SUPPLEMENTARY INFORMATION

Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium.

Vera van Noort[1]‡, Jan Seebacher[1]‡, Samuel Bader[1]‡, Shabaz Mohammed[2], Ivana Vonkova[1], Matthew J. Betts[3], Sebastian Kühner[1], Runjun Kumar[1], Tobias Maier [4], Martina O'Flaherty[2], Vladimir Rybin[1], Arne Schmeisky[5], Eva Yus[4], Jörg Stülke[5], Luis Serrano[4,6], Robert B. Russell[3], Albert J.R. Heck[2], Peer Bork[1]* and Anne-Claude Gavin[1]*

[1]Structural and Computational Biology Unit, European Molecular Biology Laboratory, EMBL, Heidelberg, Germany.

[2]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Institute for Pharmaceutical Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands.

[3]Cell Networks, University of Heidelberg, D-69120 Heidelberg, Germany.

[4]EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG) and UPF, Dr. Aiguader 88, 08003 Barcelona, Spain.

[5]Department of General Microbiology, Georg-August University Göttingen, D-37077 Göttingen, Germany.

[6]ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain.

‡ These authors contributed equally to the work

* *To whom correspondence should be addressed.* E-mail: gavin@embl.de and bork@embl.de

Table of Content

Supplementary text

Among the proteins that show complex patterns of PTMs, we found many structural proteins and chaperones (Fig.2B and Supplementary Table 1). For example, the translational GTPase EF-Tu (Tuf, Mpn665) associates with many protein complexes in *M. pneumoniae* (Kuhner et al, 2009) and has been proposed to act as a cytoskeletal factor involved in cell shape maintenance (Defeu Soufo et al, 2010) and adhesion to extracellular matrix of epithelial cells (Balasubramanian et al, 2009). We found two phosphorylation and 12 lysine acetylations for EF-Tu (Fig. 2C and Supplementary Table 1), including a conserved Thr383 located within its carboxy-terminal domain that is required for adhesion. Such extensive pattern of modification may account for EF-Tu multi-functionality and may be important for *M. pneumoniae* pathogenicity.

Eight of nine protein chaperones in *M. pneumoniae*, were lysine acetylated and four carried additional phosphorylation sites (Fig. 2A; COG class O). For example eight acetyl lysines were characterized for ClpB that localize in the substrate-binding (K286, K323, K368) and -discriminating (K28) domains and two regions, NBD1 (K191) and NBD2 (K417, K561), responsible for ClpB hexamer oligomerization. This suggests that lysine acetylation exerts pleiotropic effect on chaperonin activity, reminiscent of the complex regulation observed for the eukaryotic chaperone Hsp90 (Scroggins et al, 2007).

Four proteins of previously unknown functions, Mpn256, Mpn387, Mpn400 and Mpn454, cluster together with the sets of cytadherence proteins that are downregulated upon PknB (Mpn248) deletion (Supplementary Fig. 5). Consistent with a possible role in cell adhesion, both Mpn400 and Mpn454 have predicted transmembrane domains. Furthermore, *mpn454* belongs to the same operon as the known cytadherence proteins Hmw3 (Mpn452) and P30 (Mpn453) and Mpn400 was found to copurify in the Triton X100 insoluble fraction together with other cytoskeletal proteins (Regula et al, 2001). Also, genetic studies suggested a role for Mpn387 in maintenance of the attachment organelle and gliding motility(Hasselbring et al, 2006).

Supplementary references

Balasubramanian S, Kannan TR, Hart PJ, Baseman JB (2009). Amino acid changes in elongation factor Tu of Mycoplasma pneumoniae and Mycoplasma genitalium influence fibronectin binding. Infect Immun 77: 3533-3541

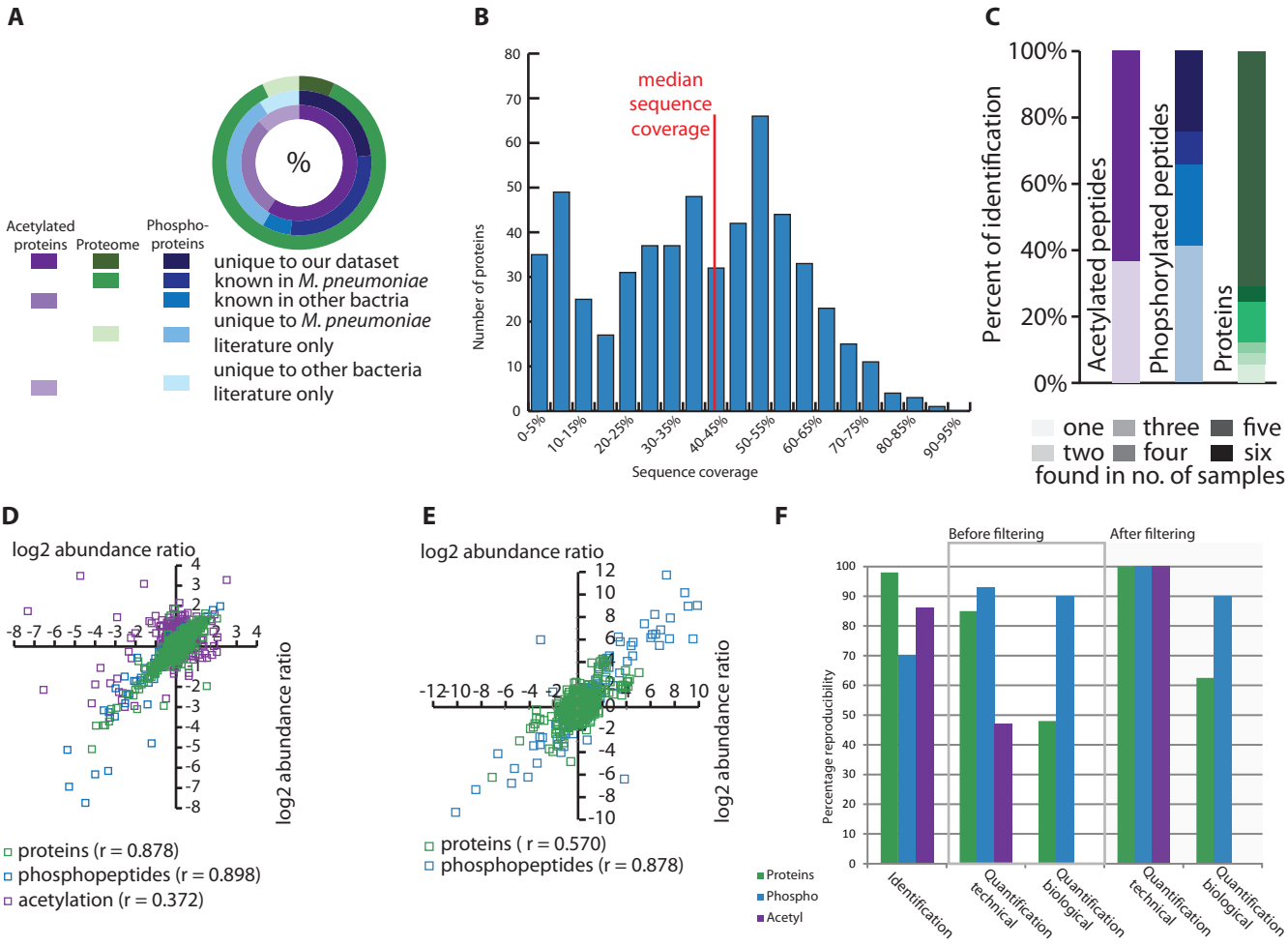
Defeu Soufo HJ, Reimold C, Linne U, Knust T, Gescher J, Graumann PL (2010). Bacterial translation elongation factor EF-Tu interacts and colocalizes with actin-like MreB protein. Proc Natl Acad Sci U S A 107: 3163-3168

