Protein phosphorylation in bacterial signal transduction

Ahasanul Kobić, Lei Shi, Ana Bosković, Christophe Grangeasse, Damjan Franjević, Ivan Mijaković

A hasanul Kobic a, Lei Shi a, Ana Boskovic a, Christophe Grangeasse b, Damjan Franjevic c, Ivan Mijakovic a,a,

a Micals, AgroParisTech-INRA UMR 1319, Domaine de Vilvert, 78352 Jouy en Josas, France
b Institut de Biologie et Chimie des Protéines, CNRS-UCBL UMR 5086, 7 passage du Vercors, 69367 Lyon cedex 07, France
c Division of Biology, Faculty of Science, Zagreb University, Roosevelt trg 6, 10000 Zagreb, Croatia

A R T I C L E   I N F O

Article history:
Received 8 October 2010
Received in revised form 15 December 2010
Accepted 18 January 2011
Available online 22 January 2011

Keywords:
Protein phosphorylation
Protein kinase
Phosphoproteomics
Phosphorylation predictor
Signal transduction

A B S T R A C T

Background: Protein phosphorylation has emerged as one of the major post translational modifications in bacteria, involved in regulating a myriad of physiological processes. In a complex and dynamic system such as the bacterial cell, connectivity of its components accounts for a number of emergent properties. This article is part of a Special Issue entitled: Systems Biology of Microorganisms.

Scope of review: This review focuses on the implications of bacterial protein phosphorylation in cell signaling and regulation and highlights the connections and cross talk between various signaling pathways: bacterial two-component systems and serine/threonine kinases, but also the interference between phosphorylation and other post-translational modifications (methylation and acetylation).

Major conclusions: Recent technical developments in high accuracy mass spectrometry have profoundly transformed proteomics, and today exhaustive site-specific phosphoproteomes are available for a number of bacterial species. Nevertheless, prediction of phosphorylation sites remains the main guide for many researchers, so we discuss the characteristics, limits and advantages of available phosphorylation predictors.

General significance: The advent of quantitative phosphoproteomics has brought the field on the doorstep of systems biology, but a number of challenges remain before the bacterial phosphorylation networks can be efficiently modeled and their physiological role understood. This article is part of a Special Issue entitled: Systems Biology of Microorganisms.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Post-translational modifications are a ubiquitous means of rapidly and reversibly modifying the physico-chemical properties of a protein, triggering a number of possible consequences: change of enzyme activity, oligomerization state, interaction with other proteins, sub-cellular localization or half-life. Cells of multicellular organisms are known for their extensive networks of post-translational modifications, in which different modifications converge in signal integration [1]. By contrast, bacteria have often been erroneously considered as simple sacs of metabolites, optimized for fast growth and devoid of the regulatory network based on post-translational modifications. However, bacteria in their natural habitat often cope with the hostile environment, lack of nutrients and constant biochemical warfare against other microbes (and eukaryal hosts in case of pathogens), interspersed with very rare bursts of rapid growth under favorable conditions. Therefore, it is not surprising that bacteria also use an extensive network of post-translational modifications to transmit signals and to coordinate cellular functions.

Phosphorylation is arguably the most extensively studied post-translational modification in bacteria [2–4], but other modifications, such as N-glycosylation and O-glycosylation [5], methylation [6] and acetylation [6] are present as well, and should not be ignored. In fact, bacteria employ a wider variety of monosaccharides than Eukarya to glycosylate their proteins [5]. To date, protein glycosylation has not been related to bacterial signal transduction, but is recognized as an important pathogenicity determinant. In Pseudomonas aeruginosa, flagellin glycosylation plays a major role in virulence [7], and in Streptococcus parasanguinis and Streptococcus gordonii O-glycosylation of serine-rich adhesion proteins modulates the attachment to host cells [8]. As exemplified by the case study of lysine acetylation in Salmonella [9], bacteria can also acetylate key enzymes to tune metabolic fluxes. A number of enzymes in the central metabolism (for example enzymes controlling the glycolysis/gluconeogenesis switch and the glyoxylate shunt) are differentially acetylated in response to different carbon sources, accompanying changes in growth rate and metabolic fluxes. Similarly to numerous cases of cross-talk described in Eukarya, where different modifications occur on the same or neighboring residues in mutually exclusive fashion, a recent study on bacterial chemotaxis in Escherichia coli describes such interplay in the case of the two-component response regulator CheY [10]. The output of the response regulator is coupled to environmental signals via a rapid phosphorylation by the histidine kinase. On a much slower
time-scale it responds to acetylation, which transmits the overall
energetic state of the cell since the acetate donors (acetate and acetyl-
CoA) and the synthesis of the acetylating enzymes are dependent on
 cellular metabolism [10]. The two signaling pathways, transmitting
information at very different speeds, thus converge on the same
regulator. Remaining with the example of chemotaxis, protein
methylation also comes heavily into play, in a reciprocal regulation
with phosphorylation. Signal terminating phosphatase CheC interacts
with a receptor-modifying deamidase CheD to terminate the signal by
receptor dephosphorylation, deamidation and methylation in Thermoto-
ga maritima and Bacillus subtilis [11]. Receptor methylation is
involved in chemotaxis in all bacteria studied so far [12], even though
a homologous system has not been found in E. coli.

