

RESEARCH ARTICLE

Open Access



# Site-specific His/Asp phosphoproteomic analysis of prokaryotes reveals putative targets for drug resistance

Shu-Jung Lai<sup>1</sup>, I-Fan Tu<sup>1</sup>, Wan-Ling Wu<sup>1</sup>, Jih-Tian Yang<sup>2</sup>, Louis Y. P. Luk<sup>3</sup>, Mei-Chin Lai<sup>2,4</sup>, Yu-Hsuan Tsai<sup>3</sup> and Shih-Hsiung Wu<sup>1,5,6\*</sup>

## Abstract

**Background:** Phosphorylation of amino acid residues on proteins is an important and common post-translational modification in both eukaryotes and prokaryotes. Most research work has been focused on phosphorylation of serine, threonine or tyrosine residues, whereas phosphorylation of other amino acids are significantly less clear due to the controversy on their stability under standard bioanalytical conditions.

**Results:** Here we applied a shotgun strategy to analyze the histidine and aspartate phosphorylations in different microbes. Our results collectively indicate that histidine and aspartate phosphorylations frequently occur also in proteins that are not part of the two-component systems. Noticeably, a number of the modified proteins are pathogenesis-related or essential for survival in host. These include the zinc ion periplasmic transporter ZnuA in *Acinetobacter baumannii* SK17, the multidrug and toxic compound extrusion (MATE) channel YeeO in *Klebsiella pneumoniae* NTUH-K2044, branched amino acid transporter AzlC in *Vibrio vulnificus* and the RNA-modifying pseudouridine synthase in *Helicobacter pylori*.

**Conclusions:** In summary, histidine and aspartate phosphorylation is likely to be ubiquitous and to take place in proteins of various functions. This work also sheds light into how these functionally important proteins and potential drug targets might be regulated at a post-translational level.

**Keywords:** Proteomics, Post-translational modification, Histidine phosphorylation, Aspartate phosphorylation, Pathogenic bacteria, Drug resistance

## Background

Protein phosphorylation is a ubiquitous chemical event found in both eukaryotes and prokaryotes. Regulated by kinases and phosphatases, this type of post-translational modification is highly dynamic, changes protein function in the due course of the cell cycle and plays critical roles in signal transduction. Due to its associations with many human disorders, the molecular networks of mammalian protein phosphorylation have been extensively investigated [1]. On the other hand, protein phosphorylation also plays critical roles in bacterial pathogenesis [2–4]. Since there are numerous unmet clinical needs caused by

bacterial infections, it is of fundamental importance to fully expand the research of bacterial phosphoproteomics which likely holds valuable information for future drug development.

In bacteria, various amino acid residues within a protein, including serine, threonine, tyrosine, histidine and aspartate, can be modified with a phosphate group [5, 6]. A plethora of data related to serine, threonine and tyrosine phosphorylations have already been reported. They are shown to be involved in changing metabolic behaviors, inducing capsule formation and initiating sporulation [5]. On the other hand, the knowledge of histidine and aspartate phosphorylations are mostly limited to the two-component systems [7], in which phosphor-relaying signaling events were induced by external environmental stimuli, including

\* Correspondence: shwu@gate.sinica.edu.tw

<sup>1</sup>Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

<sup>5</sup>Chemical Biology and Molecular Biophysics Program, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan

Full list of author information is available at the end of the article



changes in pH, osmolality and oxygen, all of which are essential for the bacterial pathogenesis in human. Nevertheless, little is known about any histidine and aspartate phosphorylation events outside of the two-component systems [8].

Protein modification status is often investigated by gel electrophoresis coupled with liquid chromatography (LC) and high-resolution mass spectrometry analysis [9]. As useful biochemical information can be revealed only when the covalently modified residues are isolated and characterized experimentally, the proteomic data available for each type of phosphorylated residues is directly proportional to its chemical stability. Among various types of protein phosphorylations, the stability of acid-labile phosphorylations under standard mass spectrometric conditions remains somewhat controversial. Jensen and co-workers investigated the stability of a peptide containing a phosphorylated histidine and concluded that a fast LC method or non-acidic solvent system is needed for well-resolved electrospray ionization mass spectrometry analysis [10]. In contrast, a study by Hohenester et al. showed that typical proteomic methods that use acidic solvent systems and collision-induced dissociation for fragmentation can be applied [11]. These differences may originate from the fact that peptides with different sequences were used [10, 11]. On the other hand, the analysis of peptides containing a phosphorylated aspartate has not been conducted. In essence, the possibility to study acid-labile phosphorylation by standard mass spectrometry methods has not been ruled out.

Here, we aim to expand the horizon of bacterial phosphoproteomic research and present mass spectrometric analysis of histidine and aspartate phosphorylations in different key prokaryotes at the exponential phase, including pathogenic bacteria (*Acinetobacter baumannii* SK17, *Helicobacter pylori*, *Klebsiella pneumoniae* NTUH-K2044, *Vibrio vulnificus*), cyanobacterium (*Arthrospira platensis* C1), thermophilic bacteria (*Meiothermus taiwanensis* WR220, *Thermus thermophilus* HB27) and methanogenic archaea (*Methanosarcina mazei* N2 M9705, *Methanohalophilus portucalensis* FDF1<sup>T</sup>) (Table 1). Our results showed that, under optimal growing conditions, numerous proteins beyond the two-component systems are phosphorylated at the aspartate and histidine residues. Notably, many proteins that contain aspartate and histidine phosphorylations are involved in cell survival and antibiotic resistance.

## Methods

### Bacterial strains and growth conditions

The isolate of cyanobacterium, *A. platensis* C1, was a kind gift from Prof. Apiradee Hongsthong (National Center for Genetic Engineering and Biotechnology, Thailand). Axenic cultures of this bacterium were grown at 35 °C under

**Table 1** Characterized unique pHis/pAsp sites and phosphopeptides in the nine organisms

	pHis	pAsp	p-peptides
<i>A. baumannii</i>	31	15	40
<i>H. pylori</i>	22	6	24
<i>K. pneumoniae</i>	7	0	7
<i>V. vulnificus</i>	30	18	40
<i>A. platensis</i>	12	2	14
<i>M. taiwanensis</i>	15	11	21
<i>T. thermophilus</i>	11	3	11
<i>M. mazei</i>	18	9	25
<i>M. portucalensis</i>	13	5	15
Total	159	69	197

illumination by 100  $\mu\text{E m}^{-2} \text{s}^{-1}$  fluorescent light with continuous stirring in 1.5 L of Zarrouk's medium [12]. The cultures were grown to the mid-exponential phase ( $\text{OD}_{560 \text{ nm}} = 0.4$ ).

