Published in final edited form as:

J Proteomics. 2013 February 21; 79: 60–71. doi:10.1016/j.jprot.2012.12.001.

Differential lysine acetylation profiles of *Erwinia amylovora* strains revealed by proteomics

Xia Wu^{a,1}, Adaikkalam Vellaichamy^{b,2}, Dongping Wang^{c,3}, Leonid Zamdborg^b, Neil L. Kelleher^{b,d}, Steven C. Huber^{a,e,*}, and Youfu Zhao^{c,**}

^aDepartment of Plant Biology, University of Illinois, Urbana, IL 61801, USA

^bDepartment of Chemistry and Institute for Genomic Biology, University of Illinois, Urbana, IL 61801, USA

^cDepartment of Crop Sciences, University of Illinois, Urbana, IL 61801, USA

^dDepartment of Chemistry and Molecular Biosciences and the Feinberg School of Medicine, Northwestern University, Evanston, IL 60208, USA

^eU.S. Department of Agriculture - Agricultural Research Service, Urbana, IL 61801, USA

Abstract

Protein lysine acetylation (LysAc) has recently been demonstrated to be widespread in *E. coli* and *Salmonella*, and to broadly regulate bacterial physiology and metabolism. However, LysAc in plant pathogenic bacteria is largely unknown. Here we first report the lysine acetylome of *Erwinia amylovora*, an enterobacterium causing serious fire blight disease of apples and pears.

Immunoblots using generic anti-lysine acetylation antibodies demonstrated that growth conditions strongly affected the LysAc profiles in *E. amylovora*. Differential LysAc profiles were also observed for two *E. amylovora* strains, known to have differential virulence in plants, indicating translational modification of proteins may be important in determining virulence of bacterial strains. Proteomic analysis of LysAc in two *E. amylovora* strains identified 141 LysAc sites in 96 proteins that function in a wide range of biological pathways. Consistent with previous reports, 44% of the proteins are involved in metabolic processes, including central metabolism, lipopolysaccharide, nucleotide and amino acid metabolism. Interestingly, for the first time, several proteins involved in *E. amylovora* virulence, including exopolysaccharide amylovoran biosynthesis- and type III secretion-associated proteins, were found to be lysine acetylated, suggesting that LysAc may play a major role in bacterial virulence. Comparative analysis of LysAc sites in *E. amylovora* and *E. coli* further revealed the sequence and structural commonality for LysAc in the two organisms. Collectively, these results reinforce the notion that LysAc of proteins is widespread in bacterial metabolism and virulence.

© 2012 Elsevier B.V. All rights reserved.

*Correspondence to: S.C. Huber, 1201 W. Gregory Drive, 197 ERML, Urbana, IL 61801, USA. Tel.: +1 217 265 0909; fax: +1 217 244 4419; schuber1@illinois.edu (S.C. Huber). **Correspondence to: Y.F. Zhao, 1201 W. Gregory Drive, 288A ERML, Urbana, IL 61801, USA. Tel.: +1 217 265 9822; fax: +1 217 333 4582; zhao888@illinois.edu (Y. Zhao).

¹Current address: Department of Genome Sciences, University of Washington, Seattle, WA 98109, USA.

²Current address: Department of Nanomedicine, The Methodist Hospital Research Institute, Houston, TX 77030, USA.

³Current address: Department of Plant Pathology and Microbiology, Texas A&M University, College Station, TX 77843, USA.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jprot.2012.12.001>.

Keywords

Acetylome; Metabolism; Fire blight; Virulence; Type III secretion system; Exopolysaccharide

1. Introduction

Post-translational modifications (PTM) frequently occur to proteins and mediate their biological functions. Lysine acetylation (LysAc), a dynamic and reversible PTM, has emerged as a major PTM in both eukaryotes and prokaryotes [1,2]. Lysine acetylation normally refers to N^ε-acetylation, the transfer of an acetyl group from the acetyl donor acetyl-CoA to the ε-amino group of a specific lysine residue in a specific protein, resulting in acetyl-lysine (AcK) [3,4]. LysAc was first discovered in histone proteins about half a century ago [5,6], and has since been known to be crucial in regulating the functions of histones and transcription factors [7,8]. Comprehensive lysine acetylomics studies have significantly expanded the scope of LysAc beyond histone proteins and a more complete lysine acetylome has been reported from human, mouse, drosophila, plant, protozoan, and bacteria [1,2,9–14]. These proteomics studies revealed that LysAc is an evolutionarily conserved and widespread PTM [15,16] and demonstrated new roles of LysAc [1,2,7,17,18]. It is also well established that acetylation of proteins not only results in stimulation of DNA binding and thus gene expression, but also in regulating protein–protein interactions, protein stability and mRNA stability [3,4,15,19,20]. Furthermore, new findings point to an unexpected importance of LysAc in metabolic control and coordination of different metabolic pathways [2,21]. It is also expected that studies of LysAc will shed light on disease therapy or prevention by targeting lysine acetyltransferase and deacetylase, the enzymes that mediate the reversible protein LysAc [15].

Since the discovery of N^ε-acetylation of *Salmonella enterica* acetyl-CoA synthase (Acs) in 2002, several proteomics studies of α and γ proteobacteria, including *Escherichia coli* and *Salmonella enterica*, have identified several hundred acetylated proteins [2,13,22,23], indicating that a wide range of prokaryote proteins can be acetylated. Interestingly, about 50% of those identified proteins in bacteria are enzymes participating in multiple metabolic pathways and are important for the control of central metabolism, particularly energy, fatty acids and nucleotide metabolism. In addition, biochemical analysis of lysine acetylation on bacterial enzymes Acs and chemotaxis protein CheY demonstrated that site specific lysine acetylation directly modulated activities of both enzymes [24–26], indicating LysAc in bacteria can be functionally important. Moreover, given the fact that mitochondria, which are evolutionarily derived from bacteria, contain many acetylated proteins in mammals, LysAc in bacteria may be ubiquitous across genera and families [1,3,20,21]. However, no studies of protein LysAc on plant pathogenic bacteria have been reported so far.

Erwinia amylovora, an enterobacterium belonging to γ proteobacteria, causes fire blight disease, an economic important plant disease of the Rosaceae crops, including apples, pears and raspberries. The disease costs millions of dollars of crop losses annually around the world and its control has become a major concern for apple and pear industry [27]. Genetics studies in *E. amylovora* indicate that hypersensitive response and pathogenicity (*hrp*)-type III secretion system (T3SS) and the exopolysaccharide (EPS) amylovoran production are the

two major virulence factors [28,29]. The genomic sequences of at least four *E. amylovora* isolates have been reported, which share more than 99.99% sequence identity, and more than 98% of proteins are identical [27,30,31]. However, differential virulence has been observed for *Erwinia* isolates, while the underlying mechanism for the differences in virulence was largely unclear [32]. On the other hand, amylovoran is a carbon compound derived from primary carbon metabolites [33], and LysAc was recently shown to dynamically regulate enzymes in carbon metabolism [2,21]. We thus hypothesized that LysAc on metabolic enzymes in *E. amylovora* may play a regulatory role in its amylovoran production, and therefore may affect its virulence.

