara et al., 2002; Hasegawa et al., 2002; Anderson et al., 2006). The S129 phosphorylation of  $\alpha$ -syn in aggregates has also been observed in animal models of PD (Kahle et al., 2000; Neumann et al., 2002; Takahashi et al., 2003). Mechanistic studies have shown that aggregated forms of  $\alpha$ -syn are more prone to phosphorylation and that pS129 phosphorylated aggregates accumulate as the disease progresses (Waxman and Giasson, 2008; Mbefo et al., 2010; Paleologou et al., 2010; Waxman and Giasson, 2011), suggesting that the degree of  $\alpha$ -syn pS129 phosphorylation is an indicator of disease progression.

The link between S129 phosphorylation and PD pathology has fueled an interest in modulating  $\alpha$ -syn phosphorylation at S129 as a potential therapy for PD (Vancraenenbroeck et al., 2011). Multiple kinases have been identified which phosphorylate  $\alpha$ -syn at S129, with most evidence pointing to polo-like kinase 2 (PLK2) as the primary phosphorylated of  $\alpha$ -syn S129 [for an extended up to date review of  $\alpha$ -syn phosphorylation, please refer to Tenreiro et al. (2014)]. A straightforward therapeutic approach based on reducing  $\alpha$ -syn phospho-S129 would be to inhibit PLK2 kinase activity; however some contradictory findings should be taken into account. For instance, overexpression of PLK2 in rat brain using adeno-associated viral vectors can suppress  $\alpha$ -syn toxicity by promoting autophagy-mediated degradation of phospho-S129  $\alpha$ -syn (Oueslati et al., 2013). Therefore, therapies based on modulating  $\alpha$ -syn phospho-S129 appears to require an optimal phosphorylation level rather than a complete dephosphorylation.

## PHOSPHATASES OF $\alpha$ -SYN

Few studies have sought to identify phosphatases of  $\alpha$ -syn (Braithwaite et al., 2012); however, concurring data point to PP2A as a major phosphatase of the S129 site. For instance, PP2A enzyme but not PP1 is shown to dephosphorylate  $\alpha$ -syn-pS129 *in vitro* (Lee et al., 2011), and treatment of cells with the PP2A inhibitor okadaic acid (OA) but not the PP1 inhibitor tautomycin leads to an increased level of  $\alpha$ -syn-pS129. Further characterization showed that  $\alpha$ -syn-pS129 was increased upon knockdown of the PP2A catalytic subunit and when PP2A enzyme is methylated.



**FIGURE 1 | Schematic of  $\alpha$ -synuclein and its phosphorylation sites.**  $\alpha$ -syn is a small protein of 140 amino acids in length. It is subdivided into three domains, an N-terminal alpha-helical domain, a central NAC domain (standing for non-Abeta-component) and an acidic C-terminal

domain. The S129 site is hyperphosphorylated in disease and is regulated by a phosphatase of the heterotrimeric PP2A class. Please refer to **Table 1** for an overview of studies on phosphatases regulating  $\alpha$ -syn phosphorylation.

It is important to note that the phosphatases of the PP2A class function in complexes called holoenzymes, which are composed of regulatory and catalytic phosphatase subunits. In the case of PP2A phosphatases, these are composed of a catalytic subunit (PPP2CA or PPP2CB) together with a scaffold subunit ( $A\alpha$  or  $A\beta$ ) and a regulatory subunit (of which there are four families, B, B', B'', and B''', each with 2–5 different members). The precise heterotrimeric composition of the holoenzyme guides PP2A to specific substrate sites. Accordingly, the testing of four different holoenzyme compositions shows that holoenzymes with regulatory subunits of the B family are more efficient at dephosphorylating  $\alpha$ -syn-pS129 than those of the B' and B'' families. Interestingly,  $\alpha$ -syn may function in a feedback loop with PP2A, with studies reporting that  $\alpha$ -syn has the ability to activate PP2A activity (Peng et al., 2005) and that phospho-S129  $\alpha$ -syn is less efficient at activating PP2A (Lou et al., 2010).

PP2A enzymes have the particularity that their enzymatic activity is positively regulated by its methylation which is itself regulated via the opposing activities of a PP2A-specific methyltransferase and a PP2A-specific methyltransferase (PME). Accordingly, treatment of mice with the PME inhibitor eicosanoyl-5-hydroxytryptamide (EHT) increases PP2A methylation as well as decreased  $\alpha$ -syn-pS129 levels in brain and a concurrent reduction in synuclein pathology (Lee et al., 2011). Related to this, the diabetes drug metformin was shown to reduce  $\alpha$ -syn phosphorylation at S129 through activation of PP2A and inhibition of mammalian target of rapamycin (mTOR; Perez-Revuelta et al., 2014). These studies confirm a primordial role of PP2A in the phosphoregulation of  $\alpha$ -syn at S129 and also provide a proof-of-principle that phosphorylation levels of  $\alpha$ -syn can be modulated by targeting its phosphatases. Thus far, no phosphatases have been described for the tyrosine phosphorylation sites of  $\alpha$ -syn. An overview of  $\alpha$ -syn phosphatases is given in **Table 1**.

## LRRK2

LRRK2 is a complex protein of 2527 amino acids containing several predicted functional domains (**Figure 2**). Several arguments underline the importance of LRRK2 for PD. First, LRRK2 is one of the most prevalent causes of monogenic PD. Furthermore, LRRK2

mutations are present in apparently sporadic cases of PD, with prevalences of 2% to up to 40% in certain population groups and LRRK2 was recently genetically associated to PD in several independent GWAS (Satake et al., 2009; Simon-Sanchez et al., 2009). Finally, PD patients carrying the LRRK2 mutations show a clinical and neuropathological profile which is virtually indistinguishable to sporadic PD (Healy et al., 2008), indicating that LRRK2 may contribute to a PD disease pathway common to both familial and sporadic PD.

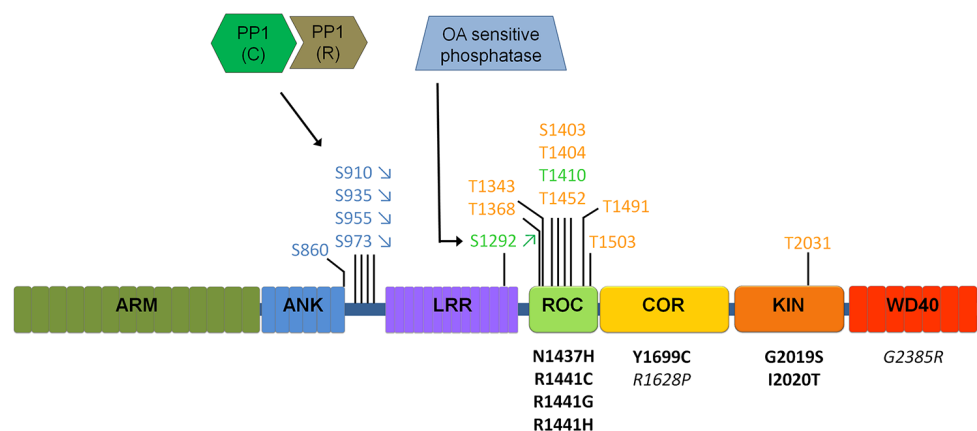
LRRK2 is a highly phosphorylated protein and phosphosite mapping studies have distinguished two notable clusters of phosphorylation sites, one in or near the Ras of complex proteins (ROC) domain (Greggio et al., 2009; Kamikawaji et al., 2009) and another in the interdomain region between the ankyrin repeat (ANK) and leucine rich repeat (LRR) domains (see **Figure 2**; West et al., 2007; Gloeckner et al., 2010; Nichols et al., 2010; see reference Lobbstaël et al., 2012 for a detailed overview of studies reporting LRRK2 phosphorylation sites).

The physiological and pathological relevance of LRRK2 phosphorylation has only just begun to be described. For instance, phosphosites of the ANK-LRR interdomain region, including the S910/S935/S955/S973 sites, are sites phosphorylated by upstream kinases. These sites are detectable in basal conditions in multiple cellular and tissular systems, including for endogenous LRRK2 in Swiss 3T3 or NIH3T3 cells (Nichols et al., 2010; Lobbstaël et al., 2013), mouse primary neurons (Lobbstaël et al., 2013), mouse brain, kidney, and lung (Deng et al., 2011; Choi et al., 2012; Zhang et al., 2012; Delbroek et al., 2013), mouse embryonic fibroblasts (Dzamko et al., 2012), mouse bone marrow derived macrophages (Dzamko et al., 2012) and human peripheral blood mononuclear cells (Dzamko et al., 2013). The S910 and S935 sites mediate an interaction of LRRK2 with 14-3-3 proteins and regulate LRRK2 cellular localization (Nichols et al., 2010). The search for the kinases responsible for phosphorylating LRRK2 at this cluster is still ongoing. Studies *in vitro* and in COS-7 cells have suggested a role for protein kinase A as an upstream kinase of the S910–S935 sites (Muda et al., 2014), however, these findings are not confirmed in other cell types such as HEK293T cells (Reynolds et al., 2014), suggesting cell-specific mechanisms of phosphorylation. This is

**Table 1 | Overview of studies reporting phosphatases regulating  $\alpha$ -syn phosphorylation.**

Site	Identified phosphatase	Comments, references
S129	PP2A	<i>In vitro</i> incubation with PP2A [catalytic subunit C $\alpha$ (PPP2CA), regulatory subunit B $\alpha$ (PPP2R2A)] dephosphorylates $\alpha$ -syn S129 better than with B' $\alpha$ (PPP2R5A), B' $\gamma$ (PPP2R5C), or B'' (PPP2R3A). Knockdown of PPP2CA or PPP2R2A increases $\alpha$ -syn S129 phosphorylation (Lee et al., 2011).
	PPP2R1A	siRNA mediated knockdown of PPP2R1A (Protein phosphatase 2 (formerly 2A), regulatory subunit A, alpha isoform) reported to decrease $\alpha$ -syn S129 phosphorylation levels in neuroblastoma derived cells (Henderson-Smith et al., 2013).
	PTPRA	siRNA mediated knockdown of PPP3R1 (Protein tyrosine phosphatase, receptor type, A) reported to decrease $\alpha$ -syn S129 phosphorylation levels in neuroblastoma derived cells (Henderson-Smith et al., 2013).
	PPP3R1	siRNA mediated knockdown of PPP3R1 (Protein phosphatase 3 (formerly 2B), regulatory subunit B, alpha isoform) found to increase $\alpha$ -syn S129 phosphorylation levels in neuroblastoma derived cells (Henderson-Smith et al., 2013).

See text for references to reviews discussing  $\alpha$ -syn kinases.



**FIGURE 2 | Schematic of LRRK2 domain structure and currently identified LRRK2 phosphatases.**

Phosphorylation sites are indicated above the LRRK2 schematic, for those sites confirmed independently by at least two different groups. Sites represented in blue indicate cellular transphosphorylation sites, in orange *in vitro* autophosphorylation sites and in green autophosphorylation sites confirmed in cells. Disease mutants are indicated below the schematic in black, mutations in italics

are risk factor mutations. Arrows (for S910/S935/S955/S973/S1292) indicate the most common regulation of phosphorylation (up or down) observed across most disease mutants. Different sites are likely regulated in different ways, including through PP1 and a yet to be defined phosphatase sensitive to okadaic acid (OA). Please refer to **Table 2** comparing the regulation of phosphorylation at S935 compared to S1292.

further supported by the work of Dzamko et al. (2012) who show that inhibitor of kappa B kinases (IKKs) phosphorylate LRRK2 in bone marrow-derived macrophages upon activation of toll-like receptor signaling which is specific to immune cells. The phosphorylation pattern of LRRK2 is completely different from that of its closest homolog LRRK1 which does not contribute to PD, suggesting that phosphorylation regulation of LRRK2 is a potential mechanism distinguishing LRRK2 from LRRK1 functionally.

Phosphorylation levels of the ANK-LRRK interdomain phosphosites are reduced for several pathogenic mutants such as R1441C/G, Y1699C, I2020T (Nichols et al., 2010; Li et al., 2011; Lobbstaël et al., 2013). This observation suggests that the reduction in LRRK2 phosphorylation levels may be involved in the pathogenic mechanism of LRRK2 PD. A corollary of that conclusion is that reduced LRRK2 phosphorylation may be used as a biomarker, however, there are some caveats. For instance, the most prevalent LRRK2 variant in patients, G2019S, does not display reduced phosphorylation levels and a recent study reported no differences in LRRK2 S935 phosphorylation in PBMCs of PD patients compared to matched healthy individuals (Dzamko et al., 2013). Nevertheless, the striking phosphorylation reduction at the ANK-LRR sites seen in all other confirmed LRRK2 pathogenic mutants warrants further evaluation as a disease or diagnostic biomarker.

The other major group of phosphorylation sites for LRRK2 is comprised of autophosphorylation sites. These are sites which were initially identified on LRRK2 protein after *in vitro* incubation with ATP to allow the protein to autophosphorylate itself (Greggio et al., 2009; Kamikawaji et al., 2009; Gloeckner et al., 2010). The majority of these sites cluster in the ROC domain and studies with phospho-mimicking mutants show that at least some of these modifications (T1491D, T1503D) alter LRRK2 GTP-binding properties (Kamikawaji et al., 2009; Webber et al., 2011).

The precise physiological relevance of autophosphorylation sites is unknown since the majority of these sites are undetectable in cells, even on overexpressed protein. The notable exceptions are the T1410 site located in the ROC domain identified in overexpressed LRRK2 in HEK293T cells (Pungaliya et al., 2010), the S1058 (Reyniers et al., 2014), and the S1292 site (Gloeckner et al., 2010; Pungaliya et al., 2010; Sheng et al., 2012), located in the 3rd and 13th of the 14 leucine-rich repeats of the leucine-rich repeat domain, respectively, just outside the ROC domain (Vancaenenbroeck et al., 2012). These reports suggest that at least some autophosphorylation events are occurring in cells. Specifically, the S1292 site has been characterized in more detail and displays a number of interesting features. The S1292 site is one of the few autophosphorylation sites located outside ROC. The site is phosphorylated at detectable levels in basal conditions in LRRK2 overexpressed in cell lines or in transgenic mice (Sheng et al., 2012) as well as on endogenous LRRK2 in lymphocytes (Reynolds et al., 2014). In contrast to what is described for the ANK-LRR interdomain sites, the S1292 levels are increased in cells for the majority of pathogenic mutants and decreased in LRRK2 kinase-dead variants (Sheng et al., 2012; Reynolds et al., 2014). Increased phosphorylation levels at S1292 may therefore be indicative of LRRK2's pathogenic state; however, this remains to be tested in PD patients. Because the kinase activating mutant G2019S shows increased phospho-S1292 and kinase dead mutants show reduced phospho-S1292, it may be suggested that phospho-S1292 levels are indicative of LRRK2 kinase activity in cells, however, some discrepancies appear. For instance, other mutants, such as N1437H, R1441C, or R1441G which are significantly less active in their kinase activity than G2019S nevertheless display similar phospho-S1292 levels relative to the G2019S (Sheng et al., 2012; Reynolds et al., 2014). Two other pathogenic mutants, Y1699C and I2020T, which display kinase activity similar to WT or slightly increased, show varying phospho-S1292



levels depending on the system tested. Indeed, the Y1699C mutant displays increased phospho-S1292 levels in stable overexpression HEK293T cells (Reynolds et al., 2014) and unchanged phospho-S1292 levels in transfected HEK293T cells (Sheng et al., 2012), while the inverse is true for the I2020T mutant. Further work will be required to elucidate the precise regulation of phospho-S1292 levels in LRRK2. It also remains to be shown whether the conclusions here for S1292 (an autophosphorylation site which occurs in cells) also hold true for other autophosphorylation sites or whether this is a new 'class' of sites (besides the ROC autophosphorylation cluster and the ANK-LRR interdomain cluster).

Of importance for the development of LRRK2 kinase inhibitors is that the majority of these sites is downregulated by kinase inhibitors and may therefore be used to assess inhibitor activity. For instance, the phosphorylation of autophosphorylation sites such as S1491 is inhibited by LRRK2 kinase inhibitors *in vitro* (Doggett et al., 2012). Also, cellular treatment with kinase inhibitors leads to a dephosphorylation of the ANK-LRR interdomain sites S910/S935/S955/S973 (Dzamko et al., 2010; Deng et al., 2011; Doggett et al., 2012) as well as the S1292 autophosphorylation site (Sheng et al., 2012; Reynolds et al., 2014). Intriguingly, although the ANK-LRR interdomain phosphorylations are not autophosphorylation sites, the observed dephosphorylation of LRRK2 by kinase inhibitors can be attributed to the activity of the inhibitors on LRRK2 itself.

In sum, the emerging picture of LRRK2 phosphorylation is that LRRK2 is a highly phosphorylated protein where at least two, perhaps three, classes of phosphorylation sites can be discerned (see **Figure 2**). A first class of phosphosites in LRRK2 is the *in vitro* autophosphorylation site class. These sites appear after *in vitro* autophosphorylation by purified LRRK2; however, their presence in cells is not confirmed. While these sites offer opportunities to develop assays of LRRK2 kinase activity, their physiological relevance is unclear. Further work will be required to determine whether these sites occur in physiological systems under specific conditions of activation, or whether the appearance of these phosphorylations is an *in vitro* phenomenon. A second class of LRRK2 phosphorylation sites may be termed 'cellular' phosphorylation sites, including those sites of the ANK-LRR interdomain region introduced above, exemplified by the S935 site. Finally, the third and most recent class of LRRK2 phosphorylation sites is the class of cellular autophosphorylation sites, exemplified by the S1292 site, which are detected in cells and which also increase after *in vitro* autophosphorylation. The cellular sites (S935) and cellular autophosphorylation sites (S1292) are the most physiologically relevant and the comparison of both sites (summarized in **Table 2**) suggests that these may be useful indicators of LRRK2 activity or pathology. With a few exceptions, pathogenic mutants of LRRK2 display decreased phospho-S935 levels and increased phospho-S1292 levels. It remains to be confirmed whether these changes can be used as diagnostic or disease biomarkers, either individually or together. Interestingly, cellular treatment with kinase inhibitors leads to a reduction of both S935 and S1292. Therefore, both sites are also useful as pharmacodynamic marker to assess activity of kinase inhibitors in cellular and animal models.

**Table 2 | Comparison of the two different LRRK2 phosphorylation sites pS935 and pS1292.**

Parameter	pS935	pS1292
Present in cells	YES	YES
Change after <i>in vitro</i> autophosphorylation	NO CHANGE	INCREASE
Change after cellular kinase inhibition	DECREASE	DECREASE
N1437H vs. WT	/	HIGHER
R1441C/G vs. WT	LOWER	HIGHER
Y1699C vs. WT	LOWER	SIMILAR
G2019S vs. WT	NO CHANGE	HIGHER
I2020T vs. WT	LOWER	HIGHER
G2385R vs. WT	LOWER	LOWER
Regulation	IKK, PP1, +tbd	AutoP, +tbd
Physiological role	14-3-3 binding, subcellular localization	Putative role in neurite length regulation

Indicated is the regulation of phosphorylation levels at these sites by autophosphorylation, cellular kinase inhibition, or disease mutants. AutoP, autophosphorylation; tbd, to be discovered.

## PHOSPHATASES OF LRRK2

We recently reported that protein phosphatase 1 (PP1) is a main phosphatase of the LRRK2 ANK-LRR interdomain sites (Lobbestael et al., 2013). The study first shows that of a panel of recombinant serine/threonine phosphatases, only protein phosphatase 1 can efficiently dephosphorylate LRRK2 *in vitro*. *In vitro* dephosphorylation was demonstrated on purified LRRK2 protein which was previously metabolically labeled by radioactive phosphates, showing that PP1 is responsible for dephosphorylation at the majority of LRRK2's phosphosites, a finding confirmed for 4 sites with phospho-specific antibodies, i.e., S910, S935, S955, and S973. Upon pharmacological inhibition of cells with either PP1 or PP2A phosphatase inhibitors, it was observed that PP1 but not PP2A inhibition could reverse LRRK2 dephosphorylation.

Interestingly, the effects of PP1 in LRRK2 phosphorylation could be confirmed in multiple cell types including HEK293T, SH-SY5Y neuroblastoma cells, mouse primary cortical neurons, U2OS osteosarcoma cells, NIH3T3 mouse fibroblast cells and A549 human lung cancer cells. This shows that PP1 is active as a LRRK2 phosphatase independent of the cell type tested, and it may be predicted that PP1 can dephosphorylate LRRK2 throughout the body.

Similar to PP2A phosphatases, PP1 class phosphatases are holoenzymes which are composed of one catalytic subunit, responsible for catalyzing the actual dephosphorylation event, and one regulatory subunit, responsible for directing the holoenzyme to its specific substrates. There are more than 150 PP1 regulatory subunits reported, allowing several 100 possible holoenzyme compositions (Bollen et al., 2010). This mode of functioning is necessary given that only three PP1 catalytic subunits are expressed in mammalian cells (PP1 $\alpha$ , PP1 $\beta$ , and PP1 $\gamma$ ; HGNC codes

PPP1CA, PPP1CB, and PPP1CC) which on their own are insufficiently diverse to account for the specificity in the huge volume of phosphatase activity mediated by PP1. Indeed, PP1 together with PP2A (which is represented by only two catalytic subunits, PPP2CA and PPP2CB, see below) account for more than 90% of the protein phosphatase activity in eukaryotes (Moorhead et al., 2007; Virshup and Shenolikar, 2009). This is in stark contrast to the diversity for instance of kinases, of which there are ~400 serine/threonine kinases (Manning et al., 2002). Therefore, a key issue is to identify the composition of the PP1 holoenzyme by identifying the LRRK2-specific PP1 regulatory subunit which associates with the PP1 catalytic subunit.

There is little data available on the phosphatases involved in the regulation of LRRK2 phosphosites outside of the ANK-LRR interdomain region, however, initial evidence suggests that other phosphatases are at play. Work done at the S1292 site shows that inhibitor induced dephosphorylation of LRRK2 at S1292 is insensitive to the phosphatase inhibitors calyculin A (mixed PP1 and PP2A inhibitor) and OA (selective PP2A inhibitor; Reynolds et al., 2014), in contrast to what is observed at the S935 site where inhibitor induced dephosphorylation is inhibited by calyculin A (Lobbestael et al., 2013). However, the low basal S1292 phosphorylation levels of the R1441G mutant is upregulated by both calyculin A and OA treatment (Reynolds et al., 2014), while the S935 phosphorylation levels of the same mutant is only upregulated by calyculin A (Lobbestael et al., 2013). These findings suggest the hypothesis that PP2A, rather than PP1, is the phosphatase system regulating R1441G LRRK2 at S1292.

