Clostridioides difficile phosphoproteomics shows an expansion of phosphorylated proteins in stationary growth phase

Wiep Klaas Smits[†], Y. Mohammed[‡], Arnoud de Ru[‡], Valentina Cordó[‡#], Annemieke Friggen[†], Peter A. van Veelen[‡], Paul J. Hensbergen[‡*]

[†]Department of Medical Microbiology, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands

[‡]Center for Proteomics and Metabolomics, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands

[#]Current address: Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands

[*]Correspondence to:

P.J. Hensbergen, Center for Proteomics and Metabolomics, Leiden University Medical Center, PO Box 9600, 2300 RC, Leiden, The Netherlands. Tel.: +31-71-5266394. Fax: +31 71 5266907. E-mail: P.J.Hensbergen@lumc.nl

or

W.K. Smits, Department of Medical Microbiology, Section Experimental Bacteriology, Leiden University Medical Center, PO Box 9600, 2300 RC, Leiden, The Netherlands. Tel.: +31-71-5261229. E-mail: W.K.Smits@lumc.nl

LIST OF ABBREVIATIONS

HPLC = High Performance Liquid Chromatography

LC = Liquid Chromatography

LC-MS = Liquid Chromatography – Mass Spectrometry

nLC-MS/MS = Nano Liquid Chromatography -Tandem Mass Spectrometry

LP: localisation probability

aa: amino acid

ABSTRACT

Phosphorylation is a post-translational modification that can affect both house-keeping functions and virulence characteristics in bacterial pathogens. In the Gram-positive enteropathogen *Clostridioides difficile* the extent and nature of phosphorylation events is poorly characterized, though a protein-kinase mutant strain demonstrates pleiotropic phenotypes. Here, we used an immobilized metal affinity chromatography strategy to characterize serine, threonine and tyrosine phosphorylation in *C. difficile*. We find limited protein phosphorylation in the exponential growth phase but a sharp increase in the number of phosphopeptides after the onset of stationary growth phase. Among the overall more than 1500 phosphosites, our approach identifies expected targets and phosphorylation sites, including the protein kinase PrkC, the anti-sigma-F factor antagonist (SpoIIAA), the anti-sigma-B factor antagonist (RsbV) and HPr kinase/phosphorylase (HprK). Analysis of high-confidence phosphosites shows that phosphorylation on serine residues is most common, followed by threonine and tyrosine phosphorylation. This work forms the basis for a further investigation into the contributions of individual kinases to the overall phosphoproteome of *C. difficile* and the role of phosphorylation in *C. difficile* physiology and pathogenesis.

Importance

In this manuscript, we present a comprehensive analysis of protein phosphorylation in the Gram-positive enteropathogen *Clostridioides difficile*. To date, only limited evidence on the role of phosphorylation in regulation in this organism has been published; the current study is expected to form the basis for research on this post-translational modification in *C. difficile*.

INTRODUCTION

Proteins in organisms from all domains of life can be functionally altered by post-translational modifications (PTMs) that are often species-dependent (1). PTMs in bacteria generally occur at much lower levels than in eukaryotes (2). PTMs serve to regulate activity in response to environmental conditions, as they process generally faster than transcription-translation (2), but can also be used to tune phenotypic diversity (3). Amongst the high variety of PTMs (1), protein phosphorylation is one of the best-studied examples, partially attributable to major developments in phosphoproteomics.

To date, phosphorylation has been identified on the side chains of different amino acids: serine (Ser), threonine (Thr), tyrosine (Tyr), histidine (His), arginine (Arg), lysine (Lys), aspartate (Asp) and cysteine (Cys). The chemistry involved in the phosphorylation of different amino acid is different. Phosphorylation on Ser/Thr/Tyr results in more stable phosphoester bonds, whereas His/Lys/Arg phosphorylation results in thermodynamically unstable phosphoamidates (1). Phosphorylation on Asp leads to a mixed phosphoacylanhydride and modification of cysteine (Cys), finally, leads to a phosphothiolate (1). Ser/Thr/Tyr phosphorylation appears to be the most common in bacteria (4), and may relate to their thermodynamic stability.

In canonical protein phosphorylation, a phosphate group is transferred to a protein through the action of a kinase, and can be removed by a phosphatase; sometimes the kinase and phosphatase functions are encoded by the same bifunctional enzyme (1). Oftentimes, the response of bacteria to environmental stimuli is dependent on so called two component systems, that consist of a membrane-associated two-component sensor histidine kinase (HK) and a cytosolic response regulator (RR). When triggered, autophosphorylation of the HK on a

histidine residue results in the transfer of the phosphate group to an aspartate residue on the response regulator. Subsequently, the phosphorylated RR can bind to DNA to regulate transcription of target genes (1). Other His-phosphorylated proteins include the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS) proteins, with the exception of the EIIB component that can be phosphorylated on Cys residues (5). As the name suggests, the donor for the phosphorylation of PTS systems is not a kinase, but phosphoenol pyruvate (PEP). The PTS phosphorylation cascade includes an unusual bifunctional kinase/phosphatase, called HPr, that in its S-phosphorylated form can interact with the transcription factor CcpA to regulate metabolic activity (5).

Phosphorylation on Ser/Thr is mediated by eukaryotic-type Ser/Thr Kinases (eSTK), also known as Hanks-type kinases (6). Interestingly there is some evidence that eSTKs may outnumber two component systems (7). Phosphorylation on Tyr in bacteria is mediated by a unique family of proteins known as BY kinases (2, 8, 9).

The exploration of bacterial phosphorylation has been greatly stimulated by the development of mass spectrometry-based phosphoproteomics techniques that provide a snapshot of the site-resolved phosphorylation state of proteins under given conditions (10, 11). These techniques rely on strategies to enrich for phosphorylated peptides, through for instance anti-phospho antibodies, strong cation exchange chromatography, chemical modifications or immobilized metal affinity chromatography (IMAC) (2, 4). Overall, the picture that emerges from these experiments is that both the number of phosphorylation sites and the relative abundance of different types of phosphorylation differs between organisms (2, 4, 11, 12). Additionally, it was found that one protein can be dynamically phosphorylated on multiple different sites (4, 13).

In bacteria, phosphorylation has been implicated in diverse processes such as cell cycle regulation, cellular differentiation, morphogenesis, and metabolism, persistence and virulence (1). Whereas phosphoproteomics initially focused on model organisms such as the Gram-negative *Escherichia coli* (14) and Gram-positive *Bacillus subtilis* (15), phosphoproteomes have been determined for a large number of pathogens as well (16, 17). Though some of the effects on pathogenicity can be related back to pleiotropic effects on cellular integrity, for instance, phosphorylation is also directly implicated in virulence by modulating virulence gene expression or factors involved in host establishment, as well as affecting the host immune system (10, 16). Finally, phosphorylation has also been found to affect antimicrobial resistance in multiple pathogens (16, 18). Together, these findings suggest that targeting phosphorylation might be a viable strategy for novel antimicrobial or anti-virulence strategies (16, 18).

Despite the wealth of phosphoproteomic studies (16, 17), phosphorylation in Clostridia is largely unexplored. A single study has been performed in *Clostridium acetobutylicum*, a solventogenic species with biotechnological relevance, that identified a total of 61 phosphorylated proteins (19). To our knowledge, no studies have been carried out in pathogenic species such as *Clostridioides difficile*.

C. difficile is a major cause of healthcare-associated diarrhea, and can cause a potentially fatal disease (*C. difficile* infection, or CDI) as a result of the production of toxins that affect the integrity of the colon epithelium (20-22). The development of CDI is clearly linked to reduced diversity of the microbiota of the host and consequently restoring this diversity for instance by fecal microbiota transplantations has proven to be an effective treatment (23, 24). Transmission and persistence of *C. difficile* is dependent on its ability to form endospores, as in a mutant unable to form spores these processes are severely inhibited

(25). All the processes above are subject to complex regulation and - based on homology with the Gram-positive model organism *B. subtilis* – can be expected to involve phosphorylation of key proteins (26-30). However, important differences in these regulatory pathways exist between *B. subtilis* and *C. difficile* (31, 32). Importantly, genetic investigations have revealed that the eSTK protein PrkC of *C. difficile* is reported to affect cell wall homeostasis and resistance to antimicrobials, though the molecular mechanism through which this occurs have not been elucidated (33). In contrast, it was recently shown that PrkC directly affects cell division through phosphorylation of the peptidoglycan hydrolase CwlA (34). To date, this is the only reported substrate of the PrkC kinase.

Here, we report the first phosphoproteomic analysis of *C. difficile* based on an IMAC-based enrichment strategy. We show that phosphorylation is growth phase-dependent and abundant in stationary growth phase, and that several aspects of phosphorylation-dependent regulation appear to be conserved. Our results contribute to our understanding of phosphorylation in pathogens and pave the way for functional dissection of the role of kinases and phosphatases in *C. difficile*.

MATERIALS AND METHODS

Chemicals

Unless noted otherwise, chemicals were obtained from Sigma Aldrich Chemie.

Cell culture

C. difficile 630Δ*erm* cells (35, 36) were grown at 37 °C in brain-heart infusion medium (BHI; Oxoid) supplemented with yeast extract (YE) in a Don Whitley VA-1000 workstation (10% CO_2, 10% H_2 and 80% N_2 atmosphere). An initial pre-culture was grown until an optical density at 600 nm (OD_{600nm}) of 0.36. Three independent cultures were started by adding 31 mL of the pre-culture to 400 mL fresh BHI-YE for each new culture. Samples were taken at mid-exponential (3.25h post inoculation), early stationary (5.25h post inoculation) and late stationary phase (24h post inoculation) (**Supplemental Figure 1**). To get a roughly equivalent amount of cells, 100 mL was taken at the mid-exponential time point (OD_{600nm} 0.96), while 50 mL was taken at the two later time points (OD_{600nm} ~2).

JY cells were grown at 37 °C in Iscove's modified Dulbecco's medium, supplemented with 10% heat-inactivated fetal bovine serum and L-glutamine. Both eukaryotic and bacterial cells were harvested by centrifugation (3184 x g, 10 min), and washed three times with 25 mL of phosphate buffered saline (PBS; Fresenius Kabi Nederland BV). Cell pellets were subsequently stored at -20 °C until further use.

