



Cite this: DOI: 10.1039/d6sc00084c

 All publication charges for this article have been paid for by the Royal Society of Chemistry

# Artificial intelligence-assisted phenotyping of drug-resistant bacteria using a monosaccharide-based fluorescent sensor array

Zhuo-Fan Zhang,<sup>†a</sup> Wen-Zhen Gui,<sup>†a</sup> Yi-Fan Tang,<sup>†c</sup> Hui-Qi Gan,<sup>a</sup> Xi-Le Hu,<sup>\*a</sup> Tony D. James,<sup>ib</sup>\*fg Shengbing Yang,<sup>\*e</sup> Qian Liu<sup>\*d</sup> and Xiao-Peng He<sup>ib</sup>\*ab

Chemical tools capable of effectively phenotyping drug-resistant bacteria can help improve therapeutic efficacy toward bacterial infections. While conventional techniques rely on labor-intensive procedures for the determination of bacterial susceptibility to antibiotics, here we developed a sensor array based on fluorogen-labelled monosaccharides to accurately phenotype drug-resistant bacteria with the assistance of artificial intelligence (AI). D-Glucose, D-galactose, L-fucose and D-mannose, which are common monomeric building blocks of natural glycans, were labelled with a "conformationally-adaptive" fluorophore (DPAC) with two different linkers, giving rise to a sensor array that consists of eight fluorescent glycoprobes. Using homogeneous high-throughput screening, we found that all the glycoprobes exhibited sensitive ratiometric fluorescence changes in the presence of *Pseudomonas aeruginosa* (*P. aeruginosa*) expressing bacterial lectins (LecA and LecB) selective for D-galactose, L-fucose and D-mannose. However, minimal fluorescence changes were seen when the glycoprobes were incubated with other bacterial strains lacking lectin expression. The use of ensemble learning to process the acquired sensing signals further enabled the accurate discrimination of clinically isolated, drug-resistant *P. aeruginosa* from drug-sensitive strains. Interestingly, using AI-assisted array sensing, we also achieved the phenotyping of *P. aeruginosa* after long-term exposure to mechanistically different antibiotics, thus highlighting the effectiveness of this approach for precision medicine.

Received 5th January 2026  
Accepted 5th May 2026

DOI: 10.1039/d6sc00084c

rsc.li/chemical-science

## Introduction

Bacterial infections are a global challenge for public health, with multidrug-resistant (MDR) pathogens posing serious threats.<sup>1</sup> A representative example is *Pseudomonas aeruginosa* (*P. aeruginosa*), a Gram-negative, opportunistic pathogen designated by the World Health Organization (WHO)<sup>2</sup> as a "critical

priority" MDR bacterium due to its diverse drug-resistant mechanisms and rapid adaptive evolution. Infection with *P. aeruginosa* can lead to many clinical complications including pneumonia,<sup>3</sup> bacteremia,<sup>4</sup> burn-wound sepsis,<sup>5</sup> and refractory chronic pulmonary colonization in cystic fibrosis patients.<sup>6</sup> Notably, *P. aeruginosa* develops extensive drug resistance (XDR) or pandrug resistance (PDR) through mechanisms such as biofilm formation, efflux pump overexpression, and acquired  $\beta$ -lactamases, with a global carbapenem resistance rate of >30%.<sup>1,7,8</sup> These emerging resistance patterns underscore the critical need for effective diagnostic approaches to accurately identify drug resistance.<sup>9</sup>

Current gold-standard methods-including bacterial culture, microscopic examination, biochemical assays, serological tests, nucleic acid hybridization, and PCR-face substantial limitations.<sup>10</sup> They are often time-consuming (24–72 hours of culturing), exhibit suboptimal sensitivity/specificity (particularly for polymicrobial infections), require technical expertise, and cannot differentiate viable from non-viable bacteria.<sup>11</sup> Such constraints compromise timely clinical interventions, especially in managing MDR infections where early appropriate antimicrobial therapy significantly impacts outcomes.<sup>12,13</sup>

In recent years, fluorescent sensor arrays have emerged as a promising analytical tool for the discrimination of

<sup>a</sup>Key Laboratory for Advanced Materials and Joint International Research Laboratory of Precision Chemistry and Molecular Engineering, Feringa Nobel Prize Scientist Joint Research Centre, School of Chemistry and Molecular Engineering, East China University of Science & Technology, Shanghai 200237, China. E-mail: xlhu@ecust.edu.cn; xphe@ecust.edu.cn

<sup>b</sup>National Center for Liver Cancer, The International Cooperation Laboratory on Signal Transduction, Eastern Hepatobiliary Surgery Hospital, Shanghai 200438, China

<sup>c</sup>Academy of Pharmacy Xi'an-Jiaotong Liverpool University, Suzhou, 215123, China

<sup>d</sup>Department of Laboratory Medicine, Ren Ji Hospital, 160 Pujian Road, Shanghai 200127, China. E-mail: liuqian\_rj@shsmu.edu.cn

<sup>e</sup>Shanghai Key Laboratory of Orthopedic Implants, Department of Orthopedic Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200011, China. E-mail: shengbingyang@shsmu.edu.cn

<sup>f</sup>Department of Chemistry, University of Bath, Bath BA2 7AY, UK. E-mail: t.d.james@bath.ac.uk

<sup>g</sup>School of Chemistry and Chemical Engineering, Henan Normal University, Xinxiang 453007, China

<sup>†</sup> Equal contribution.



environmental and biological samples.<sup>14,15</sup> Sensor arrays offer significant advantages, such as high sensitivity, strong anti-interference ability, and reliable visual detection results, compensating for the shortcomings of traditional diagnostic methods. A variety of fluorescent sensor arrays based on different principles have been developed. One popular approach is the construction of an array of polymeric or small-molecule fluorescent probes that exhibit moderate affinities to multiple binding targets. The subsequent use of dimension-reduction analyses to process the generated data allows for the differentiation of analytes. For example, dual-fluorescence sensor arrays based on poly(arylene ethynylene) and fluorophores characteristic of aggregation-induced emission (AIE),<sup>16</sup> fluorescent probes based on 3-hydroxyflavone derivatives,<sup>17</sup> and nitroreductase-based “turn-on” fluorescent probes<sup>18</sup> have been successfully used to establish sensor arrays that enable the classification of Gram-positive and Gram-negative bacteria. As such, we constructed a sensor array based on fluorogens that exhibit dual-fluorescence emission, for the phenotyping of methicillin-resistant *Staphylococcus aureus* (MRSA).<sup>19</sup> Additionally, sensor arrays based on nanomaterials have been applied for bacterial sensing. For instance, Phillips and co-workers<sup>20</sup> used gold nanoparticles and poly(phenylene ethynylene) to establish a sensor array capable of classifying twelve types of bacteria. Wang and co-workers<sup>21</sup> developed a paper-based sensor array using functionalized carbon dots to achieve rapid identification of five types of bacteria with the aid of machine learning.

