

1 SyntheMol-RL: a flexible reinforcement learning framework for designing novel and
2 synthesizable antibiotics

3

4 **Authors**

5 Kyle Swanson ^{1,*}, Gary Liu ^{2,*}, Denise B. Catacutan ^{2,*}, Stewart McLellan ², Autumn Arnold ²,
6 Megan M. Tu ², Eric D. Brown ², James Zou ^{1,3,**}, Jonathan M. Stokes ^{2,**}

7

8 **Affiliations**

9 ¹ Department of Computer Science, Stanford University, Stanford, CA, USA

10 ² Department of Biochemistry and Biomedical Sciences, Michael G. DeGroote Institute of
11 Infectious Disease Research, David Braley Centre for Antibiotic Discovery, McMaster
12 University, Hamilton, Ontario, Canada

13 ³ Department of Biomedical Data Science, Stanford University, Stanford, CA, USA

14 *These authors contributed equally

15 **Correspondence: stokesjm@mcmaster.ca, jamesz@stanford.edu

16

17 **Abstract**

18 The rise of antibiotic-resistant pathogens such as *Staphylococcus aureus* has created an urgent
19 need for new antibiotics. Generative artificial intelligence (AI) has shown promise in drug
20 discovery, but existing models often fail to propose compounds that are both effective and
21 synthetically tractable. To address these challenges, we introduce SyntheMol-RL, a
22 reinforcement learning-based generative model that can rapidly design synthetically accessible
23 small molecule drug candidates from a massive chemical space of 46 billion compounds.

24 SyntheMol-RL improves upon our prior Monte Carlo Tree Search (MCTS)-based SyntheMol
25 model by generalizing across chemically similar building blocks and enabling multi-parameter
26 optimization. We applied SyntheMol-RL to generate candidate antibiotics against *S. aureus* by

27 optimizing for both antibacterial activity and aqueous solubility, and we found that SyntheMol-RL
28 generated molecules with improved predicted properties compared to both the previous MCTS
29 version of SyntheMol as well as an AI-based virtual screening baseline. We synthesized 79
30 SyntheMol-RL compounds that were unique relative to the training dataset and found that 13
31 showed potent *in vitro* activity, of which seven were structurally novel after detailed literature
32 searches. Furthermore, one hit compound, synthecin, demonstrated efficacy in a murine wound
33 infection model of methicillin-resistant *S. aureus* (MRSA). These results validate SyntheMol-
34 RL's ability to generate novel and synthetically accessible candidate antibiotics and position
35 SyntheMol-RL as a powerful tool for drug design across therapeutic domains.

36

37 **Introduction**

38 The rapid spread of antibiotic resistance is a critical challenge facing modern medicine, with
39 global implications for public health. In 2019, approximately 4.95 million deaths were linked to
40 infections caused by drug-resistant bacteria, and this number is expected to rise to 10 million
41 annually by 2050 if the emergence of antimicrobial resistance (AMR) continues to outpace the
42 development of new antibiotics¹. Among the most concerning are the ESKAPE pathogens, six
43 bacterial species notorious for their virulence and resistance to multiple drugs². Among these,
44 *Staphylococcus aureus* stands out as a significant cause of morbidity and mortality in both
45 community and healthcare settings^{3,4}. In high-income regions, *S. aureus* is the leading cause of
46 deaths associated with AMR (25.4% of such deaths), and globally, in 2019, over 700,000
47 deaths associated with AMR were linked to *S. aureus* infections¹. Furthermore, of drug-resistant
48 pathogens, methicillin-resistant *S. aureus* (MRSA) had the greatest increase in attributable
49 burden from 1990 to 2021, doubling from 57,200 to 130,000 attributable deaths over this period.
50 The World Health Organization highlights *S. aureus* as a high priority for the development of
51 new antibiotics, particularly due to the dwindling number of treatment options available for

52 antibiotic-resistant *S. aureus* infections⁵. Thus, there is an urgent need for novel therapeutic
53 approaches to combat *S. aureus* and its increasing antibiotic resistance.

54

55 Artificial intelligence (AI) is playing an important role in drug development broadly and in
56 antibiotic discovery in particular^{6–9}. One AI method commonly employed in drug discovery is
57 property prediction models, which are trained to predict the properties of molecules such as
58 antibacterial activity or aqueous solubility¹⁰. These models can be applied in a virtual
59 screening¹¹ approach to evaluate large chemical libraries to identify compounds that are
60 predicted to be effective. However, since property prediction models must evaluate molecules
61 one-by-one, they do not scale well to the chemical spaces of billions of molecules that are
62 increasingly used in industrial and academic settings. Indeed, in recent years, multiple ultra-
63 large synthetically accessible chemical databases have been assembled, including Enamine's
64 REAL Space, WuXi's GalaXi, and OTAVA's CHEMriya, all exceeding 10^{10} molecules.
65 Therefore, more recently, generative AI methods have been developed to directly design
66 molecules with promising properties without a slow evaluation process^{12–15}.

67

68 Generative AI methods for drug discovery are promising, but a major limitation of many of these
69 methods is the synthesizability of the compounds proposed by these models¹⁶. Although many
70 of the AI-generated molecules are predicted to be effective, they are often synthetically
71 intractable, thus preventing experimental validation and limiting the real-world impact of these
72 models. To address this challenge, we previously developed SyntheMol¹⁷, a generative AI
73 method for drug design that uses a Monte Carlo tree search (MCTS) guided by a trained
74 antibacterial activity property predictor to explore a space of 30 billion easy-to-synthesize
75 compounds. This method enabled us to rapidly synthesize 58 SyntheMol-generated compounds
76 and experimentally validate their antibacterial efficacy, with six compounds proving to have high
77 potency against *Acinetobacter baumannii* and several other ESKAPE species.

78

79 SyntheMol successfully designed compounds with *in vitro* antibacterial activity, but two
80 important limitations were observed. First, the MCTS algorithm that powers SyntheMol is not the
81 most effective method for exploring chemical space. Indeed, SyntheMol builds compounds by
82 selecting molecular building blocks and combining them with a pre-defined set of chemical
83 reactions to form molecules, and MCTS is responsible for scoring and selecting the most
84 promising building blocks. However, MCTS treats every building block independently and
85 cannot learn patterns of antibacterial activity that are shared across chemically similar building
86 blocks, thereby limiting its search efficiency. Second, SyntheMol can only optimize a single
87 molecular property at a time, even though real-world drug discovery requires identifying
88 compounds with many druglike properties simultaneously. This proved to be a particular
89 challenge in our prior work, where only two of our six potent antibacterial compounds were
90 sufficiently soluble to be administered for *in vivo* toxicity experiments.

91

92 In this work, we introduce SyntheMol-RL, a new and significantly improved version of SyntheMol
93 that overcomes the two limitations of the original version of SyntheMol (hereafter referred to as
94 SyntheMol-MCTS). First, SyntheMol-RL replaces the MCTS algorithm with a reinforcement
95 learning model that can generalize across chemically similar building blocks to more rapidly and
96 effectively search massive combinatorial chemical spaces for promising, easy-to-synthesize
97 compounds. Second, SyntheMol-RL introduces the ability to perform multi-parameter
98 optimization to generate compounds that simultaneously possess multiple relevant molecular
99 properties for drug discovery applications. Moreover, in this work we expand the chemical space
100 SyntheMol-RL explores from 30 billion compounds to 46 billion compounds.

101

102 We then apply SyntheMol-RL to design easily synthesizable compounds against the target
103 bacterium *S. aureus* with simultaneous optimization for antibacterial activity against *S. aureus*

104 and aqueous solubility (Fig. 1). Importantly, SyntheMol-RL markedly outperforms SyntheMol-
105 MCTS and a virtual screening approach that uses property prediction models *in silico*, finding
106 more compounds that are novel, diverse, and predicted to be antibacterial. Next, we
107 synthesized and experimentally tested 79 compounds designed by two variants of SyntheMol-
108 RL and found two and 11 potent hits, respectively, compared to zero hits for SyntheMol-MCTS
109 and two hits for virtual screening. We further investigated one particularly potent and structurally
110 novel *de novo* generated compound, which we call synthecin, and we demonstrated its ability to
111 fully arrest growth of MRSA in a murine wound infection model. These results demonstrate that
112 SyntheMol-RL is an effective and flexible framework for novel drug design applications.

