Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Proteomic and Machine Learning Analyses Reveal the Mechanism of Baicalin in Inhibiting Biofilm Formation of Escherichia coli

Lei Wei
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Kang Zhang
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Jingyan Zhang
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Zhiting Guo
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Xiaorong Lu
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Guowei Xu
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Ziyi Wang
Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Lei Wang
wanglei03@caas.cn

Lanzhou Institute of Husbandry and Pharmaceutical Sciences
Jianxi Li
Lanzhou Institute of Husbandry and Pharmaceutical Sciences

Research Article

Keywords: Baicalin, Escherichia coli, biofilm, proteomics, flagella, machine learning

Posted Date: August 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10663623/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Microbiology on September 28th, 2026. See the published version at https://doi.org/10.1186/s12866-026-05706-0.

Abstract

Biofilm formation is a key survival strategy by which *Escherichia coli* (*E. coli*) resists environmental stress and sustains persistent infections. Baicalin, a natural flavonoid compound, has been reported to possess antibacterial and antivirulence activities; however, its antibiofilm mechanism remains incompletely understood. This study integrated phenotypic assays, DIA-based quantitative proteomics, bioinformatics, machine learning and scanning electron microscopy to investigate the effect of baicalin on *E. coli* biofilm formation. Baicalin markedly inhibited biofilm formation in a concentration-dependent manner. Importantly, viable cell counts under biofilm-inducing culture conditions showed no significant reduction, indicating that the inhibitory effect of baicalin on biofilm formation was not caused by changes in the total number of viable bacteria in the wells. Proteomic analysis indicated that baicalin treatment was associated with changes in energy metabolism, chemotaxis and flagellar assembly. Consistently, baicalin increased intracellular ROS levels, reduced intracellular ATP levels, whereas extracellular ATP remained unchanged, suggesting metabolic and redox remodeling rather than nonspecific membrane leakage. Baicalin also impaired swimming and swarming motility and reduced flagellar abundance. Machine learning further identified the flagellin-associated protein A0A140NAX2 as a candidate feature linked to the antibiofilm phenotype. Collectively, these findings suggest that baicalin suppresses *E. coli* biofilm formation mainly by altering metabolic state and impairing motility- and flagella-associated early adhesion processes, rather than by directly reducing total bacterial viability under biofilm-inducing conditions.

1. Introduction

E. coli is a widely distributed opportunistic pathogen commonly encountered in clinical settings, food systems, and livestock environments [1]. In addition to planktonic growth, *E. coli* can form biofilms on host tissues and abiotic surfaces, including medical devices, thereby increasing tolerance to antimicrobial agents, disinfectants, and host immune clearance[2]. Biofilm formation is therefore an important contributor to persistent infection and treatment failure[3].Given the limited efficacy of conventional antibacterial strategies against biofilm-associated infections, targeting biofilm formation rather than bacterial viability alone has emerged as an attractive anti-infective approach[4]. Such antibiofilm or antivirulence strategies may reduce bacterial colonization and persistence while imposing lower selective pressure for resistance development [5].

Baicalin, a naturally occurring flavonoid, has attracted increasing attention because of its broad availability, low toxicity, and multi-target biological activities[6]. Previous studies have demonstrated its antibacterial, anti-inflammatory, and antioxidant properties[7]. However, the molecular mechanisms by which baicalin inhibits *E. coli* biofilm formation remain poorly understood, particularly at the levels of global proteomic alterations and functional network regulation. Given the multifactorial nature of biofilm formation, quantitative proteomics is essential for comprehensively capturing treatment-induced changes in protein expression and associated biological pathways, whereas machine learning provides a reliable approach for identifying biologically meaningful patterns and prioritizing key molecular

determinants from complex datasets [8, 9]. Therefore, the integration of proteomics and machine learning provides an effective approach for elucidating the molecular mechanisms underlying antibiofilm activity [8, 10].

Here, we integrated phenotypic assays, DIA-based quantitative proteomics, bioinformatics, machine learning, and scanning electron microscopy to systematically characterize the inhibitory effects of baicalin on *E. coli* biofilm formation and to identify candidate proteins and pathways associated with this phenotype.

2. Materials and Methods

2.1 Bacterial Strain Preservation and Culture

The bacterial strains used in this study included *E. coli* D5 (PX494327), isolated from uterine mucus of a dairy cow with endometritis at a cattle farm in Datong, Shanxi, China, and *E. coli* CMCC 44113 (purchased from the China Medical Culture Collection Center).

All bacterial strains were stored at − 80°C in a cryoprotective freezing medium. For use, they were inoculated into LB broth and cultured overnight at 37°C with shaking at 220 rpm.

2.2 Materials and reagents

Baicalin (B110210) and ampicillin sodium (A105484) were purchased from Shanghai Aladdin Biochemical Technology Co., Ltd. (Shanghai, China). Crystal violet was obtained from Shanghai Macklin Biochemical Co., Ltd. (Shanghai, China). Methanol and glacial acetic acid were purchased from Sinopharm Chemical Reagent Co., Ltd. (China). The ATP assay kit (S0026) was purchased from Beyotime Biotechnology (Shanghai, China), and the ROS assay kit (E004-1-1) was obtained from Nanjing Jiancheng Bioengineering Institute (Nanjing, China). LB broth (028320) and yeast extract (050090) were obtained from Guangdong Huankai Microbial Sci. & Tech. Co., Ltd. (Guangzhou, China). Tryptic soy broth (TSB; CM0129B) was purchased from Thermo Fisher Scientific (Oxoid, UK). Sterile 96-well polystyrene microplates (11510) were purchased from Beijing Lanjieke Technology Co., Ltd. (Beijing, China).

2.3 Determination of minimum inhibitory concentration (MIC) and minimum bactericidal concentration (MBC)

The MIC and MBC were determined using the broth microdilution method in accordance with Clinical and Laboratory Standards Institute (CLSI) guidelines [11]. Overnight cultures of *E. coli* were diluted in Mueller–Hinton (MH) broth to a final concentration of 2.4×10^5 CFU/mL. Equal volumes (100 μL) of the diluted bacterial suspension and serially diluted baicalin or ampicillin sodium solutions were added to 96-well plates and incubated at 37°C for 24 h in a constant-temperature shaking incubator (HNY-2102C; Tianjin Ounuo Instrument Co., Ltd., Tianjin, China). After incubation, the optical density at 600 nm (OD_{600}) was measured using an EnSpire™ Multilabel Reader (Model 2300, PerkinElmer, Singapore). The MIC was

determined as the lowest concentration showing inhibition of bacterial growth compared with the negative control (OD_{600} nm ≤ 1.5×OD_{600} nm negative control).

For MBC determination, 10 μL from each well was plated onto MH agar plates after thorough mixing and incubated at 37°C for 24 h. The lowest concentration at which no visible colonies were observed was defined as the MBC. All experiments were performed in six replicates, and the most frequently observed value was reported.

2.4 Assessment of Biofilm Formation, MBIC Determination, and Viable Cell Counts

Biofilm formation and the inhibitory effect of baicalin were evaluated using the crystal violet staining assay [12]. Overnight *E. coli* cultures were diluted in LB broth to 1.5 × 10^7 CFU/mL and inoculated into 96-well plates (n = 6). Plates were incubated at 26°C for 24 h. After incubation, the supernatant was gently removed, and wells were washed with phosphate-buffered saline (PBS). Biofilms were stained with crystal violet, and the bound dye was subsequently solubilized using methanol and glacial acetic acid. Absorbance was measured at 600 nm to quantify biofilm biomass.

To determine the minimum biofilm inhibitory concentration (MBIC), bacterial suspensions at a final concentration of 1.5 × 10^7 CFU/mL were treated with different concentrations of baicalin prepared in LB medium and incubated under the same culture conditions. Biofilm inhibition rates were determined at 12, 24, 48, and 72 h as described above. Each concentration at each time point was tested in six replicate wells. The MBIC was defined as the lowest concentration of baicalin that achieved a biofilm inhibition rate of at least 99% at all four time points.

To determine whether the inhibitory effect of baicalin on *E. coli* biofilm formation was attributable to bactericidal activity, treatment groups were established with final baicalin concentrations of 0, 1/4, 1/2, 1, and 2× MBIC. The initial bacterial concentration in each well was adjusted to 1.5×10^7 CFU/mL. After incubation at 26°C for 24 h, the 96-well plates were subjected to ultrasonic treatment for 1 min to thoroughly disperse the bacteria in each well. Subsequently, 100 μL of the bacterial suspension was collected from each well, serially diluted 10-fold, plated onto agar medium, and incubated at 37°C for 24 h. Plates containing 30–300 colonies were selected for colony counting, and viable bacterial counts were calculated for each treatment group. The experiment was performed in triplicate.

2.5 Observation of *E. coli* biofilms by scanning electron microscopy (SEM)

Scanning electron microscopy (SEM) was employed to examine the effects of baicalin on *E. coli* biofilms[13]. Overnight cultures of *E. coli* D5 were diluted in LB broth to 3×10^7 CFU/mL and inoculated onto sterile coverslips placed in 24-well plates at a volume of 500 μL per well. Subsequently, 500 μL of baicalin solution was added to each well to achieve final concentrations of 1/4, 1/2, and 1× MBIC, with 0 MBIC used as the control, followed by incubation at 26°C for 24 h under static conditions.

After incubation, the culture medium was removed, and samples were gently washed three times with phosphate-buffered saline (PBS, pH 7.4). Cells were fixed with 2.5% glutaraldehyde for 2 h and then washed three times with PBS. Dehydration was performed using a graded ethanol series (30%, 50%, 70%, 85%, 95%, 100%, and 100%), with each step lasting 15 min. After drying, samples were sputter-coated with gold and subsequently observed under SEM for imaging.