Despite the rapid progress in sensors for bacterial classification, the binding interaction between the developed probes and bacteria are mainly based on electrostatic attraction, which could easily be interfered with by electrolytes present in the test solutions. Therefore, the exploitation of more specific biological features of bacteria as sensing targets might improve diagnostic accuracy. As a proof-of-concept, here we developed a fluorescent sensor array that consists of monosaccharides that target the virulence-associated lectins of *P. aeruginosa*, LecA (PA-IL, galactose-selective) and LecB (PA-IIL, fucose- and mannose-selective).<sup>22,23</sup> These lectins have been shown to modulate pathoadaptive resistance of *P. aeruginosa* by facilitating biofilm formation and immune evasion, and can coordinate with resistance mechanisms such as efflux pumps (e.g., MexAB-OprM) and carbapenemases, particularly in high-risk clones (e.g., ST235/ST111).<sup>8</sup> Clinically, these adaptations lead to the generation of MDR, XDR or even PDR phenotypes, underscoring the need for rapid and accurate diagnostics to guide antibiotic selection, thus overcoming treatment failures in critical care.<sup>24–27</sup>

To the best of our knowledge, this study reports the first fluorescent sensor array that targets bacterial lectin-mediated resistance signatures, offering a useful tool for the early intervention of bacterial infections as well as outbreak containment. We used previously developed DPAC (*N,N'*-disubstituted-dihydrodibenzo[*a,c'*]phenazines)-based fluorogen characteristic of “conformational adaptivity”<sup>28,29</sup> to label four common monosaccharides, D-glucose, D-galactose, L-fucose and D-mannose, using two different linkers and click chemistry (Fig. 1a). These

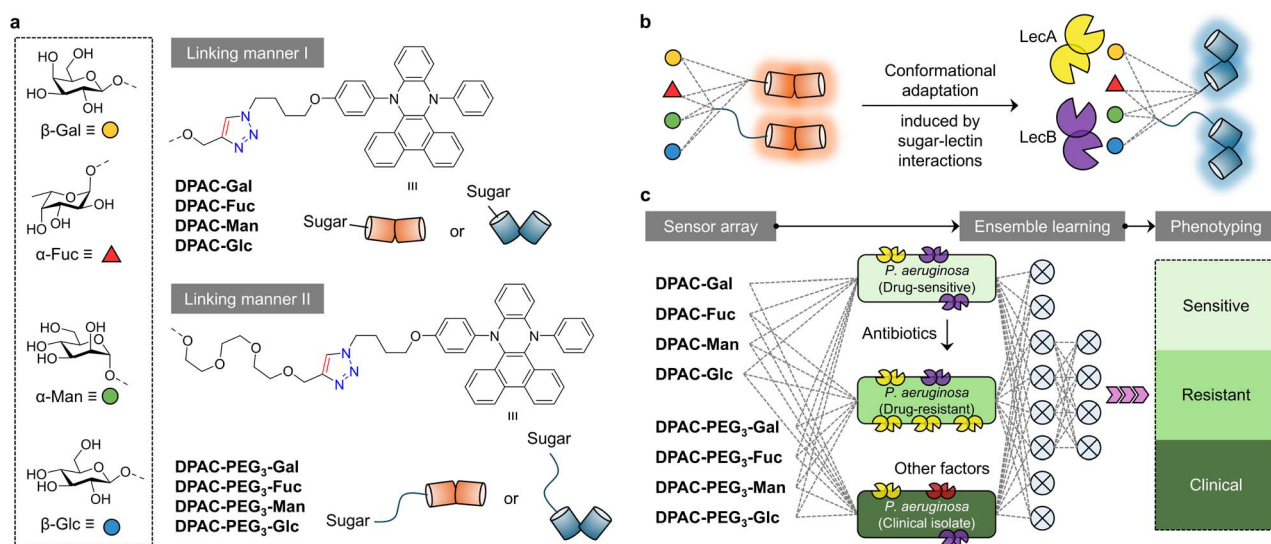
glycophores exhibited ratiometric fluorescence changes upon lectin binding (Fig. 1b). Subsequently, a sensor array that consists of the eight glycophores was subsequently established for bacterial classification. We determined that targeting their endogenous LecA and LecB expression, as well as the up or downregulation of both lectins under different antibiotic stress gave rise to the accurate phenotyping of drug-sensitive, drug-resistant and clinically isolated *P. aeruginosa* strains with the assistance of artificial intelligence (AI) (Fig. 1c).

## Results and discussion

The sensor array constructed in this study consists of eight monosaccharide-based glycophores (DPAC-Gal, DPAC-PEG<sub>3</sub>-Gal, DPAC-Fuc, DPAC-PEG<sub>3</sub>-Fuc, DPAC-Man, DPAC-PEG<sub>3</sub>-Man, DPAC-Glc, DPAC-PEG<sub>3</sub>-Glc) (Fig. 1a). Given that the sugar-binding properties of bacteria can differ largely, we anticipated that these glycophores would generate distinct sensing signals for different bacterial species, thus facilitating classification (Fig. 1b). The probes were synthesized by conjugating D-glucose, D-galactose, D-mannose and L-fucose through linkers of varying lengths with DPAC using the Cu(I)-catalyzed azide-alkyne cycloaddition click reaction. As shown in Schemes S1 and S2, the four monosaccharides were first glycosylated with propargyl alcohol or an O-alkynyl trimeric polyethylene glycol (PEG<sub>3</sub>) under rigorously anhydrous and anaerobic conditions using boron trifluoride diethyl ether (BF<sub>3</sub>·Et<sub>2</sub>O) for 6 hours. The resulting eight intermediates (G1–G8) were then treated with an azide-modified DPAC derivative (DPAC-N<sub>3</sub>) in the presence of Cu<sub>2</sub>SO<sub>4</sub>·H<sub>2</sub>O and sodium ascorbate at room temperature for 6 hours to afford G9–G16 in 85–95% yields, respectively. Finally, deacetylation of G9–G16 gave the desired glycophores in almost quantitative yields. Besides the glycophores, two control fluorogens without sugar modification (DPAC-OH and DPAC-PEG<sub>3</sub>-OH) were also synthesized as negative controls for subsequent binding assays (Schemes S1 and S2).

With the probes in hand, we first evaluated their photo-physical properties using UV-vis and fluorescence spectroscopy in phosphate buffered saline (PBS) with 0.02% Tween 20 solution. We found that the glycophores exhibited UV-vis absorption maximum over a range of 354–365 nm (Fig. S1a–e), characteristic of DPAC.<sup>30</sup> Two fluorescence emission bands centered at 452–493 nm and 590–616 nm, which are assigned to a constrained and planar conformation of DPAC,<sup>30</sup> respectively, were detected in their representative fluorescence spectra (Fig. S1f–j). The observation of dual fluorescence emission suggests that the amphiphilic DPAC-based glycophores, with the presence of the hydrophilic sugar groups, maintain equilibrium between the constrained and planar conformation of DPAC. Notably, the red (*I*<sub>601–616</sub>)/blue (*I*<sub>452–493</sub>) fluorescence intensity ratios of all probes were determined to be similar irrespective of variations in sugar structure and linker chain length. This permits a minimal analytical bias in the following sensing assays that might be caused by variations in the initial conformational state (i.e., large differences in fluorescence intensity ratios) of the probes. The control probes (DPAC-PEG<sub>3</sub>-OH and DPAC-OH) exhibited similar UV-vis absorption and dual-fluorescence emission





**Fig. 1** (a) Structure of the DPAC-based fluorescent glycoprobes with different linkers between sugars and the fluorogen. Schematic illustration of (b) the conformational adaptation of the glycoprobes upon sugar–lectin interactions to produce ratiometric fluorescence changes, and (c) the construction of a homogenous sensor array for the ensemble learning-assisted phenotyping of *P. aeruginosa* variants (drug-sensitive, drug-resistant and clinically isolated) with varying expression levels of LecA and LecB.

profiles to those of the glycoprobes (Fig. S1e and j). This observation confirms that the conjugation of monosaccharide moieties does not alter the inherent photophysical properties and the initial conformational equilibrium of the DPAC core.