Hasselbring BM, Page CA, Sheppard ES, Krause DC (2006). Transposon mutagenesis identifies genes associated with Mycoplasma pneumoniae gliding motility. J Bacteriol 188: 6335-6345

Kuhner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, Rode M, Yamada T, Maier T, Bader S, Beltran-Alvarez P, Castano-Diez D, Chen WH, Devos D, Guell M, Norambuena T, Racke I, Rybin V, Schmidt A, Yus E, Aebersold R, et al. (2009). Proteome organization in a genome-reduced bacterium. Science 326: 1235-1240

Regula JT, Boguth G, Gorg A, Hegermann J, Mayer F, Frank R, Herrmann R (2001). Defining the mycoplasma 'cytoskeleton': the protein composition of the Triton X-100 insoluble fraction of the bacterium Mycoplasma pneumoniae determined by 2-D gel electrophoresis and mass spectrometry. Microbiology 147: 1045-1057

Scroggins BT, Robzyk K, Wang D, Marcu MG, Tsutsumi S, Beebe K, Cotter RJ, Felts S, Toft D, Karnitz L, Rosen N, Neckers L (2007). An acetylation site in the middle domain of Hsp90 regulates chaperone function. Mol Cell 25: 151-159



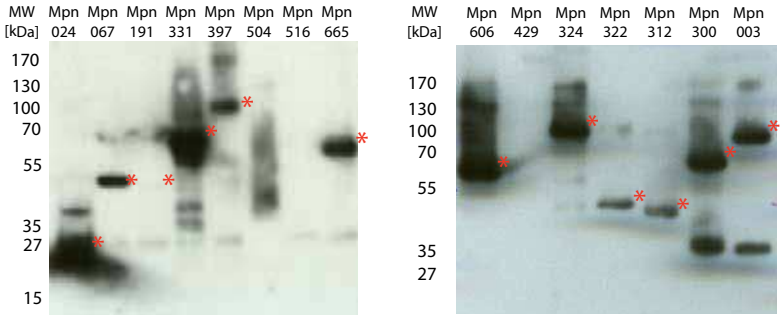
Supplementary Figure 1 | Quality measure of the dataset. A) the fraction of the proteins unique to our dataset, unique to the literature, either found in *Mycoplasma pneumoniae* or in other bacteria, and the overlap between the literature and our dataset. **B)** shows the distribution of the sequence coverage of the identified proteins including the median of the sequence coverage. **C)**shows the distribution in how many sample the lysine-acetylated peptide, the phosphorylated peptide or identified proteins. Only two samples were enriched for lysine-acetylated peptides, four samples were used for phosphoproteome analysis and all six samples were used for protein identification. The correlation between quantification between technical duplicates (panel **D**) and biological duplicates (panel **E**) for lysine acetylated, phosphorylated peptides and quantified proteins. **F)** Reproducibility of identifications of proteins and peptides between mixes with the same k.o. strains (mixed biological and technical duplicates) and up- and down-regulation of technical duplicates (same culture samples with swapped labels) and biological duplicates (different culture samples). If technical duplicates of quantifications were available, only those were both ratios are in the same direction, above the threshold and significant are considered in the final table of regulated proteins/peptides. Thus technical reproducibility after filtering is 100%, biological reproducibility after this filtering of up-/down-regulations is higher than before filtering.