113

114 **Results**

115 *Property Prediction Model Development*

116 SyntheMol-RL is guided by molecular property prediction models that evaluate the properties of
117 generated molecules and thereby provide feedback to improve the generative model's ability to
118 design molecules with desired properties. Therefore, the first step toward building SyntheMol-
119 RL is to train molecular property prediction models on the properties that we require of our
120 molecules. Here, we focus on two key properties: (1) antibacterial activity, which we define here
121 as the ability of a compound to effectively inhibit the growth of *S. aureus* *in vitro*, and (2)
122 aqueous solubility, which is necessary for the compound to be efficiently delivered and tested *in*
123 *vivo*. For antibacterial activity, we performed our own assay to build a training set, as described
124 below. For aqueous solubility, we took the AqSolDB dataset of 9,982 molecules with log
125 solubility values as curated by the Therapeutics Data Commons¹⁸. While there are many drug-
126 like properties that we could prioritize at this stage, we chose aqueous solubility to help
127 circumvent issues of low solubility that we observed when we previously applied SyntheMol-
128 MCTS to antibiotic discovery¹⁷.

129

130 To create our antibacterial activity training set, we physically screened 10,716 compounds
131 against *S. aureus* RN4220 at 50 μ M in LB medium (Fig. 2a and Supplementary Data 1)¹⁹.
132 Experiments were conducted in biological duplicate, with end-point growth being measured at
133 600 nm optical density (OD_{600}) after 16 hours of incubation at 37°C (Extended Data Fig. 1a). We
134 computed the mean μ and standard deviation σ OD_{600} value across the library and used $\mu - 2\sigma$
135 as a threshold for binarizing the OD_{600} values into active (low OD_{600}) and inactive (high OD_{600})
136 molecules. After removing duplicate compounds based on SMILES (see Methods), we obtained
137 a set of 10,658 unique molecules with 1,137 active compounds (10.7%) and 9,521 (89.3%)
138 inactive compounds. A t-SNE visualization of our training dataset compounds and 1,007
139 molecules from ChEMBL²⁰ with known antibacterial activity (Supplementary Data 2) shows that
140 our active compounds cover both known and novel antibiotic chemical space (Fig. 2b).

141
142 After acquiring these two training datasets, we leveraged two different property prediction model
143 architectures for each dataset: (1) Chemprop-RDKit and (2) MLP-RDKit (Fig. 2c). The
144 Chemprop-RDKit architecture consists of the graph neural network (GNN) model Chemprop¹⁰
145 augmented with 200 molecular features computed by the cheminformatics package RDKit²¹.
146 Chemprop-RDKit combines the GNN representation of the molecule with the 200 RDKit features
147 using a multilayer perceptron (MLP). The MLP-RDKit architecture is nearly identical to the MLP
148 contained within the Chemprop-RDKit model, but MLP-RDKit only uses the 200 RDKit features
149 as input and does not have a GNN component. We chose to experiment with both model
150 architectures to compare the importance of accuracy and speed; Chemprop-RDKit learns a
151 more comprehensive molecular representation and is thus potentially more accurate, while
152 MLP-RDKit is potentially less accurate but is faster without the GNN component.

153
154 We trained Chemprop-RDKit and MLP-RDKit on both the *S. aureus* antibacterial activity and
155 aqueous solubility datasets. For each of the four model-dataset combinations, we trained the

156 model using 10-fold cross-validation with splits containing 80% train, 10% validation, and 10%
157 test data. Each of the four model-dataset combinations trained in less than 80 minutes on a
158 machine with 8 CPUs and 1 GPU. The Chemprop-RDKit models slightly outperformed the MLP-
159 RDKit models, although both showed high performance (Fig. 2d, Table 1, and Supplementary
160 Data 3), including on more challenging splits based on molecular scaffold (Extended Data Fig.
161 1b-c).

162

163 **Table 1. Property Prediction Model Performance**

Model	Antibiotic ROC-AUC	Antibiotic PRC-AUC	Solubility MAE	Solubility R ²
Chemprop-RDKit	0.875 +/- 0.013	0.570 +/- 0.050	0.656 +/- 0.023	0.822 +/- 0.021
MLP-RDKit	0.873 +/- 0.018	0.553 +/- 0.040	0.688 +/- 0.020	0.817 +/- 0.013

164

165 *Generative Model Development*

166 SyntheMol¹⁷ is a generative model that designs easily synthesizable molecules with desirable
167 properties, such as inhibition of bacterial growth *in vitro*. SyntheMol ensures straightforward
168 synthesis of generated molecules by constraining the generative process to a combinatorial
169 chemical space, where it builds molecules using small molecular building blocks that are readily
170 purchasable and chemical reactions that are well-validated. To generate a molecule, SyntheMol
171 selects two or three building blocks and combines them using a chemical reaction to form a
172 molecule, which is then scored by a property prediction model. This score provides feedback to
173 SyntheMol to guide its generative process toward higher scoring molecules, resulting in a set of
174 generated molecules that are easy to synthesize and likely to possess the desired properties.
175 We previously applied SyntheMol to generate antibiotic candidates against the gram-negative
176 bacterium *A. baumannii*, and we successfully synthesized 58 generated compounds and

177 identified six highly potent hits, demonstrating the ability of SyntheMol to design effective
178 antibacterial compounds.

179

180 In its original form, SyntheMol employed a Monte Carlo tree search (MCTS) algorithm^{22,23} to
181 guide its selection of building blocks and reactions to design molecules. SyntheMol-MCTS treats
182 the generative process as a tree search, where nodes in the tree are combinations of one or
183 more building blocks. SyntheMol-MCTS computes a value for each node based on statistics
184 about the node that measure both exploration (how many times the node has been selected)
185 and exploitation (the average property score of molecules built using that node). During
186 generation, SyntheMol-MCTS sequentially explores nodes with one, two, or three building
187 blocks, at each step choosing the node with the highest value. When two or three building
188 blocks in a node are compatible with one of the specified chemical reactions, a new node is
189 created with the product of the chemical reaction applied to those building blocks. Once a node
190 with a product molecule is selected, SyntheMol-MCTS updates its statistics for each node
191 involved in generating that molecule based on the property prediction score of the molecule,
192 concluding one rollout. While MCTS is a simple and effective algorithm, it suffers from a key
193 inefficiency by treating each node independently rather than leveraging the chemical similarity
194 between molecular building blocks in different nodes to determine the nodes' values.

195

196 Here, we developed a new version of SyntheMol that overcomes this inefficiency by replacing
197 MCTS with a reinforcement learning (RL) algorithm^{24–27} (Fig. 1, see Methods). SyntheMol-RL
198 uses a deep learning model as a value function that determines the value of a node based on
199 the chemical structures of the building block(s) in that node. Specifically, the RL value function
200 learns to predict the expected property score of molecules constructed using the building
201 block(s) in a given node. During each step of generation, SyntheMol-RL computes the value of
202 nodes containing one, two, or three building blocks. SyntheMol-RL then samples a node

203 proportional to the node's value, which encourages exploration (by not always selecting the
204 highest value node) and exploitation (by sampling higher value nodes more frequently). To
205 simultaneously optimize for multiple properties (in our case, antibacterial activity and aqueous
206 solubility), the value of a node is computed as a weighted combination of multiple RL models –
207 one for each property. SyntheMol-RL automatically adjusts the weights dynamically over time to
208 explore the Pareto frontier of the multiple properties and identify where the greatest density of
209 'hit' molecules exist (those that pass user-defined thresholds for all properties). The RL models
210 are trained during the generative process by taking nodes and molecules from previous
211 iterations and training the RL models to predict the property scores of each molecule from the
212 nodes that created it. Thus, as the RL model generates more molecules and creates more
213 training data, the accuracy of the RL models and the quality of the generated molecules
214 improve.

215

216 For antibiotic generation with SyntheMol-RL, we used our trained Chemprop-RDKit models for
217 antibacterial activity and aqueous solubility as fixed property predictors to score designed
218 molecules since those models were the most accurate. For the RL value function, we used
219 either a Chemprop-RDKit model architecture or an MLP-RDKit model architecture to evaluate
220 tradeoffs between accuracy and speed. In both cases, we initialized the RL value models with
221 the parameters from our trained property prediction models, but then we further trained them
222 during generation on the RL task of predicting the value of nodes (i.e., combinations of building
223 blocks).