The objective of this study was to investigate the abundance of protein LysAc in *E. amylovora*, and compare the LysAc profiles of two natural isolates of *E. amylovora* strains, Ea273 from USA and Ea1189 from Germany, which have differential virulence in different host plants [32]. Using proteomic approaches, differential LysAc profiles for two *E. amylovora* isolates were documented and for the first time, several proteins involved in *E. amylovora* virulence were found to be lysine acetylated, including EPS amylovoran biosynthesis- and T3SS-associated proteins. The proteomics data of this study was acquired at a high resolution LTQ-FTICR mass spectrometry, which fully distinguishes LysAc from lysine trimethylation.

2. Materials and methods

2.1. Bacterial strains and culture media

The *E. amylovora* strains Ea1189 and Ea273 were either grown in LB medium or in MBMA medium (3 g KH₂PO₄, 7 g K₂HPO₄, 1 g [NH₄]₂SO₄, 2 ml glycerol, 0.5 g citric acid, 0.03 g MgSO₄) plus 1% sorbitol [34] as described previously [32]. Bacterial growth was monitored by measuring OD₆₀₀ and harvested at log or stationary phase as indicated. Samples for *E. amylovora* strains were processed simultaneously to allow side-by-side comparison.

2.2. Protein extraction

The *E. amylovora* cells were harvested by centrifugation and cell pellets were directly lysed by boiling in 2× SDS sample buffer and protein concentration was measured by Bradford assay (Bio-Rad, Hercules, CA). Thirty micrograms of total soluble proteins were loaded in each lane for SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) analysis. Alternatively, cells were lysed by sonication in a buffer containing competitive protease and deacetylase inhibitors (100 mM Tris pH 8.0, 5 mM caproic acid, 1 mM para-amino-benzamidine, 2 mM leupeptin, 5 μM PTACH and 2 μg/ml apicidin) (Sigma-Aldrich). Protein extracts were further fractionated by differential centrifugation at 20,000g and 100,000g (Beckman Coulter) for mass spectrometry analysis.

2.3. Anti-lysine acetylation immunoblots and acetyl-lysine peptide preparations

The generic anti-acetyl lysine antibodies (ImmuneChem Pharmaceuticals, Burnaby, CA) were used at a 1:1250 dilution for both 1D and 2D immunoblots to detect the overall LysAc. In acetyl-BSA competition assays, 30 μg/ml acetylated BSA (Ambion, Austin, TX) was

added during the primary antibody incubation, and simultaneously processed with other immunoblots.

Fractionated proteins were dissolved and denatured in the buffer containing 6 M urea and 50 mM ammonium bicarbonate. Proteins were reduced by dithiothreitol (DTT), alkylated by iodoacetamide (IAA), and further diluted to reduce the urea concentration to less than 1 M for enzymatic digestion. Aliquots of samples with equal total proteins were digested separately by trypsin (Promega, Madison, WI) or endo-protease Glu-C (Roche, Atlanta, GA) overnight. Digested peptides were collected in C18 SPE column (Alltech, Deerfield, IL) and dissolved in immune-precipitation buffer (50 mM MOPS pH 7.2, 10 mM sodium phosphate, 50 mM sodium chloride) [1]. Lysine acetylated peptides were affinity captured by anti-acetyl lysine agarose beads (ImmuneChem Pharmaceuticals, Burnaby, CA) following an overnight incubation, before elution in 0.1% TFA (pH 3.0). Peptides were dried down and resuspended in 0.1% formic acid for tandem mass spectrometry analysis.

2.4. LC-MS/MS (nanoLC and tandem mass spectrometry) analysis

The 12 T LTQ-FT Ultra (Thermo Fisher Scientific) interfaced with a 1D NanoLC (Eksigent Technologies, Dublin, CA) was used in this study. Briefly, peptides were loaded by autosampler and separated in reverse phase columns, which comprised a 3 cm trap column (C18, 200 Å, New Objective) and a 15 cm analytical column (C18, 100 Å, New Objective) in a 90 min LC gradient (0–60 min, 5–35% B; 60–75 min, 35–80% B; 75–90 min, 5% B equilibration; B=0.1% formic acid and 99.9% acetonitrile). Peptides were ionized by electrospray, followed by mass spectrometry identification. Data acquisition on the LTQ-FT instrument consisted of a full scan event on the FT analyzer (100,000 resolving power) and the data-dependent CID MS/MS scans on LTQ of the five most abundant peaks from the previous full scan. The dynamic exclusion was enabled with a repeat count of 3, an exclusion duration of 180 s, and a repeat duration of 30 s [35]. Analyses for each strain consisted of two biological and six technical runs.

The resulting raw files were converted to mzML and mgf files in Proteome Discover 1.2 (Thermo Fisher Scientific). The search database included the target protein sequence of *E. amylovora* ATCC 49946 (3565 entries) and the corresponding randomized decoy sequence. Search parameters included missed cleavage=3, precursor error tolerance=5 ppm, fragmentation error tolerance=0.7 Da, fixed modification of carbamidomethylation on cysteine, variable modifications of acetylation on lysine and oxidation on methionine and histidine. Peptide hits were sorted by peptide Expectation value (for Mascot hits) or XCorr (for Sequest hits) from high to low, and LysAc peptides with the 1% false discovery rate (FDR) cutoff were presented. Peptides identified with acetylation on the C-terminal lysine residues were manually verified. The plots of mass error distribution and expectation score distribution for all identified lysine acetylated peptides are presented in Fig. S1. For the spectral counting analysis, the data collected from two biological replicates and a total of six technical replicates were included for each strain. The parameters of total number of LysAc spectra, total number of unique LysAc peptides, and average of spectra number per unique LysAc peptides were compared to evaluate the relative abundance of LysAc between *E. amylovora* strains.

2.5. Two-dimensional gel electrophoresis (2DE-Gel) and MALDI-TOF analysis

Soluble proteins for isoelectric focusing electrophoresis (IEF) analysis were prepared by the phenol-tris and ammonium acetate precipitation method as described previously [12]. To minimize *in vitro* carbamylation on lysine residues, the protein dissolving step did not involve any sample heating and was performed in a timely manner. Equal 250 µg of total proteins were loaded on each IPG strip (pH 3–10, 13 cm, GE Healthcare) and proteins were separated by the in-gel rehydration method for 30 V for 360 Vhrs, 500 V for 500 Vhrs, 1000 V for 1000 Vhrs, and 8000 V for 36,000 Vhrs at 20 °C. Replicate IPG strips were prepared in an identical IEF run to achieve high reproducibility of the pI resolution. 12% (w/v) SDS-PAGE was used for second dimensional separation, and replicate strips created replicate gels for the subsequent analysis of anti-LysAc immunoblots or coomassie stain followed by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometer (MALDI-TOF) protein identification.

The abundant LysAc protein spots of Ea273 and Ea1189 in 2DE immunoblots were selected. Gel spots were digested with trypsin and released peptides were collected and concentrated using C18 zip-tips (Millipore, Billerica, MA) before analysis in the Voyager DE-STR (Applied Biosystems, Carlsbad, CA). The mass spectral between 500 and 4000 m/z were collected at the linear mode, and peaks were smoothened with the Gaussian method. The mass fingerprints of each spot was searched in Profound [36], with the settings of (database: NCBI nr; taxonomy: enterobacteria; protein mass: 0–250 kDa; protein pI: 0–14; using average mass; mass tolerance: 1.5 Da; Enzyme: Trypsin; allowing 1 missed cleavages; including partial modification of methionine oxidation). The protein hits that matched the taxa *E. amylovora* were reported in Table S1. “Remove the matched mass and search again” option was used to search the redundant proteins from a protein spot. The identified protein sequence was validated to insure that the predicted pI and MW matched its position in the 2DE gel.