A

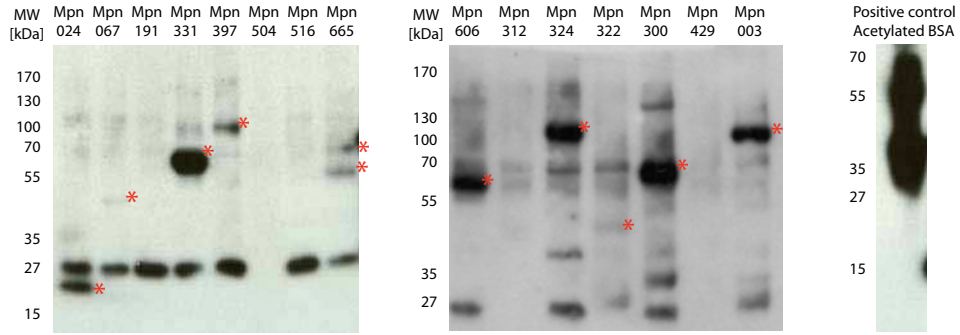
Systematic name	Mpn 003	Mpn 024	Mpn 067	Mpn 191	Mpn 300	Mpn 312	Mpn 322	Mpn 324	Mpn 331	Mpn 397	Mpn 429	Mpn 504	Mpn 516	Mpn 606	Mpn 665
Protein name	GyrB	RpoE	Mpn 067	RpoA	ScpA	Mpn 312	NrdF	NrdE	Tig	SpoT	Pgk	Mpn 504	RpoB	Eno	Tuf
Molecular weight [kDa]	73	17	36	37	60	24	40	82	51	86	44	15	156	49	43
identified acetylated by MS/MS															
identified acetylated in WesternBlot															
identified purified protein (α-CBP signal)				weak											

