Novel bile acid-mediated cysteine modification regulates bacterial virulence

Yanqi Wang[1,2#], Chang Zheng[3,4#], Ying Chen[2], Yicheng Liu[2], Ziming Tang[2], Shibo Yin[2], Yuan Liu[2], Yuan-Fei Zhou[2], Shouli Yuan[1], Peng Xu[3], Biao Yu[3*], Chu Wang[1,2,4*]

[1]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China.

[2]Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, College of Chemistry and Molecular Engineering, Peking University, Beijing, China.

[3]State Key Laboratory of Chemical Biology, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200032, China

[4]Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies, Chengdu 610041, China

* Correspondence should be addressed to chuwang@pku.edu.cn; byu@sioc.ac.cn

# These authors contribute equally to this work

Abstract

Bile acids (BAs) are a class of endogenous metabolites with important roles in fat digestion, vitamin absorption and body signaling. Synthesized from cholesterol in the liver and extensively transformed by the gut microbiota, BAs undergo enterohepatic circulation that places them at a key interface of host–microbe interactions. Current research has focused exclusively on how bile acids interact with the key protein targets in the host and bacteria to transduce signals, however, their potential of forming post-translational modifications (PTMs) in proteomes remains entirely unexplored. We herein report the discovery of a novel cysteine PTM mediated by lithocholic acid (LCA) that is highly conserved across diverse bacterial species. Structural characterization using a series of LCA analogue probes revealed that this modification is formed selectively through C6 of the steroid ring. Functionally, LCA modification disrupts the early stage of *Salmonella* infection by suppressing its virulence effector SipA and targeting the bacterial protein translocation system. Collectively, our study unveils a previously unrecognized covalent mechanism of BA-mediated regulation, which will open a new venue for exploring the host-microbe interaction and co-metabolism as well as related disorders.

Key words: Bile acid; Cysteine; Novel PTM; Bacterial Virulence; Chemoproteomics

Introduction

Bile acids (BAs) are a class of important endogenous metabolites that consist of a steroid core and a side chain with a carboxyl group. As amphipathic metabolites, primary BAs such as cholic acid (CA) and chenodeoxycholic acid (CDCA) are synthesized in liver and secreted into intestinal lumen, where they facilitate the emulsification and absorption of dietary lipids[1,2]. Beyond their well-established role in nutrient digestion, BAs also function as signaling molecules by activating nuclear receptors[3] and G protein-coupled receptors (GPCRs)[4], thereby orchestrating lipid and glucose homeostasis.

Following their secretion into intestinal lumen, primary BAs undergo extensive microbial biotransformation through deconjugation, dehydrogenation and dehydroxylation, generating secondary BAs[5] such as deoxycholic acid (DCA) and lithocholic acid (LCA). These metabolites can exert pleiotropic effects on the intestinal ecosystem including inducing genomic instability[6], regulating energy metabolism[7,8], maintaining immune homeostasis[9-11] and balancing the cecal microbiota to defend enteric pathogens[12,13]. Recent discovery of new BA conjugates has highlighted the capacity of gut microbiota to remodel the BA pool and generate bioactive metabolites that modulate the host physiology[14-18]. Meanwhile, these microbiota-derived BAs also contribute to shaping the composition and function of the gut microbial community[15], establishing a dynamic bidirectional communication axis between the host and microbiota.

In light of the increasingly recognized roles of BAs in diverse physiological processes, chemoproteomic studies have been performed to capture BA-interacting proteins in both host and bacteria proteomes for functional receptor discovery. In particular, photo-affinity probes of different BA species have been synthesized to study the mechanisms underlying the anti-infective effects of BAs as well as the bacterial tolerance to BAs [19-23]. However, such studies have so far focused exclusively on the

non-covalent interactions between BA and proteomes, and to our best knowledge, covalent modification of proteins by BA has never been reported. Inspired by the pioneering study that cholesterol, a metabolic precursor of BAs, regulates biological processes through covalent modification of key proteins in the Hedgehog pathway[24,25], we are motivated to explore the possibility of covalent BA modifications in proteomes as well as their potential functions.

In the current study, we applied a series of bioorthogonal BA analogue probes to systematically profile their covalent modifications in proteomes. It was revealed that an LCA-based probe could covalently modify cysteine residues and yield a fixed modification that targeted hundreds of proteins in living bacteria. Importantly, native LCA exhibited identical cysteine-modifying activity and the modification event was highly conserved across diverse bacterial species. By synthesizing an array of LCA derivatives and crosschecking with samples after exhaustive proteome hydrolysis, we unambiguously determined that LCA modifies cysteine residues specifically via the C6 position of its steroid skeleton. Functionally, we demonstrated that the LCA modification plays an important role in modulating bacterial protein secretion and counteracting pathogenic infection.

Results

Proteomic discovery of an LCA-mediated modification

During the course of applying photo-affinity BA probes to identify their bacterial receptors[20], we serendipitously observed that an LCA-based probe, 3-x-LCA-Galk, exhibited strong labeling in bacterial proteomes even without UV irradiation **(Fig. S1)**, suggesting that it may directly mediate covalent modification of bacterial proteins. To investigate the origin of such labeling, the photo-affinity moiety was removed to generate an "LCA-Galk" probe by conjugating the propargylglycine group with LCA **(Fig. 1a)**. Similarly, the conjugation was performed with another three highly abundant BA species in humans including CA, CDCA and DCA **(Fig. 1a)**. We hypothesized this panel of "BA-Galk" probes could closely mimic their native counterparts to enable profiling of any possible covalent modification events.

We initially characterized the labeling efficiency of these four BA-Galk probes following their metabolic labeling of living bacteria via in-gel fluorescence **(Fig. S2a)**. Interestingly, the results demonstrated that only LCA-Galk exhibited a robust labeling profile **(Fig. 1b)**, which could also be outcompeted by the addition of excess native LCA **(Fig. 1c)**. Furthermore, this proteome-wide labeling was both time- and dose-dependent **(Fig. S3)**. Notably, the labeling was achieved exclusively in living bacteria but not in cell lysates **(Fig. S2b)**.

We next applied LCA-Galk in a chemoproteomic workflow to identify its modified proteins and sites **(Fig. 1d)**. After metabolic labeling with 150 μM of LCA-Galk for 6 h, bacteria were lysed and the probe-labeled proteomes were conjugated with a photo-cleavable azido-biotin tag **(Fig. S4a)**[26] via click chemistry. A tandem orthogonal proteolysis (TOP)-ABPP[27] platform was then employed to capture and enrich the modified peptides. Following mass spectrometry (MS) analysis, an open-search strategy by MSfragger[28] was implemented to simultaneously determine the mass shifts and corresponding modification sites by the LCA-Galk probe. Excitingly, the open-

search analysis revealed a defined mass shift of 512.37 Da only on cysteine residues (**Fig. 1e-f**). Such a mass shift was fully recapitulated when a different acid-cleavable azido-biotin tag was used **(Fig. S4a-c)** [26], supporting the robust nature of this modification product by LCA-Galk on bacterial proteins. Furthermore, incorporation of an isotopically labeled TEV tag[29] resulted in a unique isotopic signature on the modified peptides **(Fig. S4a,d)**, confirming that the observed mass increment is genuinely derived from the probe modification. After deducting the part from the biotin tag, the exact mass for the modification mediated by LCA-Galk was 469.32 Da, which corresponds to the molecular weight of the probe (471.33 Da) minus two hydrogen atoms **(Fig. S5)**.