Next, we evaluated the sensing properties of the glycoprobes for plant lectins that selectively bind monosaccharides. Concanavalin A (ConA – glucose and mannose selective), peanut agglutinin (PNA – galactose selective) and lentil agglutinin (LCA – core-fucose and high-mannose selective) were selected. After the addition of increasing concentrations (1–10  $\mu\text{M}$ ) of lectins to the glycoprobes, a concentration-dependent increase in their blue fluorescence emission intensity (465–487 nm) was observed when a selective lectin was mixed with the corresponding glycoprobe (*i.e.*, ConA with **DPAC-Man** and **DPAC-PEG<sub>3</sub>-Man**, **DPAC-Glc** and **DPAC-PEG<sub>3</sub>-Glc**; PNA with **DPAC-Gal** and **DPAC-PEG<sub>3</sub>-Gal**) (Fig. 2). In contrast, the red emission intensities (600–620 nm) of the probes remained relatively unchanged (Fig. 2). These results suggest that the selective recognition between monosaccharides and lectins draws the adjacent DPAC core into a peripheral domain of the protein. Supramolecular interactions (*e.g.*, hydrophobic and electrostatic forces) with the surrounding amino acid residues restrict the intramolecular vibrations of DPAC, thus stabilizing its bent, constrained conformation. This binding-induced conformational restriction effectively triggers the ratiometric red-to-blue fluorescence shift, a mechanism that is consistent with our previous Small-Angle X-ray Scattering (SAXS) investigations on DPAC-protein interactions<sup>19,31</sup>

Furthermore, a range of other physiological proteins, including human serum albumin (HSA), bovine serum albumin (BSA), lysozyme (Lys), human immunoglobulin G (hIgG), ribonuclease A (RNase A), pepsin (PeP), and trypsin (Tryp), were used as negative controls. When treating the glycoprobes with

an unselective lectin and other control proteins, their red/blue emission ratios were only slightly changed, underscoring that the spectral changes seen arose mainly from selective lectin–monosaccharide interactions (Fig. S2–S9). This specificity was further supported by using the control probes (**DPAC-OH** and **DPAC-PEG<sub>3</sub>-OH**), which exhibited negligible ratiometric changes upon incubation with the same lectins (Fig. S10–S12), precluding the interference of non-specific hydrophobic interactions.

Having established the ratiometric fluorescence responses of the DPAC-based glycoprobes upon lectin recognition, we proceeded to construct a fluorescent sensor array for differential sensing of bacterial species. This design adopts the sensor-array concept pioneered by Anslyn and co-workers,<sup>32</sup> and employs a panel of receptors exhibiting moderate cross-reactivity toward multiple analytes. The resulting response patterns were then processed *via* multivariate statistical analyses (*e.g.*, PCA) to generate “fingerprints” for individual analytes<sup>33</sup>—an approach having been successfully implemented for differentiating biologically relevant species.<sup>34</sup> We therefore envisioned that the differential changes in the blue and red emission intensities of the glycoprobes upon bacterial binding could enable classification.

The bacterial strains used include *P. aeruginosa* (ATCC 27853 and clinically-isolated PA 3887 and PA A258), *Klebsiella oxytoca* (*K. oxytoca*, ATCC 13883), *Acinetobacter baumannii* (*A. baumannii*, ATCC 19606), *Escherichia coli* (*E. coli*, ATCC 25922 and ATCC 35218), *Staphylococcus aureus* (*S. aureus*, ATCC 29213 and ATCC 43300), *Staphylococcus epidermidis* (*S. epidermidis*, CMCC 26069) and *Salmonella typhi* (*S. Typhi*, CMCC 32206). Of note are *P. aeruginosa* and *A. baumannii*, which are known to express bacterial lectins. *P. aeruginosa* expresses LecA (galactose-selective) and LecB (fucose and mannose-selective), two



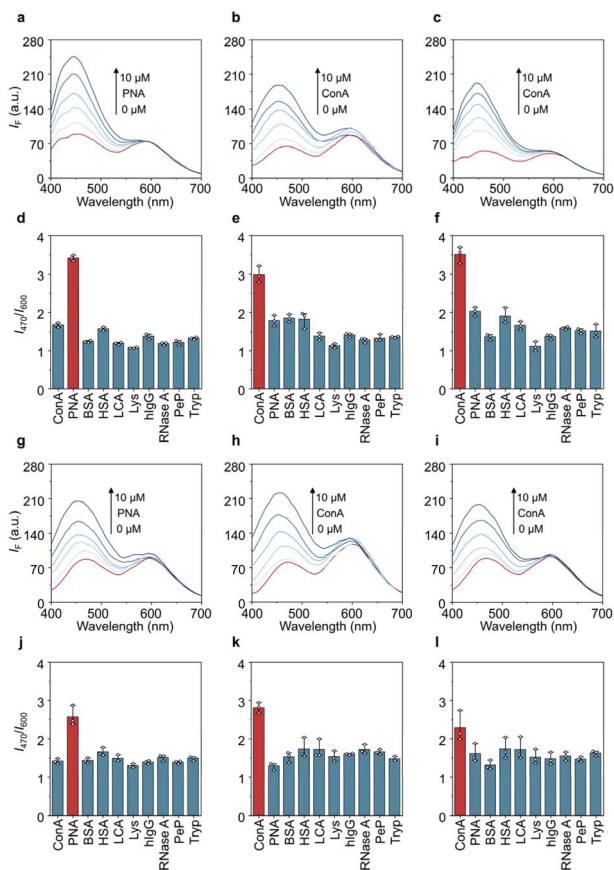


Fig. 2 Fluorescence emission spectra of (a) DPAC-PEG<sub>3</sub>-Gal with increasing concentrations of PNA, (b) DPAC-PEG<sub>3</sub>-Glc with increasing concentrations of ConA, (c) DPAC-PEG<sub>3</sub>-Man with increasing concentrations of ConA, (g) DPAC-Gal with increasing concentrations of PNA, (h) DPAC-Glc with increasing concentrations of ConA, and (i) DPAC-Man with increasing concentrations of ConA. Changes in fluorescence intensity ratio ( $I_{470}/I_{600}$ ) of (d) DPAC-PEG<sub>3</sub>-Gal, (e) DPAC-PEG<sub>3</sub>-Glc, (f) DPAC-PEG<sub>3</sub>-Man, (j) DPAC-Gal, (k) DPAC-Glc, and (l) DPAC-Man in the presence of different proteins (10  $\mu$ M). All measurements were done in PBS (0.01 M, pH 7.4, containing 0.02% (v/v) Tween 20) at a probe concentration of 10  $\mu$ M and an excitation at 360 nm.

virulent lectins that mediate biofilm formation and host infection. *A. baumannii* expresses the Acinetobacter trimeric auto-transporter adhesin (Ata, galactose-selective) that functions as a virulence factor mediating bacterial adhesion to host cells and facilitating biofilm formation.<sup>35</sup> These two bacteria are thus considered to be ideal targets for our monosaccharide-based glycoprobes. Other test bacteria that lack lectin expression were used as controls. To initiate the sensing, the eight glycoprobes were added to a 96-well microplate, followed by an orthogonal incubation with the bacteria. Then, the fluorescence intensity changes of the resulting mixtures at 470 nm and 600 nm were recorded in a high-throughput manner.

We first found that the ratiometric changes ( $I_{470}/I_{600}$ ) of all glycoprobes changed most significantly in the presence of *P. aeruginosa* strains over time, with the standard ATCC 27853 strain inducing the strongest signal variations. Generally,

probes incorporating a shorter linker elicited higher signal responses than the PEG-modified equivalents, which suggests that the shorter linker favors binding between glycoprobes and bacterial lectins. Among the LecB-targeting glycoprobes, the fucosylated variant outperformed the mannosylated one, which agrees with a previous study suggesting a higher affinity of fucosides than mannosides to LecB.<sup>36</sup> An unexpected observation was the ratiometric changes of glucose-modified probes (DPAC-PEG<sub>3</sub>-Glc and DPAC-Glc) with ATCC 27853. This may be attributed to the structural similarity between glucose and mannose (C2 epimer to each other), which results in unselective binding with bacterial lectins.