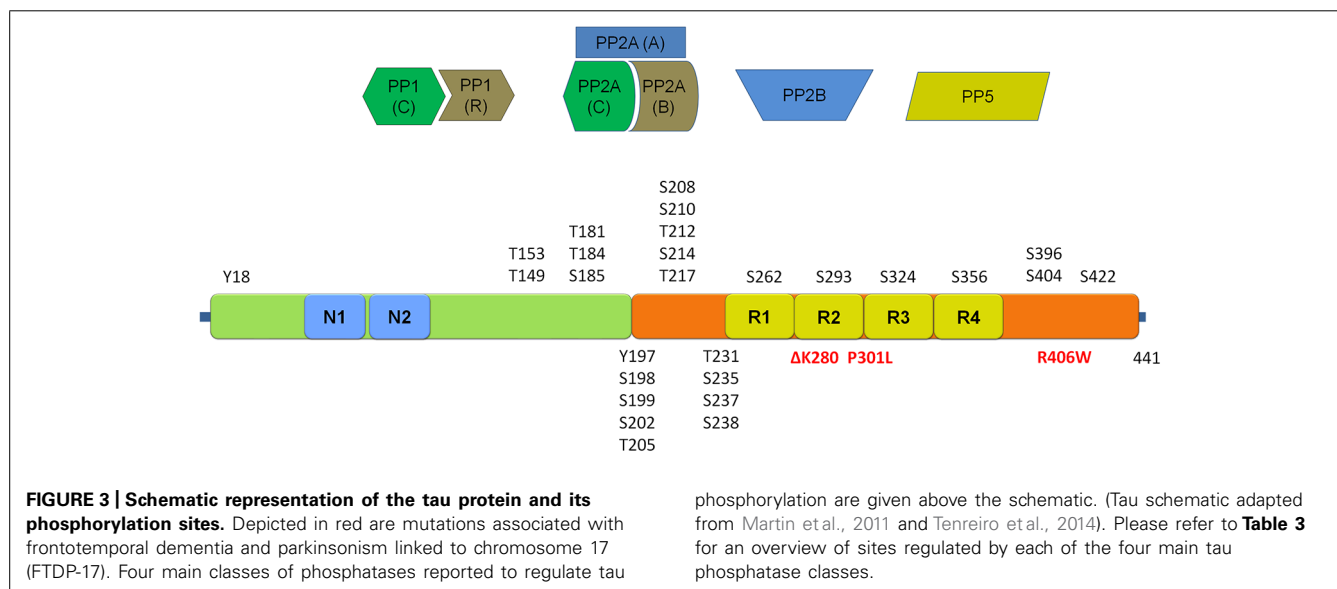
## TAU

It may seem surprising to discuss Tau in relation to PD pathogenesis, as this protein has a relatively long history as a protein involved in neurodegenerative dementias; however, accumulating evidence puts this protein to the forefront in PD with a number of reports pointing to specific properties of Tau in PD distinguishing it from Tau in other neurodegenerative disorders. Tau is a microtubule

associating protein which is involved in several neurodegenerative diseases including AD, progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), and some cases of frontotemporal lobar dementia (FTLD). Although tau is mostly associated to dementias, the tau gene has also been identified as a risk factor for PD via genome wide association studies (GWAS; Sharma et al., 2012). The genetic association of MAPT locus variants with PD is a striking finding, and is in stark contrast with the fact that no genetic associations of the MAPT locus are observed in AD (Lambert et al., 2013), showing that tau contributes to both AD and PD but via separate mechanisms.

Tau can occur in six different splice isoforms ranging in size from 352 to 441 amino acids and at least 45 potential tau phosphorylation sites have been reported, including serine, threonine and tyrosine phosphorylation sites [see figure schematic representation of tau protein and the localization of phosphorylation sites, **Figure 3**; for reviews of tau phosphorylation, please refer to Martin et al. (2013a,b) and Tenreiro et al. (2014)]. Several kinases have been reported to phosphorylate tau. These include proline directed kinases [glycogen synthase kinase 3 (GSK3), cyclin dependent kinase 5 (CDK5), and 5' adenosine monophosphate-activated protein kinase (AMPK)], non-proline directed kinases [casein kinase 1 (CK1), microtubule affinity regulating kinases (MARKs), death associated protein kinase 1 (DAPK1), cyclic AMP-dependent protein kinase A (PKA) and dual specificity tyrosine-phosphorylation regulated kinase 1A (DYRK-1A)] as well tyrosine kinases including Fyn, Abl, and Syk. The inhibition of tau phosphorylation has been proposed as a therapeutic strategy in tauopathies including early phase clinical testing of GSK-3 $\beta$  inhibition (Del Ser et al., 2013).

Tau phosphorylation is important for tau function/dysfunction as hyperphosphorylation of tau is generally correlated with the formation of tau protein aggregates which are major components of neurofibrillary tangles, one of the main neuropathological hallmarks of AD which is also observed in other tauopathies including PD. Phosphorylated tau also more specifically influences



its affinity for microtubules and high phosphorylation levels of tau have been reported to negatively influence cytoskeleton, synaptic functions as well as cell viability (Buee et al., 2000). Recently, some work has appeared showing some differences in the tau phosphorylation pattern in post-mortem brain of tauopathies, including PD (Duka et al., 2013). This work revealed notable differences in phosphorylation patterns between PD and other tauopathies. For instance, S202, T205, S262, S409 are hyperphosphorylated in AD and DLB, but are unchanged in PD, while others such as T181, S184, S195, S198, S237, S400 are hyperphosphorylated in all three, AD, DLB, and PD. It remains to be confirmed that these varying phosphorylation patterns are indeed related to the varying pathology observed in these different diseases.

## PHOSPHATASES OF TAU

In light of the importance of tau phosphorylation for its pathology, several studies have sought to identify phosphatases dephosphorylating tau (reviewed in Martin et al., 2013a). These studies reveal a predominant role for PP2A, which is the most efficient phosphatase to dephosphorylate tau *in vitro* at S202, S262, and S356, but not S396 (Bennecib et al., 2000; Kuszczak et al., 2009; Martin et al., 2009). This finding has been related to AD where to PP2A activity is reduced by up to 50% in AD brains (Voronkov et al., 2011; Martin et al., 2013a), observations pointing towards the potential of targeting PP2A for therapy in tauopathies, although this relation has yet to be explored in PD. Besides the major role for PP2A, other phosphatases have also been found to act on tau phosphorylation. For instance, PP1 has been reported to act on tau at a limited number of sites in AD brains (T212, T217, S262, S396, S422; Rahman et al., 2005). Similarly, PP2B (aka calcineurin) is able to dephosphorylate tau at S262 and S396 (Rahman et al., 2006) while PP5 is reported to dephosphorylate tau at sites S198–S199–S202, T231–S235, S262–S356, S396–S404, and S422 (Gong et al., 2004). An overview of these four classes of phosphatases and the precise sites reported to be regulated by each is given in **Table 3**.

Thus far, the precise regulatory subunits, which may render specificity of these phosphatases to specific tau phosphosites, have yet to be elucidated and confirmed. Some evidence points to the importance of the PP2A regulatory B $\alpha$  subunit in PP2A mediated phosphoregulation of tau. Expression of the B $\alpha$  subunit of PP2A is decreased in frontal and temporal cortices of AD brain (Sontag et al., 1999, 2004), although it is not known if B $\alpha$  subunit expression is altered in PD brain. In *in vitro* assays, the B $\alpha$  subunit is found to direct the PP2A holoenzyme to microtubules (Xu et al., 2008; Virshup and Shenolikar, 2009), consistent with a role for the B $\alpha$  subunit in the PP2A holoenzyme dephosphorylating tau.

## POTENTIAL INTERPLAY BETWEEN PHOSPHORYLATION OF LRRK2, $\alpha$ -SYN, AND TAU

As LRRK2,  $\alpha$ -syn, and tau are all three involved in dominant Parkinsonism, the hypothesis has been put forward that these three proteins interact in pathological pathways (Taymans and Cookson, 2010), most notably with synuclein and tau acting as toxic proteins and LRRK2 acting as an upstream modulator.

Experimental evidence from animal models has begun to support the hypothesis of an interaction between these proteins in PD pathology. For instance, toxicity induced by high  $\alpha$ -syn levels in mouse brain is attenuated in LRRK2 knockout mice, both in transgenic mice with high  $\alpha$ -syn (inducible CaMKII promoter) expression levels (Lin et al., 2009) as well as after viral delivery of  $\alpha$ -syn (Daher et al., 2014), although this has not been replicated in other transgenic mice using other promoters to drive LRRK2 and  $\alpha$ -syn expression such as the Thy1 promoter (Herzig et al., 2012) or PrP promoter (Daher et al., 2012) suggesting that the effect is dependent on expression patterns or levels. Several reports show that LRRK2 overexpression affects tau expression or tau phosphorylation (see below), supporting the hypothesis that LRRK2 modulates tau although whether LRRK2 is required for tau toxicity has yet to be tested. Interestingly, the injection of  $\alpha$ -syn fibrillar strains into mouse brain is shown to induce tau aggregation (Guo et al., 2013), suggesting an interconnection between  $\alpha$ -syn and tau pathological mechanisms. Studies of relationships between LRRK2,  $\alpha$ -syn, and tau showing pairwise interactions between these proteins suggest that 3-way interaction studies, which are currently still lacking, are warranted. More information on the overall interplay between these three proteins in PD pathology can be found in recent reviews covering this topic (Greggio et al., 2011; Tenreiro et al., 2014). With regards to the topic of the present review, we highlight below the relationships between LRRK2,  $\alpha$ -syn, and tau with regard to their phosphorylation regulation.

As LRRK2 is a kinase, it has been hypothesized that LRRK2 may phosphorylate  $\alpha$ -syn or tau. Direct phosphorylation of  $\alpha$ -syn by LRRK2 *in vitro* has been tested, however, this led to negative results [reference (Khan et al., 2005) and JMT, VB unpublished results]. Also, in LRRK2 overexpressing mice, phosphorylation levels of  $\alpha$ -syn at S129 were found to be unchanged (Herzig et al., 2012) or even reduced compared to controls (Lin et al., 2009), countering the hypothesis that LRRK2 phosphorylates  $\alpha$ -syn. Qing and colleagues reported that crude lysates of *E. coli* expressing LRRK2 as source of enzyme could phosphorylate  $\alpha$ -syn at S129, suggesting that LRRK2 may regulate  $\alpha$ -syn phosphorylation in conjunction with bacterial proteins (Qing et al., 2009), although this is unlikely to be representative of a human physiological situation. The direct phosphorylation of  $\alpha$ -syn by LRRK2 can therefore be excluded.

There is, however, evidence that LRRK2 may be involved in regulating tau phosphorylation. First, tau pathology has been observed in post-mortem brain of LRRK2 mutation carriers, including in carriers of the I1371V (Biernacka et al., 2011), N1437H (Puschmann et al., 2012), R1441C (Zimprich et al., 2004), Y1699C (Khan et al., 2005), G2019S (Lin et al., 2010), and I2020T (Ujiie et al., 2012) LRRK2 mutations. Interestingly, genetic studies suggest that tau variants influence LRRK2 disease, more specifically by influencing the age of disease onset (Gan-Or et al., 2012), although another study found that interactions between LRRK2 and tau were at the limit of statistical significance (Biernacka et al., 2011). In cells and *in vivo*, several pieces of evidence point to LRRK2 in regulating tau phosphorylation. MacLeod et al. (2006) found that overexpressed LRRK2 G2019S or I2020T in primary neurons colocalized with phospho-tau punctae in axons. Overexpression of LRRK2 via

**Table 3 | Overview of different classes of phosphatases reported to regulate tau phosphorylation sites, including PP1, PP2A, PP2B, and PP5.**

Site	PP1	PP2A	PP2B	PP5
S198				Gong et al. (2004)
S199	Liu et al. (2005)	Liu et al. (2005)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Gong et al. (2004), Liu et al. (2005)
S202	Liu et al. (2005)	Liu et al. (2005), Kuszczuk et al. (2009), Martin et al. (2009)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Gong et al. (2004), Liu et al. (2005)
T205	Liu et al. (2005)	Liu et al. (2005), Kuszczuk et al. (2009), Martin et al. (2009)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Liu et al. (2005)
T212	Liu et al. (2005), Rahman et al. (2005)	Liu et al. (2005)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Liu et al. (2005)
S214	Liu et al. (2005)	Liu et al. (2005)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Liu et al. (2005)
T217	Rahman et al. (2005)		Saito et al. (1995), Bennecib et al. (2000), Rahman et al. (2006)	
T231		Saito et al. (1995), Chen et al. (2008)		Gong et al. (2004)
S235	Liu et al. (2005)	Saito et al. (1995), Liu et al. (2005), Chen et al. (2008)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Gong et al. (2004), Liu et al. (2005)
S262	Liu et al. (2005), Rahman et al. (2005)	Saito et al. (1995), Liu et al. (2005), Chen et al. (2008), Martin et al. (2009)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Gong et al. (2004), Liu et al. (2005)
S293				
S324				
S356		Saito et al. (1995), Chen et al. (2008), Martin et al. (2009)		Gong et al. (2004)
S396	Liu et al. (2005), Rahman et al. (2005)	Liu et al. (2005)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Gong et al. (2004), Liu et al. (2005)
S404	Liu et al. (2005)	Saito et al. (1995), Liu et al. (2005), Chen et al. (2008)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Gong et al. (2004), Liu et al. (2005)
S409	Liu et al. (2005)	Liu et al. (2005)	Saito et al. (1995), Bennecib et al. (2000), Liu et al. (2005), Rahman et al. (2006)	Liu et al. (2005)
S422	Rahman et al. (2005)		Saito et al. (1995), Bennecib et al. (2000), Rahman et al. (2006)	Gong et al. (2004)

References of reports showing activity of phosphatases per phosphorylation site mentioned in the first cell of each table row are given in the column of the corresponding phosphatase class. See text for a list of reviews discussing non-phosphatase tau phosphorylation regulators.

the ThyI promoter left tau and phospho-tau (S202/T205) levels unchanged in mouse brains (Herzig et al., 2012), although S202 phosphorylation levels were found to be enhanced in brains of BAC transgenic mice expressing LRRK2 G2019S (Melrose et al., 2010). In *Drosophila*, overexpression of LRRK2 G2019S is reported to affect tau dendritic localization and promote tau phosphorylation at T212 through the recruitment of GSK3 $\beta$  (Lin et al., 2010). Interestingly, LRRK2 has been shown to phosphorylate tubulin-associated tau but not free tau (Kawakami et al., 2012). This last finding may be related to the observations in cells that LRRK2 can in certain conditions translocate to skein-like cytoplasmic pools. Although these skein-like cytoplasmic pools have yet to be fully characterized, at least a portion of these are associated to microtubules (Kett et al., 2012). This points to a mechanism whereby LRRK2 may be recruited to microtubules where it may

regulate tau phosphorylation with other kinase or phosphatase partners.

In relation to our knowledge of phosphatases of LRRK2,  $\alpha$ -syn, and tau and the pathogenic nature of phosphorylations in these proteins, a strategy to target phosphatases in a way that will counteract PD-associated phosphorylation can be proposed. To target LRRK2 disease-related phosphorylations, PP1 would be targeted in order to modulate phosphorylation at its ANK-LRR interdomain phosphorylation sites to a 'healthy' level. For instance, pharmacological inhibition of PP1 was shown to increase the S910/S935/S955/S973 phosphorylation to levels comparable to WT. Also, PP1 inhibition was able to reduce the prevalence of LRRK2 presence in skein-like structures observed with several hypophosphorylated disease mutants (Lobbestael et al., 2013). This effect is consistent with a reduction of microtubule associated



LRRK2 and therefore reduced risk of LRRK2-mediated tau hyperphosphorylation, suggesting that PP1 inhibition may be a viable therapeutic strategy to inhibit LRRK2 mediated pathology. Thus far, the specific composition of the PP1 holoenzyme targeting the ANK-LRR interdomain sites remains to be elucidated prior to developing LRRK2-specific phosphoregulation strategies. Also, the elucidation of phosphatases regulating other sites, such as S1292, may also reveal other potential phosphatase targets for potential therapies targeting LRRK2 phosphorylation.

For  $\alpha$ -syn, targeting PP2A holo-enzymes is a preferred strategy, more specifically through the activation of PP2A to reduce  $\alpha$ -syn-S129 phosphorylation levels. The potential of this approach has been shown by pharmacologically enhancing PP2A activity (see above; Lee et al., 2011). The precise PP2A holoenzyme composition for  $\alpha$ -syn-S129 has begun to be elucidated, with a preference for the B regulatory subunit above B' and B'' subunits. Further elucidation of the preferred C (catalytic) and A (scaffolding) subunits for  $\alpha$ -syn-S129 will allow development of phosphatase activation by targeting the PP2A holoenzyme.

For tau, several phosphatases may be targeted, including PP2A, PP1, PP2B, or PP5. The predominant role of PP2A suggests that activation of PP2A may also be a beneficial therapeutic strategy. One issue is whether separate PP2A holoenzymes are responsible for dephosphorylation of tau at the separate phosphosites. Specifically for PD, it may be expected that dephosphorylation of tau at PD hyperphosphorylation sites should be targeted prioritarily, including T181, S184, S195, S198, S237, S400 (Duka et al., 2013). Further research will be needed to describe the PD-tau phosphorylations and their regulation.

## TARGETING OF PHOSPHATASES

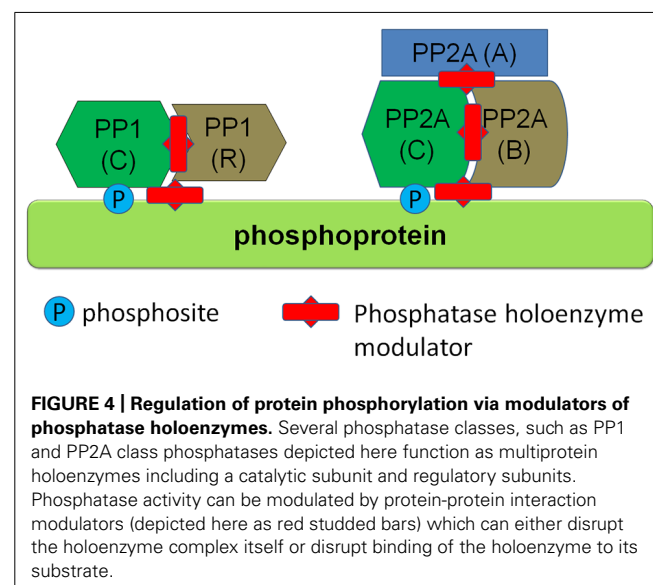
Modulation of phosphorylation levels of disease proteins is an attractive approach to develop disease modifying therapies. For instance, much effort has been spent on targeting kinases responsible for pathogenic phosphorylations. Kinases are also attractive drug targets. Indeed, attrition rates during development of drugs acting on kinases are lower than for most other classes of drugs (Walker and Newell, 2009). An important factor to bear in mind with pharmacological inhibition of kinases is that this will influence the phosphorylation levels of all of the kinase's substrates, both in the disease protein as well as for all of its other substrates. Therefore, for key phosphorylations, it may be necessary to target phosphorylations through other means than by targeting kinases. One possibility is to target phosphatases.

Phosphatases are divided into different classes, including phosphoprotein phosphatases (PPP), metal-dependent protein phosphatases and protein tyrosine phosphatases (PTP). Phosphatases from the PPP class are responsible for the vast majority of dephosphorylations of central nervous system proteins. PPPs, such as PP1 and PP2A class phosphatases, have the characteristic of functioning as a holoenzyme composed of two or more subunits, including a catalytic subunit as well as regulatory subunits. Interestingly, there are few catalytic subunits in the PPP family and the regulatory subunits have therefore generally been thought to confer substrate specificity to the holoenzyme. Emerging structural evidence is supporting this view, namely that phosphatase holoenzymes of the PPP family associate in a structured way

and that this structure determines substrate specificity of the dephosphorylation. Some examples are the interaction between PP1 and spinophilin, a complex found in neurons (Ragusa et al., 2010), and the PP1 $\gamma$ -MYPT (myosin phosphatase; PPP1R12A) complex that acts as a myosin phosphatase in muscle (Terrak et al., 2014; Yamashiro et al., 2008). Based on this knowledge, it has been proposed that by targeting the holoenzyme structure, for instance by molecules which disrupt substrate-phosphatase holoenzyme complex, one can specifically modulate phosphorylation levels of proteins (McConnell and Wadzinski, 2009; Tsaytler and Bertolotti, 2013; see conceptual schematic in **Figure 4**). Based on the knowledge of the phosphatases responsible for dephosphorylation of specific phosphosites of proteins, such as emerging knowledge of the composition of phosphatase holoenzymes dephosphorylating specific sites in  $\alpha$ -syn, LRRK2, and tau, the specific modulation of phosphorylation levels of single phosphosites is theoretically possible. Further work, including further identification and characterization of specific phosphatase holoenzyme compositions for  $\alpha$ -syn, LRRK2, and tau and development of modulators of phosphatase holoenzyme complexes with their specific substrates, is required to test the efficacy of this approach.

## CONCLUSION

The three main proteins linked to Parkinsonism,  $\alpha$ -syn, tau, and LRRK2 all display phosphorylations which are modified in disease. The phosphorylation changes are the result of a balance between activity of kinases and phosphatases. Emerging evidence points to phosphatases regulating different pathological phosphorylations in these three proteins, primarily PP1 for LRRK2 and PP2A for  $\alpha$ -syn and tau. It appears now feasible to target phosphatases of  $\alpha$ -syn, tau, and LRRK2 in order to alleviate pathology mediated by pathological phosphorylation of these disease proteins. In order for this approach to reach its full potential, additional research will be required further linking individual or clustered phosphorylation events in these three proteins to disease. Also,



refined knowledge of the precise holoenzyme compositions for each pathological phosphorylation is necessary to develop highly specific phosphatase holoenzyme modulators.

