Deep learning-guided discovery of an antibiotic targeting *Acinetobacter baumannii*

Gary Liu[1,10], Denise B. Catacutan[1,10], Khushi Rathod[1,10], Kyle Swanson [2], Wengong Jin[2], Jody C. Mohammed[1], Anush Chiappino-Pepe [3,4], Saad A. Syed[5], Meghan Fragis [1,6], Kenneth Rachwalski [1], Jakob Magolan [1,6], Michael G. Surette[5], Brian K. Coombes[1], Tommi Jaakkola [2], Regina Barzilay[2,7], James J. Collins [3,7,8,9] & Jonathan M. Stokes [1]

Acinetobacter baumannii is a nosocomial Gram-negative pathogen that often displays multidrug resistance. Discovering new antibiotics against *A. baumannii* has proven challenging through conventional screening approaches. Fortunately, machine learning methods allow for the rapid exploration of chemical space, increasing the probability of discovering new antibacterial molecules. Here we screened ~7,500 molecules for those that inhibited the growth of *A. baumannii* in vitro. We trained a neural network with this growth inhibition dataset and performed in silico predictions for structurally new molecules with activity against *A. baumannii*. Through this approach, we discovered abaucin, an antibacterial compound with narrow-spectrum activity against *A. baumannii*. Further investigations revealed that abaucin perturbs lipoprotein trafficking through a mechanism involving LolE. Moreover, abaucin could control an *A. baumannii* infection in a mouse wound model. This work highlights the utility of machine learning in antibiotic discovery and describes a promising lead with targeted activity against a challenging Gram-negative pathogen.

Acinetobacter baumannii is a nosocomial Gram-negative pathogen that often displays multidrug resistance due to its robust outer membrane and its ability to acquire and retain extracellular DNA[1] that frequently encodes antibiotic resistance genes. Moreover, it can survive for prolonged durations on surfaces and is resistant to desiccation. Discovering fundamentally new antibiotics against *A. baumannii* has proven challenging through conventional screening approaches. Indeed, most new antibiotics that achieve clinical use are analogs of existing classes[2]. However, while structural analogs of existing antibiotic classes may satisfy short-term clinical needs, their long-term efficacy is inherently

limited due to the high prevalence of existing resistance determinants[3]. Ideally, new antibiotic discovery efforts should focus on identifying new chemotypes with mechanisms of action that are unique relative to existing antibiotics. Such compounds are likely to have prolonged utility, given that the probability of pre-existing clinical resistance is low.

Fortunately, machine learning methods allow for the rapid exploration of vast chemical/sequence spaces in silico, increasing the probability of discovering desirable new chemotypes with antibacterial activity, particularly against challenging pathogens like *A. baumannii*. As a reference, typical high-throughput screening programs are limited to testing a few million molecules for antibacterial activity at the largest scales[4]. Contrarily, contemporary algorithmic approaches can assess hundreds of millions to billions of molecules for antibacterial properties. For example, Stokes et al.[5] applied a message-passing neural network trained on growth inhibition of lab strain *Escherichia coli* to discover new broad-spectrum small molecule antibacterial compounds. In a complementary application, Ma et al.[6] applied multiple natural language processing neural network models to predict broad-spectrum antimicrobial peptides encoded in the human gut microbiome. These, and other important studies (detailed in ref. [7]) showcase the importance of machine learning approaches toward the discovery of structurally and functionally new antibiotic candidates.

Beyond simply discovering structurally and functionally new antibiotics, a largely unmet need exists for the application of narrow-spectrum therapies that target specific bacterial species. Such antibiotics are beneficial for the following two reasons[8]: first, the rate at which resistance to narrow-spectrum agents would disseminate is likely lower than conventional broad-spectrum agents, because narrow-spectrum drugs do not impose a universal selective pressure that favors the wide propagation of resistance determinants; second, narrow-spectrum antibiotics would not disrupt the ecology of the microbiota during treatment. Indeed, dysbiosis has been associated with a wide array of poor health outcomes, including infectious diseases[9], inflammatory bowel diseases[10], metabolic diseases[11], neuropsychiatric disorders[12] and cancer[13]. For instance, *Clostridioides difficile* infections are prime examples of opportunistic infections resulting from antibiotic-induced dysbiosis, causing upwards of 224,000 infections in hospitalized patients and 13,000 deaths in the US alone in 2017 (ref. [14]).

An opportunity exists to apply contemporary machine learning methods to discover structurally and functionally new antibiotics that specifically target challenging pathogens, with *A. baumannii* being a prime candidate. Here we screened ~7,500 molecules for those that inhibited the growth of *A. baumannii* in vitro. We trained a message-passing neural network with this growth inhibition dataset and performed predictions on the Drug Repurposing Hub[15] for structurally new molecules with activity against *A. baumannii*. Through this approach, we discovered abaucin, an antibacterial compound with narrow-spectrum activity against *A. baumannii*, which could overcome intrinsic and acquired resistance mechanisms in clinical isolates. Further mechanistic investigations revealed that abaucin perturbs lipoprotein trafficking through a mechanism involving LolE, a functionally conserved protein that contributes to shuttling lipoproteins from the inner membrane to the outer membrane[16,17]. Moreover, abaucin was able to control an *A. baumannii* infection in a mouse wound model. This study highlights the utility of machine learning in discovering new antibiotics and describes a promising lead molecule with specific activity against a challenging Gram-negative pathogen.

Results

Machine learning-guided discovery of abaucin

Our recent work has highlighted the utility of machine learning in discovering new antibacterial molecules using *E. coli* K12 as a model organism[5]. Building off this prior research, here we applied a message-passing deep neural network[18] to discover new antibiotics against *A. baumannii*,

a problematic nosocomial Gram-negative pathogen that commonly displays multidrug resistance and, increasingly, pan-drug resistance (Fig. 1a). We first screened a diverse collection of 7,684 small molecules at 50 μM for those that inhibited the growth of *A. baumannii* ATCC 17978 in Lysogeny Broth (LB) medium (Fig. 1b and Extended Data Fig. 1a). This chemical collection consisted of both off-patent drugs (2,341 molecules) and synthetic chemicals (5,343 molecules) curated from various high-throughput screening sub-libraries at the Broad Institute. Using a conventional hit cutoff of one standard deviation below the mean growth of the entire dataset resulted in 480 molecules being defined as 'active' and 7,204 being defined as 'inactive' (Supplementary Data 1).

Next, this dataset was used to train a binary classifier to predict whether structurally new molecules may display activity against *A. baumannii*. Briefly, we leveraged a directed message-passing neural network architecture, which translates the graph structure of a molecule into a continuous vector[18] (Fig. 1a).

This type of model operates by iteratively exchanging information of local chemistry between adjacent atoms and bonds in a series of 'message-passing' steps. Each iteration of message passing propagates information about local chemistry across the molecule, thereby allowing the model to build a more holistic representation of the molecule. After a defined number of message-passing steps, the vector representations of various local chemical regions of a molecule are summed into a single continuous vector that captures the complexity of the entire compound. This learned final vector is then supplemented with fixed molecular features computed using RDKit[19]. A final vector containing both learned and computed features is then used as an input vector for a feed-forward neural network that predicts antibacterial properties. The model was further optimized by using an ensemble of ten classifiers, increasing its robustness. Our final model achieved an area under the precision-recall curve of 0.337 ± 0.088 and an area under the receiver-operating characteristic curve of 0.792 ± 0.042, providing confidence in leveraging the model for predictions in new chemical spaces.

After model training, we applied our ensemble of ten RDKit-augmented models to identify antibacterial molecules with activity against *A. baumannii* from the updated Drug Repurposing Hub[15] consisting of 6,680 molecules. This chemical library was selected as a proof-of-concept due to the structural diversity inherent to this collection, as well as the favorable cytotoxicity and drug-like properties that are observed for many Drug Repurposing Hub molecules. Our ensemble of ten classifiers returned a prediction score for each compound, representing the probability of growth inhibition against *A. baumannii* (Fig. 1c and Supplementary Data 2). Molecules that were both strongly predicted to be antibacterial and structurally unique from training set 'actives' were prioritized for in vitro testing (Supplementary Data 3). Notably, this process of performing predictions and prioritizing molecules for validation was completed within a couple of hours. We note here that we also trained and applied a message-passing neural network model that was not augmented with RDKit features as a baseline reference (Supplementary Data 2). This model did not perform as well as the RDKit-augmented primary model, achieving an area under the precision-recall curve of 0.266 ± 0.070 and an area under the receiver-operating characteristic curve of 0.756 ± 0.050. The reduced performance of the model omitting RDKit features highlights the importance of these computable molecular features in maximizing predictive utility in the context of our training dataset.

Using a prediction score threshold of >0.2 (Fig. 1c) and subsequently filtering these predicted compounds based on a Tanimoto nearest neighbor similarity of <0.3 to molecules that were 'active' in the training dataset, we identified 240 priority molecules that met these highly stringent criteria (Supplementary Data 3). These molecules were acquired and tested against *A. baumannii* at a concentration of 50 μM in LB medium—the same conditions in which the training data were acquired (Extended Data Fig. 1b and Supplementary Data 3). Using
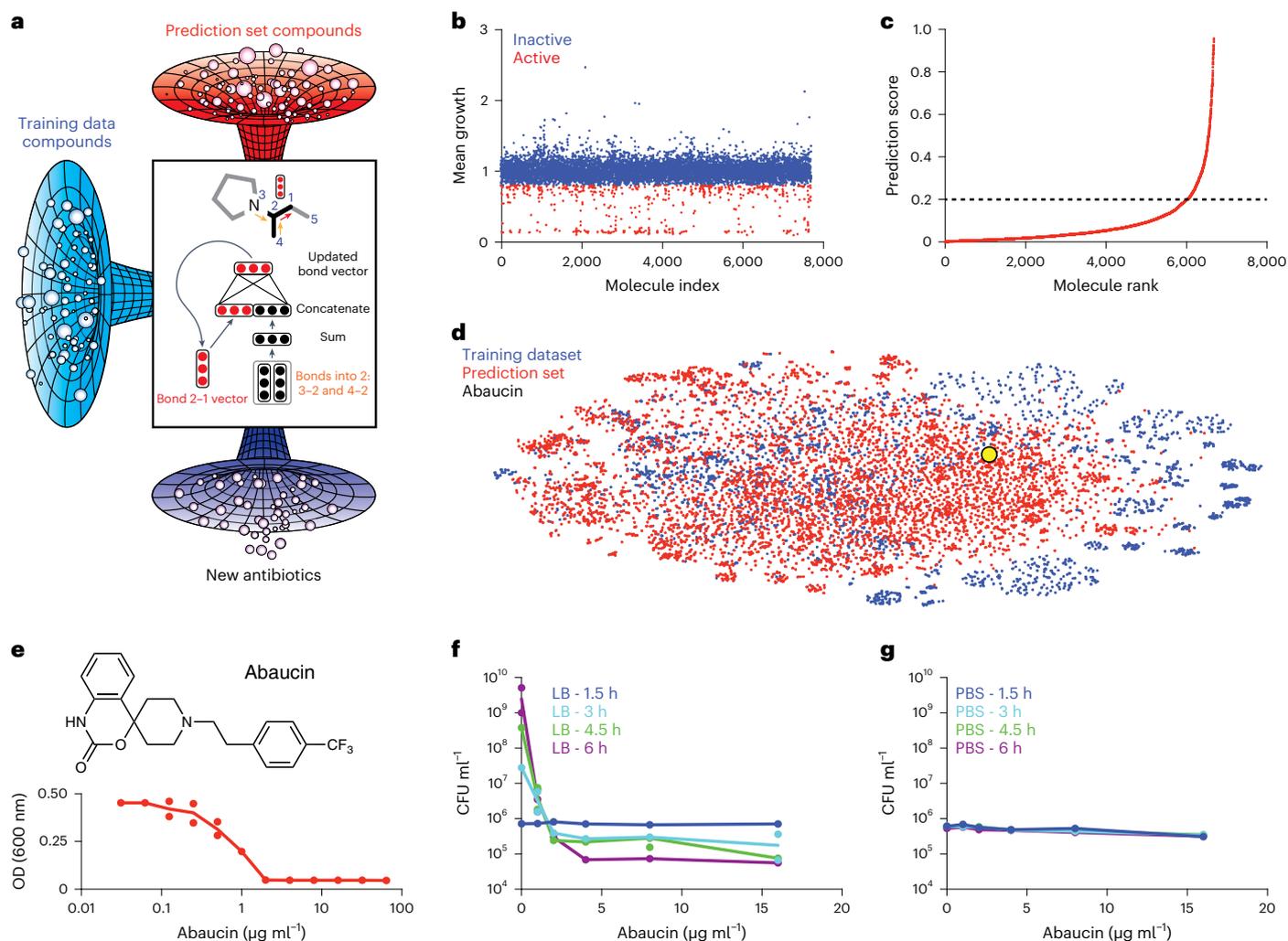
Fig. 1 | Machine learning-guided discovery of abaucin. a, Machine learning approaches allow for the rapid exploration of chemical space, increasing the probability of discovering new chemical matter with antibacterial activity. We screened ~7,500 molecules for those that inhibited the growth of *A. baumannii* (blue), trained a directed message-passing deep neural network with this growth inhibition dataset and performed predictions (red) on the Drug Repurposing Hub for structurally new molecules with activity against *A. baumannii* (purple). **b**, Growth inhibition of *A. baumannii* ATCC 17978 by a collection of 7,684 small molecules at 50 μM. The mean of two biological replicates is shown. Red dots represent actives, with the hit cutoff defined as one standard deviation below the mean of the dataset. **c**, Rank-ordered prediction scores of molecules within the Drug Repurposing Hub by our trained model. Molecules with prediction scores > 0.2 were considered preliminary candidates for experimental validation.