Using the close-search mode in MSFragger[28], we identified a total of 523 LCA-Galk modified cysteines across three biological replicates **(Fig. 1g, Table S1)**, many of which were supported by MS/MS spectra with unique reporter ions generated by the probe fragmentation **(Fig. 1h)**. To confirm that native LCA intrinsically possesses such a cysteine reactivity, we performed a competitive labeling experiment on the proteomes of LCA-pretreated bacteria using a cysteine-reactive probe, IA-alkyne **(Fig. S6a)**. In-gel fluorescence demonstrated that LCA treatment could efficiently outcompete the labeling signal of IA-alkyne **(Fig. 1i)**. To achieve the site-specific quantification, we performed the competitive rdTOP-ABPP profiling[30] with IA-alkyne[30,31], and quantified a total of 1782 cysteines in bacterial proteomes by stable isotope dimethyl labeling. We assigned 507 cysteines with a competition ratio over 1.5 from two biological replicates as the targets that are sensitive to LCA treatment **(Table S2)**. They constitute 28% of all the quantified cysteine sites **(Fig. S6b)** and 97 of them overlap with the list of cysteines captured directly by the LCA-Galk probe **(Fig. S6c).** In addition, when bacteria were co-treated with N-acetylcysteine (NAC) and LCA-Galk, the proteome-wide labeling signal progressively decreased with the increasing NAC concentrations **(Fig. S7)**. These results collectively supported that both LCA-Galk and the native LCA possess cysteine reactivity to form direct modifications in bacterial proteomes.

Structural validation of the LCA modification

We next aimed to investigate the molecular structure of such an LCA-mediated modification based on the detected mass. Our initial hypothesis was that the side chain of LCA-Galk undergoes a process resembling the fatty acid β-oxidation[32] in bacteria and yields a reactive intermediate with a conjugated double bond to modify cysteine residues via Michael addition **(Fig. 2a)**. To test this idea, we synthesized a probe derivative named "22,23-dehydro-LCA-Galk" based on the putative intermediate core **(Fig. 2b)** and evaluated its proteomic labeling by both in-gel fluorescence and TOP-ABPP analysis. The results showed that 22,23-dehydro-LCA-Galk did not directly induce any visible labeling in bacterial lysates **(Fig. S8)**. While its labeling in living bacteria was weaker than that of LCA-Galk, it was still significantly stronger as compared to the background control **(Fig. 2c)**, which yielded a mass shift of 467.30 Da **(Fig. 2d, Table S1)**. This corresponding reduction of 2 Da in mass shift excludes the possibility of LCA-Galk to modify cysteines via Michael addition through a side-chain oxidation intermediate.

To test if the modification occurs via a thioester formation at the carboxyl group or a thioether formation via the 3-hydroxyl elimination **(Fig. 2a)**, we synthesized the corresponding "LCA-alk" and "3-methoxy-LCA-Galk" probes to block the carboxyl and 3-hydroxyl groups in LCA, respectively **(Fig. 2b)**. Both probe derivatives exhibited comparable labeling efficiencies to that of LCA-Galk in living bacteria **(Fig. 2c)**, resulting in corresponding mass shifts (-58 Da and +114 Da) that are equal to the mass differences between LCA-Galk and the probe derivatives, respectively **(Fig. 2d, Table S1)**. These results demonstrate that the 3-hydroxyl and carboxyl groups in LCA are not directly involved in the modification. Furthermore, we moved the bioorthogonal alkyne handle from the carboxyl terminus to the 3-hydroxyl group to synthesize another derivative named "3alkLCA" **(Fig. 2b)**. The observation that 3alkLCA could still mediate the modification suggested the reaction does not rely on the propargylglycine moiety or the alkyne handle either **(Fig. 2c-d, Table S1)**.

We lastly focused on the steroid scaffold of LCA to pinpoint the reaction site. Previous studies have reported that LCA can be metabolically converted into 3-oxo-LCA and isoLCA[9,10], and it can also undergo hydroxylation by the cytochrome P450 family enzymes during reabsorption to form either the 6-hydroxylated species (e.g. HDCA or MDCA) or the 7-hydroxylated species (e.g. CDCA)[33,34] **(Fig. S9)**. To minimally perturb the steroid core, we therefore introduced the deuterium labeling at the C3, C6, and C7 positions of the LCA-Galk probe to investigate their involvement in the modification process **(Fig. 2b)**. Both the "C3-d-LCA-Galk" and "isoC3-d-LCA-Galk" probe derivatives exhibited the similar labeling efficiencies as LCA-Galk and resulted in a mass shift variation of +1 Da **(Fig. 2c-d, Table S1)**, which indicates that the deuterium at the C3 position was retained during the reaction and the C3 position is not directly involved in the modification. Similarly, the successful labeling by the "C7-d-LCA-Galk" probe excluded the C7 position from participating in the modification process **(Fig. 2c-d, Table S1)**.

In contrast, the proteomic labeling of the C6-deuterated derivative, "C6-d_2-LCA-Galk", was completely abolished **(Fig. 2c-d)**, suggesting that LCA is very likely to react with cysteine at the C6 position. One possible explanation is that the dual deuteration elevates the activation energy required for the bond cleavage, thereby abolishing the reaction, which is the underlying principle for deuterium labeling in medicinal chemistry to extend drugs' half-lives due to the kinetic isotope effect (KIE)[35]. To validate the structure of LCA-Cys adduct, we subjected the lysates of LCA-treated bacteria to exhaustive proteome hydrolysis[36] and compared the sample mixture of single amino acids after full hydrolysis to the synthetic standard of LCA-C6(α)-Cys by parallel reaction monitoring (PRM) **(Fig. 2e-f)**. Satisfyingly, the results showed that the sample of proteome hydrolysis contained a peak that precisely matched the synthetic LCA-C6(α)-Cys standard in terms of both retention time and fragmentation patterns **(Fig. 2g)**. Collectively, we confirmed the structure of this modification as an LCA-cysteine conjugation adduct through C6 of the steroid core.

LCA modification impairs the infection ability of *S.* Typhimurium

Besides *E. coli*, LCA-Galk could successfully label multiple pathogenic bacteria, including *S. aureus*, *S.* Typhimurium and *P. aeruginosa* **(Fig. S10a)**, and deuteration at the C6 position of the probe completely abolished the labeling signals **(Fig. S10b)**, which suggested that the LCA-mediated modification is conserved across these bacteria species. The observation also enabled the C6-deuterated LCA derivative, "C6-d_2-LCA" **(Fig. S10c)**, as a perfect negative control to differentiate the covalent and non-covalent interactions between LCA and proteins, which allowed us to perform phenotypic screening in pathogens to identify biological responses selectively triggered by the LCA modification **(Fig. 3a)**.

We initially investigated the possible role of bacterial growth inhibition by the LCA modification as previous studies have reported that LCA inhibited the growth of multiple gram-positive pathogens, including *C. difficile* and *S. aureus*[13,22,23]. Through antibacterial assays, we demonstrated that the LCA-Galk probe retained an antibacterial efficacy comparable to that of the native LCA **(Fig. S11a-c)**. Unfortunately, C6-d_2-LCA-Galk also showed such an activity, suggesting that LCA did not inhibit the bacterial growth through the cysteine modification **(Fig. S11c)**. Given that infecting host cells is another major aspect of the bacterial virulence, we next examined the impact of LCA treatment on the intracellular pathogen *S.* Typhimurium in terms of its infectivity. Immunofluorescence imaging demonstrated that LCA, but not C6-d_2-LCA, could effectively inhibit the bacterial infection at the early stage and impair the colocalization between *S.* Typhimurium and F-actin in host cells **(Fig. 3b)**. The results could be also confirmed by the quantification of intracellular colony forming unit (CFU) **(Fig. 3c)**.