224

225 We provided SyntheMol-RL with molecular building blocks and chemical reactions from two
226 easily synthesizable combinatorial chemical spaces: the Enamine REAL Space and the WuXi
227 GalaXi (Fig. 2e, Extended Data Fig. 1d-e). The Enamine REAL Space²⁸ (2022 version) contains
228 31 billion molecules that can be produced using ~139,000 molecular building blocks and 169

229 chemical reactions (Supplementary Data 4). We used 13 of the most common reactions, which
230 can build 30 billion molecules. The WuXi GalaXi²⁹ (2022 version) contains 16 billion molecules
231 that can be produced with ~15,000 building blocks and 36 chemical reactions (Supplementary
232 Data 5). Since these two chemical spaces are relatively distinct (Fig. 2b), this gives SyntheMol-
233 RL the capability to generate a structurally diverse set of easily synthesizable molecules within a
234 chemical space of 46 billion molecules.

235

236 *Generative Model Application*

237 We applied SyntheMol-RL to generate antibiotic candidates against *S. aureus*. We ran two
238 versions of SyntheMol-RL: (1) RL-Chemprop, which uses a Chemprop-RDKit model architecture
239 as the value function, and (2) RL-MLP, which uses an MLP-RDKit model architecture as the
240 value function. In both cases, we ran the generative model for 10,000 rollouts, resulting in the
241 generation of roughly 10,000 unique molecules. To demonstrate the improved efficacy of RL
242 over MCTS, we also ran MCTS for 10,000 rollouts. Furthermore, to illustrate the benefit of using
243 generative modeling over virtual screening approaches, we applied our Chemprop-RDKit
244 property prediction models in a virtual screening¹¹ manner (VS-Chemprop) to evaluate the
245 antibacterial activity and aqueous solubility of 21 million randomly sampled molecules (14
246 million REAL and 7 million GalaXi to match the relative sizes of the chemical spaces). We chose
247 21 million since it took VS-Chemprop about 7 days to evaluate that many compounds, which
248 roughly matched the time of our slowest generative model, RL-Chemprop, on the slowest of our
249 GPUs and CPUs. This gives VS-Chemprop the best possible chance to discover promising
250 compounds compared to the generative models, normalized for time.

251

252 We first evaluated the compounds generated by RL-Chemprop, RL-MLP, and MCTS, as well as
253 those screened by VS-Chemprop, based on their predicted antibacterial activity and aqueous
254 solubility. Defining a hit as any molecule with an antibacterial prediction score ≥ 0.5 and a log

255 solubility ≥ -4 (which constitutes high solubility³⁰), RL-Chemprop generated 11.6% hits (1,273 of
256 10,983; Supplementary Data 6) and RL-MLP generated 5.3% hits (493 of 9,228; Supplementary
257 Data 7) compared to 3.0% hits for MCTS (347 of 11,630; Supplementary Data 8) and 0.006%
258 hits for VS-Chemprop (1,290 of 21,000,000; Supplementary Data 9) (Fig. 3a). The distributions
259 of antibacterial prediction scores (Fig. 3b) and log solubility prediction scores (Fig. 3c) show that
260 all methods can similarly identify soluble molecules, but the generative models – and particularly
261 the RL models – excel at designing molecules with high predicted antibacterial activity. This
262 illustrates the benefit of using RL for molecule design. Of note, we perform a set of ablation
263 experiments that elucidate the effect of individual components of the RL model on generation
264 performance (see Methods and Extended Data Fig. 2).

265
266 To obtain a set of promising compounds for synthesis and testing *in vitro*, we established a set
267 of four filters to prioritize compounds. First, as above, we identified hits and only kept molecules
268 with an antibacterial prediction score ≥ 0.5 and a log solubility ≥ -4 . Second, we selected novel
269 hits by computing the Tversky similarity³¹ between each hit molecule and all known antibiotics in
270 our training set and the ChEMBL antibiotics set, only keeping molecules with a maximum
271 Tversky similarity of ≤ 0.6 . Third, we selected a diverse set of hits by computing the largest set
272 of novel hits such that no two compounds in that set had a Tanimoto similarity³² greater than 0.6
273 (see Methods). Fourth, to limit the diverse, novel hits to a practical number of compounds to test
274 in the wet lab, we ranked the remaining compounds by antibacterial prediction score and
275 selected the top 150 compounds from each method. This gave us 600 molecules: 150 each
276 from RL-Chemprop, RL-MLP, MCTS, and VS-Chemprop. Finally, to provide an unbiased
277 estimate of the baseline antibacterial activity and solubility of compounds in the REAL and
278 GalaXi chemical spaces, we selected 150 random compounds (100 REAL, 50 GalaXi;
279 Supplementary Data 10), for a total of 750 molecules for testing (Fig. 3d).

280

281 Among these 750 selected compounds, the RL-Chemprop and RL-MLP compounds had higher
282 average antibacterial prediction scores (median of 0.758 and 0.712, respectively) compared to
283 MCTS, VS-Chemprop, and random (median of 0.605, 0.694, and 0.007, respectively) (Fig. 3e)
284 while all of the compounds had similar predicted log solubilities (Fig. 3f). This computational
285 evidence shows that SyntheMol-RL is the most adept at generating likely antibiotic candidates.

286

287 We then requested the availability of these 750 compounds (747 unique compounds due to
288 three duplicates between RL-MLP and MCTS) from Enamine (659 unique compounds) and
289 WuXi (88 unique compounds), depending on the chemical space each compound belongs to.
290 Among those compounds, 435 / 659 (66%) of the Enamine compounds and 49 / 88 (56%) of the
291 WuXi compounds were available for potential synthesis (Supplementary Data 11). Notably, RL-
292 Chemprop, RL-MLP, and MCTS demonstrated a strong preference for Enamine compounds
293 and selected very few WuXi compounds (3, 1, and 0, respectively), of which none were
294 available for synthesis, so WuXi compounds were only available for VS-Chemprop and random.
295 This result may be because the Enamine REAL Space has roughly 10x the number of building
296 blocks as the WuXi GalaXi, despite having only 3x as many total molecules. Since SyntheMol
297 makes building block-level decisions, it has disproportionately more opportunities to select
298 Enamine building blocks.

299

300 To further narrow down these available compounds to the most promising set of potential
301 antibiotic candidates, we ranked the available compounds by predicted clinical toxicity according
302 to ADMET-AI^{33,34}, a multi-task Chemprop-RDKit model for ADMET property prediction. We then
303 selected the 50 compounds with lowest predicted clinical toxicity for each model type (RL-
304 Chemprop, RL-MLP, MCTS, and VS-Chemprop). For the random baseline, we randomly
305 selected 50 available compounds (33 REAL, 17 GalaXi) from the random set.

306 Altogether, we submitted 250 compounds (50 compounds for each of five settings, 248 unique
307 compounds total with two duplicates between RL-MLP and MCTS) for synthesis (Fig. 3g,
308 Supplementary Data 11). Enamine successfully synthesized 177 / 224 (79%) unique
309 compounds, while WuXi successfully synthesized 17 / 24 (71%) unique compounds. Altogether,
310 we obtained 194 unique compounds for experimental validation (38 RL-Chemprop, 41 RL-MLP,
311 38 MCTS, 37 VS-Chemprop, and 44 random).

312

313 A final computational evaluation was conducted prior to *in vitro* testing, which benchmarked
314 SyntheMol-RL against GFlowNet^{35,36}, a state-of-the-art-generative model for drug design. While
315 the GFlowNet architecture generated molecules with higher antibacterial activity and aqueous
316 solubility scores than the SyntheMol-RL molecules, the GFlowNet compounds are bulky with
317 complex multi-ring structures and would likely suffer from poor whole cell activity and
318 synthesizability (Extended Data Fig. 3, Supplementary Data 12). Additionally, medicinal
319 chemists at Enamine and WuXi offered limited synthesis of these molecules with >30x higher
320 associated costs (see Extended Discussion). This comparison illustrates the advantages of
321 SyntheMol-RL for molecule generation.