detected not detected

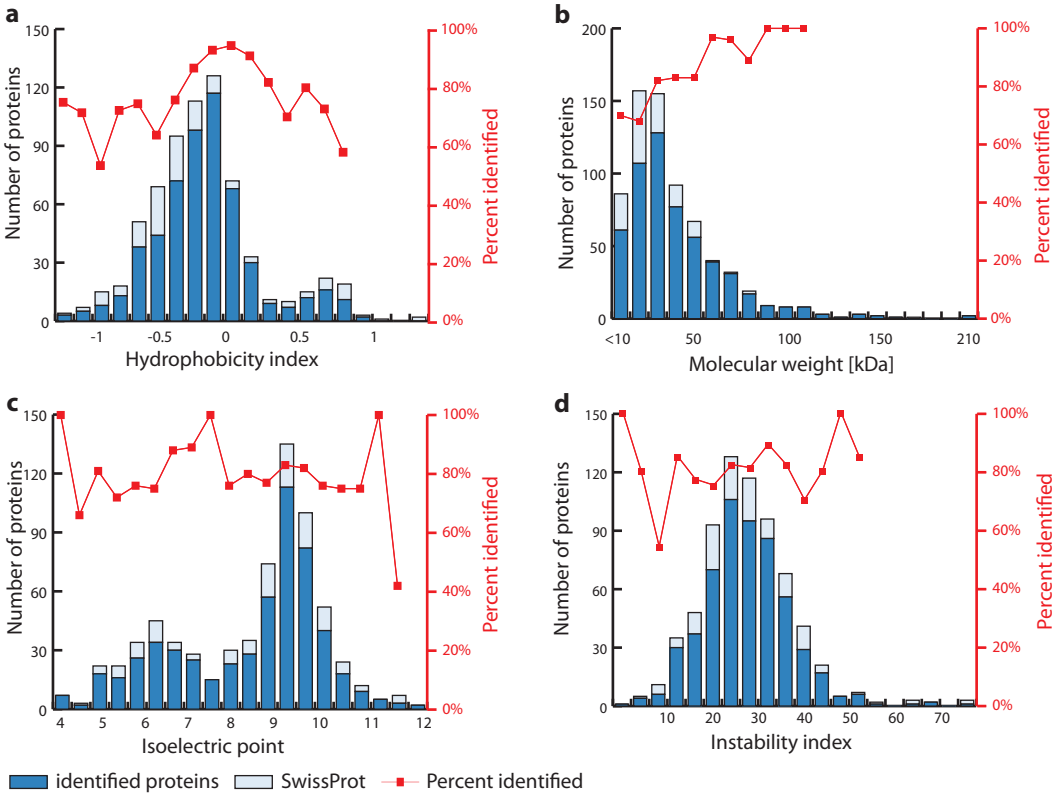
B



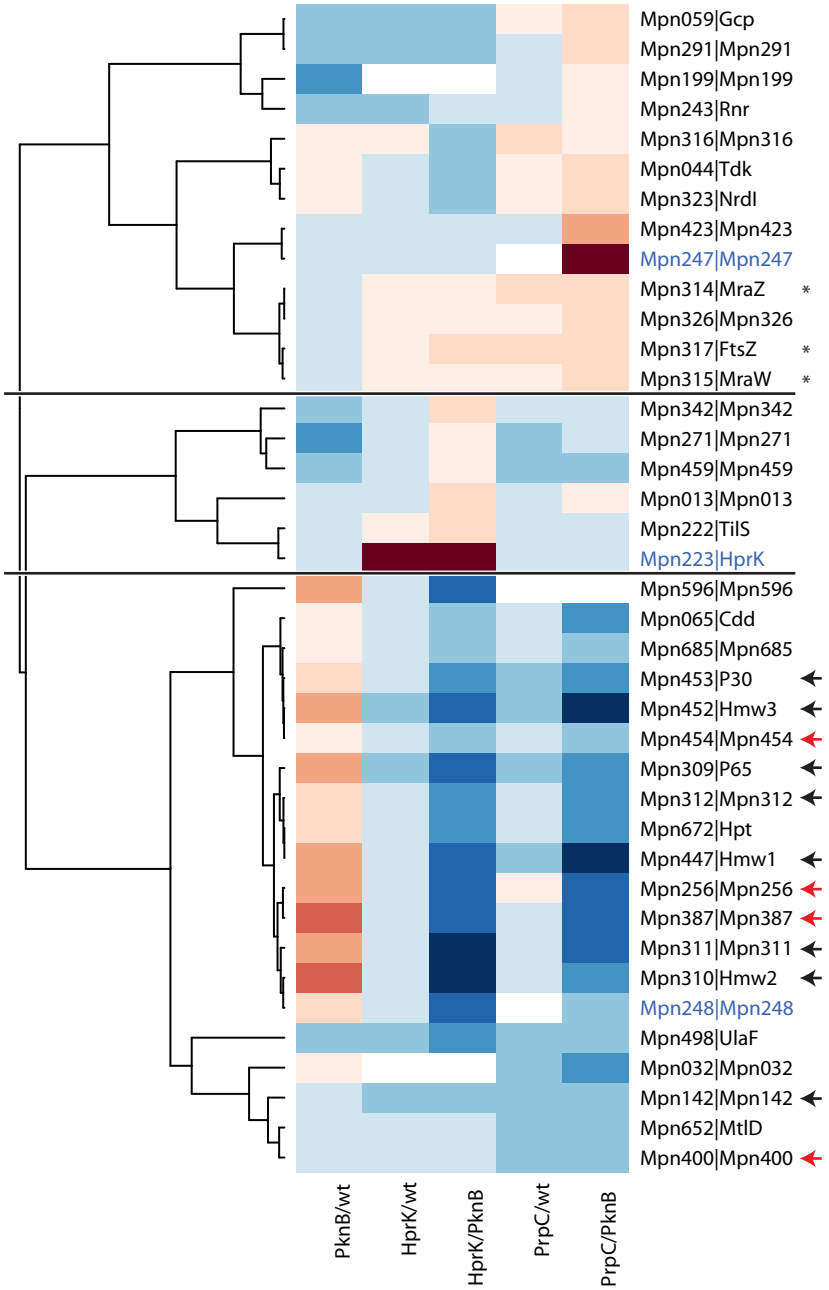
C



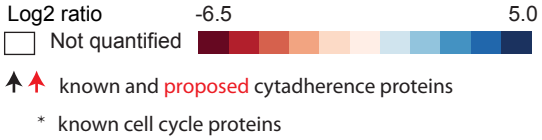
Supplementary Figure 2 | Validation of acetylated proteins by Western Blot. A) Result matrix of validation experiments. The columns give the systematic name, protein name and the molecular weight (MW) of each validation candidate. There color coded matrix rows below the validation candidate names indicate, whether a protein was (i) identified as being acetylated in the mass-spectrometric analysis, (ii) identified being acetylated by Western blot and (iii) identified being purified prior to Western blot acetlyation analysis. Green indicates a detected signal, red represents no detected signal. **B)** is Western blot raw data supporting the detection of the affinity purified validation candidate proteins via an α-calmodulin binding peptide antibody. **C)** is Western blot raw data supporting the detection of the acetylated validation candidate proteins via an α-acetyllysine antibody. As positive control a Western blot of acetylated bovine serum albumin detected via the α-acetyllysine antibody is shown.



Supplementary Figure 3 | The protein identification has not bias for physio-chemical properties such as isoelectric point (panel a), hydrophobicity (panel b), molecular weight (panel c) and instability index (panel d). The identified proteins (dark blue) are compared to the M. pneumoniae proteins annotated in SwissProt. Each parameter was calculated for each protein using the protparam tool from ExPASy (http://www.expasy.ch/tools/protparam.html) and binned into 20 equally sized ranges. Identified proteins are shown in dark blue, in SwissProt annotated proteins in light blue. The red line indicated the coverage in each bin.

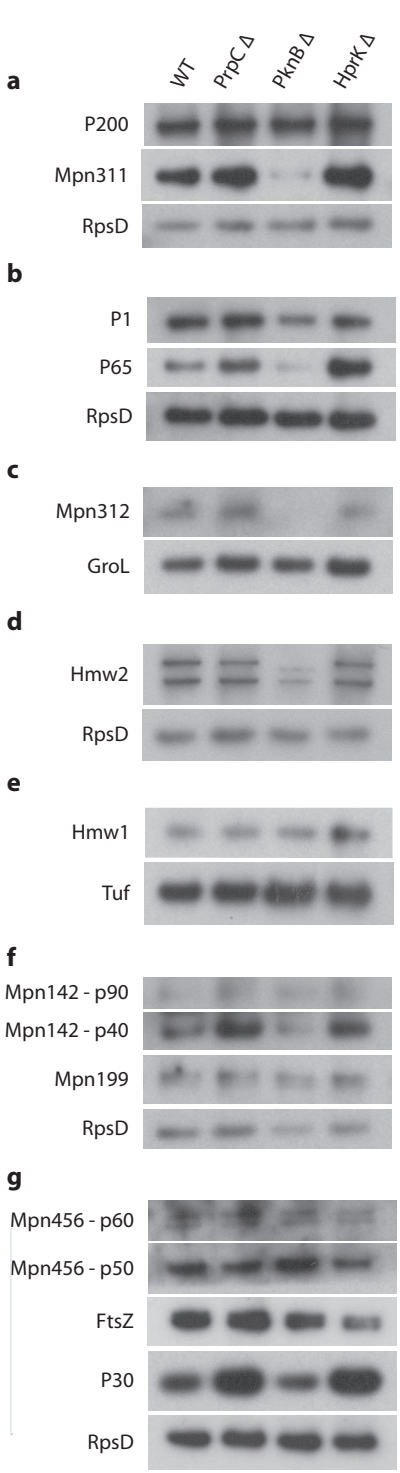


Supplementary Figure 4 | Clustering of protein abundance changes reveals new cytadherence proteins and functional implications of PrpC in cell cycle. The regulated proteins were clustered according to their relative abundance changes between the different strains. We propose four uncharacterized proteins (red arrows), which cluster together with the known cytadherence proteins (black arrows), to be involved in cytadherence. Proteins with a function related to the cell cycle (asterisk) cluster together with the PrpC, suggesting that these proteins are downregulaetd in *prpC* k.o. strain. The enzymes, which are depleted in either of the strains, are shown in blue.