## REFERENCES

- Anderson, J. P., Walker, D. E., Goldstein, J. M., De Laat, R., Banducci, K., Caccavello, R. J., et al. (2006). Phosphorylation of Ser-129 is the dominant pathological modification of alpha-synuclein in familial and sporadic Lewy body disease. *J. Biol. Chem.* 281, 29739–29752. doi: 10.1074/jbc.M600933200
- Benneicib, M., Gong, C. X., Grundke-Iqbal, I., and Iqbal, K. (2000). Role of protein phosphatase-2A and -1 in the regulation of GSK-3, cdk5, and cdc2 and the phosphorylation of tau in rat forebrain. *FEBS Lett.* 485, 87–93. doi: 10.1016/S0014-5793(00)02203-1
- Biernacka, J. M., Armasu, S. M., Cunningham, J. M., Ahlskog, J. E., Chung, S. J., and Maraganore, D. M. (2011). Do interactions between SNCA, MAPT, and LRRK2 genes contribute to Parkinson's disease susceptibility? *Parkinsonism Relat. Disord.* 17, 730–736. doi: 10.1016/j.parkreldis.2011.07.001
- Bollen, M., Peti, W., Ragusa, M. J., and Beullens, M. (2010). The extended PP1 toolkit: designed to create specificity. *Trends Biochem. Sci.* 35, 450–458. doi: 10.1016/j.tibs.2010.03.002
- Braithwaite, S. P., Voronkov, M., Stock, J. B., and Mouradian, M. M. (2012). Targeting phosphatases as the next generation of disease modifying therapeutics for Parkinson's disease. *Neurochem. Int.* 61, 899–906. doi: 10.1016/j.neuint.2012.01.031
- Buee, L., Bussiere, T., Buee-Scherrer, V., Delacourte, A., and Hof, P. R. (2000). Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. *Brain Res. Brain Res. Rev.* 33, 95–130. doi: 10.1016/S0165-0173(00)00019-9
- Chartier-Harlin, M. C., Kachergus, J., Roumier, C., Mouroux, V., Douay, X., Lincoln, S., et al. (2004). Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* 364, 1167–1169. doi: 10.1016/S0140-6736(04)17103-1
- Chen, L., and Feany, M. B. (2005). Alpha-synuclein phosphorylation controls neurotoxicity and inclusion formation in a *Drosophila* model of Parkinson disease. *Nat. Neurosci.* 8, 657–663. doi: 10.1038/nn1443
- Chen, L., Periquet, M., Wang, X., Negro, A., Mclean, P. J., Hyman, B. T., et al. (2009). Tyrosine and serine phosphorylation of alpha-synuclein have opposing effects on neurotoxicity and soluble oligomer formation. *J. Clin. Invest.* 119, 3257–3265. doi: 10.1172/JCI39088
- Chen, S., Li, B., Grundke-Iqbal, I., and Iqbal, K. (2008). I1PP2A affects tau phosphorylation via association with the catalytic subunit of protein phosphatase 2A. *J. Biol. Chem.* 283, 10513–10521. doi: 10.1074/jbc.M709852200M709852200
- Choi, H. G., Zhang, J., Deng, X., Hatcher, J. M., Patricelli, M. P., Zhao, Z., et al. (2012). Brain penetrant LRRK2 inhibitor. *ACS Med. Chem. Lett.* 3, 658–662. doi: 10.1021/ml300123a
- Daher, J. P., Pletnikova, O., Biskup, S., Musso, A., Gellhaar, S., Galter, D., et al. (2012). Neurodegenerative phenotypes in an A53T alpha-synuclein transgenic mouse model are independent of LRRK2. *Hum. Mol. Genet.* 21, 2420–2431. doi: 10.1093/hmg/dd5057
- Daher, J. P., Volpicelli-Daley, L. A., Blackburn, J. P., Moehle, M. S., and West, A. B. (2014). Abrogation of alpha-synuclein-mediated dopaminergic neurodegeneration in LRRK2-deficient rats. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9289–9294. doi: 10.1073/pnas.14032151111403215111
- Delbroek, L., Van Kolen, K., Steegmans, L., Da Cunha, R., Mandemakers, W., Daneels, G., et al. (2013). Development of an enzyme-linked immunosorbent assay for detection of cellular and in vivo LRRK2 S935 phosphorylation. *J. Pharm. Biomed. Anal.* 76, 49–58. doi: 10.1016/j.jpba.2012.12.002
- Del Ser, T., Steinwachs, K. C., Gertz, H. J., Andres, M. V., Gomez-Carrillo, B., Medina, M., et al. (2013). Treatment of Alzheimer's disease with the GSK-3 inhibitor tideglusib: a pilot study. *J. Alzheimers. Dis.* 33, 205–215. doi: 10.3233/JAD-2012-120805
- Deng, X., Dзамко, N., Prescott, A., Davies, P., Liu, Q., Yang, Q., et al. (2011). Characterization of a selective inhibitor of the Parkinson's disease kinase LRRK2. *Nat. Chem. Biol.* 7, 203–205. doi: 10.1038/nchembio.538
- Doggett, E. A., Zhao, J., Mork, C. N., Hu, D., and Nichols, R. J. (2012). Phosphorylation of LRRK2 serines 955 and 973 is disrupted by Parkinson's disease mutations and LRRK2 pharmacological inhibition. *J. Neurochem.* 120, 37–45. doi: 10.1111/j.1471-4159.2011.07537.x
- Duka, V., Lee, J. H., Credle, J., Wills, J., Oaks, A., Smolinsky, C., et al. (2013). Identification of the sites of tau hyperphosphorylation and activation of tau kinases in synucleinopathies and Alzheimer's diseases. *PLoS ONE* 8:e75025. doi: 10.1371/journal.pone.0075025
- Dзамко, N., Chua, G., Ranola, M., Rowe, D. B., and Halliday, G. M. (2013). Measurement of LRRK2 and Ser910/935 phosphorylated LRRK2 in peripheral blood mononuclear cells from idiopathic Parkinson's disease patients. *J. Parkinsons Dis.* 3, 145–152. doi: 10.3233/JPD-130174
- Dзамко, N., Deak, M., Hentati, F., Reith, A. D., Prescott, A. R., Alessi, D. R., et al. (2010). Inhibition of LRRK2 kinase activity leads to dephosphorylation of Ser(910)/Ser(935), disruption of 14-3-3 binding and altered cytoplasmic localization. *Biochem. J.* 430, 405–413. doi: 10.1042/BJ20100784
- Dзамко, N., Inesta-Vaquera, F., Zhang, J., Xie, C., Cai, H., Arthur, S., et al. (2012). The IkkappaB kinase family phosphorylates the Parkinson's disease kinase LRRK2 at Ser935 and Ser910 during Toll-like receptor signaling. *PLoS ONE* 7:e39132. doi: 10.1371/journal.pone.0039132
- Fujiwara, H., Hasegawa, M., Dohmae, N., Kawashima, A., Masliah, E., Goldberg, M. S., et al. (2002). alpha-Synuclein is phosphorylated in synucleinopathy lesions. *Nat. Cell Biol.* 4, 160–164. doi: 10.1038/ncb748
- Gan-Or, Z., Bar-Shira, A., Mirelman, A., Gurevich, T., Giladi, N., and Orr-Urtreger, A. (2012). The age at motor symptoms onset in LRRK2-associated Parkinson's disease is affected by a variation in the MAPT locus: a possible interaction. *J. Mol. Neurosci.* 46, 541–544. doi: 10.1007/s12031-011-9641-0
- Gasser, T. (2009). Molecular pathogenesis of Parkinson disease: insights from genetic studies. *Expert Rev. Mol. Med.* 11, e22. doi: 10.1017/S1462399409001148
- Gloeckner, C. J., Boldt, K., Von Zweydford, F., Helm, S., Wiesent, L., Sarioglu, H., et al. (2010). Phosphopeptide analysis reveals two discrete clusters of phosphorylation in the N-terminus and the Roc domain of the Parkinson-disease associated protein kinase LRRK2. *J. Proteome Res.* 9, 1738–1745. doi: 10.1021/pr9008578
- Gong, C. X., Liu, F., Wu, G., Rossie, S., Wegiel, J., Li, L., et al. (2004). Dephosphorylation of microtubule-associated protein tau by protein phosphatase 5. *J. Neurochem.* 88, 298–310. doi: 10.1111/j.1471-4159.2004.02147.x
- Greggio, E., Bisaglia, M., Civiero, L., and Bubacco, L. (2011). Leucine-rich repeat kinase 2 and alpha-synuclein: intersecting pathways in the pathogenesis of Parkinson's disease? *Mol. Neurodegener.* 6:6. doi: 10.1186/1750-1326-6-6
- Greggio, E., Taymans, J. M., Zhen, E. Y., Ryder, J., Vancraenenbroeck, R., Beilina, A., et al. (2009). The Parkinson's disease kinase LRRK2 autophosphorylates its GTPase domain at multiple sites. *Biochem. Biophys. Res. Commun.* 389, 449–454. doi: 10.1016/j.bbrc.2009.08.163
- Guo, J. L., Covell, D. J., Daniels, J. P., Iba, M., Stieber, A., Zhang, B., et al. (2013). Distinct alpha-synuclein strains differentially promote tau inclusions in neurons. *Cell* 154, 103–117. doi: 10.1016/j.cell.2013.05.057
- Hasegawa, M., Fujiwara, H., Nonaka, T., Wakabayashi, K., Takahashi, H., Lee, V. M.-Y., et al. (2002). Phosphorylated alpha-synuclein is ubiquitinated in alpha-synucleinopathy lesions. *J. Biol. Chem.* 277, 49071–49076. doi: 10.1074/jbc.M208046200
- Healy, D. G., Falchi, M., O'Sullivan, S. S., Bonifati, V., Durr, A., Bressman, S., et al. (2008). Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol.* 7, 583–590. doi: 10.1016/S1474-4422(08)70117-0
- Henderson-Smith, A., Chow, D., Meechoovet, B., Aziz, M., Jacobson, S. A., Shill, H. A., et al. (2013). SMG1 identified as a regulator of Parkinson's disease-associated alpha-synuclein through siRNA screening. *PLoS ONE* 8:e77711. doi: 10.1371/journal.pone.0077711
- Herzig, M. C., Bidinosti, M., Schweizer, T., Hafner, T., Stemmelen, C., Weiss, A., et al. (2012). High LRRK2 levels fail to induce or exacerbate neuronal alpha-synucleinopathy in mouse brain. *PLoS ONE* 7:e36581. doi: 10.1371/journal.pone.0036581
- Houlden, H., and Singleton, A. B. (2012). The genetics and neuropathology of Parkinson's disease. *Acta Neuropathol.* 124, 325–338. doi: 10.1007/s00401-012-1013-5
- Kahle, P. J., Neumann, M., Ozmen, L., and Haass, C. (2000). Physiology and pathophysiology of alpha-synuclein. Cell culture and transgenic animal models based on a Parkinson's disease-associated protein. *Ann. N. Y. Acad. Sci.* 920, 33–41. doi: 10.1111/j.1749-6632.2000.tb06902.x
- Kamikawaji, S., Ito, G., and Iwatsubo, T. (2009). Identification of the autophosphorylation sites of LRRK2. *Biochemistry* 48, 10963–10975. doi: 10.1021/bi9011379
- Kawakami, E., Yabata, T., Ohta, E., Maekawa, T., Shimada, N., Suzuki, M., et al. (2012). LRRK2 phosphorylates tubulin-associated tau but not the free

- molecule: LRRK2-mediated regulation of the tau-tubulin association and neurite outgrowth. *PLoS ONE* 7:e30834. doi: 10.1371/journal.pone.0030834
- Kett, L. R., Boassa, D., Ho, C. C., Rideout, H. J., Hu, J., Terada, M., et al. (2012). LRRK2 Parkinson disease mutations enhance its microtubule association. *Hum. Mol. Genet.* 21, 890–899. doi: 10.1093/hmg/ddr526
- Khan, N. L., Jain, S., Lynch, J. M., Pavese, N., Abou-Sleiman, P., Holton, J. L., et al. (2005). Mutations in the gene LRRK2 encoding dardarin (PARK8) cause familial Parkinson's disease: clinical, pathological, olfactory and functional imaging and genetic data. *Brain* 128, 2786–2796. doi: 10.1093/brain/awh667
- Kramer, T., Schmidt, B., and Lo Monte, F. (2012). Small-molecule inhibitors of GSK-3: structural insights and their application to Alzheimer's disease models. *Int. J. Alzheimers Dis.* 2012, 381029. doi: 10.1155/2012/381029
- Kuszczyk, M., Gordon-Krajcer, W., and Lazarewicz, J. W. (2009). Homocysteine-induced acute excitotoxicity in cerebellar granule cells in vitro is accompanied by PP2A-mediated dephosphorylation of tau. *Neurochem. Int.* 55, 174–180. doi: 10.1016/j.neuint.2009.02.010
- Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452–1458. doi: 10.1038/ng.2802ng.2802
- Lee, K. W., Chen, W., Junn, E., Im, J. Y., Grosso, H., Sonsalla, P. K., et al. (2011). Enhanced phosphatase activity attenuates  $\alpha$ -synucleinopathy in a mouse model. *J. Neurosci.* 31, 6963–6971. doi: 10.1523/JNEUROSCI.6513-10.2011
- Li, X., Wang, Q. J., Pan, N., Lee, S., Zhao, Y., Chait, B. T., et al. (2011). Phosphorylation-dependent 14-3-3 binding to LRRK2 is impaired by common mutations of familial Parkinson's disease. *PLoS ONE* 6:e17153. doi: 10.1371/journal.pone.0017153
- Lin, C. H., Tsai, P. I., Wu, R. M., and Chien, C. T. (2010). LRRK2 G2019S mutation induces dendrite degeneration through mislocalization and phosphorylation of tau by recruiting autoactivated GSK3 $\alpha$ . *J. Neurosci.* 30, 13138–13149. doi: 10.1523/JNEUROSCI.1737-10.2010
- Lin, X., Parisiadou, L., Gu, X. L., Wang, L., Shim, H., Sun, L., et al. (2009). Leucine-rich repeat kinase 2 regulates the progression of neuropathology induced by Parkinson's-disease-related mutant  $\alpha$ -synuclein. *Neuron* 64, 807–827. doi: 10.1016/j.neuron.2009.11.006
- Liu, F., Grundke-Iqbal, I., Iqbal, K., and Gong, C. X. (2005). Contributions of protein phosphatases PP1, PP2A, PP2B and PP5 to the regulation of tau phosphorylation. *Eur. J. Neurosci.* 22, 1942–1950. doi: 10.1111/j.1460-9568.2005.04391.x
- Lobbestael, E., Baekelandt, V., and Taymans, J. M. (2012). Phosphorylation of LRRK2: from kinase to substrate. *Biochem. Soc. Trans.* 40, 1102–1110. doi: 10.1042/BST20120128
- Lobbestael, E., Zhao, J., Rudenko, I. N., Beylina, A., Gao, F., Wetter, J., et al. (2013). Identification of protein phosphatase 1 as a regulator of the LRRK2 phosphorylation cycle. *Biochem. J.* 456, 119–128. doi: 10.1042/BJ20121772BJ20121772
- Lou, H., Montoya, S. E., Alerte, T. N., Wang, J., Wu, J., Peng, X., et al. (2010). Serine 129 phosphorylation reduces the ability of  $\alpha$ -synuclein to regulate tyrosine hydroxylase and protein phosphatase 2A in vitro and in vivo. *J. Biol. Chem.* 285, 17648–17661. doi: 10.1074/jbc.M110.100867M110.100867
- MacLeod, D., Dowman, J., Hammond, R., Leete, T., Inoue, K., and Abeliovich, A. (2006). The familial Parkinsonism gene LRRK2 regulates neurite process morphology. *Neuron* 52, 587–593. doi: 10.1016/j.neuron.2006.10.008
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298, 1912–1934. doi: 10.1126/science.1075762298/5600/1912
- Martin, L., Latypova, X., and Terro, F. (2011). Post-translational modifications of tau protein: implications for Alzheimer's disease. *Neurochem. Int.* 58, 458–471. doi: 10.1016/j.neuint.2010.12.023
- Martin, L., Latypova, X., Wilson, C. M., Magnaudeix, A., Perrin, M. L., and Terro, F. (2013a). Tau protein phosphatases in Alzheimer's disease: the leading role of PP2A. *Ageing Res. Rev.* 12, 39–49. doi: 10.1016/j.arr.2012.06.008
- Martin, L., Latypova, X., Wilson, C. M., Magnaudeix, A., Perrin, M. L., Yardin, C., et al. (2013b). Tau protein kinases: involvement in Alzheimer's disease. *Ageing Res. Rev.* 12, 289–309. doi: 10.1016/j.arr.2012.06.003
- Martin, L., Magnaudeix, A., Esclaire, F., Yardin, C., and Terro, F. (2009). Inhibition of glycogen synthase kinase-3 $\beta$  downregulates total tau proteins in cultured neurons and its reversal by the blockade of protein phosphatase-2A. *Brain Res.* 1252, 66–75. doi: 10.1016/j.brainres.2008.11.057
- Mbeof, M. K., Paleologou, K. E., Boucharaba, A., Oueslati, A., Schell, H., Fournier, M., et al. (2010). Phosphorylation of synucleins by members of the Polo-like kinase family. *J. Biol. Chem.* 285, 2807–2822. doi: 10.1074/jbc.M109.081950
- McConnell, J. L., and Wadzinski, B. E. (2009). Targeting protein serine/threonine phosphatases for drug development. *Mol. Pharmacol.* 75, 1249–1261. doi: 10.1124/mol.108.053140
- Melrose, H. L., Dachselt, J. C., Behrouz, B., Lincoln, S. J., Yue, M., Hinkle, K. M., et al. (2010). Impaired dopaminergic neurotransmission and microtubule-associated protein tau alterations in human LRRK2 transgenic mice. *Neurobiol. Dis.* 40, 503–517. doi: 10.1016/j.nbd.2010.07.010
- Moorhead, G. B., Trinkle-Mulcahy, L., and Ulke-Lemee, A. (2007). Emerging roles of nuclear protein phosphatases. *Nat. Rev. Mol. Cell Biol.* 8, 234–244. doi: 10.1038/nrm2126
- Muda, K., Bertinetti, D., Gesellchen, F., Hermann, J. S., Von Zweyendorf, F., Geerlof, A., et al. (2014). Parkinson-related LRRK2 mutation R1441C/G/H impairs PKA phosphorylation of LRRK2 and disrupts its interaction with 14-3-3. *Proc. Natl. Acad. Sci. U.S.A.* 111, E34–E43. doi: 10.1073/pnas.13127011111312701111
- Nalls, M. A., Plagnol, V., Hernandez, D. G., Sharma, M., Sheerin, U. M., Saad, M., et al. (2011). Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* 377, 641–649. doi: 10.1016/S0140-6736(10)62345-8
- Neumann, M., Kahle, P. J., Giasson, B. I., Ozmen, L., Borroni, E., Spooen, W., et al. (2002). Misfolded proteinase K-resistant hyperphosphorylated  $\alpha$ -synuclein in aged transgenic mice with locomotor deterioration and in human  $\alpha$ -synucleinopathies. *J. Clin. Invest.* 110, 1429–1439. doi: 10.1172/JCI200215777
- Nichols, R. J., Dzakmo, N., Morrice, N. A., Campbell, D. G., Deak, M., Ordureau, A., et al. (2010). 14-3-3 binding to LRRK2 is disrupted by multiple Parkinson's disease-associated mutations and regulates cytoplasmic localization. *Biochem. J.* 430, 393–404. doi: 10.1042/BJ20100483
- Oueslati, A., Schneider, B. L., Aebischer, P., and Lashuel, H. A. (2013). Polo-like kinase 2 regulates selective autophagic  $\alpha$ -synuclein clearance and suppresses its toxicity in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3945–E3954. doi: 10.1073/pnas.13099911101309991110
- Paleologou, K. E., Oueslati, A., Shakked, G., Rospigliosi, C. C., Kim, H. Y., Lamberto, G. R., et al. (2010). Phosphorylation at S87 is enhanced in synucleinopathies, inhibits  $\alpha$ -synuclein oligomerization, and influences synuclein-membrane interactions. *J. Neurosci.* 30, 3184–3198. doi: 10.1523/JNEUROSCI.5922-09.2010
- Peng, X., Tehranian, R., Dietrich, P., Stefanis, L., and Perez, R. G. (2005).  $\alpha$ -Synuclein activation of protein phosphatase 2A reduces tyrosine hydroxylase phosphorylation in dopaminergic cells. *J. Cell Sci.* 118, 3523–3530. doi: 10.1242/jcs.02481
- Perez-Revuelta, B. I., Hettich, M. M., Ciociaro, A., Rotermund, C., Kahle, P. J., Krauss, S., et al. (2014). Metformin lowers Ser-129 phosphorylated  $\alpha$ -synuclein levels via mTOR-dependent protein phosphatase 2A activation. *Cell Death Dis.* 5, e1209. doi: 10.1038/cddis.2014.175cddis2014175
- Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A., et al. (1997). Mutation in the  $\alpha$ -synuclein gene identified in families with Parkinson's disease. *Science* 276, 2045–2047. doi: 10.1126/science.276.5321.2045
- Pungaliya, P. P., Bai, Y., Lipinski, K., Anand, V. S., Sen, S., Brown, E. L., et al. (2010). Identification and characterization of a leucine-rich repeat kinase 2 (LRRK2) consensus phosphorylation motif. *PLoS ONE* 5:e13672. doi: 10.1371/journal.pone.0013672
- Puschmann, A., Englund, E., Ross, O. A., Vilarino-Guell, C., Lincoln, S. J., Kachergus, J. M., et al. (2012). First neuropathological description of a patient with Parkinson's disease and LRRK2 p.N1437H mutation. *Parkinsonism Relat. Disord.* 18, 332–338. doi: 10.1016/j.parkreldis.2011.11.019
- Qing, H., Wong, W., McGeer, E. G., and McGeer, P. L. (2009). Lrrk2 phosphorylates  $\alpha$  synuclein at serine 129: Parkinson disease implications. *Biochem. Biophys. Res. Commun.* 387, 149–152. doi: 10.1016/j.bbrc.2009.06.142
- Ragusa, M. J., Dancheck, B., Critton, D. A., Nairn, A. C., Page, R., and Peti, W. (2010). Spinophilin directs protein phosphatase 1 specificity by blocking substrate binding sites. *Nat. Struct. Mol. Biol.* 17, 459–464. doi: 10.1038/nsmb.1786
- Rahman, A., Grundke-Iqbal, I., and Iqbal, K. (2005). Phosphothreonine-212 of Alzheimer abnormally hyperphosphorylated tau is a preferred substrate of protein phosphatase-1. *Neurochem. Res.* 30, 277–287. doi: 10.1007/s11064-005-2483-9