2.6 Protein extraction, quantification, and proteomic analysis

Based on the MBIC assay described above, baicalin treatments were applied at 1/4, 1/2, 1, and 2 × MBIC, with four independent biological replicates per condition [14]. After incubation at 26°C for 2 h, bacterial cells were harvested by centrifugation at 8000 rpm for 5 min and washed three times with PBS. Total proteins were extracted by ultrasonic lysis using a lysis buffer. Protein concentration was determined using the bicinchoninic acid (BCA) assay, followed by enzymatic digestion, desalting, and mass spectrometry analysis [15].

Proteomic data were processed using R packages (ropls and ggplot2) to generate a quantitative protein matrix. Differentially expressed proteins (DEPs) were identified using the criteria of fold change ≥ 1.5 or ≤ 0.67 and P value/FDR < 0.05 and were subjected to subsequent bioinformatics analyses.

2.7 Principal component analysis and clustering analysis

Principal component analysis (PCA) and protein abundance profiling were performed on the normalized quantitative proteomic matrix to assess overall differences between groups and reproducibility among replicates. Changes in protein expression patterns across treatment groups were evaluated, and the consistency among samples within each group was examined [16].

2.8 Functional enrichment analysis

The number of DEPs under different treatment concentrations was quantified, and heatmaps were generated to visualize inter-sample variation. Differentially expressed proteins identified from K-means clustering were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses to characterize the affected biological processes, molecular functions, cellular components, and signaling pathways associated with baicalin treatment [17].

2.9 Protein–protein interaction network construction and identification of key proteins

Differentially expressed proteins were imported into the STRING database to construct a protein–protein interaction (PPI) network. Interactions with a combined score greater than the predefined threshold of 0.7 were retained. The resulting network was visualized and analyzed using Cytoscape. Topological parameters, including betweenness centrality (BC) and closeness centrality, were calculated using the Network Analyzer tool, with betweenness centrality serving as a primary metric for identifying key nodes

[18]. Functional modules within the network were identified using the MCODE algorithm, and the interconnections among modules were further examined.

2.10 Prediction of anti-*E. coli* biofilm-related proteins based on Machine Learning

Quantitative proteomic data were used as independent variables, and biofilm inhibition rates were used as the dependent variable. Least absolute shrinkage and selection operator (LASSO) regression was applied for feature selection and dimensionality reduction. The optimal penalty parameter was determined by cross-validation, and candidate proteins were selected according to the one-standard-error (1-SE) rule to obtain a parsimonious and robust model.

A random forest model was then employed to evaluate feature importance, and proteins were ranked based on IncNodePurity. The top 20 ranked proteins were intersected with the LASSO-selected candidates to identify key features. For proteins identified by LASSO, Spearman correlation analysis was performed to assess their association with biofilm inhibition rates.

To further evaluate the contribution of candidate proteins to phenotype prediction, an *in silico* virtual knockout analysis was conducted. The expression level of each candidate protein was reduced to its observed minimum value, and the biofilm inhibition rate was re-predicted using the trained LASSO model. Changes in predicted outcomes before and after perturbation were compared to assess the directional contribution of each protein.

All analyses were performed in R software, and figures were generated primarily using the ggplot2, pheatmap, and VennDiagram packages.

2.11 Determination of intracellular ATP content and ROS levels

E. coli cells were treated under the same conditions as those used for proteomic analysis. After treatment, intracellular and extracellular adenosine triphosphate (ATP) levels were measured using a commercial ATP assay kit (S0026, Beyotime, China) according to the manufacturer's instructions. Briefly, bacterial cells were collected and lysed, followed by the addition of the substrate reaction mixture. Luminescence signals were then detected and recorded using a chemiluminescence analyzer.

Intracellular reactive oxygen species (ROS) levels were determined using a fluorescence probe-based assay. According to the instructions of the assay kit (E004-1-1, Nanjing Jiancheng Bioengineering Institute, Nanjing, China), treated bacterial cells were incubated with the fluorescent probe. The excitation and emission wavelengths were set at 488 nm and 525 nm, respectively. Intracellular ROS levels in *E. coli* were evaluated based on fluorescence intensity [19].

All measurements were normalized to the corresponding OD_{600} values to account for differences in bacterial biomass, each group contained three biological replicates.

2.12 Assessment of bacterial growth, motility, and aggregation ability

Swimming and swarming motility assays were performed using agar media supplemented with different concentrations of baicalin, corresponding to final concentrations of 1/4, 1/2, 1, and 2 times the minimum biofilm inhibitory concentration (MBIC). The bacterial suspension was adjusted to 4.5×10^8 CFU/mL according to the McFarland turbidity standard. After the soft agar had solidified, 0.2 μL of the bacterial suspension was inoculated onto the center of each plate. The plates were incubated at 37°C for 24 h, after which the diameter of the turbid migration zone was measured to evaluate bacterial swimming and swarming motility [20], each group contained three biological replicates.

2.13 Observation of *E. coli* flagella by SEM

The effect of baicalin on the flagellar structure of *E. coli* was investigated using SEM [13]. The detailed experimental procedures were the same as those described in Section 4.4.

2.14 Statistical analysis

Data are presented as mean ± standard deviation(SD), with a minimum of three biological replicates per group. Statistical differences were determined using one-way analysis of variance (ANOVA), and graphs were generated using GraphPad Prism 10.6 (San Diego, USA). Differences were considered statistically significant at $P < 0.05$, with ** indicating $P < 0.01$ and *** indicating $P < 0.001$.

3. Results

3.1 Minimum inhibitory concentration (MIC) and minimum bactericidal concentration (MBC) of baicalin against *E. coli*

The MIC and MBC values of baicalin against *E. coli* were determined using the CLSI broth microdilution method. For the D5 strain, the MIC and MBC of baicalin were 60 mg/mL and 64 mg/mL, respectively, whereas those of ampicillin sodium were 512 μg/mL and 1024 μg/mL.

For the CMCC 44113 strain, the MIC and MBC of baicalin were 16 mg/mL and 32 mg/mL, respectively. In comparison, ampicillin sodium exhibited MIC and MBC values of 4 μg/mL and 8 μg/mL, respectively.

3.2 Baicalin inhibits *E. coli* biofilm formation without significantly reducing whole-well viable cell counts under the biofilm-inducing condition

The biofilm-forming capacities of strains D5 and CMCC 44113 were evaluated using the crystal violet staining assay. The results showed that strain D5 exhibited significantly stronger biofilm-forming ability than CMCC 44113 (Fig. 1a). Baicalin significantly inhibited *E. coli* biofilm formation in a concentration-dependent manner (Fig.1b, c). At 12, 24, 48, and 72 h, biofilm biomass progressively decreased with increasing concentrations of baicalin. At 64 mg/mL, the biofilm inhibition rate exceeded 99% at all four

time points, reaching 99.4%. (Fig.1d), thus determining the minimum biofilm inhibitory concentration (MBIC) as 64 mg/mL. Scanning electron microscopy revealed that, upon treatment with baicalin at 1/4, 1/2, and 1 MBIC concentrations, the biofilm biomass of *E. coli* gradually decreased (Fig. 1f).

Notably, baicalin at concentrations ranging from 1/4 to 2 MBIC did not significantly reduce viable cell counts recovered from the whole well after incubation and ultrasonic dispersion (Fig. 1e). This indicates that, under biofilm-inducing culture conditions, baicalin markedly reduced the total biofilm biomass but did not decrease the total number of viable bacteria. Therefore, under the tested biofilm-inducing condition, the weakened crystal violet staining and reduced biofilm structures observed by scanning electron microscopy were not primarily explained by a decrease in the total recoverable viable bacterial population.

3.3 Proteomic data quality and separation among treatment groups

To evaluate the variability and reproducibility of proteomic profiles across different treatment groups, principal component analysis (PCA) was performed on the normalized protein expression matrix. As shown in Fig. 2a, samples from different treatment groups exhibited clear separation in the PCA space. This trend was further supported by the violin plot (Fig. 2b), which demonstrated a consistent overall distribution of protein expression across samples, with minimal differences in median values and distribution ranges.

These results indicate that the obtained proteomic data are robust, reproducible, and comparable across groups, supporting their suitability for subsequent differential expression analysis and functional characterization.

3.4 Differential proteomic changes induced by baicalin at varying MBIC concentrations

To investigate global proteomic changes induced by different MBIC levels, differential expression analysis was performed by comparing each treatment group with the control. The number of differentially expressed proteins (DEPs) varied with baicalin concentration. The 2×MBIC group exhibited the highest number of DEPs, with 527 upregulated and 413 downregulated proteins. This was followed by the 1 × MBIC group (340 upregulated and 273 downregulated proteins), and the 1/2×MBIC group (247 upregulated and 164 downregulated proteins). In contrast, the 1/4×MBIC group showed relatively fewer changes, with 190 proteins upregulated and 60 downregulated (Fig.3a). Overall, the number of DEPs increased with increasing baicalin concentration (Fig.3c–f). Heatmap visualization combined with hierarchical clustering further demonstrated high reproducibility within groups and clear separation between groups, indicating substantial proteomic differences across treatments (Fig.3b). Venn diagram analysis identified 118 shared differentially expressed proteins across multiple MBIC treatment conditions. These commonly regulated DEPs may represent a core set of proteins involved in the antibiofilm response to baicalin, providing a foundation for subsequent identification of candidate feature (Fig. 3g).

3.5 Functional clustering and enrichment analyses reveal metabolic and motility-associated pathways

K-means clustering of differentially expressed proteins identified nine distinct expression modules (Fig. 4a). Among these, two modules did not exhibit significant enrichment, likely due to small cluster size, limited annotation coverage, or high functional heterogeneity. In addition, Cluster 4 and Cluster 9 were enriched in pathways related to cationic antimicrobial peptide resistance (eco01503) and ribosome function (eco03010), respectively, and were therefore not included in the main mechanistic analysis. The remaining five modules were subjected to GO and KEGG enrichment analyses. KEGG results revealed that the differentially expressed proteins were primarily involved in glycolysis/gluconeogenesis, the tricarboxylic acid (TCA) cycle, carbon metabolism, chemotaxis and flagellar assembly, oxidative phosphorylation, fatty acid biosynthesis, cofactor biosynthesis, and ribosomal pathways (Fig. 4b–f). These findings indicate that baicalin-mediated inhibition of biofilm formation is not attributable to a single target, but rather results from coordinated modulation of multiple biological processes, including central carbon metabolism, energy production, protein translation, as well as bacterial motility and environmental responsiveness.