Sample preparation

Cell pellets were re-suspended in 3 mL lysis buffer (8 M urea/50 mM Tris-HCl pH 7.4, 1 mM orthovanadate, 5 mM Tris(2-carboxyethyl)phosphine (TCEP), 30 mM chloroacetamide (CAA),

phosphoSTOP phosphatase inhibitor (Thermo Fischer Scientific), cOmplete™ mini EDTA free protease inhibitor, 1 mM $MgCl_2$) and incubated for 20 min at room temperature. Cells were lysed by sonication (Soniprep 150 ultrasonic disintegrator (MSE), 5x 30 s, amplitude 12 microns). In between sonication steps samples were cooled on ice for 30 s. Next, samples were centrifuged for 15 min at 7200 x g at 4 °C. The urea concentration was adjusted to 6 M by the addition of 1 mL of 50 mM Tris-HCL pH 7.4, after which 1 µL Benzonase (250 U/ µL) was added and samples were incubated for two hrs at room temperature.

Proteins were precipitated by first adding 16 mL of methanol (Actu-All Chemicals) and mixing, followed by 4 mL of chloroform (Merck Millipore) and mixing. After the addition of twelve mL of Milli-Q water (obtained from a Elga Pure Lab Chorus 1 Machine), the samples were mixed by vortexing. Samples were then centrifuged for 15 min at 11000 g, followed by another 5 min at 11000 x g (this resulted in better separation of the two phases than one round of 20 min of centrifugation). The protein precipitates at the interphase were collected and washed twice with 2 mL methanol, followed by 5 min centrifugation at 11000 x g. The protein pellets were air-dried and re-suspended in 5 mL 25 mM NH_4HCO_3 pH 8.4.

Trypsin was added at a ratio of 1:25 (w/w) and overnight digestions were performed at 37 °C. The next day, samples were centrifuged for 10 min at 11000 x g and the supernatants were collected. Desalting of the samples was performed using Oasis HLB 1 cc Vac cartidges (Waters). Briefly, the cartridge was washed once with 1 mL acetonitrile (Actu-All Chemicals)/H_2O 90/10 (v/v), equilibrated with Milli-Q H_2O/acetonitrile/formic acid (Fluka Analytical) (95/3/0.1 v/v/v) (solution A, 2x 1 mL). Following sample loading, the column was washed with solution A (3x 1 mL) after which peptides were eluted with acetonitrile/water/formic acid (30/70/0.1 v/v/v). Tryptic peptides were lyophilized (Salm en Kipp Christ RVC 2-18 CD plus) and stored at -20 °C until use.

196

197 *Phosphopeptide enrichment*

198 A one-step phosphopeptide IMAC enrichment procedure was performed using a 4 x 50 mm
199 ProPac IMAC-10 analytical column (Thermo Fisher Scientific). For the JY cells an equivalent of
200 5 mg of protein was used for one IMAC purification while for the *C. difficile* cells, the total
201 amount of peptides from each sample were applied, with an average of 11 (± 1.4 (S.D.)) mg
202 protein per sample. Prior to each IMAC run the column was stripped using 50 mM EDTA/0.5
203 M NaCl pH 4 and charged with 25mM $FeCl_3$ in 100mM acetic acid, according to the
204 manufacturer's instructions, using an offline pump system (Shimadzu). After this, the column
205 was connected to an Agilent 1200 chromatography system running at a flow rate of 0.3
206 mL/min. Prior to sample loading, the column was equilibrated in solvent A
207 (water/acetonitrile/trifluoroacetic acid 70/30/0.07 (v/v/v)) for 30 mins at a constant flow rate
208 of 0.3 ml/min.

209 Tryptic peptides (resuspended in 260 µL solvent A, of which 250 µL injected) were
210 loaded and non-phosphopeptides were removed by washing the column for 20 min at with
211 solvent A. Peptide separation was performed using a linear gradient from 0 to 45% of solvent
212 B (0.5% v/v NH_4OH). The peak fraction of phosphopeptides between 46-49 min was manually
213 collected and lyophilized prior to mass spectrometric analysis.

214

215 *LC-MS/MS analysis*

216 Lyophilized peptides were reconstituted in 100 µL water/formic acid 100/0.1 (v/v) and
217 analysed by on-line C18 nanoHPLC MS/MS with a system consisting of an Easy nLC 1200
218 gradient HPLC system (Thermo Fisher Scientific, Bremen, Germany), and an Orbitrap Fusion™
219 Lumos™ Tribrid™ mass spectrometer (Thermo Scientific™). Samples (10 µL, in duplicate)

were injected onto a homemade pre-column (100 µm × 15 mm; Reprosil-Pur C18-AQ 3 µm, Dr. Maisch, Ammerbuch, Germany) and eluted on a homemade analytical nano-HPLC column (30 cm × 75 µm; Reprosil-Pur C18-AQ 3 um). The gradient was run from 2% to 36% solvent B (water/acetonitrile/formic acid 20/80/0.1 (v/v/v)) in 120 min. The nano-HPLC column was drawn to a tip of ~5 µm which acted as the electrospray needle of the MS source.

The Lumos mass spectrometer was operated in data-dependent MS/MS mode for a cycle time of 3 seconds, with a HCD collision energy at 32 V and recording of the MS2 spectrum in the Orbitrap. In the master scan (MS1) the resolution was 120000, the scan range *m/z* 300-1500, at an AGC target of 400000 at a maximum fill time of 50 ms. Dynamic exclusion was set after n=1 with an exclusion duration of 60 s. Charge states 2-4 were included. For MS2, precursors were isolated with the quadrupole with an isolation width of 1.2 Da. First mass was set to 110 Da. The MS2 scan resolution was 30000 with an AGC target of 50000 at maximum fill time of 60 ms.

Data analysis

MaxQuant software (version 1.5.1.2) was used to process the raw data files, which were searched against the *C. difficile* strain 630 database (4103 entries) or human canonical database (67911 entries). The mass tolerance for MS1 and MS2 was set to 4.5 and 20 ppm, respectively. Trypsin was selected as the enzyme with a maximum of two missed cleavages. Carbamidomethylation of cysteines was selected as a fixed modification and phosphorylation on serine, threonine and tyrosine, as well as oxidation of methionine and acetylation of the protein N-terminus were selected as variable modifications. The match between runs option was selected with a window of 2 min. The false discovery rate at the peptide level was set to 1% and for modified peptides a score cut-off of 40 was used.

The MaxQuant output table "phospho (STY)Sites.txt" (**Supplemental Table 1**) was used for further analysis of the phosphopeptides. For the stringent phosphorylation site assignment, only peptides with a localisation probability score >0.95 were selected. Because we used the match between runs option, we only used unique peptides without considering the phosphorylation site (i.e. not considering phospho-isomers) for comparison of different samples.

Gene Ontology protein/gene set enrichment analysis was performed using a sorted list of all identified phosphoproteins in each analysis time point, i.e. mid-exponential, beginning stationary, and late stationary growth phase. The sorting was based on the abundance of the phosphoproteins identified and determined by the ion intensity as reported by the mass spectrometer in the output of MaxQuant (cut-off at score of 40 and false discovery rate at 1%). We considered the phosphoproteins that were reported at least once in any of the three replicates. Each protein identification was considered only once in the sorted list with a rank determined by the most intense results we have obtained in any of the three replicates by any protein-associated phosphopeptide. The enrichment was performed using weighted Kolmogorov-Smirnov-like statistic as implemented in the common gene set enrichment analysis. We limited our analysis to the top 30 enriched Gene Ontology terms. We generated enrichment maps that represent the relationship between the enriched terms as a network. The edges of an enrichment map correspond to the number of shared proteins between the associated terms. These maps give a higher-level overview of the enrichment analysis and allow identifying functional modules with ontology terms that are related to each other by the underlying protein set used. For the functional analysis we used all Gene Ontology annotation of *C. difficile* as reported by the Gene Ontology database hosted by EBI (as of

August 15 2021). For the functional analysis and visualization, we used R 3.6.2 and Bioconductor library Cluster Profiler.

Data availability

All data are contained within the manuscript or the associated Supplementary Material (available via the journal website), or are available from the authors on request. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (37) with the dataset identifier PXD029475.

RESULTS

A one-step IMAC procedure allows the identification of phosphopeptides from *C. difficile*

The number of phosphorylated proteins in bacteria is generally much lower than in eukaryotic cells (1). To benchmark our IMAC workflow, we therefore first analysed a human cell line (JY cells) where we were expecting a relatively high number of phosphopeptides compared to *C. difficile*. From three biological replicates, we were able to identify a total of 10620 unique phosphopeptides with a good overlap between the different sample (**Supplemental Figure 2**). These results are in line with other recent studies (38, 39), and show that our approach allows the reliable identification of phosphopeptides in complex samples.

Next, we analysed the phosphoproteome of *C. difficile* at three different time points (mid-exponential phase, start stationary phase and late stationary phase (24 hr)) from three independently grown cultures. Overall, this resulted in the identification of almost 3000 phosphopeptides (**Supplemental Table 1**). Of these, 1759 had a high localisation probability (LP; above 0.95) for the correct assignment of the phosphorylation site. These sites were found on 1604 unique tryptic peptides derived from more than 700 proteins. Within this set of peptides, 75% was found to be phosphorylated on a serine, 20% on a threonine residues and 5% on a tyrosine (**Figure 1A**). Based on these results, we conclude that serine phosphorylation is most common in *C. difficile*, as also observed for other species (4, 11).

To compare the different samples, we selected unique peptide sequences that were found to be phosphorylated, without considering phospho-isomers (i.e. phosphorylation at different sites in the same peptide). In general, we observed a progressive increase in protein phosphorylation during growth. We identified most phosphopeptides at later growth phases

with the most prominent increase between the samples taken at the onset of stationary growth phase and the samples from the 24 hr cultures (**Figure 1B**).

We compared the reproducibility of the identifications between the different samples at the various timepoints. At the mid-exponential growth phase 232 phosphopeptides were identified in all three samples, corresponding to 41% of the total number of peptides identified at this timepoint (**Figure 1C**). For the second time point, 191 phosphopeptides were found in all three samples (20% of the total number identified at this time point). This relatively low percentage was mainly due to one of the samples, in which we found a lower number of phosphopeptides (**Figure 1B**, Culture 3) as the overlap between phosphopeptides was much higher between the other two cultures at this time point (63% of the total number). Of note, this was not due to experimental variation as analysis of a second sample from the same culture at the same time point showed a similar low number of phosphopeptides. After 24h, 1422 phosphopeptides were common to all three samples (66% of the total number identified at this time point) Overall, this shows significant congruence between our phosphopeptide identifications, despite individual processing of all three biological replicates and timepoints.