- Rahman, A., Grundke-Iqbal, I., and Iqbal, K. (2006). PP2B isolated from human brain preferentially dephosphorylates Ser-262 and Ser-396 of the Alzheimer disease abnormally hyperphosphorylated tau. *J. Neural. Transm.* 113, 219–230. doi: 10.1007/s00702-005-0313-5
- Reyniers, L., Del Giudice, M.-G., Civiero, L., Belluzzi, E., Lobbstaël, E., Beilina, A., et al. (2014). Differential protein-protein interactions of LRRK1 and LRRK2 indicate roles in distinct cellular signaling pathways. *J. Neurochem.* 131, 239–250. doi: 10.1111/jnc.12798
- Reynolds, A., Doggett, E. A., Riddle, S. M., Lebakken, C. S., and Nichols, R. J. (2014). LRRK2 kinase activity and biology are not uniformly predicted by its autophosphorylation and cellular phosphorylation site atatus. *Front. Mol. Neurosci.* 7:54. doi: 10.3389/fnmol.2014.00054
- Saito, T., Ishiguro, K., Uchida, T., Miyamoto, E., Kishimoto, T., and Hisanaga, S. (1995). In situ dephosphorylation of tau by protein phosphatase 2A and 2B in fetal rat primary cultured neurons. *FEBS Lett.* 376, 238–242. doi: 10.1016/0014-5793(95)01292-0
- Satake, W., Nakabayashi, Y., Mizuta, I., Hirota, Y., Ito, C., Kubo, M., et al. (2009). Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.* 41, 1303–1307. doi: 10.1038/ng.485
- Sharma, M., Ioannidis, J. P., Aasly, J. O., Annesi, G., Brice, A., Van Broeckhoven, C., et al. (2012). Large-scale replication and heterogeneity in Parkinson disease genetic loci. *Neurology* 79, 659–667. doi: 10.1212/WNL.0b013e318264e353
- Sheng, Z., Zhang, S., Bustos, D., Kleinheinz, T., Le Pichon, C. E., Dominguez, S. L., et al. (2012). Ser1292 autophosphorylation is an indicator of LRRK2 kinase activity and contributes to the cellular effects of PD mutations. *Sci. Transl. Med.* 4, 164ra161. doi: 10.1126/scitranslmed.3004485/164/164ra161
- Simon-Sanchez, J., Schulte, C., Bras, J. M., Sharma, M., Gibbs, J. R., Berg, D., et al. (2009). Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* 41, 1308–1312. doi: 10.1038/ng.487
- Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., et al. (2003). alpha-Synuclein locus triplication causes Parkinson's disease. *Science* 302, 841. doi: 10.1126/science.1090278/302/5646/841
- Sontag, E., Luangpirom, A., Hladik, C., Mudrak, I., Ogris, E., Speciale, S., et al. (2004). Altered expression levels of the protein phosphatase 2A A $\alpha$  enzyme are associated with Alzheimer disease pathology. *J. Neuropathol. Exp. Neurol.* 63, 287–301.
- Sontag, E., Nunbhakdi-Craig, V., Lee, G., Brandt, R., Kamibayashi, C., Kuret, J., et al. (1999). Molecular interactions among protein phosphatase 2A, tau, and microtubules. Implications for the regulation of tau phosphorylation and the development of tauopathies. *J. Biol. Chem.* 274, 25490–25498. doi: 10.1074/jbc.274.36.25490
- Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997). Alpha-synuclein in Lewy bodies. *Nature* 388, 839–840. doi: 10.1038/42166
- Takahashi, M., Kanuka, H., Fujiwara, H., Koyama, A., Hasegawa, M., Miura, M., et al. (2003). Phosphorylation of alpha-synuclein characteristic of synucleinopathy lesions is recapitulated in alpha-synuclein transgenic *Drosophila*. *Neurosci. Lett.* 336, 155–158. doi: 10.1016/S0304-3940(02)01258-2
- Taymans, J. M., and Cookson, M. (2010). Mechanisms of dominant parkinsonism; the toxic triangle of LRRK2, alpha-synuclein and tau. *Bioessays* 32, 227–235. doi: 10.1002/bies.200900163
- Tell, V., and Hilgeroth, A. (2013). Recent developments of protein kinase inhibitors as potential AD therapeutics. *Front. Cell. Neurosci.* 7, 189. doi: 10.3389/fncel.2013.00189
- Tenreiro, S., Eckermann, K., and Outeiro, T. F. (2014). Protein phosphorylation in neurodegeneration: friend or foe? *Front. Mol. Neurosci.* 7, 42. doi: 10.3389/fnmol.2014.00042
- Terrak, M., Kerff, F., Langsetmo, K., Tao, T., and Dominguez, R. (2004). Structural basis of protein phosphatase 1 regulation. *Nature* 429, 780–784. doi: 10.1038/nature02582
- Tsayler, P., and Bertolotti, A. (2013). Exploiting the selectivity of protein phosphatase 1 for pharmacological intervention. *FEBS J.* 280, 766–770. doi: 10.1111/j.1742-4658.2012.08535.x
- Ujiie, S., Hatano, T., Kubo, S., Imai, S., Sato, S., Uchiyama, T., et al. (2012). LRRK2 I2020T mutation is associated with tau pathology. *Parkinsonism Relat. Disord.* 18, 819–823. doi: 10.1016/j.parkreldis.2012.03.024
- Vancraenenbroeck, R., Lobbstaël, E., De Maeyer, M., Baekelandt, V., and Taymans, J. M. (2011). Kinases as targets for Parkinson's disease; from genetics to therapy. *CNS Neurol. Disord. Drug Targets* 10, 724–740. doi: 10.2174/187152711797247858
- Vancraenenbroeck, R., Lobbstaël, E., Weeks, S. D., Strelkov, S. V., Baekelandt, V., Taymans, J. M., et al. (2012). Expression, purification and preliminary biochemical and structural characterization of the leucine rich repeat namesake domain of leucine rich repeat kinase 2. *Biochim. Biophys. Acta* 1824, 450–460. doi: 10.1016/j.bbapap.2011.12.009
- Virshup, D. M., and Shenolikar, S. (2009). From promiscuity to precision: protein phosphatases get a makeover. *Mol. Cell.* 33, 537–545. doi: 10.1016/j.molcel.2009.02.015
- Voronkov, M., Braithwaite, S. P., and Stock, J. B. (2011). Phosphoprotein phosphatase 2A: a novel druggable target for Alzheimer's disease. *Future Med. Chem.* 3, 821–833. doi: 10.4155/fmc.11.47
- Walker, I., and Newell, H. (2009). Do molecularly targeted agents in oncology have reduced attrition rates? *Nat. Rev. Drug Discov.* 8, 15–16. doi: 10.1038/nrd2758
- Waxman, E. A., and Giasson, B. I. (2008). Specificity and regulation of casein kinase-mediated phosphorylation of alpha-synuclein. *J. Neuropathol. Exp. Neurol.* 67, 402–416. doi: 10.1097/NEN.0b013e31816fc995
- Waxman, E. A., and Giasson, B. I. (2011). Characterization of kinases involved in the phosphorylation of aggregated alpha-synuclein. *J. Neurosci. Res.* 89, 231–247. doi: 10.1002/jnr.22537
- Webber, P. J., Smith, A. D., Sen, S., Renfrow, M. B., Mobley, J. A., and West, A. B. (2011). Autophosphorylation in the leucine-rich repeat kinase 2 (LRRK2) GTPase domain modifies kinase and GTP-binding activities. *J. Mol. Biol.* 412, 94–110. doi: 10.1016/j.jmb.2011.07.033
- West, A. B., Moore, D. J., Choi, C., Andrabi, S. A., Li, X., Dikeman, D., et al. (2007). Parkinson's disease-associated mutations in LRRK2 link enhanced GTP-binding and kinase activities to neuronal toxicity. *Hum. Mol. Genet.* 16, 223–232. doi: 10.1093/hmg/ddl471
- Xu, Y., Chen, Y., Zhang, P., Jeffrey, P. D., and Shi, Y. (2008). Structure of a protein phosphatase 2A holoenzyme: insights into B55-mediated Tau dephosphorylation. *Mol. Cell.* 31, 873–885. doi: 10.1016/j.molcel.2008.08.006
- Yamashiro, S., Yamakita, Y., Totsukawa, G., Goto, H., Kaibuchi, K., Ito, M., et al. (2008). Myosin phosphatase-targeting subunit 1 regulates mitosis by antagonizing polo-like kinase 1. *Dev. Cell* 14, 787–797. doi: 10.1016/j.devcel.2008.02.013
- Zhang, J., Deng, X., Choi, H. G., Alessi, D. R., and Gray, N. S. (2012). Characterization of TAE684 as a potent LRRK2 kinase inhibitor. *Bioorg. Med. Chem. Lett.* 22, 1864–1869. doi: 10.1016/j.bmcl.2012.01.084
- Zimprich, A., Biskup, S., Leitner, P., Lichtner, P., Farrer, M., Lincoln, S., et al. (2004). Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* 44, 601–607. doi: 10.1016/j.neuron.2004.11.005

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 June 2014; accepted: 17 October 2014; published online: 07 November 2014.

Citation: Taymans J-M and Baekelandt V (2014) Phosphatases of  $\alpha$ -synuclein, LRRK2, and tau: important players in the phosphorylation-dependent pathology of Parkinsonism. *Front. Genet.* 5:382. doi: 10.3389/fgene.2014.00382

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Taymans and Baekelandt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Identification of developmentally-specific kinotypes and mechanisms of Varroa mite resistance through whole-organism, kinome analysis of honeybee

Albert J. Robertson<sup>1</sup>, Brett Trost<sup>2</sup>, Erin Scruten<sup>3</sup>, Thomas Robertson<sup>1</sup>, Mohammad Mostajeran<sup>1</sup>, Wayne Connor<sup>3</sup>, Anthony Kusalik<sup>2</sup>, Philip Griebel<sup>3,4</sup> and Scott Napper<sup>3,5\*</sup>

<sup>1</sup> Meadow Ridge Enterprises Ltd., Saskatoon, SK, Canada

<sup>2</sup> Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada

<sup>3</sup> Vaccine and Infectious Disease Organization, University of Saskatchewan, Saskatoon, SK, Canada

<sup>4</sup> School of Public Health, University of Saskatchewan, Saskatoon, SK, Canada

<sup>5</sup> Department of Biochemistry, University of Saskatchewan, Saskatoon, SK, Canada

## Edited by:

Andreas Zanzoni, Inserm TAGC,  
UMR1090, France  
Allegra Via, Sapienza University, Italy

## Reviewed by:

Eduardo S. Zeron, Centro de  
Investigación y de Estudios  
Avanzados del IPN Department of  
Mathematics, Mexico  
David Georges Biron, Centre  
National de la Recherche  
Scientifique, France

## \*Correspondence:

Scott Napper, Vaccine and Infectious  
Disease Organization, University of  
Saskatchewan, 120 Veterinary Road,  
Saskatoon, SK, S7N 5E3, Canada  
e-mail: scott.napper@usask.ca

Recent investigations associate *Varroa destructor* (Mesostigmata: Varroidae) parasitism and its associated pathogens and agricultural pesticides with negative effects on colony health, resulting in sporadic global declines in domestic honeybee (*Apis mellifera*) populations. These events have motivated efforts to develop research tools that can offer insight into the causes of declining bee health as well as identify biomarkers to guide breeding programs. Here we report the development of a bee-specific peptide array for characterizing global cellular kinase activity in whole bee extracts. The arrays reveal distinct, developmentally-specific signaling profiles between bees with differential susceptibility to infestation by Varroa mites. Gene ontology analysis of the differentially phosphorylated peptides indicates that the differential susceptibility to Varroa mite infestation does not reflect compromised immunity; rather, there is evidence for mite-mediated immune suppression within the susceptible phenotype that may reduce the ability of these bees to counter secondary viral infections. This hypothesis is supported by the demonstration of more diverse viral infections in mite-infested, susceptible adult bees. The bee-specific peptide arrays are an effective tool for understanding the molecular basis of this complex phenotype as well as for the discovery and utilization of phosphorylation biomarkers for breeding programs.

**Keywords:** peptide arrays, kinome, kinotype, *Varroa destructor*, honeybee, *Apis mellifera*

## 1. INTRODUCTION

In recent years, there has been an alarming worldwide decline in populations of honeybees (*Apis mellifera*) (Dietemann et al., 2013). This is of considerable concern, as approximately one-third of the human food supply depends on pollination by the honeybee (Greenleaf and Kremen, 2006; Cox-Foster et al., 2007; Vanengelsdorp et al., 2009). A number of possible causes have been suggested, including Varroa mite parasitism and associated pathogens (Martin et al., 2012; Nazzi et al., 2012), increased use of pesticides, lack of genetic diversity, and other factors (Vanengelsdorp et al., 2009; Mullin et al., 2010).

The ectoparasitic mite *Varroa destructor*, and RNA viruses that are associated with it, are a significant challenge to the honeybee. Deformed wing virus (DWV) (Martin et al., 2012, 2013), Israeli acute paralysis virus (IAPV), acute bee paralysis virus (ABPV), and Kashmir bee virus (KBV) are the major viruses vectored by Varroa (Di Prisco et al., 2011). Varroa mites continue to spread throughout the world and contribute to the decline of domesticated honeybee populations (Martin et al., 2012; Nazzi et al., 2012). Their natural host, the Asian honeybee (*Apis ceranae*), has developed protective mechanisms based on behavioral characteristics, such as grooming and hygienic traits, as well as

differences in brood development time, rather than differences in immunity (Sammataro et al., 2000; Rosenkranz et al., 2010). The western honeybee, initially exposed to Varroa mite parasitism in the mid-1960s (Sammataro et al., 2000), has yet to develop adequate resistance mechanisms. Many synthetic miticides have been deployed to combat Varroa infestations, but the mites quickly develop resistance; further, the miticides have detrimental effects on honeybee health, and can also leave dangerous residues in the wax (Lodesani and Costa, 2005).

A more attractive approach is to breed honeybees capable of resisting or controlling Varroa mite infestation. However, breeding for Varroa resistance is complicated by a lack of understanding of honeybee susceptibility to mite parasitism, a dearth of biomarkers to identify potentially resistant progeny, and the instability of resistant phenotypes. A number of groups have used natural selection to identify colony phenotypes with Varroa resistance (Le Conte et al., 2007; Seeley, 2007). The most well-characterized genetic stocks able to suppress Varroa population growth are the Varroa sensitive hygiene (VSH) lines (Harbo and Harris, 2009; Tsuruda et al., 2012). In this work, the Saskatraz natural selection project (<http://www.saskatraz.com>) selected and characterized susceptible and resistant honeybee

colony phenotypes for molecular analyses. This project focuses on recurrent natural selection of survivor colonies for honey production, wintering ability, resistance to Varroa, and overall colony health, in the absence of synthetic miticides.

There is a general consensus that understanding the cellular mechanisms of these disease-resistance phenotypes requires a global perspective on bee biology. To this end, a number of recent studies have examined the differential expression of genes (Le Conte et al., 2011) and proteins (Parker et al., 2012) in honeybees that suppress Varroa population growth. These efforts have neither provided clear insight into the cellular mechanisms of Varroa mite susceptibility nor identified reliable biomarkers. This reflects the challenges associated with deciphering complex biology, in particular within the context of a mixed genetic population.

Similar challenges have been overcome in other livestock species through the development and application of species-specific peptide arrays for analysis of global cellular kinase (kinome) activity (Arsenault et al., 2012, 2013b; Trost et al., 2013a). Kinase-mediated protein phosphorylation is critical for the regulation of cellular responses and phenotypes. Analysis of global kinome activity has provided a powerful tool to understand complex biology as well as to identify therapeutic targets and biomarkers (Eglen and Reisine, 2011). In particular, the ability to use short peptides as surrogate substrates for kinases makes it possible to monitor the kinome using high-throughput peptide arrays (Arsenault et al., 2011). While detailed descriptions of the phosphoproteome are available for only a limited number of species, it is possible to predict the sequence contexts of phosphorylation events based on genomic information, creating the opportunity to develop species-specific kinome microarrays for species whose phosphoproteomes have not been extensively characterized (Jalal et al., 2009; Trost et al., 2013a). Kinome analysis has been demonstrated to have considerable utility in understanding cellular mechanisms of host-pathogen interaction (Kindrachuk et al., 2011; Arsenault et al., 2012, 2013a; Määttä et al., 2013; Mulongo et al., 2014) as well as identifying phosphorylation biomarkers that predict or reflect phenotypic traits (Arsenault et al., 2013b). Recently, the existence of temporally-stable species and individual-specific phosphorylation profiles, or kinotypes, was reported (Trost et al., 2013c). These stable patterns within individuals likely reflect genetic, epigenetic, environmental and developmental influences and may provide mechanistic and predictive insight into complex, multifactorial phenotypes. Similarly, while kinome analysis is traditionally performed on samples of low biological complexity, such as cultured cells or purified cell populations, recent applications have extended this analysis to more complex samples, including intestinal tissue (Määttä et al., 2013) and muscle biopsies (Arsenault et al., 2013b).

Here we report the development of a bee-specific kinome array and its application to characterize honeybees with a quantified, differential susceptibility to Varroa mite infestation. Bees of the susceptible and resistant phenotypes possess distinct kinome profiles at a number of developmental stages ranging from pupae to adult, highlighting the potential to use these differences as markers for breeding programs. Kinome analysis also offers insight into

the mechanisms underlying disease susceptibility. Specifically, the kinome data indicate that the susceptibility to Varroa mite infestation does not reflect compromised immunity. There is, however, evidence for mite-mediated immune suppression within the susceptible phenotype, which may reduce the ability of these bees to counter secondary infections. Consistent with this hypothesis, an increased diversity of viral infections is observed in Varroa-infested susceptible bees. Overall, the bee-specific peptide arrays offer an effective tool for understanding the molecular basis of complex phenotypes and for analyzing specific biological responses, and may facilitate the identification of phosphorylation biomarkers for breeding programs.

## 2. MATERIALS AND METHODS

### 2.1. COLONY PHENOTYPE SELECTION

A detailed description of the honeybee breeding and selection program that was used to construct and identify the Varroa mite susceptible and resistant phenotypes can be accessed at <http://www.saskatraz.com>. Briefly, Meadow Ridge Enterprises Ltd. established a closed-population mating program in 1992, selecting from approximately 1200 colonies annually for honey production, wintering ability and chalk brood resistance. Tracheal mites were first observed in the colonies in the late 1990s, and Varroa mites were detected shortly thereafter. The selected population showed no resistance to either mite. To introduce mite resistance, Russian stock was imported as embryos from the USDA between 2001 and 2005 (Rinderer et al., 2001). Russian virgins from three different selections were close-population mated to selected colonies at the Meadow Ridge apiary. The F1 hybrids from these initial crosses were established at three different isolated apiaries, and used to backcross Russian virgins from subsequent shipments to regenerate Russian stock, and for re-selection under Canadian conditions. These apiaries served as a source of colonies for the natural selection apiary, and for drones in crosses used to increase Varroa resistance. In 2004, a natural selection apiary was established at an isolated area in Saskatchewan, called Saskatraz, using colonies from Meadow Ridge and collaborating Saskatchewan beekeepers. This apiary was established to further select for productive colonies with mite resistance and good wintering ability, without synthetic miticide treatment. Tracheal mites were introduced in the fall of 2004 by adding 200 worker bees with 60% tracheal mite infestations. Varroa mites were present in the original selections.

A colony phenotype called Saskatraz 88 (S88) was constructed by backcrossing a daughter from a Russian hybrid line selected at Saskatraz in 2006 to drones at an isolated Russian apiary (RP30) previously established at Meadow Ridge to increase Varroa tolerance. The resulting colony superseded and a daughter was mated at the RP30 apiary again, resulting in two back crosses at the RP30 apiary. Extensive screening of Varroa present on adult bee populations in both breeding populations and commercial colonies identified G4, a susceptible colony phenotype established in the summer of 2009. G4 bees showing high Varroa mite infestations during spring evaluations were selected and moved to an isolated apiary used as a Varroa nursery for experimental purposes. Susceptible colonies were not treated and left to die, serving to remove susceptible colonies from the breeding population. G4

and S88 were located in different apiaries during the course of the experiment. No queen events (swarming, supersedure) were noted in either S88 or G4 colonies during their lifespans. The S88 queen was last observed in the fall of 2010 in the Saskatraz natural selection apiary and failed in the spring of 2011.

Varroa infestations on adult bees (phoretic phase) were evaluated by washing 200–300 bees in 100% methanol. Analyses of Varroa in sealed brood (percent brood infestation and number of Varroa per cell) and natural Varroa drop onto sticky boards was also monitored. For molecular analyses, several hundred adult worker bees were collected from the brood nest and white-eyed, pink-eyed and dark-eyed pupae were collected from sealed brood of both S88 and G4 colonies in September 2010. Pupae and adult bees, either infested or not infested with Varroa mites, were collected. The samples were frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

## 2.2. DESIGN OF A HONEYBEE-SPECIFIC PEPTIDE ARRAY

To the authors' knowledge, no phosphorylation sites have been experimentally characterized in honeybee. As such, the following procedure was performed in order to identify putative honeybee phosphorylation sites. Experimentally-determined phosphorylation sites from other organisms were downloaded from the PhosphoSitePlus (Hornbeck et al., 2004, 2012) and Phospho.ELM (Diella et al., 2004, 2008; Dinkel et al., 2011) databases, and were combined into a single file. These included sites from organisms such as human, rat, mouse, cow, and *Drosophila melanogaster* (the closest honeybee relative for which phosphorylation sites are known). Phosphorylation sites were represented as 15-mer peptides, with the phosphorylated residue in the center and seven residues on either side. The honeybee proteome was constructed as follows. First, all of the honeybee proteins from UniProt (671 proteins) and GenBank (12,050 proteins) were downloaded. Second, the honeybee genome (Honeybee Genome Sequencing Consortium, 2006) was downloaded in the form of 16,501 contigs, and genes (along with their translations) were predicted using the program GeneMark.hmm (Lukashin and Borodovsky, 1998), giving 27,730 predicted proteins. Proteins from these three sources were then combined to create a final honeybee proteome consisting of 40,451 proteins. Using the DAPPLE program (Trost et al., 2013a), the 15-mer peptides from PhosphoSitePlus and Phospho.ELM were searched using BLAST against the honeybee proteome to find homologous sites. DAPPLE produced a table designed to facilitate the process of selecting honeybee peptides for inclusion on the array. Each row of the output table corresponded to a phosphorylation site from PhosphoSitePlus or Phospho.ELM. In addition to the sequence of the best hit in the honeybee proteome, the table contained the number of sequence differences between the query peptide and the honeybee peptide, with honeybee peptides having few sequence differences being preferred. The table also included the position (e.g., Y128) of the phosphoacceptor residue for both the query peptide and the hit peptide, with honeybee peptides where the position was similar for both query and hit being preferentially selected. In addition, peptide sequences contained within proteins from UniProt or GenBank were preferred over those from proteins predicted by GeneMark.hmm. Using the above

criteria, this list was manually curated to select appropriate honeybee phosphorylation sites for inclusion on the array. Peptides were selected that represent phosphorylation events associated with a broad spectrum of signaling pathways, but with specific emphasis on proteins and processes associated with innate immunity. A total of 299 peptides were ultimately selected. Each of these peptides was spotted in triplicate within each block. Further, each block was printed in triplicate, providing nine technical replicates for each peptide. Peptide synthesis, array spotting and quality control were performed as a commercial service (JPT Peptide Technologies, Berlin, Germany).