The immunoblots were scanned in LI-COR Odyssey, with exposure setting consistent in membranes (700 channel, sensitivity: linear manual 5; brightness, 50; contrast, 50). Densitometry of LysAc protein spots was quantitated in Odyssey software (V 3.0). The densitometry was further normalized based on the intensity of the protein standard and a few most abundant LysAc protein spots (TufA, IcdA and LeuS) before statistical analysis. The coomassie stain membrane was scanned on Typhoon 8600 (GE Healthcare), using the settings of (fluorescence mode, 400 V normal sensitivity, 560 LP/Green (532 nm), focal plane=Platen). Spots densitometry was quantitated in ImageQuant software (V 5.2). Two technical replicates for each strain were analyzed.

3. Results

3.1. Growth conditions altered protein lysine acetylation profile in *Erwinia amylovora*

To examine the diversity and relative abundance of LysAc proteins in *E. amylovora*, immunoblots of bacterial cell lysates were probed with generic anti-lysine acetylation antibodies (Fig. 1). Bacterial strains Ea1189 and Ea273 were grown in both rich LB and MBMA minimal media and harvested at stationary phase. Strong immunoblot signals were

observed for both Ea1189 and Ea273 with equal protein loading (Fig. 1A and B), indicating that LysAc was abundant in both strains at both growth conditions. However, acetylation signals were much stronger when bacteria were grown in minimal medium than those in LB medium (Fig. 1B), suggesting that the growth medium dramatically altered the LysAc profile. Interestingly, strain Ea273 displayed higher LysAc level than that of Ea1189 when grown under the same conditions (Fig. 1B). This observation was confirmed by 2DE immunoblots of anti-lysine acetylation antibodies (Fig. 2A and B) and further characterized in detail by LC-MS/MS and MALDI-TOF (See below). In order to test the specificity of anti LysAc antibodies, acetyl-BSA competition assays were conducted. The majority of the protein LysAc signals were eliminated by premixing the antibodies with acetyl-BSA, indicating that the anti-lysine acetylation antibodies used were acetylation specific (Fig. 1C). Because of higher abundance of LysAc, *E. amylovora* strains grown in minimal medium at stationary phase were used for the mass spectrometry analyses.

3.2. Acetylated proteins in *Erwinia amylovora* strains identified by MALDI-TOF MS

To further characterize proteins that are acetylated in the two *Erwinia* strains, we compared LysAc proteins from both strains in IEF-2DE anti-LysAc immunoblots followed by identification of acetylated proteins by MALDI-TOF MS (Fig. 2). First, overall higher anti-LysAc signals from Ea273 than Ea1189 observed in the 1 DE analysis were recapitulated in the 2DE immunoblots. As shown in Fig. 2, majority of the proteins were resolved into distinct protein spots and the 2DE anti-LysAc immunoblots consistently detected overall higher abundance of LysAc for Ea273 over Ea1189, despite equal total protein loading. Second, to identify LysAc proteins, abundant LysAc protein spots were excised from the corresponding pI and MW positions in coomassie stained gels for MALDI-TOF peptide mass fingerprinting. Overall, a total of 76 LysAc proteins with protein IDs were identified by MALDI-TOF (Table S1). Many of these LysAc proteins are abundant at protein level, such as isocitrate dehydrogenase (IcdA, spot 34), elongation factor Tu (TusfA, spot 36), and glyceraldehyde-3-phosphate dehydrogenase A (GapA, spot 51). Other major LysAc proteins included phosphoenol pyruvate synthase (PpsA), sorbitol-6-phosphate dehydrogenase (SrlD), carbamoyl phosphate synthase large chain (CarB), surface protein (InlA) and stringent starvation protein A (SspA), heat shock protein (HtpG), and leucyl tRNA synthase (LeuS).

3.3. Identification of lysine acetylome in *Erwinia amylovora* by LC-MS/MS

To identify the LysAc sites for the LysAc proteins in *E. amylovora*, LysAc peptides were immunoaffinity enriched [1] and subject to LC-MS/MS analysis. The mass spectrometry analysis was performed in a high resolution LTQ-FTICR to fully distinguish LysAc from lysine trimethylation. A total of 141 unique LysAc sites on 96 LysAc proteins in *E. amylovora* strains was identified (Fig. 3, Table S2). Among them, 100 LysAc sites on 67 LysAc proteins were identified in strain Ea273, while 61 LysAc sites on 46 LysAc proteins were identified from strain Ea1189, and 20 LysAc sites on 17 proteins were common for both strains (Fig. 3A and B, Table S3). Compared to LysAc proteins identified by MALDI-TOF analysis above, 17 proteins were commonly identified by LC-MS/MS (Fig. 3C, Table S3). All of our identified peptides were confirmed within 5 ppm error tolerance. More than 85% of our identifications matched their theoretical mass within 1 ppm, which highlighted

the precision of FTICR mass spectrometry and the quality of the dataset (Fig. S1). In addition, spectral counting analysis was used to compare the relative abundance of LysAc between *E. amylovora* strains (Fig. 3D). Several lines of evidence suggest a higher abundance of LysAc in strain Ea273 compared to Ea1189. First, we identified a higher number of unique LysAc peptides and proteins in Ea273 than in Ea1189. Second, 444 total spectral counts for LysAc peptides were obtained for Ea273 compared to only 205 spectral counts for Ea1189. Third, Ea273 samples averaged 2.5 spectral counts per LysAc peptide per run, which was also higher than the averaged 2.0 spectral counts in Ea1189 samples. Therefore, the greater abundance of LysAc in strain Ea273 than that of Ea1189 was supported by total spectral counts for LysAc peptides, the number of identifications of unique LysAc peptides, and averaged counts per peptide. These observations are consistent with the anti-LysAc immunoblot analysis (Figs. 1 and 2).

To understand the biological significance of the LysAc proteins in *E. amylovora* strains, gene ontology analyses were performed (Fig. 4). The lysine acetylome in *E. amylovora* consisted of diverse proteins functioning in a wide variety of biological pathways. In particular, 44% of the identified proteins are involved in metabolic processes, including central carbon, lipopolysaccharide, nucleotide and amino acid metabolism. Such an extensive acetylation on metabolic enzymes was also found in previous studies in *E. coli* and *S. enterica* [2,23]. Statistical assessment revealed that categories of central carbon metabolism, protein translations and folding, polysaccharide and lipid metabolic process, ribosomal proteins and redox regulation proteins were overrepresented in the lysine acetylome, compared with the whole genome in *E. amylovora* (Fig. 4, Fig. S2). Most importantly, several acetylated proteins associated with *E. amylovora* virulence were identified as LysAc proteins, which included three proteins involved in EPS amylovan production (AmsJ, ThiI, and GalE), two LPS biosynthesis (KdsA and WzzE) and two T3SS-related proteins (HsvB and YopH) (Table 1, Fig. S2). In addition, several redox regulation proteins were also acetylated in *E. amylovora*, including ferredoxin-NADP reductase, thioredoxin, and glutaredoxin.