GO enrichment analysis further corroborated these results, showing that the differentially expressed proteins were mainly involved in biosynthetic processes, organic nitrogen compound metabolism, and cellular metabolic processes. Most proteins were localized to the cytoplasm and exhibited binding functions, including ion binding, protein binding, and small molecule binding, which are characteristic features of enzymes involved in metabolic and biosynthetic pathways (Fig. 5g).

Collectively, these results suggest that baicalin primarily exerts its effects by reprogramming energy metabolism, motility-associated systems, and biosynthetic pathways in *E. coli*, rather than targeting a single structural component such as the outer membrane.

3.6 PPI network reveals functional modules associated with the antibiofilm activity of baicalin

Differentially expressed proteins from the five functionally relevant clusters were subjected to protein–protein interaction (PPI) network analysis using the STRING database. Based on betweenness centrality and MCODE clustering, multiple hub proteins with high centrality were identified, revealing a modular organization within the network. These modules included a membrane and motility regulation module (e.g., *OmpA* [A0A0C5L2N0], *FliA* [A0A0E0XXM1], *FliC* [Q1PAE5]), a core metabolic module (e.g., TCA cycle–related proteins GltA [A0A0E0Y3S5], *Icd* [B1B4B8], *AtpA* [A0A0E0Y7M5]), and a translation/transcription module (e.g., ribosomal protein *RpsL* [A0A0E0XWU7], *Tig* [A0A0E0Y4G5], and RNA polymerase *RpoA* [A0A0E0XUZ9]), each forming highly interconnected clusters. Notably, significant interactions were observed between these functional modules, suggesting tight regulatory coordination among energy metabolism, protein synthesis, and biofilm-associated functions. Network topology and node centrality analysis further indicated that several hub proteins, including *OmpA, RpoA, FliA, Icd, GltA*, and *AtpA*, may play critical roles in cross-module regulation.

Integration of PPI network analysis with functional enrichment results suggested a systems-level model underlying baicalin-mediated antibiofilm activity. First, baicalin was associated with changes in metabolic and energy-generating pathways centered on proteins such as *AtpA, GltA* and *Icd*. Second, it influenced protein homeostasis and transcription/translation-associated components, including *Tig* and *RpoA*, potentially reflecting stress adaptation. Third, it affected flagellar assembly and chemotaxis networks via motility-associated proteins such as *FliA* and *FliC*, thereby reducing early-stage adhesion and collective migration. Concurrent alterations in membrane-associated proteins, including *OmpA*, further suggested changes in cell surface structure and environmental adaptability. These coordinated changes are consistent with a transition from an active motility- and adhesion-competent state toward a lower-motility adaptive state, ultimately resulting in inhibition of biofilm formation.

3.7 Integrated machine learning identifies candidate protein features associated with the antibiofilm phenotype

To identify protein features most closely associated with the antibiofilm phenotype from high-dimensional proteomic data, an integrated machine learning framework was established. The quantitative proteomic matrix was used as input features, and biofilm inhibition rate was defined as the response variable. LASSO regression, random forest modeling, correlation analysis and in silico feature-perturbation analysis were combined to screen candidate proteins associated with baicalin-mediated biofilm inhibition.

LASSO regression was first applied for dimensionality reduction and feature selection. Cross-validation results indicated that a parsimonious model could be obtained under the one-standard-error criterion (Fig. 6a), yielding four candidate proteins: A0A024L5J1, A0A061Y676, A0A0B0VUB5 and A0A140NAX2 (Fig. 6b). Heatmap visualization demonstrated distinct expression patterns of these proteins across samples, suggesting their association with the response induced by baicalin (Fig. 6b). Spearman correlation analysis revealed that A0A024L5J1 was positively correlated with biofilm inhibition, whereas A0A061Y676, A0A0B0VUB5 and A0A140NAX2 showed negative correlations (Fig. 6c).

To further evaluate feature importance from a complementary modeling perspective, a random forest model was employed using IncNodePurity as the feature importance metric. Intersection analysis between the LASSO-selected proteins and the top 20 random forest features identified A0A140NAX2 as an overlapping candidate (Fig. 6e). A0A140NAX2 is annotated as a flagellin-associated protein and is therefore biologically consistent with the enrichment of flagellar assembly and chemotaxis pathways. In silico feature-perturbation analysis showed that reducing the expression value of A0A140NAX2 to its observed minimum increased the model-predicted biofilm inhibition rate, supporting a negative association between this feature and the antibiofilm phenotype (Fig. 6f).

Because these analyses are computational and based on proteomic associations, A0A140NAX2 should be interpreted as a candidate feature linked to baicalin-mediated antibiofilm activity rather than as a validated direct molecular target. Nevertheless, its identification across LASSO regression, random

forest intersection analysis and in silico perturbation analysis is consistent with the PPI network results and subsequent phenotypic validation of impaired motility and reduced flagellar abundance.

3.8 Baicalin impairs energy metabolism, motility, and flagellar abundance in *E. coli*

To validate pathways identified by proteomic enrichment, PPI network analysis and machine learning-based screening, we assessed bacterial energy metabolism, motility and flagellar structure. With increasing baicalin concentrations, intracellular ATP decreased, intracellular ROS increased, whereas extracellular ATP remained largely unchanged(Fig.7a–c), indicating that energy depletion was primarily due to impaired energy production rather than ATP leakage.

Swimming and swarming motility assays showed that, compared with the control group, both bacterial motility and swarming ability were significantly reduced, and this inhibitory effect was concentration-dependent, as evidenced by the marked decrease in migration zone diameter (Fig. 7d,e). Combined with the proteomic and functional enrichment analyses, these results suggest that baicalin impairs bacterial motility by weakening energy metabolism, chemotaxis, and flagellar function, thereby hindering the initial adhesion of *E. coli* and ultimately reducing biofilm formation[21]. This conclusion is consistent with the experimentally observed reduction in swimming and swarming abilities.

Machine learning analysis identified A0A140NAX2 as a candidate feature associated with the antibiofilm phenotype. This protein is annotated as flagellin, a key structural component of the flagellar filament. SEM was therefore performed to examine bacterial surface structures. Baicalin-treated cells exhibited a marked reduction in visible flagellar structures compared with untreated controls (Fig.7f–g). The SEM results were consistent with the machine learning screening results, indicating that baicalin may inhibit flagellar synthesis or assembly, thereby impairing bacterial motility, weakening initial adhesion, and ultimately exerting antibiofilm activity. This finding is also consistent with the enriched pathways identified by the proteomic analysis.

4. Discussion

Biofilm formation is a critical survival strategy by which *E. coli* adapts to environmental stress, enhances persistent colonization, and increases tolerance to antimicrobial agents[22]. It is also a major contributor to chronic infections and treatment failure[2, 23]. Therefore, compared with strategies that simply inhibit bacterial proliferation, anti-infective approaches targeting the process of biofilm formation have emerged as a promising research direction.

In the present study, phenotypic assays, proteomics, protein–protein interaction network analysis, machine learning, and scanning electron microscopy collectively showed that baicalin significantly inhibited biofilm formation in the *E. coli* D5 strain by interfering with pathways associated with energy metabolism, chemotaxis, and flagellar assembly. Notably, even at a baicalin concentration of 2×MBIC, the total bacterial count within the biofilm culture system did not significantly decrease, indicating that

the inhibitory effect of Baicalin on *E. coli* biofilm formation is not dependent on direct inhibition of bacterial growth.

Baicalin at 2 × MBIC primarily affected biofilm formation, with no significant difference in the number of total bacterial within the biofilm culture system, The effect of baicalin on bacterial counts in biofilm culture system of LB medium differs from that on planktonic bacteria in MH medium. It may be explained by their dependence on experimental conditions, such as medium composition, pH, inoculum size, incubation temperature, and bacterial physiological state[24–26]. In the present study, biofilm-related experiments were performed in LB medium at 26°C under high-inoculum conditions, while the determination of MIC and MBC was carried out in MH medium at 37℃ under low-inoculum conditions in accordance with CLSI guidelines. Previous studies have also shown that results obtained under standard antimicrobial susceptibility testing conditions do not necessarily reflect bacterial responses in other culture systems. For example, compared with cation-adjusted Mueller–Hinton broth, the MIC of colistin against mcr-1-positive *E. coli* was significantly increased in artificial urine, most likely due to differences in divalent cation concentrations and pH between the two systems [25]. Similarly, the antibiotic susceptibility of Bartonella differed markedly among culture media: antibiotics with low MICs in enrichment medium failed to effectively inhibit bacterial growth in standard insect medium, suggesting that medium composition and potential interactions between antibiotics and medium components can substantially affect susceptibility results [24]. Temperature is another important factor affecting antimicrobial activity. Studies on acid stress have demonstrated that antimicrobial effects vary with incubation temperature, within a certain range with stronger activity at higher temperatures and greater bacterial survival at lower temperatures [26].