## 2.3. KINOME ANALYSIS

Application of the peptide arrays was based upon a previously reported protocol with modifications (Määttä et al., 2013). Briefly, individual frozen whole bees were placed in a sealed plastic bag in the presence of 300  $\mu\text{l}$  of lysis buffer. The bees were struck repeatedly with a rubber mallet and the suspension was centrifuged at  $10,000 \times g$  for 10 min. Supernatants were used for kinome analysis.

## 2.4. DATA ANALYSIS

The dataset for each array contained the signal intensities associated with the nine technical replicates for each of the 299 peptides for the whole body extracts of honeybee pupae or adults either uninfested or infested with Varroa mites. Those treatments were labeled "G4–" (susceptible and uninfested), "G4+" (susceptible and infested), "S88–" (resistant and uninfested), and "S88+" (resistant and infested). Kinome data were processed through PIIKA 2, a pipeline for processing kinome array data (Li et al., 2012; Trost et al., 2013b), with the following study specifics.

### 2.4.1. Consistency of technical replicates

For each peptide within a given array, a chi-square test was performed to determine whether the degree of variability among the technical replicates for that peptide was greater than would be expected by chance. Any peptide that had a *P*-value according to the chi-square test of less than 0.01 was considered to be inconsistently phosphorylated among the technical replicates.

### 2.4.2. Treatment-treatment variability analysis and pathway analysis

For each peptide, a paired *t*-test was used to compare its normalized signal intensity values under a treatment condition with those under a control condition. Three tests were performed for each peptide: G4+ versus G4–, S88+ versus S88–, and G4– versus S88–. Peptides with significant (*P*-value  $< 0.10$ ) changes in phosphorylation were identified. This level of significance was chosen to retain as much data as possible in order to facilitate subsequent pathway analysis (Li et al., 2012). Pathway and gene ontology (GO) analysis was performed as described previously (Kindrachuk et al., 2011; Määttä et al., 2013) using InnateDB (Lynn et al., 2008).

### 2.4.3. Cluster analysis

The pre-processed data were subjected to hierarchical clustering and principal component analysis (PCA) to cluster peptide

response profiles across arrays. Only peptides that were consistently phosphorylated among the technical replicates for all arrays were included in the clustering analysis. For each consistently-phosphorylated peptide on a given array, the average was taken over the nine replicates before performing clustering. For hierarchical clustering, the distance metric used was (1–Pearson correlation), while the linkage method used was that of McQuitty (1966). Subsets of peptides that could discriminate between resistant and susceptible bees were identified as described previously (Trost et al., 2013b).

## 2.5. VIRUS DETECTION

Bees were stored at  $-80^{\circ}\text{C}$  until RNA was extracted. Individual pupa were placed in small plastic bags, pulverized on dry ice, and solubilized in 700  $\mu\text{l}$  Trizol (Invitrogen Canada, Burlington, ON). RNA was purified using RNeasy Mini-columns (Qiagen Canada Inc., Mississauga, ON) and RNA concentration quantified with an Agilent 2100 Bioanalyzer using RNA 6000 Nano kits (Agilent Technologies Canada Inc., Mississauga, ON). RNA pellets were re-suspended in DEPC water and converted to cDNA using qScript cDNA Supermix (Quanta Biosciences, Gaithersburg, MD). qRT-PCR was performed using PerfeCta SYBR Green Supermix for IQ (Quanta Biosciences) on a BioRad IQ5 thermocycler. Deformed wing virus was detected using primers CA GTAGCTTGGGCGATTGTT (forward) and AGCTTCTGGAAC GGCAGATA (reverse) (Cox-Foster et al., 2007). Israeli acute paralysis virus was detected using primers GCGGAGAATATA AGGCTCAG (forward) and CTTGCAAGATAAGAAAGGGGG (reverse) (Di Prisco et al., 2011). Kashmir bee virus was detected using primers GATGAACGTCGACCTATTGA (forward) and TGTGGGTGTTGGCTATGAGTCA (reverse) (Cox-Foster et al., 2007). The presence of a single PCR product of the expected size was confirmed in 2% agarose gels (Invitrogen). Detection of DWV, IAPV, and KBV was performed using an end-point PCR protocol with Phusion polymerase (New England Biolabs, Whitby, ON) with amplification at  $98^{\circ}\text{C}$  for 30 s, then 30 cycles of:  $98^{\circ}\text{C}$  for 10 s,  $60^{\circ}\text{C}$  for 15 s, and  $72^{\circ}\text{C}$  for 20 s followed by 20 s at  $72^{\circ}\text{C}$ . Amplified products were visualized with ethidium bromide staining of 2% agarose gels. The real time cycling protocol for quantification of DWV was  $95^{\circ}\text{C}$  for 2 min, then 40 cycles of  $95^{\circ}\text{C}$  for 15 s,  $60^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 30 s, followed by a melt curve to confirm amplification of a single product.

## 3. RESULTS

### 3.1. CHARACTERIZATION OF VARROA MITE SUSCEPTIBLE AND RESISTANT BEE PHENOTYPES

Varroa mite infestation was quantified yearly between 2007 and 2011 for the resistant (S88) colony and in 2010 for the susceptible (G4) colony (Figure 1A). In 2009, the average Varroa infestation rates for S88 remained below 10 per 100 bees (PHB) but ranged as high as 19 PHB. In 2010, eight samples were analyzed between May and October showing an average infestation of three to five PHB in the S88 colony. Adult bee samples with and without Varroa were sampled in September for kinome analyses, when phoretic mite levels were four PHB (Figure 1A). S88 died in April 2011 with a Varroa mite population of nine PHB after a colony lifespan of 58 months. This colony resisted Varroa mite population growth throughout its lifetime, although significant

levels of Varroa mites persisted in the colony from establishment. High levels of phoretic Varroa were detected in May 2010 in G4 and reached as high as 67 PHB. Varroa mite population growth was very rapid in this colony (Figure 1A). Adult bees with and without Varroa were sampled for kinome analyses when phoretic Varroa populations were highest (September 2010). G4 died in October with a lifespan of 17 months.

These resistant and susceptible colonies were further defined by evaluating Varroa infestation in the sealed brood at the same time as adult bee samples were collected for molecular analyses. Honeybee colonies during September in Western Canada decrease brood rearing and the adult population begins to decline. Varroa increase migration into the brood, and brood Varroa levels can quickly increase. Scoring sealed G4 brood cells ( $n = 500$ ) revealed that 88%, 84%, and 70% of the white-eyed, pink-eyed and dark-eyed pupae, respectively, were Varroa-infested (Figure 1B). The phoretic mite levels on adult G4 bees (67 PHB) was similar to the infestation rate for dark-eyed pupae. In contrast, S88 brood infestation levels were much lower, with dark-eyed pupae infestation levels dropping to 17% from 44% and adult phoretic levels to four PHB (Figure 1B). These results imply that S88 resists Varroa population growth by removing Varroa from the brood. In addition, fewer Varroa per cell were detected in dark-eyed pupae and pre-emergent pupae in S88 than G4 at July 2010 sampling dates. G4 showed  $2.7 \pm 2.0$  Varroa per cell ( $\pm$  standard error of the mean,  $n = 70$ ), and S88 showed  $1.5 \pm 1.0$  Varroa per cell ( $n = 9$ ).

### 3.2. DEVELOPMENT OF A BEE-SPECIFIC PEPTIDE ARRAY

The bee-specific peptide array was developed using the DAPPLE program (Trost et al., 2013a) as described in section 2. DAPPLE predicted nearly 10,000 phosphorylation events within the honeybee proteome. Of the predicted phosphorylation events, approximately 0.6% were exactly conserved over a peptide of 15 amino acids (seven residues flanking each side of the phosphoacceptor site) (Supplementary Table 1). The low degree of conservation highlights the importance of developing species-specific arrays as opposed to simply translating commercially available arrays across species.

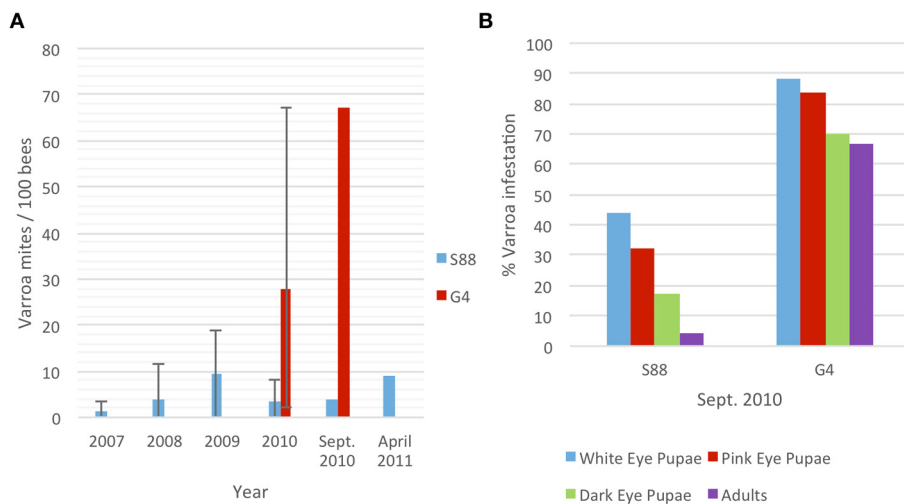
From this panel, 299 unique phosphorylation events were selected using the criteria described in section 2. Peptides were selected to represent phosphorylation events associated with a broad spectrum of signaling pathways (to facilitate novel discovery) but with emphasis on pathways and processes associated with insect innate immunity. A GenePix Array List (GAL) file containing the exact layout and content of the array used in this study is provided (Supplementary File 1).

An image highlighting the format of the arrays as well as the consistency and reproducibility of peptide spotting is presented (Figure 2A). An image of a data scan of a representative array used for analysis of a whole-bee lysate is also provided (Figure 2B). All of the arrays used in this study were of comparable quality with respect to the clarity and consistency of peptide phosphorylation.

### 3.3. KINOME PROFILING OF BEE PHENOTYPE AT DIFFERENT DEVELOPMENTAL STAGES

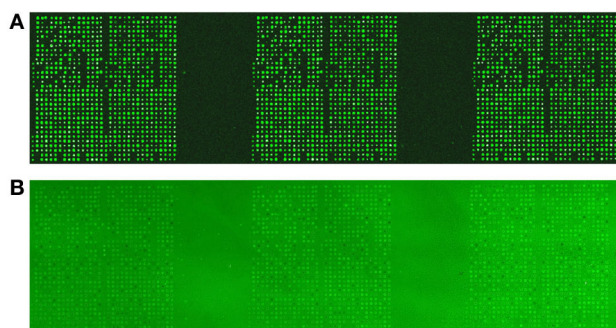
Uninfested bees ( $n = 3$ ) of each phenotype (G4 and S88) were considered at each of three developmental stages (pink-eyed





**FIGURE 1 | Quantification of Varroa mite infestation of G4 and S88 bees.** (A) Average phoretic Varroa infestations per 100 bees in S88 and G4 colonies. Bars show the range of yearly phoretic Varroa infestations in S88 (2007–2010)

and G4 (2010). (B) Percent Varroa infestation in sealed brood at different stages of development. Over 500 sealed brood cells were analyzed for each colony and scored for presence of Varroa.



**FIGURE 2 | Printing and validation of the bee-specific peptide array.** (A) The arrays were printed by a commercial provider (JPT Peptide Technologies, Berlin, Germany). For each array, each spot was printed in triplicate within each block. Each block was then printed in triplicate, for a total of nine technical replicates of each peptide. This image, taken as a quality control step in array production, illustrates the consistency and reproducibility of peptide spotting. (B) An image of a data scan of a representative array used for analysis of a whole-bee sample. A clear and consistent pattern of peptide phosphorylation is apparent across the three printed blocks.

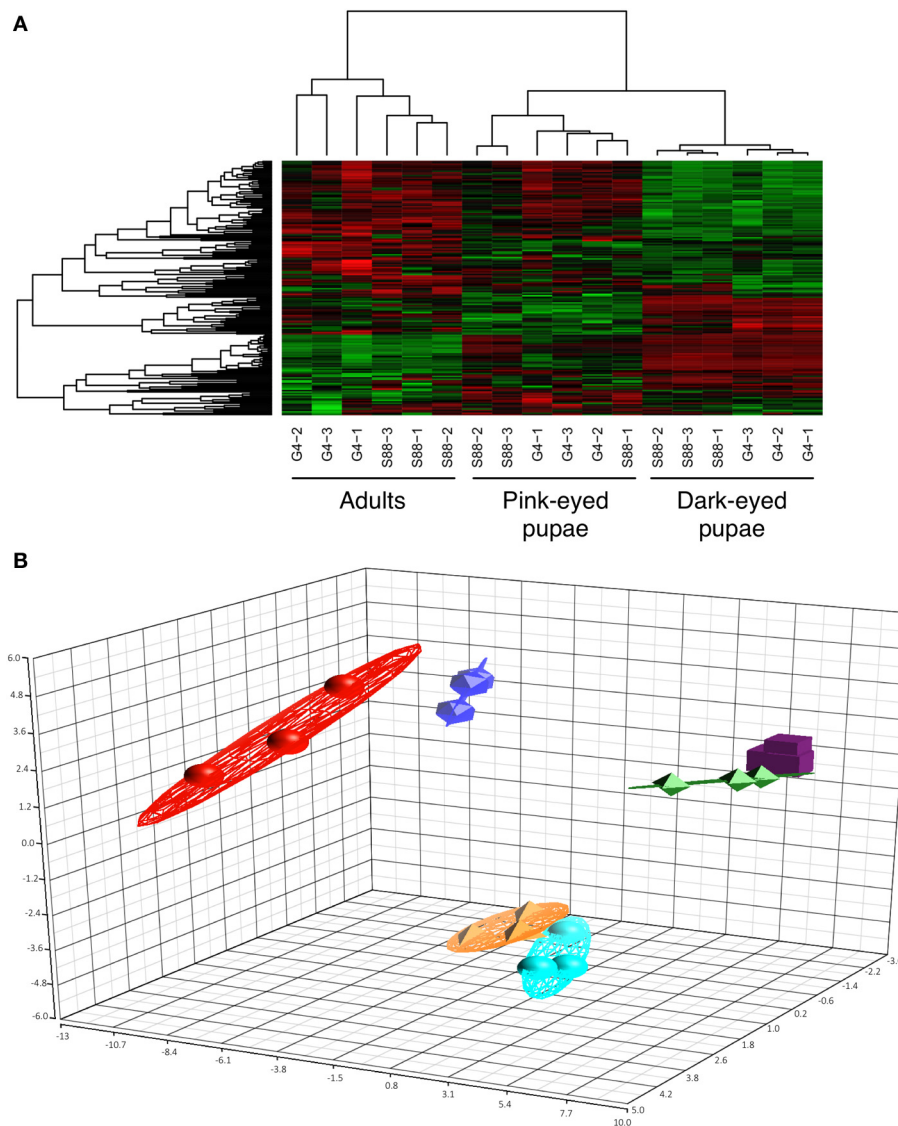
pupae, dark-eyed pupae and adult). In each case, kinome analysis was performed with lysate extracted from the whole organism. Morphologically, there was a clear distinction between each developmental stage. There was, however, no obvious difference in bee morphology when comparing between G4 and S88 within each developmental stage. The relationships among the 18 kinome datasets were evaluated through hierarchical clustering (Figure 3A) and three-dimensional PCA (Figure 3B). There was a clear indication of distinct developmentally-specific kinome profiles. Further, within each developmental stage, there was strong evidence of distinct kinome profiles for the G4 and S88 bees, indicating that Varroa mite

susceptibility or resistance is reflected at the level of signal transduction.

### 3.4. PHOSPHOMARKERS OF VARROA MITE SUSCEPTIBILITY IN DARK-EYED PUPAE

The ability of the arrays to detect distinct kinome profiles (kinotypes) corresponding to each phenotype suggests that the arrays may represent a valuable tool for identification of kinase activity biomarkers that are associated with resistance or the response to Varroa mite infestation. Specifically, the bee-specific peptide array, representing 299 phosphorylation events, was able to discriminate between each developmental stage, and between the two phenotypes within each developmental stage (Figure 3).

To determine whether smaller sets of peptides could also discriminate between the phenotypes, the peptide subset analysis described by Trost et al. (2013b) was performed on the bees at the dark-eyed pupae stage. This procedure was used to identify subsets of peptides having the property that, when samples were clustered using these peptides, bees of the same phenotype clustered together as closely as possible. This was done for peptide subsets of size 3–200. For subsets of selected cardinalities (5, 10, 25, 50, 100, 150, and 200), the random tree analysis described by Trost et al. (2013b) was performed to determine whether that set of peptides discriminated between the susceptible and resistant phenotypes better than would be expected by chance. It was discovered that subsets of as few as five peptides could discriminate the resistant and susceptible bee phenotypes with a high degree of confidence ( $P$ -value < 0.001) (Table 1). Given this, it may be possible to create a smaller, more targeted array that could provide unique kinomic profiles for each phenotype. Such a peptide subset could serve as a minimal array of practical value for screening bees within breeding programs as well as for assurance of phenotype in the sales and marketing of commercial bees.



**FIGURE 3 | Clustering of the kinome profiles of bees of different phenotypes at different developmental stages. (A)** Hierarchical clustering of kinome datasets. (1—Pearson correlation) was used as the distance metric, while McQuitty linkage was used as the linkage method. Each column represents the kinome activity of individual bees ( $n = 3/\text{treatment}$ ). The kinome profiles of the bees segregated first by developmental stage and then largely by colony phenotype (S88: resistant; G4: susceptible). Colors indicate the average (over nine

intra-array replicates) normalized phosphorylation intensity of each target, with red indicating greater amounts of phosphorylation and green indicating lesser amounts of phosphorylation. **(B)** Principal component analysis. The first three principal components are shown. The points are as follows: red, adult G4; dark blue, adult S88; green, dark-eyed G4; purple, dark-eyed S88; orange, pink-eyed G4; light blue, pink-eyed S88. The proportions of variance explained by the first, second, and third principal components were 29.1%, 15.3%, and 7.5%, respectively.

### 3.5. KINOMIC RESPONSES OF SUSCEPTIBLE AND RESISTANT DARK-EYED PUPAE TO VARROA MITE CHALLENGE

Kinome profiles were determined for individual dark-eyed pupae ( $n = 3$ ) of both the G4 and S88 colony phenotypes in the presence and absence of Varroa mite infestation. Hierarchical clustering analysis of the kinome data demonstrated distinct clustering on the basis of Varroa mite susceptibility, indicating distinct patterns of phosphorylation-mediated signal transduction within the two phenotypes (Figure 4A). This was confirmed with PCA, in which distinct clustering of samples

corresponding to the phenotypes was also observed (Figure 4B). For both hierarchical clustering and PCA, there was further sub-clustering based on the infestation status of the samples within the susceptible phenotype. This sub-clustering was not observed within the resistant samples, except for one S88 infested pupae which showed some overlap with the susceptible G4 phenotype. These observations imply Varroa parasitism induced a more pronounced change in intracellular physiology within Varroa susceptible bees compared to resistant bees.

**Table 1 | Ability of subsets of peptides to discriminate susceptible and resistant bees at the dark-eyed pupae stage.**

Number of peptides	P-value
200	0.0006
150	0.0002
100	0.0007
50	0.0001
25	0.0002
10	0.0004
5	0.0004

Subsets of peptides were determined that best differentiated susceptible and resistant dark-eyed pupae. For selected subsets, a statistical test (Trost et al., 2013b) was used to determine whether those peptides could discriminate between the two phenotypes better than would be expected by chance. The first column of the table contains the size of the peptide subset, while the second column contains the P-value associated with this statistical test.

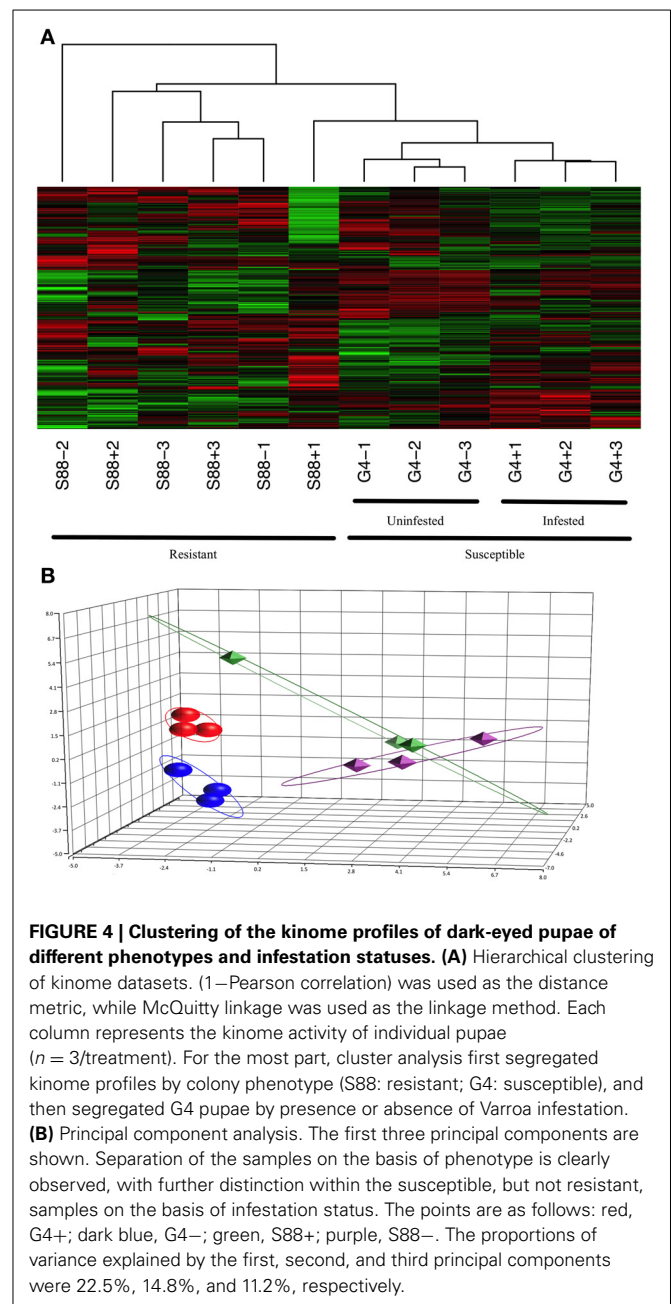
### 3.6. CELLULAR MECHANISMS OF VARROA MITE SUSCEPTIBILITY

The kinome data were interrogated to define the biological differences between bee phenotypes at the dark-eyed pupae stage of development. Many peptides were differentially phosphorylated between phenotypes or treatments. For instance, in the uninfested samples of each phenotype, there were 153 peptides (over half of the peptides on the array) for which there were significant ( $P$ -value < 0.1) differences in phosphorylation between the phenotypes. This is consistent with resistance to Varroa mite infestation being a complex and multi-faceted process.