To globally profile the targets of LCA modification in *S.* Typhimurium, we employed the LCA-Galk probe for metabolic labeling and identified a total of 378 LCA-modified proteins by TOP-ABPP from three biological replicates **(Fig. 3d, Table S1)**. Screening these targets for virulence effectors led us to focus on SipA, a key effector tightly linked

to the early-stage bacterial infectivity **(Fig. 3e)**. As an essential substrate of the Type III secretion system (T3SS), SipA is translocated into host cells during the initial phases of infection, where it interacts with and stabilizes the host cytoskeleton to promote bacterial internalization[37,38]. We constructed a *sipA*-knockout strain of *S.* Typhimurium (*ΔsipA*) and observed that the depletion of SipA significantly disrupted the bacterial infectivity, especially at the early stage of infection **(Fig. S12a)**. To our delight, LCA treatment of the wild-type bacteria led to an infection-suppressive phenotype similar to that observed for the *ΔsipA* strain **(Fig. S12b)**.

Individual mutagenesis of the three cysteine residues in SipA pinpointed that C590 is the major modification site by LCA-Galk **(Fig. 3f),** which is also supported by the MS/MS spectrum **(Fig. S13a)**. Given that C590 is located in the actin-binding core domain of SipA **(Fig. S13b)**, we next investigated the impact of the LCA modification on the actin-binding capacity of SipA. We expressed and purified the recombinant C-terminal domain of the wild-type SipA, and subsequently introduced the LCA-mimicking modification at the C590 residue through reaction with a C6-chloroacetyl-LCA derivative named "6-AcCl-LCA" **(Fig. S14)**. Through co-immunoprecipitation with F-actin in HeLa cell lysates, we confirmed that such an LCA-mimicking modification indeed impaired the interaction between SipA and actin **(Fig. 3g)**. Additionally, immunoblotting and quantitative proteomic analysis of the *S.* Typhimurium's secretome demonstrated that LCA treatment significantly suppressed the secretion of SipA while the treatment of C6-d_2-LCA had much less impact **(Fig. 3h-i)**. These results collectively suggested that the LCA modification could impair the function of SipA by simultaneously suppressing its secretion and disrupting its interaction with F-actin in host cells.

LCA modification globally disrupts protein localization of *S.* Typhimurium by targeting the translocase complex

During the analysis of *S.* Typhimurium's secretome, we noticed that LCA treatment

led to the downregulation of multiple secreted proteins compared to the deuterated derivative **(Fig. S15a)**. To systematically investigate the impact of LCA modification on protein subcellular localization, we extracted the secretome and membrane proteome of *S.* Typhimurium and subjected them to data-independent acquisition (DIA) mass spectrometry for proteomic quantification **(Fig. S15b, Table S3)**. We set a ratio cutoff of <0.67 or >1.5 between the LCA and C6-d_2-LCA treated groups to filter for proteins undergoing significant localization changes mediated by the LCA modification **(Fig. 4a)**. In total, we quantified 1184 proteins in the secretome sample, which constituted 34% of the proteins that were significantly downregulated after the LCA treatment **(Fig. 4a, Table S3)**. This subset included flagellar-associated proteins (FliD, FlgE, FlgK, and FlgB), Type III secretion system (T3SS) assembly machinery components (SipC, SipD, InvJ, and PrgI), and key virulence effectors (SipA, SopA, SopB, SopD2, and PipB2) **(Fig. 4b)**. Conversely, 23% of the membrane proteome (3256 proteins) exhibited a significant increase in abundance **(Fig. 4a, Table S3)**. Notably, proteins that are critically involved in T3SS assembly and translocation, including SipB, SipC, and SipD, were significantly enriched in the membrane fraction **(Fig. 4a-b)**. Together, these findings indicate that the LCA modification triggered a marked accumulation of essential transport components in the bacterial membrane, consequently leading to a substantial impairment of their secretion.

To further elucidate the molecular mechanism by which LCA inhibits protein translocation in *S.* Typhimurium, we cross-referenced the protein targets of LCA modification with those undergoing subcellular localization shifts. Integration of these two datasets prioritized SecA, a key component of the bacterial protein translocase complex, as the potential functional target **(Fig. S16a)**. Previous studies have reported that SecA mediates the transmembrane translocation and assembly of the T3SS inner and outer ring structural elements (InvG, PrgK, and PrgH) [39-41]. Within this translocase complex, LCA modification exclusively upregulated the membrane localization of SecA without altering its total level of protein expression **(Fig. 4c)**. MS/MS analysis

and mutagenesis of individual cysteine residues revealed that LCA primarily modifies SecA within its C terminal zinc-binding domain **(Fig. 4d, Fig. S16b)**, where C885, C887, C896, and H897 coordinate a zinc ion to facilitate substrate delivery **(Fig. 4e)**[42,43].

We therefore hypothesized that LCA modification could impair SecA's function, thereby compromising the secretion of virulence effectors. Given that *secA* is an essential housekeeping gene and cannot be knocked out, we performed the overexpression experiments to evaluate its functional role. Indeed, SecA's overexpression partially but significantly restored the secretion of SipA, SipB, SipC and SipD that were suppressed by the LCA treatment **(Fig. 4f)**. Take together, these results supported a model that LCA modification impairs the secretion of T3SS assembly components and virulence effectors including SipA by inhibiting the translocase SecA. Although a minor fraction of SipA can still be secreted and translocated into host cells, the LCA modification on its Cys590 imposed a second line of defense by directly disrupting SipA's interaction with F-actin **(Fig. 4g)**.

Discussion

In this study, we have discovered a previously unrecognized cysteine PTM mediated by LCA, a secondary BA produced by gut microbiota. We showed that LCA modified cysteine residues via the C6 position of its steroid core during bacterial metabolism. Functionally, LCA modification reshaped the global protein localization in *S.* Typhimurium by targeting the protein translocase SecA. In addition, LCA could also covalently modify the virulence effector SipA to impair its F-actin binding ability. Together, these functional events led to suppression of *S.* Typhimurium's infection at the early stage.

During infection, *S.* Typhimurium relies on the Salmonella pathogenicity island 1 (SPI-1)-encoded T3SS to deliver virulence effectors into host cells and promote bacterial invasion. By disrupting protein trafficking and compromising the function of a key T3SS effector, LCA modification constrains this critical stage of infection. Our findings uncover a previously unrecognized mode of bile acid-mediated regulation of bacterial virulence beyond their detergent-like effects as currently established and further reveal the post-translational modification as an important regulatory layer in bacterial pathogenesis.