d, A t-SNE plot showing the chemical relationship between the training dataset (blue), the prediction set (red) and abaucin (yellow). **e**, Growth inhibition of *A. baumannii* by abaucin in LB medium. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. The structure of abaucin is shown. **f**, Killing of *A. baumannii* by abaucin in nutrient-replete conditions (LB) at varying concentrations after incubation for 1.5 h (blue), 3 h (teal), 4.5 h (green) and 6 h (purple). Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. **g**, Killing of *A. baumannii* by abaucin in nutrient-deplete conditions (PBS) at varying concentrations after incubation for 1.5 h (blue), 3 h (teal), 4.5 h (green) and 6 h (purple). Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted.

another stringent cutoff of >80% growth inhibition, we observed that nine of the tested molecules displayed antibacterial activity against *A. baumannii*, emphasizing the ability of our model to generalize to compounds that are highly divergent from the chemical space on which the model was trained. We note here that >80% growth inhibition is a statistically more stringent cutoff than that used for model training, which was a conventional one standard deviation below the mean growth of the dataset, or ~20% growth inhibition. This stringent >80% growth inhibition cutoff was applied to efficiently prioritize the most potent predicted molecules on which to conduct further experimentation. For reference, if we defined a hit cutoff of 20% growth inhibition for prediction selection—similar to that used for model training—we would acquire 41 molecules classified as active predictions (Supplementary Data 3). Notably, we also tested the 240 molecules

with the lowest prediction scores and observed that none displayed antibacterial activity as defined by >80% growth inhibition, emphasizing the discriminatory utility of our model (Extended Data Fig. 1b and Supplementary Data 3). Furthermore, testing the 240 molecules with the highest prediction scores, without considering Tanimoto nearest neighbor similarity to training set 'actives', resulted in 40 molecules that passed the stringent >80% growth inhibition cutoff, indicating that our model has a high predictive value over traditional chemical screening (Extended Data Fig. 1b and Supplementary Data 3).

The nine priority molecules were subsequently assessed to remove (1) those with major structural features that are observed in known antibiotics; (2) those with reported antibacterial activity from the scientific or patent literature; and (3) those with possible nonspecific membrane activity as assessed by the presence of acyclic aliphatic
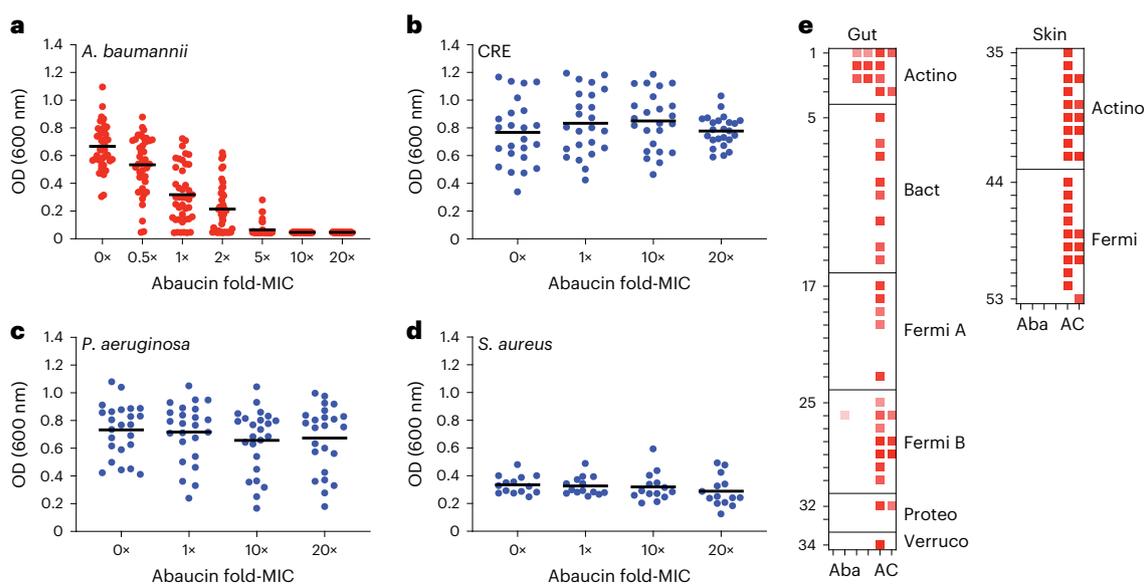
Fig. 2 | Abaucin has a narrow phylogenetic spectrum of antibacterial activity. **a**, Growth inhibition of 41 multidrug-resistant clinical *A. baumannii* isolates by abaucin. Drug concentrations are shown as fold-MIC based on the MIC of abaucin against *A. baumannii* ATCC 17978. Each dot represents a different clinical strain (mean OD of two biologically independent replicates). Black bars show the mean growth inhibition of the entire panel at each abaucin concentration. **b**, Same as **a**, except using 24 carbapenem-resistant Enterobacteriaceae clinical isolates. **c**, Same as **a**, except using 24 multidrug-resistant *P. aeruginosa* clinical isolates. **d**, Same as **a**, except using 14 *S. aureus* clinical isolates. **e**, Heat map showing mean growth inhibition of a panel of human gut commensal species (left) and human skin commensal species (right) by abaucin at 0×, 1×, 10× and 20× MIC (left to right), ampicillin (A) at 128 µg ml⁻¹, and ciprofloxacin (C) at 2 µg ml⁻¹. Darker red indicates more growth inhibition as analyzed by OD. Strains used are described in Supplementary Tables 5 and 6. Experiments were conducted in biological duplicate. Note that abaucin displays minimal activity against commensal species, whereas ampicillin and ciprofloxacin have broad-spectrum activity against commensals. Actino, Actinobacteriota; Bact, Bacteroidota; Fermi, Fermicutes; Proteo, Proteobacteria; Verruco, Verrucomicrobiota.

moieties. This structure-based filtering resulted in the retention of two molecules—RS102895 and serdemetan. RS102895 is a well-studied CCR2⁻ selective chemokine receptor antagonist[20,21] that displayed a minimum inhibitory concentration (MIC) of ~2 µg ml⁻¹ against *A. baumannii* ATCC 17978 (Fig. 1d,e); serdemetan[22,23] is an antagonist of the transcription factor HDM2 that displayed MIC of ~32 µg ml⁻¹ (Extended Data Fig. 1c). Given that RS102895 was substantially more potent at inhibiting the growth of *A. baumannii*, we focused our subsequent investigations on this molecule.

Upon further experimentation with RS102895 to assess *A. baumannii* viability after treatment, we observed modest bactericidal activity against *A. baumannii* in LB medium (Fig. 1f). Indeed, upon removal of RS102895 from *A. baumannii* cultures in vitro after 6 h of treatment, we observed *A. baumannii* regrowth, wherein the apparent lag period increased with increasing concentrations of abaucin (Extended Data Fig. 1d). We observed no discernable activity in nutrient-deplete PBS (Fig. 1g). Collectively, these data suggest that RS102895—renamed abaucin for its activity against *A. baumannii*—displayed its antibacterial efficacy through inhibition of a biological process that was maximally active during growth and division[24], consistent with most known antibiotics[25]. Notably, these data also suggest that abaucin is not membrane active via physical disruption of the phospholipid bilayer, a mechanism of action that we deliberately attempted to avoid during our prediction filtering process. Indeed, membrane-active molecules generally retain bactericidal efficacy in nutrient-deplete conditions[24].
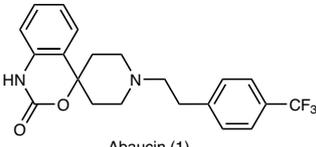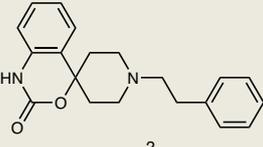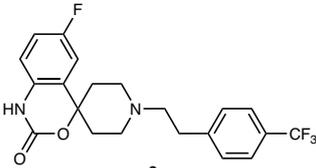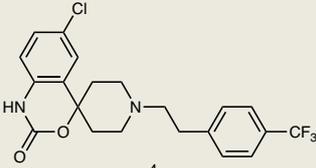
Abaucin has a narrow spectrum of activity

After probing the activity of abaucin against *A. baumannii* ATCC 17978, we next tested for growth inhibitory activity against clinical isolates of *A. baumannii*. Here we acquired 41 strains of *A. baumannii* from the Center for Disease Control and Prevention Antibiotic Resistance Isolate Bank (ARIsolate Bank; Supplementary Table 1) and tested abaucin at a range of concentrations below and above MIC. Remarkably, we

observed that abaucin could overcome all intrinsic and acquired resistance mechanisms within the *A. baumannii* isolates from this diverse clinical strain library (Fig. 2a and Supplementary Table 1).

Next, we investigated the phylogenetic spectrum of activity displayed by abaucin by testing this compound against 24 carbapenem-resistant Enterobacteriaceae strains (Supplementary Table 2), 24 *Pseudomonas aeruginosa* strains (Supplementary Table 3) and 14 *Staphylococcus aureus* strains (Supplementary Table 4), all from the ARIsolate Bank. In stark contrast to our observations with *A. baumannii*, abaucin did not display any growth inhibitory activity against these pathogenic species up to 20× the MIC in *A. baumannii* ATCC 17978 (Fig. 2b–d). Excitingly, in the context of clinical bacterial pathogens, these data provide strong evidence that abaucin has narrow-spectrum antibacterial activity, which is advantageous for decreasing the inter-pathogen dissemination of resistance. We note here that a modest collection of structural analogs of abaucin displayed varying levels of activity against *A. baumannii* but avoided any discernable activity against lab strains of the bacterial species mentioned above (Table 1). This initial structure–activity relationship investigation provides strong support that exploration of the chemical space around abaucin toward developing a medicinal chemistry-optimized narrow-spectrum antibiotic against *A. baumannii* is feasible without being hampered by broad-spectrum antibacterial activity.