Because data dependent acquisition (DDA) in mass spectrometry tends to bias the identification towards abundant proteins, we extracted the top 100 most abundant proteins as reported in *E. coli* [37,38] and examined their gene ontology (Fig. 4). We found that the top 100 abundant proteins were also overrepresented in categories for metabolic enzymes and ribosomal proteins. Thus, the identified LysAc proteins were enriched in genes functioning in carbon metabolism and ribosomal proteins in *E. amylovora* may be partially due to the higher protein abundance of these proteins, and the preferential identification for these proteins in mass spectrometry with DDA method.

3.4. Comparison of *Erwinia amylovora* lysine acetylome to those of *E. coli*

To understand the sequence and structural commonality of LysAc between organisms, the lysine acetylome of *E. amylovora* identified by LC-MS/MS was first compared to the known LysAc proteins reported in *E. coli* [23]. Nine LysAc sites from eight proteins were commonly identified in both *E. coli* and *E. amylovora*, and another 24 proteins were commonly identified as LysAc proteins, but with different LysAc sites (Fig. 5A and B).

When both sets of data (LC-MS/MS and MALDI-TOF) were compared to those identified in *E. coli*, 32 common LysAc proteins were shared by both *E. coli* and *E. amylovora* (Fig. 5C, Table S3). The non-overlapped LysAc peptides/proteins may be attributed to different bacteria and their growth conditions as well as the different analytical approaches in the mass spectrometry [4]. In addition, both bacteria contained many lysine acetylated proteins with more than one LysAc site (Fig. 5D). On the other hand, the acetylated peptides of *E. amylovora* and *E. coli* shared a similar trend in terms of LysAc location in protein secondary structures. The majority of identified LysAc sites from both bacteria were found in loops or helices, with a slight preference for loops (Fig. 5E). This preference was also observed for the human lysine acetylome [1]. Two proteins (inorganic pyrophosphatase (Ppa) and glyceraldehyde-phosphate dehydrogenase (GapA)) had five LysAc sites identified in *E. amylovora*, which is similar to those reported for CheY in *E. coli* [39]. In terms of the acetylation motif, the preference for H, L, Y amino acids at the +1 position in *E. coli* was not observed in *E. amylovora* (Fig. 5F, Table S4) [23]. The sequence of acetylation sites was further classified based on secondary structures and the sequence motif was also analyzed. However, no clear motifs emerged from subgroups in *E. amylovora* (Fig. S3).

To compare the divergence of AcK residues, all the lysine acetylated proteins of *E. amylovora* were searched against the *E. coli* genome (Fig. S4A, Table S4). We found that 77% of the *E. amylovora* AcK residues were conserved with a lysine residue in *E. coli*; 12% of the *E. amylovora* acetylated proteins did not contain a homolog in *E. coli*, and 11% AcK was found to match a non-synonymous mutation in *E. coli*. Interestingly, these non-synonymous substitutions had the preference for E, Q, R, D residues, which correspond to a LysAc mimic (Q), reverse (D, E) or conservative substitution (R) (Fig. S4A, Table S4). To confirm that the observation was evolutionarily selective, a reverse search for *E. coli* LysAc proteins against the *E. amylovora* genome was performed. The majority (62%) of the *E. coli* sites of LysAc had a corresponding lysine residue in *E. amylovora*. With the exception of the 20% of the LysAc proteins that did not have an orthologous protein in *E. amylovora*, the rest of the non-conserved AcK residues were also more frequently substituted with E, Q, R, D, A residues (Fig. S4B, Table S4). These results might suggest the biological significance of the AcK and how organisms can maneuver the function of this PTM by amino acid substitutions.

4. Discussion

LysAc of proteins is a highly conserved, but reversible, dynamic, and regulated PTM in bacteria [15,20]. Recent proteomic studies indicate that LysAc is widespread from human to bacteria and affects cellular functions, especially carbohydrate and energy metabolism [40,41]. In this study, we present evidence that LysAc is also a robust PTM in plant pathogenic bacteria. Specifically, we demonstrated that LysAc profiles in *E. amylovora* were strongly altered under different growth conditions. In addition, differential LysAc profiles were observed for two *E. amylovora* strains that differ in their ability to cause disease on different apple genotypes. We confirmed that many (44%) acetylated proteins are involved in metabolic processes, ranging from central metabolism, lipopolysaccharide to nucleotide and amino acid metabolism. Several proteins involved in *E. amylovora* virulence, including amylovoran biosynthesis- and T3SS-associated proteins, were found to be lysine

acetylated. Our findings are consistent with the idea that the primary role of LysAc of proteins is to control carbon and energy metabolism, and also suggest that LysAc of proteins may play a role in regulating important bacterial virulence factors, such as EPS and T3SS.

4.1. Majority of acetylated proteins are involved in metabolism in bacteria identified by similar proteomics approaches

Trypsin-generated peptides analyzed with high performance LC-MS/MS is the typical approach used to identify LysAc sites of proteins. In theory, when trypsin is used as protease, acetylated lysine residues should only locate at internal positions of tryptic peptides, because trypsin cleaves after lysine/arginine residues by a series of nucleophilic reactions, that requires a positively charged residue [42]. While this rule was generally true (95% of acetylated lysine residues were located at internal positions), 8 tryptic peptides (5% of the total) that contained acetylated lysine at the C-terminus were identified in this study. However, their spectra and high mass accuracy relationship suggested that these peptides were not false positives (Fig. S2). Thus, while not a preferred reaction, it is clear that trypsin can cleave at acetylated lysine residues.

Similar proteomic studies of the acetylome have been performed in *E. coli*, *S. enterica*, and *Rhodopseudomonas palustris* [2,13,22,23]. Yu et al. [22] identified 125 acetylation sites on 85 proteins when *E. coli* strain W3110 was grown in LB medium at both log and stationary phases. Similarly, Zhang et al. [23] identified 138 acetylation sites on 91 proteins when *E. coli* strain MG1655 was grown in LB medium at log phase. Combining the results from both studies, a total of 260 acetylation sites on 144 proteins were found to be acetylated [3,41]. Further analysis indicated that only 9 sites on six proteins were common in both studies (Table S3), although 32 common proteins were found with different acetylation sites [4]. In addition, Wang et al. reported that 235 acetylated sites on 191 proteins were under LysAc modification in *S. enterica* grown in glucose-rich medium [2]. In this study, we identified 141 LysAc sites on 96 proteins in *E. amylovora* strains Ea273 and Ea1189 grown in minimal medium at stationary phase and 20 LysAc sites on 17 proteins were common in both strains (Table S3). It has been suggested that growth phase and bacterial strains may contribute to the difference of acetylated proteins or sites identified in *E. coli* [4,22,41]. Our findings confirm this assumption and further suggest that growth condition is another major factor, indicating the dynamic nature of protein LysAc [3].