Within one culture, most phosphopeptides observed at the earlier time points were also observed at later stages (as exemplified for Culture 2 in **Figure 1D**), suggesting a gradual increase in the repertoire of phosphoproteins rather than a large scale reprogramming. This is also clear from the fact that 155 phosphopeptides were found in all samples at all time points. Nevertheless, we also observed a limited set of proteins that appear to be phosphorylated predominantly, if not exclusively, in the exponential growth phase (**Supplemental Table 1**) and are absent from late stationary growth phase samples. One such example is discussed further below.

We performed protein/gene set enrichment analysis of the phosphorylated proteins identified and determined the molecular functions and biological processes associated with these proteins. **Figure 2** represents the enrichment maps for each of the three time points, providing an overview by aggregating the enriched terms according to the shared protein set. Comparing the enrichment maps side by side revealed that during the first two stages there were concentrated, interconnected clusters of functions/processes, i.e. performed by same set of proteins; a trend that was lost in the late stationary growth phase. At the mid-exponential growth phase, these clusters are largely associated with metabolic and biosynthesis processes including few related to phosphorylation, as well as translation-associated functions and processes. At the beginning of the stationary growth phase, additional functions/processes start to be visible, including those related to ion/cation channel activities, ATP related processes, as well as cell motility and carbohydrate metabolic processes. At the late stationary growth phase, multiple parallel-unconnected functions and processes like DNA repair, transferases, catalytic activities were enriched, but also nucleotide biosynthetic processes that were already visible at the earlier time points. Of note, only the top 30 terms are shown in these maps, i.e. the functions and processes that are visible on top in the early stages likely are still active at the late stage, but they do not appear in the top 30.

Phosphorylation related to annotated kinases

We used the UniProt annotation of the *C. difficile* 630 genome (search with "cd630+kinase" in June 2019, 182 hits) to identify potential kinases and evaluate auto-phosphorylation or phosphorylation of substrates of these kinases on the basis of findings in other organisms. Here, we focus on the putative protein kinases that are expected to phosphorylate serine,

threonine and tyrosine residues: SpoIIAB (CD630_07710)(32), RsbW (CD630_00100)(31), HPr kinase (CD630_34090)(30), PrkC/Stk (CD630_25780)(33) and CD630_21480. Other kinases identified by their annotation phosphorylate small molecules (nucleotides, metabolites), or are part of two-component regulatory systems and fall outside the scope of the present study.

SpoIIAB is an anti-sigma-factor which in other bacteria is known to phosphorylate the anti-sigma factor F antagonist SpoIIAA during the process of sporulation (26, 32). In line with a role during sporulation, SpoIIAA phosphopeptides were exclusively identified in the 24 hr sample, but not in samples taken at mid-exponential or early stationary growth phase. Phosphorylation of SpoIIAA was found on Ser56 (LP=0.83), Ser57 (LP=0.55), Ser77 (LP=1.00), Ser82 (LP=0.84), Ser83 (LP=0.97) (**Supplemental Table 1**). Of these, Ser56 corresponds to the equivalent serine of *B. subtilis* SpoIIAA (Ser58) which is known to be phosphorylated by SpoIIAB in that organism (40). The corresponding peptide was found in all *C. difficile* cultures after 24 hrs. Manual inspection of the MS/MS spectrum **(Figure 3A)** confirmed phosphorylation of Ser-56 (corresponding to Ser14 in the tryptic peptide), because we observed neutral loss of phosphoric acid starting from the y_{10}.

We also found phosphorylation of the kinase SpoIIAB itself (**Supplemental Table 1**): Ser13 (LP=0.98), Ser17 (LP=1.0) and Ser106 (LP=1.00). Lower confidence sites in SpoIIAB include Ser123 (LP=0.50), Ser124 (LP=0.49) and Thr111 (LP=0.62). Of all these sites, only Ser106 was found consistently in all three 24hr cultures and the assignment of the site was confirmed by manual inspection of the MS/MS data (**Supplemental Figure 3A)**.

RsbW is the anti-sigma factor for σ^B, an important regulator of the stress response in gram-positive bacteria (27, 31). The interaction of RsbW with σ^B is inhibited through the binding of the anti-σ-factor antagonist RsbV, to RsbW (27, 31). The interaction between RsbV and RsbW is negative regulated by RsbW-dependent phosphorylation of RsbV (27). We found

two high-LP phosphopeptides for RsbV, indicative of phosphorylation of Ser57 (LP=0.93) and Ser84 (LP=1.00) and a low-LP phosphorylation of Thr58 (LP=0.56). Work in *B. subtilis* has identified three phosphorylation events and two of these (Ser 56 and Thr57) are on the residues equivalent to the ones identified here in *C. difficile* RsbV (15). On the basis of manual inspection of the MS/MS data, we provide strong evidence for phosphorylation of Ser57 in *C. difficile* RsbV **(Figure 3B).** We also identified a significant number of phosphopeptides derived from the kinase RsbW itself, in most cultures, at the 24h timepoint (**Supplemental Table 1**): Ser87 (LP=0.87), Ser89 (LP=0.83), Thr90 (LP=0.79) and Ser 99 (LP=1.00).

HPr kinase (HprK/ScoC) can phosphorylate the phosphocarrier protein HPr (PtsH), a protein involved in the import of monosaccharides as part of the PTS system (5). In *B. subtilis*, HPr kinase-mediated phosphorylation of HPr occurs on Ser46 (41, 42). We found phosphorylation of the homologous residue (Ser45, LP=1.00) of *C. difficile* HPr in all samples, and additionally identified high confidence phosphorylation of Ser11 (LP=1.00), and lower confidence phosphorylation of Ser30 (LP=0.85) and Thr31 (LP=0.82) (**Supplemental Table 1**). The equivalent site of Ser11 in *B. subtilis* HPr (Ser12) was also identified in a phosphoproteomics experiment (15). Importantly, the corresponding tryptic peptide of *C. difficile* HPr also contains a conserved histidine (His14), which is known to be phosphorylated through the phosphotransferase activity of the PTS-system (5, 43). However, the MS/MS spectrum of the tryptic peptide clearly demonstrates neutral losses of phosphoric acid from the b_3-ion and a series of non-phosphorylated y-ions until y_{12}. Based on this, and the absence of the characteristic pHis immonium ion in the MS/Ms spectrum, we conclude that Ser12 rather than His14 of HPr is phosphoryalted (**Supplemental Figure 3B**). We also found a single phosphopeptide derived from the HprK kinase itself: Thr299 (LP=1.00) appears to be phosphorylated primarily in stationary growth phase, with most peptides identified after 24h.

Phosphorylation of the Hanks-type serine-threonine kinases *C. difficile* PrkC (33) and CD630_21480 is hitherto undescribed and their substrates are virtually uncharacterized. We did not identify phosphopeptides derived from CD630_2148, but found extensive phosphorylation of PrkC in the region between Ala158 and Lys183 (tryptic peptide AV**S**N**S**T**T**M**T**NIG**S**IIG**S**VHYFSPEQAK, LP>0.95 in **red** and 0.70<LP<0.95 in **bold**) throughout growth. Notably, the conserved residues Thr163 and Thr165 were also found to be phosphorylated in *B. subtilis* (44). Thr290 of *B. subtilis* PrkC, another residue that was identified in multiple studies to be phosphorylated (15, 44), is not conserved in *C. difficile*. However, two residues in the equivalent region of *C. difficile* PrkC (Thr287 and Thr302, both LP=1.00) were identified, suggesting that phosphorylation in this region is structurally conserved. We also found phosphopeptides that indicate phosphorylation close to the N-terminus of PrkC (Thr4 and Thr35, LP=1.00) (**Supplemental Table 1**).

A single substrate for PrkC has been characterized: CwlA (CD1135) (34). For this protein, phosphorylation was reported on Ser136 and Thr405, with the latter being specific for PrkC. In our analyses, we find multiple Ser-phosphorylated peptides, including Ser136 (LP=1.00), and Thr405 (LP=1.00) (**Supplemental Table 1**). Additionally, it appears that PrkC is involved in cell wall homeostasis and antimicrobial resistance based on phenotypes of a *prkC* knockout strain (33). The authors postulate that part of the effects could be mediated by CD630_12830, a homolog of IreB of *Enterococcus faecalis.* In this organism, IreB negatively controls cephalosporin resistance and is phosphorylated by the serine/threonine kinase IreK on a conserved N-terminal threonine residue (45). We confirmed phosphorylation of the equivalent residue (Thr8, LP=1.00) across all time points, in all samples (**Supplemental Table 1**).

Dynamic phosphorylation within a protein

PrkC may also regulate cell division, such as DivIVA (33, 46-49). We identified high-confidence phosphopeptides derived from *C. difficile* DivIVA (CD630_26190), indicating phosphorylation of Ser75 (LP=1.00), Ser82 (LP=1.00) and Thr160 (LP=0.98). The phosphorylation of DivIVA is one of the few examples of consistent phosphorylation at a specific site (Thr160) at early time points, but not after 24 hr of culturing (**Supplemental Table 1**). Interestingly, the cell division protein SepF (CD630_26220), which is encoded in the same gene cluster as *divIVA*, shows the same pattern of phosphorylation for Thr56 (LP=0.96). In contrast, the other sites in DivIVA (Ser75 and Ser82) were found to be consistently phosphorylated in stationary growth phase, but never at the mid-exponential phase. It therefore appears that the pattern of phosphorylation of DivIVA differs throughout growth. Exponential growth phase phosphorylation is not observed for all cell division proteins, as for instance ZapA (CD630_07910) and FtsZ (CD630_26460) are predominantly phosphorylated in stationary growth phase (**Supplemental Table 1**).

We were interested to see if more proteins demonstrate dynamic phosphorylation. Manual inspection of the list of phosphopeptides identified two hypothetical proteins that show different patterns of phosphorylation in exponential compared to stationary growth phase: CD630_27170 and CD630_28470 (**Supplemental Table 1**). Phosphorylation of CD630_27170 occurs exclusively on Ser573 (LP=1.00) in exponential growth phase, at Ser573 and Ser286 (LP=1.00) at the onset of stationary growth phase, and at multiple other sites at 24h. CD630_28470 is phosphorylated at Ser186 (LP=1.00) in exponential growth phase and onset of stationary phase but not at late stationary phase (**Supplemental Table 1**). Together, these examples indicate that post-translational modification within a particular protein in *C. difficile* can change in a growth phase dependent manner.