Given these observations, we hypothesized that abaucin would display minimal growth inhibitory activity against human commensal species. Indeed, currently employed antibiotics often induce dysbiosis during treatment[26], resulting in a wide array of complications, including secondary infections caused by opportunistic pathogens—*C. difficile* gut infections being a common example[27]. To test this hypothesis, we curated panels of 34 diverse human gut commensal isolates (Supplementary Table 5) and 19 diverse human skin commensal isolates (Supplementary Table 6), then tested the growth inhibitory properties of abaucin at varying concentrations, as well as ampicillin and

Table 1 | MICs of abaucin analogs against a panel of bacterial species

	E. coli	P. aeruginosa	S. aureus	A. baumannii
Abaucin (1)	>128 µg ml⁻¹	>128 µg ml⁻¹	>128 µg ml⁻¹	2 µg ml⁻¹
2	>128 µg ml⁻¹	>128 µg ml⁻¹	>128 µg ml⁻¹	8 µg ml⁻¹
3	>128 µg ml⁻¹	>128 µg ml⁻¹	>128 µg ml⁻¹	64 µg ml⁻¹
4	>128 µg ml⁻¹	>128 µg ml⁻¹	>128 µg ml⁻¹	128 µg ml⁻¹

Growth inhibition of bacterial species by abaucin and analogs thereof. Cells were grown in LB in the presence of varying concentrations of each molecule at 37 °C and the MIC of each was determined. The strains shown are *E. coli* BW25113, *P. aeruginosa* PAO1, *S. aureus* RN4220 and *A. baumannii* ATCC 17978. Experiments were conducted in biological duplicate and resulted in identical MIC values.

ciprofloxacin at their respective MIC concentrations in *A. baumannii* ATCC 17978 (Extended Data Fig. 2a). As expected, ampicillin and ciprofloxacin displayed antibacterial activity across a wide range of commensal isolates, whereas abaucin largely avoided growth inhibition of commensal species, even up to 20× MIC (Fig. 2e). Indeed, of the 53 isolates tested, abaucin only displayed *bona fide* growth inhibition against the gut isolates *Bifidobacterium breve* and *Bifidobacterium longum*, and this occurred above the MIC observed in *A. baumannii* ATCC 17978 (Extended Data Fig. 2b–d). Given that *Bifidobacterium* is a Gram-positive genus that is phylogenetically divergent from Gram-negative *Acinetobacter*, it is likely that the lower potency activity of abaucin against *Bifidobacterium* is through a mechanism that is unrelated to that in *Acinetobacter*. This statement is elaborated below.

Abaucin inhibits lipoprotein trafficking in *A. baumannii*

Given the *A. baumannii* selectivity displayed by abaucin, we next ventured to elucidate the mechanism underlying its narrow-spectrum functionality. To this end, we first selected abaucin-resistant mutants using wild-type *A. baumannii* growing on solid media supplemented with varying concentrations of abaucin. Using 4 µg ml⁻¹ and 5 µg ml⁻¹, we were able to isolate abaucin-resistant clones that did not display cross-resistance to functionally diverse antibiotics (Fig. 3a and Extended Data Fig. 3a–h). Whole-genome sequencing of four independent isolates revealed mutations in or upstream of the gene encoding LolE, an essential inner membrane protein involved in lipoprotein trafficking that has become a target of strong interest for new

Gram-negative antibiotic development[16,17,28]. Two mutants (Y394F and an upstream G to A mutation) displayed fourfold resistance to abaucin relative to wild-type *A. baumannii*, and two independent mutants contained an identical A362T variant that resulted in 16-fold resistance. We note that the frequency of resistance to abaucin in vitro is 10^{-8}–10^{-7} (Supplementary Table 7), largely consistent with known antibiotics that target a single protein[29]. Interestingly, *A. baumannii* LolE position A362 is homologous to *E. coli* LolE position I365, which resides near the acyl chains of the nascent lipoprotein during transport[17]. We predicted the structure of *A. baumannii* LolE using RoseTTAFold[30,31] (Extended Data Fig. 3i,j) and observed that *Ab*LolE position A362 is near *Ec*LolE position I365 in space, and both are adjacent to the accommodated acyl chains of the nascent lipoprotein (Fig. 3b).

These mutational data, combined with in silico structural insights, provided evidence suggesting that abaucin may disrupt lipoprotein accommodation and transport facilitated by LolE. To further explore the hypothesis that abaucin was interfering with lipoprotein trafficking, we treated wild-type *A. baumannii* with 5× MIC of abaucin, or no compound, for varying durations and subjected cells to RNA sequencing. A transcriptomics approach would afford systems-level insight that could provide additional indirect evidence to support or negate our initial mechanistic hypothesis. After sequencing, we performed differential expression analyses between the no-drug control cultures and the abaucin-treated cultures to quantify up- and down-regulated transcripts. Differentially abundant transcripts were then clustered based on Gene Ontology (GO) term to identify the biological processes
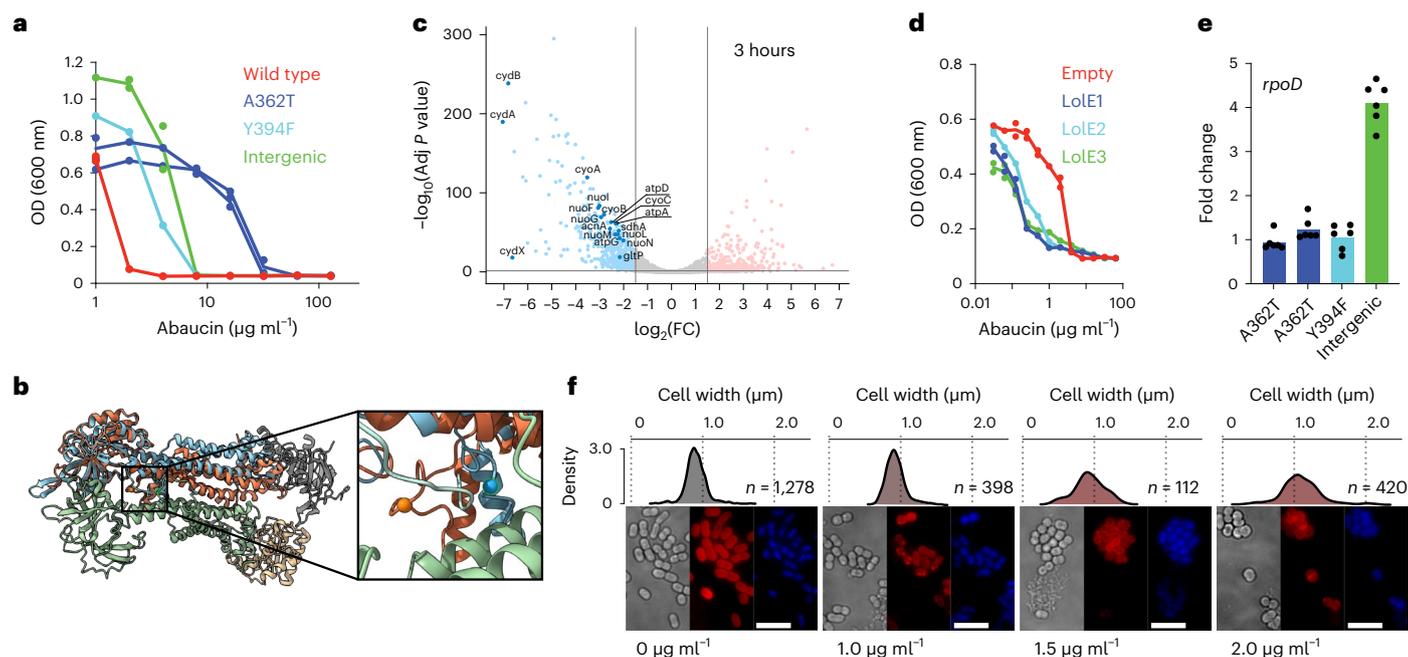
Fig. 3 | Abaucin inhibits lipoprotein trafficking in *A. baumannii*. a, Growth inhibition of wild-type *A. baumannii* (red), two independent abaucin-resistant mutants with the A362T variant in LolE (2,901,674C→T; blue), a mutant with the Y394F variant in LolE (2,901,577T→A; teal) and a mutant with an intergenic mutation (2,902,955G→A) upstream of LolE (green). All strains were grown in LB medium. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. **b,** Structure of the LolCDE complex from *E. coli* (*Ec*LolC is green, *Ec*LolD is gray and beige and *Ec*LolE is orange), with the predicted structure of *A. baumannii* LolE overlayed (*Ab*LolE is blue). The highlighted region shows the position of the A362T variant of *Ab*LolE and the homologous position in *E. coli* (I365). Note how *Ab*LolE position 362 is predicted to reside near *Ec*LolE I365 and both are adjacent to the acyl chain. **c,** RNA sequencing of wild-type *A. baumannii* treated with 5× MIC abaucin for 3 h. Data are the mean of biological duplicates. Transcript abundance in drug-treated samples is normalized to no-drug control cultures grown in identical conditions.

Vertical black lines show statistical significance cutoff values. Note the significant downregulation of genes involved in the electron transport chain and transmembrane ion transport. **d,** Growth inhibition of *A. baumannii* harboring an empty CRISPRi vector (red), or three distinct sgRNAs targeting *lolE* (blue, teal and green). All strains were grown in LB medium with induction using aTc to knock down *lolE*. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. **e,** qPCR quantifying the expression of *lolE* relative to the housekeeping gene *rpoD* in all four abaucin-resistant mutants, normalized to wild-type *A. baumannii*. Experiments were conducted in biological duplicates with technical triplicates. Bar height represents the mean normalized expression. **f,** Fluorescence micrographs of wild-type *A. baumannii* treated with abaucin. From left to right, cells were treated with no drug, or abaucin (1 μg ml⁻¹, 1.5 μg ml⁻¹ or 2 μg ml⁻¹) before imaging. Scale bars, 5 μm. Note the loss of native cell morphology and nucleoid condensation as a function of abaucin concentration.

that most statistically changed due to abaucin exposure. Through this method, we observed that abaucin caused downregulation of genes involved in the aerobic electron transport chain and transmembrane ion transport (Fig. 3c, Extended Data Fig. 3k and Supplementary Data 4). Interestingly, this transcriptional response is consistent with the activation of the Cpx two-component envelope stress response that has been well-characterized in the model bacterium *E. coli*[32,33]. Indeed, CpxAR has a principal role in monitoring lipoprotein trafficking from the inner membrane to the outer membrane in Gram-negative bacteria[34]. When lipoprotein trafficking is perturbed, CpxA autophosphorylates upon associated membrane stress, before phosphotransfer to the transcriptional regulator CpxR and subsequent transcriptional remodeling to restore envelope homeostasis[32]. These transcriptomics data further strengthen the hypothesis that abaucin displays antibacterial efficacy through perturbation of Lol complex-mediated lipoprotein trafficking. We note that the *A. baumannii* Lol system has not yet been thoroughly explored, necessitating our interpretation of these data in the context of the better-studied *E. coli* model.