Specific consideration of these differentially phosphorylated peptides from the perspective of gene ontology and pathway overrepresentation analysis revealed a number of points of biological difference between uninfested bees of the resistant and susceptible phenotypes (Table 2 and Supplementary Table 2), between infested and uninfested bees of the susceptible phenotype (Table 3 and Supplementary Table 3), and between infested and uninfested bees of the resistant phenotype (Table 4 and Supplementary Table 4). When comparing uninfested bees from the two phenotypes, there were no clear differences in pathways and processes associated with immune function (Table 2 and Supplementary Table 2). An interesting exception is that within the G4 pupae, there was a trend toward the down-regulation of innate immunity ( $P$ -value < 0.1) in response to Varroa mite infestation (Table 3). Down-regulation of innate immune processes in response to Varroa mite infestation was not observed in the resistant phenotype (Table 4).

### 3.7. DETECTION OF SECONDARY VIRAL INFECTIONS

For bees of both phenotypes, at the dark-eyed pupae stage of development and in the absence of Varroa mites, there was a shared presence of detectable, but low levels of DWV (Figure 5). However, in the presence of Varroa mites there was an approximately 10,000-fold increase in DWV RNA relative to the Varroa mite-free pupae (Figure 5). There was also no detectable IAPV and KBV RNA in pupae of both phenotypes, regardless of the presence or absence of mite infestation (data not shown). These observations support the hypothesis that Varroa mites serve as



a vector for virus transmission and that both phenotypes experience equal levels of viral infection following mite infestation. This observation supports the conclusion that kinotypic differences between pupae from the two phenotypes reflect differences in host responses to the Varroa mite and not viral infection.

The presence of immunosuppression was suggested by kinome data analysis of susceptible bees at the dark-eyed pupae stage of development. If this immunosuppression persists throughout the life of a bee, then the ability of bees to counter further infection by secondary pathogens may be compromised. Consistent with this hypothesis, screening for two additional viral bee pathogens, IAPV and KBV, confirmed higher rates of infection in the susceptible adult bees in the face of Varroa mite infestation (Table 5).

**Table 2 | Gene ontology analysis of uninfested resistant and susceptible dark-eyed pupae (S88-/G4-).**

Category	Name	ID	1	2	3	4	5
Biological process	Cell cycle arrest	GO:0007050	5	5	0.040	0	1
	Response to peptide hormone stimulus	GO:0043434	4	4	0.078	0	1
	ATP biosynthetic process	GO:0006754	4	0	1	4	0.03
	Positive regulation of neuron apoptosis	GO:0043525	4	0	1	4	0.03
	Cytoskeleton organization	GO:0007010	6	1	0.99	5	0.05
Cellular component	Cell surface	GO:0009986	7	6	0.082	1	0.98
	Golgi apparatus	GO:0005794	7	1	0.99	6	0.02
	Plasma membrane	GO:0005886	33	14	0.96	19	0.04

Based on levels of differential expression or phosphorylation, InnateDB (Lynn et al., 2008) can predict upregulated or downregulated pathways that are consistent with the experimental data. Pathways are assigned a *P*-value based on the number of proteins present for a particular pathway. The numbered columns are as follows: 1, total genes uploaded for that pathway; 2, number of genes up-phosphorylated; 3, *P*-value for up-phosphorylation; 4, number of genes down-phosphorylated; 5, *P*-value for down-phosphorylation.

**Table 3 | Gene ontology analysis of susceptible dark-eyed pupae (G4+/G4-).**

Category	Name	ID	1	2	3	4	5
Biological process	Transport	GO:0006810	4	0	1	4	0.081
	Innate immune response	GO:0045087	27	9	0.945	18	0.098
	Cell cycle	GO:0007049	9	1	0.99	8	0.03
	DNA repair	GO:0006281	5	0	1	5	0.04
	Mitotic cell cycle	GO:0000278	5	0	1	5	0.04
	Glycolysis	GO:0006096	7	6	0.031	1	0.99
	Phosphatidylinositol-mediated signaling	GO:0048015	4	4	0.039	0	1
	Multicellular organismal development	GO:0007275	6	5	0.064	1	0.992
Cellular component	Nucleoplasm	GO:0005654	22	7	0.95	15	0.106
	Plasma membrane	GO:0005886	30	17	0.099	13	0.94
	Golgi apparatus	GO:0005794	8	1	0.99	7	0.05
	Integral to membrane	GO:0016021	8	6	0.081	2	0.98
	Basolateral plasma membrane	GO:0016323	4	4	0.038	0	1
Molecular function	ATPase activity	GO:0016887	4	0	1	4	0.081
	RNA binding	GO:0003723	4	0	1	4	0.081
	RNA pol. II transcription factor activity	GO:0003705	4	4	0.038	0	1
	GTP binding	GO:0005525	6	5	0.064	1	0.992

For details, see the caption and footnote of **Table 2**.

#### 4. DISCUSSION

There is a clear and emerging priority for the ability to define global host responses at the level of phosphorylation-mediated signal transduction. As technologies advance, there is greater opportunity to apply these approaches to a broader range of species as well as samples of increasing biological complexity. Kinome analysis is often performed on cellular samples of low complexity, such as cultured cells, or purified primary cell populations, such as monocytes. Recently, there have been demonstrations of kinome analysis of samples of greater biological complexity, such as organ samples (Arsenault et al., 2013b) and intestinal tissue (Määttä et al., 2013). The current report, to the best of our knowledge, represents the first development of an insect-specific peptide kinome array as well as the first application

of kinome analysis at the whole-organism level. The incentive to push the technology in this direction was to develop a research tool of value in the understanding of colony collapse disorder of bees. Specifically, we sought to apply the bee-specific array to populations with differing resistance to Varroa mite infestation, in the presence and absence of this critical pathogen, to provide insight into mechanisms of disease resistance as well as biomarkers for strategic bee breeding programs.

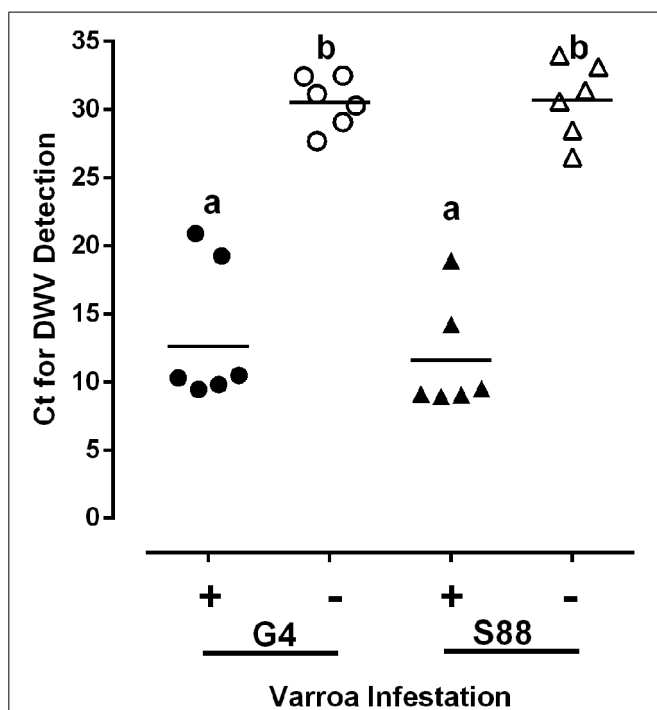
The kinome data emerging from analysis of distinct phenotypes (susceptible and resistant) at three developmental stages (pink-eyed pupae, dark-eyed pupae, and adults) provided clear evidence of a phenotype-associated kinotype. As might be anticipated, each stage of development was also associated with a different global pattern of signal transduction activity. Within



**Table 4 | Gene ontology analysis of resistant dark-eyed pupae (S88+/S88–).**

Category	Name	ID	1	2	3	4	5
Biological process	RNA metabolic process	GO:0016070	5	1	0.99	4	0.067
	mRNA metabolic process	GO:0016071	5	1	0.99	4	0.067
	Nerve growth factor receptor signaling	GO:0048011	7	7	0.028	0	1
	Positive regulation of apoptotic process	GO:0043065	5	5	0.082	0	1
	Peptidyl-serine phosphorylation	GO:0018105	7	2	0.99	5	0.069
Molecular function	Phosphor-transferase activity	GO:0016772	15	12	0.087	3	0.978
	Protein kinase binding	GO:0019901	4	0	1	4	0.018
	RNA binding	GO:0003723	5	1	0.99	4	0.067
	DNA binding	GO:0003677	7	2	0.99	5	0.069
	Kinase activity	GO:0016301	14	6	0.97	8	0.094

For details, see the caption and footnote of **Table 2**.



**FIGURE 5 | Virus presence in honeybee populations.** The level of deformed wing virus (DWV) present in dark-eyed pupae was compared in the presence (+) or absence (–) of a detectable Varroa mite infestation. DWV was detected using qRT-PCR and the level of viral infection was measured as the threshold cycle (Ct) for viral RNA amplification. Ct values are inversely proportional to the abundance of viral RNA. Data presented are values for individual pupae ( $n = 6/\text{group}$ ). Significant differences ( $P\text{-value} < 0.05$ ) among treatment groups are denoted by different letters above each column.

these development-specific patterns of clustering, there was clear evidence for distinct sub-profiles corresponding to each of the Varroa mite susceptibility phenotypes. This suggests the potential to translate the arrays into a tool that could be utilized to inform commercial aspects of bee production, such as sales and breeding. Phosphosignatures that reflect important phenotypes, such

**Table 5 | Percentage of resistant and susceptible adult bees with detectable virus.**

Virus	G4 (%)	S88 (%)
DWV	100	100
IAPV	60	0
KBV	15	0

Bees ( $n = 20/\text{group}$ ) were sampled in September 2010 (see **Figure 1A**). Viruses were detected using 30 cycles of amplification in qRT-PCR, and amplified products were visualized by agarose gel electrophoresis. Specific primer pairs were used to detect deformed wing virus (DWV), Israeli acute paralysis virus (IAPV), and Kashmir bee virus (KBV).

as disease resistance or production value, could be incorporated into a second generation honeybee-specific array.

In the absence of Varroa mite infestation, there were clear and consistent differences in the signaling profiles of the susceptible and resistant bees. The magnitude of these differences suggests that resistance is a complex, multifactorial process. Interestingly, for the uninfested bees there were no obvious differences between the two phenotypes that relate to pathways or processes immediately associated with immunity. This is consistent with a previous investigation of the biological basis of Varroa mite susceptibility phenotypes through gene expression approaches, which suggested that differences in behavior, rather than immune function, underlie Varroa resistance (Navajas et al., 2008). The most well-defined traits associated with Varroa resistance are hygienic behavior and grooming behavior that function to maintain lower Varroa populations (Harbo and Harris, 2009; Tsuruda et al., 2012). The S88 phenotype also showed better grooming behavior (unpublished observations). However, in our breeding efforts, it is difficult to stabilize Varroa resistant phenotypes, and the progeny of selected colony phenotypes are highly variable. Colony phenotypes can also change over time within the same colony. The survival of a resistant phenotype may be due to combinations of grooming and hygienic behavior as well as undefined mechanisms that restrict the propagation of viral pathogens. This combination of traits may be critical for bee survival in the presence of a persistent Varroa infestation. Elucidation of the

mechanisms involved in this resistance to colony collapse may be critical for breeding bees able to tolerate low levels of persistent Varroa parasitism while maintaining colony health.

The responses of the two bee phenotypes to Varroa mite infestation in the current study were also investigated using pathway over-representation and gene ontology analysis. For the resistant bees, a small number of pathways were found to be activated in response to Varroa infestation. Specifically, there was robust activation of MAPK signaling, which may represent the most effective host response through induction of stress response pathways. Activation of MAPK signaling has been linked to successful management of pathogenic challenge in a number of species, including insects (Arthur and Ley, 2013). In contrast, within the susceptible bees, there were more far-reaching consequences to Varroa mite challenge, including evidence for a down-regulation of innate immune responses.

There are conflicting opinions in the literature regarding the significance of host immunity, and the potential ability of Varroa mites to compromise host immunity. For example, some investigations have reported that Varroa mites, or virus associated with mites, compromise honeybee immunity (Gregory et al., 2005) and promote amplification of bee viruses (Yang and Cox-Foster, 2005). From a more global perspective, a number of ectoparasites immunosuppress their vertebrate hosts and increase susceptibility to infectious disease (Yang and Cox-Foster, 2005). Varroa mites may contribute to colony collapse by suppressing bee immunity and promoting secondary viral infections (Yang and Cox-Foster, 2005; Evans and Schwarz, 2011). Given the conserved transmission route associated with many bee parasites, co-infection of individual bees and colonies by multiple viral pathogens is a common occurrence that can have direct and indirect interactions that may be additive, synergistic or neutral in consequences to the host (Evans and Schwarz, 2011). Varroa mites are associated with a number of honeybee RNA viruses. In this capacity, the mites are known to contribute to colony failure both by acting as a reservoir and incubator for the viruses as well as facilitating their spread among bees (Nazzi et al., 2012). Our work adds another layer to this synergy by suggesting that infestation by the mite renders the bee host more susceptible to viral infection by compromising the innate immune system.

Our kinome data strongly indicate that differences in immune capabilities are likely not involved in Varroa susceptibility; rather, this phenotype may reflect primarily behavioral differences. Following Varroa mite infestation, however, the immunosuppression observed in the susceptible bees may influence their ability to counter further infestation by mites as well as secondary viral pathogens. This hypothesis is supported by greater diversity of secondary viral infections in the susceptible bees following Varroa mite infestation. This could occur at the level of the individual bees as well as the entire colony. The ultimate collapse of these colonies may represent the collective toll of these combined infections, as well as other potential stressors. This suggests that bees are not susceptible to Varroa mite infestation because they are immunocompromised; rather, they are immunocompromised because they are infested with Varroa mites. This understanding, in concert with the use of the arrays to identify appropriate biomarkers, may enable strategic breeding

and management efforts to deal with the problem of Varroa parasitism and honeybee colony loss worldwide.

This initial kinome-wide analysis of honeybees has generated a number of important questions that motivate further experimental investigation. For example, more targeted investigation of the host-pathogen interaction between honeybees and Varroa mites may confirm the hypothesis that the vulnerability of the susceptible bees reflects consequences of Varroa mite infestation, as well as evidence of the molecular mechanisms involved. Unknown factors may be acting at the cellular level in Varroa resistant bees identified by natural selection (survival colonies), which may or may not be present in bees showing behavioral characteristics for expression of Varroa resistance. These factors may protect against the fatal effects associated with viruses (DWV, IAPV, KBV) vectored by Varroa, or may reduce the ability of Varroa to cause deficiencies in innate immune or stress responses. Experiments are in progress using honeybee kinome analyses to investigate these possibilities in individual bees from inbred colony lines showing varying degrees of resistance and susceptibility to Varroa. Additionally, the ability of the proposed phosphorylation-associated biomarkers of Varroa mite susceptibility should be evaluated in large-scale investigations of honeybees representing a spectrum of susceptibilities. The ability of these markers to effectively discriminate and predict this important phenotype within the context of naturally occurring variance will be important for determining the value of these markers. Ultimately, a methodology for using specific, targeted subsets of the peptide array probes (just 5–10 of them) to identify Varroa resistant and susceptible phenotypes needs to be developed.

## AUTHOR CONTRIBUTIONS

Albert J. Robertson designed the breeding program, helped plan the kinome array experiments, participated in data analysis, supervised the research, and wrote parts of the manuscript. Brett Trost helped design the honeybee-specific kinome arrays, participated in data analysis, and wrote parts of the manuscript. Erin Scruten performed the kinome array experiments. Thomas Robertson and Mohammad Mostajeran performed bee selections and helped define the resistant and susceptible phenotypes. Wayne Connor performed the virus quantification experiments. Anthony Kusalik and Philip Griebel participated in data analysis and supervised the research. Scott Napper helped design the honeybee-specific arrays, participated in data analysis, wrote parts of the manuscript, and supervised the research. All authors participated in revising the manuscript, with particular contributions by Albert J. Robertson, Brett Trost, Anthony Kusalik, and Scott Napper.

## FUNDING

This work was funded in part by grants from Saskatchewan Agriculture (Agriculture Development Fund) and the Agriculture Council of Saskatchewan to Albert J. Robertson and by Meadow Ridge Enterprises Ltd. Philip Griebel holds a Tier I Canada Research Chair funded by the Canadian Institutes of Health Research. Brett Trost and Anthony Kusalik received funding from the Natural Sciences and Engineering Research Council of Canada (NSERC).

## ACKNOWLEDGMENTS

We thank the Saskatchewan Beekeepers Association for administration of funds, Sanjie Jiang and Syed Shah for help with sample collection, John Pedersen and Neil Morrison for help with breeding work, Dr. Bob Danka (USDA) for critically reviewing the manuscript, and all Saskatchewan beekeepers who supported the Saskatraz breeding program.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fgene.2014.00139/abstract>

## REFERENCES

- Arsenault, R., Griebel, P., and Napper, S. (2011). Peptide arrays for kinome analysis: new opportunities and remaining challenges. *Proteomics* 11, 4595–4609. doi: 10.1002/pmic.201100296
- Arsenault, R. J., Li, Y., Bell, K., Doig, K., Potter, A., Griebel, P. J., et al. (2012). Mycobacterium avium subsp. paratuberculosis inhibits gamma interferon-induced signaling in bovine monocytes: insights into the cellular mechanisms of Johne's disease. *Infect. Immun.* 80, 3039–3048. doi: 10.1128/IAI.00406-12
- Arsenault, R. J., Li, Y., Maattanen, P., Scruten, E., Doig, K., Potter, A., et al. (2013a). Altered toll-like receptor 9 signaling in Mycobacterium avium subsp. paratuberculosis-infected bovine monocytes reveals potential therapeutic targets. *Infect. Immun.* 81, 226–237. doi: 10.1128/IAI.00785-12
- Arsenault, R. J., Napper, S., and Kogut, M. H. (2013b). Salmonella enterica Typhimurium infection causes metabolic changes in chicken muscle involving AMPK, fatty acid and insulin/mTOR signaling. *Vet. Res.* 44, 35. doi: 10.1186/1297-9716-44-35
- Arthur, J. S. C., and Ley, S. C. (2013). Mitogen-activated protein kinases in innate immunity. *Nat. Rev. Immunol.* 13, 679–692. doi: 10.1038/nri3495
- Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., et al. (2007). A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318, 283–287. doi: 10.1126/science.1146498
- Di Prisco, G., Pennacchio, F., Caprio, E., Boncristiani, H. F. Jr., Evans, J. D., and Chen, Y. (2011). *Varroa destructor* is an effective vector of Israeli acute paralysis virus in the honeybee, *Apis mellifera*. *J. Gen. Virol.* 92, 151–155. doi: 10.1099/vir.0.023853-0
- Diella, F., Cameron, S., Gemünd, C., Linding, R., Via, A., Kuster, B., et al. (2004). Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* 5:79. doi: 10.1186/1471-2105-5-79
- Diella, F., Gould, C. M., Chica, C., Via, A., and Gibson, T. J. (2008). Phospho.ELM: a database of phosphorylation sites—update 2008. *Nucleic Acids Res.* 36, D240–D244. doi: 10.1093/nar/gkm772
- Dietemann, V., Nazzi, F., Martin, S. J., Anderson, D. L., Locke, B., Delaplane, K. S., et al. (2013). Standard methods for varroa research. *J. Apicult. Res.* 52, 1. doi: 10.3896/IBRA.1.52.1.09
- Dinkel, H., Chica, C., Via, A., Gould, C. M., Jensen, L. J., Gibson, T. J., et al. (2011). Phospho.ELM: a database of phosphorylation sites—update 2011. *Nucleic Acids Res.* 39, D261–D267. doi: 10.1093/nar/gkq1104
- Eglen, R., and Reisine, T. (2011). Drug discovery and the human kinome: recent trends. *Pharmacol. Ther.* 130, 144–156. doi: 10.1016/j.pharmthera.2011.01.007
- Evans, J. D., and Schwarz, R. S. (2011). Bees brought to their knees: microbes affecting honey bee health. *Trends Microbiol.* 19, 614–620. doi: 10.1016/j.tim.2011.09.003
- Greenleaf, S. S., and Kremen, C. (2006). Wild bees enhance honey bees' pollination of hybrid sunflower. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13890–13895. doi: 10.1073/pnas.0600929103
- Gregory, P. G., Evans, J. D., Rinderer, T., and de Guzman, L. (2005). Conditional immune-gene suppression of honeybees parasitized by Varroa mites. *J. Insect Sci.* 5:7. Available online at: [http://www.bioone.org/doi/abs/10.1672/1536-2442\(2005\)005%5B0001%3ACISOHP%5D2.0.CO%3B2](http://www.bioone.org/doi/abs/10.1672/1536-2442(2005)005%5B0001%3ACISOHP%5D2.0.CO%3B2)
- Harbo, J. R., and Harris, J. W. (2009). Responses to Varroa by honey bees with different levels of varroa sensitive hygiene. *J. Apic. Res.* 48, 156–161. doi: 10.3896/IBRA.1.48.3.02
- Honeybee Genome Sequencing Consortium (2006). Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443, 931–949. doi: 10.1038/nature05260
- Hornbeck, P. V., Chabra, I., Kornhauser, J. M., Skrzypek, E., and Zhang, B. (2004). PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* 4, 1551–1561. doi: 10.1002/pmic.200300772
- Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., et al. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261–D270. doi: 10.1093/nar/gkr1122
- Jalal, S., Arsenault, R., Potter, A. A., Babiuk, L. A., Griebel, P. J., and Napper, S. (2009). Genome to kinome: species-specific peptide arrays for kinome analysis. *Sci. Signal.* 2, pl1. doi: 10.1126/scisignal.254pl1
- Kindrachuk, J., Scruten, E., Attah-Poku, S., Bell, K., Potter, A., Babiuk, L. A., et al. (2011). Stability, toxicity, and biological activity of host defense peptide BMAP28 and its inversed and retro-inversed isomers. *Biopolymers* 96, 14–24. doi: 10.1002/bip.21441
- Le Conte, Y., Alaux, C., Martin, J.-F., Harbo, J. R., Harris, J. W., Dantec, C., et al. (2011). Social immunity in honeybees (*Apis mellifera*): transcriptome analysis of varroa-hygienic behaviour. *Insect Mol. Biol.* 20, 399–408. doi: 10.1111/j.1365-2583.2011.01074.x
- Le Conte, Y., de Vaublanc, G., Crauser, D., Jeanne, F., Rousselle, J.-C., and Bécard, J. M. (2007). Honey bee colonies that have survived *Varroa destructor*. *Apidologie* 38, 566–572. doi: 10.1051/apido:2007040
- Li, Y., Arsenault, R. J., Trost, B., Slind, J., Griebel, P. J., Napper, S., et al. (2012). A systematic approach for analysis of Peptide array kinome data. *Sci. Signal.* 5, pl2. doi: 10.1126/scisignal.2002429
- Lodesani, M., and Costa, C. (2005). Limits of chemotherapy in beekeeping: development of resistance and the problem of residues. *Bee World* 86, 102–109.
- Lukashin, A. V., and Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26, 1107–1115. doi: 10.1093/nar/26.4.1107
- Lynn, D. J., Winsor, G. L., Chan, C., Richard, N., Laird, M. R., Barsky, A., et al. (2008). InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Mol. Syst. Biol.* 4, 218. doi: 10.1038/msb.2008.55
- Määttänen, P., Trost, B., Scruten, E., Potter, A., Kusalik, A., Griebel, P., et al. (2013). Divergent immune responses to Mycobacterium avium subsp. paratuberculosis infection correlate with kinome responses at the site of intestinal infection. *Infect. Immun.* 81, 2861–2872. doi: 10.1128/IAI.00339-13
- Martin, S. J., Ball, B. V., and Carreck, N. L. (2013). The role of deformed wing virus in the initial collapse of varroa infested honey bee colonies in the UK. *J. Apicult. Res.* 52, 251–258. doi: 10.3896/IBRA.1.52.5.12
- Martin, S. J., Highfield, A. C., Brettell, L., Villalobos, E. M., Budge, G. E., Powell, M., et al. (2012). Global honey bee viral landscape altered by a parasitic mite. *Science* 336, 1304–1306. doi: 10.1126/science.1220941
- McQuitty, L. L. (1966). Similarity analysis by reciprocal pairs for discrete and continuous data. *Educ. Psychol. Meas.* 26, 825–831. doi: 10.1177/001316446602600402
- Mullin, C. A., Frazier, M., Frazier, J. L., Ashcraft, S., Simonds, R., Vanengelsdorp, D., et al. (2010). High levels of miticides and agrochemicals in North American apiaries: implications for honey bee health. *PLoS ONE* 5:e9754. doi: 10.1371/journal.pone.0009754
- Mulongo, M., Prysliak, T., Scruten, E., Napper, S., and Perez-Casal, J. (2014). *In vitro* infection of bovine monocytes with *Mycoplasma bovis* delays apoptosis and suppresses production of gamma interferon and tumor necrosis factor alpha but not interleukin-10. *Infect. Immun.* 82, 62–71. doi: 10.1128/IAI.00961-13
- Navajas, M., Migeon, A., Alaux, C., Martin-Magniette, M., Robinson, G., Evans, J., et al. (2008). Differential gene expression of the honey bee *Apis mellifera* associated with *Varroa destructor* infection. *BMC Genomics* 9:301. doi: 10.1186/1471-2164-9-301
- Nazzi, F., Brown, S. P., Annoscia, D., Del Piccolo, F., Di Prisco, G., Varricchio, P., et al. (2012). Synergistic parasite-pathogen interactions mediated by host immunity can drive the collapse of honeybee colonies. *PLoS Pathog.* 8:e1002735. doi: 10.1371/journal.ppat.1002735
- Parker, R., Guarna, M. M., Melathopoulos, A. P., Moon, K.-M., White, R., Huxter, E., et al. (2012). Correlation of proteome-wide changes with social immunity behaviors provides insight into resistance to the parasitic mite, *Varroa destructor*, in the honey bee (*Apis mellifera*). *Genome Biol.* 13, R81. doi: 10.1186/gb-2012-13-9-r81