Proteomic data further showed that the number of differentially expressed proteins increased with increasing baicalin concentration, indicating that baicalin broadly affected the *E. coli* protein expression profile in a concentration-associated manner. GO and KEGG enrichment analyses showed that the differentially expressed proteins were mainly involved in the TCA cycle, carbon metabolism, flagellar assembly, chemotaxis, protein translation and folding. These results are consistent with previous evidence that *E. coli* biofilm formation is not determined by a single regulatory factor, but instead results from coordinated regulation of motility, adhesion, metabolism and stress adaptation during the transition from planktonic growth to surface-associated colonization [22, 23, 27]. With respect to energy metabolism, several candidate proteins identified in the PPI network, including *GltA*, *Icd* and *AtpA*, are closely linked to the TCA cycle, central carbon metabolism and ATP synthesis. Recent studies have highlighted central metabolism as a major determinant of *E. coli* biofilm physiology, with perturbations in carbon flux, aerobic respiration and cellular redox homeostasis being sufficient to alter biofilm-associated phenotypes [28, 29].. Consistent with this view, intracellular ATP levels decreased with increasing baicalin concentration, while extracellular ATP levels remained largely unchanged. This pattern suggests that baicalin-induced ATP depletion is more likely associated with impaired energy production or metabolic remodeling than with nonspecific membrane disruption and ATP leakage. Since biofilm formation requires sustained energetic investment for initial attachment, extracellular matrix

production and community-level coordination, reduced ATP availability may contribute to the impaired biofilm-forming capacity observed after baicalin treatment [21, 28, 29] .

After baicalin treatment, intracellular reactive oxygen species (ROS) levels increased, indicating that baicalin also disrupts bacterial redox homeostasis. These findings suggest that baicalin interferes with central carbon metabolism and respiration, leading to reduced ATP synthesis and concomitant intracellular oxidative stress. In contrast, extracellular ATP levels remained largely unchanged, thereby excluding nonspecific membrane damage as the primary cause of this stress response. Given the close relationship among redox balance, bacterial motility and biofilm formation, baicalin-induced ROS accumulation may further impair the physiological processes required for early surface colonization and biofilm maturation [30, 31]. Taken together, our findings support the following working model: baicalin suppresses *E. coli* biofilm formation by disrupting chemotaxis and flagellar biosynthesis, reducing ATP availability and aggravating oxidative stress.

Altered motility represents another central feature of baicalin-mediated antibiofilm activity. Classic studies in *E. coli* demonstrated that flagella and motility are required for efficient early-stage biofilm formation, especially during surface approach and initial attachment[20]. More recent work has further emphasized that the flagellum acts not only as a locomotive organelle but also as a mechanosensory structure involved in the “swim-or-stick” transition associated with surface colonization[32, 33]. In line with these concepts, our soft agar assays showed that baicalin significantly reduced swimming and swarming, while proteomic and PPI analyses identified motility-related proteins such as FliA (A0A0E0XXM1) and FliC (Q1PAE5)/flagellin-associated components as central candidates. SEM further confirmed a marked reduction in flagellar structures following baicalin treatment. Collectively, these results indicate that baicalin interferes with flagellar assembly and motility regulation, thereby impairing bacterial surface approach, initial adhesion, and subsequent collective migration, all of which are critical early events in biofilm establishment[20, 32, 33].

Particularly noteworthy is that A0A140NAX2, annotated as flagellin, was consistently retained as a high-confidence candidate by LASSO, random forest intersection analysis, and virtual knockout validation. The convergence of machine learning, proteomics, and SEM phenotyping strengthens the inference that flagellar structure is a biologically meaningful candidate feature associated with the antibiofilm phenotype. This observation also extends previous reports on baicalin, which have mainly focused on quorum sensing inhibition, virulence attenuation, or general antibacterial effects[34, 35]. Our data suggest that, at least in *E. coli*, baicalin can further inhibit biofilm formation by modulating pathways involved in energy metabolism, chemotaxis, and flagellar assembly, thereby refining the current understanding of the mechanism of action of this natural compound.

In summary, the inhibitory effect of baicalin on *E. coli* biofilm formation is unlikely to arise from a single antibacterial mechanism. Instead, it appears to result from coordinated regulation across multiple pathways, including suppression of central carbon metabolism and ATP production, disruption of protein homeostasis and redox adaptation, and impairment of flagellar assembly, chemotaxis, and surface

adhesion. These combined effects would be expected to hinder the physiological transition from planktonic growth to stable surface-associated community formation. Therefore, this study provides a systems-level framework for understanding the antibiofilm activity of baicalin and offers a basis for future target validation and the development of metabolism- and motility-directed antibiofilm strategies. A schematic overview of the proposed antibiofilm mechanism of baicalin is presented in Fig. 8.

Declarations

Ethics approval and consent to participate

No animal experiments were performed in the present study. All experiments were conducted in vitro using *E. coli* bacterial cultures. The *E. coli* D5 strain used in this study was obtained from a pre-existing laboratory strain collection and had been isolated and preserved prior to the initiation of the present study. Although this strain was originally isolated from the uterine mucus of a dairy cow with endometritis, no animals were recruited, sampled, experimentally manipulated, anesthetized, or euthanized for the purposes of the present study. The original isolation of the D5 strain occurred prior to the initiation of this study and was not part of the experimental procedures reported in this manuscript. Therefore, animal ethics approval, informed consent from animal owners, and permissions or licenses for animal specimen collection were not applicable to the experiments reported in the present study.

Funding

This work was supported by National Natural Science Foundation of China, grant number 32473082, The China Agriculture Research System, grant number CARS-36-04; Major Agricultural Science and Technology Project of Gansu Province, grant number 24ZDNA001.

Author Contribution

Kang Zhang: Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Jingyan Zhang: Writing – review & editing, Resources, Investigation. Jingyan Zhang: Writing – review & editing, Resources, Investigation. Zhiting Guo, Guowei Xu: Writing – review & editing, Resources, Investigation. Xiaorong Lu: Writing – review & editing, Resources, Investigation. Xiaorong Lu, Ziyi Wang: Writing – review & editing, Investigation. Lei Wang: Writing – review & editing, Supervision, Resources, Project administration, Conceptualization, Funding acquisition. Jianxi Li: Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.Lei Wei: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Acknowledgement

The authors gratefully acknowledge Shanghai Biotree Biomedical Technology Co., Ltd. for its technical support in proteomic analysis and the Lanzhou Institute of Chemical Physics, Chinese Academy of Sciences, for its assistance with scanning electron microscopy imaging. The authors also sincerely thank all laboratory colleagues for their valuable assistance and support throughout the experimental work and data analysis.

Data Availability

Escherichia coli strain D5 genome sequence data were deposited into the NCBI Nucleotide database under accession number PX494327 and are available at the following URL: https://www.ncbi.nlm.nih.gov/nuccore/PX494327.1/. Protein-protein interaction (PPI) analysis was performed via the STRING database, accessible at the following URL: https://cn.string-db.org/.The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Elbehiry A, Marzouk E. From Farm to Fork: Antimicrobial-Resistant Bacterial Pathogens in Livestock Production and the Food Chain. Vet Sci. 2025;12(9). 10.3390/vetsci12090862.
2. Hall CW, Mah TF. Molecular mechanisms of biofilm-based antibiotic resistance and tolerance in pathogenic bacteria. FEMS Microbiol Rev. 2017;41(3):276–301. 10.1093/femsre/fux010.
3. Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an emergent form of bacterial life. Nat Rev Microbiol. 2016;14(9):563–75. 10.1038/nrmicro.2016.94.
4. Koo H, Allan RN, Howlin RP, Stoodley P, Hall-Stoodley L. Targeting microbial biofilms: current and prospective therapeutic strategies. Nat Rev Microbiol. 2017;15(12):740–55. 10.1038/nrmicro.2017.99.
5. Dickey SW, Cheung GYC, Otto M. Different drugs for bad bugs: antivirulence strategies in the age of antibiotic resistance. Nat Rev Drug Discov. 2017;16(7):457–71. 10.1038/nrd.2017.23.
6. Qiu X, Huang R, Xie J, Luo S, Cheng X, Cui J, et al. Recent Advances in the Therapeutic Effects and Molecular Mechanisms of Baicalin. Biology (Basel). 2025;14(6). 10.3390/biology14060637.
7. Hu Q, Zhang W, Wu Z, Tian X, Xiang J, Li L, et al. Baicalin and the liver-gut system: Pharmacological bases explaining its therapeutic effects. Pharmacol Res. 2021;165:105444. 10.1016/j.phrs.2021.105444.
8. Torun FM, Virreira Winter S, Doll S, Riese FM, Vorobyev A, Mueller-Reif JB, et al. Transparent Exploration of Machine Learning for Biomarker Discovery from Proteomics and Omics Data. J Proteome Res. 2023;22(2):359–67. 10.1021/acs.jproteome.2c00473.