With mutational, in silico structural and transcriptional data providing support that abaucin perturbs lipoprotein trafficking, we next investigated whether modulation of LolE expression would change abaucin potency. First, we hypothesized that the MIC of abaucin would decrease with decreased LolE expression[35]. To test this, we engineered *A. baumannii* with an inducible CRISPRi system using three distinct guide RNAs targeting *Ab*LolE. As expected, we observed that induction

of the CRISPRi construct resulted in fourfold to eightfold decreased abaucin MIC relative to an empty vector control with no guide RNA (Fig. 3d). Notably, uninduced cells with or without LolE-targeting guide RNAs all displayed identical abaucin sensitivities (Extended Data Fig. 3l). Next, we hypothesized that the abaucin-resistant mutant with the upstream intergenic mutation may be conferring resistance through increased expression of *lolE*—multicopy suppression-mediated resistance[36]. This is the inverse experiment to that described above. To test this hypothesis, we performed qPCR on all four independent abaucin-resistant mutants, as well as the wild-type parent strain of *A. baumannii*, and observed that the mutant carrying the upstream G to A mutation displayed ~4-fold increased expression of *lolE* relative to all other strains tested (Fig. 3e and Extended Data Fig. 3m). Intriguingly, this fourfold increase in *lolE* expression parallels the fourfold increase in abaucin MIC observed relative to wild-type *A. baumannii*. Collectively, these data are consistent with the mechanistic model that abaucin targets LolE-mediated lipoprotein trafficking.

Prior work has shown that inhibition of lipoprotein transport in *E. coli* results in abnormal cell morphology characterized by significant bacterial cell swelling and loss of nucleoid condensation[37]. We hypothesized that abaucin treatment would result in similar morphological features in *A. baumannii*. Therefore, we subjected wild-type *A. baumannii* to increasing concentrations of abaucin and imaged these cells using fluorescence microscopy. Here we stained *A. baumannii* cells with DAPI and FM4-64, respectively, to visualize DNA and the
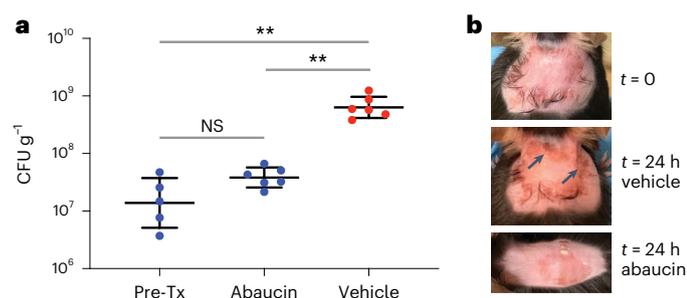
Fig. 4 | Abaucin can suppress _A. baumannii_ infection in a wound model.
a, In a dorsal wound infection model, mice were infected with _A. baumannii_ ATCC 17978 (~6.5 × 10^6 CFU). After 1 h of infection, mice were treated with vehicle (red; $n = 6$) or 4% abaucin (blue; $n = 6$) over 24 h. Bacterial load from wound tissue after treatment (25 h postinfection) was determined by selective plating. Pre-Tx represents the bacterial load at the time of initial treatment ($n = 5$). Black lines represent the mean ± s.d. of the bacterial load for each group. NS means not statistically significant; **$P < 0.005$ using unpaired two-sided t-test with Welch's correction (pre-Tx versus 4% abaucin, $P = 0.0704$; pre-Tx versus vehicle, $P = 0.0034$; vehicle versus 4% abaucin, $P = 0.0039$). **b**, Representative images of the dorsal surface of mice before infection ($t = 0$), after 24 h of treatment with vehicle, and after 24 h of treatment with abaucin. Note the inflammation seen in the vehicle control (arrows) that is absent in the abaucin-treated mouse.

cell envelope. Consistent with prior work in _E. coli_, abaucin-treated _A. baumannii_ cells displayed increased swelling and a loss of intracellular nucleoid condensation as a function of concentration (Fig. 3f). These data provide compelling phenotypic support that abaucin disrupts lipoprotein trafficking.

Abaucin can suppress _A. baumannii_ in a wound infection model

A. baumannii is a problematic nosocomial pathogen that survives for prolonged periods on surfaces and has the ability to accumulate extracellular DNA[1], including antibiotic resistance genes[38]. Indeed, _A. baumannii_ is a major cause of multidrug-resistant infections in wounded military personnel[39]. Our data showing that abaucin inhibited the growth of a wide array of multidrug-resistant clinical isolates of _A. baumannii_—including pan-resistant strains—provided confidence that abaucin may be used to treat such problematic wound infections. To test the in vivo efficacy of abaucin, we established a wound infection on the dorsal surface of neutropenic C57BL/6 mice using _A. baumannii_ ATCC 17978 (~6.5 × 10^6 CFU inoculum) and allowed the bacterial population to expand for 1 h. Mice were subsequently treated with Glaxal Base Moisturizing Cream supplemented with vehicle (1.65% DMSO) or abaucin (4% wt/vol). Application of vehicle or abaucin occurred at 2 h, 3 h, 4 h, 6 h, 10 h, 21 h and 24 h postinfection. Mice were killed at 25 h postinfection, and tissue was aseptically dissected and then plated to quantify _A. baumannii_ viability. In vehicle-treated mice, we retrieved ~6.9 × 10^8 CFU g^{-1} at the experimental endpoint and observed that the wounded tissues displayed significant inflammation. Contrarily, abaucin-treated mice carried ~4.0 × 10^7 CFU g^{-1}—nearly identical to the pretreated infection control mice—and abaucin-treated tissues displayed markedly less inflammation (Fig. 4). These data show that abaucin can effectively suppress an _A. baumannii_ wound infection, consistent with its effect on _A. baumannii_ viability in vitro (Fig. 1f). We note, however, that conventional clinically employed bactericidal antibiotics would likely result in enhanced _A. baumannii_ clearance in our wound model relative to abaucin. However, due to widespread resistance to conventional antibiotics, as well as the broad-spectrum activity of these medicines, abaucin and derivatives thereof may represent a structurally and functionally new class of _A. baumannii_-specific antibiotics that can overcome existing resistance determinants. Indeed, given that abaucin

is readily amenable to medicinal chemistry optimization (Table 1), we emphasize the existing opportunity to develop structural analogs of abaucin with enhanced in vivo activity.

Discussion

Structurally and functionally new antibiotics are urgently needed for _A. baumannii_, which is notoriously difficult to eradicate due to its ability to uptake and retain antibiotic resistance determinants[1]. Moreover, species-selective antibiotics hold promise to limit the horizontal dissemination of resistance determinants and decrease the likelihood of dysbiosis during treatment[8]. The machine learning-guided discovery of abaucin highlights the utility of algorithmic approaches to discover new antibacterial molecules against _A. baumannii_ and provides the field with a promising new narrow-spectrum molecular scaffold to address one of the world's most challenging Gram-negative pathogens. Lipoprotein trafficking is a highly sought-after antibiotic target that has yet to be perturbed by clinically used antibacterial drugs, which is advantageous toward increasing the duration of clinical utility of molecules that disrupt this process. Notably, our observations that abaucin displays narrow-spectrum activity through perturbation of LolE-mediated lipoprotein trafficking can be explained, at least in part, due to the divergence of the _A. baumannii_ Lol system relative to most other Gram-negative species[16]. Specifically, in most Gram-negative organisms, the inner membrane-associated lipoprotein transport machinery consists of LolC, LolD and LolE in an asymmetric multiprotein complex. _A. baumannii_, on the other hand, encodes a symmetric inner membrane complex containing LolD and two copies of LolE (also termed LolF), without LolC.

The work described herein represents an advancement toward the validation of machine learning for new narrow-spectrum antibiotic discovery and adds to the growing body of literature that supports the utility of computational approaches to accelerate drug discovery more generally[7]. Indeed, over the past half-decade, investigators have successfully applied classic molecular fingerprint-based machine learning methods, as well as more sophisticated neural network approaches, to predict antibiotics[5], antiviral compounds[40] and anticancer therapies[41]. Furthermore, recent advancements in chemical generation algorithms have resulted in the design of new molecules with desired properties of interest[42].

Expectedly, there exist opportunities for improvement of molecular property prediction methods. For example, future endeavors can readily expand on this work by increasing the size of the _A. baumannii_ antibiotic training dataset and performing predictions on larger in silico chemical libraries. Given that our prediction filtering method was exceptionally stringent to enrich for chemicals that were structurally divergent from our training dataset, it is likely that similarly aggressive filtering approaches could be leveraged going forward to manage the number of molecules for experimental validation, even when predictions are performed on the scale of hundreds of millions to billions of compounds. However, we emphasize that algorithmic approaches may be applied to in silico chemical libraries of any size, serving to decrease the time and costs required to identify valuable molecules by avoiding the strict necessity for successive rounds of resource-intensive screening.

Moreover, investigators can apply multiproperty optimization models that simultaneously predict numerous properties of interest. In the context of antibiotic drug discovery specifically, future models could be trained on (1) growth inhibition of a pathogen of interest and (2) mammalian cell toxicity, given that a sufficiently robust training dataset exists. In this manner, forthcoming predictive models can be applied to wide regions of unexplored chemical space, with more confidence that prioritized molecules will satisfy multiple properties that are required of new clinical antibiotics. Indeed, with the increasing availability of high-quality datasets on which to train, we posit that machine learning methods are now well positioned to become widely

employed tools to more efficiently identify structurally and functionally new antibacterial leads.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-023-01349-8.

