Journal of proteome research

Article

pubs.acs.org/jpr

Systematic Analysis of the Lysine Acetylome in *Vibrio parahemolyticus*

Jianyi Pan,*[,†] Zhicang Ye,[†] Zhongyi Cheng,[‡] Xiaojun Peng,[§] Liangyou Wen,[†] and Fukun Zhao[†]

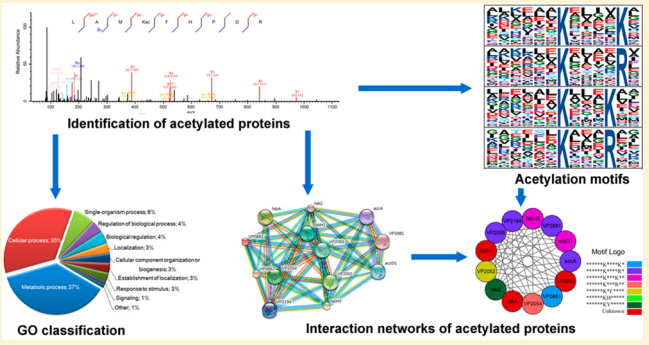
[†]Institute of Proteomics and Molecular Enzymology, School of Life Sciences, Zhejiang Sci-Tech University, Hangzhou 310018, China

[‡]Advanced Institute of Translational Medicine, Tongji University, Shanghai 200092, China

[§]Jingjie PTM Biolab (Hangzhou) Co., Ltd., Hangzhou 310018, China

Supporting Information

ABSTRACT: Lysine acetylation of proteins is a major post-translational modification that plays an important regulatory role in almost every aspect of cells, both eukaryotes and prokaryotes. *Vibrio parahemolyticus*, a model marine bacterium, is a worldwide cause of bacterial seafood-borne illness. Here, we conducted the first lysine acetylome in this bacterium through a combination of highly sensitive immune-affinity purification and high-resolution LC−MS/MS. Overall, we identified 1413 lysine acetylation sites in 656 proteins, which account for 13.6% of the total proteins in the cells; this is the highest ratio of acetyl proteins that has so far been identified in bacteria. The bioinformatics analysis of the acetylome showed that the acetylated proteins are involved in a wide range of cellular functions and exhibit diverse subcellular localizations. More specifically, proteins related to protein biosynthesis and carbon metabolism are the preferential targets of lysine acetylation. Moreover, two types of acetylation motifs, a lysine or arginine at the +4/+5 positions and a tyrosine, histidine, or phenylalanine at the +1/+2 positions, were revealed from the analysis of the acetylome. Additionally, protein interaction network analysis demonstrates that a wide range of interactions are modulated by protein acetylation. This study provides a significant beginning for the in-depth exploration of the physiological role of lysine acetylation in *V. parahemolyticus*.



KEYWORDS: *Lysine acetylation, acetylome, lysine acetylation motif, interaction network, V. parahemolyticus*

INTRODUCTION

The acetylation of lysine residues in proteins is a dynamic and reversible post-translational modification (PTM) that was first discovered in histone proteins nearly 50 years ago.[1] Early studies of this PTM mainly focused on histones and other transcription factors in the nucleus. After the discovery of lysine acetylation in non-histone proteins, it is believed that the extent of this modification is not restricted to nuclei.[2] In recent years, lysine acetylation have been found to occur in almost every compartment of a cell, such as the cytoplasm and mitochondria,[3−5] and plays a central role in various cellular processes. Therefore, lysine acetylation is believed to rival phosphorylation as a major signaling modality because of its ubiquitous occurrence and diverse biological function.[6,7]

Prokaryotic cells, particularly bacteria, were presumed to exhibit large-scale protein acetylation because of the successful identification of the acetylome in the mitochondrion and its close relationship to proteobacteria.[8] At present, lysine acetylation has been found to be widespread in proteobacteria. To date, the proteome-wide lysine acetylation profile in three members of the proteobacteria phylum, namely, *Escherichia coli*,[9−11] *Salmonella enterica*,[12] and *Thermus thermophilus*,[13] has been established. These studies have provided definite evidence that lysine acetylation is a common post-translational modification in bacteria. The identified acetyl proteins are associated with almost every aspect of cellular physiology, such as metabolism, translation, and folding, suggesting a close association between acetylation and diverse cellular functions. As a result, investigators have remarked that lysine acetylation can provide a new target for the development of effective drugs or vaccines based on an understanding of the regulatory functions of this modification.[14,15]

Vibrio parahemolyticus, a Gram-negative marine bacterium, occurs naturally in coastal marine waters and estuaries. It is a worldwide essential cause of bacterial seafood-borne illness, such as acute gastroenteritis, diarrhea, and abdominal pain, as a result of the consumption of contaminated raw or partially cooked seafood. In addition, septicemia and mortality occasionally occur during bacterial exposure to open wounds and in immunocompromised individuals.[16] Although diseases caused by *V. parahemolyticus* are sporadic cases, clinical strains with

Received: February 11, 2014
Published: May 30, 2014

   dx.doi.org/10.1021/pr500133t | *J. Proteome Res.* 2014, 13, 3294−3302

O3:K6 and several other serotypes have caused an increasing number of worldwide outbreaks of gastroenteritis.[17] In contrast, the mechanism controlling global virulence in these strains has not yet been uncovered. Interestingly, the lysine acetylation of kinases has been reported to be a novel mechanism through which this bacterium can manipulate host signaling to cause infection,[18] which suggests that acetylation plays a very important role not only in biological processes but also in the virulence and pathogenicity of *V. parahemolyticus*. Therefore, a proteome-scale analysis of the lysine acetylation sites and acetylated proteins in *V. parahemolyticus* is required to elucidate thoroughly the regulatory functions of this modification in this organism.

In this article, we present the systematic identification of the lysine acetylome of *V. parahemolyticus* serotype O3:K6, which was obtained using an integrated proteomic method. Overall, we identified 1413 unique lysine acetylation sites in 656 acetylated proteins with remarkably diverse biological functions and cellular localizations in this bacterium. In addition, several conserved acetylation motifs similar to those extracted from eukaryotic cells were found through a bioinformatic analysis of the sequences surrounding the acetylation sites. The interaction networks indicate that a very wide range of regulatory roles are modulated by acetylation. These results provide the first comprehensive view of the acetylome of *V. parahemolyticus*.

MATERIALS AND METHODS

Bacterial Strains and Culture

V. parahemolyticus serotype O3:K6 (strain RIMD 2210633) was grown overnight in high-salt LB medium (containing 3% NaCl) with shaking. One milliliter of overnight cultures was inoculated in 100 mL of high-salt LB medium and then incubated at 25 °C with shaking at 200 rpm. The bacterial cells were harvested at an OD_{600} of 0.8 (about 2.5 h after seed bacteria were inoculated into cultures) by centrifugation at 5000 *g* and 4 °C for 10 min and then washed twice with sterile phosphate buffered saline (PBS).