9. Li X, Tian F, Zhang B, Zhang L, Chen X, Lin X, et al. Quantitative proteomics analysis reveals an important role of the transcriptional regulator UidR in the bacterial biofilm formation of Aeromonas hydrophila. Front Cell Infect Microbiol. 2024;14:1380747. 10.3389/fcimb.2024.1380747.
10. Kang S, Kong F, Liang X, Li M, Yang N, Cao X, et al. Label-Free Quantitative Proteomics Reveals the Multitargeted Antibacterial Mechanisms of Lactobionic Acid against Methicillin-Resistant Staphylococcus aureus (MRSA) using SWATH-MS Technology. J Agric Food Chem. 2019;67(44):12322–32. 10.1021/acs.jafc.9b06364.
11. Clinical L, Standards I. Performance Standards for Antimicrobial Susceptibility Testing. CLSI supplement M100. 36th ed. Malvern, PA: Clinical and Laboratory Standards Institute; 2026.
12. Wang L, Wang J, Zhang K, Zhang J, Cui D, Wang J, et al. Linalool as a potential agent for inhibiting Escherichia coli biofilm formation and exopolysaccharide production. BMC Vet Res. 2025;21(1):235. 10.1186/s12917-025-04681-4.
13. Kang J, Liu L, Wu X, Sun Y, Liu Z. Effect of thyme essential oil against Bacillus cereus planktonic growth and biofilm formation. Appl Microbiol Biotechnol. 2018;102(23):10209–18. 10.1007/s00253-018-9401-y.
14. Weng Z, Zeng F, Wang M, Guo S, Tang Z, Itagaki K, et al. Antimicrobial activities of lavandulylated flavonoids in Sophora flavences against methicillin-resistant Staphylococcus aureus via membrane disruption. J Adv Res. 2024;57:197–212. 10.1016/j.jare.2023.04.017.
15. Shi Y, Li Y, Yang K, Wei G, Huang A. Antimicrobial Peptide BCp12 Inhibits Staphylococcus aureus Growth by Altering Lysine Malonylation Levels in the Arginine Synthesis Pathway. J Agric Food Chem. 2022;70(1):403–14. 10.1021/acs.jafc.1c05894.
16. Callister SJ, Barry RC, Adkins JN, Johnson ET, Qian WJ, Webb-Robertson BJ, et al. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. J Proteome Res. 2006;5(2):277–86. 10.1021/pr050300l.
17. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. 2016;13(9):731–40. 10.1038/nmeth.3901.
18. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–13. 10.1093/nar/gky1131.
19. Song M, Liu Y, Li T, Liu X, Hao Z, Ding S, et al. Plant Natural Flavonoids Against Multidrug Resistant Pathogens. Adv Sci (Weinh). 2021;8(15):e2100749. 10.1002/advs.202100749.
20. Pratt LA, Kolter R. Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili. Mol Microbiol. 1998;30(2):285–93. 10.1046/j.1365-2958.1998.01061.x.
21. Ballén V, Cepas V, Ratia C, Gabasa Y, Soto SM. Clinical Escherichia coli: From Biofilm Formation to New Antibiofilm Strategies. Microorganisms. 2022;10(6). 10.3390/microorganisms10061103.
22. Jamal M, Ahmad W, Andleeb S, Jalil F, Imran M, Nawaz MA, et al. Bacterial biofilm and associated infections. J Chin Med Assoc. 2018;81(1):7–11. 10.1016/j.jcma.2017.07.012.

23. Costerton JW, Stewart PS, Greenberg EP. Bacterial biofilms: a common cause of persistent infections. Science. 1999;284(5418):1318–22. 10.1126/science.284.5418.1318.
24. Gadila SKG, Embers ME. Antibiotic Susceptibility of Bartonella Grown in Different Culture Conditions. Pathogens. 2021;10(6). 10.3390/pathogens10060718.
25. Loose M, Naber KG, Coates A, Wagenlehner FME, Hu Y. Effect of Different Media on the Bactericidal Activity of Colistin and on the Synergistic Combination With Azidothymidine Against mcr-1-Positive Colistin-Resistant Escherichia coli. Front Microbiol. 2020;11:54. 10.3389/fmicb.2020.00054.
26. Sorrells KM, Enigl DC, Hatfield JR. Effect of pH, Acidulant, Time, and Temperature on the Growth and Survival of Listeria monocytogenes. J Food Prot. 1989;52(8):571–3. 10.4315/0362-028x-52.8.571.
27. Hassan RM, Abd El-Maksoud MS, Ghannam IAY, El-Azzouny AA, Aboul-Enein MN. Synthetic non-toxic anti-biofilm agents as a strategy in combating bacterial resistance. Eur J Med Chem. 2023;262:115867. 10.1016/j.ejmech.2023.115867.
28. Yaeger LN, French S, Brown ED, Côté JP, Burrows LL. Central metabolism is a key player in E. coli biofilm stimulation by sub-MIC antibiotics. PLoS Genet. 2023;19(11):e1011013. 10.1371/journal.pgen.1011013.
29. Ranson TM, Barton ME, McLean RJC. Influence of central metabolism disruption on Escherichia coli biofilm formation. Can J Microbiol. 2023;69(11):432–8. 10.1139/cjm-2023-0096.
30. Kalashnikova TV, Sutormina LV, Samoilova ZY, Oktyabrsky ON, Smirnova GV. Effect of Changes in the Redox Status on Biofilm Formation in Escherichia coli. Bull Exp Biol Med. 2023;174(4):451–4. 10.1007/s10517-023-05727-z.
31. da Cruz Nizer WS, Adams ME, Allison KN, Montgomery MC, Mosher H, Cassol E, et al. Oxidative stress responses in biofilms. Biofilm. 2024;7:100203. 10.1016/j.bioflm.2024.100203.
32. Zhai Y, Tian W, Chen K, Lan L, Kan J, Shi H. Flagella-mediated adhesion of Escherichia coli O157:H7 to surface of stainless steel, glass and fresh produces during sublethal injury and recovery. Food Microbiol. 2024;117:104383. 10.1016/j.fm.2023.104383.
33. Belas R. Biofilms, flagella, and mechanosensing of surfaces by bacteria. Trends Microbiol. 2014;22(9):517–27. 10.1016/j.tim.2014.05.002.
34. Luo J, Dong B, Wang K, Cai S, Liu T, Cheng X, et al. Baicalin inhibits biofilm formation, attenuates the quorum sensing-controlled virulence and enhances Pseudomonas aeruginosa clearance in a mouse peritoneal implant infection model. PLoS ONE. 2017;12(4):e0176883. 10.1371/journal.pone.0176883.
35. Peng LY, Yuan M, Wu ZM, Song K, Zhang CL, An Q, et al. Anti-bacterial activity of baicalin against APEC through inhibition of quorum sensing and inflammatory responses. Sci Rep. 2019;9(1):4063. 10.1038/s41598-019-40684-6.

Figures

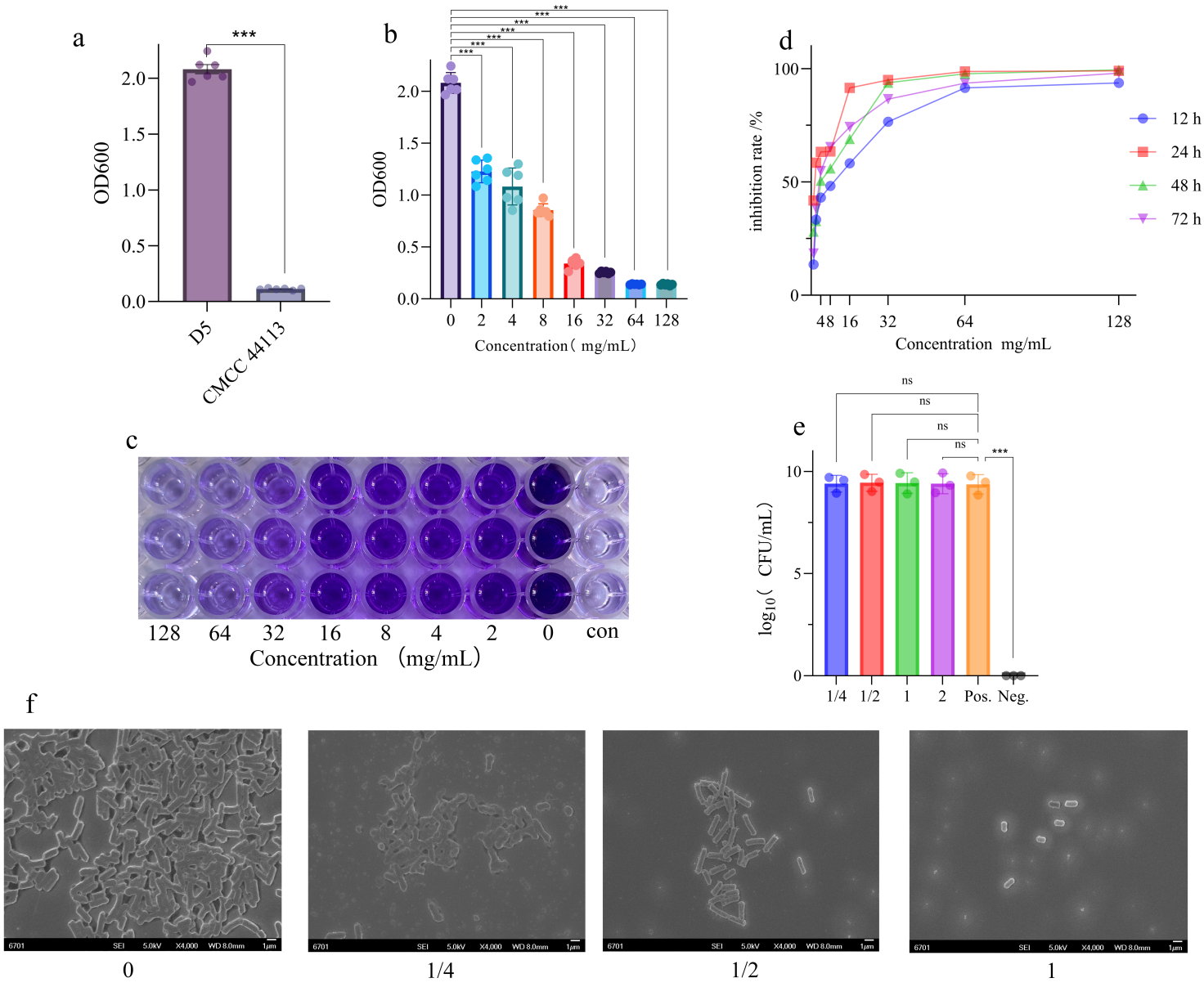


Figure 1

Baicalin inhibits *E. coli* biofilm formation without significantly reducing whole-well viable cell counts under biofilm-inducing conditions.

(a) Comparison of biofilm-forming capacities between *Escherichia coli* D5 and CMCC 44113 strains as assessed by crystal violet staining, n=6. (b) Concentration-dependent inhibitory effect of baicalin on *E. coli* biofilm formation after 24 h of incubation (n = 6). (c) Representative images of crystal violet-stained biofilms after 24 h of incubation with increasing concentrations of baicalin, n=6. (d) Biofilm inhibition rates at different concentrations of baicalin after 12, 24, 48, and 72 h of incubation (n = 6). (e) Effect of baicalin on bacterial growth in wells, expressed as colony-forming units (CFU/mL). No significant differences were observed among the 1/4–2 MBIC treatment groups, n=3. (f) Scanning electron microscopy images showing biofilm inhibition by baicalin at 1/4 to 1 MBIC concentrations. Imaging parameters: 5.0 kV; magnification 4,000×; scale bar = 1 μm; working distance (WD) = 8.0 mm. Data are

presented as mean ± standard deviation. Statistical significance is indicated as ***P < 0.001; ns denotes no significant difference, n=3.