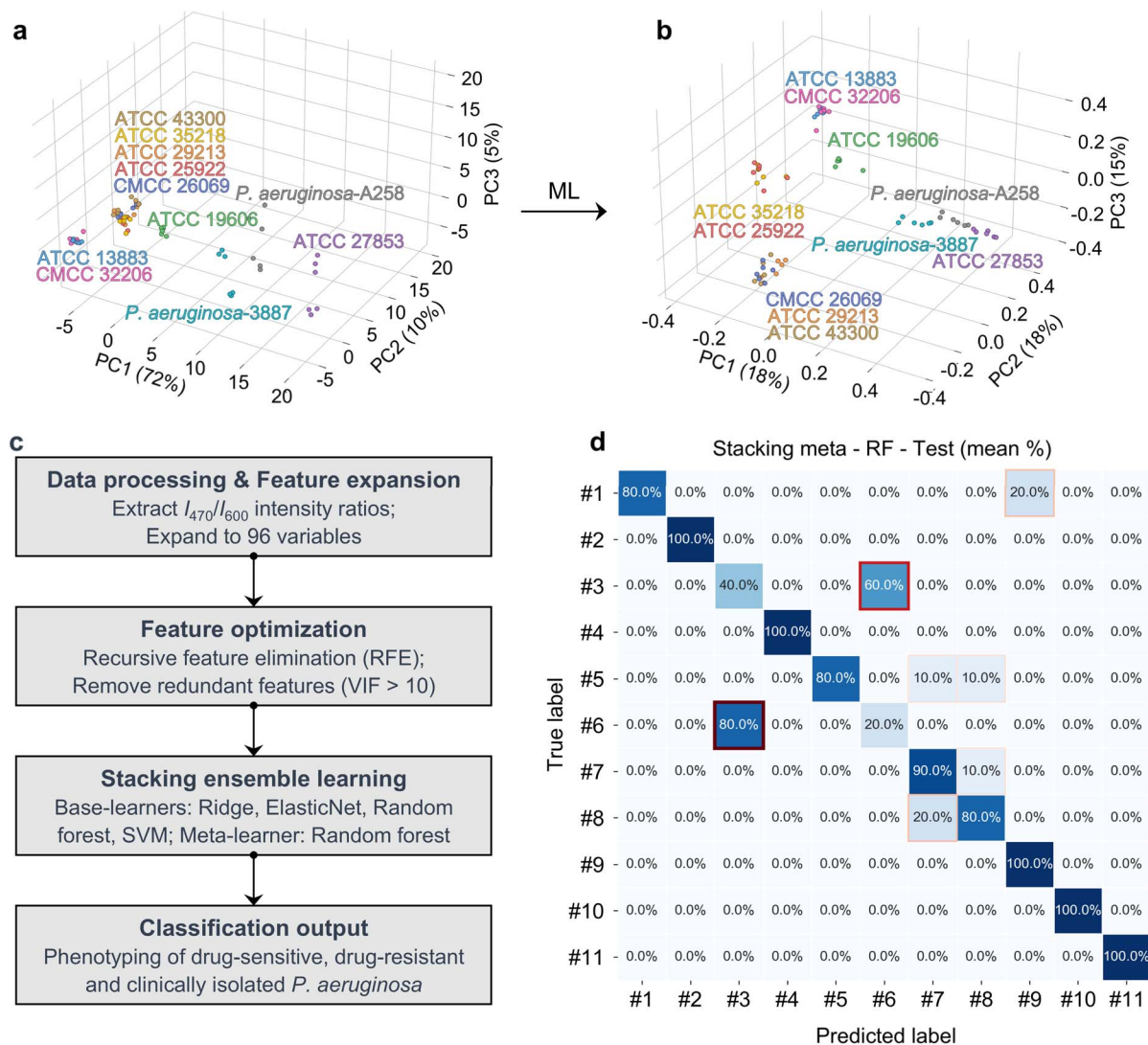
To further corroborate the binding, a fluorescence imaging experiment was carried out using confocal laser-scanning microscopy with Syto-9 co-staining. We found that the galactose-based glycoprobes (DPAC-Gal and DPAC-PEG<sub>3</sub>-Gal) effectively adhered to *P. aeruginosa* (ATCC 27853) as evidenced by the good colocalization of the blue/red dual emission of DPAC with the green fluorescence of Syto-9. Furthermore, these glycoprobes visibly induced the aggregation of the bacteria. In sharp contrast, the control probes (DPAC-OH and DPAC-PEG<sub>3</sub>-OH) exhibited minimal fluorescence on the bacteria and failed to induce bacterial clustering (Fig. S13). This agrees with earlier reports that glycosylated materials trigger bacterial clustering.<sup>37</sup>

To further verify that the fluorescence responses on the bacterial surface were exclusively mediated by specific sugar-lectin interactions, a competitive binding assay was performed. *P. aeruginosa* (ATCC 27853) was pre-incubated with varying concentrations (1, 5, and 10 mM) of free monosaccharides (D-galactose, L-fucose, D-mannose, and D-glucose) for 1 h prior to adding the corresponding glycoprobes (DPAC-Gal, DPAC-Fuc, DPAC-Man, and DPAC-Glc). As depicted in Fig. S14, the fluorescence enhancements in the blue emission channel were significantly attenuated in a concentration-dependent manner, confirming that the free sugars competitively occupied the lectin binding sites on the bacteria.

Similarly, time-dependent ratiometric changes were seen for the glycoprobes with *A. baumannii* (ATCC 19606), which expresses the galactose-selective Ata protein and mannose-selective chaperone-usher pilus. However, the signal variations of the probes for this bacterium appeared to be weaker than those for *P. aeruginosa*. Finally, we found that strains lacking bacterial lectins showed minimal responses over time for the PEG-containing probes, and the probes with the shorter linker exhibited a gradual signal decline (Fig. S15). These data collectively constitute a characteristic “fingerprint” for each bacterial strain.

Next, we used a series of different techniques including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Multidimensional Scaling (MDS) and Factor Analysis (FA) to process the acquired fluorescence-based data for bacterial classification. Both  $I_{470}/I_{600}$  ratios and independent changes in  $I_{470}$  and  $I_{600}$  were processed. The results showed a better classification outcome from analyzing independent fluorescence intensity changes of the glycoprobes (Fig. S16 and S17), probably because of the inclusion of more variables. However, the overall classification result appeared to be





**Fig. 3** (a) Initial 3D PCA score plot for the discrimination of different bacterial strains. (b) Enhanced 3D PCA plot demonstrating improved cluster separation following machine learning (ML) optimization. (c) Schematic illustration of the optimized machine-learning workflow for classification models. (d) Confusion matrix of the stacking ensemble model (meta-learner: random forest (RF)) evaluating the final classification accuracy. Labels #1–#11 correspond to ATCC 13883, ATCC 19606, ATCC 25922, ATCC 27853, ATCC 29213, ATCC 35218, ATCC 43300, CMCC 26069, CMCC 32206, PA 3887, and PA A258, respectively.

suboptimal as different *P. aeruginosa* strains failed to be sufficiently differentiated. This prompted us to resort to AI to better differentiate those strains.