In this review we will focus on an overview of the recent advances
in the field of bacterial phosphorylation, highlighting the impact of
protein phosphorylation on cellular physiology, connectivity of the
phosphorylation networks, recent methods in quantitative phospho-
proteomics and bioinformatic tools for prediction of bacterial
phosphorylation sites. However, one should not forget that the
bacteria use a plethora of post-translational modifications in their
signaling, that may yet prove to be as rich and as versatile as the one in
eukaryal signaling.

2. Histidine and aspartate kinases

When it comes to protein phosphorylation, Bacteria and Eukarya
modify the tyrosine, serine and threonine residues [2]. However, bacteria
(and some plants) also possess the two component systems, which rely on histidine autophosphorylation of the sensory kinases
(the first components) and aspartate phosphorylation of the response
regulators (the second components). The basic mechanism of two
component system action is the specific signal recognition by the
sensory kinase, activation of its autokinase activity, phosphotransfer
to the cognate response regulator, which in turns triggers the cellular
response, most often by binding to DNA sequences regulating gene
transcription. Besides this canonical mode of functioning, bacteria
have evolved a number of variations, which have all been extensively
reviewed recently [13]. One notable variation is the Pseudomonas
aeruginosa response regulator GacA that transmits the signal exclusively by directly controlling the expression of two small
regulatory RNAs: rsmY and rsmZ. These sRNAs in turn affect the
transcript stability of target genes, accounting for the entire GacA regulon [14]. In their role of sensing the immediate environment, two
component sensory kinases are often functionally coupled to the
transmembrane transporters that serve as co-sensors [15]. The key
feature of two component systems is the high fidelity of recognition
between the kinase and response regulator (almost no cross-talk with
other two-component systems) during phosphotransfer [16]. An
emerging feature in the field are the auxiliary regulators (distinct
from sensory kinases and response regulators), that interfere with
these highly specific phosphotransfer reactions [17]. They act directly
on the sensory kinases, or the response regulators, and their purpose
is to connect the given two-component system to other signaling
pathways, i.e. accomplish signal integration. This notion of signal
integration and connectivity has prompted some authors to use terms
such as “cognition” and “knowledge” in association with bacterial
signaling networks based on two-component systems [18], which
Despite the obvious semantic arguments, does underline their
complexity. The kinetics and feedback (signal extinction) of two
component systems are far from fully understood. Some two
component systems function like rheostats, others have adopted the
“all or nothing” approach (with a threshold level for activation) and
some function in the so-called mixed mode [19]. An interesting recent
study compared these output modes with the known kinetic
parameters of the histidine kinases [19]. A stochastic kinetic model
of two component signaling, based on published data, allowed the
authors to conclude that two-component systems positively auto-
regulated at the transcription level exhibit “all or nothing” responses, whereas those without a positive feedback loop exhibit graded and
mixed mode responses. In the latter cases, the graded response
changes to mixed mode by an increase of the translation initiation rate
of the histidine kinase. There are exceptions to this rule, for example some positively autoregulated systems show graded behavior [20].
This in fact demonstrates that two-component system overall design
can allow the bacterium to choose between deterministic regulation
and stochastic switching using the same pathway architecture. The
level of cross-talk between two-component systems is either
nonexistent or very low, and bacteria actively seek to minimize it
[21,22]. But when detected, is the cross talk just a noise, imperfect
insulation, or does it have biological consequences? B. subtilis two
component system WalRK (previously known as YycFG) coordinates the
cell wall synthesis with the cell cycle [23], but it is not fully insulated. The normal trigger for the response regulator WalR is the
septal localization of the kinase WalK. Under the specific conditions
of phosphate limitation, the cell must go into induced stationary phase,
concomitant to activation of phosphate scavenging functions. There-
fore, the phosphate starvation master regulator, the sensory kinase
PhoR, bypasses the WalK input and phosphotransferates directly WalR,
thus relaying the signal to cease cell wall synthesis [24]. There is also
accumulating evidence of cross-talk between two-component sys-
tems and serine–threonine kinases. The two signaling pathways can
converge on the same transcription regulator, as exemplified by the regulator MrpC in M. xanthus [25]. Alternatively, the serine/threonine
kinase can directly phosphorylate the response regulator eliciting a
direct response. Such is the case of the response regulator CovR from
Streptococcus agalactiae [26], RitR from Streptococcus pneumoniae [27],
and DosR from Mycobacterium tuberculosis [28]. Besides the two
component systems, bacteria possess a number of proteins of the
phosphoenol-pyruvate phosphotransferase system (PTS), which are
also transiently phosphorylated on histidine residues. The PTS
consists of a cascade of phosphotransferases that transmit the
phosphatase from phosphoenol pyruvate to the sugar that is being
imported. The level of histidine phosphorylation in the cascade
depends on the flux of incoming sugars and participates actively in
many facets of cellular regulation [29]. In the study of protein
phosphorylation at the global level, modern tools of site-specific
 phosphoproteomics have become indispensable for detection of
phosphorylated serines, threonines and tyrosines [30]. However,
these methods do not allow the routine detection of phospho-
histidines and phospho-aspartates, due to their inherent instability
under acidic conditions used in the experimental pipeline. Neverthe-
less, advanced transcriptomics tools, developments in FRET, immu-
nodetection, bioinformatic analyses and other technologies hold
promise of rapid advances in two-component systems research, as
outlined in this methodological review [31].