As previously described [13–15], we cultured and analyzed the phosphoproteome of *A. baumannii* SK17 [15], *K. pneumoniae* NTUH-K2044 [14] and *T. thermophilus* HB27 [13], where were kind gifts from Dr. Te-Li Chen (Taipei Veterans General Hospital, Taiwan), Prof. Jin-Town Wang (National Taiwan University, Taiwan) and Prof. Guang-Huey Lin (Tzu-Chi University, Taiwan), respectively. *A. baumannii* SK17 and *K. pneumoniae* used in this study were originally isolated from human patients [16, 17].

*H. pylori* reference strain 26695 (ATCC 700392) was obtained from the Food Industry Research and Development Institute, Taiwan (BCRC 17219). Cells were grown at 37 °C under a standard microaerobic atmosphere (5% O<sub>2</sub>, 10% CO<sub>2</sub>, 85% N<sub>2</sub>) on Columbia agar base (CAB; Oxoid) containing 10% horse blood for 24 h.

*V. vulnificus* was obtained from the Food Industry Research and Development Institute, Taiwan (BCRC 12B0001). A single clone of this bacterium was grown in LB medium at 37 °C with vigorous shaking. The overnight culture was diluted to  $\text{OD}_{600 \text{ nm}} = 0.01$  into fresh LB medium contained 2% NaCl and 10 nM FeSO<sub>4</sub>. The culture was grown to the mid-exponential phase ( $\text{OD}_{600 \text{ nm}} = 0.6$ ).

*M. taiwanensis* WR220 was obtained from the Food Industry Research and Development Institute, Taiwan (BCRC 17171). The thermophilic bacterium was grown under aerobic conditions at 55 °C in TM medium [18] to the mid-exponential phase ( $\text{OD}_{600 \text{ nm}} = 0.8$ ).

For methanogenic archaea, *M. portucalensis* FDF1<sup>T</sup> was purchased from the Leibniz Institute DSMZ—German Collection of Microorganisms and Cell Cultures (DSM No. 7471), and *M. mazei* N2 M9705 was originally isolated from an aquaculture fishpond near Wang-gong,

Taiwan [19] and has been deposited with the Food Industry Research and Development Institute, Taiwan (BCRC 16179). *M. portucalensis* FDF1<sup>T</sup> was cultured in defined medium containing 120 g L<sup>-1</sup> NaCl and 20 mM trimethylamine as the sole carbon and energy source [20], whereas *M. mazei* N2 M9705 were cultured in MB medium containing 5 g L<sup>-1</sup> NaCl and 20 mM of methanol. Sterile medium was prepared under 20% CO<sub>2</sub> in N<sub>2</sub> atmosphere by a modification of the Hungate technique [21]. The growth rates were monitored by removing 1 mL of the culture with a N<sub>2</sub>-flushed syringe into a Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> containing cuvette [22]. The cultures were grown anaerobically at 37 °C to the mid-exponential phase (OD<sub>540 nm</sub> = 0.5).

#### Protein extraction

Cells at the mid-exponential phase were harvested by centrifugation at 6000 g for 15 min at 4 °C and washed twice with PBS. The resulting pellets were resuspended in freshly prepared lysis buffer containing 25 mM ammonium bicarbonate, phosphatase inhibitor (PhosSTOP, Roche), 6 M urea and 2 M thiourea. Cells were disrupted by sonication on ice. Cellular debris were removed by centrifugation at 12,000 g for 30 min at 4 °C. The supernatant was recovered, and the protein concentrations were determined by the Bradford assay (Bio-Rad).

#### In-solution protein digestion

About 10 mg of total protein extracts was reduced with 10 mM dithiothreitol (DTT) at 37 °C for 1 h and was alkylated with 55 mM iodoacetamide at room temperature in the absence of light for 1 h. The sample was diluted to 1:3 (v:v) ratio with 25 mM ammonium bicarbonate buffer (pH 8.5) and then digested with TPCK trypsin (1:50 w/w) (Pierce) at 37 °C for overnight. The tryptic peptides were desalted using SDB-XC StageTip with SDB-XC Empore disc membranes (3 M) [23] and eluted in a buffer containing 0.1% trifluoroacetic acid (TFA) and 80% acetonitrile (ACN). The eluates were dried in a SpeedVac concentrator (Thermo Electron Corporation) and stored at -20 °C.

#### In-gel protein digestion

About 2 mg of the protein extracts were resolved on a 12.5% SDS-PAGE gel (1.5-mm-thick) and stained with Coomassie Brilliant Blue G-250. The gel was sliced into eight parts, and the individual slices were further diced into small pieces (0.5–1 mm<sup>3</sup>) for in-gel digestion. Each piece was first washed with water and completely destained with 50% ACN in 25 mM ammonium bicarbonate buffer (pH 8.5). ACN was removed from gel slices followed by re-equilibration in 25 mM ammonium bicarbonate buffer (pH 8.5), reduced with 30 mM DTT at 37 °C for 1 h, and then alkylated with 60 mM iodoacetamide at room temperature in the dark for 1 h. Before trypsin digestion was performed, 100% ACN covered the

gel pieces until the gel pieces became white and shrunken, then dried in a vacuum centrifuge. The gel pieces were partially swollen with 25 mM ammonium bicarbonate buffer (pH 8.5). The TPCK trypsin (1:50 w/w) (Pierce) dissolved in the same solution was added to recuperate their original size. The digestion was carried out overnight at 37 °C and stopped by the extraction buffer containing 5% TFA and 50% ACN. The resulting peptides were recovered twice with the extraction buffer by 10 min incubation in the sonication bath. The extracts were dried in a vacuum centrifuge and stored at -20 °C.

#### Phosphopeptide enrichment and MS analysis

Phosphopeptides from the digested peptides were enriched by custom-made HAMMOC tips, which were prepared using 0.5 mg TiO<sub>2</sub> beads (GL Sciences) packed into 10-μL C8-StageTips, as described previously [23, 24]. The HAMMOC tips were washed with solution A (0.1% TFA, 80% ACN) followed by the equilibration of the solution A containing 300 mg/mL lactic acid as a selectivity enhancer (solution B). About 100 μg of the dry tryptic digest was re-dissolved in solution A and diluted with an equal volume of solution B before sample loading. After washing the bound phosphopeptides twice with solution B and solution A individually, 0.5 and 5% piperidine were used for elution. The other enrichment strategy was performed with pre-incubated TiO<sub>2</sub> beads by loading buffer (1 M glycolic acid in 80% ACN and 5% TFA). Total peptide solution was then mixed and incubated with 2 mg TiO<sub>2</sub> beads for 1 h at 4 °C. Nonspecific binding peptides were washed with loading buffer and wash buffer (80% ACN and 5% TFA), and the bound phosphopeptides were eluted with 1% NH<sub>4</sub>OH in 40% ACN, pH > 10.5. The eluate was acidified with 20% phosphoric acid to pH 2.5, and desalted using SDB-XC StageTip as described above. The resulting phosphopeptides were concentrated in a vacuum centrifuge before subsequent nanoLC-MS/MS analysis as previously described [13–15].