References

1. Antunes, L. C. S., Visca, P. & Towner, K. J. *Acinetobacter baumannii*: evolution of a global pathogen. *Pathog. Dis.* **71**, 292–301 (2014).
2. *2020 Antibacterial Agents in Clinical and Preclinical Development: An Overview and Analysis* (World Health Organization, 2021); https://www.who.int/publications/i/item/9789240021303
3. Walsh, C. Where will new antibiotics come from? *Nat. Rev. Microbiol.* **1**, 65–70 (2003).
4. Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I. & Miller, A. A. ESKAPEing the labyrinth of antibacterial discovery. *Nat. Rev. Drug Discov.* **14**, 529–542 (2015).
5. Stokes, J. M. et al. A deep learning approach to antibiotic discovery. *Cell* **180**, 688–702 (2020).
6. Ma, Y. et al. Identification of antimicrobial peptides from the human gut microbiome using deep learning. *Nat. Biotechnol.* **40**, 921–931 (2022).
7. Lluka, T. & Stokes, J. M. Antibiotic discovery in the artificial intelligence era. *Ann. N. Y. Acad. Sci.* **1519**, 74–93 (2023).
8. Melander, R. J., Zurawski, D. V. & Melander, C. Narrow-spectrum antibacterial agents. *MedChemComm* **9**, 12–21 (2018).
9. Theriot, C. M. et al. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. *Nat. Commun.* **5**, 3114 (2014).
10. Willing, B. P. et al. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* **139**, 1844–1854 (2010).
11. Turnbaugh, P. J. et al. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
12. Kelly, J. R. et al. Transferring the blues: depression-associated gut microbiota induces neurobehavioural changes in the rat. *J. Psychiatr. Res.* **82**, 109–118 (2016).
13. Wu, N. et al. Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb. Ecol.* **66**, 462–470 (2013).
14. Lee, H. S., Plechot, K., Gohil, S. & Le, J. *Clostridium difficile*: diagnosis and the consequence of over diagnosis. *Infect. Dis. Ther.* **10**, 687–697 (2021).
15. Corsello, S. M. et al. The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.* **23**, 405–408 (2017).
16. Kaplan, E., Greene, N. P., Crow, A. & Koronakis, V. Insights into bacterial lipoprotein trafficking from a structure of LolA bound to the LolC periplasmic domain. *Proc. Natl Acad. Sci. USA* **115**, E7389–E7397 (2018).
17. Tang, X. et al. Structural basis for bacterial lipoprotein relocation by the transporter LolCDE. *Nat. Struct. Mol. Biol.* **28**, 347–355 (2021).
18. Yang, K. et al. Analyzing learned molecular representations for property prediction. *J. Chem. Inf. Model.* **59**, 3370–3388 (2019).
19. Landrum, G. RDKit: a software suite for cheminformatics, computational chemistry, and predictive modeling. https://www.rdkit.org/RDKit_Overview.pdf
20. Seok, S. J. et al. Blockade of CCL2/CCR2 signalling ameliorates diabetic nephropathy in db/db mice. *Nephrol. Dial. Transplant.* **28**, 1700–1710 (2013).
21. Cerri, C. et al. The chemokine CCL2 mediates the seizure-enhancing effects of systemic inflammation. *J. Neurosci.* **36**, 3777–3788 (2016).
22. Chargari, C. et al. Preclinical assessment of JNJ-26854165 (Serdemetan), a novel tryptamine compound with radiosensitizing activity in vitro and in tumor xenografts. *Cancer Lett.* **312**, 209–218 (2011).
23. Lehman, J. A. et al. Serdemetan antagonizes the Mdm2-HIF1α axis leading to decreased levels of glycolytic enzymes. *PLoS ONE* **8**, e74741 (2013).
24. Stokes, J. M., Lopatkin, A. J., Lobritz, M. A. & Collins, J. J. Bacterial metabolism and antibiotic efficacy. *Cell Metab.* **30**, 251–259 (2019).
25. Zheng, E. J., Stokes, J. M. & Collins, J. J. Eradicating bacterial persisters with combinations of strongly and weakly metabolism-dependent antibiotics. *Cell Chem. Biol.* **27**, 1544–1552 (2020).
26. Francino, M. P. Antibiotics and the human gut microbiome: dysbioses and accumulation of resistances. *Front. Microbiol.* **6**, 1543 (2016).
27. Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. *Clostridium difficile* infection. *Nat. Rev. Dis. Prim.* **2**, 16020 (2016).
28. Sharma, S. et al. Mechanism of LolCDE as a molecular extruder of bacterial triacylated lipoproteins. *Nat. Commun.* **12**, 4687 (2021).
29. Nicholson, W. L. & Maughan, H. The spectrum of spontaneous rifampin resistance mutations in the *rpoB* gene of *Bacillus subtilis* 168 spores differs from that of vegetative cells and resembles that of *Mycobacterium tuberculosis*. *J. Bacteriol.* **184**, 4936–4940 (2002).
30. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
31. Buel, G. R. & Walters, K. J. Can AlphaFold2 predict the impact of missense mutations on structure? *Nat. Struct. Mol. Biol.* **29**, 1–2 (2022).
32. Raivio, T. L., Leblanc, S. K. D. & Price, N. L. The *Escherichia coli* Cpx envelope stress response regulates genes of diverse function that impact antibiotic resistance and membrane integrity. *J. Bacteriol.* **195**, 2755–2767 (2013).
33. Guest, R. L., Wang, J., Wong, J. L. & Raivio, T. L. A bacterial stress response regulates respiratory protein complexes to control envelope stress adaptation. *J. Bacteriol.* **199**, e00153-17 (2017).
34. Delhaye, A., Laloux, G. & Collet, J.-F. The lipoprotein NlpE is a Cpx sensor that serves as a sentinel for protein sorting and folding defects in the *Escherichia coli* envelope. *J. Bacteriol.* **201**, e00611-18 (2019).
35. Peters, J. M. et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* **165**, 1493–1506 (2016).
36. Pathania, R. et al. Chemical genomics in *Escherichia coli* identifies an inhibitor of bacterial lipoprotein targeting. *Nat. Chem. Biol.* **5**, 849–856 (2009).
37. McLeod, S. M. et al. Small-molecule inhibitors of Gram-negative lipoprotein trafficking discovered by phenotypic screening. *J. Bacteriol.* **197**, 1075–1082 (2015).
38. Manchanda, V., Sanchaita, S. & Singh, N. Multidrug resistant acinetobacter. *J. Glob. Infect. Dis.* **2**, 291–304 (2010).
39. Davis, K. A., Moran, K. A., McAllister, C. K. & Gray, P. J. Multidrug-resistant Acinetobacter extremity infections in soldiers. *Emerg. Infect. Dis.* **11**, 1218–1224 (2005).
40. Jin, W. et al. Deep learning identifies synergistic drug combinations for treating COVID-19. *Proc. Natl Acad. Sci. USA* **118**, e2105070118 (2021).
41. Preuer, K. et al. DeepSynergy: predicting anti-cancer drug synergy with Deep Learning. *Bioinformatics* **34**, 1538–1546 (2018).

42. Jin, W., Barzilay, R. & Jaakkola, T. Junction tree variational autoencoder for molecular graph generation. In *Proc. 35th International Conference on Machine Learning* (eds Dy, J. & Krause, A.) 2323–2332 (PMLR, 2018).

Methods

Training data acquisition

A. baumannii ATCC 17978 was grown in 2 ml LB medium (Becton, Dickinson and Company) overnight at 37 °C with shaking. Cells were diluted 1/10,000 into fresh LB and 99 μl of cells was added to each well of a 96-well flat-bottom plate (Corning). Next, 1 μl of a 5 mM stock of each molecule from a collection of 7,684 small molecules (FDA-approved drugs and molecules from screening collections from the Broad Institute) was added, in duplicate, using an Agilent Bravo liquid handling system. The final concentration was 50 μM. Plates were incubated in sealed plastic bags at 37 °C for 16 h and read at 600 nm using a SpectraMax M3 plate reader. Plate data were normalized based on the interquartile mean of each plate before binarization into 'active' and 'nonactive' categories for model training. Active molecules were defined as those that resulted in growth at least 1σ below the mean growth of the entire dataset.

Model training and predictions

A directed message-passing neural network (Chemprop), like other message-passing neural networks, learns to predict molecular properties directly from the graph structure of molecules, where atoms are nodes and bonds are edges. For every molecule in our training dataset, we reconstructed the molecular graph corresponding to each compound's SMILES string and determined the set of atoms and bonds using RDKit[19]. Next, we initialized a feature vector for each atom and bond, as described previously[5].

The model applies a series of message-passing steps where it aggregates information from neighboring atoms and bonds to build a representation of local chemistry. On each step of message passing, every bond's featurization is updated by summing the featurization of neighboring atoms and bonds, applying a single neural network layer, adding the bond's previous featurization, and then applying ReLU activation. After a defined number of message-passing steps, the learned featurizations across the molecule are summed to produce a single featurization for the whole molecule. Lastly, this featurization is subjected to a feed-forward neural network that outputs a prediction of the property of interest—in our case, antibiotic activity against *A. baumannii*. To augment the architecture described here, we employed the following two model enhancements: molecule-level features and ensembling.

Molecule-level features. The message-passing approach is ideal for extracting features of local chemistry within a larger molecule. However, it can struggle to extract global molecular features for larger molecules. To address this limitation, we concatenated the molecular representation that is learned during message passing with 200 additional molecule-level features computed using RDKit.

Ensembling. Ensembling was used to further improve model performance, where several copies of the same model architecture with different random initial weights are trained and their predictions are averaged. Here we used an ensemble of ten models, with each model trained on a unique split of the training dataset.

After model building and training on the 7,684-molecule training dataset (with ~6.2% active examples), we applied our model to the updated Drug Repurposing Hub, consisting of 6,680 compounds, many of which occupy a unique chemical space relative to that on which the model was trained. Here we randomly split the dataset into 80% training data, 10% validation data and 10% test data. We trained our model on these data for 30 epochs and evaluated the model on the validation data at the end of each epoch. Once training was complete, we used the model parameters that performed best on the validation data and tested the model on the test data. We ran tenfold cross-validation by repeating this procedure with ten different splits by systematically dividing training, validation and test sets such that all molecules appeared at equal proportions across all sets over the course of training

iterations. After we achieved model performance that was acceptable for our prediction task, we conducted predictions on the Drug Repurposing Hub.

Our model had the following hyperparameters: number of message-passing steps = 3; neural network hidden size = 300; number of feed-forward layers = 2; dropout probability = 0.

Structural analysis

We used Tanimoto similarity to quantify the chemical relationship between molecules in our training dataset and prediction dataset. The Tanimoto similarity between two molecules is a measure of the proportion of shared chemical substructures. To compute Tanimoto similarity, we first made Morgan fingerprints for each molecule using radius = 2 and 2,048-bit fingerprint vectors using RDKit. Tanimoto similarity was then computed as the number of chemical substructures contained in both molecules divided by the total number of unique chemical substructures in either molecule. The Tanimoto similarity is a number between 0 and 1, where 0 represents no shared substructures and 1 represents all substructures that are shared. For t-distributed stochastic neighbor embedding (t-SNE) analysis, plots were created using the implementation of t-distributed stochastic neighbor embedding by scikit-learn[43]. Here we first used RDKit to compute Morgan fingerprints for each molecule as described above. We then used t-SNE with the Jaccard distance metric to reduce the data from 2,048 dimensions to the two dimensions that are plotted. Jaccard distance is Tanimoto distance, which is defined as follows: Tanimoto distance = 1 − Tanimoto similarity. We used the scikit-learn default values for all t-SNE parameters besides the distance metric.

Pathogen growth inhibition

Cells were grown overnight in 2 ml LB medium and diluted 1/10,000 into fresh LB. In 96-well flat-bottom plates (Corning), cells were then introduced to compound at a final concentration of 50 μM, or to compound at twofold serial dilutions, in final volumes of 100 μl. Plates were then incubated at 37 °C until untreated control cultures reached the stationary phase, at which time plates were read at 600 nm using a SpectraMax M3 plate reader. We note here that the incubation time required to reach the stationary phase differed slightly between strains. For abaucin analog spectrum of coverage assays, the laboratory strains were *E. coli* BW25113, *S. aureus* RN4220 and *P. aeruginosa* PAO1. Clinical isolate panels were curated from the CDC AR Isolate Bank and assayed in LB as described above.

Bacterial cell killing

Cells were grown overnight in 2 ml LB medium and diluted 1/10,000 into fresh LB. In 96-well flat-bottom plates (Corning), cells were grown to the required density (100 μl final volume) at 37 °C, at which time compound was added at the indicated concentration and cultures were incubated for the required duration. Cells were then pelleted in plates by centrifugation at 4,000*g* for 15 min at 4 °C and washed in ice-cold PBS. After washing, cells were tenfold serially diluted in PBS and plated on LB. In experiments where cells were incubated with antibiotic in nutrient-deplete conditions, cells were grown to the required density in LB media at 37 °C, washed in PBS and subsequently resuspended in PBS before the addition of compound (100 μl final volume). After cultures were incubated for the required duration at 37 °C, cells were pelleted in plates by centrifugation at 4,000*g* for 15 min at 4 °C and washed in ice-cold PBS. After washing, cells were tenfold serially diluted in PBS and plated on LB. For post-treatment regrowth experiments, cells were treated with abaucin and incubated in LB for 6 hours as described above (100 μl final volume). Cells were then pelleted in plates by centrifugation at 4,000*g* for 15 min at 4 °C and washed in ice-cold PBS. After washing, cells were resuspended in 100 μl fresh LB, diluted 1/1,000 in fresh 100 μl LB and grown at 37 °C in a Biotek Synergy H1 plate reader. Plates were read at 600 nm every 20 min for 16 h.