3. Serine and threonine kinases

For serine and threonine phosphorylation, Bacteria mostly use the
so-called Hanks type serine/threonine kinases, but also some atypical
bi-functional kinases/phosphorylases. For example, the Hpr kinase/
phosphorylase [32] that mediates the CcpA-dependent carbon
catabolite regulation in Firmicutes phosphorylates the small regula-

tory protein HPr. In E. coli, the evolutionarily unrelated isocitryate
dehydrogenase kinase/phosphatase [33] participates in regulating the
glyoxylate shunt by phosphorylating the isocytrate dehydrogenase. Nevertheless, most bacterial serine/threonine kinases belong to
Hanks kinases [2], a protein family that seems to be deeply rooted
in the tree of life, spanning all kingdoms. Despite the fact that Hanks
kinases are present in eukaryal and bacterial genomes alike, researchers often refer to the bacterial ones as “eukaryotic type
kinases” [34]. This is probably due to historical reasons: Hanks kinases
were first discovered in Eukarya and later in Bacteria. Since it is very unlikely that the genes encoding these kinases arrived to almost all bacterial genomes by horizontal transfer from Eukarya, we would like to argue here that the name “eukaryotic type kinases” is somewhat misleading. Hanks kinases have received ample attention in bacterial pathogens, where they were often shown to participate in virulence functions [35]. Due to their homology to eukaryal kinases, bacterial Hanks kinases (and associated phosphatases) are often employed in the biochemical warfare, scrambling the host defenses via phosphorylation (or dephosphorylation) of host signaling proteins. Probably the best studied bacterial system in this respect is M. tuberculosis, where the structure/function relationship of Hanks kinases, and their role as signal receptors has been probed extensively [36,37]. Apart from their involvement in virulence functions, bacterial Hanks kinases are also involved in housekeeping functions: threonine kinases have recently been shown to regulate bacterial protein secretion [38,39] and a regulation of the cell shape via coordination of envelope synthesis and cell division was described in Actinobacteria [40,41].