Current antibacterial therapies for *Salmonella* infections largely rely on conventional antibiotics that inhibit DNA replication[44], protein synthesis[45] or cell-wall biosynthesis[46], but their effectiveness are increasingly threatened by antimicrobial resistance. More recently, the T3SS has emerged as an attractive anti-virulence target, with inhibitors being developed to interfere with the effector secretion[47,48]. Our finding that LCA covalently modifies Cys590 of SipA and inhibits its interaction with F-actin established this functionally important cysteine residue as a potential anchor point for the development of covalent anti-virulence agents. Moreover, owing to its broad cysteine-labelling profile in bacterial proteomes, the LCA-Galk probe might serve as an enabling

tool for competitive screening of covalent antibacterial compounds by chemoproteomic technologies.

Although we here reported for the first time a novel PTM mediated by bile acid, it should be explicitly acknowledged that the detailed chemistry mechanism underlying the LCA modification remains unresolved in the current study. While many metabolites have been shown to react with cysteine residues in proteins through their electrophilic functional groups, the C6 carbon of LCA resides within its steroid scaffold is not intrinsically reactive at all. The fact that the addition of LCA directly to native cell lysates failed to generate such a modification suggested that LCA must be metabolically activated under a stringent condition requiring an intact cellular environment. One possibility is that LCA is metabolically converted by a series of enzymes in living bacteria into a transient and highly reactive intermediate, potentially involving radical species, whereas the disruption of cellular metabolism and subtle local microenvironment upon membrane lysis completely abolishes this reaction from occurring. In addition, these metabolic enzymes must work on LCA with high specificity as none of these other three common BAs showed cysteine reactivity. Future studies should be directed to identify the molecular machinery responsible for installing this modification and define its underlying chemical mechanism by an integrated approach of chemical synthesis, genetic perturbation, structural biology and multi-dimensional omics analyses.

As an expanding repertoire of microbiota-derived secondary bile acids has been identified in recent years, the possibility exists that such a cysteine modification may extend beyond LCA. Notably, isoLCA is an abundant isomer of LCA and was also capable of modifying cysteine residues with even greater efficiency according to our data **(Fig. 2e)**. In addition, many of the LCA-Galk derivative probes could still covalently label cysteine despite that their carboxyl terminus and 3-hydroxyl group blocked. These observations suggest that cysteine reactivity may represent a broader chemical property of structurally related secondary bile acids rather than an isolated

feature of LCA. Considering that LCA can also be converted through bile acid conjugation pathways into a variety of derivatives such as glycolithocholic acid (GLCA)[49], whether these endogenous metabolites can retain a similar cysteine-modifying potential is worth investigating in future.

More broadly, as a key metabolite positioned at the interface between host and microbiota, LCA also exerts regulatory effects on epithelial cell metabolism[33] and T cell differentiation[9]. Beyond the bacterial proteome, our preliminary test indicated that such a LCA modification is also detectable in host tissues (data not shown). As no cysteine modification by primary BAs (e.g. CA and CDCA) was identified, these observations raise the possibility that microbial conversion of primary BAs into secondary BAs creates a previously unrecognized route through which the gut microbiota modulates the host signaling. The LCA-mediated cysteine modification discovered in this study may therefore constitute a molecular link in host–microbiome communication, with potentially important roles in both physiology and pathology.

References

1 Russell, D. W. The Enzymes, Regulation, and Genetics of Bile Acid Synthesis. *Annu. Rev. Biochem.* **72**, 137-174 (2003).

2 Thomas, C. *et al.* Targeting Bile-Acid Signalling for Metabolic Diseases. *Nat. Rev. Drug Discov.* **7**, 678-693 (2008).

3 Parks, D. J. *et al.* Bile Acids: Natural Ligands for an Orphan Nuclear Receptor. *Science* **284**, 1365-1368 (1999).

4 Kawamata, Y. *et al.* A G Protein-Coupled Receptor Responsive to Bile Acids*. *J. Biol. Chem.* **278**, 9435-9440 (2003).

5 Funabashi, M. *et al.* A Metabolic Pathway for Bile Acid Dehydroxylation by the Gut Microbiome. *Nature* **582**, 566-570 (2020).

6 Payne, C. M. *et al.* Hydrophobic Bile Acids, Genomic Instability, Darwinian Selection, and Colon Carcinogenesis. *Clin. Exp. Gastroenterol.* **1**, 19-47 (2008).

7 Qu, Q. *et al.* Lithocholic Acid Phenocopies Anti-Ageing Effects of Calorie Restriction. *Nature* **643**, 192-200 (2025).

8 Qu, Q. *et al.* Lithocholic Acid Binds Tulp3 to Activate Sirtuins and Ampk to Slow Down Ageing. *Nature* **643**, 201-209 (2025).

9 Hang, S. *et al.* Bile Acid Metabolites Control T(H)17 and T(Reg) Cell Differentiation. *Nature* **576**, 143-148 (2019).

10 Paik, D. *et al.* Human Gut Bacteria Produce Tη17-Modulating Bile acid Metabolites. *Nature* **603**, 907-912 (2022).

11 Jin, W.-B. *et al.* Microbiota-Derived Bile Acids Antagonize the Host Androgen Receptor and Drive Anti-Tumor Immunity. *Cell* **188**, 2336-2353.e2338 (2025).

12 Islam, K. B. *et al.* Bile Acid Is a Host Factor That Regulates the Composition of the Cecal Microbiota in Rats. *Gastroenterology* **141**, 1773-1781 (2011).

13 Thanissery, R. *et al.* Inhibition of Spore Germination, Growth, and Toxin Activity of Clinically Relevant C. difficile Strains by Gut Microbiota Derived Secondary Bile Acids. *Anaerobe* **45**, 86-100 (2017).

14 Gentry, E. C. *et al.* Reverse Metabolomics for the Discovery of Chemical Structures from Humans. *Nature* **626**, 419-426 (2024).

15 Guzior, D. V. *et al.* Bile Salt Hydrolase Acyltransferase Activity Expands Bile Acid Diversity. *Nature* **626**, 852-858 (2024).

16 Mohanty, I. *et al.* The Underappreciated Diversity of Bile Acid Modifications. *Cell* **187**, 1801-1818.e1820 (2024).

17 Yang, J. *et al.* Structure-Guided Discovery of Bile Acid Derivatives for Treating Liver Diseases without Causing Itch. *Cell* **187**, 7164-7182.e7118 (2024).

18 Lin, J. *et al.* A Microbial Amino-Acid-Conjugated Bile Acid, Tryptophan-Cholic Acid, Improves Glucose Homeostasis Via the Orphan Receptor Mrgpre. *Cell* (2025).

19 Zhuang, S. *et al.* Chemoproteomic Profiling of Bile Acid Interacting Proteins. *ACS Cent. Sci.* **3**, 501-509 (2017).