Protein Extraction and Digestion

The cultured bacterial cells were lysed in 8 M urea supplemented with 2 mM EDTA, 1 mM DTT, protease inhibitor cocktail (Protease Inhibitor Cocktail Set III; Calbiochem), and HDAC inhibitor (50 mM sodium butyrate, 30 mM nicotinamide, and 3 μM Trichostatin A) and then sonicated via 12 short 10 s bursts with 30 s cooling intervals. Unbroken cells and debris were removed by centrifugation at 4 °C and 20 000*g* for 10 min. The protein content in the supernatant was determined using the 2-D Quant kit (GE Healthcare) according to the manufacturer's instructions and precipitated with 20% trichloroacetic acid overnight at 4 °C. The resulting precipitate was washed three times with ice-cold acetone. The air-dried precipitate was resuspended in 100 mM NH_4HCO_3 and digested with trypsin (Promega) at an enzyme-to-substrate ratio of 1:50 for 12 h at 37 °C. The tryptic peptides were reduced with 5 mM dithiothreitol at 56 °C for 45 min and then alkylated with 15 mM iodoacetamide for 30 min at room temperature in the dark. The reaction was terminated by incubation with 30 mM cysteine at ambient temperature for 20 min. To ensure complete digestion, additional trypsin at an enzyme-to-substrate ratio of 1:100 was added, and the mixture was incubated for an additional 4 h. The digested peptides were lyophilized in a SpeedVac (Thermo Scientific).

Enrichment of Lysine Acetylated Peptides

The tryptic digest was redissolved in NETN buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris, and 0.5% Nonidet P-40, pH 8.0) and incubated with antiacetyl lysine agarose beads (catalog no. PTM-104, PTM Biolabs) in a ratio of 15 μL of beads/mg protein at 4 °C overnight with gentle shaking. After incubation, the beads were carefully washed three times with NETN buffer, twice with ETN buffer (100 mM NaCl, 1 mM EDTA, and 50 mM Tris, pH 8.0), and once with water. The bound peptides were eluted from the beads with 1% trifluoroacetic acid and dried in a SpeedVac. The resulting peptides were cleaned with C_{18} Zip Tips (Millipore) according to the manufacturer's instructions and then subjected to LC−MS/MS analysis.

LC−MS/MS Analysis by Q Exactive

The peptides were resuspended in buffer A (2% ACN, 0.1% FA) and centrifuged at 20 000*g* for 2 min. The supernatant was transferred into a sample tube and loaded onto an Acclaim PepMap 100 C18 trap column (Dionex, 75 μm × 2 cm) by EASY nLC1000 nano-UPLC (Thermo Scientific), and the peptide was eluted onto an Acclaim PepMap RSLC C18 analytical column (Dionex, 50 μm × 15 cm). A 34 min gradient was run at 300 nL/min; the mobile phase was changed from 5 to 30% B (80% ACN, 0.1% FA) to 48% B over a 2 min period, then to 80% B over a 2 min period, and then maintained at 80% B for 4 min.

The peptides were subjected to NSI source followed by tandem mass spectrometry (MS/MS) using a Q Exactive instrument (Thermo Scientific) coupled online to the UPLC instrument. Intact peptides were detected in the Orbitrap at a resolution of 70 000. The peptides were selected for MS/MS using 25% NCE with 4% stepped NCE, and ion fragments were detected in the Orbitrap at a resolution of 17 500. A data-dependent procedure that alternated between one MS scan and 15 continuous MS/MS scans was applied for the analysis of the top 15 precursor ions above a threshold ion count of 4×10^4, as determined through an MS survey scan with a dynamic exclusion of 2.5 s. The electrospray voltage applied was 1.8 kV. Automatic gain control (AGC) was used to prevent overfilling of the ion trap, and 2×10^5 ions were accumulated for the generation of their MS/MS spectra. For the MS scans, the m/z scan range was 350 to 1800 Da.

Database Search

The protein and acetylation site identification was performed through MaxQuant with an integrated Andromeda search engine (ver. 1.3.0.5). The tandem mass spectra were searched against the UniProt *V. parahemolyticus* serotype O3:K6 protein database (4823 sequences) concatenated with a reverse decoy database and protein sequences of common contaminants. Trypsin/P was specified as a cleavage enzyme, and the search allowed up to three missing cleavages, four modifications per peptide, and five charges. The mass error was set to 6 ppm for precursor ions and 0.02 Da for fragment ions. Carbamidomethylation on Cys was specified as a fixed modification, and oxidation on Met, acetylation on Lys, and acetylation on protein N-terminus were specified as variable modifications. The false discovery rate (FDR) thresholds for proteins, peptides, and modification sites were specified to be 0.01. The minimum peptide length was set to 7. All of the other parameters in the MaxQuant analysis were set to the default values. Lys acetylation sites identified with a localization probability of less than 0.75 compared with that of reverse or contaminant protein sequences were removed.

dx.doi.org/10.1021/pr500133t | J. Proteome Res. 2014, 13, 3294−3302

Bioinformatics Analysis

GO Annotation. The Gene Ontology (GO) annotation proteome was derived from the UniProt-GOA database (http://www.ebi.ac.uk/GOA/). The proteins were classified by Gene Ontology annotation based on three categories: biological process, cellular compartment, and molecular function.

KEGG Pathway Annotation. We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) to annotate the pathways. We first annotated the proteins using the KEGG online service tool KAAS and then mapped the annotation result on the KEGG pathway database using the KEGG online service tool KEGG Mapper.

GO/KEGG Pathway Functional Enrichment Analysis. Fisher's exact test was used to test for the enrichment or depletion (two-tailed test) of specific annotation terms among members of the resulting protein clusters. The derived p values were further adjusted to address multiple hypotheses through the method proposed by Benjamini and Hochberg. Any terms with adjusted p values of less than 0.05 in any of the clusters were treated as significant.

GO/KEGG Pathway Functional Enrichment-Based Clustering of Protein Groups Based on Protein Quantification. For further hierarchical clustering based on categories, we first collated all of the categories obtained after the enrichment and their p values and then identified those categories that were enriched in at least enriched one of the clusters with a p value of less than 0.05. This filtered p-value matrix was transformed by the function $x = -\log(p \text{ value})$, and the x values for each category were z-transformed. These z scores were then clustered by one-way hierarchical clustering (Euclidean distance, average linkage clustering) using Genesis. The cluster membership was visualized by a heat map using the heatmap.2 function in the ggplot2 R-package

Model of Sequences Surrounding Acetylation Sites. Motif-X software was used to analyze the model of the sequences with amino acids in specific positions of acetyl-13-mers (six amino acids upstream and downstream of the acetylation site) in all of the protein sequences. In addition, all of the database protein sequences were used as the background database parameter, and the other parameters were set to the default values.

Protein−Protein Interaction Analysis. We analyzed the protein−protein interactions for the identified acetylated proteins using Cytoscape software. The protein−protein interaction network was obtained from the STRING database, which defines a metric called the "confidence score" to define the interaction confidence; we obtained all of the interactions with a confidence score of at least 0.7 (high confidence).