- Rinderer, T. E., de Guzman, L. I., Delatte, G. T., Stelzer, J. A., Lancaster, V. A., Kuznetsov, V., et al. (2001). Resistance to the parasitic mite *Varroa destructor* in honey bees from far eastern Russia. *Apidologie* 32, 381–394. doi: 10.1051/apido:2001138
- Rosenkranz, P., Aumeier, P., and Ziegelmann, B. (2010). Biology and control of *Varroa destructor*. *J. Invertebr. Pathol.* 103, S96–S119. doi: 10.1016/j.jip.2009.07.016
- Sammataro, D., Gerson, U., and Needham, G. (2000). Parasitic mites of honey bees: life history, implications, and impact. *Annu. Rev. Entomol.* 45, 519–548. doi: 10.1146/annurev.ento.45.1.519
- Seeley, T. D. (2007). Honey bees of the Arnot Forest: a population of feral colonies persisting with *Varroa destructor* in the northeastern United States. *Apidologie* 38, 19–29. doi: 10.1051/apido:2006055
- Trost, B., Arsenault, R., Griebel, P., Napper, S., and Kusalik, A. (2013a). DAPPLE: a pipeline for the homology-based prediction of phosphorylation sites. *Bioinformatics* 29, 1693–1695. doi: 10.1093/bioinformatics/btt265
- Trost, B., Kindrachuk, J., Määttänen, P., Napper, S., and Kusalik, A. (2013b). PIKA 2: an expanded, web-based platform for analysis of Kinome microarray data. *PLoS ONE* 8:e80837. doi: 10.1371/journal.pone.0080837
- Trost, B., Kindrachuk, J., Scruten, E., Griebel, P., Kusalik, A., and Napper, S. (2013c). Kinotypes: stable species- and individual-specific profiles of cellular kinase activity. *BMC Genomics* 14:854. doi: 10.1186/1471-2164-14-854
- Tsuruda, J. M., Harris, J. W., Bourgeois, L., Danka, R. G., and Hunt, G. J. (2012). High-resolution linkage analyses to identify genes that influence *Varroa* sensitive hygiene behavior in honey bees. *PLoS ONE* 7:e48276. doi: 10.1371/journal.pone.0048276
- Vanengelsdorp, D., Evans, J. D., Saegerman, C., Mullin, C., Haubruge, E., Nguyen, B. K., et al. (2009). Colony collapse disorder: a descriptive study. *PLoS ONE* 4:e6481. doi: 10.1371/journal.pone.0006481
- Yang, X., and Cox-Foster, D. L. (2005). Impact of an ectoparasite on the immunity and pathology of an invertebrate: evidence for host immunosuppression and viral amplification. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7470–7475. doi: 10.1073/pnas.0501860102

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 February 2014; paper pending published: 02 April 2014; accepted: 28 April 2014; published online: 21 May 2014.

Citation: Robertson AJ, Trost B, Scruten E, Robertson T, Mostajeran M, Connor W, Kusalik A, Griebel P and Napper S (2014) Identification of developmentally-specific kinotypes and mechanisms of *Varroa* mite resistance through whole-organism, kinome analysis of honeybee. *Front. Genet.* 5:139. doi: 10.3389/fgene.2014.00139

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Robertson, Trost, Scruten, Robertson, Mostajeran, Connor, Kusalik, Griebel and Napper. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Turnover of protein phosphorylation evolving under stabilizing selection

Christian R. Landry<sup>1,2,3 \*</sup>, Luca Freschi<sup>1,2,3</sup>, Taraneh Zarin<sup>4</sup> and Alan M. Moses<sup>4,5,6</sup>

<sup>1</sup> Département de Biologie, Université Laval, Québec, QC, Canada

<sup>2</sup> Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

<sup>3</sup> Network for Research on Protein Function, Structure, and Engineering (PROTEO), Université Laval, Québec, QC, Canada

<sup>4</sup> Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada

<sup>5</sup> Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada

<sup>6</sup> Center for Analysis of Genome Evolution and Function, University of Toronto, Toronto, ON, Canada

## Edited by:

Andreas Zanzoni, Inserm TAGC  
UMR1090, France  
Allegria Via, Sapienza University, Italy

## Reviewed by:

Manuela Helmer-Citterich, University  
of Rome Tor Vergata, Italy  
Pedro Beltrao, European Molecular  
Biology Laboratory, European  
Bioinformatics Institute, UK

## \*Correspondence:

Christian R. Landry, Département de  
Biologie, Université Laval, 1030  
Avenue de la Médecine, Québec, QC  
G1V 0A6, Canada  
e-mail: christian.landry@bio.ulaval.ca

Most proteins are regulated by posttranslational modifications and changes in these modifications contribute to evolutionary changes as well as to human diseases. Phosphorylation of serines, threonines, and tyrosines are the most common modifications identified to date in eukaryotic proteomes. While the mode of action and the function of most phosphorylation sites remain unknown, functional studies have shown that phosphorylation affects protein stability, localization and ability to interact. Two broad modes of action have been described for protein phosphorylation. The first mode corresponds to the canonical and qualitative view whereby single phosphorylation sites act as molecular switches that either turn on or off specific protein functions through direct or allosteric effects. The second mode is more akin to a rheostat than a switch. In this case, a group of phosphorylation sites in a given protein region contributes collectively to the modification of the protein, irrespective of the precise position of individual sites, through an aggregate property. Here we discuss these two types of regulation and examine how they affect the rate and patterns of protein phosphorylation evolution. We describe how the evolution of clusters of phosphorylation sites can be studied under the framework of complex traits evolution and stabilizing selection.

**Keywords: protein phosphorylation, evolutionary turnover, molecular switches, molecular rheostats, protein evolution, molecular evolution, cell signaling**

## INTRODUCTION

The rate of discovery of new protein forms is increasing with the growing sensitivity of biochemical, analytical and bioinformatics tools (Smith et al., 2013). We now contemplate the idea that a large fraction of biological diversity originates in mechanisms that regulate protein expression and functions posttranscriptionally and posttranslationally. Among the major sources of posttranslational regulation and cellular complexity are posttranslational modifications (PTMs; Jensen, 2006), which are additions of peptides, chemical groups or other complex molecules to proteins that modify their activity, stability, degradation, localization, and ability to interact (Sprang et al., 1988; Madeo et al., 1998; Vazquez et al., 2000; Khmelinskii et al., 2009). Hundreds of such modifications have been reported in the literature and some of these appear to be playing dominant roles, at least in terms of occurrence. Protein phosphorylation on serines, threonines, and tyrosines dominates by an order of magnitude the number of experimental PTMs recorded in common databases (Khoury et al., 2011; Lu et al., 2013). This domination likely derives from biases resulting from the long historical interest for protein phosphorylation (Cohen, 1982), from the more advanced state of experimental identification methods for these PTMs and also for biological reasons, for instance because of the large number of protein kinases in eukaryotic genomes that can perform the necessary enzymatic

reactions. The impact of protein phosphorylation on the regulation and deregulation of protein functions in human diseases such as Alzheimer's disease (Grundke-Iqbal et al., 1986; Alonso et al., 1996), and the numerous gains and losses of phosphorylation in cancer cells (Reimand and Bader, 2013; Reimand et al., 2013) suggest that it plays a major role in proteome regulation and in complex cellular phenotypes. For these reasons, there has been much interest in the recent years for understanding how these PTMs evolve (Moses and Landry, 2010). However, the study of phosphorylation site evolution has met several difficulties that derive from the complex mapping between PTMs and protein functions.

## MOLECULAR MECHANISMS OF PROTEIN REGULATION BY PHOSPHORYLATION

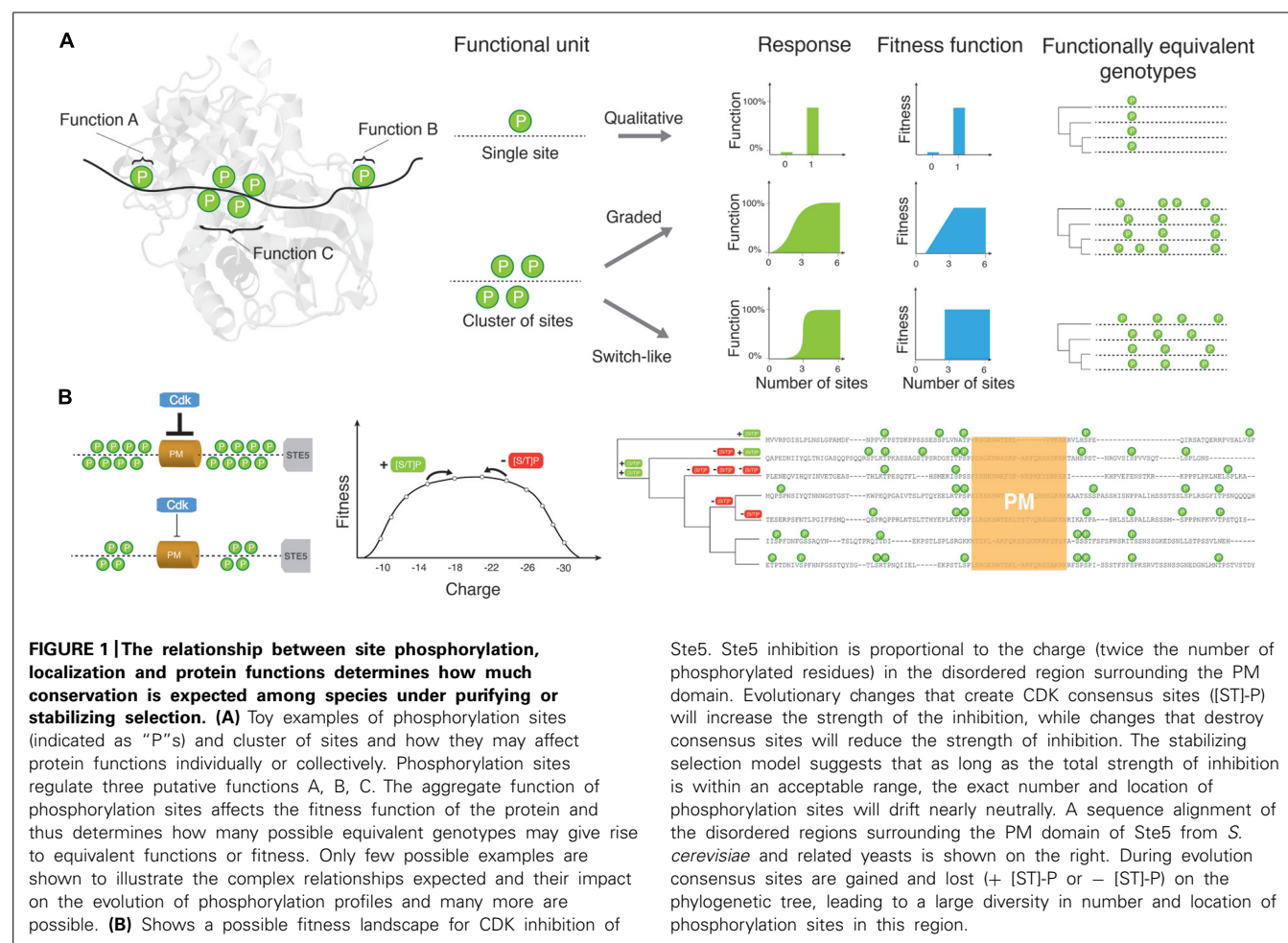
Phosphorylation sites play roles in allosteric and orthosteric regulation of proteins (Nussinov et al., 2012). Allosteric regulation acts through long-distances and involves a conformational change of the protein, while orthosteric regulation occurs at the active site of an enzyme or at the interface between a protein and another molecule. The study of protein phosphorylation has been historically centered on the role of individual phosphorylation sites. Indeed, a single phosphorylation site may have dramatic effects in regulating protein functions. For instance, the phosphorylation

of tyrosine 527 on the protein kinase and oncogene Src inactivates the protein through the interaction of this modified residue with its SH2 domain, which closes the kinase into an inactive conformation (Frame, 2002). Because single phosphorylation sites in specific cases play key roles in protein regulation, mutations at these sites may have complex organismal phenotypes. For instance, mutation of Ser47 on the *Drosophila* circadian clock protein PER modifies its interactions with other circadian proteins and lengthens adult locomotor activity from 24 to 31 h (Blau, 2008; Chiu et al., 2008). It has become clear that the simplistic view of one phosphorylation site – one function cannot be generalized to all phosphosites. Proteins are often multi phosphorylated and the different sites may affect each other's functions (Cohen, 2000).

Multisite phosphorylation is the premise of more complex types of regulation, for instance concerted regulation and modular regulation (reviewed in Salazar and Hofer, 2007). In the first case, all phosphorylation sites on a protein regulate one or more protein functions in a concerted manner (**Figure 1**). In the second case, groups of phosphorylation sites are organized into modules of multiple sites found in a short distance in a particular domain or disordered region of a protein and each cluster regulates a particular and independent function. Each of the mechanisms described above has its own complexity in terms of effects on the

protein and the dynamics of activation, and may thus affect the evolution of these sites (**Figure 1**). It is important to note that these mechanisms are not mutually exclusive. They are typically combined with various types of logic to encode patterns of substrate protein activity. For example, phosphorylation of some sites might “prime” (or increase the probability of phosphorylation) of other sites within the same region through the binding of other proteins that enhance the efficiency of the kinase (Cohen, 2000; Koivomagi et al., 2013; McGrath et al., 2013). In other cases, several sequential regulatory steps have been shown to be required for protein function (Yuan et al., 2002).

Multisite phosphorylation is associated with complex dynamic responses in signaling cascades. The number of sites being phosphorylated and their chronological order of modification can lead to graded or switch-like responses (**Figure 1**), depending on several parameters such as enzyme and substrate concentrations, binding parameters and the kinase/phosphatase processivity (Salazar and Hofer, 2007). One example of graded response involves the enhancement of p53 binding to CREB-binding protein (CBP). The p53 transactivation domain mediates the interaction with CBP and HDM2. Phosphorylation of Thr18 in this domain regulates the qualitative binding (on/off) of HDM2. However, the binding to CBP is regulated in a graded manner, with phosphorylation



events contributing additively to the binding energy of p53 to CBP (Lee et al., 2010). Accordingly, it is the sum of the effect that provides an appropriate function to the sites. Another recent example of aggregate effect comes from the circadian rhythm protein FRQ in *Neurospora*, which has more than 100 phosphorylation sites. It was recently demonstrated that the phosphorylation of FRQ is progressive and leads to a buildup of charge in one region of the protein that eventually leads to its degradation (Menet and Rosbash, 2011; Querfurth et al., 2011). A similar mechanism was found to control membrane association of Ste5 in the yeast mating pathway (Serber and Ferrell, 2007). The feature of interest here is that each phosphorylation site does not have a precise role but rather contributes to an aggregate property. The aggregate property was demonstrated experimentally in only few cases but clusters of phosphorylation sites are so abundant in proteomes that a large fraction of proteins could be regulated this way. Large-scale studies have indeed found that phosphorylation sites tend to localize in dense clusters of serines and threonines, which often tend to be phosphorylated by the same kinases (Moses et al., 2007a; Schweiger and Linial, 2010), supporting the hypothesis that a fraction of phosphorylation sites could regulate protein functions in a concerted way rather than acting as individual switches. However, these large-scale datasets need to be interpreted with caution. As it is the case for single phosphorylation sites (see below), clusters of phosphorylation sites could also appear as a result of non-functional phosphorylation. The tendency of phosphorylation sites to cluster could in this case result from the fact that regions accessible to protein kinases are not randomly distributed and form cluster of accessible sites that are susceptible to these phosphorylation events.

## EVOLUTION OF PHOSPHORYLATION SITES

The observation that a large fraction of proteins are posttranslationally modified led to the hypothesis that changes in PTM may contribute to a large fraction of phenotypic variation within species and divergence among species (Moses and Landry, 2010). Many studies have thus examined the conservation and divergence of PTMs, particularly phosphorylation sites. Given the diversity of molecular functions that can be regulated by phosphorylation sites either individually, or when they are found in a multi-site phosphorylated protein, it is expected that their patterns of evolution would also be diverse. For example, in the classical paradigm where a single phosphorylation event leads to a conformation change that alters enzyme activity (Cohen, 1982), one predicts that the phosphorylation site and its position are conserved by purifying selection, as long as the regulation of that enzyme activity is important. Indeed, some well-characterized examples are highly conserved sites over evolution (Rittenhouse et al., 1986; Landry et al., 2009; Tan et al., 2009; Beltrao et al., 2012). However, surprisingly, other well-characterized phosphorylation sites that regulate enzyme activity are not conserved (Hwang and Fletterick, 1986; Landry et al., 2009; Boulais et al., 2010; Nguyen Ba and Moses, 2010; Freschi et al., 2011, 2014; Levy et al., 2012) suggesting that regulation of those enzymes is not critical, or that other mechanisms exist that make phosphorylation sites and their positions unessential. While the phenotypic consequences of the changes in individual sites among species is still mostly unknown

– with a few spectacular exceptions (Lynch et al., 2011) – the estimates of the relative rate of phosphorylation site evolution based on large samples led to some challenging conclusions. Some studies concluded that phosphorylation sites are generally under strong evolutionary constraints, i.e., that they evolve much slower than non-modified residues (e.g., Gray and Kumar, 2011), while others estimated that the average constraint imposed on proteins by their phosphorylation is relatively weak (e.g., Landry et al., 2009; studies reviewed in Levy et al., 2012).

Beyond the fact that variation in the different methods and datasets may contribute to some of the disagreements in estimating phosphorylation site conservation, much of the debate comes from the fact that some authors focused their attention on cases where purifying selection is strong whereas others focused on cases where there is no or little purifying selection. On the first hand, in cases where phosphorylation is known to play an important role in regulating the protein, phosphorylation sites are often strongly conserved as predicted (Landry et al., 2009; Nguyen Ba and Moses, 2010; Beltrao et al., 2012). On the other hand, the more challenging observation is that large numbers of uncharacterized sites are poorly conserved among species. There are several possible reasons why sites would evolve quickly. The first is that databases reporting large-scale data on phosphorylation may be populated with a significant fraction of false-positive identifications, i.e., sites that are not actually phosphorylated in cells. Although this limitation contributes little to our understanding of the evolution of phosphorylation sites, it will be an important challenge to be addressed by investigators developing instruments and analytical tools. Another possible scenario is that the rate of evolution is elevated because a significant fraction of sites are species-specific, i.e., they have evolved only recently by directional selection and tests of conservation among species reject the hypothesis that they are under evolutionary constraint. While this hypothesis is of biological interest, there is currently very little data supporting this possibility (but see Jensen et al., 2006; Kim and Hahn, 2011; Lynch et al., 2011), at least not on a scale that would affect significantly the results of analyses performed on thousands of phosphorylation sites.