known and proposed cytadherence proteins

* known cell cycle proteins



	wt	PrpC Δ	PknB Δ	HprK Δ
P200	168	186	177	194
Mpn311	149	189	73	208
RpsD	98	100	106	114

	wt	PrpC Δ	PknB Δ	HprK Δ
P1	92	93	79	81
P65	98	136	73	187
RpsD	115	125	106	105

	wt	PrpC Δ	PknB Δ	HprK Δ
Mpn312	129	148	85	120
GroL	109	138	116	149

	wt	PrpC Δ	PknB Δ	HprK Δ
Hmw2	131	129	91	135
RpsD	93	103	94	93

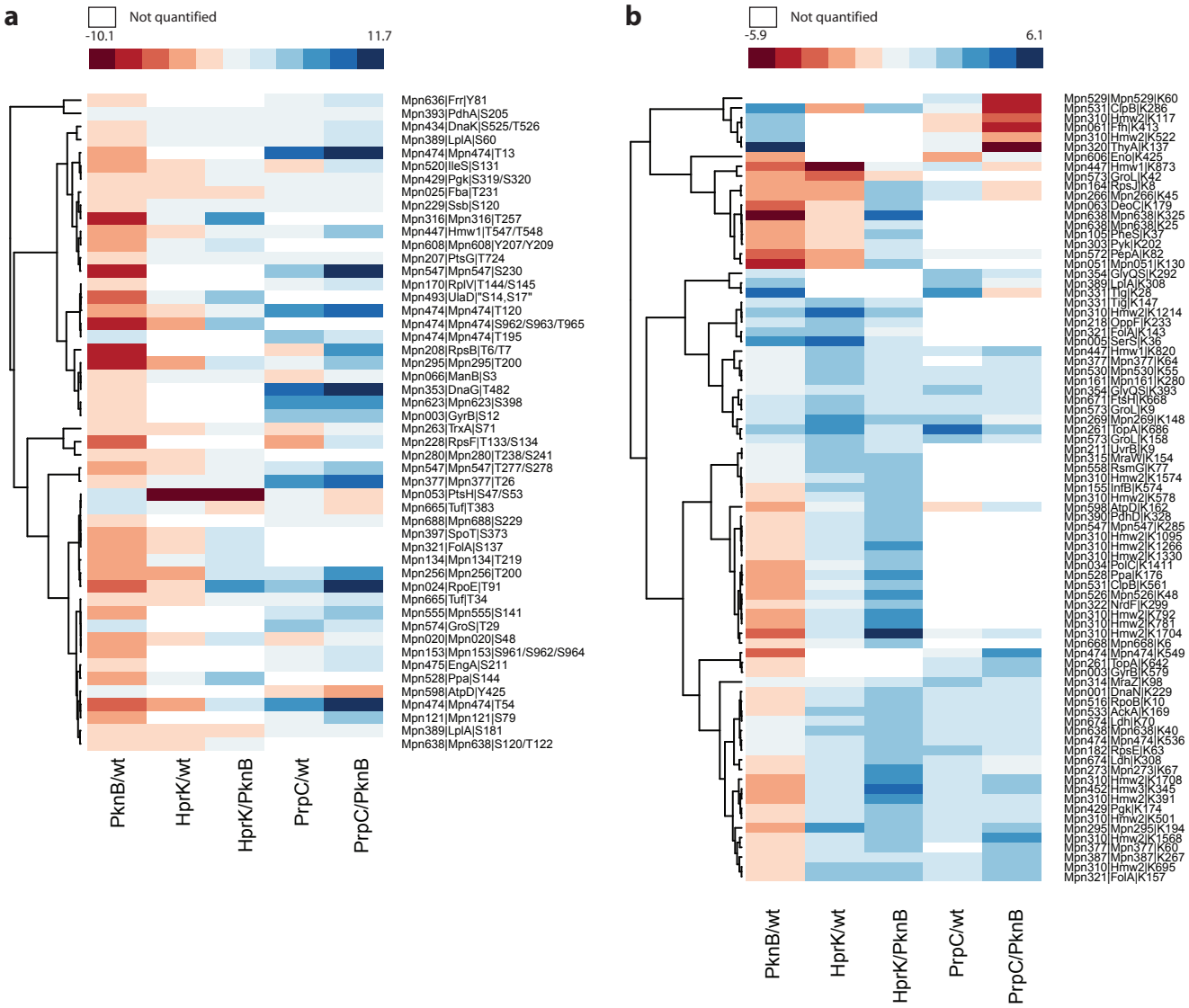
	wt	PrpC Δ	PknB Δ	HprK Δ
Hmw1	86	89	89	118
Tuf	138	160	173	159

	wt	PrpC Δ	PknB Δ	HprK Δ
Mpn142 - p90	100	114	104	108
Mpn142 - p40	158	193	114	166
Mpn199	105	109	100	106
RpsD	103	108	87	96