Interestingly, the majority of bacterial LysAc proteins identified to date are involved in metabolism. 49% and 53% of acetylated proteins reported in two *E. coli* studies, respectively, are enzymes involved in metabolism of energy, fatty acids and nucleotides [22,23], whereas about 50% of the 191 acetylated proteins identified in *S. enterica* are also involved in metabolism [2]. In our case, about 44% acetylated proteins identified were enzymes involved in central metabolism, lipopolysaccharide to nucleotide and amino acid metabolism. Therefore, it appears that LysAc in bacteria often happens on proteins involved in glycolysis, gluconeogenesis, the TCA cycle, the glyoxylate bypass, glycogen biosynthesis, amino acid biosynthesis, fatty acid and nucleotide metabolism [2,22,23]. Since both growth medium and growth phase affect the extent of protein acetylation, it also appears that bacteria have developed or evolved LysAc to monitor energy generation and

consumption by either regulating the expression of central metabolic enzymes or their stability [3,4]. These results are consistent with previous findings that glucose and citrate feeding caused significant changes of the acetylation profile of metabolic enzymes in *Salmonella* [2], and that food intake or alcohol consumption can alter the LysAc profile in mouse liver [9,43]. It has been proposed that protein acetylation can act as regulatory mechanism as well as energy-storage mechanism, when energy is in excess [44]. In energy poor or rich medium, protein acetylation may govern how bacterial cells balance glycolysis versus oxidative metabolism, therefore determining storage or utilization of carbon energy [44]. However, over-representation of metabolic enzymes in the acetylome may also be partially attributed to the method of data-dependent-acquisition (DDA) of mass spectrometry in most of these studies. DDA detects the most abundant acetylated peptides, and metabolic enzymes generally are relatively abundant so that they are easier to be detected by mass spectrometry.

4.2. Protein acetylation may play a role in virulence

The virulence of *E. amylovora* is mainly attributed to two factors: (i) protein secretion and translocation system of *hrp*-T3SS and (ii) EPS amylovoran production [29,45]. However, we could not explain differential virulence or different host ranges of *E. amylovora* strains that occur in nature, as these strains share 99.99% genome identity and more than 98% of proteins are identical, including all known major virulence factors [27,32]. This study suggested that LysAc or other PTM regulation may play a role in this natural phenomenon. Indeed, seven proteins directly involved in pathogenesis in *E. amylovora* were identified as LysAc proteins (Table 1, Fig. S2), including two proteins associated with T3SS, an effector protein YopH and a *hrp*-associated systemic virulence B (HsvB) protein. HsvB is a conserved virulence protein, and is required for systemic virulence of *E. amylovora* in apples [29]. YopH encodes a tyrosine-protein phosphatase and dephosphorylates host proteins once translocated inside the host [46,47]. Interestingly, *Erwinia* YopH contained two tyrosine-protein phosphatase domains (Fig. S5). The identified acetyl lysine K451 is located at the second phosphatase domain and conserved with K342 of *Yersinia* YopH. The crystal structure of *Yersinia* YopH demonstrates that K342 directly interacts with the phospho-tyrosine substrate peptide through ionic interactions [48]. It will be interesting to test whether acetylation of K342/K451 in *Yersinia/Erwinia* YopH affects substrate binding or enzymatic activity.

Five acetylated proteins identified in this study are directly or indirectly involved in amylovoran biosynthesis (Table 1, Fig. S2) [33,49–51]. Both AmsJ and GalE are required for *E. amylovora* virulence. Interestingly, an AmsJ homolog enzyme (WcaA) involved in EPS production and a response regulator RcsB of EPS production are also reported to be acetylated in *E. coli* [52]. GalE, a UDP-glucose 4-epimerase, is involved in production of UDP-galactose, the building block for amylovoran biosynthesis [51]. The acetylated residue K245 (equivalent to human GalE K253) is very close to a natural GalE mutant in the human protein (K257R), which results in galactose epimerase deficiency, or GalE-deficiency galactosemia [53]. The K257R mutation inhibits UDP-glucose regeneration, preventing the formation of glucose-1-phosphate and leading to the accumulation of galactose and

galactose-1-phosphate [53]. It will be interesting to determine what effect K245 acetylation has on the function of GalE.

4.3. Conserved residues specific to protein structure are acetylated, but no acetylation motif is commonly found in acetylated proteins

The importance of LysAc was supported by the localization of specific sites in the protein structure. A number of AcK residues of proteins identified in *E. amylovora* were localized at critical regions in those proteins, including substrate binding, nucleotide binding or catalytic sites, as supported by previous biochemical analysis (Table 2). The chemotaxis response regulator CheY is acetylated *in vivo* at six lysine residues, which are clustered at the surface that binds CheA and CheZ [39]. Previous studies have also demonstrated that enzyme activities are controlled by acetylation both negatively and positively, such as promoting glyceraldehyde 3-phosphate dehydrogenase (GapA) to convert glyceraldehyde 3-phosphate to glycerate-1,3-phosphates [2], limiting phosphoglycerate kinase (Pkg) and phosphoenolpyruvate carboxykinase (PckA) [11,21]. On the other hand, the findings that many proteins are acetylated at multiple sites suggest acetylation may act in a hierarchical way to alter the function of target proteins[41].

Our findings showed that at positions where LysAc was found, a lysine residue was conserved in orthologous proteins of different organisms, and this has been noted in other studies as well [2,13]. While there appears to be specificity in acetylation, no acetylation motifs were found in *E. amylovora*, although some preference for His, Leu, Tyr residues at the +1 position surrounding the acetyl lysine in *E. coli* was previously reported [23]. Residues N-terminal to the modification site can also play a role, and a putative acetylation motif (PxxxxGK) for acetyl-CoA synthetase (ACS) was recently reported [13]. ACS is one of the central metabolic enzymes, which is well known to be acetylated in bacteria [3,4], and is important for conversion of acetate into high energy intermediate acetyl-CoA, which is essential for many cellular processes including lipid, amino acid biosynthesis and energy generation [54].

5. Conclusions

In summary, previous studies demonstrated that reversible LysAc is a conserved and dynamic protein modification in bacteria. Our findings reinforce the notion that LysAc of proteins is regulated and crucial in bacterial metabolism [40], and further suggest that LysAc may also play a role in bacterial virulence and survival under stress conditions [55,56]. The accumulated data on bacterial acetylome provides a platform for elucidating the molecular mechanisms of LysAc and understanding how bacteria adapt to a changing environment and host [57]. However, because of the reversible nature of LysAc, only a few studies have attempted to verify the results of proteomic studies, such as those conducted with ACS, response regulator CheY, GapA, isocitrate lyase (AceA) and isocitrate dehydrogenase kinase/phosphatase (AceK) [2,13,52]. Therefore, future work should be focused on verification and function studies of protein LysAc and the role of acetylation in pathogenic bacteria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs. Furong Sun and Haijun Yao for the help on MALDI-TOF analysis and Dr. P. Yau at the proteomics facility, Keck Center of the University of Illinois for helpful discussions. We also thank the Campus Research Board, University of Illinois for providing partial funding for this research (Award #10192). This project was supported by the Agriculture and Food Research Initiative Competitive Grants Program Grant no. 2010-65110-20497 from the USDA National Institute of Food and Agriculture (YFZ). LZ is supported by NIDA grant 1F30DA026672.