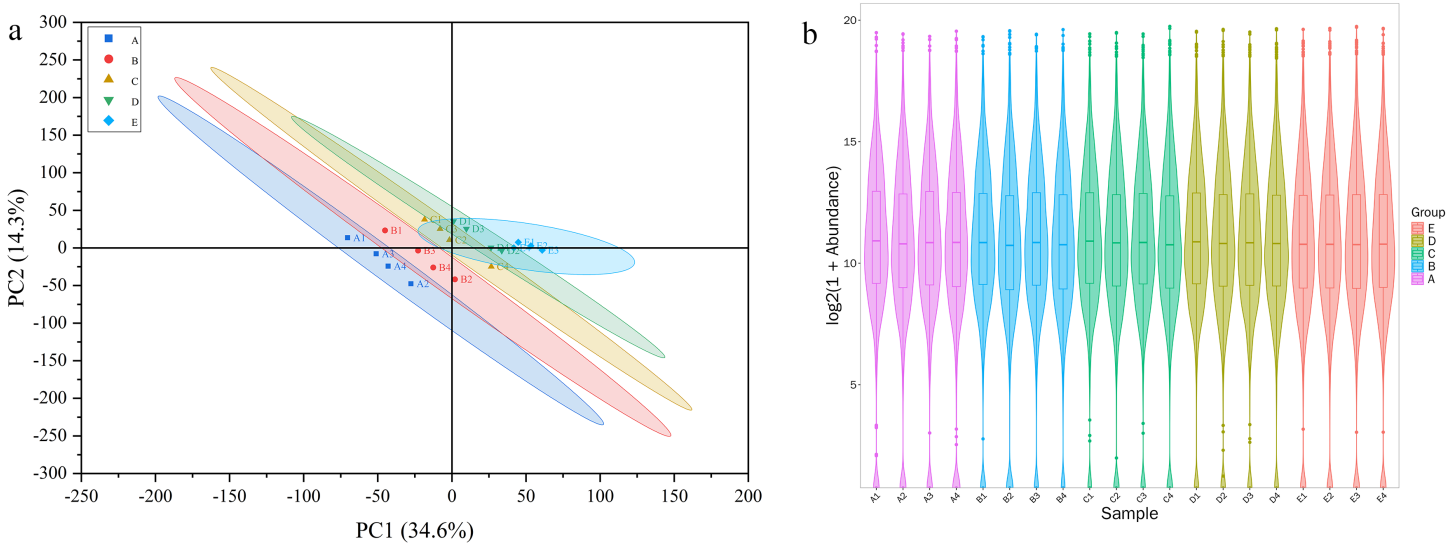


Figure 2

Proteomic profiles are robust and clearly separated across baicalin treatment groups.
(a) PCA of normalized proteomic data across treatment groups. (b) Violin plots showing the distribution of $\log_2$ protein abundance across samples.

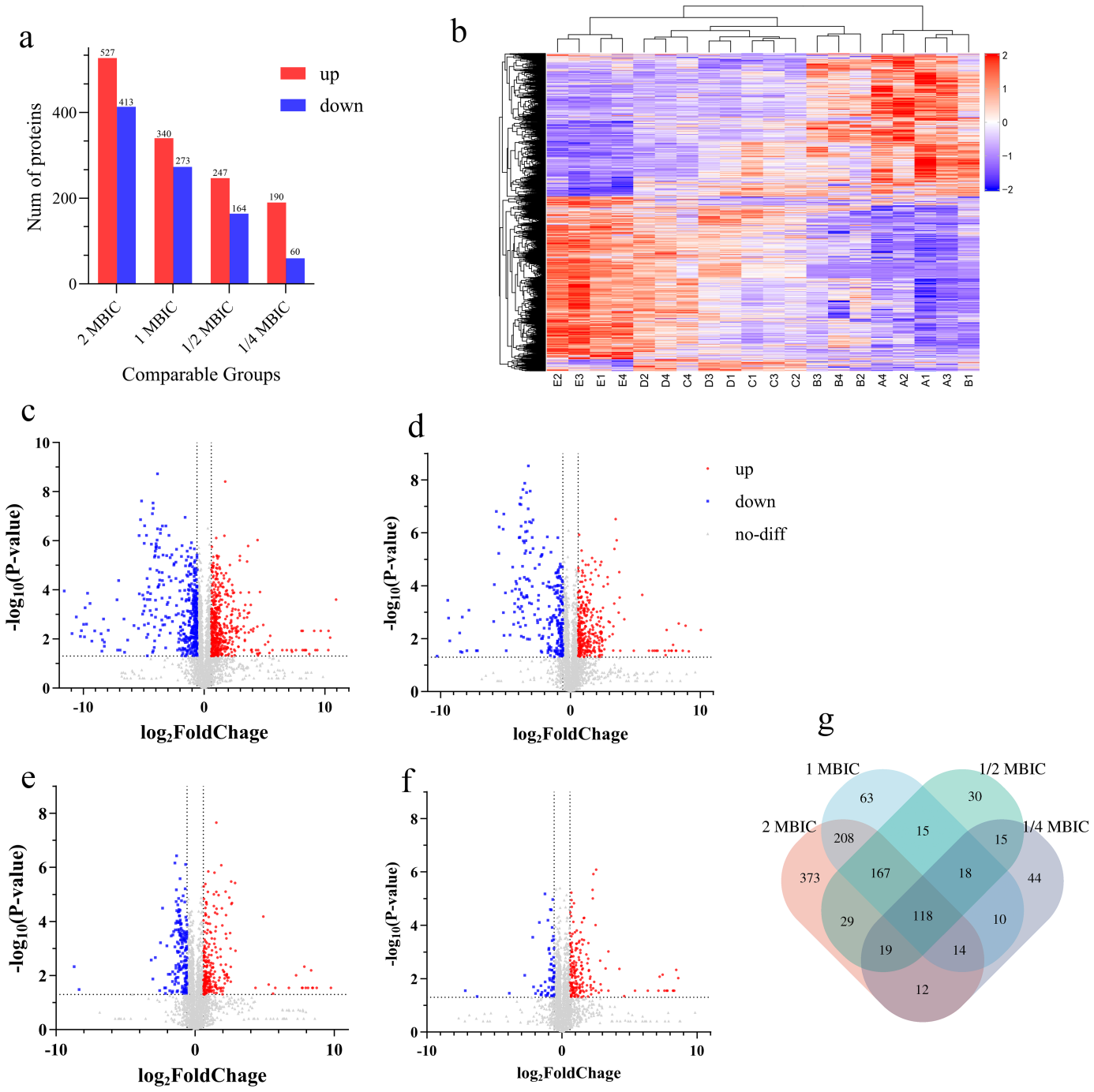


Figure 3

Baicalin induces concentration-associated changes in the *E. coli* proteome.

(a) Numbers of upregulated and downregulated proteins identified in each baicalin-treated group compared with the untreated control. Red indicates upregulated proteins, and blue indicates downregulated proteins. (b) Hierarchical clustering heatmap of differentially expressed proteins across all samples. (c–f) Volcano plots of differentially expressed proteins in *E. coli* treated with baicalin at 2×MBIC (c), 1×MBIC (d), 1/2×MBIC (e), and 1/4×MBIC (f), compared with the untreated control. Red dots indicate upregulated proteins, blue dots indicate downregulated proteins, and gray dots indicate proteins

without significant changes. (g) Venn diagram showing the overlap of differentially expressed proteins among the four baicalin-treated groups. A total of 118 proteins were commonly detected across multiple MBIC conditions.