Despite a mountain of evidence on Hanks-kinase downstream actions, phosphorylation of protein substrates and consequent consequences for the bacterial physiology, we know precious little of the upstream part of the cascade, the ligand binding. In B. subtilis, the Hanks kinase PrkC has been shown to sense muropeptides of the cell wall and to control spore germination [42]. A recent structural study of the M. tuberculosis PrkC homolog, PknB prompted the authors to propose a mechanism whereby muropeptides binding by the extracellular PASTA domains of the kinase leads to dimerization and subsequent activation of the enzyme [43]. Appealing as the hypothesis may be in its similarity to the activation of Hanks kinases in Eukarya, further structure-functional analysis will be necessary to substantiate it. In parallel, first attempts at classifying bacterial Hanks kinases into sub-families are under way [44]. This is a particularly important task, having in mind that such classification was precious in constructing sub-family-specific phosphorylation predictors for Eukarya, where the choice of substrates seems to be constrained within kinase sub-families. In parallel to global phosphoproteomic approaches (described in Section 5), there are interesting alternative initiatives for screening Hanks kinase substrates in vivo. Molle et al. [45] have proposed a pipeline comprising of co-expression of kinase and substrate genes in E. coli, followed by identification of substrate phosphorylation by mass spectrometry. This approach for pair wise testing of kinase–substrate relations certainly has advantages over in vitro phosphorylation studies, and can be applied also for other types of kinases.

4. Tyrosine kinases

When it comes to protein-tyrosine phosphorylation, bacteria opted against using kinases of the Hanks family. Very few cases of Hanks kinase phosphorylating tyrosines have been found in bacteria [46,47]. Canonical bacterial protein–tyrosine kinases (BY-kinases) [48] are unique enzymes that exploit the ATP/GTP-binding Walker motif to catalyze phosphorylation of tyrosine residues. BY-kinases seem widespread, if not ubiquitous, in the bacterial kingdom [49], but are limited to only a few copies (typically one or two) per bacterial genome. Initially, BY-kinases were considered to be only autophosphorylating enzymes, involved in exopolysaccharide production, since they are usually encoded by genes in the large operons involved in biosynthesis and export of sugar polymers [50]. In fact, interaction of BY-kinases with the polysaccharide translocation machinery has been clearly established in several species, and kinase autophosphorylation was seen to have an impact on the process [51]. However, it rapidly became apparent that BY-kinases can also phosphorylate endogeneous substrates, leading to enzyme activation in the case of UDP-glucose dehydrogenases (in E. coli and B. subtilis) [52,53] and single-stranded DNA-binding proteins (in B. subtilis) [54], or inactivation, in case of phage integrases in E. coli [55]. BY-kinases are involved in pathogenicity functions via capsular polysaccharide synthesis [56,57], but also figure prominently in house-keeping functions, such as cell cycle control [58] or heat shock response [59]. Interestingly, activity changes are not the only consequences of tyrosine phosphorylation: B. subtilis BY-kinase PtkA can alter the sub-cellular localization of its proteins targets [60]. For some substrates this was related to kinase–substrate co-localization, and for other substrates, the localization pattern was dependent on phosphorylation and not interaction with the kinase itself. Recent structural studies gave important insights into oligomeric structures of transmembrane BY-kinases [61–63], suggesting autophosphorylation as the mechanism for oligomer dissociation [62]. Dissociation of BY-kinase oligomers upon autophosphorylation could arguably represent a step in signal transduction. However, the upstream part of the cascade is still obscure. Despite having a large extracellular loop, to this date no extracellular ligands have been assigned to BY-kinases, and this remains the weakest point in our knowledge concerning these enzymes. BY-kinases account for the vast majority of phosphorylated tyrosine residues in bacteria, but there are a few exceptions (unusual tyrosine-phosphorylating sensory kinases [64], tyrosine-phosphorylating Hanks kinases [46,47]), most notable of which is the kinase McsB from B. subtilis, homologous to arginine-phosphotransferases. This enzyme was found to tyrosine-phosphorylate the repressor of heat-shock genes, CtsR [65]. However, in 2009, McsB had its status revised to that of an arginine kinase [66], and it is presently not entirely clear whether the protein harbors both the activities, or just the latter one. Currently, the field of tyrosine kinases, which is the youngest one in bacteria, faces the challenge of assigning the kinases to their physiological substrates. Unlike two-component systems, where kinase and response regulator genes are often adjacent on the chromosome, tyrosine kinase and substrate genes rarely exhibit co-localization. Moreover, BY-kinases in general are much more promiscuous than histidine kinases, phosphorylating up to a dozen cognate proteins, as shown in the case of B. subtilis BY-kinase PtkA [60]. Finally, this formidable task is greatly facilitated by advanced quantitative phosphoproteomics approaches, discussed in Section 5.