Abbreviations

2DE	two-dimensional gel electrophoresis
AcK	acetyl-lysine
IEF	isoelectric focusing electrophoresis
LC-MS/MS	liquid chromatography and tandem mass spectrometry
LysAc	lysine acetylation
MALDI-TOF	matrix-assisted laser desorption/ionization-time-of-flight mass spectrometer
MW	molecular weight
pI	Isoelectric point
PTM	post-translational modification
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis

References

1. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*. 2009; 325:834–40. [PubMed: 19608861]
2. Wang Q, Zhang Y, Yang C, Xiong H, Lin Y, Yao J, et al. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science*. 2010; 327:1004–7. [PubMed: 20167787]
3. Hu LI, Lima BP, Wolfe AJ. Bacterial protein acetylation: the dawning of a new age. *Mol Microbiol*. 2010; 77:15–21. [PubMed: 20487279]
4. Thao S, Escalante-Semerena JC. Control of protein function by reversible N^ε-lysine acetylation in bacteria. *Curr Opin Microbiol*. 2011; 14:200–4. [PubMed: 21239213]
5. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci U S A*. 1964; 51:786–94. [PubMed: 14172992]
6. Phillips DM. The presence of acetyl groups of histones. *Biochem J*. 1963; 87:258–63. [PubMed: 13943142]
7. Kouzarides T. Acetylation: a regulatory modification to rival phosphorylation? *EMBO J*. 2000; 19:1176–9. [PubMed: 10716917]
8. Sterner DE, Berger SL. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*. 2000; 64(2):435–59. [PubMed: 10839822]

9. Yang L, Vaitheesvaran B, Hartil K, Robinson AJ, Hoopmann MR, Eng JK, et al. The fasted/fed mouse metabolic acetylome: N6-acetylation differences suggest acetylation coordinates organ-specific fuel switching. *J Proteome Res.* 2011; 10:4134–49. [PubMed: 21728379]
10. Weinert BT, Wagner SA, Horn H, Henriksen P, Liu WR, Olsen JV, et al. Proteome-wide mapping of the *Drosophila* acetylome demonstrates a high degree of conservation of lysine acetylation. *Sci Signal.* 2011; 4:ra48. [PubMed: 21791702]
11. Finkemeier I, Laxa M, Miguet L, Howden AJ, Sweetlove LJ. Proteins of diverse function and subcellular location are lysine acetylated in *Arabidopsis*. *Plant Physiol.* 2011; 155:1779–90. [PubMed: 21311031]
12. Wu X, Oh MH, Schwarz EM, Larue CT, Sivaguru M, Imai BS, et al. Lysine acetylation is a widespread protein modification for diverse proteins in *Arabidopsis*. *Plant Physiol.* 2011; 155:1769–78. [PubMed: 21311030]
13. Crosby HA, Pelletier DA, Hurst GB, Escalante-Semerena JC. System-wide studies of N-lysine acetylation in *Rhodospseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *J Biol Chem.* 2012; 287:15590–601. [PubMed: 22416131]
14. Jeffers V, Sullivan WJ Jr. Lysine acetylation is widespread on proteins of diverse function and localization in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell.* 2012; 11:735–42. [PubMed: 22544907]
15. Kim GW, Yang XJ. Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends Biochem Sci.* 2011; 36:211–20. [PubMed: 21075636]
16. Waterborg JH. Plant histone acetylation: in the beginning. *Biochim Biophys Acta.* 2011; 1809:353–9. [PubMed: 21371578]
17. Kim SC, Sprung R, Chen Y, Xu Y, Ball H, Pei J, et al. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell.* 2006; 23:607–18. [PubMed: 16916647]
18. Yang XJ, Seto E. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Mol Cell.* 2008; 31:449–61. [PubMed: 18722172]
19. Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin - Second in review series on chromatin dynamics. *EMBO Rep.* 2002; 3:224–9. [PubMed: 11882541]
20. Close P, Creppe C, Gillard M, Ladang A, Chapelle J, Nguyen L, et al. The emerging role of lysine acetylation of non-nuclear proteins. *Cell Mol Life Sci.* 2010; 67:1255–64. [PubMed: 20082207]
21. Zhao SM, Xu W, Jiang WQ, Yu W, Lin Y, Zhang TF, et al. Regulation of cellular metabolism by protein lysine acetylation. *Science.* 2010; 327:1000–4. [PubMed: 20167786]
22. Yu BJ, Kim JA, Moon JH, Ryu SE, Pan JG. The diversity of lysine-acetylated proteins in *Escherichia coli*. *J Microbiol Biotechnol.* 2008; 18:1529–36. [PubMed: 18852508]
23. Zhang JM, Sprung R, Pei JM, Tan XH, Kim S, Zhu H, et al. Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol Cell Proteomics.* 2009; 8:215–25. [PubMed: 18723842]
24. Starai VJ, Escalante-Semerena JC. Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica*. *J Mol Biol.* 2004; 340:1005–12. [PubMed: 15236963]
25. Ramakrishnan R, Schuster M, Bourret RB. Acetylation at Lys-92 enhances signaling by the chemotaxis response regulator protein CheY. *Proc Natl Acad Sci U S A.* 1998; 95:4918–23. [PubMed: 9560203]
26. Barak R, Eisenbach M. Acetylation of the response regulator, CheY, is involved in bacterial chemotaxis. *Mol Microbiol.* 2001; 40:731–43. [PubMed: 11359578]
27. Zhao Y, Qi M. Comparative genomics of *Erwinia amylovora* and related *Erwinia* species—What do we learn? *Genes.* 2011; 2:627–39. [PubMed: 24710213]
28. Zhao Y, Sundin GW, Wang D. Construction and analysis of pathogenicity island deletion mutants of *Erwinia amylovora*. *Can J Microbiol.* 2009; 55:457–64. [PubMed: 19396246]
29. Oh CS, Kim JF, Beer SV. The Hrp pathogenicity island of *Erwinia amylovora* and identification of three novel genes required for systemic infection. *Mol Plant Pathol.* 2005; 6:125–38. [PubMed: 20565644]