Commensal species growth inhibition

For analysis of commensal species from the human gut, strains were grown from frozen stock on brain heart infusion (BHI) agar supplemented with 0.5 g l^{-1} L-cysteine, 10 mg l^{-1} hemin and 1 mg l^{-1} vitamin K (BHI3). Single colonies were picked, transferred to 96-well plates containing liquid BHI3 and grown for 24 h at 37 °C. Antibiotics (ampicillin, Sigma-Aldrich; ciprofloxacin, Sigma-Aldrich; abaucin; Supplementary Note) were arrayed in BHI3 broth at twofold the final desired concentrations, and prereduced in an anaerobic environment overnight. Liquid bacterial cultures were standardized to twofold the desired optical density (OD$_{600}$ = 0.02). The twofold concentrated cultures were then inoculated into 96-well plates (50 µl) containing twofold concentrated antibiotics (50 µl). The final antibiotic concentrations were ampicillin (128 µg ml^{-1}), ciprofloxacin (0.25 µg ml^{-1}) and abaucin (2 µg ml^{-1}, 20 µg ml^{-1} and 40 µg ml^{-1}). The final culture volume was 100 µl. After inoculation, plates were sealed with breathable membranes (Breathe-Easy) and incubated at 37 °C without shaking for 24 h. After incubation, the plates were read at 600 nm using Biotek Synergy H1 plate reader. All bacterial strains were grown and incubated in a Coy Laboratory Vinyl Anaerobic Chamber (5% CO$_2$, 2% H$_2$ and 93% N$_2$). All media was prereduced under anaerobic conditions for at least 4 h before use.

For analysis of commensal skin species in aerobic conditions, strains were grown from frozen on BHI3 agar and incubated at 37 °C for 24 h. Single colonies were picked and transferred to 96-well plates containing liquid BHI3, which were then incubated at 37 °C without shaking for 24 h. Liquid cultures were then standardized to OD$_{600}$ = 0.02 (twofold the final OD). Antibiotics (ampicillin, ciprofloxacin and abaucin) were arrayed in BHI3 broth at twofold the final concentrations (the final antibiotic concentrations were ampicillin (128 µg ml^{-1}), ciprofloxacin (0.25 µg ml^{-1}) and abaucin (2 µg ml^{-1}, 20 µg ml^{-1} and 40 µg ml^{-1})). The twofold concentrated cultures (50 µl) were then inoculated into 96-well plates containing the twofold concentrated antibiotics (50 µl) to achieve a final volume of 100 µl. Plates were sealed with breathable membranes (Breathe-Easy) and incubated at 37 °C without shaking for 48 h under 5% CO$_2$. After incubation, plates were read at 600 nm using a Biotek Synergy H1 plate reader.

Suppressor mutant evolution and sequencing

A. baumannii ATCC 17978 was grown in 2 ml LB medium overnight at 37 °C with shaking. -10^8 CFU in 100 µl liquid LB was deposited onto solid LB plates supplemented with abaucin at varying concentrations ranging from 2 µg ml^{-1} to 10 µg ml^{-1}. Plates were incubated at 37 °C and monitored every 24 h for the emergence of abaucin mutants. On the emergence of colonies, these were purified by restreaking onto solid LB and solid LB supplemented with abaucin at the same concentration from which the colonies were originally collected. Four independent abaucin-resistant strains were subsequently selected for whole-genome sequencing. Chromosomal DNA from each mutant and wild-type *A. baumannii* ATCC 17978 was purified using a Gentra Puregene DNA isolation kit (Qiagen). DNA (chromosomal and plasmid) was sequenced on an Illumina MiSeq platform and reads were aligned to the reference *A. baumannii* ATCC 17978 genome (CP053098.1, CP053099.1 and CP053100.1) using Breseq[44]. For frequency of resistance quantification, 1.1×10^8 CFU of *A. baumannii* ATCC 17978 in 100 µl liquid LB was deposited onto solid LB plates supplemented with abaucin at the noted concentrations. After 24 h, 48 h and 72 h of incubation at 37 °C, colonies were counted, and these values were divided by 1.1×10^8 to quantify the frequency of resistance.

Transcriptomic analysis

A. baumannii ATCC 17978 was grown in 2 ml LB medium overnight at 37 °C with shaking. Cells were diluted 1/10,000 into 50 ml fresh LB and grown to mid-log at 37 °C with shaking, at which time cultures were supplemented with 5× MIC (10 µg ml^{-1}) abaucin or no drug and grown for an

additional 3 h, 4.5 h or 6 h. After the required durations of incubation post-treatment, 2 ml samples were collected and flash-frozen on liquid nitrogen. cDNA library construction and sequencing for these samples were performed by Genewiz. Paired-end sequence data in FASTQ file format were aligned to the transcripts of CP053098.1 (*A. baumannii* ATCC 17978 chromosome), CP053099.1 (*A. baumannii* ATCC 17978 plasmid 1) and CP053100.1 (*A. baumannii* ATCC 17978 plasmid 2). mRNA abundances were then quantified as transcripts per million using the kallisto (version 0.46.2)[45] 'quant' function, and the parameter '-b 100' with paired-end reads. All downstream analyses were performed using R (version 4.1.0)[46]. Expression differences between abaucin-treated samples and nontreated controls were quantified using the DESeq2 pipeline (version 1.34.0)[47]. We defined a cutoff of log$_2$(FC) ≥ 1.5 and padj ≤ 0.01 to identify differentially expressed genes. For GO enrichment, differentially expressed *A. baumannii* transcripts for which no gene name was identified in CP053098.1, CP053099.1 or CP053100.1 were subject to blastp analysis against *E. coli* MG1655 (https://blast.ncbi.nlm.nih.gov/). *A. baumannii* proteins with E values < 10^{-50}, percent identify >40% and query coverage >80% to annotated *E. coli* proteins were binned and GO-enriched with EcoCyc Pathway Tools[48–50] using the Fisher exact test for statistical significance of GO enrichment.

CRISPRi LolE knockdown

Three independent 20 base pair sgRNAs targeting *lolE* (sgRNA*lolE*-1: 5′-TAAACGTAAGCCAAGCGAAT-3′; sgRNA*lolE*-2: 5′-CAAATTCTTCAC AATGTCAT-3′; sgRNA*lolE*-3: 5′-TTTAAGTGAGTCGAGGCTAC-3′) were designed using predictive software[51] to maximize on-target activity and minimize off-target binding in *A. baumannii* ATCC 17978. sgRNAs were then cloned into pFD152 (Addgene plasmid 125546; provided by D. Bikard) using a single-step golden gate assembly reaction, as described previously[52]. sgRNA sequences were verified using Sanger sequencing and constructs were transformed into *A. baumannii* ATCC 17978. pFD152 was selected in *A. baumannii* using spectinomycin (250 µg ml^{-1} in liquid media and 500 µg ml^{-1} on solid media; Sigma-Aldrich). To determine whether LolE knockdown enhanced the growth inhibitory potency of abaucin, *A. baumannii* ATCC 17978 harboring pFD152, pFD152-sgRNA*lolE*-1, pFD152-sgRNA*lolE*-2 or pFD152-sgRNA*lolE*-3 were grown overnight in LB medium, diluted 1/10,000 into fresh LB, and grown mid-log phase (OD - 0.2) at 37 °C. Anhydrotetracycline (aTc; Sigma-Aldrich) was added to the subcultures at a final concentration of 1 µg ml^{-1} to induce knockdown of *lolE* expression and cultures were incubated for an additional 2.5 h at 37 °C with shaking. Following this induction period, cultures were back diluted to the equivalent OD of 1/10,000 of a dense culture and introduced to varying concentrations of abaucin and 0.5 µg ml^{-1} of aTc. MICs were then determined as described above.

Quantitative reverse transcription (qRT)–PCR

Wild-type *A. baumannii* ATCC 17978 and the resistant mutants were grown in 3 ml LB medium overnight at 37 °C with shaking. Cells were diluted 1/10,000 into 15 ml fresh LB and grown to mid-log at 37 °C with shaking, at which time cells were pelleted by centrifugation at 4,000g for 15 min at 4 °C and frozen for -2 h. Cells were lysed in 100 µl of 1 mg ml^{-1} lysozyme and incubated at room temperature for 10 min. RNA was isolated using the RNEasy Mini Kit (Qiagen), with some modifications. Briefly, to each sample, 3.5 µl of 2-mercaptoethanol, 250 µl of RLT buffer and 250 µl of 95% ethanol was added. Seven hundred microliters of each sample was then loaded onto a spin column and incubated at room temperature for 5 min. The sample was centrifuged at 8,000g for 1 min. The column was washed with 700 µl RW1 buffer and centrifuged at 8,000g for 1 min. The column was washed with 700 µl RPE buffer and centrifuged at 8,000g for 2 min. The column was washed again with 500 µl RPE buffer and centrifuged at 8,000g for 2 min. The column was transferred to a new collection tube and incubated at room temperature for 5 min. The empty column was

centrifuged at 13,000*g* for 2 min and incubated at room temperature for 5 min. The column was then placed into a fresh 1.5 ml microcentrifuge tube and 45 µl of RNase-free water was added to the center of the column and incubated at room temperature for 5 min. To elute RNA, the column was centrifuged at 8,000*g* for 2 min. RNA was subject to DNase I treatment for 10 min at room temperature. The resulting RNA was quantified using a Nanodrop. Next, the RNA integrity was assessed through agarose gel electrophoresis. For cDNA synthesis, the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) was used as specified by the manufacturer's instructions. The cDNA was then amplified with iTaq Universal SYBR Green Supermix (Bio-Rad) via a Bio-Rad CFX96 (Bio-Rad) following the manufacturer's instructions. The target amplification efficiency was evaluated by generating a standard curve with dilutions of cDNA (>95% amplification efficiency for each primer pair). Primers lolE195F + lolE195R were used to amplify *lolE*; rpoD182F + rpoD182R, glt188F + glt188R and gyrB196F + gyrB196R were used to amplify the housekeeping genes *rpoD*, *gltA* and *gyrB*. Triple technical replicates were examined per biological sample, with controls omitting reverse-transcriptase to verify a lack of genomic DNA contamination, and cDNA to verify the absence of external nucleic acid contamination and primer dimer formation. Transcript level of *lolE* from each mutant was evaluated using the comparative $2^{-\Delta\Delta Ct}$ method to the wild-type strain, normalizing to the $2^{-\Delta\Delta Ct}$ value of housekeeping genes.

> *lolE*195F−5′-CAGCTGGAGATGGAGTCGCT-3′
> *lolE*195R−5′-TGGTGACGGGTATTGACCCT-3′
> *rpoD*182F−5′-ACATGGCCGTAATAGCCCTGA-3′
> *rpoD*182F−5′-CCATACGGCCACGACGTACT-3′
> *glt*188F−5′-GCGACAGCTTCATGCGAGTC-3′
> *glt*188R−5′-TGAGCACGAACTTTCGCATCG-3′
> *gyrB*196F−5′-ATTAGTGCTGATGCGCCTGC-3′
> *gyrB*196R−5′-CTTGATCCGCCCATTGCTGG-3′

Structural protein prediction

We aligned the wild-type *A. baumannii* protein sequence of interest against the UniProt/SwissProt database[53] using blastp (version 2.13.0+)[54,55] through the NCBI blast platform (https://blast.ncbi.nlm.nih.gov/) and obtained a multiple-sequence alignment. This alignment identified the *E. coli* K12 LolE protein as a homolog with *E* value = 6×10^{-88}. Next, we aligned the *E. coli* LolE protein, the *A. baumannii* wild-type LolE protein and the *A. baumannii* abaucin-resistant A362T variant LolE protein using Geneious Prime software (version 2021.2.2; https://www.geneious.com); we applied a 'global alignment with free end gaps' as the alignment type and 'blosum62' as the cost matrix, with parameters gap open penalty = 12; gap extension penalty = 3 and refinement iterations = 2. We then predicted the structure of the wild-type *A. baumannii* LolE protein and A362T variant protein sequences with RoseTTAfold[30] through the Robetta platform (https://robetta.bakerlab.org/). We did not provide a multiple-sequence alignment for this computation. We obtained the protein structure of the wild-type *E. coli* LolCDE in complex with a lipoprotein from ref. 17, Protein Data Bank: 7ARH. We then aligned the structures of the wild-type *A. baumannii* LolE protein and the LolCDE complex with Maestro (version 12.9.137) from Schrödinger and ChimeraX (version 1.3rc202112020528)[56,57]. In Maestro, we used the default structural alignment parameters. With ChimeraX, we used the 'Needleman-Wunsch' alignment algorithm with 'blosum62' as the similarity matrix, with parameters 'bb' as chain pairing; 0.3 as SS fraction; 1 as gap extend; 18/18/6 as gap open (HH/SS/other); and 2 as iteration cutoff. Final figures of structural alignments were generated with ChimeraX (version 1.3rc202112020528)[56,57].