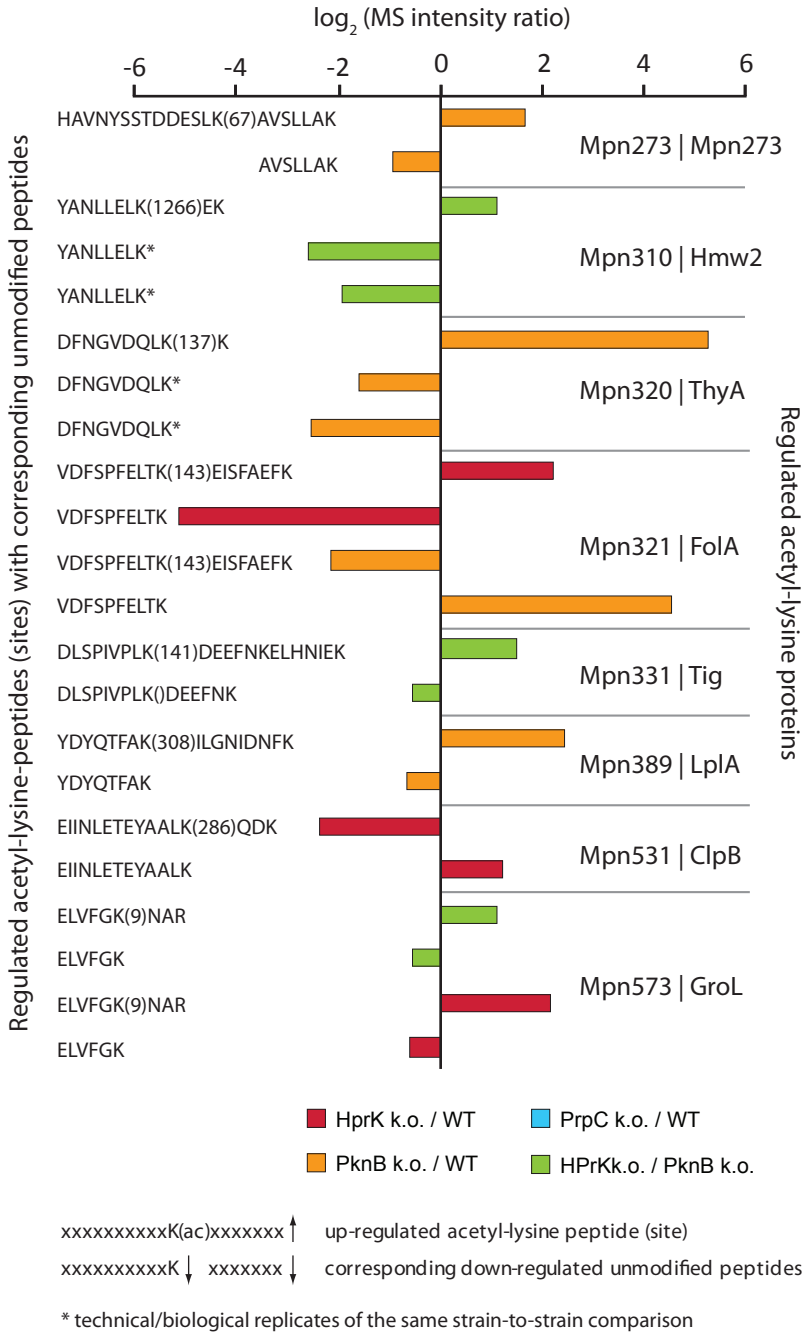
	wt	PrpC Δ	PknB Δ	HprK Δ
Mpn456 - p60	164	183	166	142
Mpn456 - p50	182	181	207	162
FtsZ	106	120	92	73
P30	70	92	70	99
RpsD	169	179	184	173

	PrpC Δ / wt	PknB Δ / wt	HprK Δ / wt	HprK Δ / PknB Δ	PrpC Δ / PknB Δ
P200	→	→	→	→	→
Mpn311	→	↓	→	↑	↑
RpsD	→	→	→	→	→
P1	→	→	→	→	→
P65	→	↓	→	↑	↑
RpsD	→	→	→	→	→
Mpn312	→	↓	→	↑	↑
GroL	→	→	→	→	→
Hmw2	→	↓	→	↑	↑
RpsD	→	→	→	→	→
Hmw1	→	↓	→	↑	↑
Tuf	→	→	→	→	→
Mpn142	→	→	→	→	↑
Mpn199	→	↑	nq	nq	→
RpsD	→	→	→	→	→
Mpn456	→	→	→	→	→
FtsZ	→	→	↓	↓	↓
P30	→	→	→	↑	↑
RpsD	→	→	→	→	→

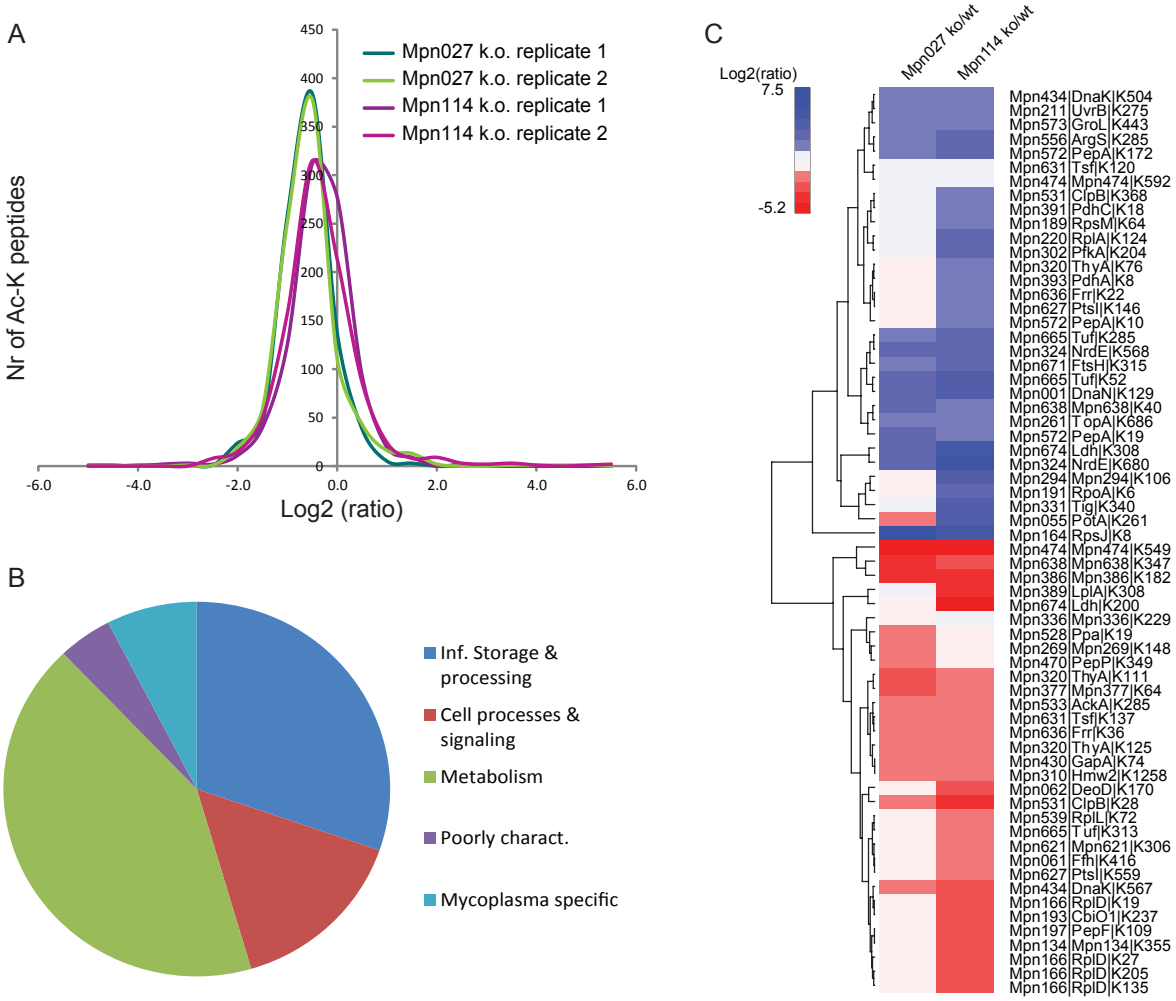
Supplementary Figure 5 | Validation of protein abundance changes by western blot. In each of the panel the signals from western blot membrane, the quantification of the western blot signals and the regulation expected from mass spectrometric data (Supplementary table 1) are shown. The proteins analyzed were **a,** P200 (Mpn567), Mpn311 and RpsD(Mpn446), **b,** P1 (Mpn141), P65 (Mpn309) and RpsD (Mpn446), **c,** Mpn312 and GroL (Mpn573), **d,** Hmw2 (Mpn310) and RpsD (Mpn446), **e,** Hmw1 (Mpn447) and Tuf (Mpn665), **f,** the two cleavage products of Mpn142, Mpn199 and RpsD (Mpn446), **g,** Mpn456, FtsZ (Mpn317), P30 (Mpn453) and RpsD (Mpn446).