RESULTS AND DISCUSSION

Proteome-Wide Analysis of Lysine Acetylation Sites and Proteins in *V. parahemolyticus*

The whole genome of *V. parahemolyticus* serotype O3:K6 was sequenced approximately a decade ago,[19] and this facilitated the systematic analysis of the lysine acetylated proteins in this species. To determine the protein acetylome in *V. parahemolyticus*, a proteomic method based on affinity purification and LC−MS/MS was applied, and tandem mass spectra were searched against the UniProt *V. parahemolyticus* serotype O3:K6 protein database (4823 sequences).[19] The mass error for precursor ions was set to 6. After applying these parameters, we identified 1413 acetylation peptides with a peptide score greater than 30 (Table S1 and Figure S1 in the Supporting Information). These peptides, which exhibit different abundances depending on their length (Figure 1A), match 656 acetylated proteins containing different acetylation sites (from 1 to 15) (Figure 1B and Table S1 in the Supporting Information). Figure 2 shows three MS/MS spectra of acetylated peptides of chaperone protein DnaJ (Q87RX2) with their acetylation sites at K31, K46, and K360, respectively. The identified acetyl proteins account for 13.6% of the total proteins in the bacteria. Therefore, we performed the first large-scale analysis of lysine acetylated proteins in *V. parahemolyticus*.

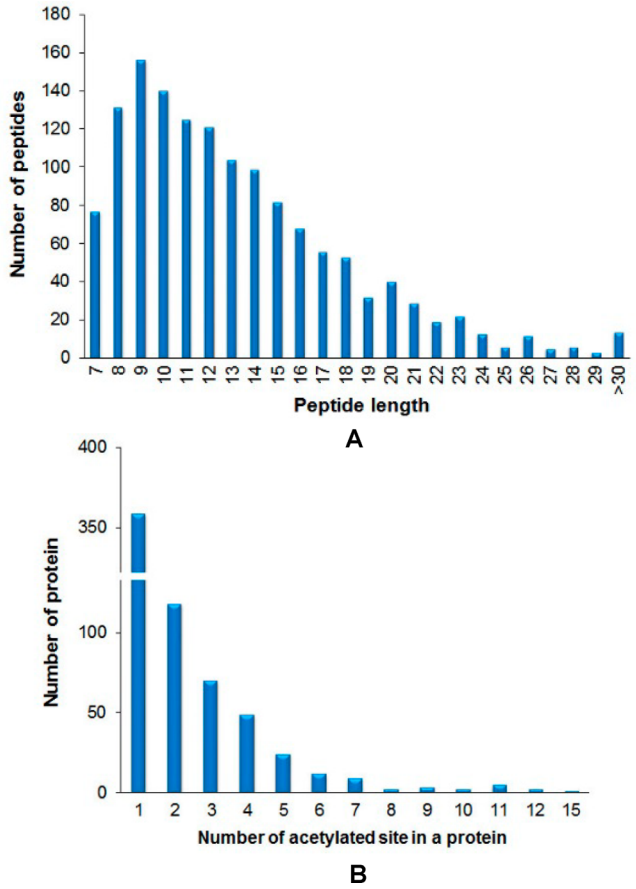


Figure 1. (A) Distribution of acetylated peptides based on their length. (B) Distribution of acetylated proteins based on their number of acetylation peptides.

To date, only three bacterial species, namely, *E. coli*,[9−11] *S. enterica*,[12] and *T. thermophilus*,[13] have been subjected to proteome-wide identification of lysine acetylation. In all of these studies, lysine acetylation was investigated using a similar strategy: the acetylated peptides were enriched by immune-affinity purification, and the acetyl proteins were then identified through LC−MS/MS. However, the amounts of identified acetylation sites and matched proteins are significantly different, as shown in Table 1. In our opinion, there are two potential reasons that may account for these differences. First, the intrinsic acetylation level of the proteins varies markedly between the different bacterial species. The percentage of identified acetyl proteins of the total proteins in each bacterium differs from 2.1% in *E. coli* to 13.6% in *V. parahemolyticus*, as determined from reported data and our results (Table 1).