5. Phosphoproteomics

Traditional phosphoproteomics, employed in bacteriology before 2007, relied on 2D-gel separation of bacterial proteomes, 32P-radiolabelling or immunodetection, and detection of phosphoproteins by low-resolution mass spectrometry, which precluded the identification of phosphorylation sites. These studies will not be reviewed here, since they offered little beyond the argument that many bacterial proteins are indeed phosphorylated. Starting from 2007, high-resolution mass spectrometry coupled to gel-free analysis lead to the publication of site-specific serine/threonine/tyrosine bacterial phosphoproteomes for a number of species: B. subtilis [67], E. coli [68], Lactococcus lactis [69], Klebsiella pneumoniae [70], Pseudomonas spp [71], S. pneumoniae [72], Mycoplasma pneumoniae [73], Streptomyces coelicolor [74], and M. tuberculosis [75]. On average, in each bacterial phosphoproteome around one hundred phosphorylation sites were reported. A notable exception is S. coelicolor [76], where the site yield was smaller (under 50 sites), probably due to a less powerful mass spectrometry analysis. In M. tuberculosis [75], the yield was significantly higher (around 500 sites), possibly due to the fact that this bacterium has a very large serine/threonine kinase complement. The phosphorylated proteins have versatile functions. The single largest sub-set of phosphorylated proteins is the enzymes involved in the central carbon metabolism. Other housekeeping proteins are also phosphorylated: helicases, chaperones, ribosomal proteins, aminoacyl tRNA-synthetases, etc. Proteins involved in signaling are over-represented in published phosphoproteomes, which also comprise a
culture (SILAC) in a case study of through the use of stable isotope labeling with amino acids in cell phosphoproteomics. For an extensive review of these techniques accomplished, we will be able to construct the phosphatases and their substrates. Once this work has been new analytical tool, kinase and phosphatase knockouts can be bacterial phosphoproteomes (and proteomes in general). With this expect the publication of a new wave of site-specific site occupancy are the obvious next step in the field of metabolism, conservation at the phosphorylation site level is in most species, such as the enzymes of the central carbon metabolism, conservation at the phosphorylation site level is nonexistent. Is serine/threonine/tyrosine phosphorylation so recent in bacteria, or does it evolve rapidly (with kinase relaxed specificities) to adapt to different environmental niches? Alternatively, are the published phosphoproteomes incomplete, are we missing many low-occupancy sites due to technical imperfections of our analysis? A recent comprehensive analysis of phosphorylated proteins of E. coli ribosomes showed that 24 ribosomal proteins are phosphorylated [77], whereas only a small number of them were detected in the E. coli phospho-proteome published only one year before [68]. Overall, the jury is still out, but a very simple experiment can show that the published phosphoproteomes are certainly incomplete. It suffices to change one parameter in the growth conditions, and in the same bacterium new sites emerge, and some of the old ones are lost, as can be seen by comparing the B. subtilis phosphoproteomes in [67] and [76]. Alternatively, if a substrate is mixed with its cognate kinase in vitro, in the presence of ATP, and subsequently analyzed by mass spectrometry, many new phosphorylation sites will appear; for example compare tyrosine phosphorylation sites in the phosphoproteome [67] to an in vitro study [60]. The biological significance of many low-occupancy phosphorylation sites is questioned by bacteriologists. Do they all have a biological role, or are they the consequence of relaxed kinase specificity, resulting in one major site and several random secondary sites on every protein substrate? Even if they are secondary, what is the optimization criterion for the bacterial cell to maintain kinases with relaxed specificity? These are important questions that we will have to tackle in order to understand bacterial signaling based on protein phosphorylation. In any case, phosphorylation is undoubtedly a dynamic event, and experiments in multiple conditions with quantitative comparisons of phosphorylation site occupancy are the obvious next step in the field of phosphoproteomics. For an extensive review of these techniques see [78]. Feasibility of such approaches has been demonstrated through the use of stable isotope labeling with amino acids in cell culture (SILAC) in a case study of B. subtilis [76], and we can soon expect the publication of a new wave of site-specific and quantitative bacterial phosphoproteomes (and proteomes in general). With this new analytical tool, kinase and phosphatase knockouts can be compared to wild type strains to infer connections between kinases/ phosphatases and their substrates. Once this work has been accomplished, we will be able to construct the first topology models of bacterial phosphorylation networks, as a stepping stone towards systems biology in this particular field.