444

445 **Indirect detection of cysteine phosphorylation**

446 In addition to protein kinase-dependent phosphorylation, in which the phosphate group is
447 donated by ATP, bacteria can also use the phosphate group from phosphoenolpyruvate (PEP)
448 in a process catalysed by a phosphotransferase system (PTS). This is of particular relevance
449 for the import of monosaccharides into the cell, during which the sugars are converted to
450 phospho-sugars (5). The initial enzymes involved in this cascade of events (Enzyme I (PtsI) and
451 Hpr (PtsH) are phosphorylated at conserved histidine residues and do not have specificity for
452 the different monosaccharides. The subunits of the Enzyme II complex (EIIA, EIIB and EIIC)
453 provide the necessary specificity for different monosaccharides. In our phosphoproteomic
454 dataset, we found that many of these subunits of the different EII complexes are
455 phosphorylated at serine and threonine residues (**Supplemental Table 1**). Activation of the
456 mono-saccharide during uptake, however, requires the transfer of phosphate from a
457 conserved cysteine residue of the EIIA and EIIB subunits, respectively. We accidently put
458 carbamidomethylation of Cys as a variable, instead of a fixed, modification during one of our
459 database searches and serendipitously observed cysteine-containing peptides from EII
460 complex subunits that did not appear to be phosphorylated (as expected based on our
461 enrichment strategy) or carbamidomethylated (as expected from our proteomics workflow).
462 Of note, when we searched data from a HeLa tryptic digest, that we use as a standard for our
463 LC-MS/MS setup, with the same parameters, we find very few free cysteines (18 out of 2753
464 peptide spectrum matches of peptides containing a cysteine).

465 As an example of a tryptic peptide with a free cysteine in our *C. difficile* data, we
466 observed the tryptic peptide ILVA**C**GAGIATSTIV**C**DRVER from the PTS system IIB component
467 (CD630_10830, aa 4-24) containing two cysteines (in bold and underlined). In total more than

300 peptide spectral matches for this peptide were identified in all our LC-MS/MS analyses of *C. difficile* phospho-samples. In all these identifications, the second cysteine (aa 16 in the tryptic peptide) was found to be carbamidomethylated, while the first cysteine at position 5 was a free cysteine **(Supplemental Figure 4**). Importantly, the first cysteine corresponds to the active site cysteine that is transiently phosphorylated in the P-loop of EII subunits (50). We therefore hypothesize that the identification of such a peptide is due to the fact that these cysteines were phosphorylated during the initial steps of our workflow. This would allow for enrichment during the IMAC procedure and – if the phosphorylation persists during the reduction and alkylation steps, preclude carbamidomethylation. As no phosphorylated cysteine modifications were found in the LC/MS-MS, it appears that the phosphate group is lost during the later stages of our workflow (due to instability of the phosphorylated cysteine under our experimental conditions).

In conclusion, it appears that a one-step Fe-IMAC enrichment procedure can indirectly detect at least a subset of cysteine-phosphorylated proteins.

DISCUSSION

In this study we characterized the phosphoproteome of a laboratory strain of *C. difficile* at three different timepoints. We show that many proteins of *C. difficile* are phosphorylated in stationary growth phase, and an analysis of subset of the identified phosphoproteins indicates that at least some of the phosphorylation sites are conserved between phylogenetically distinct organisms.

Many different methods have been used to analyse bacterial phosphoproteomes (1, 2, 11). A crucial step in these analyses is the method used to enrich for phosphopeptides (2, 4). In our study, we employed Fe-IMAC as this has recently been shown to be a very effective way for analyzing bacterial phosphoproteomes (38, 39). The number of phosphoproteins identified here is in line with these studies, and indicates this is a suitable method for the analysis of phosphoproteins in *C. difficile*. We identified Ser-, Thr-, and Tyr-phosphorylation and provided indirect evidence for Cys-phosphorylation (**Figure 1A, Supplemental Table 1**). Our workflow does not allow the identification of His-phosphorylation, but recent studies have shown that the Fe-IMAC method can be adapted to detect the thermodynamically less stable His-phosphorylation for future experiments (51, 52). The identification of other forms of phosphorylation, such as Arg phosphorylation, requires different approaches (53). We also observed multiply phosphorylated proteins (**Supplemental Table 1**), as has seen in other organisms as well (12, 54-56).

Our work provides a starting point for experimental validation of the identified phosphoproteins, as well as a dissection of the contribution of individual protein kinases and phosphatases. Generally, the action of protein kinases is better understood than that of phosphatases (16). We identified site-specific phosphorylation in the protein kinase PrkC (**Supplemental Table 1**). A knockout of the gene encoding this kinase demonstrates

pleiotropic effects (33), and comparing wild type and *prkC* negative cells using our approach may help to elucidate the substrates of this protein.

The identified phosphopeptides also offer a possibility of characterizing the effect of phosphorylation on substrate proteins, by mutating residues in these proteins to non-phosphorylatable analogs (alanine for serine/threonine and phenylalanine for tyrosine) or phosphomimetic negatively charged amino acids (aspartate and glutamate) (2, 10). Such an approach may be preferable over mutating the kinases and phosphatases, as these can have overlapping specificities (57).

In previous phosphoproteomic analyses, a large divergence was often observed in the lists of phosphorylated proteins even within a single genus or species (2, 58, 59). In our experiments, we observed apparently conserved phosphorylation events on several proteins, including SpoIIA, SpoIIAB, RsbV and others (**Figure 3**, **Supplemental Table 1**), which may appear in contrast with these observations. Part of the reason for the limited overlap in previous experiments may have been the diversity in enrichment methods and experimental protocols, limited sensitivity of the workflows and variability in sampling method and timepoint. Our workflow allows the reliable identification of a large set of phosphoproteins in comparison with other studies (11, 12, 16, 17). Moreover, we applied rigorous quality control of phosphosite assignment. It is likely that developments in both mass spectrometry instrumentation, data acquisition and analysis pipelines will reveal that many more processes in fact may be regulated by conserved phosphorylation events. For example, a recent phosphoproteomic analysis of *Staphylococcus aureus* increased the number of phosphosite identifications over 20-fold compared to earlier analyses, thereby also increasing the number of processes affected (38). Cross-species comparisons may be facilitated by dedicated

530 databases of phosphorylated proteins, such as PHOSIDA and dbPSP 2.0 (60, 61) and ultimately
531 lead to improved prediction of phosphorylation in bacteria (62).

532 Our work indicates that protein phosphorylation in *C. difficile* may differ both spatially
533 (within a protein) and temporally (between timepoints). Specifically, we observe low levels of
534 phosphorylation in exponential growth phase, whereas phosphorylation increases markedly
535 upon entry into stationary growth phase. Of note, we observed a highly variable number of
536 phosphoproteins at the entry into logarithmic growth phase, with two out of three samples
537 showing clearly elevated levels, whereas a single biological replicate resembled the profile
538 observed in exponential growth phase (**Figure 1B**). This suggests that the phosphorylation
539 state could change rapidly, and supports the notion that sampling timepoint could explain
540 differences between phosphoproteomic datasets. We also report that specific
541 phosphorylation patterns can be observed within a protein, as exemplified by DivIVA and SepF
542 (**Supplemental Table 1**). This has also been observed for other organisms; for instance, in
543 *Streptomyces coelicolor*, 58/85 phosphorylation sites were differentially phosphorylated
544 during differentiation (13). Interestingly, Thr160 of *C. difficile* DivIVA is located in the putative
545 tetramerization region of the protein (63) and phosphorylation at the C-terminus of DivIVA
546 has also been observed for *S. pneumoniae*, where is it mediated by the eSTK protein StkP (49).
547 Mutations of this region lead to aberrant cell morphology, suggesting that phosphorylation in
548 of Thr160 in *C. difficile* might have functional consequences. Thr56 of *C. difficile* SepF is
549 directly adjacent to the predicted compact domain that mediates self-interaction and
550 interaction with the cell division protein FtsZ (64). The functional consequences of the
551 dynamic phosphorylation of DivIVA and SepF remain to be elucidated.

552 We focused our in-depth analyses on those proteins for which the phosphosite
553 assignment could be done with high probability; where necessary, we manually verified the

554 automatic identifications. Nevertheless, we also observed peptides for which it was not
555 possible to assign the modification to a particular residue, similar to others (59). For those
556 proteins, alternative fragmentation techniques (65) or processing of the sample – using for
557 instance proteases different than trypsin – might yield better results.

558 The phosphoproteome described here complements existing omics approaches for *C.*
559 *difficile*, such as genomics, transcriptomics, proteomics and metabolomics. Integration of
560 these data may lead to a systems-level understanding of *C. difficile* physiology and – more
561 broadly – may contribute to our understanding of the role of phosphorylation in the
562 regulation of bacterial pathogenesis (16).

563

564

Acknowledgements

The authors wish to acknowledge J. Corver for helpful discussions, and I. Martin-Verstraete for communicating results prior to publication.

Funding

This study was supported by the research program Investment Grant NWO Medium (project number 91116004), which is partially financed by ZonMw.