dx.doi.org/10.1021/pr500133t | J. Proteome Res. 2014, 13, 3294−3302

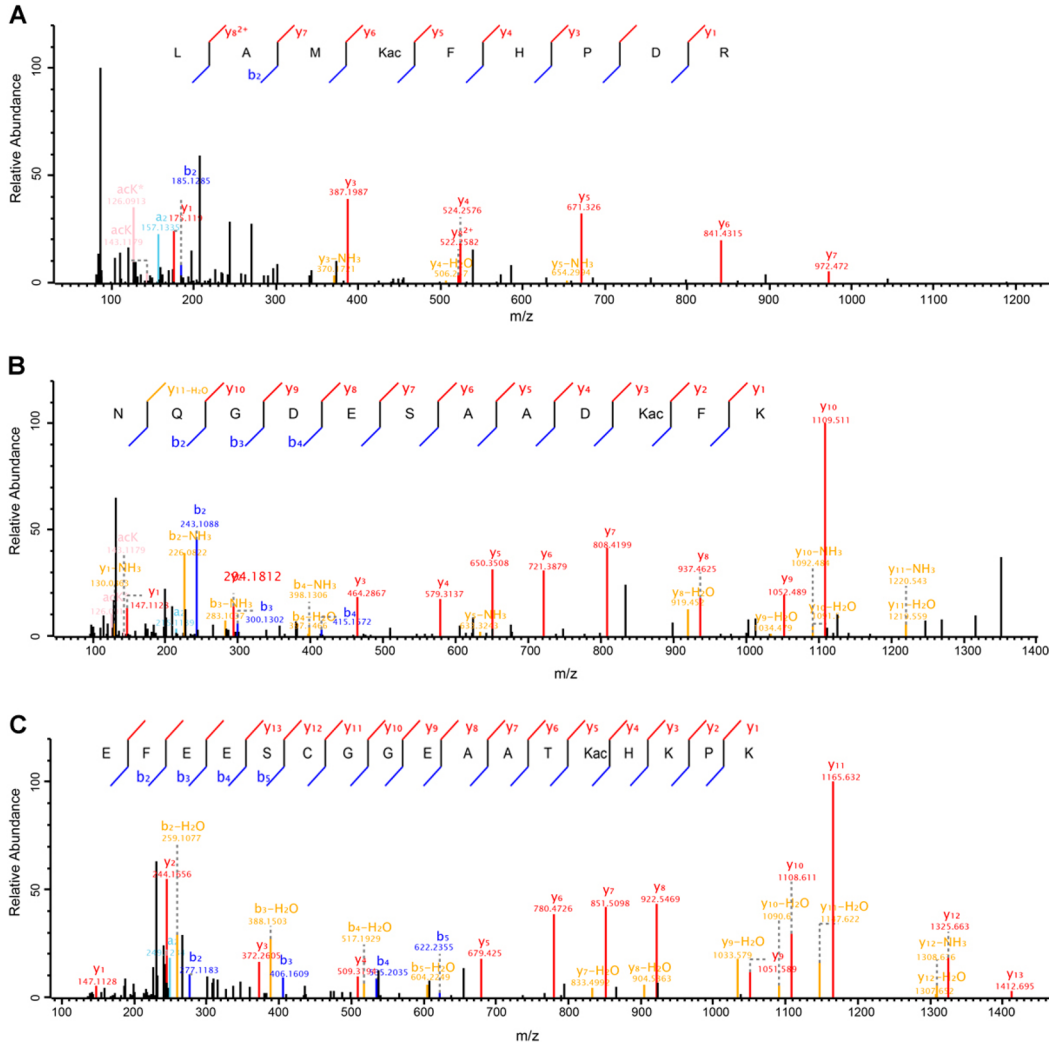


Figure 2. Representative MS/MS spectra of acetylpeptides from chaperone protein DnaJ (Q87RX2): (A) acetylpeptide LAMK^{ac}FHPDR with an acetylation site at K31, (B) acetylpeptide NQGDESAADK^{ac}FK with an acetylation site at K46, and (C) acetylpeptide EFEESCGGEAATK^{ac}HKPK with an acetylation site at K460.

Table 1. Acetylation Sites and Acetylated Proteins of *V. parahemolyticus* Identified in This Study Compared with Those Found in Other Bacterial Species

species	no. of acetylation sites	no. of acetylated proteins	percentage of acetyl proteins of the total proteins (%)	ref
V. parahemolyticus	1413	656	13.6	this study
E. coli	113	91	2.2	10
	125	85	2.1	9
	1070	349	8.2	11
S. enterica	235	191	4.2	12
T. thermophilus	197	128	5.9	13

However, our inspection of the results from three studies on *E. coli* performed in different years revealed that this is not the case. In fact, the percentages of *E. coli* acetyl proteins have been found to be 2.1, 2.2, and 8.2%, as determined by Yu et al.,[9] Zhang et al.,[10], and Zhang et al.,[11] respectively. It is obvious that the identification of a higher number of acetylated proteins may be attributed to improvements in antibody specificity and mass spectrometry sensitivity.[11] Herein, the enhancement of the proteomic technologies for the identification of lysine acetylation are the second cause that led to the divergence of acetyl proteins. Therefore, a higher number of acetylated proteins in bacterial cells should be identified with the development of novel proteomics technologies.

Characterization of Lysine Acetylome of *V. parahemolyticus*

To better understand the lysine acetylome in *V. parahemolyticus*, we investigated the GO functional classification of all of the identified acetyl proteins based on their biological process and molecular function (Figure 3 and Supporting Information Table S1). The classification results for biological process and molecular function both showed that the largest protein group of acetyl proteins is composed of enzymic proteins associated with metabolism, which accounts for 37 and 45% of the total acetyl proteins, respectively (Figures 3A,B). These findings are consistent with previous results that show that approximately half of the lysine acetylated proteins are categorized as metabolic proteins in other bacteria.[9−13] Another large acetyl protein group determined by their molecular function classification comprises binding proteins, which account for 40% of all of the identified proteins (Figure 3B). These findings indicate the essential regulatory roles of

dx.doi.org/10.1021/pr500133t | J. Proteome Res. 2014, 13, 3294−3302

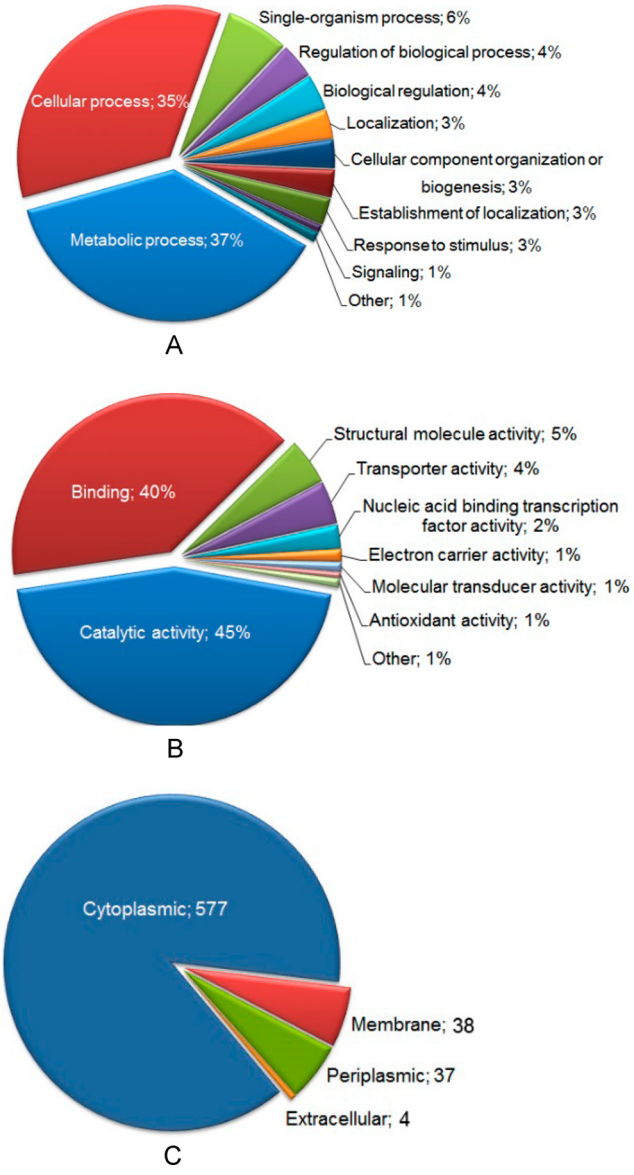


Figure 3. Gene Ontology functional classification of the identified acetylated proteins in terms of (A) biological process, (B) molecular function, and (C) and subcellular location.

acetylated proteins in cells. In contrast, the second largest group in terms of biological process is composed of proteins associated with cellular processes, and the number of proteins in this group is 35% of all of the identified proteins (Figure 3A). The GO analysis of the acetylome suggests that the acetylated proteins have wide ranges of biological processes and molecular functions in *V. parahemolyticus*.