#### MS/MS database searching and phosphorylation site analysis

All previously published (*A. baumannii* SK17, *K. pneumoniae* NTUH-K2044, *T. thermophilus* HB27) [13–15] and newly acquired MS and MS/MS raw data were analyzed using MaxQuant (version 1.5.1.2, <http://www.coxdocs.org/doku.php?id=maxquant:start>) [25] with the built-in search engine Andromeda [26] for phosphopeptide identification and phosphorylation site analysis. The protein sequences for MS/MS database search consisted of published genome databases from National Center for Biotechnology Information, including *A. baumannii* SK17, *A. platensis* C1, *H. pylori* 26695, *K. pneumoniae* NTUH-K2044, *M.*

*mazei* Go1, *M. taiwanensis* DSM 14542, *T. thermophilus* HB27 and *V. vulnificus* YJ106, and an in-house draft genome sequence of *M. portucalensis* FDF1<sup>T</sup> containing in total 2131 protein sequences constructed from a well annotated genome of *M. mahii* DSM 5219 [27], which shares 99.58% 16S rRNA sequence identity with *M. portucalensis* FDF1<sup>T</sup>. The protein-encoding genes from *M. portucalensis* FDF1<sup>T</sup> genome sequence were predicted previously by Glimmer 2.13 [28], GeneMark 2.4, and GeneMark.hmm 2.1 [29] and annotated with the RefSeq Microbial Genomes database [30] using BLASTP in standard settings ( $E$ -value  $<10^{-5}$ , identity  $>40\%$ , and matched length  $>30\%$ ).

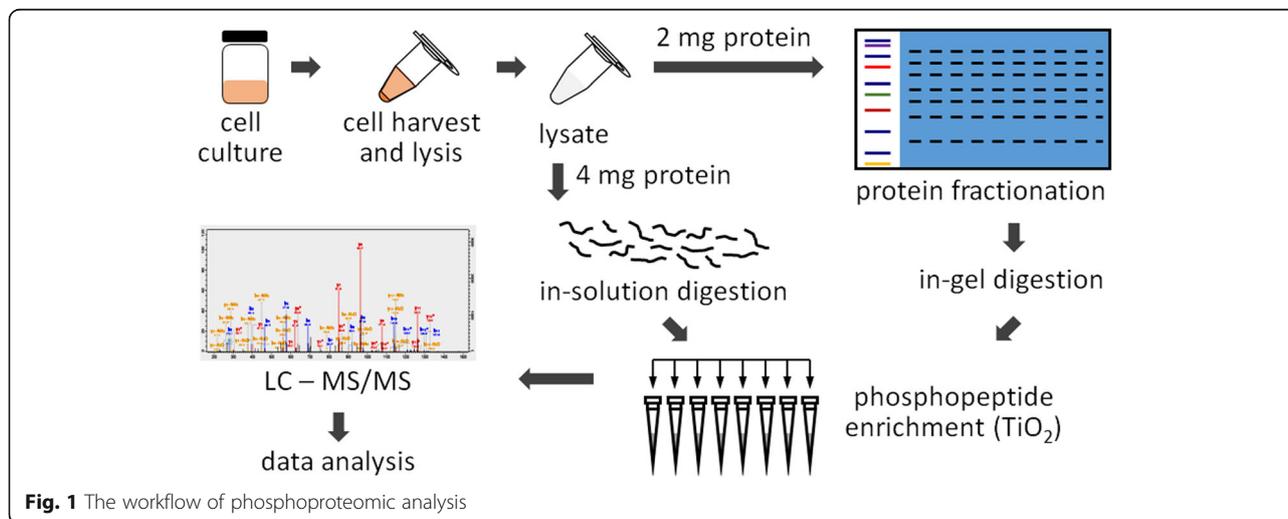
The search criteria used for phosphopeptide and phosphosite analysis were as follows: trypsin digestion; cysteine carboxyamidomethylation (+ 57.0214 Da) as the fixed modification; methionine oxidation (+ 15.9949 Da), phosphorylation of amino acid (Ser, Thr, Tyr, His, Asp) residues, and protein N-terminal acetylation as variable modifications; up to two missed cleavage allowed; minimum seven amino acids per peptide; and mass accuracy of 10 ppm for the parent ion and 0.6 Da for the fragment ions. False discovery rate was estimated from the target-decoy strategy to distinguish correct and incorrect identification. For the identification, false discovery rate was set to 0.01 for sites, peptides and proteins. We calculated the localization probabilities of all phosphorylation sites using the PTM score algorithm as previously described [31]. The phosphorylation site of at least 0.75 probability of phosphorylation is defined as reliable. MS/MS spectra of peptides containing reliable phosphohistidine and phosphoaspartate were manually inspected to confirm unambiguous identification. Only the identified phosphoproteins matched to protein sequences of the corresponding organism are reported.

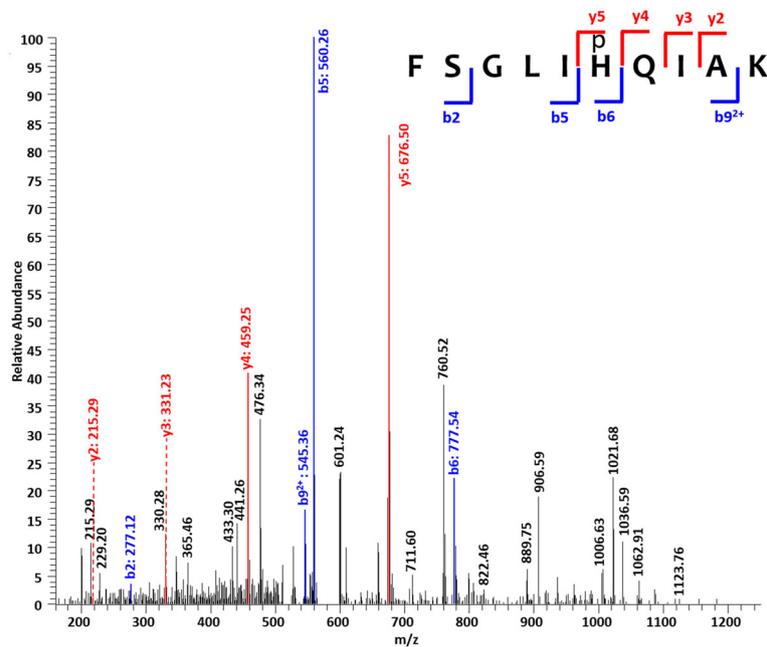
## Results

We previously performed phosphoproteome analyses of various bacterial species [13–15], and showed that trypsin-digested phosphorylated peptides could be enriched by titanium dioxide (TiO<sub>2</sub>) chromatography [23, 24] followed by LC-MS/MS analysis (Fig. 1). We showed that Ser/Thr/Tyr phosphorylations can be clearly characterized, but phosphorylation of other amino acid residues remains unclear. Nevertheless, histidine and aspartate phosphorylations play critical roles in signaling transductions [7], and thus deciphering the proteome-wide extent of such modifications is of great interest.