572 **REFERENCES**

573

574 1. Macek B, Forchhammer K, Hardouin J, Weber-Ban E, Grangeasse C, Mijakovic I. 2019. Protein
575 post-translational modifications in bacteria. Nat Rev Microbiol 17:651-664.
576 2. Mijakovic I, Macek B. 2012. Impact of phosphoproteomics on studies of bacterial physiology.
577 FEMS Microbiol Rev 36:877-92.
578 3. Kamino K, Keegstra JM, Long J, Emonet T, Shimizu TS. 2020. Adaptive tuning of cell sensory
579 diversity without changes in gene expression. Sci Adv 6.
580 4. Ge R, Shan W. 2011. Bacterial phosphoproteomic analysis reveals the correlation between
581 protein phosphorylation and bacterial pathogenicity. Genomics Proteomics Bioinformatics
582 9:119-27.
583 5. Deutscher J, Ake FM, Derkaoui M, Zebre AC, Cao TN, Bouraoui H, Kentache T, Mokhtari A,
584 Milohanic E, Joyet P. 2014. The bacterial phosphoenolpyruvate:carbohydrate
585 phosphotransferase system: regulation by protein phosphorylation and phosphorylation-
586 dependent protein-protein interactions. Microbiol Mol Biol Rev 78:231-56.
587 6. Pereira SF, Goss L, Dworkin J. 2011. Eukaryote-like serine/threonine kinases and phosphatases
588 in bacteria. Microbiol Mol Biol Rev 75:192-212.
589 7. Kannan N, Taylor SS, Zhai Y, Venter JC, Manning G. 2007. Structural and functional diversity of
590 the microbial kinome. PLoS Biol 5:e17.
591 8. Grangeasse C, Cozzone AJ, Deutscher J, Mijakovic I. 2007. Tyrosine phosphorylation: an
592 emerging regulatory device of bacterial physiology. Trends Biochem Sci 32:86-94.
593 9. Jadeau F, Bechet E, Cozzone AJ, Deleage G, Grangeasse C, Combet C. 2008. Identification of
594 the idiosyncratic bacterial protein tyrosine kinase (BY-kinase) family signature. Bioinformatics
595 24:2427-30.
596 10. Jers C, Soufi B, Grangeasse C, Deutscher J, Mijakovic I. 2008. Phosphoproteomics in bacteria:
597 towards a systemic understanding of bacterial phosphorylation networks. Expert Rev
598 Proteomics 5:619-27.
599 11. Macek B, Mijakovic I. 2011. Site-specific analysis of bacterial phosphoproteomes. Proteomics
600 11:3002-11.
601 12. Sun X, Ge F, Xiao CL, Yin XF, Ge R, Zhang LH, He QY. 2010. Phosphoproteomic analysis reveals
602 the multiple roles of phosphorylation in pathogenic bacterium Streptococcus pneumoniae. J
603 Proteome Res 9:275-82.
604 13. Rioseras B, Shliaha PV, Gorshkov V, Yague P, Lopez-Garcia MT, Gonzalez-Quinonez N,
605 Kovalchuk S, Rogowska-Wrzesinska A, Jensen ON, Manteca A. 2018. Quantitative Proteome
606 and Phosphoproteome Analyses of Streptomyces coelicolor Reveal Proteins and
607 Phosphoproteins Modulating Differentiation and Secondary Metabolism. Mol Cell Proteomics
608 17:1591-1611.
609 14. Macek B, Gnad F, Soufi B, Kumar C, Olsen JV, Mijakovic I, Mann M. 2008. Phosphoproteome
610 analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation.
611 Mol Cell Proteomics 7:299-307.
612 15. Macek B, Mijakovic I, Olsen JV, Gnad F, Kumar C, Jensen PR, Mann M. 2007. The
613 serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis. Mol
614 Cell Proteomics 6:697-707.
615 16. Bonne Kohler J, Jers C, Senissar M, Shi L, Derouiche A, Mijakovic I. 2020. Importance of protein
616 Ser/Thr/Tyr phosphorylation for bacterial pathogenesis. FEBS Lett doi:10.1002/1873-
617 3468.13797.
618 17. Yague P, Gonzalez-Quinonez N, Fernanez-Garcia G, Alonso-Fernandez S, Manteca A. 2019.
619 Goals and Challenges in Bacterial Phosphoproteomics. Int J Mol Sci 20.

18. Lin MH, Potel CM, Tehrani K, Heck AJR, Martin NI, Lemeer S. 2018. A New Tool to Reveal Bacterial Signaling Mechanisms in Antibiotic Treatment and Resistance. Mol Cell Proteomics 17:2496-2507.
19. Bai X, Ji Z. 2012. Phosphoproteomic investigation of a solvent producing bacterium Clostridium acetobutylicum. Appl Microbiol Biotechnol 95:201-11.
20. Smits WK, Lyras D, Lacy DB, Wilcox MH, Kuijper EJ. 2016. Clostridium difficile infection. Nat Rev Dis Primers 2:16020.
21. Abt MC, McKenney PT, Pamer EG. 2016. Clostridium difficile colitis: pathogenesis and host defence. Nat Rev Microbiol 14:609-20.
22. Leffler DA, Lamont JT. 2015. Clostridium difficile infection. N Engl J Med 372:1539-48.
23. Khoruts A, Staley C, Sadowsky MJ. 2021. Faecal microbiota transplantation for Clostridioides difficile: mechanisms and pharmacology. Nat Rev Gastroenterol Hepatol 18:67-80.
24. Baktash A, Terveer EM, Zwittink RD, Hornung BVH, Corver J, Kuijper EJ, Smits WK. 2018. Mechanistic Insights in the Success of Fecal Microbiota Transplants for the Treatment of Clostridium difficile Infections. Front Microbiol 9:1242.
25. Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ, West MR, Wren BW, Fairweather NF, Dougan G, Lawley TD. 2012. The Clostridium difficile spo0A gene is a persistence and transmission factor. Infect Immun 80:2704-11.
26. Hoch JA. 1993. Regulation of the phosphorelay and the initiation of sporulation in Bacillus subtilis. Annu Rev Microbiol 47:441-65.
27. van Schaik W, Abee T. 2005. The role of sigmaB in the stress response of Gram-positive bacteria -- targets for food preservation and safety. Curr Opin Biotechnol 16:218-24.
28. Martin-Verstraete I, Peltier J, Dupuy B. 2016. The Regulatory Networks That Control Clostridium difficile Toxin Synthesis. Toxins (Basel) 8.
29. Pompeo F, Foulquier E, Galinier A. 2016. Impact of Serine/Threonine Protein Kinases on the Regulation of Sporulation in Bacillus subtilis. Front Microbiol 7:568.
30. Antunes A, Martin-Verstraete I, Dupuy B. 2011. CcpA-mediated repression of Clostridium difficile toxin gene expression. Mol Microbiol 79:882-99.
31. Kint N, Alves Feliciano C, Hamiot A, Denic M, Dupuy B, Martin-Verstraete I. 2019. The sigma(B) signalling activation pathway in the enteropathogen Clostridioides difficile. Environ Microbiol 21:2852-2870.
32. Pereira FC, Saujet L, Tome AR, Serrano M, Monot M, Couture-Tosi E, Martin-Verstraete I, Dupuy B, Henriques AO. 2013. The spore differentiation pathway in the enteric pathogen Clostridium difficile. PLoS Genet 9:e1003782.
33. Cuenot E, Garcia-Garcia T, Douche T, Gorgette O, Courtin P, Denis-Quanquin S, Hoys S, Tremblay YDN, Matondo M, Chapot-Chartier MP, Janoir C, Dupuy B, Candela T, Martin-Verstraete I. 2019. The Ser/Thr Kinase PrkC Participates in Cell Wall Homeostasis and Antimicrobial Resistance in Clostridium difficile. Infect Immun 87.
34. Garcia-Garcia T, Poncet S, Cuenot E, Douché T, Gianetto QG, Peltier J, Courtin P, Chapot-Chartier MP, Matondo M, Dupuy B, Candela T, Martin-Verstraete I. 2020. Ser/Thr kinase-dependent phosphorylation of the peptidoglycan hydrolase CwlA controls its export and modulates cell division in Clostridioides difficile. bioRxiv doi:10.1101/2020.10.29.360313.
35. Hussain HA, Roberts AP, Mullany P. 2005. Generation of an erythromycin-sensitive derivative of Clostridium difficile strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. J Med Microbiol 54:137-141.
36. van Eijk E, Anvar SY, Browne HP, Leung WY, Frank J, Schmitz AM, Roberts AP, Smits WK. 2015. Complete genome sequence of the Clostridium difficile laboratory strain 630Deltaerm reveals differences from strain 630, including translocation of the mobile element CTn5. BMC Genomics 16:31.

37. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Perez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaino JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res 47:D442-D450.
38. Prust N, van der Laarse S, van den Toorn H, van Sorge NM, Lemeer S. 2021. In depth characterization of the Staphylococcus aureus phosphoproteome reveals new targets of Stk1. Mol Cell Proteomics doi:10.1074/mcp.RA120.002232:100034.
39. Potel CM, Lin MH, Heck AJR, Lemeer S. 2018. Defeating Major Contaminants in Fe(3+)-Immobilized Metal Ion Affinity Chromatography (IMAC) Phosphopeptide Enrichment. Mol Cell Proteomics 17:1028-1034.
40. Najafi SM, Willis AC, Yudkin MD. 1995. Site of phosphorylation of SpoIIAA, the anti-anti-sigma factor for sporulation-specific sigma F of Bacillus subtilis. J Bacteriol 177:2912-3.
41. Boel G, Mijakovic I, Maze A, Poncet S, Taha MK, Larribe M, Darbon E, Khemiri A, Galinier A, Deutscher J. 2003. Transcription regulators potentially controlled by HPr kinase/phosphorylase in Gram-negative bacteria. J Mol Microbiol Biotechnol 5:206-15.
42. Poncet S, Mijakovic I, Nessler S, Gueguen-Chaignon V, Chaptal V, Galinier A, Boel G, Maze A, Deutscher J. 2004. HPr kinase/phosphorylase, a Walker motif A-containing bifunctional sensor enzyme controlling catabolite repression in Gram-positive bacteria. Biochim Biophys Acta 1697:123-35.
43. Schrecker O, Stein R, Hengstenberg W, Gassner M, Stehlik D. 1975. The staphylococcal PEP dependent phosphotransferase system, proton magnetic resonance (PMR) studies on the phosphoryl carrier protein HPr: evidence for a phosphohistidine residue in the intact phospho-HPr molecule. FEBS Lett 51:309-12.
44. Madec E, Stensballe A, Kjellstrom S, Cladiere L, Obuchowski M, Jensen ON, Seror SJ. 2003. Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from Bacillus subtilis. J Mol Biol 330:459-72.
45. Hall CL, Tschannen M, Worthey EA, Kristich CJ. 2013. IreB, a Ser/Thr kinase substrate, influences antimicrobial resistance in Enterococcus faecalis. Antimicrob Agents Chemother 57:6179-86.
46. Ni H, Fan W, Li C, Wu Q, Hou H, Hu D, Zheng F, Zhu X, Wang C, Cao X, Shao ZQ, Pan X. 2018. Streptococcus suis DivIVA Protein Is a Substrate of Ser/Thr Kinase STK and Involved in Cell Division Regulation. Front Cell Infect Microbiol 8:85.
47. Kang CM, Abbott DW, Park ST, Dascher CC, Cantley LC, Husson RN. 2005. The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. Genes Dev 19:1692-704.
48. Novakova L, Bezouskova S, Pompach P, Spidlova P, Saskova L, Weiser J, Branny P. 2010. Identification of multiple substrates of the StkP Ser/Thr protein kinase in Streptococcus pneumoniae. J Bacteriol 192:3629-38.
49. Fleurie A, Cluzel C, Guiral S, Freton C, Galisson F, Zanella-Cleon I, Di Guilmi AM, Grangeasse C. 2012. Mutational dissection of the S/T-kinase StkP reveals crucial roles in cell division of Streptococcus pneumoniae. Mol Microbiol 83:746-58.
50. Volpon L, Young CR, Matte A, Gehring K. 2006. NMR structure of the enzyme GatB of the galactitol-specific phosphoenolpyruvate-dependent phosphotransferase system and its interaction with GatA. Protein Sci 15:2435-41.
51. Potel CM, Lin MH, Heck AJR, Lemeer S. 2018. Widespread bacterial protein histidine phosphorylation revealed by mass spectrometry-based proteomics. Nat Methods 15:187-190.
52. Potel CM, Lin MH, Prust N, van den Toorn HWP, Heck AJR, Lemeer S. 2019. Gaining Confidence in the Elusive Histidine Phosphoproteome. Anal Chem 91:5542-5547.