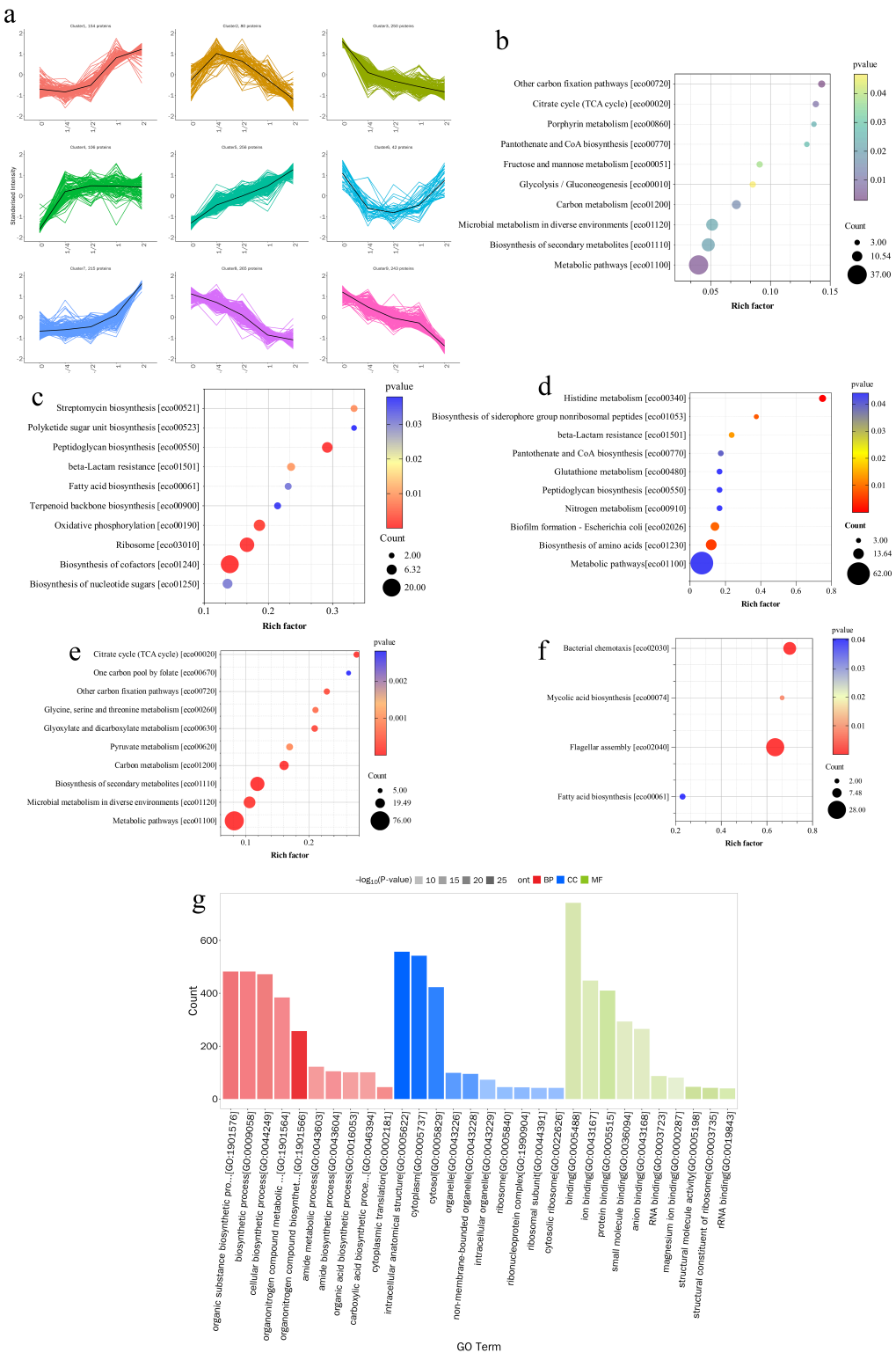


Figure 4

K-means clustering and functional enrichment analyses of differentially expressed proteins following baicalin treatment.

(a) K-means clustering identified nine expression modules with distinct abundance patterns across baicalin-treated groups. Bold lines represent the mean expression profile of each cluster, and thin lines indicate individual proteins within each module. (b–f) KEGG pathway enrichment analysis of five functionally relevant clusters. Dot size indicates the number of enriched proteins, color represents the enrichment *P* value, and the x-axis shows the Rich factor. (g) Gene Ontology (GO) enrichment analysis, with the y-axis indicating the number of mapped differentially expressed proteins and the x-axis showing the annotated GO terms.

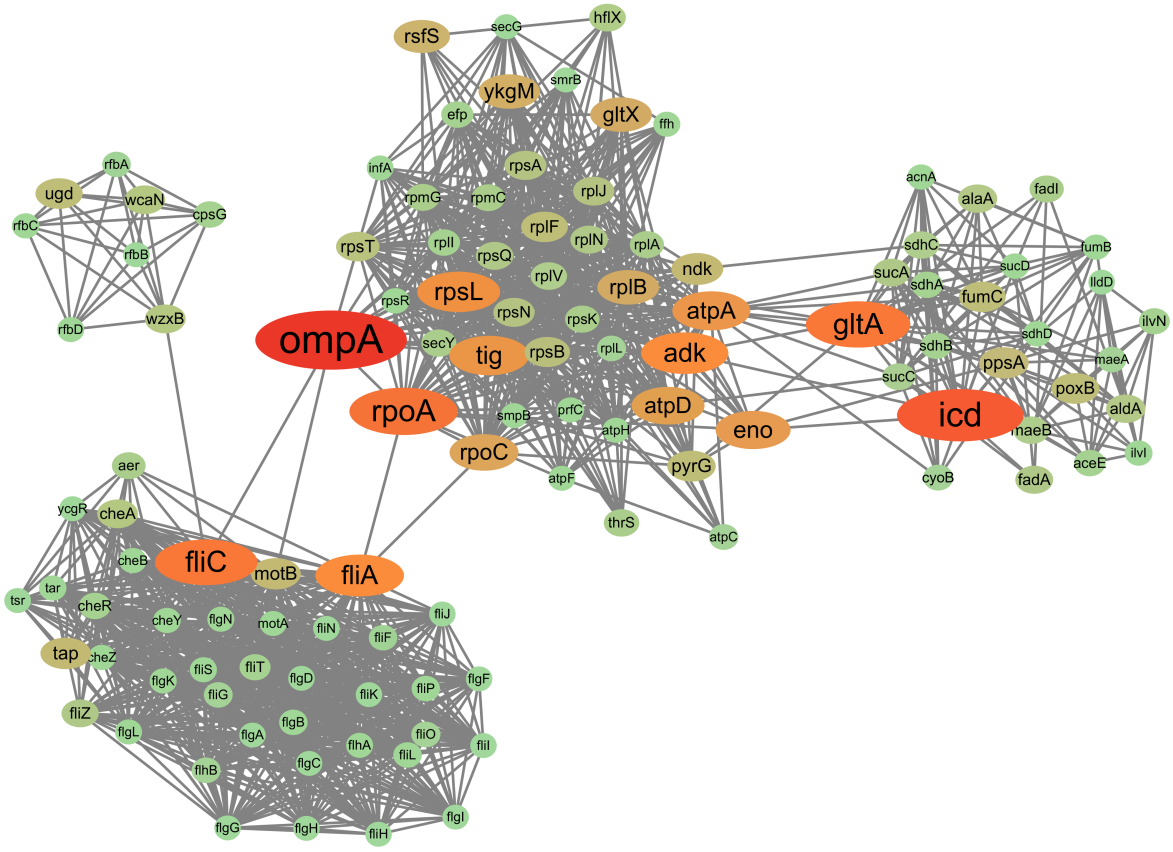


Figure 5

Protein–protein interaction network identifies hub proteins associated with metabolism, protein homeostasis, and motility

The protein–protein interaction (PPI) network of differentially expressed proteins from the five functionally relevant clusters was constructed using the STRING database and visualized with

Cytoscape. The network displayed a clear modular structure. Based on betweenness centrality and MCODE analysis, key hub proteins, including *OmpA*, *FliA*, *FliC*, *RpoA*, *RpsL*, *Tig*, *AtpA*, *GltA*, and *Icd*, were identified and highlighted in the network.

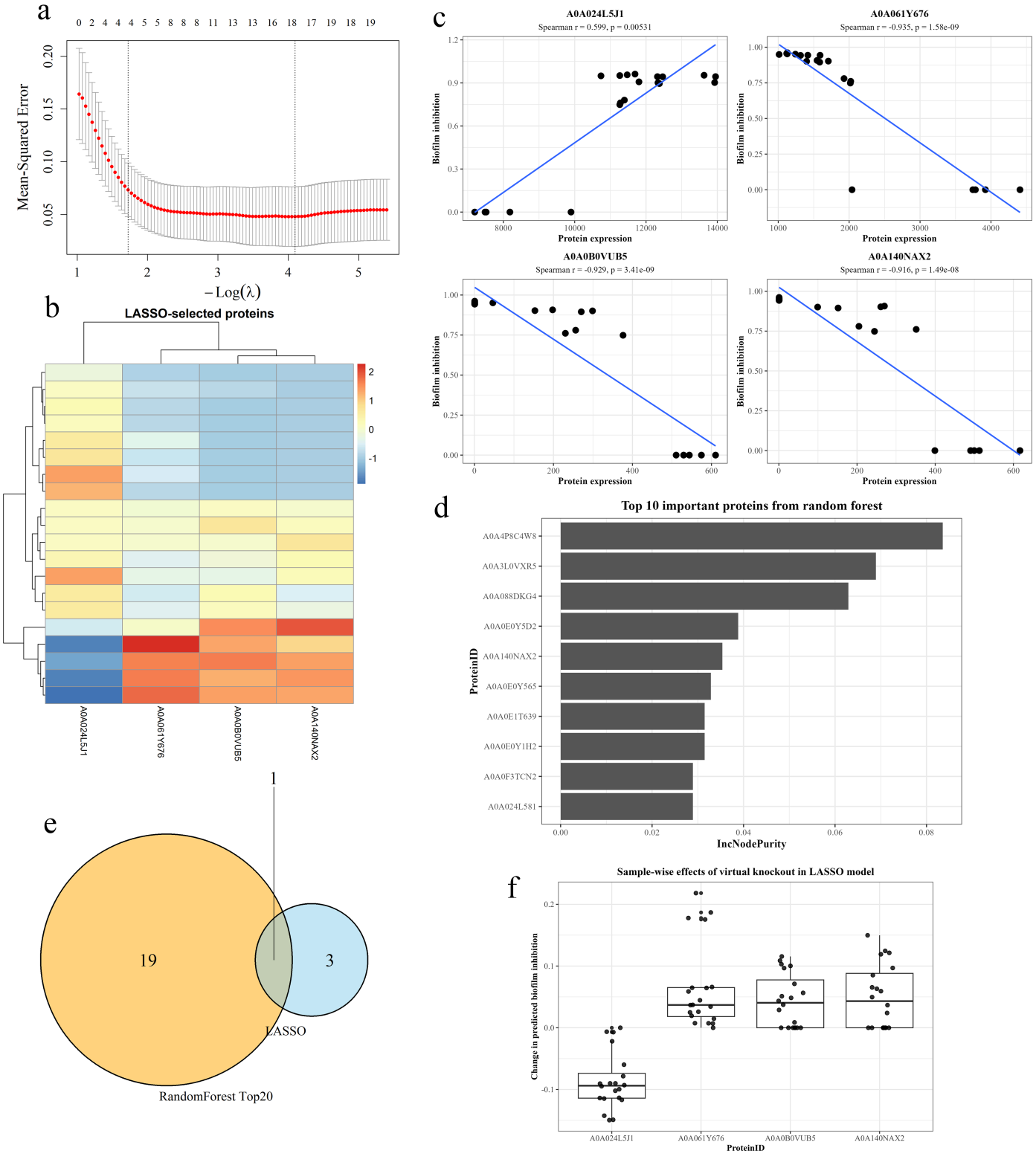


Figure 6

Integrated machine learning workflow for identification of candidate proteins associated with treatment response.