In this work, we switched our focus and collected samples from nine different species harvested at the mid-exponential phase. These include the clinically isolated pathogens, *A. baumannii* SK17, *H. pylori* 26695, *K. pneumoniae* NTUH-K2044 and *V. vulnificus*, as well as naturally isolated prokaryotes, including cyanobacterium *A. platensis* C1, thermophiles *M. taiwanensis* WR220 and *T. thermophilus* HB27 and methanogenic archaea *M. mazei* N2 M9705 and *M. portucalensis* FDF1<sup>T</sup>.

Upon in-solution and in-gel tryptic digestion, the samples were enriched and analyzed using the shotgun LC-MS/MS approach. The phosphopeptides derived from these samples were scrutinized with at least seven amino acids and a minimum five fragmented b-/y-ions using high-confidence identification (false discovery rate  $<0.01$ , localization probability  $>0.75$ , score of phosphopeptides  $>40$ ). To unambiguously identify phosphorylated histidine and aspartate residues, we also inspected the individual MS/MS spectra (Fig. 2 and Additional file 1) to assure the presence of critical b-/y-ions to support phosphorylation of only the designated histidine or aspartate residue but none of the other amino acids (i.e. Ser, Thr, Tyr, Arg, Lys, Glu, Cys) [5, 6]. For example, for peptide





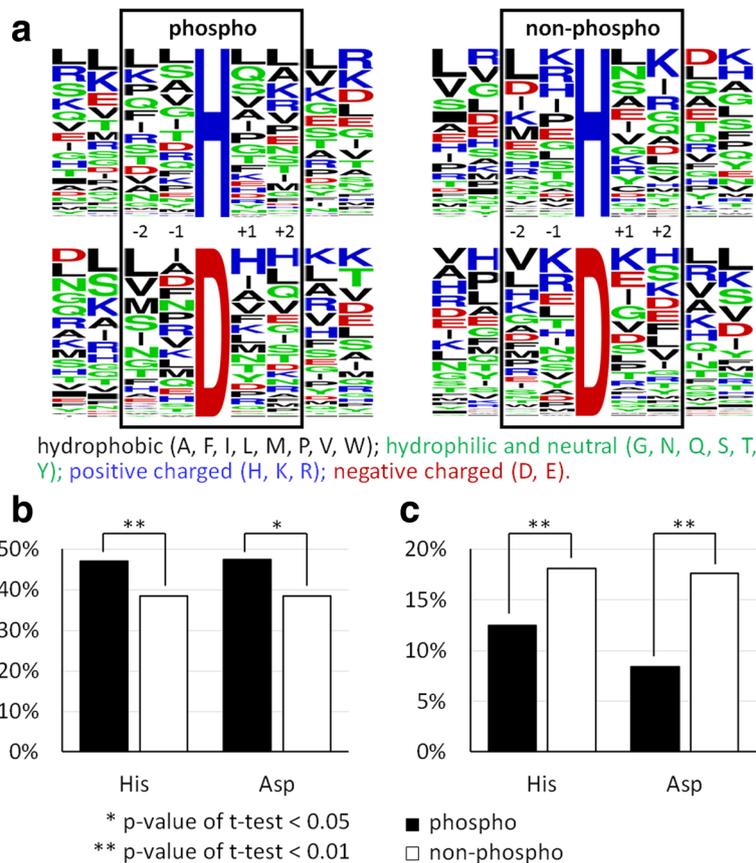
**Fig. 2** MS/MS spectra of a representative phosphohistidine peptide from *H. pylori*. The presence of y4-, y5, b5- and b6-ions enables pinpointing the phosphorylation at the indicated histidine residue. Spectra of other phosphopeptides can be found in Additional file 1

FSGLIphQIAK (Fig. 2), phosphorylation of the histidine residue is unambiguously supported by the presence of y4-, y5, b5- and b6-ions. Phosphopeptides that lack critical b-/y-ions were discarded during manual inspection. Under these stringent conditions, we found that histidine and aspartate phosphorylations are common phenomena with nearly 200 post-translationally modified peptides identified. We noticed that the majority of the identified phosphopeptides was from in-solution trypsin digestion followed by HAMMOC tip enrichment (Additional file 2).

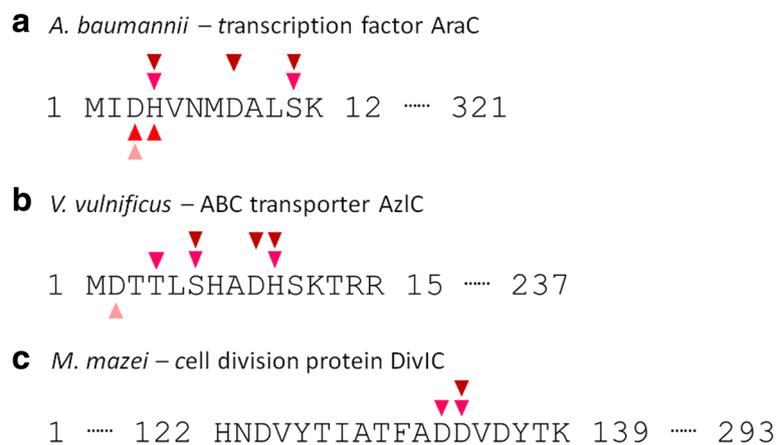
Relative abundances of amino acids flanking the phosphorylated or non-phosphorylated His and Asp residues identified in this study were compared (Fig. 3a and Additional file 3). In the immediate positions (-2, -1, +1, +2), the frequency of hydrophobic amino acid is higher around the phosphorylated versus non-phosphorylated sites (Fig. 3b). On the other hand, the probability of Arg and Lys residues is significantly higher around the non-phosphorylated than phosphorylated sites (Fig. 3c). Since trypsin cleaves exclusively at the C-terminal of Arg and Lys residues [32], sites in close proximity to Arg and Lys will end up close to peptide terminus upon trypsin digestion. Hence, our results indicated that phosphorylated sites are more distant from the peptide terminus. Consequently, we conclude that the identified phosphorylated His and Asp residues preferentially reside in an internal position of the peptide surrounded by hydrophobic amino acid residues.