30. Sebaihia M, Bocsanczy AM, Biehl BS, Quail MA, Perna NT, Glasner JD, et al. Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. *J Bacteriol.* 2010; 192:2020–1. [PubMed: 20118253]
31. Smits TH, Rezzonico F, Kamber T, Blom J, Goesmann A, Frey JE, et al. Complete genome sequence of the fire blight pathogen *Erwinia amylovora* CFBP 1430 and comparison to other *Erwinia* spp. *Mol Plant Microbe Interact.* 2010; 23:384–93. [PubMed: 20192826]
32. Wang DP, Korban SS, Zhao YF. Molecular signature of differential virulence in natural isolates of *Erwinia amylovora*. *Phytopathology.* 2010; 100:192–8. [PubMed: 20055653]
33. Nimtz M, Mort A, Domke T, Wray V, Zhang Y, Qiu F, et al. Structure of amylovoran, the capsular exopolysaccharide from the fire blight pathogen *Erwinia amylovora*. *Carbohydr Res.* 1996; 287:59–76. [PubMed: 8765060]
34. Bellemann P, Bereswill S, Berger S, Geider K. Visualization of capsule formation by *Erwinia amylovora* and assays to determine amylovoran synthesis. *Int J Biol Macromol.* 1994; 16:290–6. [PubMed: 7537077]
35. Lee JE, Atkins N, Hatcher NG, Zamdborg L, Gillette MU, Sweedler JV, et al. Endogenous peptide discovery of the rat circadian clock: a focused study of the suprachiasmatic nucleus by ultrahigh performance tandem mass spectrometry. *Mol Cell Proteomics.* 2010; 9:285–97. [PubMed: 19955084]
36. Zhang WZ, Chait BT. Profound: an expert system for protein identification using mass spectrometric peptide mapping information. *Anal Chem.* 2000; 72:2482–9. [PubMed: 10857624]
37. Taniguchi Y, Choi PJ, Li GW, Chen HY, Babu M, Hearn J, et al. Quantifying *E. coli* proteome and transcriptome with single molecule sensitivity in single cells. *Science.* 2010; 329:533–8. [PubMed: 20671182]
38. Ishihama Y, Schmidt T, Rappsilber J, Mann M, Hartl FU, Kerner MJ, et al. Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics.* 2008; 9:102. [PubMed: 18304323]
39. Eisenbach M. A hitchhiker's guide through advances and conceptual changes in chemotaxis. *J Cell Physiol.* 2007; 213:574–80. [PubMed: 17708539]
40. Guan KL, Xiong Y. Regulation of intermediary metabolism by protein acetylation. *Trends Biochem Sci.* 2011; 36:108–16. [PubMed: 20934340]
41. Jones JD, O'Connor CD. Protein acetylation in prokaryotes. *Proteomics.* 2011; 11:3012–22. [PubMed: 21674803]
42. Leiros HKS, Brandsdal BO, Andersen OA, Os V, Leiros I, Helland R, et al. Trypsin specificity as elucidated by LIE calculations, X-ray structures, and association constant measurements. *Protein Sci.* 2004; 13:1056–70. [PubMed: 15044735]
43. Fritz KS, Galligan JJ, Hirschey MD, Verdin E, Petersen DR. Mitochondrial acetylome analysis in a mouse model of alcohol-induced liver injury utilizing SIRT3 knockout mice. *J Proteome Res.* 2012; 11:1633–43. [PubMed: 22309199]
44. Guarente L. The logic linking protein acetylation and metabolism. *Cell Metab.* 2011; 14:151–3. [PubMed: 21803285]
45. Zhao Y, Wang D, Nakka S, Sundin GW, Korban SS. Systems level analysis of two-component signal transduction systems in *Erwinia amylovora*: role in virulence, regulation of amylovoran biosynthesis and swarming motility. *BMC Genomics.* 2009; 10:245. [PubMed: 19470164]
46. Bliska JB, Guan KL, Dixon JE, Falkow S. Tyrosine phosphate hydrolysis of host proteins by an essential *Yersinia* virulence determinant. *Proc Natl Acad Sci U S A.* 1991; 88:1187–91. [PubMed: 1705028]
47. Guan KL, Dixon JE. Protein tyrosine phosphatase activity of an essential virulence determinant in *Yersinia*. *Science.* 1990; 249:553–6. [PubMed: 2166336]
48. Ivanov MI, Stuckey JA, Schubert HL, Saper MA, Bliska JB. Two substrate-targeting sites in the *Yersinia* protein tyrosine phosphatase co-operate to promote bacterial virulence. *Mol Microbiol.* 2005; 55:1346–56. [PubMed: 15720545]
49. Berry MC, McGhee GC, Zhao Y, Sundin GW. Effect of a waaL mutation on lipopolysaccharide composition, oxidative stress survival, and virulence in *Erwinia amylovora*. *FEMS Microbiol Lett.* 2009; 291:80–7. [PubMed: 19076232]

50. McGhee GC, Sundin GW. Thiamin biosynthesis and its influence on exopolysaccharide production: a new component of virulence identified on *Erwinia amylovora* plasmid pEA29. *Acta Hortic.* 2008; 293:271–7.
51. Metzger M, Bellemann P, Bugert P, Geider K. Genetics of galactose metabolism of *Erwinia amylovora* and its influence on polysaccharide synthesis and virulence of the fire blight pathogen. *J Bacteriol.* 1994; 176:450–9. [PubMed: 7507102]
52. Thao S, Chen CS, Zhu H, Escalante-Semerena JC. N^ε-lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PLoS One.* 2010; 5:15123.
53. Timson DJ. Functional analysis of disease-causing mutations in human UDP-galactose 4-epimerase. *FEBS J.* 2005; 272(23):6170–7. [PubMed: 16302980]
54. Garrity J, Gardner JG, Hawse W, Wolberger C, Escalante-Semerena JC. N-Lysine propionylation controls the activity of propionyl-CoA synthetase. *J Biol Chem.* 2007; 282:30239–45. [PubMed: 17684016]
55. Ma Q, Wood TK. Protein acetylation in prokaryotes increases stress resistance. *Biochem Biophys Res Commun.* 2011; 410:846–51. [PubMed: 21703240]
56. Lima BP, Antelmann H, Gronau K, Chi BK, Becher D, Brinsmade SR, et al. Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. *Mol Microbiol.* 2011; 81:1190–204. [PubMed: 21696463]
57. Mischerikow N, Heck AJ. Targeted large-scale analysis of protein acetylation. *Proteomics.* 2011; 11:571–89. [PubMed: 21246731]
58. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003; 13:2498–504. [PubMed: 14597658]
59. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics.* 2005; 21:3448–9. [PubMed: 15972284]
60. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res.* 2004; 14:1188–90. [PubMed: 15173120]
61. Heikinheimo P, Pohjanjoki P, Helminen A, Tasanen M, Cooperman BS, Goldman A, et al. A site-directed mutagenesis study of *Saccharomyces cerevisiae* pyrophosphatase - Functional conservation of the active site of soluble inorganic pyrophosphatases. *Eur J Biochem.* 1996; 239:138–43. [PubMed: 8706698]
62. Lahti R, Salminen T, Latonen S, Heikinheimo P, Pohjanoksa K, Heinonen J. Genetic engineering of *Escherichia coli* inorganic pyrophosphatase - Tyr55 and Tyr141 are important for the structural integrity. *Eur J Biochem.* 1991; 198:293–7. [PubMed: 1645654]
63. Doyle SA, Beernink PT, Koshland DE Jr. Structural basis for a change in substrate specificity: crystal structure of S113E isocitrate dehydrogenase in a complex with isopropylmalate, Mg²⁺, and NADP. *Biochemistry.* 2001; 40:4234–41. [PubMed: 11284679]
64. Cherbavaz DB, Lee ME, Stroud RM, Koshland DE. Active site water molecules revealed in the 2.1 angstrom resolution structure of a site-directed mutant of isocitrate dehydrogenase. *J Mol Biol.* 2000; 295:377–85. [PubMed: 10623532]
65. Boel G, Pichereau V, Mijakovic I, Maze A, Poncet S, Gillet S, et al. Is 2-phosphoglycerate-dependent automodification of bacterial enolases implicated in their export? *J Mol Biol.* 2004; 337:485–96. [PubMed: 15003462]
66. Ladine JR, Cross RL. The Adenine nucleotide-binding site on yeast 3-phosphoglycerate kinase - Affinity labeling of Lys-131 by pyridoxal 5'-diphospho-5'-adenosine. *J Biol Chem.* 1991; 266:7194–8. [PubMed: 1901864]
67. Czworkowski J, Wang J, Steitz TA, Moore PB. The crystal structure of elongation factor G complexed with GDP at 2.7-angstrom resolution. *EMBO J.* 1994; 13:3661–8. [PubMed: 8070396]

Fig. 1. Anti-lysine acetylation immunoblots showing that strains and growth medium strongly altered acetylation profiles of *Erwinia amylovora*. (A) Blots stained with Coomassie Brilliant Blue showing the consistency of protein loading. (B) The overall LysAc profile in Ea1189 and Ea273 grown in LB and MBMA medium detected by probing with generic anti-lysine acetylation antibodies. (C) Specificity of acetylation signals validated by acetyl-BSA competition assay. 30 μ g of total proteins was analyzed in each lane.