The subcellular location of the acetylated proteins was also analyzed, and the results show that 577 of the proteins are distributed in the cytoplasm (88%), 37 proteins are located in the periplasmic space (6%), and 38 proteins are distributed in the membrane (6%), as shown in Figure 3C. The overall trend in the subcellular location of the acetylated proteins is similar to previous results obtained in *E. coli*.[11]

Furthermore, to determine which types of proteins are preferred targets for lysine acetylation, we evaluated the acetylation data for enrichment in two GO categories: molecular function and biological process (Figure 3A,B and Supporting Information Table S2). In the biological process category, the processes related to translation, cellular protein metabolism, tRNA aminoacylation for protein translation, and amino acid activation were found to be significantly enriched (Figure 4A). This pattern suggests that those proteins

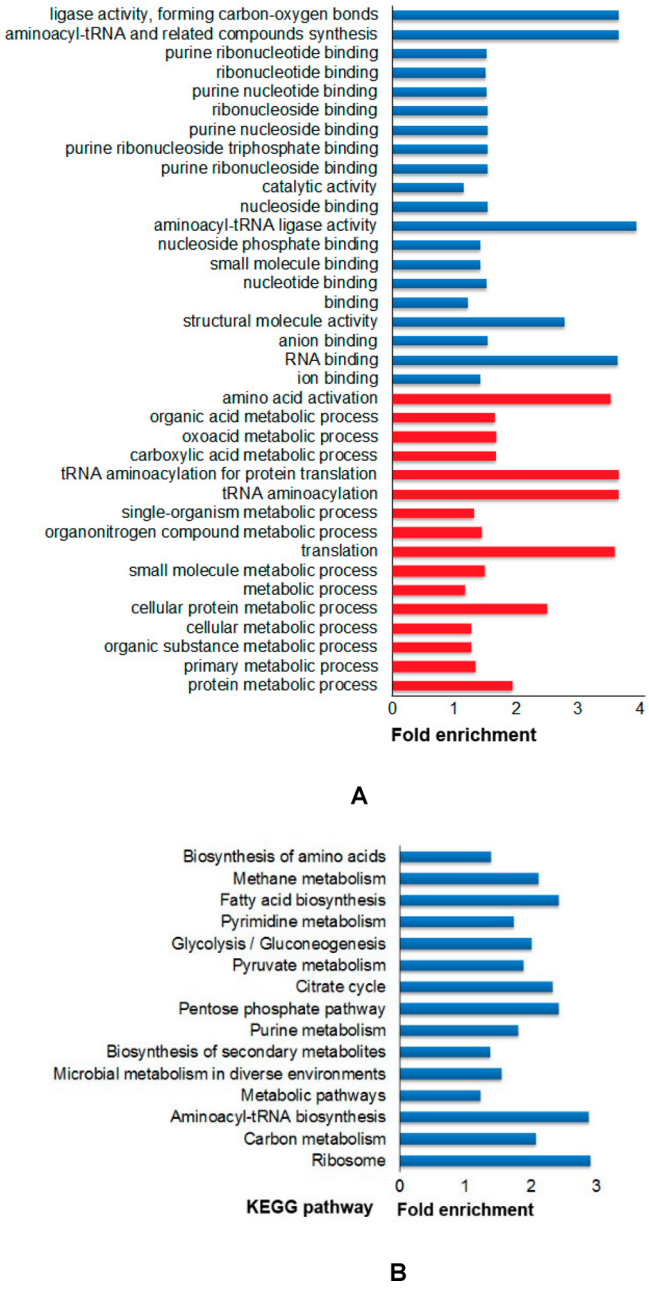


Figure 4. Enrichment analysis of the acetylated proteins in *V. parahemolyticus*. (A) The acetylated proteins in each group are significantly overrepresented, as classified by their GO annotation in terms of biological process (blue bars) and molecular function (red bars) ($p < 10^{-9}$). (B) KEGG pathway enrichment analysis ($p < 0.01$).

associated with protein biosynthesis are most likely acetylated in *V. parahemolyticus*. In agreement with this observation, the enrichment analysis based on molecular function showed that proteins with aminoacyl−tRNA ligase activity have a higher tendency to be acetylated (Figure 4A). Moreover, proteins involved in ribosome and aminoacyl−tRNA biosynthesis were also the most enriched, as determined through the KEGG

dx.doi.org/10.1021/pr500133t | *J. Proteome Res.* 2014, 13, 3294−3302

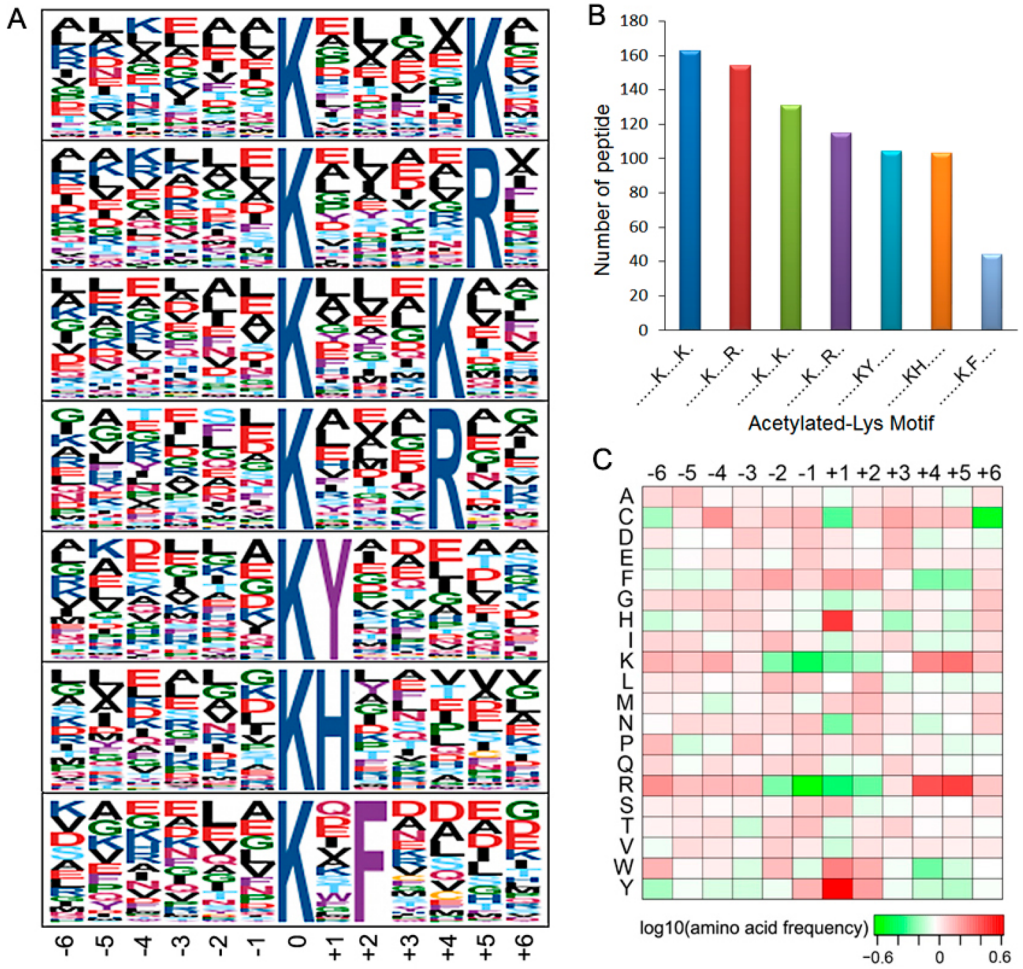


Figure 5. Properties of the acetylated peptides. (A) Acetylation motifs and conservation of acetylation sites. The height of each letter corresponds to the frequency of that amino acid residue in that position. The central K refers to the acetylated lysine. Acetylated lysine motif analysis using Motif-X software. (B) Number of identified peptides containing acetylated lysine in each motif. (C) Heat map of the amino acid compositions of the acetylation sites showing the frequency of the different types of amino acids around the acetylated lysine.