322

323 *In Vitro Validation of AI-Generated Molecules*

324 We used standard growth inhibition assays to determine all synthesized molecules' antibacterial
325 activity against *S. aureus* RN4220, defined using the minimum inhibitory concentration (MIC).
326 Remarkably, 11 / 38 (29%) molecules synthesized from RL-Chemprop and 2 / 41 (5%) from RL-
327 MLP displayed the ability to completely inhibit bacterial growth at MIC ≤ 8 µg/ml (Fig. 4a). For
328 comparison, our control set of 44 randomly selected molecules synthesized from both the REAL
329 space and GalaXi were also tested. We observed a complete absence of any activity up to our
330 limit of detection of 128 µg/ml. This may be evidence that the ultra-large chemical databases are

331 non-conducive for antibiotic discovery screens and that common benchmarks from existing
332 physical chemical libraries (~1% hit rate for a standard HTS screen) may not be applicable.
333
334 For all compounds deemed a “hit” as defined by MIC ≤ 8 µg/ml, an in-depth manual literature
335 search was conducted to ensure structural novelty from reported antibacterial activity from
336 literature. We note that this manual literature search was performed after initial *in vitro* testing to
337 enable maximally detailed literature analyses on a small number of potent molecules. Of the 13
338 identified hits from both SyntheMol-RL models, six compounds were found to contain or have
339 structural overlap with salicylanilides³⁷, a small structural class with known antibacterial activity
340 against Gram-positive bacteria. Specifically, niclosamide^{38,39}, a salicylanilide derivative, is
341 currently being investigated for topical treatment of *S. aureus* and is currently in Phase II clinical
342 trials. A similar literature search was conducted for the two hit molecules sourced from VS-
343 Chemprop; both compounds were found to be similar to molecules with antibacterial activity in
344 the literature. All molecules that failed the manual novelty check were removed from further
345 investigations, leaving six novel compounds from RL-Chemprop, one novel compound from RL-
346 MLP, and no novel compounds from VS-Chemprop (Fig. 4b).
347
348 The remaining seven compounds were assessed for their respective spectrum activity against a
349 phylogenetically diverse panel of ESKAPE isolates^{2,5}, a group of highly virulent and antibiotic-
350 resistant Gram-positive and Gram-negative bacteria that often cause challenging hospital-
351 acquired infections. Molecules displayed no significant potency across all the pathogens (Fig.
352 4c), including the other Gram-positive pathogen *E. faecium*, indicating all SyntheMol-RL
353 compounds are primarily narrow-spectrum against *S. aureus*.
354
355 Importantly, however, the laboratory strain *S. aureus* RN4220 applied until this point in the study
356 may not be representative of the strains commonly responsible for infection. Therefore, we

357 conducted growth inhibition assays against a set of antibiotic-resistant *S. aureus* strains⁴⁰
358 especially relevant to the clinic. We screened all seven hit compounds against the most
359 prevalent community-associated MRSA strain for skin and soft tissue infections (SSTIs),
360 USA300, as well as a collection of multidrug-resistant vancomycin-intermediate *S. aureus*
361 (VISA) isolates from the CDC AR Isolate Bank⁴¹, which covers all resistance mechanisms found
362 within the VISA panel (Fig. 4c). For all 7 compounds, no significant loss of antibacterial potency
363 was found, with all compounds retaining a potent MIC ≤ 8 µg/ml. This provides strong evidence
364 that all seven molecules are able to overcome a wide array of prevalent antibiotic resistance
365 determinants, likely owing to their structural novelty relative to clinical antibiotics.

366

367 *In Vivo Validation of SyntheMol-RL-Generated Molecules*

368 *S. aureus* causes most SSTIs in humans, including infected wounds and ulcers, cellulitis, and
369 folliculitis⁴⁰. Not only is MRSA prevalent in hospitals globally, but community-associated MRSA
370 (CA-MRSA) is responsible for a significant proportion of SSTIs in Europe (<1% to 32%), Asia
371 (~17%), and the United States (>50%)⁴. SSTIs caused by virulent *S. aureus* (including MRSA) is
372 a risk factor for invasive infections^{40,42}, leading to conditions such as bacteremia and
373 endocarditis, which when caused by *S. aureus* are associated with mortality rates of 15% and
374 30-40% respectively⁴³. Due to the burden MRSA imposes, we elected to use the most prevalent
375 CA-MRSA isolate in the United States, USA300, as we continued to investigate the
376 translatability of the hit molecules⁴.

377

378 Given their potent *in vitro* activities, we evaluated whether each of the seven hit molecules
379 generated by SyntheMol-RL may be amenable to formulation to treat MRSA infected wounds.
380 To investigate this possibility, we developed the Chemical Release Evaluation on Agar Media
381 (C.R.E.A.M.) assay, in which a control (10% DMSO) and test molecules (2% w/v) were
382 formulated in Glaxal Base for *in vivo* application and subsequently applied to MRSA-inoculated

383 LB agar plates, prior to overnight incubation at 37°C (Fig. 5a). This *in vivo* mimicking model
384 serves to evaluate the formulation's efficacy as a topical antibacterial treatment. It is worth
385 noting that all seven compounds were readily soluble in the Glaxal Base solution, indicating that
386 the *in silico* optimization of solubility along with antibacterial activity during generation helped
387 remove potential barriers at this stage in development. We observed that MLP-01 resulted in the
388 largest zone of growth inhibition, indicating that it may demonstrate the most promising *in vivo*
389 activity in a mouse wound infection model. Interestingly, despite similar growth inhibitory activity
390 in liquid media, there is almost a 2-fold difference in area between MLP-01 and Chemprop-03,
391 showcasing the imperfect correlation between liquid MIC values and growth inhibitory effect
392 when formulated into a cream for *in vivo* applications. Given its potency and success in the
393 C.R.E.A.M assay, we continued to characterize the activity of MLP-01, which we named
394 synthecin due to its origin of discovery. Synthecin was tested to determine whether its activity
395 was bactericidal or bacteriostatic against *S. aureus* USA300. We observed bacteriostatic activity
396 in LB medium at 2x and 4x MIC (Fig. 5b). For reference, linezolid, a commonly prescribed
397 antibiotic for the treatment of *S. aureus* skin infections, is also bacteriostatic⁴⁴.

398

399 Based on these encouraging *in vitro* results, we tested whether synthecin retains its
400 antibacterial efficacy against MRSA USA300 in a mouse wound infection model, a common
401 infection caused by MRSA. To test the *in vivo* efficacy of synthecin, we established a wound
402 infection in cyclophosphamide pre-treated C57BL/6N mice using *S. aureus* USA300 (~2.15 ×
403 10⁷ CFU) (Fig. 5c). The *S. aureus* infection was allowed to establish for 1 h prior to topical
404 treatments at 1, 4, 8, 12, and 20 h post-infection (hpi) with Glaxal Base supplemented with
405 vehicle (10% DMSO) or synthecin (2% w/v). Mice were then euthanized 24-hpi, and the skin
406 was aseptically dissected and plated to quantify *S. aureus* burden. Vehicle-treated mice carried
407 ~6.39 × 10² CFU/g at the experimental endpoint, with wounded tissues exhibiting significant
408 inflammation (Fig. 5d-e, Extended Data Fig. 4). In contrast, mice treated with synthecin showed

409 a significantly lower bacterial burden, with $\sim 5.14 \times 10^0$ CFU/g, similar to the bacterial burden
410 observed in pre-treated infection control mice (Supplementary Table 1). The synthecin treated
411 tissues also displayed no signs of inflammation (Fig. 5e). Collectively, these data demonstrate
412 that synthecin can effectively suppress an MRSA wound infection, aligning with its exceptional
413 *in vitro* activity in the C.R.E.A.M. assay.

414

415 **Discussion**

416 SyntheMol-RL is an effective generative AI model for designing easily synthesizable small
417 molecule drug candidates. SyntheMol-RL's use of reinforcement learning enables it to rapidly
418 explore massive combinatorial chemical spaces with tens of billions of molecules for promising
419 compounds that are easily synthesizable by design. Furthermore, the multiparameter
420 optimization abilities of SyntheMol-RL allow it to identify compounds that simultaneously
421 possess multiple drug-like properties – in our case, antibacterial activity and aqueous solubility –
422 which is a necessity in real-world drug development applications. When applied to antibiotic
423 discovery, SyntheMol-RL generated a superior set of antibiotic candidates for *S. aureus*
424 according to *in silico* metrics and *in vitro* experiments relative to both SyntheMol-MCTS (our
425 previous generative model) and VS-Chemprop, a machine learning-based virtual screening
426 method. Indeed, SyntheMol-RL's utility as a flexible tool for readily translatable drug discovery is
427 showcased by the discovery of synthecin, a structurally novel molecule effective at treating a
428 wound infection caused by MRSA in a mouse model.