Initially, we observed that the traditional PCA-only workflow resulted in heavily overlapping principal-component clouds, which failed to achieve effective discrimination among the different bacterial strains (Fig. 3a). In contrast, the 3D score plot generated from the stacking-based ensemble learning model demonstrated clear separation among *P. aeruginosa* strains (ATCC 27853, PA A258 and PA 3887), *A. baumannii* ATCC 19606, and the lectin-negative panel (ATCC 13883, ATCC 25922, ATCC 29213, ATCC 35218, ATCC 43300, CMCC 26069 and CMCC 32206) (Fig. 3b). This analysis further enabled differentiation between the drug-sensitive ATCC 27853 strain and the other two clinically isolated *P. aeruginosa* strains (PA A258 and PA 3887).

The confusion matrix derived from this ML processing indicated nearly 100% accuracy, suggesting high model reliability (Fig. 3c).

In the method development phase, we first screened a set of eight different machine-learning (ML) algorithms—namely, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Ridge Regression, Lasso Regression, XGBoost, ElasticNet, LightGBM, and Random Forest—to process the dual fluorescence emission intensities produced by the glycoprobes in the presence of different bacterial strains (Fig. S18). These algorithms, grounded in distinct mathematical principles and intrinsic logic (distance-based, margin-maximising, regularised-linear, ensemble-boosting and bagging), generated a spectrum of classification outcomes for bacterial classification. Subsequently, we applied a three-level data fusion strategy,



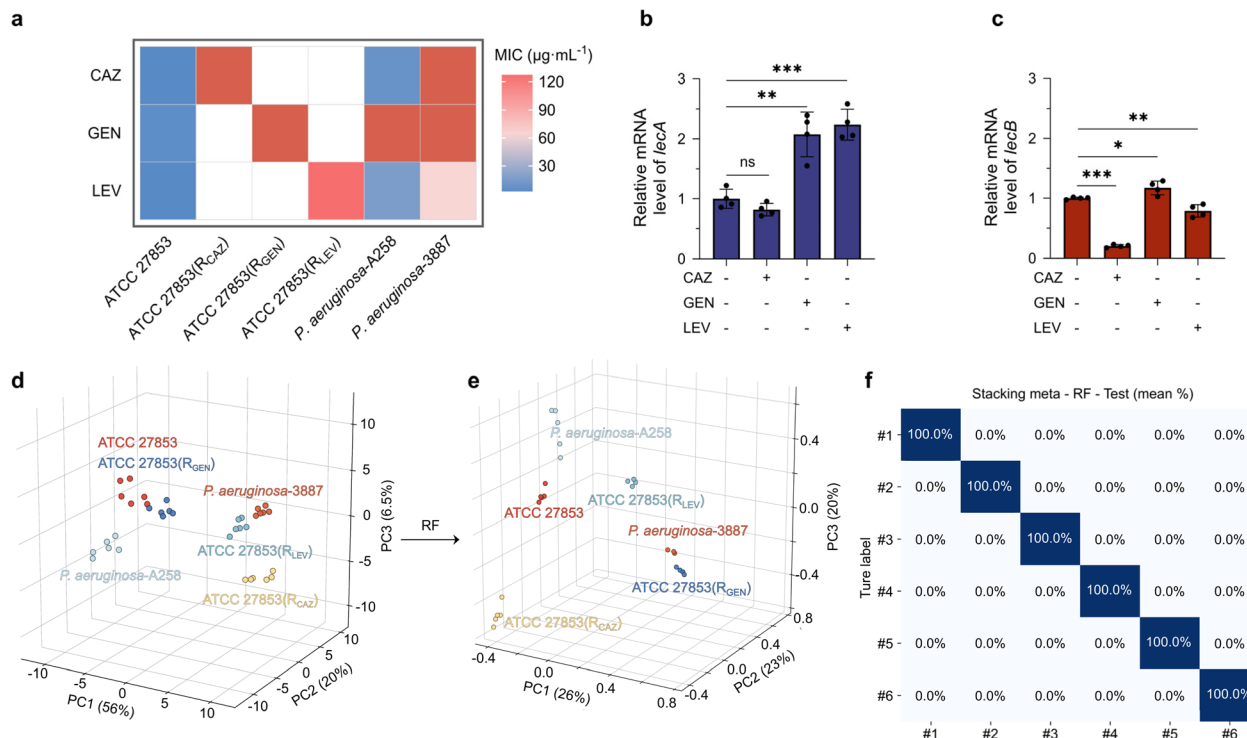


Fig. 4 (a) Drug-resistance profiles of the laboratory-induced mutants and two clinical isolates. Relative mRNA levels of (b) *lecA* and (c) *lecB* across the different resistant strains. (d) Initial 3D PCA plot of the sensitive, induced-resistant, and clinical strains. (e) Enhanced 3D PCA classification performance based on the optimized RF algorithm. (f) Confusion matrix of the RF algorithm classification (Labels #1–#6 correspond to ATCC 27853, ATCC 27853(R<sub>CAZ</sub>), ATCC 27853(R<sub>GEN</sub>), ATCC 27853(R<sub>LEV</sub>), PA 3887, and PA A258, respectively).

which significantly enhanced the performance of the typing model. Through a comprehensive comparison of macro F1 scores and classification accuracy across both training and test sets, we identified a stacking-based ensemble learning method as the optimal approach for differentiating all 11 bacterial strains. This ensemble framework employed Random Forest (RF) as the meta-learner, trained on the probability outputs of four base learners—Ridge, ElasticNet, RF and SVM—each of which was first optimized using feature sets selected *via* Recursive Feature Elimination (RFE) combined with Variance Inflation Factor (VIF) analysis. The overall workflow is summarized in Fig. 3d, and the confusion matrices used for ensemble selection are provided in Fig. S19.

To further validate that the discriminatory power of our sensor array inherently originates from selective monosaccharide-lectin interactions rather than non-specific background signals, the sensing performance of the control probes (DPAC-OH and DPAC-PEG<sub>3</sub>-OH) towards the four lectin-expressing strains (ATCC 27853, PA A258, PA 3887 and ATCC 19606) was evaluated. As expected, these non-targeting control probes exhibited negligible time-dependent fluorescence responses. In addition, when these non-specific signals were processed using the same optimized ensemble learning workflow, the model completely failed to differentiate the bacterial strains, yielding severely overlapping clusters and poor classification accuracies in both training and test sets (Fig. S20). This

experiment suggests that the sugar moieties are indispensable for the AI-assisted phenotyping to be successful.

Given that the method is capable of differentiating drug-sensitive and drug-resistant *P. aeruginosa*, we sought its utility for monitoring the generation of drug resistance of bacteria. For this purpose, we treated ATCC 27853 with Gentamicin (GEN), Levofloxacin (LEV) and Ceftazidime (CAZ) to induce drug resistance. We determined that the resulting variants ATCC 27853(R<sub>CAZ</sub>), ATCC 27853(R<sub>GEN</sub>) and ATCC 27853(R<sub>LEV</sub>) were resistant against the corresponding antibiotics with a minimum inhibitory concentration (MIC) of >128 µg mL<sup>-1</sup>. MIC tests were also done for PA A258 and PA 3887, and the result showed that PA A258 was mainly GEN-resistant and PA 3887 was insensitive to all three antibiotics tested (Fig. 4a). A subsequent measurement by real-time quantitative polymerase chain reaction (RT-qPCR) indicated changes in lectin expression levels in these strains (Fig. 4b and c). While a down-regulation in *LecB* expression was seen for ATCC 27853(R<sub>CAZ</sub>), an upregulated *LecA* expression was consistently found for ATCC 27853(R<sub>GEN</sub>) and ATCC 27853(R<sub>LEV</sub>), with respect to ATCC 27853. These results agree with previous reports revealing the close relationship between antibiotic resistance and lectin expression.<sup>38</sup>