53. Trentini DB, Fuhrmann J, Mechtler K, Clausen T. 2014. Chasing Phosphoarginine Proteins: Development of a Selective Enrichment Method Using a Phosphatase Trap. Mol Cell Proteomics 13:1953-64.
54. Ge R, Sun X, Xiao C, Yin X, Shan W, Chen Z, He QY. 2011. Phosphoproteome analysis of the pathogenic bacterium Helicobacter pylori reveals over-representation of tyrosine phosphorylation and multiply phosphorylated proteins. Proteomics 11:1449-61.
55. Soufi B, Gnad F, Jensen PR, Petranovic D, Mann M, Mijakovic I, Macek B. 2008. The Ser/Thr/Tyr phosphoproteome of Lactococcus lactis IL1403 reveals multiply phosphorylated proteins. Proteomics 8:3486-93.
56. Rosen R, Becher D, Buttner K, Biran D, Hecker M, Ron EZ. 2004. Highly phosphorylated bacterial proteins. Proteomics 4:3068-77.
57. Prisic S, Dankwa S, Schwartz D, Chou MF, Locasale JW, Kang CM, Bemis G, Church GM, Steen H, Husson RN. 2010. Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. Proc Natl Acad Sci U S A 107:7521-6.
58. Soufi B, Jers C, Hansen ME, Petranovic D, Mijakovic I. 2008. Insights from site-specific phosphoproteomics in bacteria. Biochim Biophys Acta 1784:186-92.
59. Ravichandran A, Sugiyama N, Tomita M, Swarup S, Ishihama Y. 2009. Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic Pseudomonas species. Proteomics 9:2764-75.
60. Shi Y, Zhang Y, Lin S, Wang C, Zhou J, Peng D, Xue Y. 2020. dbPSP 2.0, an updated database of protein phosphorylation sites in prokaryotes. Sci Data 7:164.
61. Gnad F, Gunawardena J, Mann M. 2011. PHOSIDA 2011: the posttranslational modification database. Nucleic Acids Res 39:D253-60.
62. Miller ML, Soufi B, Jers C, Blom N, Macek B, Mijakovic I. 2009. NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. Proteomics 9:116-25.
63. van Baarle S, Celik IN, Kaval KG, Bramkamp M, Hamoen LW, Halbedel S. 2013. Protein-protein interaction domains of Bacillus subtilis DivIVA. J Bacteriol 195:1012-21.
64. Duman R, Ishikawa S, Celik I, Strahl H, Ogasawara N, Troc P, Lowe J, Hamoen LW. 2013. Structural and genetic analyses reveal the protein SepF as a new membrane anchor for the Z ring. Proc Natl Acad Sci U S A 110:E4601-10.
65. Frese CK, Zhou H, Taus T, Altelaar AF, Mechtler K, Heck AJ, Mohammed S. 2013. Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (EThcD). J Proteome Res 12:1520-5.

LEGENDS

Figure 1. Global overview of phosphoproteins in *C. difficile*. A. Pie chart indicating the percentages of Ser, Thr and Tyr phosphorylation. **B**. Number of identified phosphopeptides over time for the three biological replicates analysed. For the definition of timepoints, see Supplemental Figure 1. **C**. Concordance between the phosphopeptide identifications between the three biological replicates. **D**. Changes in the identified phosphoproteome over time. Overlap in the identified phosphopeptides per timepoint for culture 2 is indicated in a Venn diagram.

Figure 2. Enrichment maps of Gene Ontology terms obtained from *C. difficile* phosphoprotein set enrichment analysis. Three enrichment maps correspond to the three time points analyzed in this study (mid-exponential, beginning stationary, and late stationary growth phase); each map is limited to the top 30 enriched terms. The leaves represent Gene Ontology terms and the edges correspond to the Jaccard similarity between the leaves based on the shared proteins.

Figure 3. Spectra of conserved phosphorylation in regulatory proteins of *C. difficile*. A. MS/MS spectrum of the tryptic phosphopeptide NVVFNFENINFMDpSSGIGVIIGR (pS=phosphoserine) from SpoIIAA, demonstrating phosphorylation of Ser-14 (Ser-56 in the full length protein, indicated in red) **B**. MS/MS spectrum of the tryptic phosphopeptide DLDYIDpSTGLGAMIGVLKK (pS=phosphoserine) from RsbV, demonstrating phosphorylation of Ser-7 (Ser-57 in the full length protein, indicated in red). Ions assigned as "-P" have lost H_3PO_4

779 **Supplemental Figure 1. Overview of growth and sampling of the *C. difficile* cultures**. **A**.
780 Growth curve indicating obtained optical densities at 600nm and sampling timepoints for this
781 study. **B**. Table summarizing protein yield from 100 mL (time point 1) and 50 mL (time points
782 2 and 3) of sample.
783
784 **Supplemental Figure 2**. Summary of total phosphopeptides retrieved from JY cells. Overlap
785 between the three biological replicates is represented with the Venn diagram.
786
787 **Supplemental Figure 3**. **Spectra of conserved phosphorylation in regulatory proteins of *C.***
788 ***difficile*.** A. MS/MS spectrum of the tryptic phosphopeptide AMEPLYTSKPELDRpSGMGFTVMK)
789 from SpoIIAB, demonstrating phosphorylation of Ser-15 (Ser-106 in the full length protein,
790 indicated in red) B. MS/MS spectrum of the tryptic phosphopeptide NApSGLHARPAGMFVK
791 from HPr, demonstrating phosphorylation of Ser-3 (Ser-11 in the full length protein, indicated
792 in red). pS=phosphoserine. Ions assigned as "-P" have lost H_3PO_4
793
794 **Supplemental Figure 4. Indirect identification of cysteine phosphorylation**
795 MS/MS spectrum of the tryptic peptide LVACGAGIATSTIVC_{cam}DRVER from the
796 phosphotransfer system (PTS) component IIB (CD630_10830) in which the first cysteine (in
797 red) is a free cysteine while the second is carbamidomethylated (cam).
798
799 **Supplemental Table 1.** Full data set of the phosphoproteome analysis and site assignment for
800 Ser, Thr and Tyr phosphorylation in *C. difficile*.
801
802