429

430 SyntheMol-RL overcomes limitations present in other generative AI models^{12,45} for molecular
431 design. The reinforcement learning component of SyntheMol-RL follows a line of research in
432 reinforcement learning algorithms for molecular generation^{24–27}, with the sampling of diverse but
433 high scoring building blocks within the reinforcement learning algorithm inspired by a similar
434 sampling mechanism in GFlowNets³⁵. However, the SyntheMol-RL algorithm is explicitly

435 designed for synthesizable molecule generation, which requires learning to explore a
436 combinatorial chemical space of tens of billions of molecules constructed from pre-defined
437 molecular building blocks and chemical reactions. In contrast, many reinforcement learning and
438 GFlowNet models either do not incorporate synthesizability at all^{27,35} or only include it as a
439 heuristic to be optimized^{36,46}, resulting in the generation of many synthetically infeasible
440 compounds, as seen both in our GFlowNet experiments and in other work⁴⁷. While some
441 models have been designed to generate molecules following known synthetic routes^{48–56} like
442 SyntheMol-RL, these studies typically synthesize and experimentally validate few – if any –
443 generated compounds, making it challenging to evaluate those models in terms of real-world
444 synthesizability and biological efficacy.

445
446 Interestingly, the molecules generated by SyntheMol-RL tend to form clusters of a particular
447 chemotype. Specifically, in each cluster, SyntheMol-RL selects one shared building block while
448 exploring diverse “second” building blocks (Extended Data Fig. 5). This behavior naturally arises
449 from the design of SyntheMol-RL, since the reinforcement learning policy balances both finding
450 building blocks that consistently lead to high-scoring molecules (thus, often repeating a
451 promising building block) while also exploring diverse compounds, which is further reinforced by
452 the post-hoc diversity filtering (thus, selecting diverse second building blocks). The resulting
453 clusters of generated molecules that can be selected for laboratory testing are therefore ideal
454 for both maximizing the probability of finding unique hits across distinct clusters while also
455 naturally exploring the structure-activity relationship (SAR) landscape⁵⁷ within each cluster.
456 When a cluster of compounds includes multiple hits, as with our RL-Chemprop model, the
457 cluster is likely amenable to SAR optimization based on the presence of those hits and any
458 closely related inactive compounds. In contrast, an isolated hit within a cluster would provide
459 little information about the molecular substructures that are responsible for its activity and may
460 suggest challenges with downstream medicinal chemistry optimization.

461
462 The *in vitro* results highlight the potential of SyntheMol-RL-generated compounds, particularly in
463 addressing the global burden of MRSA. Indeed, MRSA has seen a dramatic rise in mortality
464 over the past three decades, with deaths associated with MRSA increasing from 261,000 in
465 1990 to 550,000 in 2021, while attributable deaths doubled to 130,000 in the same period¹.
466 Additionally, MRSA accounts for 26.1% of attributable deaths, making it the foremost cause of
467 deaths attributed to antimicrobial resistance in high-income regions. To address this urgent and
468 growing unmet need, we designed novel antibacterial molecules that display promising activity
469 in such resistant *S. aureus* strains. Serendipitously, we discovered molecules with apparent
470 narrow-spectrum activity, which is useful to avoid the widespread disruption of host microbiota
471 and decrease the rate of resistance dissemination^{58–60}. Among the 79 total SyntheMol-RL
472 compounds synthesized, 15 highly potent active compounds were identified against *S. aureus*
473 RN4220, a remarkable hit rate of 19%. We identified seven structurally novel compounds that
474 retained potent activity against *S. aureus* USA300, as well as VISA strains that cover all
475 resistance mechanisms present for *S. aureus* in the CDC AR Isolate Bank, further supporting
476 their robust efficacy in targeting burdensome *S. aureus* infections.
477
478 Moreover, the C.R.E.A.M. assay provided a simple and informative *in vitro* platform for
479 evaluating the release and antibacterial activity of molecules formulated for topical
480 administration. The correlation between the zone of inhibition observed in this assay and the
481 subsequent *in vivo* efficacy of synthecin highlights its utility for prioritizing candidates for *in vivo*
482 validation. In the mouse wound infection model, synthecin treatment markedly suppressed
483 bacterial proliferation in wounds, resulting in a bacterial burden comparable to pre-treated
484 control mice and mitigated tissue inflammation. These findings demonstrate the C.R.E.A.M.
485 assay as a promising *in vitro* tool for prioritizing topical antibiotics and the utility of synthecin to
486 effectively treat MRSA-infected wounds.

487

488 Building on the successful application to antibiotic discovery, the capabilities of SyntheMol-RL
489 can be further explored and extended. The reinforcement learning algorithm itself can be refined
490 for even more accurate prediction of which molecular building blocks will produce the most ideal
491 compounds for a given application. Additionally, while we only ran SyntheMol-RL with two
492 objectives, the model design allows for generating molecules that optimize an arbitrary number
493 of objectives simultaneously. Future work can explore the number and type of objectives that
494 SyntheMol-RL is best suited for to maximize the drug-like characteristics of generated
495 compounds. Notably, since SyntheMol-RL is compatible with any property predictor and
496 combinatorial chemical space, it can be readily extended to a wide variety of drug discovery and
497 molecular design problems. Thus, SyntheMol-RL is a powerful and flexible tool for designing
498 promising compounds that can be rapidly synthesized, thereby bridging the gap between
499 computational design and laboratory validation.

500

501 **Methods**

502 *Antibiotic Training Set Curation*

503 The training set consists of the “Bioactives 2” chemical library, a structurally and functionally
504 diverse 10,716 compound library housed in the Centre for Microbial Chemical Biology at
505 McMaster University. The library was screened in two biological replicates against *S. aureus*
506 RN4220 for growth inhibitory activity. Cells were grown overnight at 37°C in 3 ml Luria-Bertani
507 (LB) medium and then diluted 1/10,000 in fresh LB. 99 µL of cells were added to each well of
508 Costar 96-well flat-bottom plates manually. Each compound was then added (1 µL) to a final
509 screening concentration of 50 µM in a final volume of 100 µL. Plates were then incubated at
510 37°C with shaking (900 rpm) for 16 hours. Plates were then read at 600 nm using a BioTek
511 Synergy Neo2 plate reader and data were normalized by plate using interquartile mean (IQM)
512 prior to data compiling and hit identification (see below).

513

514 To assign binary activity labels to the screened compounds, we first calculated the average
515 normalized optical density at 600 nm (OD_{600}) for each compound using two biological replicates.
516 We then determined the mean and standard deviation of these average normalized OD_{600}
517 values across all compounds in the dataset. A threshold of $\mu - 2\sigma$ was applied to binarize these
518 values – values below this threshold were labeled as active, while those at or above the
519 threshold were labeled as inactive. Subsequently, we canonicalized the SMILES for each
520 compound using RDKit version 2023.9.1. For data points with identical SMILES and binary
521 activity labels, we retained one data point and discarded the others, resulting in 10,660 data
522 points. For data points with identical SMILES but conflicting binary activity labels (at least one
523 labeled as active and one as inactive), we removed all such samples to avoid noise in the
524 activity labels. This process yielded a final *S. aureus* antibiotic training dataset of 10,658 unique
525 molecules, of which 1,137 (10.7%) are active and 9,521 (89.3%) are inactive.

526

527 *ChEMBL Antibiotics Curation*

528 To compile a set of known antibiotics for structural comparison, we queried the ChEMBL
529 database on November 8, 2023, using the search terms "antibiotic" and "antibacterial." The
530 search term "antibiotic" yielded 636 molecules
531 (https://www.ebi.ac.uk/chembl/g/#search_results/compounds/query=antibiotic), with 591 of them
532 having SMILES. The term "antibacterial" returned 611 molecules
533 (https://www.ebi.ac.uk/chembl/g/#search_results/compounds/query=antibacterial), with 590
534 having SMILES. We combined the two sets of compounds, excluded molecules with missing
535 SMILES, converted the SMILES to canonical SMILES using RDKit, and then deduplicated the
536 compounds based on the canonical SMILES. This process resulted in a set of 1,007 unique
537 molecules.

538

539 *Morgan Fingerprints*

540 Morgan fingerprints were used for molecular similarity and distance calculations. Morgan
541 fingerprints were calculated with a radius of 2 and 2,048 bits using RDKit's
542 GetMorganFingerprintAsBitVect function.

543

544 *Molecular Similarities*

545 Molecular similarities were calculated using either Tanimoto similarity or Tversky similarity.

546 Tanimoto similarity is defined as $Ta(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}$ where X and Y are Morgan fingerprints of two
547 compounds. It is a symmetric similarity measure with values between 0 and 1, where higher
548 values indicate that compounds that are more similar (i.e., they have more molecular
549 substructures in common). Tversky similarity is calculated as $Tv(X, Y) = \frac{|X \cap Y|}{|Y|}$ where X is the
550 Morgan fingerprint of a proposed (i.e., generated) compound and Y is the Morgan fingerprint of
551 a reference compound. It is an asymmetric similarity measure with values between 0 and 1,
552 where higher values indicate that the proposed compound contains a higher proportion of the
553 reference compound's substructures. We use the Tversky similarity specifically for novelty
554 calculations, where we aim to measure the proportion of known antibiotic functional groups in
555 reference compounds that are contained in proposed compounds while ignoring extraneous
556 substructures in the proposed compounds.