20 Liu, B. *et al.* Chemoproteomic Profiling Reveals the Mechanism of Bile Acid Tolerance in Bacteria. *ACS Chem. Biol.* **17**, 2461-2470 (2022).
21 Yang, X. *et al.* Anti-Infective Bile Acids Bind and Inactivate a Salmonella Virulence Regulator. *Nat. Chem. Biol.* **19**, 91-100 (2023).
22 Yang, X. *et al.* Chemical Proteomic Analysis of Bile Acid-Protein Targets in Enterococcus Faecium. *RSC Chem Biol* **3**, 1397-1402 (2022).
23 Forster, E. R. *et al.* Identification of a Bile Acid-Binding Transcription Factor in Clostridioides Difficile Using Chemical Proteomics. *ACS Chem. Biol.* **17**, 3086-3099 (2022).
24 Xiao, X. *et al.* Cholesterol Modification of Smoothened Is Required for Hedgehog Signaling. *Mol. Cell* **66**, 154-162.e110 (2017).
25 Luo, J. *et al.* Mechanisms and Regulation Of cholesterol Homeostasis. *Nat. Rev. Mol. Cell Biol.* **21**, 225-245 (2020).
26 Szychowski, J. *et al.* Cleavable Biotin Probes for Labeling of Biomolecules Via Azide−Alkyne Cycloaddition. *J. Am. Chem. Soc.* **132**, 18351-18360 (2010).
27 Weerapana, E. *et al.* Tandem Orthogonal Proteolysis-Activity-Based Protein Profiling (Top-Abpp)—a General Method for Mapping Sites of Probe Modification in Proteomes. *Nat. Protoc.* **2**, 1414-1425 (2007).
28 Kong, A. T. *et al.* Msfragger: Ultrafast and Comprehensive Peptide Identification in Mass Spectrometry–Based Proteomics. *Nat. Methods* **14**, 513-520 (2017).
29 Jia, G. *et al.* Discovery of Cysteine Carboxyalkylations by Real-Time Isotopic Signature Targeted Profiling. *J. Am. Chem. Soc.* **147**, 7513-7523 (2025).
30 Yang, F. *et al.* A Dimethyl-Labeling-Based Strategy for Site-Specifically Quantitative Chemical Proteomics. *Anal. Chem.* **90**, 9576-9582 (2018).
31 Wang, C. *et al.* A Chemoproteomic Platform to Quantitatively Map Targets of Lipid-Derived Electrophiles. *Nat. Methods* **11**, 79-85 (2014).
32 Pavoncello, V. *et al.* Degradation of Exogenous Fatty Acids in Escherichia Coli. *Biomolecules* **12** (2022).
33 Sheng, W. *et al.* The Effect of Lithocholic Acid on the Gut-Liver Axis. *Front. Pharmacol.* **13**, 910493 (2022).
34 Makishima, M. *et al.* Vitamin D Receptor as an Intestinal Bile Acid Sensor. *Science* **296**, 1313-1316 (2002).
35 Di Martino, R. M. C. *et al.* Deuterium in Drug Discovery: Progress, Opportunities and Challenges. *Nat. Rev. Drug Discov.* **22**, 562-584 (2023).
36 Baldensperger, T. *et al.* Quantitation of Reactive Acyl-Coa Species Mediated Protein Acylation by Hplc-Ms/Ms. *Anal. Chem.* **91**, 12336-12343 (2019).
37 Zhou, D. *et al.* Role of the S. Typhimurium Actin-Binding Protein Sipa in Bacterial Internalization. *Science* **283**, 2092-2095 (1999).
38 Lilic, M. *et al.* Salmonella Sipa Polymerizes Actin by Stapling Filaments with Nonglobular Protein Arms. *Science* **301**, 1918-1921 (2003).
39 Sukhan, A. *et al.* Genetic Analysis of Assembly of Thesalmonella Enterica

Serovar Typhimurium Type Iii Secretion-Associated Needle Complex. *J. Bacteriol.* **183**, 1159-1167 (2001).

40 Portaliou, A. G. *et al.* Type Iii Secretion: Building and Operating a Remarkable Nanomachine. *Trends Biochem. Sci* **41**, 175-189 (2016).

41 Dong, L. *et al.* Structural Basis of Seca-Mediated Protein Translocation. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2208070120 (2023).

42 Fekkes, P. *et al.* Zinc Stabilizes the Secb Binding Site of Seca. *Biochemistry* **38**, 5111-5116 (1999).

43 Kusters, I. & Driessen, A. J. Seca, a Remarkable Nanomachine. *Cell. Mol. Life Sci.* **68**, 2053-2066 (2011).

44 Drlica, K. & Zhao, X. DNA Gyrase, Topoisomerase Iv, and the 4-Quinolones. *Microbiol. Mol. Biol. Rev.* **61**, 377-392 (1997).

45 Wilson, D. N. Ribosome-Targeting Antibiotics and Mechanisms of Bacterial Resistance. *Nature Reviews Microbiology* **12**, 35-48 (2014).

46 Waxman, D. J. & Strominger, J. L. Penicillin-Binding Proteins and the Mechanism of Action of Beta-Lactam Antibiotics. *Annu. Rev. Biochem.* **52**, 825-869 (1983).

47 Boudrioua, A. *et al.* Discovery of Synthetic Small Molecules Targeting the Central Regulator of Salmonella Pathogenicity. *Sci. Adv.* **11**, eadr5235

48 Liu, J. *et al.* Structure-Based Identification of Salvianolic Acid B as an Inhibitor Targeting Salmonella Invc Atpase. *J. Biol. Chem.* **301**, 110722 (2025).

49 Chiang, J. Y. L. Bile Acid Metabolism and Signaling. *Compr. Physiol.* **3**, 1191-1212 (2013).

50 Hughes, C. S. *et al.* Single-Pot, Solid-Phase-Enhanced Sample Preparation for Proteomics Experiments. *Nat. Protoc.* **14**, 68-85 (2019).

51 Xu, T. *et al.* Prolucid: An Improved Sequest-Like Algorithm with Enhanced Sensitivity and Specificity. *J. Proteomics* **129**, 16-24 (2015).

52 Tabb, D. L. *et al.* Dtaselect and Contrast: Tools for Assembling and Comparing Protein Identifications from Shotgun Proteomics. *J Proteome Res* **1**, 21-26 (2002).

53 Gao, J. *et al.* Cimage2.0: An Expanded Tool for Quantitative Analysis of Activity-Based Protein Profiling (Abpp) Data. *J Proteome Res* **20**, 4893-4900 (2021).

54 Demichev, V. *et al.* Dia-Nn: Neural Networks and Interference Correction Enable Deep Proteome Coverage in High Throughput. *Nat. Methods* **17**, 41-44 (2020).

55 Datsenko, K. A. & Wanner, B. L. One-Step Inactivation of Chromosomal Genes in Escherichia Coli K-12 Using Pcr Products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640-6645 (2000).

56 D'Agostino, G. D. *et al.* Host-Microbiome Orchestration of the Sulfated Metabolome. *Nat. Chem. Biol.* **20**, 410-421 (2024).

Acknowledgments

We thank the Computing Platform of the Center for Life Science for supporting the proteomics data analysis. We thank the Public Technology Platform of Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies for supporting the LC-MS and NMR analysis for probe synthesis and compound characterization. We thank Dr. Weidi Xiao, Dr. Xiaping Fu, Prof. Changtao Jiang, Prof. Feng Shao and Prof. Yue Xu for helpful discussions.

Funding

C.W. and B.Y. acknowledges the funding support from Ministry of Science and Technology of China (2022YFA1304700). C.W also acknowledges the support from National Natural Science Foundation of China (22525071, 22361142837, 92578206), Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM507), and Beijing National Laboratory for Molecular Sciences (Grant No. BNLMS-CXTD-202401). B.Y. acknowledges the support from National Natural Science Foundation of China (22531009) and P.X. acknowledges the support from CAS Project for Young Scientists in Basic Research (YSBR-095).