The variations in lectin expression levels under antibiotic stress can be attributed to the bacterium's adaptive evolution strategies. Continuous exposure to antibiotics drives *P. aeruginosa* to enhance biofilm formation as a protective barrier,



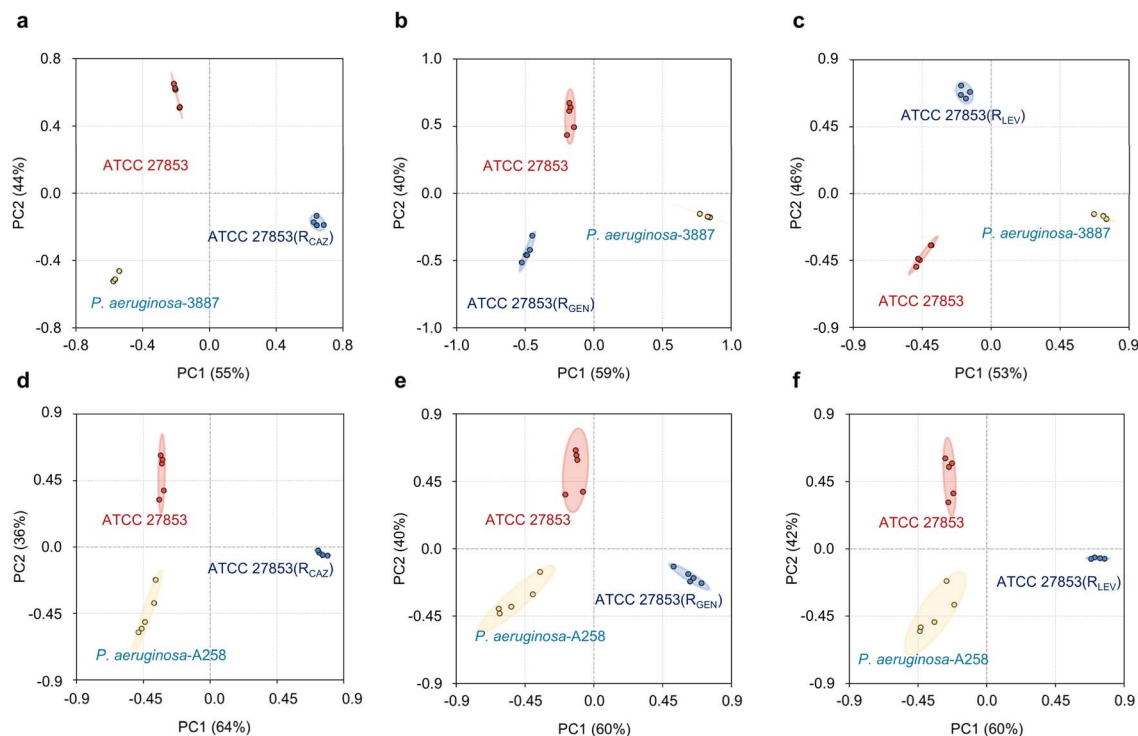


Fig. 5 2D PCA clustering based on optimized RF model analysis for discriminating drug-resistant from drug-sensitive *P. aeruginosa* strains. Panels (a–f) are different training-testing set combinations: (a) ATCC27853, *P. aeruginosa* 3887 and ATCC 27853(R<sub>CAZ</sub>), (b) ATCC27853, PA 3887 and ATCC 27853(R<sub>GEN</sub>), (c) ATCC27853, PA 3887 and ATCC 27853(R<sub>LEV</sub>), (d) ATCC27853, PA A258 and ATCC 27853(R<sub>CAZ</sub>), (e) ATCC27853, PA A258 and ATCC 27853(R<sub>GEN</sub>), and (f) ATCC27853, PA A258 and ATCC 27853(R<sub>LEV</sub>).

leading to the compensatory upregulation of structural Extracellular Matrix (ECM) components like LecA. Furthermore, the prolonged antimicrobial pressure often triggers the rewiring of the global quorum sensing (QS) networks, resulting in profound fluctuations in the expression of downstream effectors such as LecB. Such variations also reflect the metabolic fitness trade-offs as the bacteria prioritize critical resistance mechanisms over non-essential protein synthesis.<sup>39</sup>

We then incubated these strains with the glycoprobes, and obtained their “fingerprints” as shown in Fig. S18. Processing the obtained data with the RF algorithm led to differentiation of drug-resistant strains from drug-sensitive ones. Compared with clustering based on PCA (Fig. 4d), the RF algorithm produced apparently well-separated clusters for the distinct resistant subtypes (Fig. 4e). The corresponding confusion matrix indicated a classification accuracy of 100% (Fig. 4f), suggesting reliability. Generally, each ATCC 27853 variant could be well-differentiated from the drug-sensitive ATCC 27853 as well as a clinical strain, suggesting their different monosaccharide-binding profiles. This differentiation is likely a result of the differential lectin expression of the bacteria, a consequence of divergent resistance mechanisms. The distinct fingerprint patterns between the resistance-induced strains and clinical isolates suggest that the drug-resistance nature of the latter is more complicated. Furthermore, resistance may not only involve changes in surface lectin expression but also other structural and biochemical features, which collectively

contribute to the divergent detection outcomes. Notably, after preprocessing with the RF algorithm, all six strains could be fully discriminated from one another in a two-dimensional PCA projection, further validating the effectiveness of the selected feature set (Fig. 5). To distil the most relevant features from this complex data, we employed RF for feature selection. This allowed us to identify a minimal discriminant feature set, achieving clear differentiation between the ATCC 27853 strain and its resistant variants using as few as 16 features (Fig. S22 and S23). These results suggest our monosaccharide-based sensor array, combined with ML, holds promise for differentiating drug-resistant subpopulations. The requirement of only 16 features makes this approach particularly suitable for the development of inexpensive and easy-to-use analytical kits.

While the current monosaccharide-based fluorescent sensor array, coupled with ensemble learning, demonstrates remarkable accuracy in phenotyping drug-resistant *P. aeruginosa* isolates, we acknowledge that directly analyzing complex clinical samples (e.g., sputum or blood samples) remains a challenge. The intrinsic binding affinity between a single monosaccharide and a bacterial lectin is relatively low. To facilitate future clinical translation, we aim to construct multivalent glycoclusters on biocompatible materials to exponentially amplify the binding affinity through the ‘multivalent effect’. Furthermore, optimizing the AI algorithms to decouple specific lectin-binding signals from background interference in



polymicrobial environments will be a key focus of our future work.

## Conclusions

We have developed a sensor array that consists of a series of eight glycoprobes. These glycoprobes, which use DPAC as the fluorescence reporter, exhibited ratiometric fluorescence changes upon selective lectin binding due to conformational adaptability of the fluorogen. The array also generated characteristic signals for different bacteria, and the resulting signals were used for ML-based analyses, giving rise to bacterial classification. Notably, we demonstrated that bacterial strains that express lectins including *P. aeruginosa* and *A. baumannii* could be better differentiated by this technique. In addition, *P. aeruginosa* variants with induced antibiotic resistance could also be differentiated from the drug-sensitive strain using array sensing, and the differentiation was relevant to the lectin expression of these strains. As such, this study indicates that targeting bacterial lectins is a viable strategy by which to design sensor arrays for bacterial classification. The developed AI-assisted differentiation protocol also offers a complementary tool to nucleic acid-sequencing techniques that require tedious laboratory efforts, facilitating point-of-care decisions for precision antimicrobial therapy against multidrug-resistant infections.