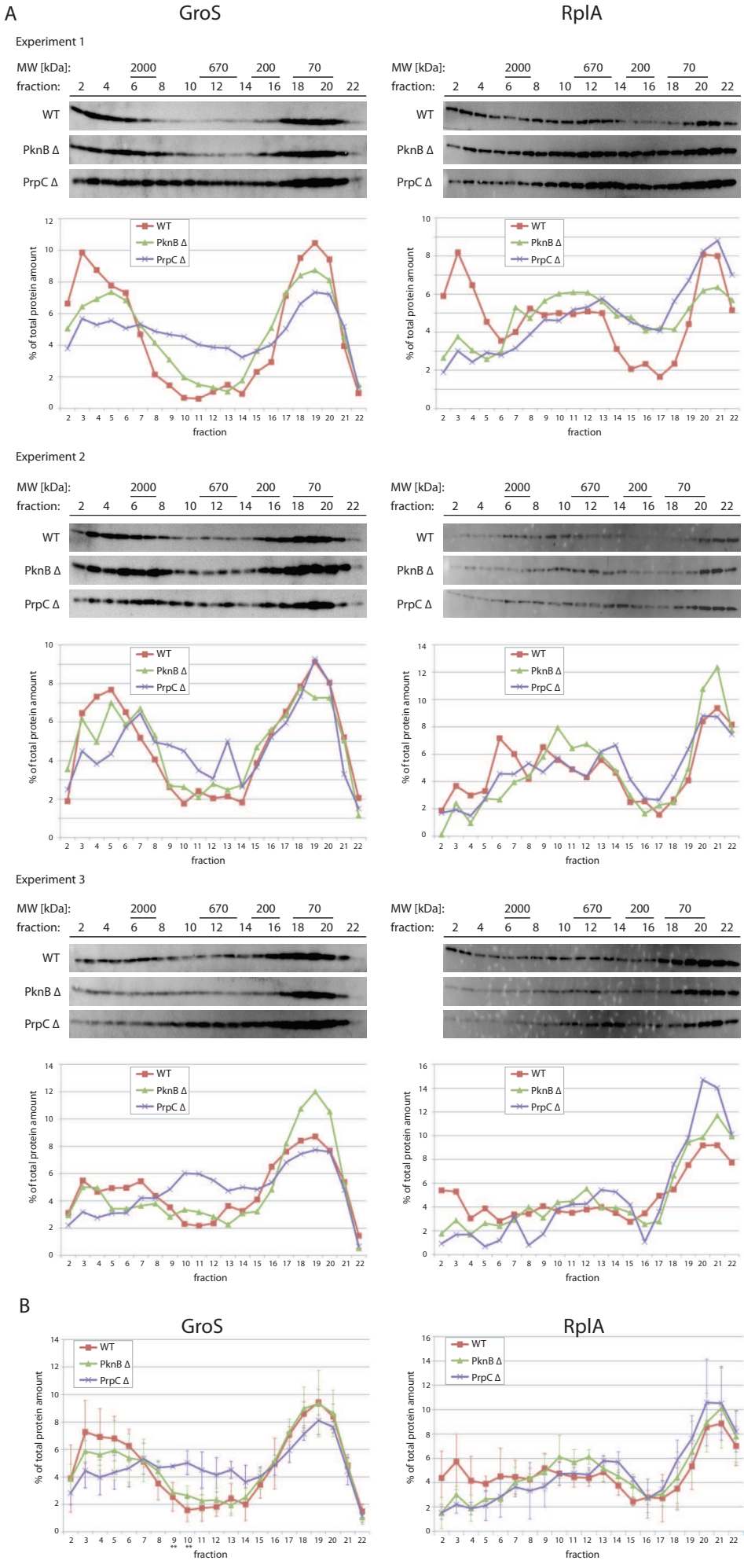
Supplementary Figure 6 | Clustering of phospho and acetyl-K sites. Only regulated sites were selected for clustering. Clustering was done with uncentered correlation and hclust as implemented in R. A clustering of phosphosites, B clustering of acetyl-K sites.