Author contributions

Y.W., B.Y. and C.W. conceived the study. Y.W. performed most of the experiments. C.Z., conducted most of the organic synthesis. Y.W., Y.C., S-B.Y. and Y-F.Z. contributed to the organic synthesis. Y-C.L., Y.Z., P.X. contributed to methodological development. Y.L. contributed to data analysis. Z.T. and S-B.Y. contributed to the plasmid construction and protein purification. S-L.Y. provided the

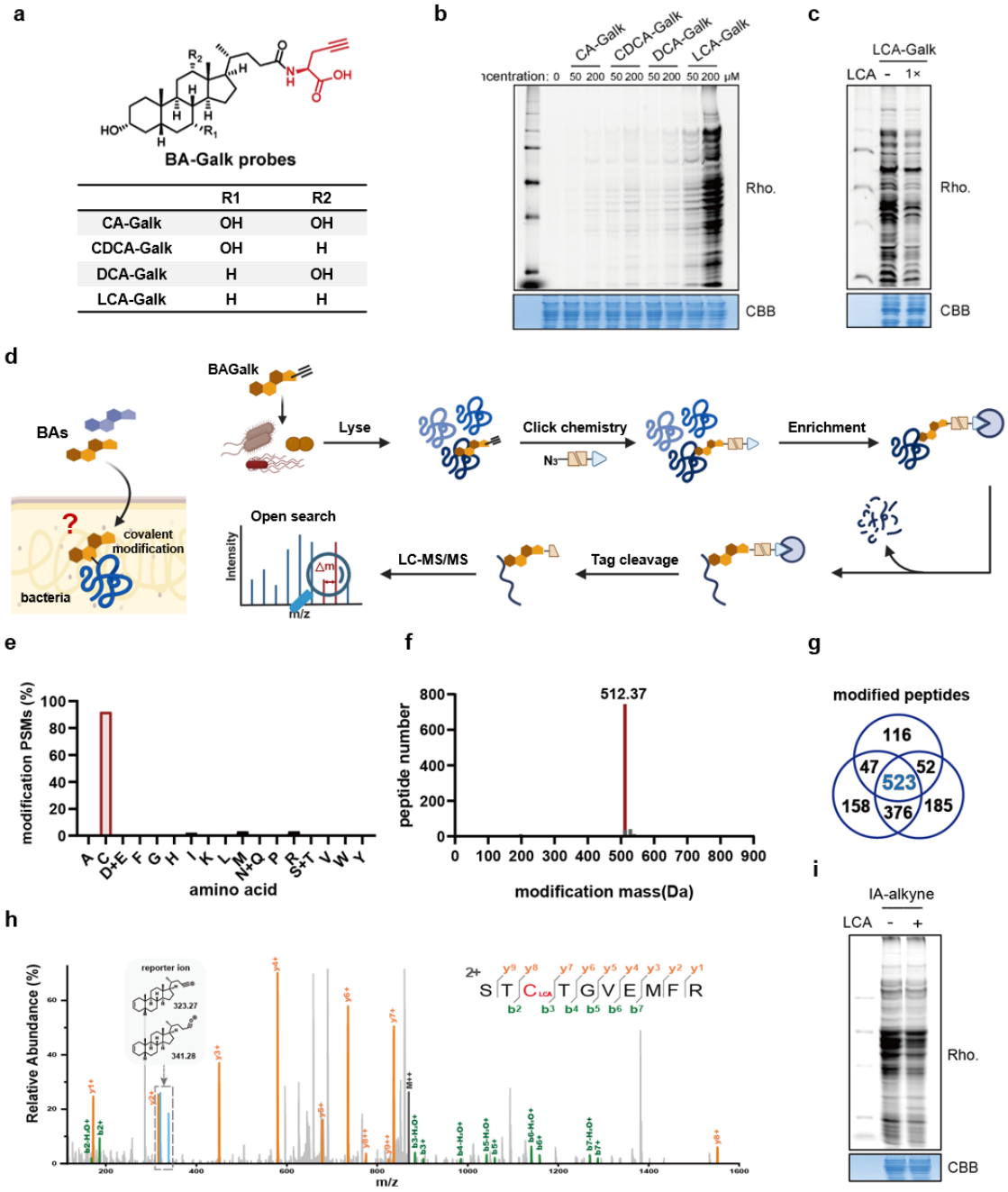
bacterial strains and contributed to the bacterial assays. Y.W., C.Z, B.Y. and C.W. wrote the paper with input from all authors. The experimental datasets used in this study are available as supplementary materials.

Competing interests

Authors declare that they have no competing interests

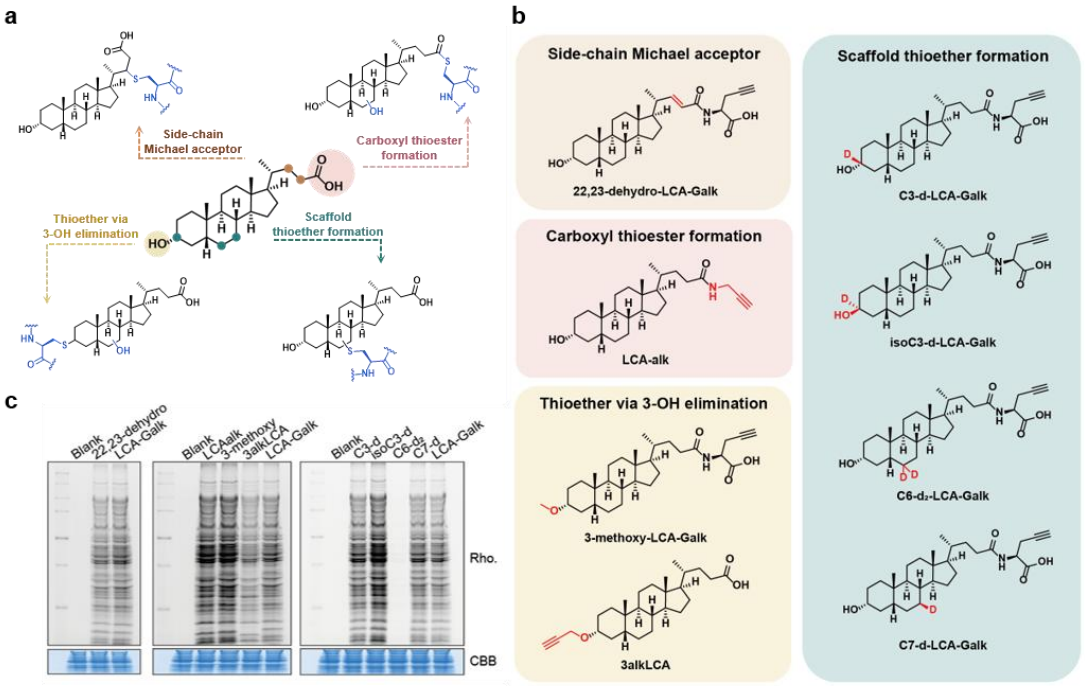
Materials availability

The raw MS data have been deposited to the ProteomeXchange Consortium (https://proteomecentral.proteomexchange.org) *via* the iProX partner repository with the data set identifier IPX0019399000. All other data are available in the main text or the supplementary materials.