One other possible mechanism that would explain why many phosphorylation sites evolve as non-modified residues would be that they provide no function to the protein, i.e., they result from non-functional encounters between kinases and their substrates (Lienhard, 2008; Landry et al., 2009). Because kinases are highly processive enzymes and their recognition motifs are highly degenerate (Ubersax and Ferrell, 2007), many collisions between kinases and proteins may result in the inconsequential phosphorylation of residues. This cause is obviously extremely difficult to demonstrate because it is almost impossible to show that some trait or molecular feature has no function, as it is impossible to test all possible parameters that could reveal its role experimentally. However, there are lines of evidence that support this model. For instance, bona fide functional phosphorylation sites are more conserved than the ones for which no evidence is available (Landry et al., 2009; Nguyen Ba and Moses, 2010; Freschi et al., 2014). In addition, the relative rate of phosphorylation site conservation decreases with protein abundance and increases with the stoichiometry of phosphorylation, another observation consistent

with a model by which the prevalence of non-functional phosphorylation increases with protein abundance, due to the increased probability of encounters between kinases and other proteins (Levy et al., 2012). However, this observation is also consistent with a higher rate of false positive phosphorylation sites being identified in mass spectrometry studies for high abundance proteins. We note that two of these explanations (false positives, non-functional encounters) posit no or limited biological significance to poorly conserved phosphorylation sites, whereas species specific directional selection would indicate great biological significance for these sites. A final possible explanation for the elevated rate of evolution of phosphorylation site is the rapid turnover of sites caused by the weak constraint on their localization on the protein. This is the case that we consider further here, where rapidly evolving phosphorylation sites do have biological functions, but function is not strongly dependent on each individual site, as discussed above in Section “Molecular Mechanisms of Protein Regulation by Phosphorylation.”

### STABILIZING SELECTION ACTING ON CLUSTERS OF PHOSPHORYLATION SITES

The turnover of phosphorylation sites found in clusters of sites is an appealing hypothesis to explain at least a fraction of rapidly evolving phosphorylation sites (Landry et al., 2009). Indeed, a large fraction of phosphorylation sites are overrepresented in disordered regions of proteins where they occur in clusters of several juxtaposed sites (Moses et al., 2007a; Landry et al., 2009; Schweiger and Linial, 2010) and could act as groups rather than individual sites as described in Section “Molecular Mechanisms of Protein Regulation by Phosphorylation” (**Figure 1**). Sites could be gained and lost within a given region with limited effect on the overall function, but contributing to an increased rate of evolution if one considers sites individually. Different clusters could also play similar roles and diverging species could lose and gain these clusters without any significant change in function (Drury and Diffley, 2009). This model, which represents a previously undescribed form of “stabilizing selection” (Burger and Lande, 1994; Charlesworth, 2013a,b), suggests that clusters of sites rather than the sites themselves are the functional unit. This model leads to several predictions that can be tested using existing data: (i) protein functions regulated by phosphorylation should be preserved despite the fact that phosphorylation sites move positions; (ii) phosphorylated residues that occur in clusters should evolve faster because any change in position would be identified as an evolutionary modification; (iii) the number rather than the position of phosphorylation sites within clusters should be preserved over evolution; (iv) sites that are gained and lost between orthologs proteins should be gained and lost within these clusters. While there are no clear experimental data supporting most of these predictions, there are several observations that are consistent with them.

Observations supporting the first prediction come from proteins involved in DNA replication that are regulated by cyclin dependent kinases (CDKs). These proteins show conserved regulation in animals and fungi but diverge in the position and number of clusters of CDK phosphorylation sites, suggesting that functions may be preserved despite changes in phosphorylation

patterns (Moses et al., 2007b). Additional evidence comes from the large-scale study of CDK-dependent sites in budding yeasts where it was suggested that phosphorylation sites present in the model species but absent in closely related species could in fact be present in these other species but at other positions (Holt et al., 2009). The second prediction has been tested by (Nguyen Ba and Moses, 2010) who examined whether gains and losses of phosphorylation sites were more permissive in proteins with a high number of phosphorylation sites but no significant evidence was found against the null hypothesis. Evidence supporting predictions (iii) and (iv) comes from the analysis of yeast and mammalian phosphoproteomes. In the first case, it was shown that despite the low rate of phosphorylation site conservation (here the phosphorylation status of the proteins was compared between species or between paralogs), the actual number of phosphorylation sites tended to be maintained over evolution between homologous proteins (Freschi et al., 2011) or group of proteins (Beltrao et al., 2009). Similarly, orthologs of bona fide targets of protein kinases were found to contain conserved clusters of phosphorylation sites, often with little amino acid sequence conservation (Lai et al., 2012). Finally, it was shown that birth and death of phosphorylation sites in proteins tended to be clustered in space (Freschi et al., 2011, 2014), supporting the last prediction of the model.

More work will be needed along these lines to test whether the stabilizing selection model can apply to clusters of phosphorylation sites. Under this model, a phosphorylation cluster would be considered as a quantitative trait where selection does not act on a particular site but rather on the number or density of sites. In this model, each site represents a single locus that contributes to a “complex trait.” Deviation from the trait optimum is associated with a fitness cost that would increase or decrease following a given function (see, e.g., Charlesworth, 2013a) when phosphorylation sites are added or removed by mutations in this region (**Figure 1**). As mutation-selection-drift equilibrium is reached after a long divergence time, there will be a large diversity of “genotypes” that encode the trait. Even if the optimal trait value is the same in all species, if the number of phosphorylation sites (loci) that contribute to the trait is large, a minimal level of conservation will be observed for the actual phosphorylation sites. For example, CDK-mediated inhibition of Ste5 membrane localization is mediated by 8 phosphorylation sites in *S. cerevisiae* (**Figure 1**; Serber and Ferrell, 2007; Strickfaden et al., 2007). The strength of the inhibition is proportional to the net charge of the Ste5 PM domain after phosphorylation by CDKs. In evolutionary terms, the net charge of the PM domain is a quantitative trait whose value simply doubles the sum of the CDK phosphorylation sites in the PM domain. If a mutation occurs at a phosphorylation site, and renders it unphosphorylatable, the net charge of the PM domain will decrease by 2. On the other hand, if a mutation creates a new phosphorylation site (by adding a serine or threonine followed by a proline) in this region, the value of the trait will increase by 2. Assuming that there are several values of the trait that are compatible with functional inhibition by CDK, the PM domain will be free to drift in sequence space through any “genotype” that has enough phosphorylation sites to retain function. At mutation-selection-drift equilibrium, a large number of genotypes will coexist (**Figure 1**),



corresponding to different positions and numbers of phosphorylation sites in the PM domain. This stabilizing selection model has been applied to explain the rapid evolution of transcription factor binding sites in highly conserved developmental enhancers in *Drosophila* (Ludwig et al., 2000). In one recent simulation study, clusters of transcription factor binding sites were shown to evolve spontaneously simply because there are a much larger number of genotypes with many binding sites that encode the trait than “simple” genotypes with few binding sites (He et al., 2012). Whether this type of model can explain the origin of phosphorylation site clusters is currently a promising open research question.

## CONCLUSION

Given the current knowledge on the function of clusters of phosphorylation sites, there is a need for the development of new evolutionary models or adaptation of existing ones that take into account the fuzziness of phosphorylation site position and density. Although to our knowledge the stabilizing selection model has not been explicitly applied to the evolution of phosphorylation sites, the distribution of kinase site consensus matches in the absence of selection has been calculated (Lai et al., 2012). This represents a neutral null hypothesis for the evolution of phosphorylation site clusters. Deviations from this model were exploited to predict substrates for several protein kinases because orthologs of bona fide substrates contain more matches to the phosphorylation site consensus than expected in the absence of selection. Knowledge on kinase recognition motifs can thus be exploited to examine the neutral evolution of phosphorylation sites and thus to derive null hypothesis regarding their evolution.

The study of protein phosphorylation will also need more experimental studies on the function of cluster of phosphorylation sites in order to learn the general principles by which their aggregate properties emerge from their combination. This will allow for instance to estimate the relationship between the number of sites, their density and positions, and protein function and organismal fitness. Most evolutionary studies performed so far, with few exceptions (Landry et al., 2009), rely on comparative data among species and not on within species polymorphisms. Detailed analysis of within species variation coupled with functional studies could help estimate the distribution of genotypes under selection-mutation-drift equilibrium for given clusters of sites. Above all, the adaptation of more complex models to the evolution of PTMs will provide a better global picture of the evolutionary forces acting on phosphoproteomes. At the same time, these models will have the potential to contribute to biochemical studies whereby evolutionary observations will guide experimental investigators in studying the code of phosphorylation site regulation.

## ACKNOWLEDGMENTS

We thank Jukka-Pekka Verta, Marie Filteau, Guillaume Diss, Samuel Rochette, and Alexandre Dubé for comments on the manuscript. This work was supported by Canadian Institute of Health Research (CIHR) Grants GMX-191597 and GMX-299432, a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant and a Human Frontier Science Program grant to Christian R. Landry. Christian R. Landry is a CIHR

New Investigator. Luca Freschi was supported by fellowships from the Fonds de Recherche du Québec – Nature et Technologies (FRQ-NT) and the Quebec Research Network on Protein Function, Structure and Engineering (PROTEO). Alan M. Moses was supported by grants from CIHR (grant MOP-119579) and NSERC.

## AUTHOR CONTRIBUTIONS

Christian R. Landry drafted the manuscript with contributions from Alan M. Moses. Luca Freschi contributed concepts, discussion and to manuscript edition. Taraneh Zarin and Luca Freschi contributed the figure.

## REFERENCES

- Alonso, D. F., Farias, E. F., Famulari, A. L., Dominguez, R. O., Kohan, S., and De Lustig, E. S. (1996). Excessive urokinase-type plasminogen activator activity in the euglobulin fraction of patients with Alzheimer-type dementia. *J. Neurol. Sci.* 139, 83–88. doi: 10.1016/0022-510X(96)00036-6
- Beltrao, P., Albanese, V., Kenner, L. R., Swaney, D. L., Burlingame, A., Villen, J., et al. (2012). Systematic functional prioritization of protein posttranslational modifications. *Cell* 150, 413–425. doi: 10.1016/j.cell.2012.05.036
- Beltrao, P., Trinidad, J. C., Fiedler, D., Roguev, A., Lim, W. A., Shokat, K. M., et al. (2009). Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol.* 7:e1000134. doi: 10.1371/journal.pbio.1000134
- Blau, J. (2008). PERSpective on PER phosphorylation. *Genes Dev.* 22, 1737–1740. doi: 10.1101/gad.1696408
- Boulais, J., Trost, M., Landry, C. R., Dieckmann, R., Levy, E. D., Soldati, T., et al. (2010). Molecular characterization of the evolution of phagosomes. *Mol. Syst. Biol.* 6, 423. doi: 10.1038/msb.2010.80
- Burger, R., and Lande, R. (1994). On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics* 138, 901–912.
- Charlesworth, B. (2013a). Stabilizing selection, purifying selection, and mutational bias in finite populations. *Genetics* 194, 955–971. doi: 10.1534/genetics.113.151555
- Charlesworth, B. (2013b). Why we are not dead one hundred times over. *Evolution* 67, 3354–3361. doi: 10.1111/evo.12195
- Chiu, J. C., Vanselow, J. T., Kramer, A., and Edery, I. (2008). The phosphoryl occupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. *Genes Dev.* 22, 1758–1772. doi: 10.1101/gad.1682708
- Cohen, P. (1982). The role of protein phosphorylation in neural and hormonal control of cellular activity. *Nature* 296, 613–620. doi: 10.1038/296613a0
- Cohen, P. (2000). The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends Biochem. Sci.* 25, 596–601. doi: 10.1016/S0968-0004(00)01712-6
- Drury, L. S., and Diffley, J. F. (2009). Factors affecting the diversity of DNA replication licensing control in eukaryotes. *Curr. Biol.* 19, 530–535. doi: 10.1016/j.cub.2009.02.034
- Frame, M. C. (2002). Src in cancer: deregulation and consequences for cell behaviour. *Biochim. Biophys. Acta* 1602, 114–130. doi: 10.1016/S0304-419X(02)00040-9
- Freschi, L., Courcelles, M., Thibault, P., Michnick, S. W., and Landry, C. R. (2011). Phosphorylation network rewiring by gene duplication. *Mol. Syst. Biol.* 7, 504. doi: 10.1038/msb.2011.43
- Freschi, L., Osseni, M., and Landry, C. R. (2014). Functional divergence and evolutionary turnover in mammalian phosphoproteomes. *PLoS Genet.* 10:e1004062. doi: 10.1371/journal.pgen.1004062
- Gray, V. E., and Kumar, S. (2011). Rampant purifying selection conserves positions with posttranslational modifications in human proteins. *Mol. Biol. Evol.* 28, 1565–1568. doi: 10.1093/molbev/msr013
- Grundke-Iqbal, I., Iqbal, K., Tung, Y. C., Quinlan, M., Wisniewski, H. M., and Binder, L. I. (1986). Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc. Natl. Acad. Sci. U.S.A.* 83, 4913–4917. doi: 10.1073/pnas.83.13.4913

- He, X., Duque, T. S., and Sinha, S. (2012). Evolutionary origins of transcription factor binding site clusters. *Mol. Biol. Evol.* 29, 1059–1070. doi: 10.1093/molbev/msr277
- Holt, L. J., Tuch, B. B., Villen, J., Johnson, A. D., Gygi, S. P., and Morgan, D. O. (2009). Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* 325, 1682–1686. doi: 10.1126/Science.1172867
- Hwang, P. K., and Fletterick, R. J. (1986). Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases. *Nature* 324, 80–84. doi: 10.1038/324080a0
- Jensen, L. J., Jensen, T. S., De Lichtenberg, U., Brunak, S., and Bork, P. (2006). Co-evolution of transcriptional and post-translational cell-cycle regulation. *Nature* 443, 594–597. doi: 10.1038/nature05186
- Jensen, O. N. (2006). Interpreting the protein language using proteomics. *Nat. Rev. Mol. Cell. Biol.* 7, 391–403. doi: 10.1038/Nrm1939
- Khmelnitskii, A., Roostalu, J., Roque, H., Antony, C., and Schiebel, E. (2009). Phosphorylation-dependent protein interactions at the spindle midzone mediate cell cycle regulation of spindle elongation. *Dev. Cell* 17, 244–256. doi: 10.1016/j.Devcel.2009.06.011
- Khoury, G. A., Baliban, R. C., and Floudas, C. A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* 1. doi: 10.1038/srep00090
- Kim, D. S., and Hahn, Y. (2011). Identification of novel phosphorylation modification sites in human proteins that originated after the human-chimpanzee divergence. *Bioinformatics* 27, 2494–2501. doi: 10.1093/Bioinformatics/Btr426
- Koivomagi, M., Ord, M., Iofik, A., Valk, E., Venta, R., Faustova, I., et al. (2013). Multisite phosphorylation networks as signal processors for Cdk1. *Nat. Struct. Mol. Biol.* 20, 1415–1424. doi: 10.1038/nsmb.2706
- Lai, A. C., Nguyen Ba, A. N., and Moses, A. M. (2012). Predicting kinase substrates using conservation of local motif density. *Bioinformatics* 28, 962–969. doi: 10.1093/bioinformatics/bts060
- Landry, C. R., Levy, E. D., and Michnick, S. W. (2009). Weak functional constraints on phosphoproteomes. *Trends Genet.* 25, 193–197. doi: 10.1016/j.tig.2009.03.003
- Lee, C. W., Ferreon, J. C., Ferreon, A. C., Arai, M., and Wright, P. E. (2010). Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19290–19295. doi: 10.1073/pnas.1013078107
- Levy, E., Michnick, S., and Landry, C. R. (2012). Protein abundance is key to distinguish promiscuous from functional phosphorylation based on evolutionary information. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2594–2606. doi: 10.1098/rstb.2012.0078
- Lienhard, G. E. (2008). Non-functional phosphorylations? *Trends Biochem. Sci.* 33, 351–352. doi: 10.1016/j.tics.2008.05.004
- Lu, C. T., Huang, K. Y., Su, M. G., Lee, T. Y., Bretana, N. A., Chang, W. C., et al. (2013). DbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic Acids Res.* 41, D295–D305. doi: 10.1093/nar/gks1229
- Ludwig, M. Z., Bergman, C., Patel, N. H., and Kreitman, M. (2000). Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* 403, 564–567. doi: 10.1038/35000615
- Lynch, V. J., May, G., and Wagner, G. P. (2011). Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB. *Nature* 480, 383–386. doi: 10.1038/nature10595
- Madeo, F., Schlauer, J., Zischka, H., Mecke, D., and Frohlich, K. U. (1998). Tyrosine phosphorylation regulates cell cycle-dependent nuclear localization of Cdc48p. *Mol. Biol. Cell* 9, 131–141. doi: 10.1091/mbc.9.1.131
- McGrath, D. A., Balog, E. R., Koivomagi, M., Lucena, R., Mai, M. V., Hirschi, A., et al. (2013). Cks confers specificity to phosphorylation-dependent CDK signaling pathways. *Nat. Struct. Mol. Biol.* 20, 1407–1414. doi: 10.1038/nsmb.2707
- Menet, J. S., and Rosbash, M. (2011). A new twist on clock protein phosphorylation: a conformational change leads to protein degradation. *Mol. Cell* 43, 695–697. doi: 10.1016/j.molcel.2011.08.015
- Moses, A. M., Heriche, J. K., and Durbin, R. (2007a). Clustering of phosphorylation site recognition motifs can be exploited to predict the targets of cyclin-dependent kinase. *Genome Biol.* 8, R23. doi: 10.1186/gb-2007-8-2-r23
- Moses, A. M., Liku, M. E., Li, J. J., and Durbin, R. (2007b). Regulatory evolution in proteins by turnover and lineage-specific changes of cyclin-dependent kinase consensus sites. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17713–17718. doi: 10.1073/Pnas.0700997104
- Moses, A. M., and Landry, C. R. (2010). Moving from transcriptional to phospho-evolution: generalizing regulatory evolution? *Trends Genet.* 26, 462–467. doi: 10.1016/j.Tig.2010.08.002
- Nguyen Ba, A. N., and Moses, A. M. (2010). Evolution of characterized phosphorylation sites in budding yeast. *Mol. Biol. Evol.* 27, 2027–2037. doi: 10.1093/molbev/msq090
- Nussinov, R., Tsai, C. J., Xin, F., and Radivojac, P. (2012). Allosteric post-translational modification codes. *Trends Biochem. Sci.* 37, 447–455. doi: 10.1016/j.tics.2012.07.001
- Querfurth, C., Diernfellner, A. C., Gin, E., Malzahn, E., Hofer, T., and Brunner, M. (2011). Circadian conformational change of the Neurospora clock protein FREQUENCY triggered by clustered hyperphosphorylation of a basic domain. *Mol. Cell* 43, 713–722. doi: 10.1016/j.molcel.2011.06.033
- Reimand, J., and Bader, G. D. (2013). Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Mol. Syst. Biol.* 9, 637. doi: 10.1038/msb.2012.68
- Reimand, J., Wagih, O., and Bader, G. D. (2013). The mutational landscape of phosphorylation signaling in cancer. *Sci. Rep.* 3, 2651. doi: 10.1038/srep02651
- Rittenhouse, J., Harrsch, P. B., Kim, J. N., and Marcus, F. (1986). Amino acid sequence of the phosphorylation site of yeast (*Saccharomyces cerevisiae*) fructose-1,6-bisphosphatase. *J. Biol. Chem.* 261, 3939–3943.
- Salazar, C., and Hofer, T. (2007). Versatile regulation of multisite protein phosphorylation by the order of phosphate processing and protein-protein interactions. *FEBS J.* 274, 1046–1061. doi: 10.1111/j.1742-4658.2007.05653.x
- Schweiger, R., and Linial, M. (2010). Cooperativity within proximal phosphorylation sites is revealed from large-scale proteomics data. *Biol. Direct* 5, 6. doi: 10.1186/1745-6150-5-6
- Serber, Z., and Ferrell, J. E. Jr. (2007). Tuning bulk electrostatics to regulate protein function. *Cell* 128, 441–444. doi: 10.1016/j.cell.2007.01.018
- Smith, L. M., Kelleher, N. L., Linial, M., Goodlett, D., Langridge-Smith, P., Ah Goo, Y., et al. (2013). Proteoform: a single term describing protein complexity. *Nat. Methods* 10, 186–187. doi: 10.1038/nmeth.2369
- Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., et al. (1988). Structural changes in glycogen phosphorylase induced by phosphorylation. *Nature* 336, 215–221. doi: 10.1038/336215a0
- Strickfaden, S. C., Winters, M. J., Ben-Ari, G., Lamson, R. E., Tyers, M., and Pryciak, P. M. (2007). A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* 128, 519–531. doi: 10.1016/j.cell.2006.12.032
- Tan, C. S., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M. O., Jorgensen, C., et al. (2009). Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci. Signal.* 2, ra39. doi: 10.1126/scisignal.2000316
- Ubersax, J. A., and Ferrell, J. E. Jr. (2007). Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* 8, 530–541. doi: 10.1038/nrm2203
- Vazquez, F., Ramaswamy, S., Nakamura, N., and Sellers, W. R. (2000). Phosphorylation of the PTEN tail regulates protein stability and function. *Mol. Cell. Biol.* 20, 5010–5018. doi: 10.1128/MCB.20.14.5010-5018.2000
- Yuan, J., Eckerd, F., Bereiter-Hahn, J., Kurunci-Csacsko, E., Kaufmann, M., and Strebhardt, K. (2002). Cooperative phosphorylation including the activity of polo-like kinase 1 regulates the subcellular localization of cyclin B1. *Oncogene* 21, 8282–8292. doi: 10.1038/sj.onc.1206011

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 June 2014; accepted: 08 July 2014; published online: 23 July 2014.

Citation: Landry CR, Freschi L, Zarin T and Moses AM (2014) Turnover of protein phosphorylation evolving under stabilizing selection. *Front. Genet.* 5:245. doi: 10.3389/fgene.2014.00245

This article was submitted to Systems Biology, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Landry, Freschi, Zarin and Moses. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read,  
for greatest visibility



## COLLABORATIVE PEER-REVIEW

Designed to be rigorous  
– yet also collaborative,  
fair and constructive



## FAST PUBLICATION

Average 85 days from  
submission to publication  
(across all journals)



## COPYRIGHT TO AUTHORS

No limit to article  
distribution and re-use



## TRANSPARENT

Editors and reviewers  
acknowledged by name  
on published articles



## SUPPORT

By our Swiss-based  
editorial team



## IMPACT METRICS

Advanced metrics  
track your article's impact



## GLOBAL SPREAD

5'100'000+ monthly  
article views  
and downloads



## LOOP RESEARCH NETWORK

Our network  
increases readership  
for your article

## Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland  
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • [info@frontiersin.org](mailto:info@frontiersin.org)  
[www.frontiersin.org](http://www.frontiersin.org)

## Find us on