557

558 *Molecular Diversity*

559 Diverse sets of generated molecules were selected using a method based on the idea of finding
560 a maximal independent set of molecules. The goal of this method is to identify the largest set of
561 molecules such that no two compounds in the set have a similarity (e.g., Tanimoto similarity)

562 greater than some user-defined threshold, which thereby provides a strict guarantee that the
563 selected set of compounds is diverse according to that similarity threshold.

564

565 More precisely, given a set of molecules A , a molecular similarity metric s , and a similarity
566 threshold t , we aim to select $D = \operatorname{argmax}_{D \subseteq A} \{D \mid \forall x, y \in D, s(x, y) \leq t\}$. If we reformulate this as
567 a graph $G = (V, E)$ where $V = A$ (i.e., every molecule is a node) and $E = \{x, y \in A \mid s(x, y) > t\}$
568 (i.e., an edge connects every pair of molecules with a similarity greater than t), then selecting
569 the largest diverse subset of molecules D is precisely equivalent to solving the maximum
570 independent set problem on G . A maximum independent set is the largest independent set in a
571 graph, where an independent set is a set of nodes such that no two nodes in the set have an
572 edge connecting them. Since the maximum independent set problem is NP-hard, solving it is
573 intractable, but we can instead calculate a maximal independent set as an approximation (i.e.,
574 an independent set that cannot be made larger by adding nodes.) Although a maximal
575 independent set is not necessarily the largest independent set in the graph, it is still an
576 independent set, which means it satisfies our diversity requirement of no two molecules
577 exceeding a given similarity threshold.

578

579 To obtain a maximal independent set, we formulate a given set of molecules as a graph with
580 nodes and edges as detailed above, and we apply the `maximal_independent_set` function from
581 `networkx` version 3.2.1. To slightly improve the approximation, we run the function 10 times with
582 different random seeds and take the largest of the maximal independent sets. This set of
583 molecules then serves as our diverse set.

584

585 *t-SNE Visualizations*

586 t-SNE visualizations were generated using scikit-learn's⁶¹ tSNE, applied to the Morgan
587 fingerprints of molecules with Jaccard (Tanimoto) as the distance metric, squared distances,
588 and a principal components analysis (PCA) initialization. For large datasets, a subset of
589 molecules was randomly sampled to accurately represent the dataset in the t-SNE visualization.

590

591 *Property Predictor Architectures*

592 We used two property prediction model architectures: (1) Chemprop-RDKit and (2) MLP-RDKit.
593 Chemprop-RDKit consists of the graph neural network (GNN) Chemprop augmented with
594 molecular features computed by the cheminformatics package RDKit. Chemprop-RDKit takes
595 as input the graph structure of a molecule with atoms as nodes and bonds as edges. Then, the
596 GNN portion of the model applies three message passing steps that aggregate simple features
597 of each atom and bond in a molecule, such as the atom type and bond type, using neural
598 network layers to build vector representations of local neighborhoods of the molecule. After the
599 message passing steps, these local representations are summed to form a single GNN vector
600 representation for the whole molecule. This GNN vector representation, which is 300-
601 dimensional, is then concatenated with a vector of 200 molecular features computed by RDKit
602 to form a 500-dimensional vector. This combined vector is passed through a multilayer
603 perceptron (MLP) with one hidden layer. The activation function on the output layer of the MLP
604 is a sigmoid layer for binary classification properties (e.g., antibacterial activity) and is a linear
605 layer for regression properties (e.g., aqueous solubility). MLP-RDKit has the same architecture
606 as the MLP at the end of the Chemprop-RDKit model, but it does not have a GNN component
607 and therefore only takes as input the vector of 200 RDKit features. Both architectures were
608 implemented using Chemprop version 1.6.1.

609

610 *Model Training*

611 Property prediction models were trained using 10-fold cross-validation with data randomly split
612 into 80% train, 10% validation, and 10% test for each fold. The models were all trained on the
613 training data and evaluated on the test data and used the validation data for early stopping. All
614 models were trained for 30 epochs using the Adam optimizer. The antibacterial activity models
615 used a binary cross-entropy loss and were evaluated using the area under the receiver
616 operating characteristic curve (ROC-AUC) and area under the precision recall curve (PRC-
617 AUC). The aqueous solubility models used a mean squared error loss and were evaluated using
618 mean squared error (MSE) and the coefficient of determination (R^2).

619

620 *Enamine REAL Space*

621 The Enamine REAL Space comprises 31 billion make-on-demand molecules that can be
622 synthesized through a single chemical reaction using a limited number of molecular building
623 blocks as reactants. For our study, we used the 11/2021 version of the REAL reactions, which
624 includes 169 chemical reactions, along with the 2022 q1-2 version of the REAL building blocks,
625 consisting of 139,517 molecular building blocks. We downloaded the 2022 q1-2 version of the
626 REAL Space on August 30, 2022, which encompasses 31,507,987,117 molecules that can be
627 produced using the specified building blocks and chemical reactions.

628

629 Following the approach used in the original SyntheMol-MCTS study, we used only 13 of the
630 most common reactions. To prepare the building blocks for our model, we first used RDKit to
631 convert the SDF file of the building blocks to SMILES format. All building blocks were converted
632 successfully. Next, we deduplicated the molecules by SMILES, reducing the number to 139,444
633 molecules. We then applied the RDKit salt remover to eliminate salts from the building blocks to
634 avoid incorrect reaction template matching during generation, and we removed 24 molecules
635 whose salts could not be properly removed. This resulted in a set of 137,656 unique molecules

636 (with 139,493 unique building block IDs due to duplicate SMILES). Using the curated reactions
637 and building blocks, we can produce 30,330,025,259 molecules, representing 96.3% of the total
638 REAL Space.

639

640 *WuXi GalaXi*

641 The WuXi GalaXi consists of 16 billion make-on-demand molecules. We used the 12/31/2022
642 version of the WuXi GalaXi, which includes 36 chemical reactions, 15,488 building blocks, and
643 16,146,071,436 molecules that can be produced using these building blocks and chemical
644 reactions. To prepare the building blocks for use in our model, we removed building blocks with
645 missing IDs (4 building blocks) and missing SMILES (4 building blocks) and then deduplicated
646 by SMILES, leaving 14,977 unique molecules. We then removed salts from the building blocks
647 using the RDKit salt remover (all salts were correctly removed).

648

649 *SyntheMol-MCTS*

650 The original implementation of SyntheMol¹⁷ used a Monte Carlo tree search (MCTS) algorithm
651 to generate molecules. Briefly, the MCTS algorithm works as follows. The input to the algorithm
652 is a chemical synthesis tree T that consists of a set of nodes N in T . Each node N contains
653 N_{mol} , which is a set of one or more molecular building blocks B in a chemical space C ($B \subset C$).

654 MCTS defines a value function on nodes, $S(N) = \frac{Q(N) + P(N)U(N)}{D(N)}$. This value function balances
655 exploiting nodes that lead to high scoring molecules via $Q(N)$, selecting nodes containing
656 building blocks with high property prediction scores via $P(N)$, exploring rarely visited nodes via
657 $U(N)$, and selecting diverse building blocks via $D(N)$ (additional details are available in ref.¹⁷).

658

659 To generate a molecule, MCTS first computes the value $S(N)$ of every node containing a single
660 building block. The MCTS policy is then to select the node with the highest value. Given this

661 choice, MCTS creates every possible node with this first building block and a second building
662 block that is synthetically compatible with the first building block in some chemical reaction r
663 contained in the set of chemical reactions R defined by the chemical space ($r \in R$). MCTS then
664 scores each of these nodes with two building blocks and applies its policy by selecting the node
665 with the highest value. Next, MCTS creates every possible node with the two selected building
666 blocks and a third synthetically compatible building block, and it also creates nodes for every
667 possible molecule that can be formed by applying one of the chemical reactions in R to just the
668 two selected building blocks to form a new molecule. MCTS scores all of these nodes – both
669 nodes with three building blocks and nodes with a single molecule formed from the two building
670 blocks – and again follows its policy by choosing the node with the highest score. If that node
671 contains three building blocks instead of a single molecule, then nodes are created using all
672 compatible chemical reactions to combine those three building blocks, and MCTS selects the
673 node with the highest value.