Interestingly, we also found peptides with different extents of phosphorylations (Fig. 4). Enrichment and subsequent identification of mono-phosphorylated peptides are generally more straightforward in comparison to multi-phosphorylated peptides [9]. Nevertheless, we found peptides with mono- and multi-phosphorylation sites, and some multi-phosphorylated peptides contain more than one phosphorylated histidine and/or aspartate residues (Additional file 2). Noticeably, we found a few peptides of the same amino acid sequence but with different phosphorylation states in several organisms (Fig. 4). The results further highlight the prevalence of protein phosphorylation in bacteria.

As shown in Table 2, the identified phosphoproteins fall in different functional classes. In all organisms, we identified proteins involved in metabolic pathways and cellular processes. We are particularly interested in proteins potentially involving in pathogenesis. Potential candidates include those participating in cellular processes, signal transduction, transport and binding, and transcription. We highlight twelve proteins here and classified our findings based on their functions (Table 3). AzlC in *V. vulnificus* [33], FecA in *H. pylori* [34], YeeO in *K. pneumoniae* [35] and ZnuA in *A. baumannii* [36] are important protein transporters linked to drug resistance. Other putative drug-resistance proteins in the pathogenic bacteria include transcription factor AraC [37], sulfate transport system permease CysW [38], twitching motility protein PilT [39], taurine-binding periplasmic protein



**Fig. 3** Sequence features of the phosphorylation sites. All His/Asp residues in the identified phosphoproteins were classified as either phosphorylated or non-phosphorylated for the analysis. **a** Relative abundances of amino acids between -4 and +4 positions of the phosphorylated or non-phosphorylated His/Asp. Sequence logos were generated using WebLogo (<http://weblogo.berkeley.edu/>). Alignment of sequences between -10 and +10 positions is shown in Additional file 3. **b** Probability of hydrophobic amino acids between -2 and +2 positions. **c** Probability of Arg and Lys between -2 and +2 positions



**Fig. 4** Peptides identified with different modification states. Unambiguously phosphorylation sites characterized by b- and y-ions are indicated. Triangle of the same color refers to phosphorylation identified from a single phosphopeptide. **a** AraC of *A. baumannii*. **b** AzIC of *V. vulnificus*. **c** DivIC of *M. mazei*

**Table 2** Distribution of identified phosphoproteins in the nine organisms by molecular function

	<i>Ab</i>	<i>Hp</i>	<i>Kp</i>	<i>Vv</i>	<i>Ap</i>	<i>Mt</i>	<i>Tt</i>	<i>Mm</i>	<i>Mp</i>
Metabolism (amino acid/carbohydrate/others)	3/-/5	4/-/6	1/-/2	-/5/2	-/2/3	2/3/-	-/4/1	-/4/-	-/-/1
Cellular process/signal transduction/transport and binding	5/1/9	2/1/1	1/-/1	8/3/2	2/1/-	3/3/-	1/1/-	5/2/-	6/-/2
Transcription/translation/protein folding and modification	3/-/-	1/1/1	-/-/-	9/-/1	-/-/2	1/3/1	-/2/-	4/3/-	3/-/-
Unknown	1	6	2	8	4	5	1	4	3
Total number of unique proteins	27	23	7	38	14	21	10	22	15

TauA [40], inner membrane protein YfiN [41], mobile element protein and plasmid mobilization protein. Lastly, pseudouridine synthase in *H. pylori* is a key nucleic acid-modifying protein that contributes to RNA stability [42].

### Discussion

Protein phosphorylation is inherently linked to bacterial pathogenesis as it plays pivotal roles in signal transduction, modulation of metabolic behavior and confers antibiotic resistance [2–7]. Standard experimental method for phosphoproteomic studies employs TiO<sub>2</sub>-based metal oxide chromatography, which enriches phosphopeptides at low pH, coupled with standard LC-MS/MS characterizations in acidic buffers [9, 23, 24]. While this approach has been used to characterize peptides containing phosphorylated serine, threonine or tyrosine residues, it is rarely used to investigate the phosphorylation events of other amino acids as it remains unclear whether certain modified residues are sufficiently stable under these conditions. In this work, we applied highly stringent criteria for the identification of phosphorylated peptides isolated from various pathogens and demonstrated that the coupling of metal oxide chromatography with LC-MS can be used to identify unprecedented phospho-histidine and -aspartate peptides.

We began our analyses with *K. pneumoniae* and *V. vulnificus*, the causative agents of pneumonia and cholera,

respectively. Interestingly, two membrane proteins are found to undergo different degrees of phosphorylation. In *K. pneumoniae*, the multidrug and toxic compound extrusion (MATE) protein YeeO was phosphorylated at His5. The biophysical properties of MATEs have been intensively studied over the past few years because of their abilities to deliver antibiotics out of the bacteria via proton/sodium-coupled molecular pumping [33]. Indeed, *K. pneumoniae* YeeO shares 78% sequence identity with the *E. coli* homologue, which has been shown to reduce the host susceptibility to several structurally diverse antibiotics [35]. Though the exact mechanism has not been clearly elucidated, crystallography studies strongly suggest that this V-shaped channel needs to undergo substantial conformational change during molecular pumping, switching from the cytoplasm-opening state to the periplasm-opening state [43]. Since the N-terminal segment is located near the center of the channel on the periplasmic side, adding a negative charge through histidine phosphorylation will likely modify the dynamic behavior of this membrane protein. On the other hand, in *V. vulnificus* the AzlC protein of ATP binding cassette (ABC) transporter was found to be heavily phosphorylated at its N-terminus, including aspartate, histidine, serine and threonine (Fig. 4b). Both MATE and ABC transporters are major families of multidrug resistance transporters [33]. In addition, ABC transporters play significant roles in drug resistance not only in bacteria

**Table 3** Identified phosphoproteins with putative link to drug resistance in pathogenic bacteria

Species	Protein	Functional class	pHis and pAsp sites
<i>Ab</i>	Periplasmic-binding zinc ABC transporter ZnuA	Transport/binding	His6
<i>Ab</i>	Transcription factor AraC	Transcription	Asp3, His4, Asp8
<i>Ab</i>	Sulfate transport system permease CysW	Transport/binding	His15, Asp17
<i>Ab</i>	Twitching motility protein PilT	Transport/binding	His73, Asp82
<i>Ab</i>	Taurine-binding periplasmic protein TauA	Transport/binding	Asp3
<i>Ab</i>	Inner membrane protein YfiN	Transport/binding	Asp224, Asp228
<i>Ab</i>	Mobile element protein	Gene mobility	His470
<i>Ab</i>	Plasmid mobilization protein	Gene mobility	His129
<i>Hp</i>	Iron(III) dicitrate transport protein FecA	Transport/binding	His21, H526
<i>Hp</i>	RNA pseudouridine synthase	metabolism	His36, His72
<i>Kp</i>	MATE family transport protein YeeO	transport/binding	His5
<i>Vv</i>	ABC transporter AzlC	transport/binding	Asp2, Asp9, His10

but also human tumors [33]. Studying the functional role of these post-translational modifications may therefore reveal new therapeutic strategy.