### 6. Phosphorylation predictors

Despite the recent technological advances in phosphoproteomics, a microbiologist working on a particular bacterial species whose phosphoproteome has not been published is not much closer to having a list of phosphorylation sites in their organism of interest. The mass spectrometry equipment involved in phosphoproteomics analyses largely surpasses standard laboratory budgets, and data acquisition and analysis usually call for expert help. Therefore, many bacteriologists are likely to use site-specific phosphorylation predictors, based on the growing phosphoproteomics datasets, to provide accurate predictions of phosphorylation events. Today, two major phosphorylation predictors conceived to predict bacterial phosphorylation sites are available on the WWW: NetPhosBac (http://www.cbs.dtu.dk/services/NetPhosBac-1.0/) [79] and Disphos (http://core.ist.temple.edu/pred/pred.html) [80]. NetPhosBac is a neural network based predictor, trained on datasets of serine/threonine phosphorylation sites in E. coli and B. subtilis. Disphos is developed on around the concept that looking for phosphorylation sites means looking for stretches of amino acids in proteins structures that are inherently disordered and accessible to solvent (i.e. kinases). It predicts serine, threonine and tyrosine, phosphorylation sites, and can be run in a bacterial-specific mode. Many readers will ask if these phosphorylation predictors can be of any help, and since we are now in possession of over 1000 bacterial phosphorylation sites, we decided to benchmark them. A predictor gives a probability value for a given residue to be phosphorylated, ranging from zero to one. Sites that score above 0.5 are likely to be phosphorylated and the ones scoring close to 1 are almost certain to be. We tested the performance of NetPhosBac and Disphos on the published datasets of phosphorylated proteins from four bacteria: P. putida [71], P. aeruginosa [71], K. pneumoniae [70] and L. lactis [69]. Phosphorylation sites scoring above 0.5 were retained as correct predictions, and the result is shown in Fig. 1. On average, NetPhosBac accurately predicted only 25% of the experimentally determined phosphorylation sites, whereas Disphos on average predicted only under 5%. These results are certainly disappointing from the point of view of a bacteriologist who would use these predictors to guide further experiments. The difference in utility of the two predictors is not as large when one takes into consideration that many NetPhosBac predictions are just above the 0.5 cut-off, and would therefore be of limited value. On the other hand, Disphos predictions are much more discerning, the true positives always

---

**Fig. 1.** NetPhosBac and Disphos were used to predict phosphorylation sites on proteins from four bacterial species: P. putida [71], P. aeruginosa [71], K. pneumoniae [70] and L. lactis [69], that were experimentally found to be phosphorylated. Experimentally determined phosphorylation sites scoring above 0.5 were retained as correct predictions, and the result for each predictor and each species is expressed as percentage of correct predictions in the entire dataset of phosphorylated proteins.
7. Concluding remarks

It is customary to finish off such review articles stating that the field is expanding exponentially (which it is) and the future is ripe with discoveries to come. However, steady exponential progress is present in most fields of biology, and real excitement comes from quantum leaps. So let us instead express hope that new thinkers will be challenged by the many mysteries still hidden in the very simple bacterial cell, and they will join our efforts to unravel some of them.

Acknowledgements

This work was supported by grants from the Institut National de Recherche Agronomique (INRA) to IM and Agence Nationale de la Recherche (ANR) to IM and CG.

References