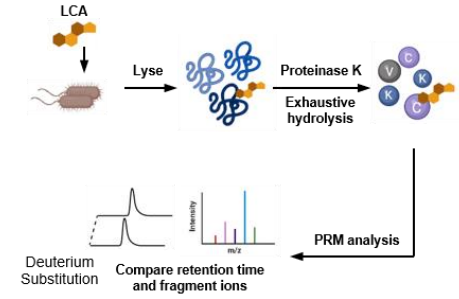
	R1	R2
CA-Galk	OH	OH
CDCA-Galk	OH	H
DCA-Galk	H	OH
LCA-Galk	H	H

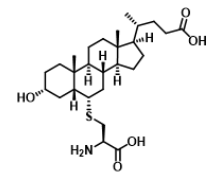
Fig. 1. Discovery of an LCA-mediated covalent modification in bacterial proteomes. a, Chemical structures of the bioorthogonal "BA-Galk" probes based on four common bile acids. CA: Cholic acid, CDCA: Chenodeoxycholic acid, DCA: Deoxycholic acid. LCA: Lithocholic acid. **b,** Comparison of the labeling efficiency of four BA-Galk probes at 50 or 200 μM in living bacteria by in-gel fluorescence. **c,** Native LCA could compete LCA-Galk labeling in bacterial proteomes. **d,** Workflow for discovering an LCA-mediated post-translational modification using LCA-Galk by TOP-ABPP. **e-f,** Open search with MSFragger showed the modification predominantly occur on cysteine **(e)** with a mass shift of 512.37 Da **(f)**. Structure of the photo cleavable tag is shown in Supplementary Figure S4a. **g,** Venn diagram showing the number of modified peptides by LCA-Galk as identified from *E. coli* proteomes in three biological replicates. **h,** A representative MS/MS spectrum of the LCA-Galk-modified peptide from EF-Tu1. The peptide sequence is shown with the modified cysteine marked in red. The b and y ions are marked with those fragmentations from the LCA-Galk probe highlighted in blue. **i,** Native LCA could compete the cysteine labeling by IA-alkyne in bacterial proteomes.

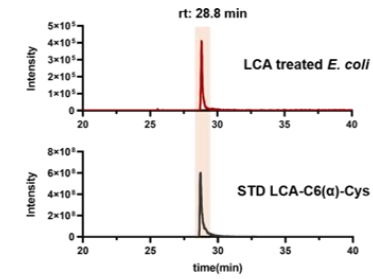


d

Analogue	Mass over LCA-Galk (Da)	Modification mass (Da)	
LCA-Galk	0	469.32	
22,23-dehydro-LCA-Galk	-2	467.30	Side-chain Michael acceptor
LCA-alk	-58	409.3	Thioester
3-methoxy-LCA-Galk	+14	483.33	Thioether via 3-OH elimination
3alk-LCA	-57	412.30	
C3-d-LCA-Galk	+1	470.33	Scaffold thioether formation
isoC3-d-LCA-Galk	+1	470.33	
C7-d-LCA-Galk	+1	470.33	
C6-d_2-LCA-Galk	**+2**	**N.D.**	







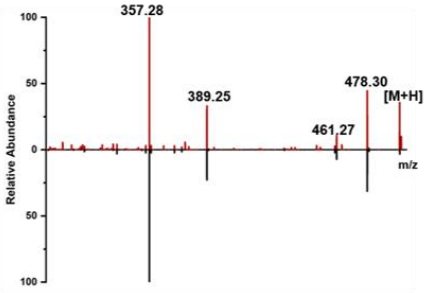


Fig. 2. Structural characterization of the LCA-cysteine modification via probe derivatization and exhaustive proteome hydrolysis. a. Proposed chemical structures of the LCA modification on cysteine through Michael addition, thioester from the carboxyl terminus, thioether from the 3-hydroxyl group or the steroid core. **b,** Structures of the LCA-Galk probe analogues, including a putative Michael addition intermediate, blockade of the carboxyl terminus or the 3-hydroxyl group, deuterium substitution at the C3, C6 or C7 positions on the steroid scaffold. **c,** Comparison of the labeling efficiency of the LCA-Galk probe analogues by in-gel fluorescence. **d,** List of mass differences between LCA-Galk and its probe analogues as well as the induced mass shifts in bacterial proteomes as analyzed by TOP-ABPP. **e,** Workflow of the exhaustive proteome hydrolysis. LCA-treated bacterial cells are lysed and subjected to the proteinase K digestion to achieve complete hydrolysis down to the single-amino-acid level, and the sample was compared with the synthetic standard in terms of chromatographic retention times and MS/MS fragmentation spectra to validate the structure of the modification adduct. **f,** Structure of the synthetic LCA-C6(α)-Cys standard. **g,** Match of the retention times (left) and fragment ions (right) between the sample of fully hydrolyzed single amino acids from LCA-treated bacteria (red) to that of the synthetic standard (black).