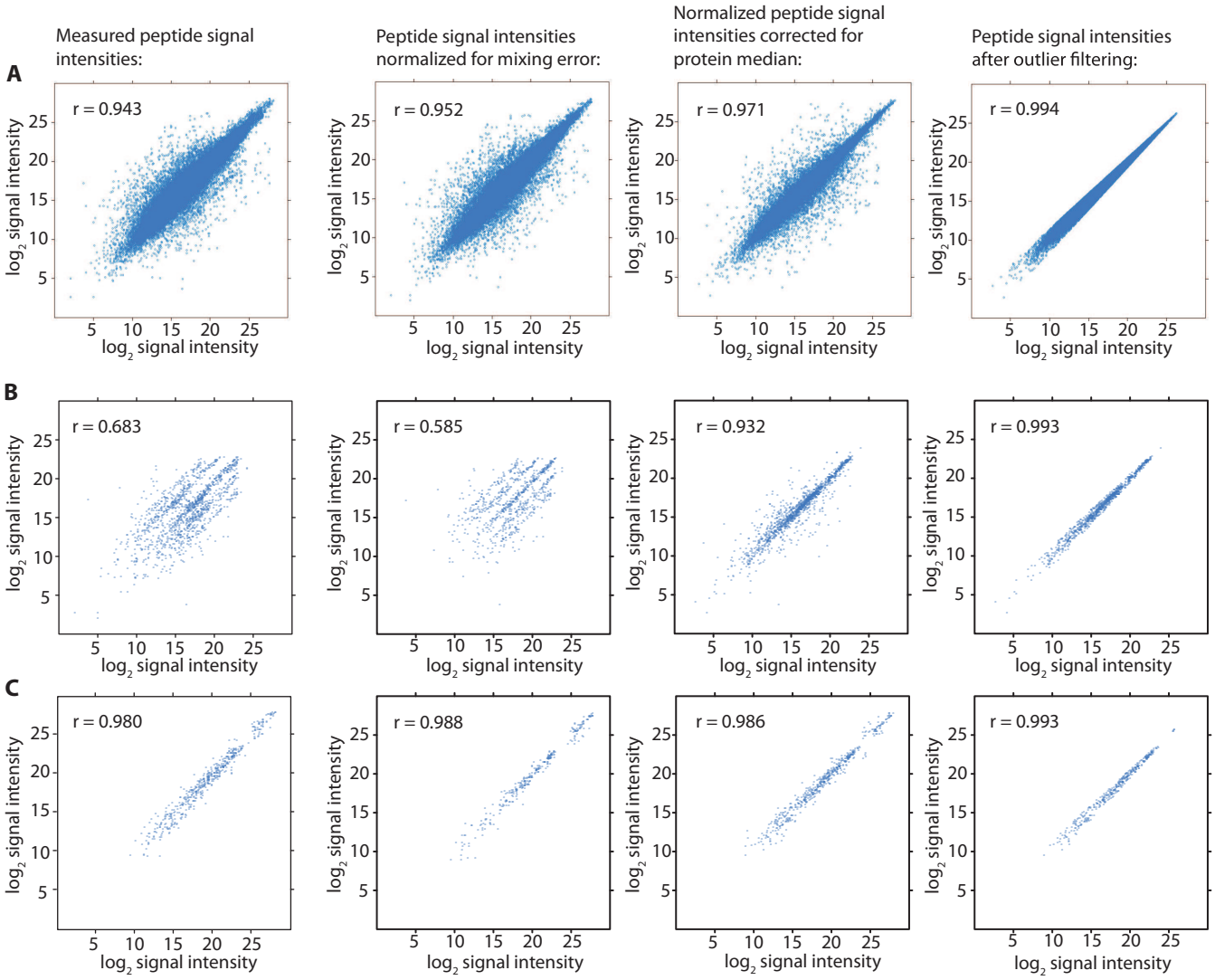
Supplementary Figure 7 | Validation of regulated acetyl-lysine sites by quantitative MS. Sets of corresponding acetyl-lysine/ unmodified peptides for nine acetyl-proteins with regulated acetyl-lysine sites in the different strains, with peptide sequences flanking the site, validated by regulation in opposite directions by $\log_2$(MS intensity ratios).



Supplementary Figure 8 | Characterization of acetylated lysines in acetyl-lysine k.o. strains. **A)** Distribution of protein weighted average normalized log2-ratios of acetylated peptides that have been quantified in both replicate experiments. **B)** Distribution of functional classifications of acetylated sites classified as regulated in acetylase k.o. strains. **C)** Heatmap of protein weighted average normalized log2 ratios of regulated acetylated lysine sites in the two acetylase k.o. strains.



Supplementary Figure 9 | Sedimentation of GroS and RplA on sucrose gradients. A) Lysis of samples from wt, PknB ko and PrpC ko was performed by supplementing the lysis buffer with the deacetylase inhibitors nicotinamid (10 mM) and butyric acid (50 mM). Volume of 30 µl of samples were layered on a top of 4 ml sucrose gradient (10-35%) and separated by 14 hours centrifugation at 130 000 g at 4°C. The gradient was subsequently divided into 22 fractions per 165 ul. Fractions were analyzed by SDS-PAGE and western blot. Polyclonal antibodies were used to detect the GroS protein (Mpn574) and as negative control 50S ribosomal protein RlpA (Mpn220); final detection was done using secondary antibodies. Band intensities were normalized for equal sample loading and in each fraction the percentage of total protein amounts of the total sample was calculated. **B)** Averages and sd's were calculated over three independent sucrose gradient separation experiments. We observed a significant difference ($p < 0.01$) of GroS in fractions 9 and 10 between wt and PrpC ko. No significant differences were found for RplA.



Supplementary Figure 10 | The effect of the peptide signal intensity normalizations in each step during the outlier detection is shown with the help of three examples. The signal intensities of all analyzed proteome combinations are displayed for all quantified peptides (panel **A**), all peptides originating from the regulated protein Hmw2 (Mpn310, panel **B**) and all the peptides of the not-regualted protein Ef-Tu (Mpn665, panel **C**). In each diagram the pearson correlation is indicated in the top left corner. For each of these examples in panel **A**-**C**, the diagrams from left to right show the comparison of these signal intensities that were measured, after they were normalized for mixing errors, after correction for protein abundance changes and after outlier removal., respectively. The signal intensities used to determine a particular abundance change are plotted against each other. The signal intensities should therefore be distributed on a diagonal.