## Author contributions

Z.-F. Zhang performed experiments, analyzed data, and drafted the original manuscript. W.-Z. Gui performed experiments and analyzed data. Y.-F. Tang performed machine learning data analysis and optimized learning methods. H.-Q. Gan analyzed data. X.-L. Hu, T. D. James, S. Yang, Q. Liu and X.-P. He secured funding, supervised the research, and wrote and edited the manuscript. All authors have approved the final version of the manuscript.

## Conflicts of interest

The authors declare no competing financial interest.

## Data availability

The data supporting this article have been included as part of the supplementary information (SI). Supplementary information: the experimental section, synthetic procedures and characterizations (NMR and HRMS) of new compounds, photophysical properties (UV-Vis and fluorescence spectra), and details in machine learning performance metrics. See DOI: <https://doi.org/10.1039/d6sc00084c>.

## Acknowledgements

The authors thank the National Natural Science Foundation of China (NSFC) (No. 92253306, 82130099 and 22477030), Science and Technology Commission of Shanghai Municipality (grant

No. 24DX1400200), the International Cooperation Program of Shanghai Science and Technology (No. 23490711600), the Fundamental Research Funds for the Central Universities (222201717003), the Programme of Introducing Talents of Discipline to Universities (B16017), the National Natural Science Foundation of Shanghai Science and Technology (No.24ZR1415400), the Shanghai Oriental Talents Youth Program (No. QNKJ2024010), the Shanghai Xuhui District Hospital Local Cooperation Project (23XHYD-20), the Open Funding Project of the State Key Laboratory of Fine Chemicals, Dalian University of Technology (KF 2402), State Key Laboratory of Chemo/Biosensing and Chemometrics, Hunan University, Changsha 410082, China, State Key Laboratory of Drug Research (SKLDR-2025-KF-09) for financial support. The Research Center of Analysis and Test of East China University of Science and Technology was gratefully acknowledged for assistance in analytical experiments. TDJ wishes to thank the University of Bath and the Open Research Fund of the School of Chemistry and Chemical Engineering, Henan Normal University (2020ZD01) for support.

## Notes and references

- 1 A. P. Magiorakos, A. Srinivasan, R. B. Carey, Y. Carmeli, M. E. Falagas, C. G. Giske, S. Harbarth, J. F. Hindler, G. Kahlmeter, B. Olsson-Liljequist, D. L. Paterson, L. B. Rice, J. Stelling, M. J. Struelens, A. Vatopoulos, J. T. Weber and D. L. Monnet, Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance, *Clin. Microbiol. Infect.*, 2012, **18**, 268–281.
- 2 E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, C. Pulcini, G. Kahlmeter, J. Kluytmans, Y. Carmeli, M. Ouellette, K. Outtersson, J. Patel, M. Cavaleri, E. M. Cox, C. R. Houchens, M. L. Grayson, P. Hansen, N. Singh, U. Theuretzbacher, N. Magrini and W. P. P. L. Workin, Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis, *Lancet Infect. Dis.*, 2018, **18**, 318–327.
- 3 S. Fujitani, H. Y. Sun, V. L. Yu and J. A. Weingarten, Pneumonia Due to *Pseudomonas aeruginosa* Part I: Epidemiology, Clinical Diagnosis, and Source, *Chest*, 2011, **139**, 909–919.
- 4 M. Bae, Y. Jeong, S. Bae, M. J. Kim, Y. P. Chong, S. H. Kim, S. O. Lee, S. H. Choi, Y. S. Kim and J. Jung, Short versus prolonged courses of antimicrobial therapy for patients with uncomplicated *Pseudomonas aeruginosa* bloodstream infection: a retrospective study, *J. Antimicrob. Chemother.*, 2022, **77**, 223–228.
- 5 D. Church, S. Elsayed, O. Reid, B. Winston and R. Lindsay, Burn wound infections, *Clin. Microbiol. Rev.*, 2006, **19**, 403.
- 6 A. Rodríguez-Rojas, A. Oliver and J. Blázquez, Intrinsic and Environmental Mutagenesis Drive Diversification and Persistence of *Pseudomonas aeruginosa* in Chronic Lung Infections, *J. Infect. Dis.*, 2012, **205**, 121–127.