674
675 At this point, MCTS has completed a single rollout and has obtained a single molecule
676 composed of two or three building blocks. This molecule $m \in C$ is scored by a weighted
677 combination of L property predictors, $M_k: C \rightarrow R$ for $k \in \{1, \dots, L\}$, using property weights w_k for
678 $k \in \{1, \dots, L\}$ (where $\sum_{k=1}^L w_k = 1$ and $0 \leq w_k \leq 1$) to obtain the molecule's overall property
679 score, $p(m) = \sum_{k=1}^L w_k * M_k(m)$. This overall property score is then used to update the $Q(N)$
680 exploit score of every node N selected during the generation of molecule m . Additionally, the
681 exploration and diversity scores, $U(N)$ and $D(N)$, are updated to reflect the nodes visited during
682 the generation of molecule m . MCTS repeats this process for a fixed number of rollouts.

683

684 *SyntheMol-RL*

685 The reinforcement learning (RL) version of SyntheMol employs the same synthesis tree T ,
686 chemical space C , chemical reactions R , and property predictors M_1, \dots, M_L with their weights
687 w_1, \dots, w_L as MCTS, but it reformulates the method for computing node values and selecting
688 nodes during each rollout. In SyntheMol-RL, RL replaces the MCTS value function $S(N)$ with an
689 RL value function $V(N)$. $V(N)$ is implemented as a deep neural network that takes as input the
690 building blocks in the node, N_{mol} , and outputs a prediction of the expected overall property
691 score of molecules that can be created from the building blocks N_{mol} by following the RL policy.
692 The RL policy is to apply the RL value function $V(N)$ to all nodes created at a given step in
693 generation and then sample a node proportional to the values of the nodes with a temperature
694 scaling, i.e., $P(N) \propto e^{V(N)/\tau}$. The temperature parameter τ can be tuned to affect the amount of
695 exploration or exploitation performed by the RL policy; high temperature prefers exploration
696 (more uniform probabilities) while low temperature prefers exploitation (more spiky
697 probabilities).

698

699 The RL value function $V(N)$ is implemented as a weighted combination of models, similar to the
700 molecule property score $p(m)$. Specifically, $V(N) = \sum_{k=1}^L w_k * Z_k(N_{mol})$ where Z_1, \dots, Z_L are
701 deep learning models and w_1, \dots, w_L are the same property weights used in $p(m)$. They are then
702 trained on the RL value function objective as follows. After each rollout constructs a molecule m ,
703 the RL algorithm stores tuples of $(N, M_1(m), \dots, M_L(m))$ for every node N in the trajectory of
704 nodes that were selected to create the molecule m . This builds an RL value function training set
705 of nodes along with the property prediction scores of the final molecules created from those
706 nodes. After every $n_{rl\ train}$ rollouts, the RL models Z_1, \dots, Z_L are trained for $n_{rl\ epochs}$ epochs to

707 take as input the building blocks of the node, N_{mol} , and predict the relevant score of the
708 generated molecule from M_1, \dots, M_L using a mean squared error loss.

709

710 *Dynamic RL Parameters*

711 The RL temperature parameter τ and property weights w_1, \dots, w_L have a large impact on the
712 diversity of generated molecules and their relative property scores, respectively. By default, the
713 temperature is set to $\tau = 0.1$ and the property weights are set to $w_1, \dots, w_L = \frac{1}{L}$. However, these
714 values may not be optimal for every molecule design problem. Rather than manually tuning
715 each parameter, we designed a dynamic tuning mechanism⁶² that automatically adjusts the
716 temperature and property weights based on pre-defined goals.

717

718 The RL temperature aims to balance exploration and exploitation, so we defined the
719 temperature goal explicitly by setting a target similarity λ^* , which is the desired maximum
720 Tanimoto similarity between each generated molecule and all previously generated molecules.
721 We then defined a dynamic tuning method to adjust the RL temperature during generation to
722 obtain a molecule similarity of λ^* on average. During generation, the dynamic temperature
723 tuning mechanism maintains a rolling average molecule similarity λ_{avg} , which is initialized as
724 $\lambda_{avg} = \lambda^*$. After each rollout i generates a molecule m_i , λ_{avg} is updated as $\lambda_{avg} = \gamma * \lambda_{avg} +$
725 $(1 - \gamma) * \max_{j=1, \dots, i-1} sim(m_i, m_j)$ where $\gamma = 0.98$ is the rolling average weight and $sim(m_i, m_j)$ is the
726 Tanimoto similarity between the molecule m_i generated on rollout i and the molecule m_j
727 generated on rollout j . Next, the percent difference between the average similarity λ_{avg} and the
728 target similarity λ^* is computed as $\lambda_{diff} = \frac{(\lambda_{avg} - \lambda^*)}{\lambda^*}$ and a new desired temperature is computed
729 as $\tau_{new} = \tau + \lambda_{diff} * \tau$. To smooth out changes to the temperature, the temperature is updated

730 via a rolling average with $\tau = \gamma * \tau + (1 - \gamma) * \tau_{new}$. Finally, the temperature is clipped within
731 reasonable limits with $\tau = \max(\tau_{min}, \min(\tau, \tau_{max}))$ where $\tau_{min} = 0.001$ and $\tau_{max} = 10$.

732

733 The property weights aim to balance the importance of the two property objectives (antibacterial
734 activity and aqueous solubility), so we defined the property weight goal as maximizing the
735 number of molecules that simultaneously possess both properties. Specifically, we defined
736 success thresholds t_1, \dots, t_L for each property, and we considered that molecule m is a hit for
737 property k if $M_k(m) \geq t_k$. In a similar manner to dynamic temperature tuning, we perform
738 dynamic property weight tuning by computing a rolling average. Here, for each property
739 $k \in \{1, \dots, L\}$, we define a rolling average success rate s_{avg}^k which is initialized with $s_{avg}^k = 0$. Our
740 aim is to maximize all success rates simultaneously, which we hypothesize will occur when the
741 success rates are equal. Therefore, we aim to adjust the property weights to make the success
742 rates match. After each rollout i generates a molecule m_i , then s_{avg}^k is updated as $s_{avg}^k = \gamma *$
743 $s_{avg}^k + (1 - \gamma) * I[M_k(m_i) \geq t_k]$ where $\gamma = 0.98$ as with the dynamic temperature and I is in the
744 indicator function that is 1 if $M_k(m_i) \geq t_k$ and 0 otherwise. We then compute the average
745 success rate across the properties: $s_{avg} = \frac{1}{L} \sum_{k=1}^L s_{avg}^k$. If $s_{avg} = 0$, meaning no successful
746 molecules have been generated, then the property weights are unchanged. If $s_{avg} > 0$, then we
747 determine the relative amount by which each individual success rate, s_{avg}^k , deviates from their
748 mean, s_{avg} , by computing $s_{avg\ diff}^k = \frac{(s_{avg}^k - s_{avg})}{s_{avg}}$. We then perform a rolling average update of the
749 property weights with $w_k = \gamma * w_k + (1 - \gamma) * (w_k - s_{avg\ diff}^k * w_k)$. These property weights are
750 normalized to sum to 1 by computing $w_k = \frac{w_k}{\sum_{k=1}^L w_k}$. Then, the property values are clipped to
751 $w_k = \max(w_k, w_{min})$ where $w_{min} = 0.001$ and are then renormalized with $w_k = \frac{w_k}{\sum_{k=1}^L w_k}$. At a
752 high level, this ensures that if the success rate for property k is higher (or lower) than the
753 average success rate across the properties, then the corresponding property weight w_k will be

754 decreased (or increased) proportional to the difference between that property's success rate
755 and the average success rate.

756

757 *Generating Molecules with SyntheMol*

758 We applied SyntheMol, both using MCTS and using RL, to generate antibacterial molecules
759 against *S. aureus*. For both RL and MCTS, we had $k = 2$ properties: antibacterial activity (binary
760 classification) and aqueous solubility (regression). Since Chemprop-RDKit performed best on
761 both properties, we set the property predictors M_1 and M_2 to be the trained Chemprop-RDKit
762 models for antibacterial activity and aqueous solubility, respectively. Specifically, M_1 and M_2
763 were the ensemble of ten models for each property from the ten folds of cross-validation, and
764 predictions consist of the average prediction across the ten models. We adapted MCTS to use
765 the same dynamic property weighting scheme as RL to set the weights w_1 and w_2 , and for both
766 models, we defined the success thresholds as $t_1 = 0.5$ for the antibacterial model and $t_2 = -4$
767 for the log solubility model to match our hit definitions for filtering generated compounds.