Fig. 2.
Resolution of LysAc proteins of *E. amylovora* strains Ea273 and Ea1189 by IEF-2DE. IEF-2DE separation for LysAc proteins in Ea273 (A) and Ea1189 (B) harvested from MBMA medium at stationary phase. LysAc proteins were detected by anti-LysAc immunoblots. Equal total proteins were loaded for both strains for each IEF strip (pI 3–10). Two technical replicates were analyzed. Abundant LysAc proteins were further excised from coomassie stain gels at the corresponding positions, and identified by MALDI-TOF mass fingerprinting.

Fig. 3. Comparative analysis of LysAc profiles of *E. amylovora* strains Ea1189 and Ea273 by LC-MS/MS. (A) Venn diagram showing the number of LysAc sites in Ea1189 and Ea273. (B) Venn diagram showing the number of LysAc proteins in Ea1189 and Ea273. (C) Venn diagram showing the total number of acetylated proteins identified by LC-MS/MS and 2DE-MALDI-TOF. (D) Spectral counting analysis for LysAc peptides in Ea1189 and Ea273 identified by LC-MS/MS. Total spectral counts: sum of the spectral counts for six LC-MS/MS runs for each strain. Ea1189: 205, Ea273: 444. Total LysAc peptides: sequence specific LysAc peptides identified. Ea1189: 62, Ea273: 105. Average counts per peptides per run: total spectral counts divided by the number of unique LysAc peptides. Ea1189: 2.0, Ea273: 2.5. The spectral analysis indicated higher abundance of LysAc in strain Ea273 compared to Ea1189.

Fig. 4.

Gene ontology analysis of lysine acetylome in *E. amylovora* strain Ea273. The analysis included 67 LysAc proteins with LysAc sites. The percentage of the categories was compared to the genes in the whole genome (labeled as Genome) or to the top 100 abundant proteins (labeled as Top 100) [37,38]. Analyses were performed in Cytoscape (V 2.8.3) BiNGO (V 2.42) [58,59]. Hypergeometric test with Benjamini and Hochberg false discovery rate correction was used for statistical analysis, and the significance cutoff was $P < 0.05$ (*) and $P < 0.01$ (**). The overrepresented gene ontology categories were indicated.

Fig. 5. Sequence and structural commonality of lysine acetylome in *E. amylovora* and *E. coli*. (A) Venn diagram showing the number of LysAc sites in *E. amylovora* by LC-MS/MS and in *E. coli*. (B) Venn diagram showing the number of LysAc proteins in *E. amylovora* by LC-MS/MS and in *E. coli*. (C) Venn diagram showing the total number of acetylated proteins identified by LC-MS/MS and 2DE-MALDI-TOF in *E. amylovora* and compared to those reported in *E. coli* [23]. (D) Number of LysAc proteins contained more than one LysAc sites in both *E. amylovora* and *E. coli*. (E) The identified LysAc sites were found to primarily occur in loops and helices in the protein structure in both *E. amylovora* and *E. coli*. (F) Sequence plot showing six amino acid residues up and downstream of all acetylated lysine residues. The plot was generated with the WebLogo tool [60].

Table 1

Lysine acetylated proteins associated with *E. amylovora* virulence.

Protein	Annotation	Sites	Peptides	Strains	Virulence function
AmsJ	Amylovoran biosynthesis protein	K241	ELAPFDK ¹ R	273	Amylovoran production
GaiE	UDP-glucose 4-epimerase	K245	GHLKALDHL ¹ SAIEGYKAYNLGAGKGYSVLE	1189	Galactose metabolism; amylovoran production
ThiI	Thiamine biosynthesis protein	K105 K122	GKTFc ² VRVRRGKHEFFSSQDVE	273	Thiamine biosynthesis
KdsA	2-Dehydro-3-deoxyphosphooctonate aldolase	K60	ASFDkANR	273	Lipopolysaccharide biosynthesis
WzzE	Lipopolysaccharide biosynthesis protein	K122	EFWLHSSYYQNRKSGNAR	1189	Lipopolysaccharide biosynthesis
HsvB	Hrp-associated systemic virulence protein	K120	MREIAQQAGIKTAR	1189	Type III associated
YopH	Tyrosine-protein phosphatase	K451	AVAYSMSISDDKVKINVPVVh ³ IYNWPVTGRPSK	273	Type III effector

Protein modifications:

¹ k: lysine acetylation,

² c: cysteine carbamidomethylation,

³ h: histidine oxidation.

Table 2

Lysine acetylated proteins with functional implications from other studies.

Protein	Annotation	Sites	Peptides	Mass (ppm)	Expectation value	Strains	Functional evidence
Ppa	Inorganic pyro-phosphatase	K30	IPANAAPIK ^a YE	0.18	8.22E-04	1189	K30R 2% activity ^b
		K143	GQISHFFEQYKALEK	-0.13	5.78E-06	273	Y142F 22% activity ^c
IcdA	Isocitrate dehydrogenase	K344	ATHGTAPKYAGQDKVNPQSVLSAE	0.35	7.29E-05	1189, 273	Residues 339–345: NADP binding site ^d
		K235	GNIMRTEGAFK	0.19	2.49E-05	273	Adjacent to substrate binding site K230. K230M decreased in activity and substrate affinity ^e
Eno	Enolase	K342	GIANSILIRFNQIGSLTETLAALK	0.73	2.68E-05	1189, 273	Substrate binding site, K342Q 1% activity ^f
Pgk	Phosphor-glycerate kinase	K117	FNKGEKKDDEILSK	0.32	6.86E-05	273	Adjacent to substrate binding site R113 ^g
FusA	Elongation factor G	K143	IAFVNkMDR	0.35	3.12E-05	273	Residues 14–145 was the GTP binding site ^h

Protein modifications:

^a k: lysine acetylation. Functional evidence:

^b K30R *ppa* mutant in *E. coli* contained only 0.4% relative activity as the WT [61].

^c Y142F *ppa* mutant in *E. coli* had decreased thermostability and contained 22% relative activity as the WT [62].

^d K344 was located at the NADP binding region of IcdA [63].

^e K235 was close to the substrate binding site K230. K230M IcdA in *E. coli* had less substrate binding affinity and decreased activity as the WT [64].

^f K342 interacted with substrate 2-phosphoglycerate, and mutant K342Q in *E. coli* retained only 1% relative activity as the WT [65].

^g K117 was adjacent to the substrate binding site R113 [66].

^h K143 was located within the GTP binding region [67].