- 7 E. Drenkard, Antimicrobial resistance of *Pseudomonas aeruginosa* biofilms, *Microb. Infect.*, 2003, **5**, 1213–1219.
- 8 E. B. M. Breidenstein, C. de la Fuente-Núñez and R. E. W. Hancock, *Pseudomonas aeruginosa*: all roads lead to resistance, *Trends Microbiol.*, 2011, **19**, 419–426.
- 9 J. Reyes, L. Komarow, L. Chen, L. Z. Ge, B. M. Hanson, E. Cober, E. Herc, T. Alenazi, K. S. Kaye, J. Garcia-Diaz, L. J. Li, S. S. Kanj, Z. Y. Liu, J. M. Oñate, R. A. Salata, K. Marimuthu, H. Gao, Z. Y. Zong, S. L. Valderrama-Beltrán, Y. S. Yu, P. Tambyah, G. Weston, S. Salcedo, L. M. Abbo, Q. Xie, K. Ordoñez, M. G. Wang, M. E. Stryjewski, J. M. Munita, D. L. Paterson, S. Evans, C. Hill, K. Baum, R. A. Bonomo, B. N. Kreiswirth, M. V. Villegas, R. Patel, C. A. Arias, H. F. Chambers, V. J. R. Fowler, Y. Doi, D. van Duin, M. J. Satlin, A. R. Leadershi and M.-D. R. O. Netw, Global epidemiology and clinical outcomes of carbapenem-resistant *Pseudomonas aeruginosa* and associated carbapenemases (POP): a prospective cohort study, *Lancet Microbe*, 2023, **4**, E159–E170.
- 10 T. S. Crofts, A. J. Gasparrini and G. Dantas, Next-generation approaches to understand and combat the antibiotic resistome, *Nat. Rev. Microbiol.*, 2017, **15**, 422–434.
- 11 M. C. Fisher, A. Alastruey-Izquierdo, J. Berman, T. Bicanic, E. M. Bignell, P. Bowyer, M. Bromley, R. Brüggemann, G. Garber, O. A. Cornely, S. J. Gurr, T. S. Harrison, E. Kuijper, J. Rhodes, D. C. Sheppard, A. Warris, P. L. White, J. P. Xu, B. Zwaan and P. E. Verweij, Tackling the emerging threat of antifungal resistance to human health, *Nat. Rev. Microbiol.*, 2022, **20**, 557–571.
- 12 O. Ari, R. Durmaz and A. Arslanturk, Performance of a molecular beacon-based qPCR method for detection of fluoroquinolone resistance in *Mycobacterium tuberculosis*: A study confirmed by sequencing, *Lancet Microbe*, 2025, **112**, 116865.
- 13 T. Schön, P. Miotto, C. U. Köser, M. Viveiros, E. Böttger and E. Cambau, drug-resistance testing: challenges, recent developments and perspectives, *Clin. Microbiol. Infect.*, 2017, **23**, 154–160.
- 14 M. M. Chen, Y. Q. Zhang, L. C. Cheng, F. J. Zhao and P. Wang, Machine learning-powered fluorescent sensor arrays for rapid detection of heavy metals and pesticides in complex environments, *Biosens. Bioelectron.*, 2025, **287**.
- 15 W. X. Zhang, S. J. You, J. Sun, Y. J. Lv, X. L. Wang, X. Li and Z. M. Su, Multi-emission fluorescent array sensors based on metal-organic frameworks: recent research progress in synthesis strategies and applications, *Coord. Chem. Rev.*, 2025, **542**.
- 16 Y. Yu, W. W. Ni, Q. Hu, H. H. Li, Y. Zhang, X. Gao, L. J. Zhou, S. M. Zhang, S. Y. Ma, Y. L. Zhang, H. Huang, F. Li and J. S. Han, A Dual Fluorescence Turn-On Sensor Array Formed by Poly(aryleneethynylene) and Aggregation-Induced Emission Fluorophores for Sensitive Multiplexed Bacterial Recognition, *Angew. Chem., Int. Ed.*, 2024, **63**, e202318483.
- 17 D. Svehkarev, M. R. Sadykov, K. W. Bayles and A. M. Mohs, Ratiometric Fluorescent Sensor Array as a Versatile Tool for Bacterial Pathogen Identification and Analysis, *ACS Sens.*, 2018, **3**, 700–708.
- 18 B. Y. Zhang, H. Chen, L. Shi, R. R. Guo, Y. Wang, Y. H. Zheng, R. Y. Bai, Y. X. Gao, B. Liu and X. F. Zhang, Nitroreductase-Based "Turn-On" Fluorescent Probe for Bacterial Identification with Visible Features, *ACS Sens.*, 2024, **9**, 4560–4567.
- 19 X.-L. Hu, H.-Q. Gan, Z.-Y. Qin, Q. Liu, M. Li, D. Chen, J. L. Sessler, H. Tian and X.-P. He, Phenotyping of Methicillin-Resistant *Staphylococcus aureus* Using a Ratiometric Sensor Array, *J. Am. Chem. Soc.*, 2023, **145**, 8917–8926.
- 20 R. L. Phillips, O. R. Miranda, C. C. You, V. M. Rotello and U. H. Bunz, Rapid and efficient identification of bacteria using gold-nanoparticle-poly(para-phenyleneethynylene) constructs, *Angew. Chem., Int. Ed.*, 2008, **47**, 2590–2594.
- 21 F. B. Wang, M. H. Xiao, J. Qi and L. Zhu, Paper-based fluorescence sensor array with functionalized carbon quantum dots for bacterial discrimination using a machine learning algorithm, *Anal. Bioanal. Chem.*, 2024, **416**, 3139–3148.
- 22 E. P. Mitchell, C. Sabin, L. Snajdrová, M. Pokorná, S. Perret, C. Gautier, C. Hofr, N. Gilboa-Garber, J. Koca, M. Wimmerová and A. Imberty, High Affinity Fucose Binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 Å Resolution Crystal Structure of the Complex Combined with Thermodynamics and Computational Chemistry Approaches, *Proteins: Struct., Funct., Bioinf.*, 2005, **58**, 735–746.
- 23 J. L. Reymond and T. Darbre, Peptide and glycopeptide dendrimer apple trees as enzyme models and for biomedical applications, *Org. Biomol. Chem.*, 2012, **10**, 1483–1492.
- 24 J. L. Reymond, M. Bergmann and T. Darbre, Glycopeptide dendrimers as *Pseudomonas aeruginosa* biofilm inhibitors, *Chem. Soc. Rev.*, 2013, **42**, 4814–4822.
- 25 T. Rasamiravaka, Q. Labtani, P. Duez and M. El Jaziri, The Formation of Biofilms by *Pseudomonas aeruginosa*: A Review of the Natural and Synthetic Compounds Interfering with Control Mechanisms, *BioMed Res. Int.*, 2015, **2015**, 759348.
- 26 E. Siebs, E. Shanina, S. Kuhaudomlarp, P. D. F. C. Gomes, C. Fortin, P. H. Seeberger, D. Rognan, C. Rademacher, A. Imberty and A. Titz, Targeting the Central Pocket of the *Pseudomonas aeruginosa* Lectin LecA, *ChemBioChem*, 2022, **23**, e202100563.
- 27 S. Kirkeby, A. K. Hansen, A. d'Apice and D. Moe, The galactophilic lectin (PA-IL, gene LecA) from *Pseudomonas aeruginosa*. Its binding requirements and the localization of lectin receptors in various mouse tissues, *Microb. Pathog.*, 2006, **40**, 191–197.
- 28 Z. Zhang, Y.-S. Wu, K.-C. Tang, C.-L. Chen, J.-W. Ho, J. Su, H. Tian and P.-T. Chou, Excited-State Conformational/Electronic Responses of Saddle-Shaped N,N'-Disubstituted-Dihydrodibenzo[a,c]phenazines: Wide-Tuning Emission from Red to Deep Blue and White Light Combination, *J. Am. Chem. Soc.*, 2015, **137**, 8509–8520.



- 29 W. Huang, L. Sun, Z. W. Zheng, J. H. Su and H. Tian, Colour-tunable fluorescence of single molecules based on the vibration induced emission of phenazine, *Chem. Commun.*, 2015, **51**, 4462–4464.
- 30 X. Y. Chen, S. D. Yuan, M. Y. Qiao, X. Jin, J. C. Chen, L. F. Guo, J. H. Su, D. H. Qu and Z. Y. Zhang, Exploring the Depth-Dependent Microviscosity inside a Micelle Using Butterfly-Motion-Based Fluorescent Probes, *J. Am. Chem. Soc.*, 2023, **145**, 26494–26503.
- 31 W. T. Dou, X. Wang, T. T. Liu, S. W. Zhao, J. J. Liu, Y. Yan, J. Li, C. Y. Zhang, A. C. Sedgwick, H. Tian, J. L. Sessler, D. M. Zhou and X. P. He, A homogeneous high-throughput array for the detection and discrimination of influenza A viruses, *Chem*, 2022, **8**, 1750–1761.
- 32 A. T. Wright, M. J. Griffin, Z. Zhong, S. C. McCleskey, E. V. Anslyn and J. T. McDevitt, Differential Receptors Create Patterns That Distinguish Various Proteins, *Angew. Chem., Int. Ed.*, 2005, **44**, 6375–6378.
- 33 S. Stewart, M. A. Ivy and E. V. Anslyn, The use of principal component analysis and discriminant analysis in differential sensing routines, *Chem. Soc. Rev.*, 2014, **43**, 70–84.
- 34 A. M. Mallet, Y. L. Liu and M. Bonizzoni, An off-the-shelf sensing system for physiological phosphates, *Chem. Commun.*, 2014, **50**, 5003–5006.
- 35 M. Weidensdorfer, M. Ishikawa, K. Hori, D. Linke, B. Djahanschiri, R. Iruegas, I. Ebersberger, S. Riedel-Christ, G. Enders, L. Leukert, P. Kraicz, F. Rothweiler, J. Cinatl, J. Berger, K. Hipp, V. A. J. Kempf and S. Göttig, The Acinetobacter trimeric autotransporter adhesin Ata controls key virulence traits of *Acinetobacter baumannii*, *Virulence*, 2019, **10**, 68–81.
- 36 A. Imberty, M. Wimmerová, E. P. Mitchell and N. Gilboa-Garber, Structures of the lectins from *Pseudomonas aeruginosa*: insights into the molecular basis for host glycan recognition, *Microb. Infect.*, 2004, **6**, 221–228.
- 37 K. Wojtczak and J. P. Byrne, Structural Considerations for Building Synthetic Glycoconjugates as Inhibitors for *Pseudomonas aeruginosa* Lectins, *ChemMedChem*, 2022, **17**, e202200081.
- 38 S. G. Pereira, A. C. Rosa and O. Cardoso, Virulence factors as predictive tools for drug resistance in *Pseudomonas aeruginosa*, *Virulence*, 2015, **6**, 679–683.
- 39 M. Letizia, S. P. Diggle and M. Whiteley, *Pseudomonas aeruginosa*: ecology, evolution, pathogenesis and antimicrobial susceptibility, *Nat. Rev. Microbiol.*, 2025, **23**, 701–717.