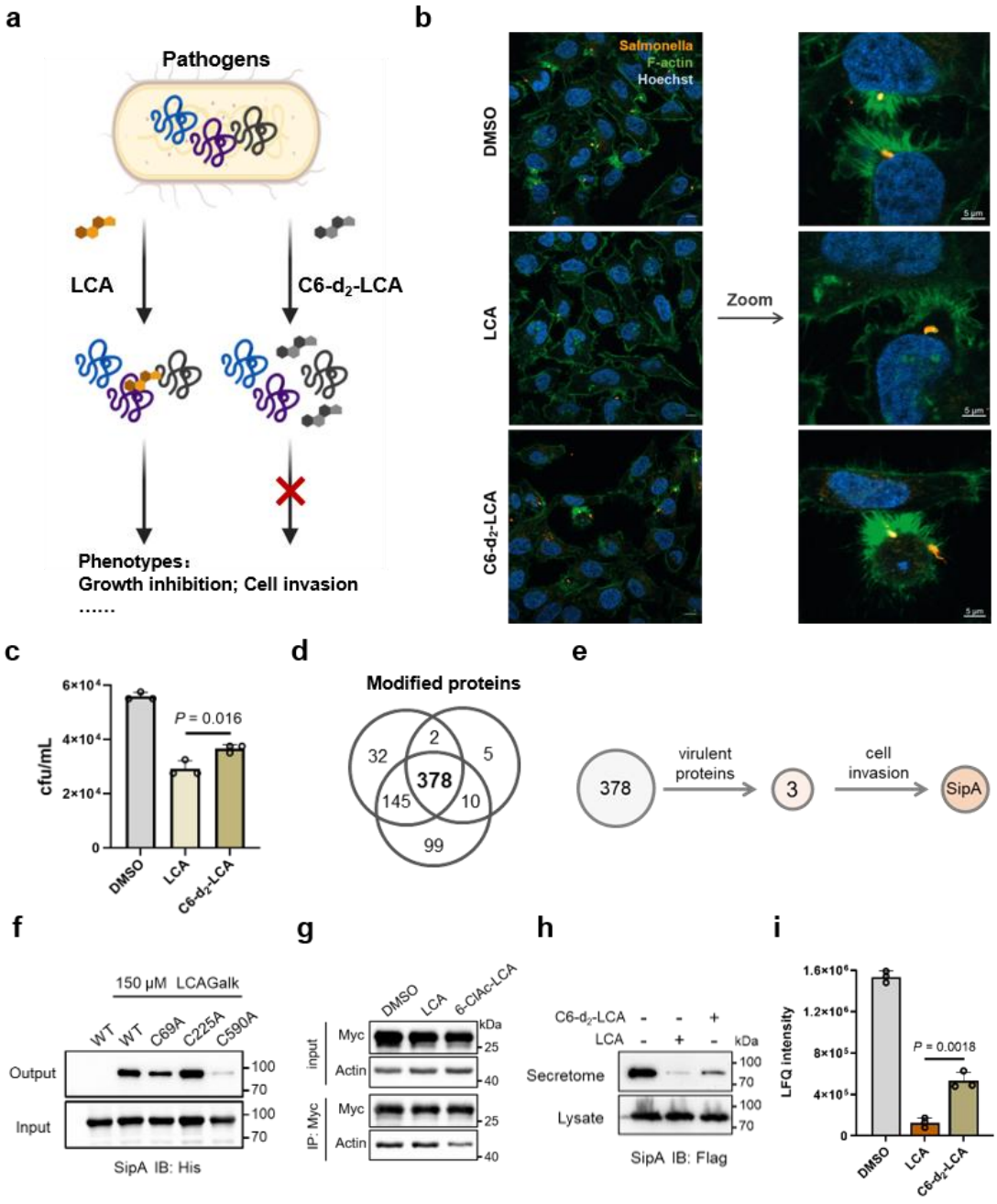


Fig. 3. LCA modification impairs the infection ability of *S.* Typhimurium. a, Workflow for screening phenotypes mediated by the LCA modification using LCA and C6-d_2-LCA. **b,** Immunofluorescent analysis of *S.* Typhimurium's infection of HeLa cells after treatment with LCA or C6-d_2-LCA. *S.* Typhimurium was visualized with an antibody against the *Salmonella* surface antigens (orange). Host cell nuclei were stained with Hoechst (blue). F-actin was stained with Phalloidin (green). Scale bar, 5 μm. **c,** Gentamicin protection assay revealed the anti-infective activity of LCA modification by quantification of CFU. *S.* Typhimurium was pretreated with LCA or C6-d_2-LCA, with DMSO treatment serving as the control, and subsequently incubated with HeLa cells at an MOI of 50. **d,** Venn diagram showing the number of LCA-Galk modified proteins identified from *S.* Typhimurium by TOP-ABPP in three biological replicates. **e,** Processing and filtering of the TOP-ABPP data to identify SipA as a functional target of LCA modification. **f,** Cysteine-to-alanine mutation pinpointed Cys590 as the principal site of LCA modification in SipA. **g,** Chemical introduction of an LCA-mimicking modification at C590 of SipA disrupted its actin-binding affinity. **h–i,** Western blot analysis **(h)** and label-free quantification **(i)** showed that LCA modification suppressed SipA secretion without affecting its expression. Data are presented as mean ± SD (n = 3 biological replicates). P value was determined via two-tailed Student's t-test.

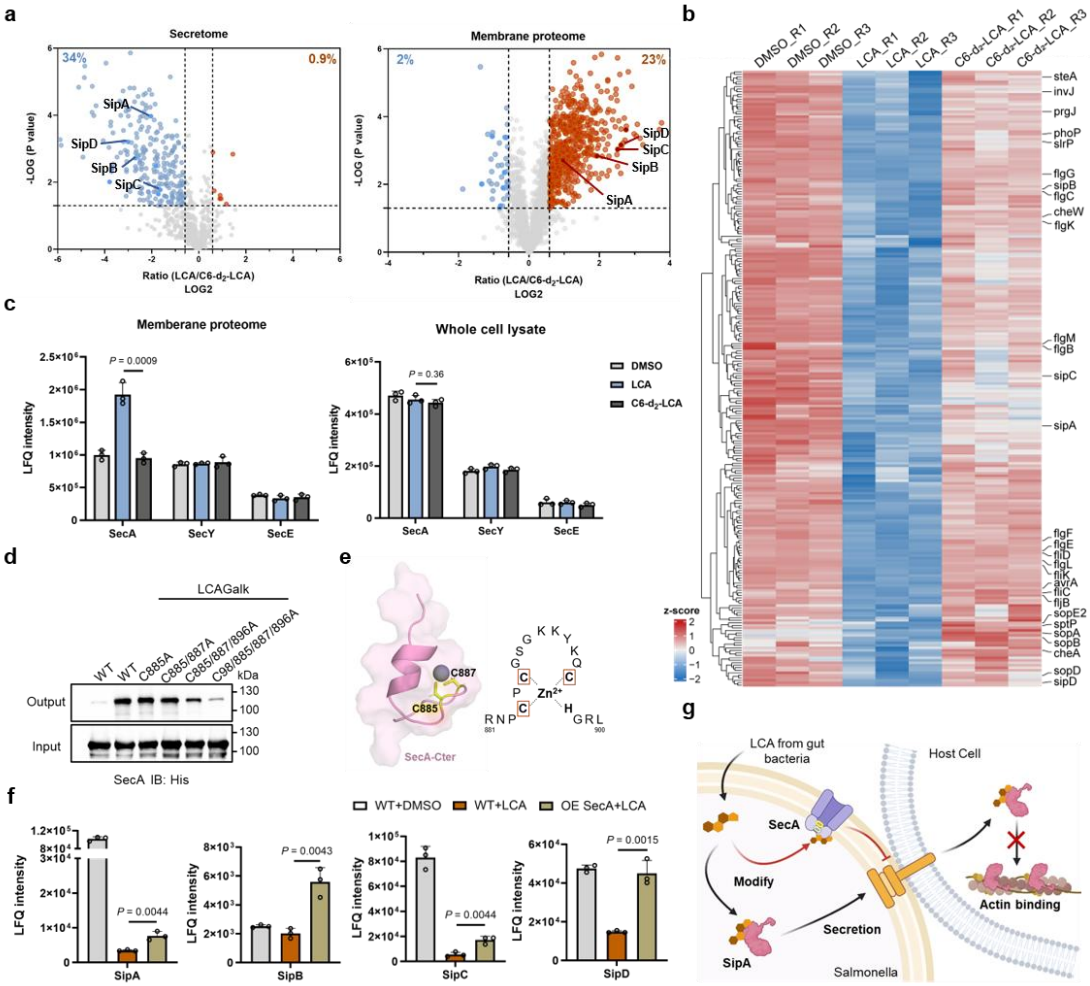


Fig. 4. LCA modification remodels the landscape of protein localization in *S.* Typhimurium by targeting the translocase complex. a, DIA-based quantitative proteomics revealed that LCA treatment of *S.* Typhimurium led to decreased and increased protein abundance in the secreted and membrane proteomes, respectively. In volcano plots, proteins significantly ($P \leq 0.05$) upregulated (ratio ≥ 1.5) or downregulated (ratio ≤ 0.67) upon LCA treatment are colored in red and blue, respectively. Representative T3SS-related proteins, including SipA, SipB, SipC and SipD, are marked. **b,** Heatmap of secreted proteins down-regulated by LCA modification. Proteins involved in virulence and motility are marked. **c,** LCA modification accumulates SecA in the membrane proteomes, but not other components of the Sec complex. **d,** Cysteine-to-alanine mutagenesis identified three zinc-binding cysteines in the C-terminal domain of SecA as the sites of LCA modification. **e,** The structure of canonical "C-X-C" zinc-binding domain at the SecA C-terminus (PDB: 1TM6). **f,** SecA's overexpression partially rescued the virulence protein secretion impaired by LCA modification. **g,** Proposed model of LCA modification-mediated inhibition of *S.* Typhimurium infection. LCA modification impairs SipA's secretion by targeting SecA, whereas direct modification of SipA further compromises its function by disrupting the interaction with F-actin. Data are presented as mean ± SD ($n = 3$ biological replicates). *P* value was determined via two-tailed Student's *t*-test.