pathway analysis (Figure 4B and Supporting Information Table S2). Overall, these data strongly suggest that the cellular process of protein biosynthesis may be strictly regulated by lysine acetylation modification and that acetylated proteins involved in protein biosynthesis are thus particularly crucial in *V. parahemolyticus*. In addition to the enrichment of acetylated proteins associated with protein biosynthesis, proteins involved in the metabolism of four types of biological macromolecules were also found to be enriched: carbohydrate metabolism (including metabolic pathways of glycolysis/gluconeogenesis, pentose phosphate, and citrate cycle), lipid metabolism (fatty acid biosynthesis), amino acid metabolism (glycine, serine, and threonine metabolism and biosynthesis of amino acids), and nucleotide metabolism (purine and pyrimidine metabolism) (Figure 5B and Supporting Information Table S2). Moreover, a special pathway termed microbial metabolism in diverse environments was also enriched with acetylated proteins, which suggests the lysine acetylation modification should contribute to the bacterium's ability to adapt to various environments.

The finding that acetylated proteins involved in translation, transcription, metabolism, and chaperones are significantly enriched has also been reported in *Plasmodium falciparum*, a type of human malaria parasite,[20] indicating the essential role underlying lysine acetylation in different organisms.

Analysis of Acetylated Lysine Motifs

Previous studies on both eukaryotic and prokaryotic cells have found preferences for amino acid residues at particular positions surrounding the acetylated lysine.[3,10,13,21] Thus, to further evaluate the nature of the acetylated lysines in *V. parahemolyticus*, we investigated the sequence motifs in all of the identified acetylated lysines using the Motif-X program, which is a software tool designed to extract overrepresented patterns from any set of sequences.[22] Of all of the acetyl-lysine peptides, 814 peptides were found to include the amino acid sequence from the −6 to the +6 positions surrounding the acetylated lysine, and these were matched to a total of seven definitively conserved motifs (Figure 5 and Supporting Information Table S3), namely, K^{ac}****K, K^{ac}****R, KK^{ac}***K, KK^{ac}***R, K^{ac}Y, K^{ac}H, and K^{ac}*F (Figure 5A), and exhibit different abundances (Figure 5B) (K^{ac} indicates the acetylated lysine, and * indicates a random amino acid residue). A survey of these motifs suggested that two distinct types of residues are found downstream of the acetylated lysine: a positively charged residue, including lysine (K), arginine (R), or histidine (H), and a residue with aromatic groups, such as tyrosine (Y) and phenylalanine (F). According to the position

dx.doi.org/10.1021/pr500133t | J. Proteome Res. 2014, 13, 3294−3302

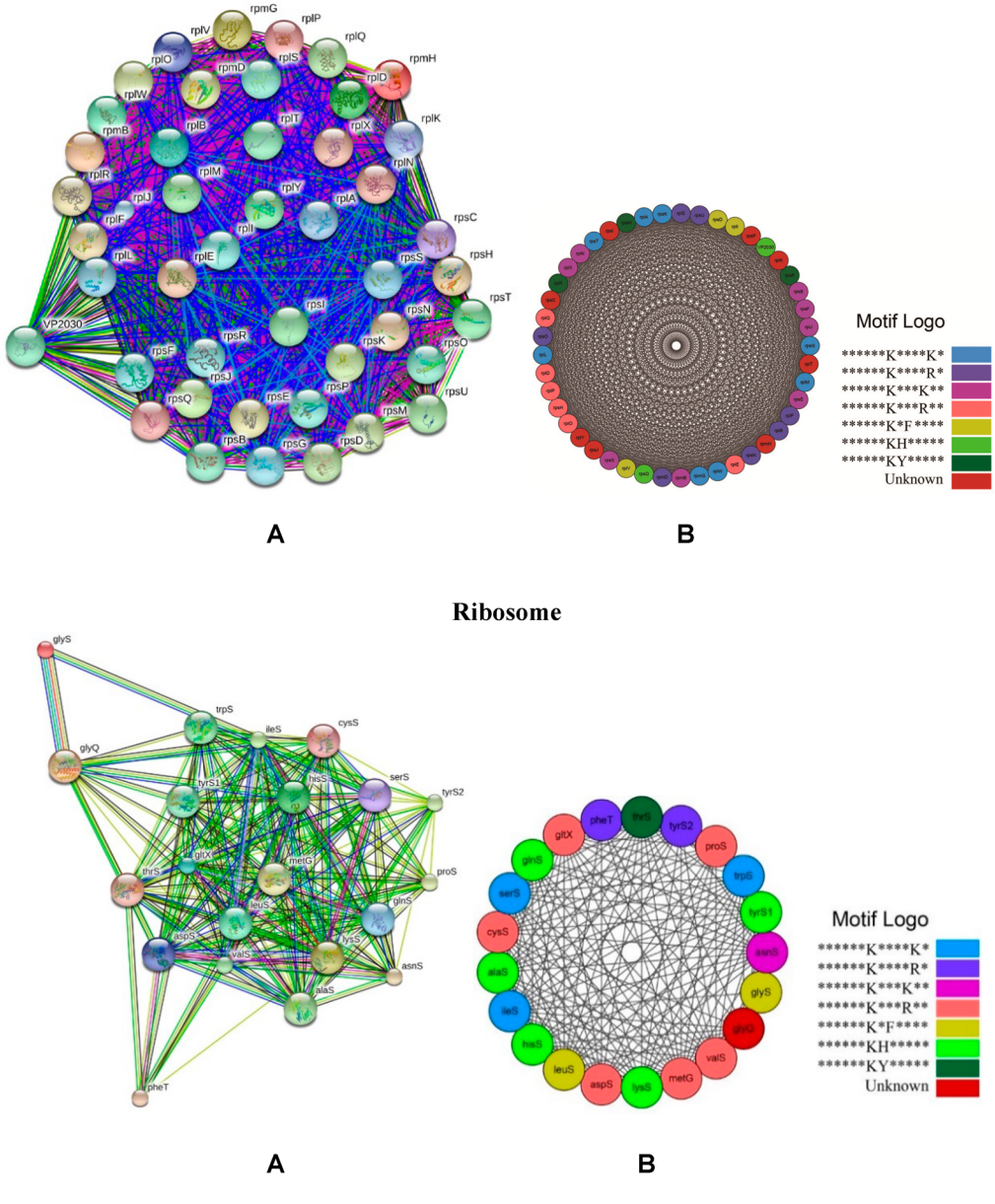


Figure 6. Interaction networks of acetylated proteins associated with ribosome and aminoacyl−tRNA biosynthesis.