Quantitative microscopy

A. baumannii ATCC 17978 was grown overnight in 2 ml LB medium and diluted 1/10,000 into fresh LB. In 96-well flat-bottom plates, cells were then introduced to the compound at the indicated concentrations, in final volumes of 100 µl. Plates were grown to mid-exponential phase in a Biotek Epoch 2 plate reader with shaking at 37 °C. Cell suspensions were transferred to a poly-lysine coated 0.17 mm glass-bottom imaging plate (Brooks Scientific) and incubated with FM 4-64 (1 µg ml⁻¹ final concentration; Invitrogen) and DAPI (0.2 µg ml⁻¹ final concentration; Invitrogen) probes for 10 min in the dark. Samples were imaged using a Nikon Eclipse Ti inverted microscope at ×100 magnification, and cell widths were assessed using Nikon Elements software. Cell widths for each sample were summarized in R using the density function to explore population shifts in morphology as a function of drug concentration.

Animal models

Mouse model experiments were conducted according to the guidelines set by the Canadian Council on Animal Care, using protocols approved by the Animal Review Ethics Board and McMaster University under Animal Use Protocol 20-12-41. Before infection, mice were relocated at random from housing cages to single-occupancy control or treatment cages. No animals were excluded from the analysis, and blinding was considered unnecessary. Six- to eight-week-old female C57BL/6N mice were pretreated with 150 mg kg⁻¹ (day T-4) and 100 mg kg⁻¹ (day T-1) of cyclophosphamide to render mice neutropenic. On day T-0, mice were anesthetized using isoflurane and administered buprenorphine as an analgesic at 0.1 mg kg⁻¹ intraperitoneally. A 2 cm² abrasion on the dorsal surface of the mouse was inflicted through tape-stripping to the basal layer of the epidermis using approximately 30–35 pieces of autoclave tape. Mice were immediately infected with ~6.5 × 10⁶ CFU *A. baumannii* ATCC 17978 directly pipetted onto the wound bed. The infection was left to establish for 1 h before the first treatment with Glaxal Base supplemented with vehicle (1.65% DMSO) or abaucin (4% wt/vol). Mice (*n* = 5 or 6) were treated 1, 2, 3, 4, 6, 10, 21 and 24 h postinfection with ~10 to 20 µl Glaxal Base with abaucin (treatment) or DMSO (control). Mice were killed at experimental endpoint (25 h postinfection) and wound tissue was collected, homogenized in phosphate-buffered saline and plated on solid LB medium supplemented with chloramphenicol to quantify bacterial load. For chemical preparation, abaucin was weighed and solubilized in 1.65% DMSO and then added to a predetermined amount of Glaxal Base to a final concentration of 4% w/v. The solution was mixed thoroughly to ensure an even distribution of the compound in the carrier. For control groups, the same amount of vehicle (DMSO) was measured and mixed through Glaxal Base.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data Availability

GenBank accession numbers for sequencing of abaucin-resistant mutants are BankIt2629921 – OP677864, OP677865, OP677866 and OP677867. GEO accession numbers for RNA sequencing datasets are GSE214305 – GSM6603484, GSM6603485, GSM6603486, GSM6603487, GSM6603488, GSM6603489 and GSM6603490. Source data are provided with this paper.

Code Availability

All custom code used for antibiotic prediction is open source and can be accessed without restriction at https://github.com/chemprop/chemprop. A cloned snapshot used for this paper is available at https://github.com/GaryLiu152/chemprop_abaucin. All commercial software used is described in Methods. Source data are provided with this paper.

References

43. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

44. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).

45. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

46. Ihaka, R. & Gentleman, R. R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299–314 (1996).

47. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

48. Karp, P. D. Pathway databases: a case study in computational symbolic theories. *Science* **293**, 2040–2044 (2001).

49. Keseler, I. M. et al. EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* **41**, D605–D612 (2013).

50. Karp, P. D. et al. Pathway tools version 19.0 update: software for pathway/genome informatics and systems biology. *Brief. Bioinform.* **17**, 877–890 (2016).

51. Calvo-Villamañán, A. et al. On-target activity predictions enable improved CRISPR–dCas9 screens in bacteria. *Nucleic Acids Res.* **48**, e64 (2020).

52. Depardieu, F. & Bikard, D. Gene silencing with CRISPRi in bacteria and optimization of dCas9 expression levels. *Methods* **172**, 61–75 (2020).

53. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).

54. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

55. Altschul, S. F. et al. Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* **272**, 5101–5109 (2005).

56. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

57. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

Author contributions

J.M.S. and J.J.C. conceptualized the study; J.M.S., G.L., K.S. and W.J. performed model building and training; J.M.S., D.B.C., K.R. and A.C-P. performed mechanistic investigations; J.M.S., K.R. and S.A.S. performed spectrum of activity experiments; J.C.M. conducted mouse model experiments; M.F. performed chemical synthesis; J.M.S. and J.J.C. wrote the paper; J.M.S., J.J.C., R.B., T.J., M.G.S., B.K.C. and J.M. supervised the research.

Competing interests

J.M.S. is cofounder and scientific director of Phare Bio. J.J.C. is cofounder and scientific advisory board chair of Phare Bio. J.J.C. is cofounder and scientific advisory board chair of Enbiotix. The other authors declare no competing interests.
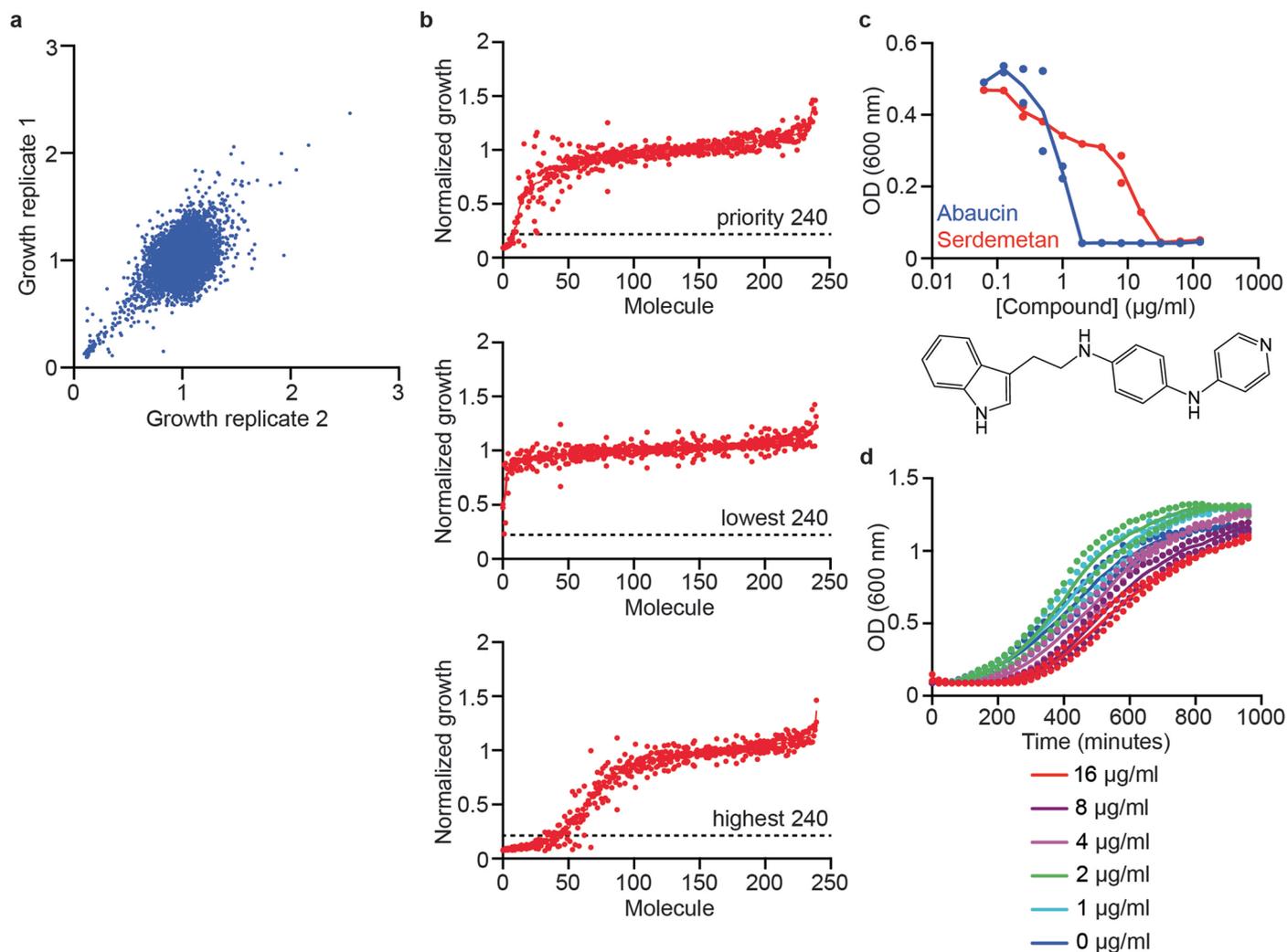
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41589-023-01349-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-023-01349-8.