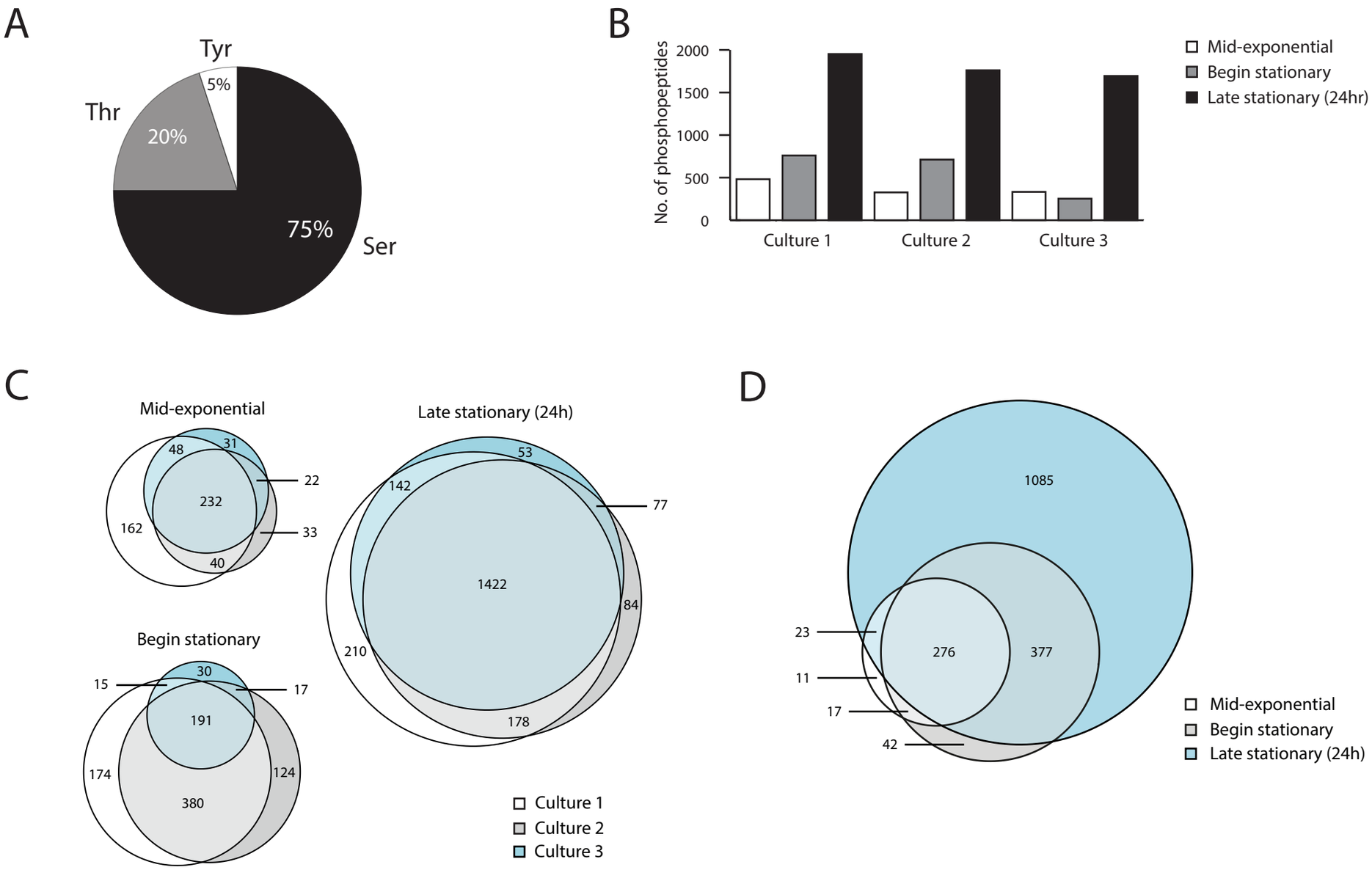
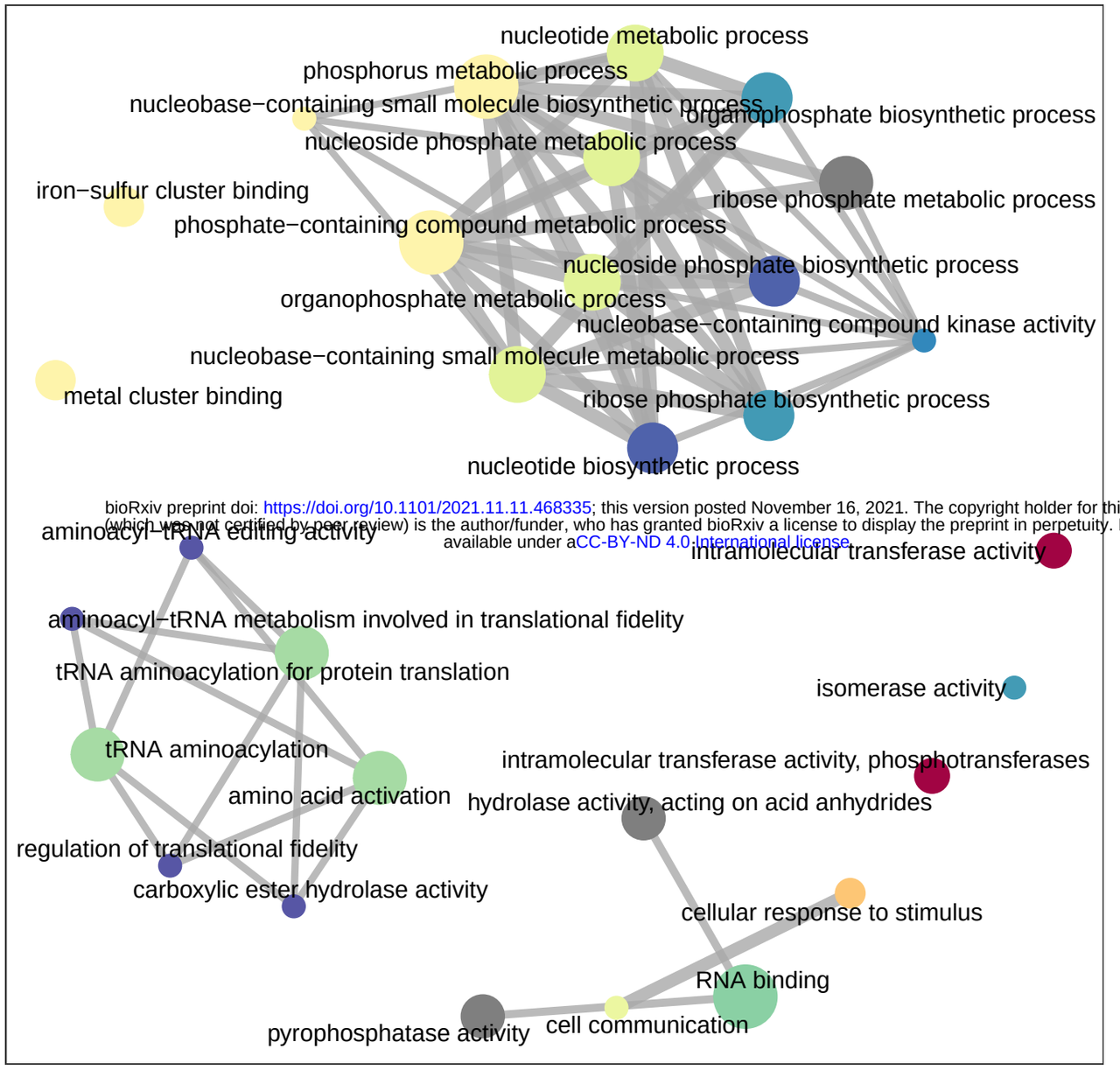
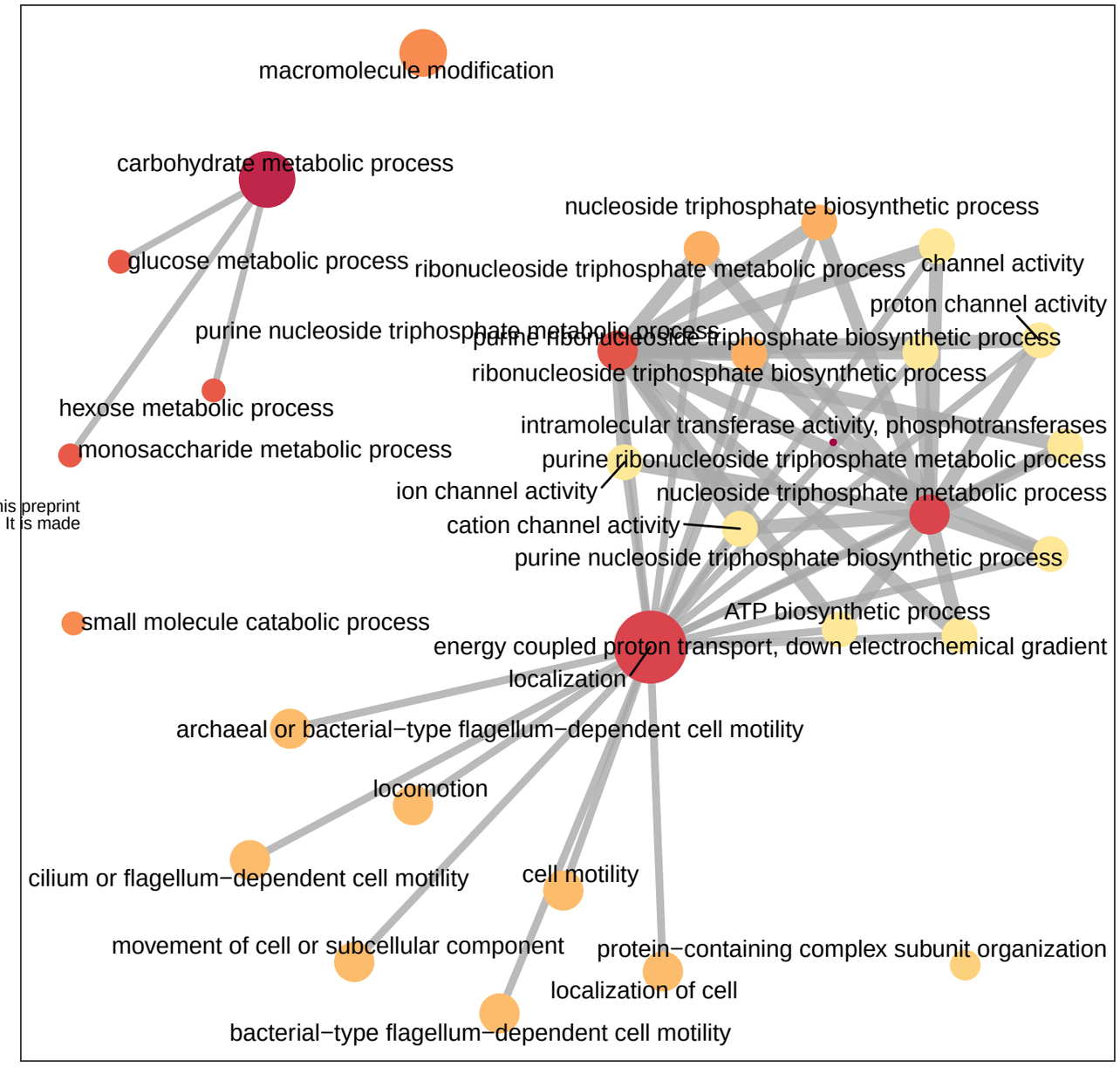

A
Tyr
5%
Thr
20%
75%
Ser
B
No. of phosphopeptides
2000
1500
1000
500
0
Culture 1
Culture 2
Culture 3
Mid-exponential
Begin stationary
Late stationary (24hr)
C
Mid-exponential
48
31
22
232
162
33
40
Late stationary (24h)
53
142
77
1422
84
210
178
Begin stationary
15
30
17
191
174
124
380
Culture 1
Culture 2
Culture 3
D
1085
23
276
377
11
17
42
Mid-exponential
Begin stationary
Late stationary (24h)