of the residues and other properties of the residues around the acetylated lysine, these motifs can be classified into two categories: the +4 or +5 position is an alkaline residue with a long side chain (K or R) and the +1 or +2 position is a residue with a large side chain (Y, H, or F). In particular, the first motif is strikingly conserved, and the acetylated peptides with this motif account for approximately 40% of all of the identified acetylated peptides. Moreover, the frequency of K and R in positions −2 to +2 in the motif is the lowest, as determined by an inspection of the heat map of the amino acid compositions surrounding the acetylation sites (Figure 5C). These results indicate that the lysine residue of a polypeptide with a K and R amino acid at the +4 or +5 position and without one of these amino acids at the −2 to +2 positions is a preferred substrate of lysine acetyltransferase. Intriguingly, some acetylated lysine motifs are also observed in eukaryotes, such as K^{ac}****K, K^{ac}***K, K^{ac}Y, K^{ac}*F, and K^{ac}H in human cells,[2,3,21] K^{ac}***K in rat,[23] and K^{ac}Y in grape,[24] which suggests that eukaryotic lysine acetyltransferases also exist in prokaryotes, particularly bacteria. These results are consistent with the previous finding that lysine acetyltransferases are broadly distributed in bacteria, as determined by genetic BLAST searches.[25,26] To some extent, our results provide evidence of the homology of these enzymes, although they are currently ill-defined. However, to date, only two motifs, namely, K^{ac}H and K^{ac}Y, have been uncovered in other prokaryotic cells (*E. coli*),[10] likely because the sizes of the acetylation data sets obtained from studies on prokaryote cells are not sufficiently large to extract other motifs. Apparently, the large data sets obtained in our study enabled us to acquire these highly specific acetylation motifs.

dx.doi.org/10.1021/pr500133t | *J. Proteome Res.* 2014, 13, 3294−3302

Analysis of Protein Interaction Networks of Acetylated Proteins in *V. parahemolyticus*

To further understand the cellular processes regulated by acetylation in *V. parahemolyticus*, we determined the protein interaction network for all of the acetylated proteins using Cytoscape software. The results showed that 502 acetylated proteins are network nodes and that 8627 direct physical interactions with a combined score higher than 0.70 were obtained from the STRING database (Supporting Information Table S4). The global network graph of these interactions is shown in Supporting Information Figure S3, and the detailed parameters of the statistical analysis of the network are shown in Supporting Information Figure S2. Prior to the current study, the complete interaction network of acetylated proteins in prokaryote cells was determined only in the model bacteria species *E. coli*, in which the number of network nodes was 331 proteins and the number of direct physical interactions was 3690.[11] Therefore, this study presents the first high-quantity interaction network of acetylated proteins in prokaryotic cells.

In addition, previous studies in eukaryotes and prokaryotes have investigated subnetworks of acetylated proteins in different categories of cellular functions, such as the ribosome and cell cycle in human leukemia cells[4] and protein biosynthesis and RNA binding in bacterial cells.[11] Therefore, a similar subnetwork pathway enrichment analysis of acetylated proteins (e.g., ribosome, carbon metabolism, aminoacyl−tRNA biosynthesis, and TCA cycle; Figure 4B) was also performed using the large acetylome data set. The results show that acetylated proteins involved in a certain pathway comprise a dense protein interaction network. The subnetwork graphs of acetylated proteins in the pathways of ribosome and aminoacyl−tRNA biosynthesis are shown in Figure 6A and Supporting Information Table S4. The networks of ribosome and aminoacyl−tRNA biosynthesis consist of 982 and 162 direct physical interactions, respectively, indicating the important regulatory roles of the lysine acetylation modification in these two processes. Other subnetworks for the pathways of carbon metabolism, TCA cycle, glycolysis/gluconeogenesis, fatty acid biosynthesis, starch and sucrose metabolism, and methane metabolism are shown in Supporting Information Figure S4A and Table S4. These networks also have a relatively high density.

Furthermore, on the basis of the extraction of several special lysine acetylation motifs from the acetylated proteins, we also constructed a distinctive atlas of the protein interaction networks for each pathway (Figure 6B and Supporting Information Figure S4B). Using these networks, we can not only evaluate the physical interactions between acetylated proteins but also clearly identify which lysine motifs these proteins possess. This provides a better approach for understanding the functions of acetylated proteins in cells.

CONCLUSIONS

In this study, we identified the acetylation sites in the proteome of *V. parahemolyticus* using a highly sensitive proteomic method and provided the first comprehensive view of the acetylome of this organism. We identified 1413 lysine acetylation sites in 656 acetylated proteins in this bacterium, and these account for 13.6% of the total proteins in the cell. In addition, through extensive characterization of the acetylome, we found that the acetylated proteins are mostly associated with cellular functions and are distributed in different cellular compartments. However, the lysine modification of proteins related to protein biosynthesis and carbon metabolism may play particularly important physiological roles in *V. parahemolyticus*. Moreover, the analysis of the amino acid sequence motifs revealed that the acetylated lysine is surrounded by an alkaline residue with a long side chain (K or R) at the +4 or +5 position or by a residue with a large side chain (Y, H, or F) at the +1 or +2 position. These motifs are also found in human cells and other bacterial cells. The protein interaction networks resulting from the acetylome clearly indicate that a very wide range of regulatory roles are modulated by lysine acetylation in *V. parahemolyticus*.

ASSOCIATED CONTENT

Supporting Information

Figure S1, peptide MaxQunat score; Figure S2, parameters of network analysis; Figure S3, global network graph of protein interactions; Figure S4, interaction networks of some special pathways; Table S1, protein and functional annotation; Table S2, enrichment of KEGG pathway and GO annotation; Table S3, motif analysis; and Table S4, interaction networks. This material is available free of charge via the Internet at http://pubs.acs.org. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[27] under data set identifier PXD000924 and DOI 10.6019/PXD000924.

AUTHOR INFORMATION

Corresponding Author

*E-mail: panjy@zstu.edu.cn; Tel: +086 571 86843748; Fax: +086 571 86843745.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (31200110), the Natural Science Foundation of Zhejiang Province (Y2090396), the Commonweal Technology Application Project Program of Zhejiang Province (2011C23067), and the Zhejiang Provincial Top Key Discipline of Biology (2011C04) to J.P. as well as by a NSFC grant (31271362) to Z.C.

REFERENCES

(1) Phillips, D. M. The presence of acetyl groups of histones. *Biochem. J.* **1963**, *87*, 258−263.

(2) Chen, Y.; Zhao, W.; Yang, J. S.; Cheng, Z.; Luo, H.; Lu, Z.; Tan, M.; Gu, W.; Zhao, Y. Quantitative acetylome analysis reveals the roles of SIRT1 in regulating diverse substrates and cellular pathways. *Mol. Cell. Proteomics* **2012**, *11*, 1048−1062.

(3) Kim, S. C.; Sprung, R.; Chen, Y.; Xu, Y.; Ball, H.; Pei, J.; Cheng, T.; Kho, Y.; Xiao, H.; Xiao, L.; Grishin, N. V.; White, M.; Yang, X. J.; Zhao, Y. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* **2006**, *23*, 607−618.