The mortality rate caused by the nosocomial multi-drug resistant *A. baumannii* has increased at an alarming rate during the past decade [44]. A wide range of proteins that contribute to antibiotic resistance are found to be phosphorylated at various positions, including histidine and aspartate. Noticeably, the transcription factor AraC is heavily phosphorylated at its N terminus. This family of transcription factor is widely distributed in diverse prokaryote genera and regulates genes involved in metabolism and virulence [37]. Phosphorylation of AraC was also identified in the nitrogen-fixing bacterium, *Sinorhizobium meliloti* [45]. Among different AraC transcription factors, the N-terminus domain is less conserved and is involved in allosteric regulation and dimerization by the co-inducer [46]. It is likely that AraC in *A. baumannii* is regulated via N-terminus phosphorylation, where the addition of a phosphate group induces other phosphorylation events as well as phosphate group transfer among the residues in vivo (Fig. 4a). Additionally, ZnuA, a key periplasmic transporter in the operon of the Zn storage system, is also heavily phosphorylated at the N-terminus [36]. Proteins in this transport system is attributed to enhance Zn ion chelation to the antibiotic calprotectin, thereby quenching its antibacterial effects. Homologues of other transport and binding proteins which have a direct link to drug resistance are also shown to be phosphorylated in *A. baumannii*. These include sulfate transport system permease CysW [38], twitching motility protein PilT [39], taurine-binding periplasmic protein TauA [40] and inner membrane protein YfiN [41]. Additionally, proteins involved in horizontal gene transfer, mobile element protein and plasmid mobilization protein, also undergo histidine and aspartate phosphorylations.

*H. pylori*, the causative agent of gastric ulcer, has been proven to enhance the development of gastric cancer, a major economic burden worldwide [47]. For the first time, we demonstrate that two non-two-component proteins undergo histidine phosphorylation in this pathogen. One of the modified proteins is FecA, an outer membrane protein responsible for the cellular uptake of iron(III) citrate [34]. Inside the human tissues, the free iron concentration is not sufficient to support bacterial growth and thus iron transport and storage are essential for the survival of *H. pylori* [34]. While the two-component EnvZ-OmpR system is known to regulate FecA at a translational level [48], other regulatory mechanisms of FecA remain unclear. Here, we demonstrated that FecA is likely to be regulated at a post-translational level by phosphorylation at His21 and His526 (Additional file 2). Homology modeling [49] showed that His21 is located at the periplasm side of the  $\beta$ -barrel transport protein, whereas His526 is close to the

ligand binding site exposed (Additional file 4). Accordingly, they might involve in an unexplored signaling network with physiologically important cross-talk between these modifications. This work also revealed that pseudouridine synthase in *H. pylori* is phosphorylated at several of the histidine, serine and threonine positions. Site-specific uridine isomerization within a RNA molecule is the most common post-transcriptional modification found in prokaryotes. By forming a hydrolysis-resilient non-canonical C5-glycosidic bond, pseudouridine synthase is capable to enhance the structural stability of the RNA molecules and consequently cellular stability [42]. In this work, we showed that the histidine phosphorylation sites (position 36 and 72) of the pseudouridine synthase are located in the S4 RNA-binding domain. The addition of negative charges by these post-translation modifications may modulate the substrate binding affinity of this enzyme.

By switching our proteomic focus, we have revealed novel protein phosphorylation of over ten pathogenesis-related proteins found in several commonly seen pathogens. As demonstrated by the current and previous work by us and others, enrichment of phosphopeptides by TiO<sub>2</sub> is highly robust, as it is highly specific toward phosphopeptides with femtomolar sensitivity and large tolerance of salts and detergents [9]. However, there was a controversy on whether the acidic condition used in standard LC-MS (typically 0.1% formic acid in running buffer) is too harsh for certain phosphorylated amino acid residues. It has been shown that phosphorylation of histidine and aspartate are less stable than those of serine, threonine and tyrosine; free phosphohistidine residue has also been shown to degrade in low-pH buffer, whereas free phosphoaspartate residue appears to be unstable in both acidic and alkali conditions [50]. However, the half-lives of the phospho-histidine and -aspartate peptides are most likely to be sequence-dependent. Hydrophobic residues situated next to the phosphorylated residue appear to be more profound in our study (Fig. 3c), thus suggesting that they may protect the post-translationally modified residue from cleavage. Furthermore, individual peptides might adopt specific secondary conformations that slow the hydrolysis process. As an example, phosphohistidine in *Escherichia coli* HPr protein is actually more resilient to hydrolysis at low pH than under neutral conditions [51]. These factors most likely contribute to why phospho-histidine and -aspartate peptides can be characterized in this work. Indeed, in our view it is rather difficult to draw a consensus on the stability of phosphorylated peptides, as their sequences dictate their structures, physical properties and consequently their chemical stabilities.

Better methods to enrich and identify phospho-histidine and aspartate peptides are needed, as phosphopeptides in

low abundance or with less chemical stability are unlikely to be found in regular characterization methods. For example, the acid-labile arginine phosphorylation was found only in *Bacillus subtilis* strain with arginine phosphatase knocked out [52]. Hence, genetic removal of other phosphatases might enhance characterizations of different phosphopeptides. Performing the MS analysis in neutral buffer may also improve the characterization of histidine phosphorylation, though the experimental condition is currently less refined than that of the acidic counterpart [10]. Lastly, novel chemical biology tools, such as reagents that target and report histidine and aspartate phosphorylations, will likely find uses in the field of phosphoproteomics. This work demonstrates that there is an additional level of protein regulations beyond our current knowledge of microbial molecular biology. Most importantly, this work illustrated new research directions for over ten proteins that are related to pathogen survival and antibiotic resistance. Hence, it will be of paramount interest to identify the functional impact of the identified phosphorylation, as well as the designated kinases responsible for the post-translational modifications. Such information may provide novel insights and targets to combat the current antibiotic resistance crisis.