Figure 1

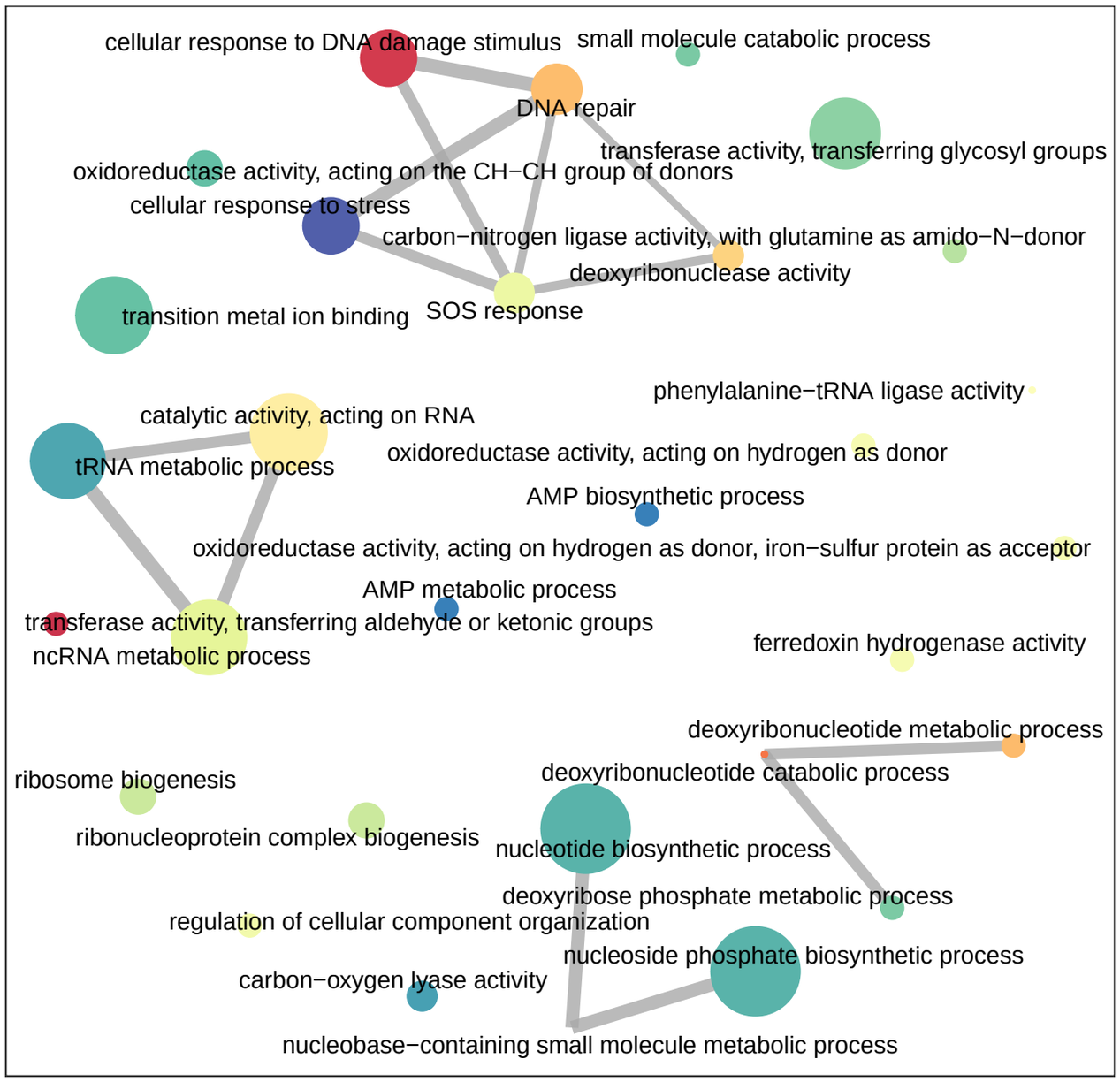
mid−exponential



begin stationary



late stationary



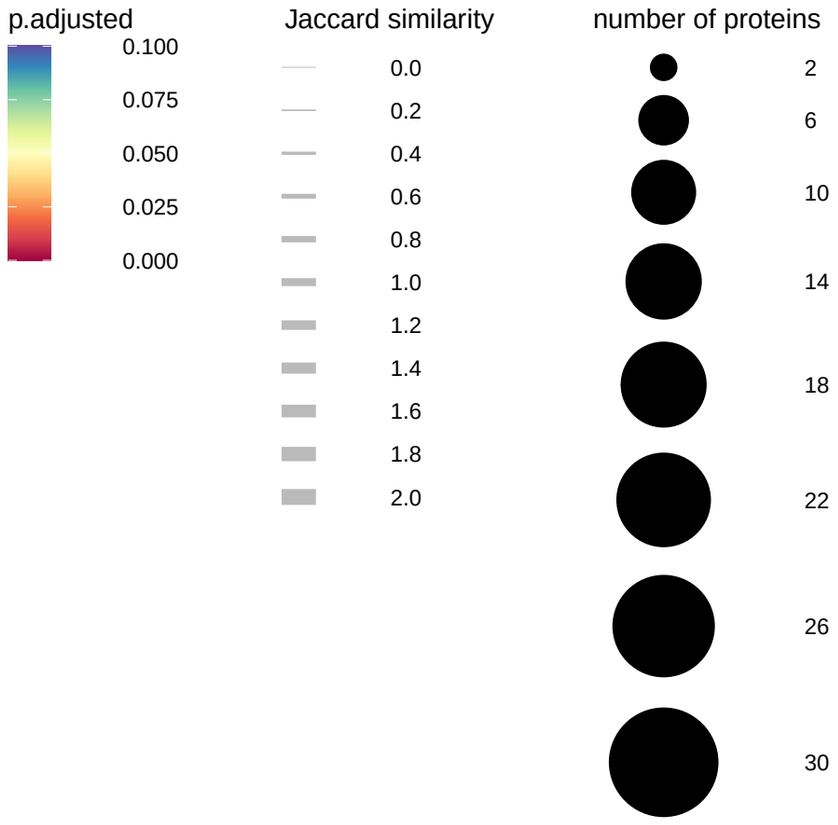
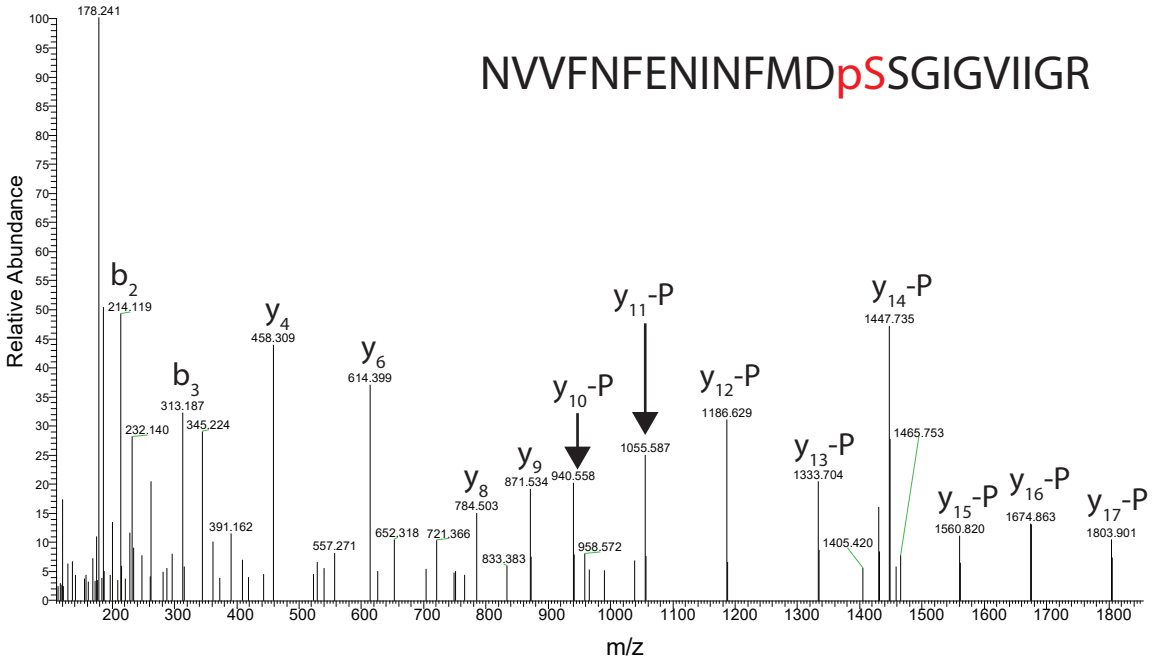


Figure 2

A

>Q18C94:Anti-sigma F factor antagonist (SpoIIAA)

MVNYSLEHKNLYIEFMCSELDHHVANEIRELIDNLLSVNQVKNVVFNFENINFMDSSGIGVIIGRYKKISNEGG
RVSVINISSRVKKIFDLSGLNKIIGIYDTYEEALSSL



B

>Q18C94:Anti-sigma factor antagonist(RsbV)

MNIDSNLDSQNKFWNVCLDGELDVSTADKLKEHLHALIEKNMLDVKINLKDLDYIDSTGL
GAMIGVLKKLKINEKEIYIVNPKSNVRKIFTITGLDKIFKVEG

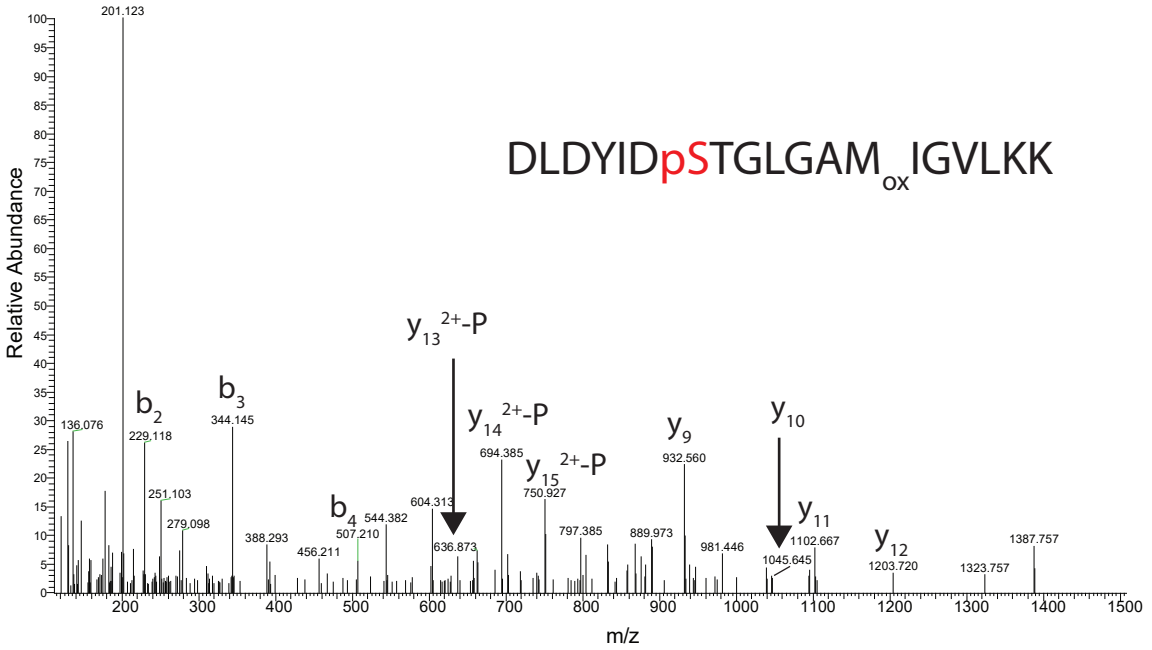


Figure 3