(4) Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V.; Mann, M. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **2009**, *325*, 834−840.

(5) Zhao, S.; Xu, W.; Jiang, W.; Yu, W.; Lin, Y.; Zhang, T.; Yao, J.; Zhou, L.; Zeng, Y.; Li, H.; Li, Y.; Shi, J.; An, W.; Hancock, S. M.; He, F.; Qin, L.; Chin, J.; Yang, P.; Chen, X.; Lei, Q.; Xiong, Y.; Guan, K. L.

Regulation of cellular metabolism by protein lysine acetylation. *Science* **2010**, *327*, 1000−1004.
(6) Norris, K. L.; Lee, J. Y.; Yao, T. P. Acetylation goes global: the emergence of acetylation biology. *Sci. Signaling* **2009**, *2*, pe76.
(7) Norvell, A.; McMahon, S. B. Cell biology. Rise of the rival. *Science* **2010**, *327*, 964−965.
(8) Gray, M. W.; Burger, G.; Lang, B. F. Mitochondrial evolution. *Science* **1999**, *283*, 1476−1481.
(9) Yu, B. J.; Kim, J. A.; Moon, J. H.; Ryu, S. E.; Pan, J. G. The diversity of lysine acetylated proteins in *Escherichia coli*. *J. Microbiol. Biotechnol.* **2008**, *18*, 1529−1536.
(10) Zhang, J.; Sprung, R.; Pei, J.; Tan, X.; Kim, S.; Zhu, H.; Liu, C. F.; Grishin, N. V.; Zhao, Y. Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol. Cell. Proteomics* **2009**, *8*, 215−225.
(11) Zhang, K.; Zheng, S.; Yang, J. S.; Chen, Y.; Cheng, Z. Comprehensive profiling of protein lysine acetylation in *Escherichia coli*. *J. Proteome Res.* **2013**, *12*, 844−851.
(12) Wang, Q.; Zhang, Y.; Yang, C.; Xiong, H.; Lin, Y.; Yao, J.; Li, H.; Xie, L.; Zhao, W.; Yao, Y.; Ning, Z. B.; Zeng, R.; Xiong, Y.; Guan, K. L.; Zhao, S.; Zhao, G. P. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **2010**, *327*, 1004−1007.
(13) Okanishi, H.; Kim, K.; Masui, R.; Kuramitsu, S. Acetylome with structural mapping reveals the significance of lysine acetylation in *Thermus thermophilus*. *J. Proteome Res.* **2013**, *12*, 3952−3968.
(14) Ravindra, K. C.; Selvi, B. R.; Arif, M.; Reddy, B. A.; Thanuja, G. R.; Agrawal, S.; Pradhan, S. K.; Nagashayana, N.; Dasgupta, D.; Kundu, T. K. Inhibition of lysine acetyltransferase KAT3B/p300 activity by a naturally occurring hydroxynaphthoquinone, plumbagin. *J. Biol. Chem.* **2009**, *284*, 24453−24464.
(15) Bougdour, A.; Maubon, D.; Baldacci, P.; Ortet, P.; Bastien, O.; Bouillon, A.; Barale, J. C.; Pelloux, H.; Ménard, R.; Hakimi, M. A. Drug inhibition of HDAC3 and epigenetic control of differentiation in *Apicomplexa* parasites. *J. Exp. Med.* **2009**, *206*, 953−966.
(16) Whitaker, W. B.; Parent, M. A.; Boyd, A.; Richards, G. P.; Boyd, E. F. The *Vibrio parahaemolyticus* ToxRS regulator is required for stress tolerance and colonization in a novel orogastric streptomycin-induced adult murine model. *Infect. Immun.* **2012**, *80*, 1834−1845.
(17) Izutsu, K.; Kurokawa, K.; Tashiro, K.; Kuhara, S.; Hayashi, T.; Honda, T.; Iida, T. Comparative genomic analysis using microarray demonstrates a strong correlation between the presence of the 80-kilobase pathogenicity island and pathogenicity in Kanagawa phenomenon-positive *Vibrio parahaemolyticus* strains. *Infect. Immun.* **2008**, *76*, 1016−1023.
(18) Trosky, J. E.; Li, Y.; Mukherjee, S.; Keitany, G.; Ball, H.; Orth, K. VopA inhibits ATP binding by acetylating the catalytic loop of MAPK kinases. *J. Biol. Chem.* **2007**, *282*, 34299−34305.
(19) Makino, K.; Oshima, K.; Kurokawa, K.; Yokoyama, K.; Uda, T.; Tagomori, K.; Iijima, Y.; Najima, M.; Nakano, M.; Yamashita, A.; Kubota, Y.; Kimura, S.; Yasunaga, T.; Honda, T.; Shinagawa, H.; Hattori, M.; Iida, T. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. *Lancet* **2003**, *361*, 743−749.
(20) Miao, J.; Lawrence, M.; Jeffers, V.; Zhao, F.; Parker, D.; Ge, Y.; Sullivan, W. J., Jr.; Cui, L. Extensive lysine acetylation occurs in evolutionarily conserved metabolic pathways and parasite-specific functions during *Plasmodium falciparum* intraerythrocytic development. *Mol. Microbiol.* **2013**, *89*, 660−675.
(21) Shao, J.; Xu, D.; Hu, L.; Kwan, Y. W.; Wang, Y.; Kong, X.; Ngai, S. M. Systematic analysis of human lysine acetylation proteins and accurate prediction of human lysine acetylation through bi-relative adapted binomial score Bayes feature representation. *Mol. BioSyst.* **2012**, *8*, 2964−2973.
(22) Chou, M. F.; Schwartz, D. Biological sequence motif discovery using motif-x. *Current Protocols in Bioinformatics*; Wiley: New York, 2001; Chapter 13, pp 15−24.
(23) Lundby, A.; Lage, K.; Weinert, B. T.; Bekker-Jensen, D. B.; Secher, A.; Skovgaard, T.; Kelstrup, C. D.; Dmytriyev, A.; Choudhary, C.; Lundby, C.; Olsen, J. V. Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. *Cell Rep.* **2012**, *2*, 419−431.
(24) Melo-Braga, M. N.; Verano-Braga, T.; León, I. R.; Antonacci, D.; Nogueira, F. C.; Thelen, J. J.; Larsen, M. R.; Palmisano, G. Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botrana* infection. *Mol. Cell. Proteomics* **2012**, *11*, 945−956.
(25) Altman-Price, N.; Mevarech, M. Genetic evidence for the importance of protein acetylation and protein deacetylation in the halophilic archaeon *Haloferax volcanii*. *J. Bacteriol.* **2009**, *191*, 1610−1617.
(26) Jones, J. D.; O'Connor, C. D. Protein acetylation in prokaryotes. *Proteomics* **2011**, *11*, 3012−3022.
(27) Vizcaíno, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Ríos, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P. A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H. J.; Albar, J. P.; Martinez-Bartolomé, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H. ProteomeXchange provides globally co-ordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, *30*, 223−226.