Correspondence and requests for materials should be addressed to James J. Collins or Jonathan M. Stokes.
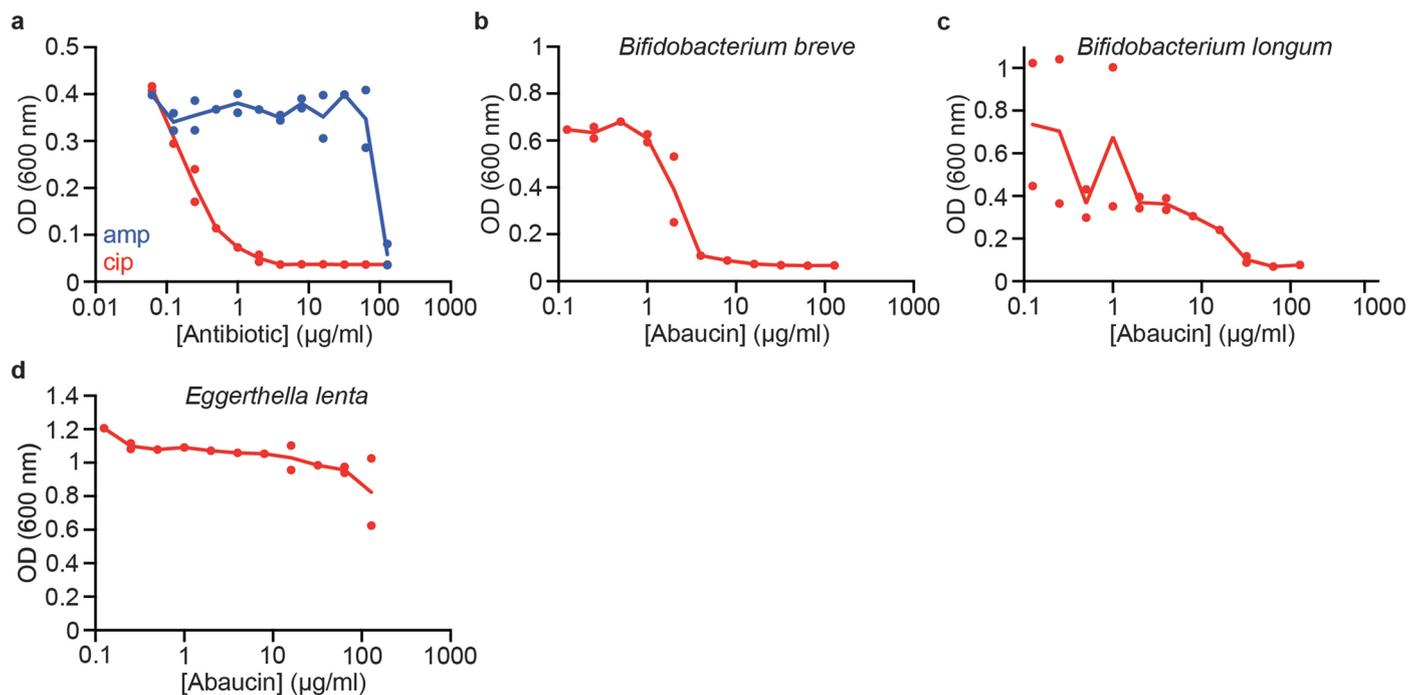
Peer review information *Nature Chemical Biology* thanks Jean Francois Collet and the other, anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
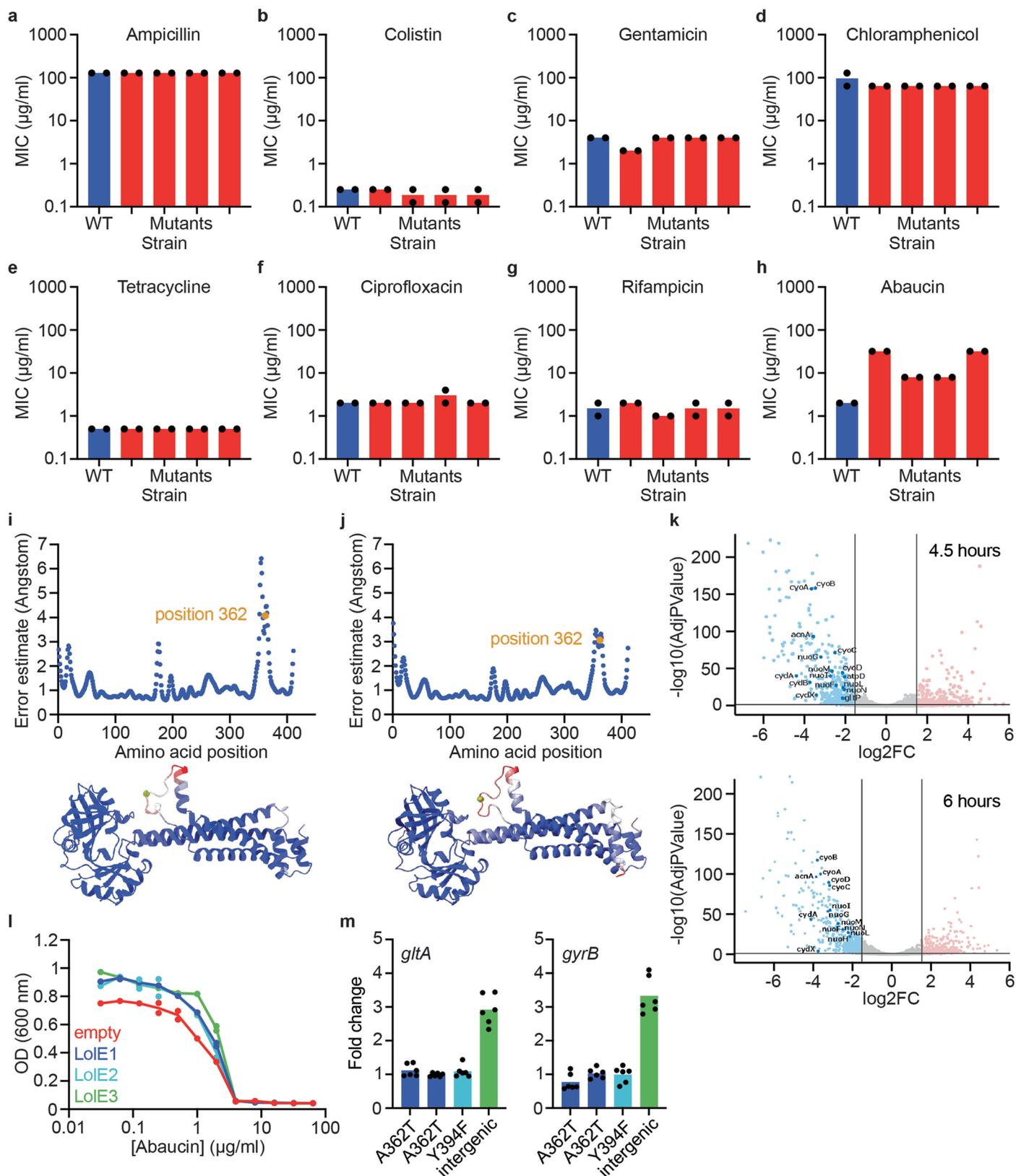
Extended Data Fig. 1 | Model training data and prediction. (a) Replicate plot showing primary screening data of 7,684 small molecules for those that inhibited the growth of *A. baumannii* ATCC 17978 in LB medium at 50 μM. **(b)** Rank-ordered growth inhibition data of the prioritized 240 molecules from our prediction set that were selected for empirical validation (top); rank-ordered growth inhibition data of the 240 predicted molecules with the lowest prediction score (middle); rank-ordered growth inhibition data of the 240 predicted molecules with the highest prediction score that were not found in the training dataset (bottom). Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. Dashed horizontal line represents the stringent hit cut-off of >80% growth inhibition at 50 μM. **(c)** Growth inhibition of *A. baumannii* by abaucin (blue) and serdemetan (red) in LB medium. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. The structure of serdemetan is shown. **(d)** Growth kinetics of *A. baumannii* cells after treatment with abaucin at varying concentrations for 6 hours. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted.

Extended Data Fig. 2 | Antibacterial activity of abaucin against human commensal species. (**a**) Growth inhibition of *A. baumannii* ATCC 17978 by ampicillin (blue) and ciprofloxacin (red) in LB medium. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. (**b**) Growth inhibition of *B. breve* by abaucin. Experiments were

conducted in biological duplicate. (**c**) Growth inhibition of *B. longum* by abaucin. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. (**d**) Non-validated (see Fig. 2e) growth inhibition of *E. lenta* by abaucin. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted.

Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Abaucin mechanism of action. (a–h) Growth inhibition of wildtype *A. baumannii* (WT) and the four independent abaucin-resistant mutants by a collection of diverse antibiotics. From left to right for each plot, the mutants are: A362T variant 1, Y394F, intergenic, and A362T variant 2. Experiments were conducted in biological duplicate. Note that the abaucin-resistant mutants do not display cross-resistance to other antibiotics. **(i)** Structural prediction of wildtype *A. baumannii* LolE using RoseTTAFold (bottom), with the structural error estimate of each amino acid (top). Position 362 is highlighted orange and resides in a disordered region of the protein. **(j)** same as (i), except with the Y362T abaucin-resistant mutant of LolE. **(k)** RNA sequencing of wildtype *A. baumannii* treated with 5x MIC abaucin for 4.5 hr (top) or 6 hr (bottom). Data are the mean of biological duplicates. Transcript abundance is normalized to no-drug control cultures grown in identical conditions. Vertical black lines show statistical significance cut-off values. Note the highly significant downregulation of genes involved in the electron transport chain and transmembrane ion transport. **(l)** Growth inhibition of *A. baumannii* harboring an empty CRISPRi vector (red), or three distinct sgRNAs targeting *lolE* (blue, teal, and green). All strains were grown in LB medium without induction. Experiments were conducted in biological duplicate. Individual replicates with means connected are plotted. **(m)** qPCR quantifying the expression of *lolE* relative to the housekeeping gene *gltA* (left) and *gyrB* (right) in all four abaucin resistant mutants, normalized to wildtype *A. baumannii*. Experiments were conducted in biological duplicate with technical triplicates. Bar height represents mean expression.

Corresponding author(s): Jonathan M. Stokes and James J. Collins

Last updated by author(s): April 14, 2023

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All custom Chemprop code for molecular property prediction (antibiotic activity against A. baumannii) is freely available to download from https://github.com/chemprop/chemprop. Additionally, a snapshot of the version used in this paper is preserved and freely available from https://github.com/GaryLiu152/chemprop_abaucin
Data analysis	Molecular feature were calculated using the open-source package RDKit (https://www.rdkit.org) (version 2020.03.1.0). t-SNE plots were generated using scikit-learn (https://scikit-learn.org/stable/) (version 0.22.2.post1). SNP identification was conducted using BreSeq (https://barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing) (v0.35.0). kallisto (https://pachterlab.github.io/kallisto/about) (v0.46.1) was used to compute mRNA abundances for RNA sequencing. DESeq2 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) was used to quantify drug-induced mRNA expression changes. EcoCyc (https://ecocyc.org) (v24.1) was used for GO term enrichment analyses. sgRNA design was performed using code from (https://gitlab.pasteur.fr/dbikard/ecowg1). A. baumannii LolE structural predictions were computed using RoseTTAfold through the Robetta platform (https://robetta.bakerlab.org/). Structural alignments were conducting using Maestro (version 12.9.137) and and ChimeraX (version 1.3rc202112020528).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

GenBank accession numbers for sequencing of abaucin resistant mutants are: BankIt2629921 – OP677864, OP677865, OP677866, OP677867. GEO accession numbers for RNA sequencing datasets are GSE214305 – GSM6603484, GSM6603485, GSM6603486, GSM6603487, GSM6603488, GSM6603489, GSM6603490.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Mouse experiment sample size was determined using field convention group sizes for antibiotic efficacy evaluation (n=5 or n=6) (for example, see Stokes et al., 2020). All in vitro experiments were conducted using biologically independent replicates, also based on field convention. Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4). https://doi.org/10.1016/j.cell.2020.01.021
Data exclusions	No data was excluded from any analyses.
Replication	All in vitro experiments were conducted using biologically independent replicates. We define biological independence as experimental replicates being conducted using identical protocols on different days.
Randomization	Randomization was only applicable to mouse studies . Before infection, mice were relocated at random from housing cages to single-occupancy control or treatment cages. For in vitro antibiotic testing, randomization is not necessary since the same bacterial culture is aliquoted and each aliquot received the desired treatment with the appropriate controls running side-by-side.
Blinding	Investigators were not blinded for any in vitro or animal experiments. We were analyzing the antibacterial properties of a new molecule against A. baumannii. Many of the analyses performed pertained to antibacterial efficacy, which is quantifiable using common laboratory approaches. Therefore, blinding was considered unnecessary.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	Human research participants
☒	Clinical data
☒	Dual use research of concern

Methods

n/a	Involved in the study
☒	ChIP-seq
☒	Flow cytometry
☒	MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Six- to eight-week-old female C57BL/6N mice were purchased from Charles River Laboratories. Animals were housed in a specific

Laboratory animals	pathogen free barrier unit under Level 2 conditions, temperature controlled (21℃) 30-50% humidity, 12h light and dark cycle environment (dark from 7pm to 7am) and were fed regular chow ad libtum.
Wild animals	No wild animals were used in the study.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	Mouse model experiments were conducted according to the guidelines set by the Canadian Council on Animal Care, using protocols approved by the Animal Review Ethics Board and McMaster University under Animal Use Protocol #20-12-41.

Note that full information on the approval of the study protocol must also be provided in the manuscript.