(a) Ten-fold cross-validation plot of the LASSO logistic regression model. The vertical dashed lines indicate the optimal penalty parameters selected according to the minimum criteria and the 1-standard-error rule. (b) Heatmap showing the expression patterns of proteins selected by the LASSO model across the analyzed samples. Rows represent proteins and columns represent samples. (c) Correlation analysis of representative LASSO-selected proteins with phenotypic or experimental variables. Each dot represents one sample, and the blue line indicates the linear regression fit. (d) Top 10 important proteins ranked by the random forest model according to feature importance. (e) Venn diagram showing the overlap between proteins identified by random forest and those selected by LASSO, highlighting the shared candidate proteins retained by both machine learning approaches. (f) Virtual knockout analysis of proteins retained in the LASSO model. Boxplots show the sample-wise effects of removing individual proteins on model performance, indicating their relative contributions to predicted biofilm inhibition rate.

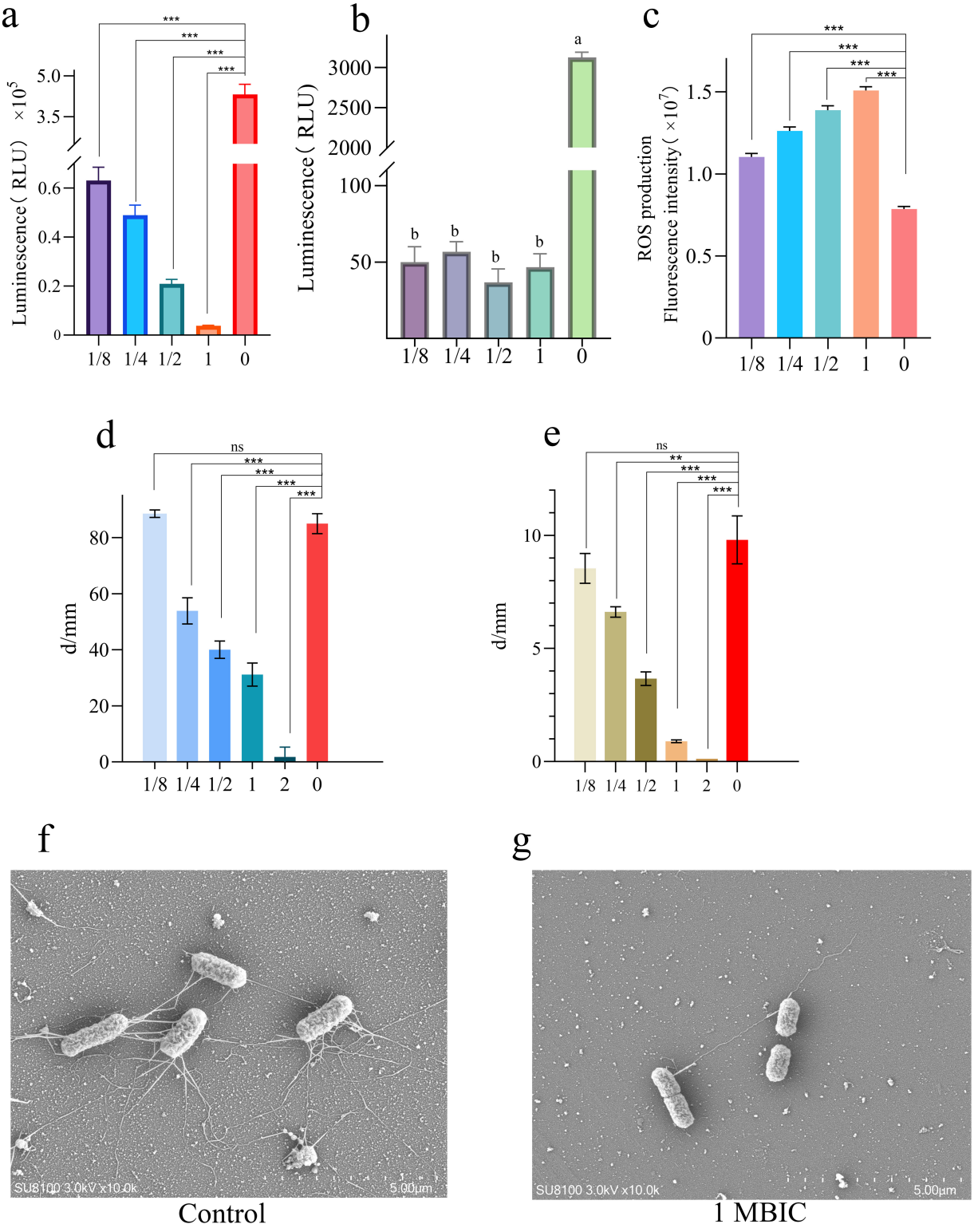


Figure 7

Baicalin impairs energy metabolism, motility, and flagellar abundance in *E. coli*.

(a, b) Changes in intracellular and extracellular adenosine triphosphate (ATP) levels in *E. coli* after treatment with different concentrations of baicalin, n=3. (c) Intracellular ROS levels following baicalin treatment. ROS levels were elevated in a concentration-dependent manner, n=3. (d, e) Swimming and

swarming motility assays showed that baicalin inhibited bacterial swimming and swarming motility, n=3. (f, g) Scanning electron microscopy images of *E. coli* D5 cells in the untreated control group (f) and after treatment with baicalin at 1 MBIC (g) Baicalin treatment showed fewer visible surface flagella, n=3.

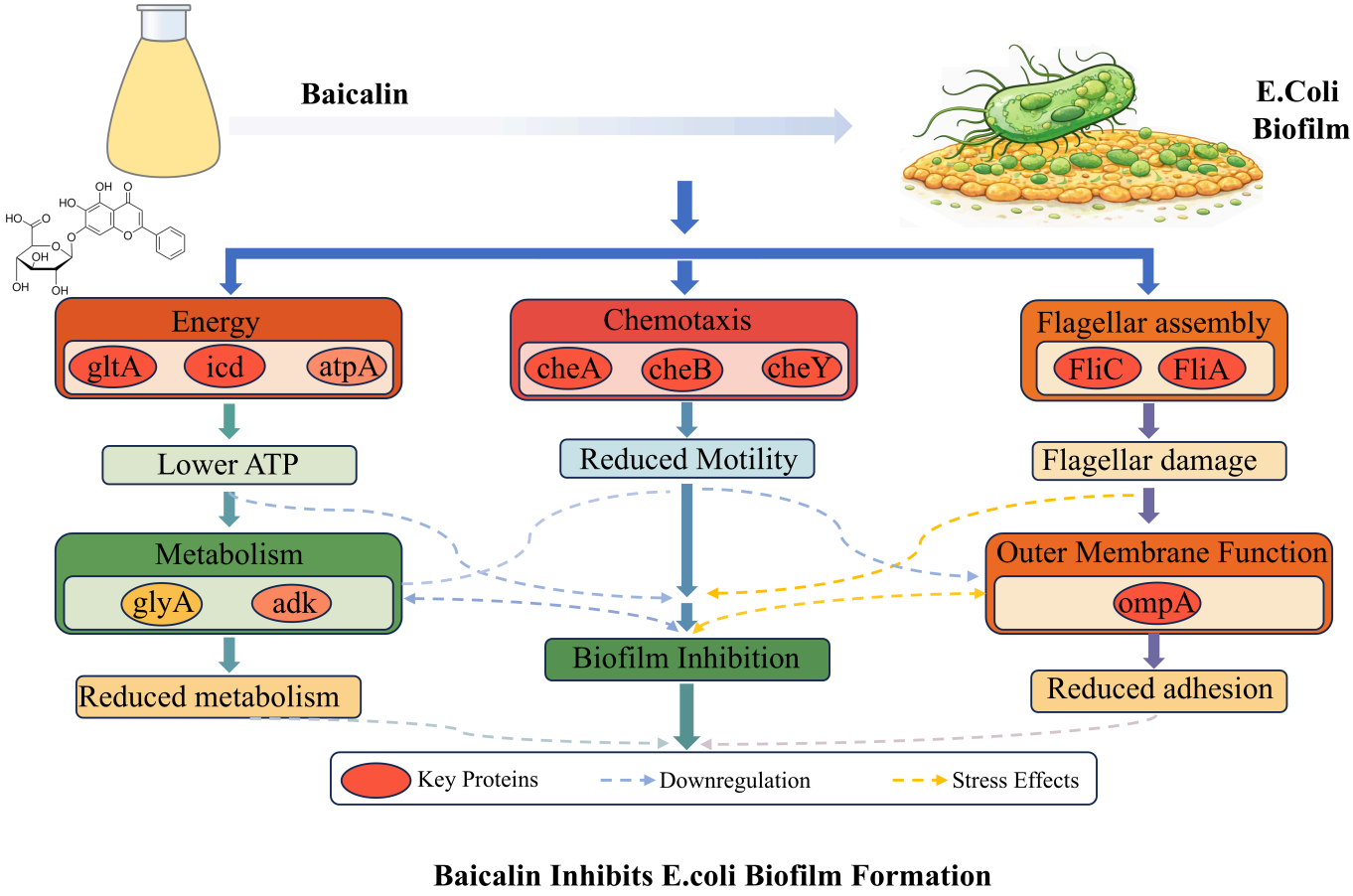


Figure 8

Schematic model of the mechanisms underlying baicalin-mediated inhibition of *E. coli* biofilm formation.

Baicalin suppresses *E. coli* biofilm formation by coordinately disrupting metabolism and energy production, chemotaxis and motility, outer membrane-associated functions, and protein homeostasis. These changes lead to reduced ATP levels, impaired motility and adhesion, increased metabolic stress, and ultimately inhibition of biofilm formation.