## Conclusions

We performed site-specific His/Asp phosphoproteomic analysis of nine prokaryotes, including four pathogenic bacteria isolated from hospital patients. The use of highly stringent conditions ensured unambiguous identification of the phosphorylated sites, including histidine and aspartate, which were previously thought to be uncharacterizable under standard acidic LC-MS conditions. This work clearly illustrates that His/Asp phosphorylation events occur beyond two-component systems. Noticeably, for the first time a number of key pathogenesis-related proteins are shown to undergo His and Asp phosphorylations. These findings have shed light on important research directions, which will eventually find uses in combating the emergence of antibiotic resistance and designing new therapeutics in the future.

## Additional files

**Additional file 1:** MS/MS Spectra. Spectra of the identified phosphopeptides (PDF 4714 kb).

**Additional file 2: Table S1.** Phospho probabilities, process method, leading protein identifier, UniProt No. (if available), protein description, and protein functional class of the identified phosphopeptides (PDF 539 kb).

**Additional file 3: Figure S1.** The relative abundances of amino acids between -10 and +10 positions of the phosphorylated or non-phosphorylated His/Asp (PDF 377 kb).

**Additional file 4: Figure S2.** Homology model of *H. pylori* FecA (PDF 359 kb).

## Abbreviations

ABC: ATP binding cassette; ACN: Acetonitrile; DTT: Dithiothreitol; LC: Liquid chromatography; MATE: Multidrug and toxic compound extrusion; MS: Mass spectrometry; PBS: Phosphate-buffered saline; TFA: Trifluoroacetic acid

## Acknowledgements

We are grateful to financial supports from and We thank the Mass Spectrometry supports from NRPGM Core Facilities at the Institute of Biological Chemistry, Academia Sinica for technical supports. We also thank Mr. Samuel Franklin for proofreading of the manuscript.

## Funding

Work in the laboratory of SHW is supported by Academia Sinica (AS-104-SS-A05) and Taiwan Ministry of Science and Technology (NSC-101-2923-B-001-005-MY3, MOST-103-2113-M-001-029-MY3, MOST-103-2325-B-001-020, MOST-104-2325-B-001-003). We also thank support from Wellcome Trust (202056/Z/16/Z to LYPL, 200730/Z/16/Z to YHT). The funders had no role in study design, data collection and interpretation, or on the decision to submit the work for publication.

## Availability of data and materials

All data and materials are available in the research group of Shih-Hsiung Wu, Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan. Original MS data generated and analyzed in this study are available upon request.

## Authors' contributions

SJL and SHW designed research; SJL, IFT, WLW and JTY performed experiments; SJL, LYPL, MCL and YHT analyzed the data; SJL, LYPL, YHT and SHW wrote the manuscript; all authors have read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan. <sup>2</sup>PhD Program in Microbial Genomics, National Chung Hsing University and Academia Sinica, Taichung, Taiwan. <sup>3</sup>School of Chemistry, Cardiff University, Cardiff, UK. <sup>4</sup>Department of Life Sciences, National Chung Hsing University, Taichung, Taiwan. <sup>5</sup>Chemical Biology and Molecular Biophysics Program, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan. <sup>6</sup>Department of Chemistry, National Taiwan University, Taipei, Taiwan.

Received: 10 February 2017 Accepted: 15 May 2017

Published online: 25 May 2017

## References

- Roskoski R. A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacol Res.* 2015;100:1–23.
- Dworkin J. Ser/Thr phosphorylation as a regulatory mechanism in bacteria. *Curr Opin Microbiol.* 2015;24:47–52.
- Mijakovic I, Macek B. Impact of phosphoproteomics on studies of bacterial physiology. *FEMS Microbiol Rev.* 2012;36:877–92.
- Hutchings MI, Hong HJ, Buttner MJ. The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. *Mol Microbiol.* 2006;59:923–35.
- Mijakovic I, Grangeasse C, Turgay K. Exploring the diversity of protein modifications: special bacterial phosphorylation systems. *FEMS Microbiol Rev.* 2016;40:398–417.
- Attwood PV, Besant PG, Piggott MJ. Focus on phosphoaspartate and phosphoglutamate. *Amino Acids.* 2011;40:1035–51.
- Casino P, Rubio V, Marina A. The mechanism of signal transduction by two-component systems. *Curr Opin Struct Biol.* 2010;20:763–71.

8. Cain JA, Solis N, Cordwell SJ. Beyond gene expression: the impact of protein post-translational modifications in bacteria. *J Proteome*. 2014;97:265–86.
9. Engholm-Keller K, Larsen MR. Technologies and challenges in large-scale phosphoproteomics. *Proteomics*. 2013;13:910–31.
10. Kleinnijenhuis AJ, Kjeldsen F, Kallipolitis B, Haselmann KF, Jensen ON. Analysis of histidine phosphorylation using tandem MS and ion-electron reactions. *Anal Chem*. 2007;79:7450–6.
11. Hohenester UM, Ludwig K, Kriegelstein J, König S. Proteomic methodologies access acid-labile histidine phosphorylation. *Biomacromol Mass Spectrom*. 2010;2:71–6.
12. Hongsthong A, Sirijuntarut M, Prommeenat P, Lertladaluck K, Porkaew K, Cheevadhanarak S, et al. Proteome analysis at the subcellular level of the cyanobacterium *Spirulina platensis* in response to low-temperature stress conditions. *FEMS Microbiol Lett*. 2008;288:92–101.
13. Wu WL, Liao JH, Lin GH, Lin MH, Chang YC, Liang SY, et al. Phosphoproteomic Analysis Reveals the Effects of PiIF Phosphorylation on Type IV Pilus and Biofilm Formation in *Thermus thermophilus* HB27. *Mol Cell Proteomics*. 2013;12:2701–13.
14. Lin MH, Hsu TL, Lin SY, Pan YJ, Jan JT, Wang JT, et al. Phosphoproteomics of *Klebsiella pneumoniae* NTUH-K2044 Reveals a Tight Link between Tyrosine Phosphorylation and Virulence. *Mol Cell Proteomics*. 2009;8:2613–23.
15. Lai JH, Yang JT, Chern J, Chen TL, Wu WL, Liao JH, et al. Comparative Phosphoproteomics Reveals the Role of AmpC beta-lactamase Phosphorylation in the Clinical Imipenem-resistant Strain *Acinetobacter baumannii* SK17. *Mol Cell Proteomics*. 2016;15:12–25.
16. Chen TL, Lee YT, Kuo SC, Hsueh PR, Chang FY, Siu LK, et al. Emergence and Distribution of Plasmids Bearing the bla(OXA-51)-Like Gene with an Upstream ISAbal1 in Carbapenem-Resistant *Acinetobacter baumannii* Isolates in Taiwan. *Antimicrob Agents Chemother*. 2010;54:4575–81.
17. Fang CT, Chuang YP, Shun CT, Chang SC, Wang JT. A novel virulence gene in *Klebsiella pneumoniae* strains causing primary liver abscess and septic metastatic complications. *J Exp Med*. 2004;199:697–705.
18. Koyama Y, Hoshino T, Tomizuka N, Furukawa K. Genetic Transformation of the Extreme Thermophile *Thermus-Thermophilus* and of Other *Thermus* Spp. *J Bacteriol*. 1986;166:338–40.
19. Lai MC, Shu CM, Chiou MS, Hong TY, Chuang MJ, Hua JJ. Characterization of *Methanosarcina mazei* N2M9705 isolated from an aquaculture fishpond. *Curr Microbiol*. 1999;39:79–84.
20. Lai MC, Sowers KR, Robertson DE, Roberts MF, Gunsalus RP. Distribution of Compatible Solutes in the Halophilic Methanogenic Archaeobacteria. *J Bacteriol*. 1991;173:5352–8.
21. Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS. Methanogens - Re-Evaluation of a Unique Biological Group. *Microbiol Rev*. 1979;43:260–96.
22. Lai SJ, Lai MC. Characterization and Regulation of the Osmolyte Betaine Synthesizing Enzymes gSMT and SDMT from Halophilic Methanogen *Methanohalophilus gosalensis*. *PLoS One*. 2011;6:e25090.
23. Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc*. 2007;2:1896–906.
24. Sugiyama N, Masuda T, Shinoda K, Nakamura A, Tomita M, Ishihama Y. Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics*. 2007;6:1103–9.
25. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26:1367–72.
26. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J Proteome Res*. 2011;10:1794–805.
27. Spring S, Scheuner C, Lapidus A, Lucas S, Del Rio TG, Tice H, et al. The Genome Sequence of *Methanohalophilus mahii* SLPT Reveals Differences in the Energy Metabolism among Members of the Methanosarcinaceae Inhabiting Freshwater and Saline Environments. *Archaea*. 2010;2010:690737.
28. Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*. 2007;23:673–9.
29. Borodovsky M, Mills R, Besemer J, Lomsadze A. Prokaryotic gene prediction using GeneMark and GeneMark.hmm. *Curr Protoc Bioinformatics*. 2003; Chapter 4:Unit4.5. (<https://www.ncbi.nlm.nih.gov/pubmed/18428700>).
30. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2007;35:D61–D5.
31. Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. 2006;127:635–48.
32. Olsen JV, Ong SE, Mann M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics*. 2004;3:608–14.
33. Lubelski J, Konings WN, Driessen AJ. Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiol Mol Biol Rev*. 2007;71:463–76.
34. Pich OQ, Merrell DS. The ferric uptake regulator of *Helicobacter pylori*: a critical player in the battle for iron and colonization of the stomach. *Future Microbiol*. 2013;8:725–38.
35. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, et al. Phenotypic landscape of a bacterial cell. *Cell*. 2011;144:143–56.
36. Hood MI, Mortensen BL, Moore JL, Zhang Y, Kehl-Fie TE, Sugitani N, et al. Identification of an *Acinetobacter baumannii* zinc acquisition system that facilitates resistance to calprotectin-mediated zinc sequestration. *PLoS Pathog*. 2012;8:e1003068.
37. Ibarra JA, Perez-Rueda E, Segovia L, Puente JL. The DNA-binding domain as a functional indicator: the case of the AraC/XylS family of transcription factors. *Genetica*. 2008;133:65–76.
38. Wang X, Wang H, Xie J. Genes and regulatory networks involved in persistence of *Mycobacterium tuberculosis*. *Sci China Life Sci*. 2011;54:300–10.
39. Han X, Kennan RM, Davies JK, Reddacliff LA, Dhungyel OP, Whittington RJ, et al. Twitching motility is essential for virulence in *Dichelobacter nodosus*. *J Bacteriol*. 2008;190:3323–35.
40. Fuentes-Hernandez A, Plucain J, Gori F, Pena-Miller R, Reding C, Jansen G, et al. Using a sequential regimen to eliminate bacteria at sublethal antibiotic dosages. *PLoS Biol*. 2015;13:e1002104.
41. Giardina G, Paiardini A, Ferriccola S, Franceschini S, Rinaldo S, Stelitano V, et al. Investigating the allosteric regulation of YfiN from *Pseudomonas aeruginosa*: clues from the structure of the catalytic domain. *PLoS One*. 2013;8:e81324.
42. Hamma T, Ferre-D'Amare AR. Pseudouridine synthases. *Chem Biol*. 2006;13:1125–35.
43. Lu M. Structures of multidrug and toxic compound extrusion transporters and their mechanistic implications. *Channels*. 2016;10:88–100.
44. Gulen TA, Guner R, Celikbilek N, Keske S, Tasyaran M. Clinical importance and cost of bacteremia caused by nosocomial multi drug resistant *Acinetobacter baumannii*. *Int J Infect Dis*. 2015;38:32–5.
45. Liu T, Tian CF, Chen WX. Site-Specific Ser/Thr/Tyr Phosphoproteome of *Sinorhizobium meliloti* at Stationary Phase. *PLoS One*. 2015;10:e0139143.
46. Ninfa AJ. In: Neidhardt FC, Curtiss R, Ingraham JL, ECC L, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umberger HE, editors. *Escherichia coli* and *Salmonella typhimurium*, Cellular and Molecular Biology. Washington, DC: American Society for Microbiology; 1996. p. 1246–62.
47. Hardbower DM, Peek RM, Wilson KT. At the Bench: *Helicobacter pylori*, dysregulated host responses, DNA damage, and gastric cancer. *J Leukoc Biol*. 2014;96:201–12.
48. Noinaj N, Guillier M, Barnard TJ, Buchanan SK. TonB-Dependent Transporters: Regulation, Structure, and Function. *Annu Rev Microbiol*. 2010;64:43–60.
49. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014;42:W252–8.
50. Duclos B, Marcandier S, Cozzone AJ. Chemical properties and separation of phosphoamino acids by thin-layer chromatography and/or electrophoresis. *Methods Enzymol*. 1991;201:10–21.
51. Anderson JW, Pullen K, Georges F, Klevit RE, Waygood EB. The involvement of the arginine 17 residue in the active site of the histidine-containing protein, HPr, of the phosphoenolpyruvate:sugar phosphotransferase system of *Escherichia coli*. *J Biol Chem*. 1993;268:12325–33.
52. Elsholz AK, Turgay K, Michalik S, Hessling B, Gronau K, Oertel D, et al. Global impact of protein arginine phosphorylation on the physiology of *Bacillus subtilis*. *Proc Natl Acad Sci U S A*. 2012;109:7451–6.