768

769 For RL, we implemented the RL value function models Z_1 and Z_2 as either Chemprop-RDKit or
770 MLP-RDKit models. When Z_1 and Z_2 are Chemprop-RDKit models, we refer to the overall
771 architecture as “RL-Chemprop”, and when Z_1 and Z_2 are MLP-RDKit models, we refer to the
772 overall architecture as “RL-MLP”. The Z_1 and Z_2 models were initialized with the weights of M_1
773 and M_2 . Specifically, we set Z_1 and Z_2 to be the first of the ten models in the ensembles of M_1
774 and M_2 since Z_1 and Z_2 in our implementation are single models, not ensembles.

775

776 Notably, as an RL value function, these models must make predictions not just on single
777 molecules but on combinations of molecules (e.g., multiple molecular building blocks). For
778 Chemprop-RDKit models, the model itself is unmodified. The Chemprop GNN component is

779 applied in its usual form to the multiple molecules by treating them as a single graph with
780 disconnected components, while the RDKit component uses the average of the RDKit features
781 of the individual molecules, thereby preserving the dimensionality of the RDKit feature input. For
782 MLP-RDKit models, the model is modified by replicating the MLP weights of the first layer
783 $n_{max\ bbs}$ times, where $n_{max\ bbs}$ is the maximum number of building blocks that can be combined
784 into a single molecule ($n_{max\ bbs} = 3$). For a given combination of molecules, RDKit features are
785 computed, concatenated, and fed as input to this modified MLP. If there are fewer than $n_{max\ bbs}$
786 molecules, the remaining elements of the input vector are set to 0.

787

788 After modification as described above, the Chemprop-RDKit (Z_1) and MLP-RDKit (Z_2) RL value
789 function models were then trained during generation every $n_{rl\ train} = 10$ rollouts for $n_{rl\ epochs} =$
790 5 epochs across all the nodes and molecules that had been previously generated. To
791 dynamically tune the RL temperature τ , we set the target similarity $\lambda^* = 0.6$ to match our
792 diversity filter, which uses 0.6 as a threshold to define diverse molecules.

793

794 We ran all three SyntheMol versions – RL-Chemprop, RL-MLP, and MCTS – for 10,000 rollouts,
795 generating 10,983; 9,228; and 11,630 molecules, respectively. We used 8 CPUs for all three
796 versions, and we additionally used 1 GPU for RL-Chemprop due to the computational burden of
797 running the Chemprop models on millions of building block combinations (RL-MLP and MCTS
798 were faster with CPU only due to the relatively lightweight computations involved). RL-
799 Chemprop took 4–7 days, RL-MLP took 9–12 hours, and MCTS took 4–5 hours, depending on
800 the precise hardware that was used (GPUs were either NVIDIA A40, TITAN_Xp, TITAN_V, or
801 RTX_2080Ti).

802

803 *Virtual Screening with Chemprop-RDKit*

804 To perform virtual screening, we first selected 21 million molecules by sampling 14 million
805 molecules uniformly at random from the 31 billion REAL molecules and by sampling 7 million
806 molecules uniformly at random from the 16 billion GalaXi molecule. The final 21 million
807 compounds roughly approximated the ratio between the two chemical spaces. We then applied
808 our Chemprop-RDKit models for antibacterial activity and aqueous solubility to all 21 million
809 molecules using 1 GPU and 8 CPUs, which took roughly 7 days. This is in line with the time
810 taken by the slowest SyntheMol-RL model.

811

812 *Random Selection*

813 We randomly selected 150 molecules as a baseline for synthesis and validation by sampling
814 100 molecules uniformly at random from the 31 billion REAL molecules and by sampling 50
815 molecules uniformly at random from the 16 billion GalaXi molecules.

816

817 *Toxicity Predictions*

818 Toxicity predictions for molecules were made using ADMET-AI version 1.2.0. ADMET-AI
819 consists of Chemprop-RDKit models trained on 41 absorption, distribution, metabolism,
820 excretion, and toxicity (ADMET) datasets from the Therapeutics Data Commons. We applied
821 ADMET-AI to make predictions for all 41 properties for each molecule and then extracted the
822 clinical toxicity (ClinTox) predictions. The clinical toxicity predictions are values in the range [0,
823 1] with higher values indicating a higher probability of toxicity.

824

825 *Compound Synthesis*

826 Compound synthesis was performed by Enamine as part of their Enamine REAL Space and by
827 WuXi as part of their WuXi GalaXi. Compound purity of at least 85% was verified using liquid

828 chromatography–mass spectrometry (LC–MS) except in cases of poor solubility, compound
829 instability under LC–MS conditions, or non-informative LC–MS. In these cases, proton nuclear
830 magnetic resonance (¹H-NMR) was used to assess chemical purity.

831

832 *GFlowNet*

833 GFlowNet molecules were generated using a multi-objective GFlowNet model. We specifically
834 used the implementation of multi-objective GFlowNets from
835 <https://github.com/recursionpharma/gflownet>, which implemented a multi-objective reward
836 including binding to the target sEH, drug-likeness measure QED, molecular weight, and
837 synthetic accessibility heuristic SAScore. We replaced this reward function with a new reward
838 consisting of our antibacterial activity Chemprop-RDKit model ensemble, our aqueous solubility
839 Chemprop-RDKit model ensemble, and the same SAScore heuristic
840 (<https://github.com/swansonk14/gflownet/tree/antibiotics>). We ran the GFlowNet model with this
841 new reward function while keeping all other parameters at their default values. After generating
842 molecules, we selected a final set of molecules using the same filtering procedure applied to the
843 SyntheMol compounds, but we used an additional filter of SAScore ≤ 4 (i.e., easy
844 synthesizability according to the SAScore heuristic) to increase the likelihood of obtaining
845 synthetically tractable compounds. Furthermore, instead of selecting the top 150 compounds by
846 antibacterial prediction score, we only selected the top 14 since compounds had to be manually
847 reviewed by chemists at Enamine and WuXi for synthetic accessibility.

848

849 *SyntheMol-RL Ablations*

850 We performed a series of ablation experiments to determine the effect of each of the
851 components of SyntheMol-RL on the distribution of predicted antibacterial activity and log
852 solubility scores across the generated molecules. We ran each ablation experiment five times

853 with a different random seed each time. All SyntheMol-RL experiments were run with both RL-
854 Chemprop and RL-MLP. Below are the ablation experiments:

855

8561. Final: The final version of SyntheMol-RL. The first seed of this experiment is the one with results
857 reported in the main text.

858

8592. Fixed Property Weights: Instead of the dynamic property weights for antibacterial activity and
860 aqueous solubility, fixed weights were used. The weights were (0.00 antibacterial, 1.00
861 solubility), (0.86 antibacterial, 0.14 solubility), (0.88 antibacterial, 0.12 solubility), (0.90
862 antibacterial, 0.10 solubility), (0.92 antibacterial, 0.08 solubility), (0.94 antibacterial, 0.06
863 solubility), (0.96 antibacterial, 0.04 solubility), and (1.00 antibacterial, 0.00 solubility). The fixed
864 weights were chosen to balance the relative scale of the antibacterial activity predictions
865 (ranging from 0 to 1) and the log solubility predictions (roughly ranging from -10 to 2).

866

8673. Dynamic Temperature Target: Instead of the dynamic temperature similarity target of 0.6,
868 similarity targets of 0.4, 0.5, 0.7, and 0.8 were used.

869

8704. Fixed Temperature: Instead of the dynamic temperature with a similarity target, fixed
871 temperatures of 0.01, 0.05, 0.1, 0.5, and 1.0 were used. As a point of comparison, we also ran
872 SyntheMol-MCTS with different explore weights (the MCTS equivalent of the RL temperature
873 parameter). The explore weights were 0.5, 1.0, 5.0, and 50.0 (the default of 10.0 was used in
874 the final model reported in the main text).

875

8765. RL Training: Instead of using reinforcement learning to train the RL value models Z_1 and Z_2 ,
877 these models were kept fixed after initializing them with weights from the trained property
878 predictors M_1 and M_2 . Thus, they know how to score complete molecules but are not trained to

879 score molecular building blocks for their potential to form complete molecules with desired
880 properties (i.e., the RL objective).

881

8826. RL Pretrained Weights: Instead of initializing the RL value models Z_1 and Z_2 with weights from
883 the trained property predictors M_1 and M_2 , these models are initialized randomly and trained
884 from scratch using the RL objective.

885

886 *Antibacterial Potency Analyses*

887 *S. aureus* RN4220, *S. aureus* USA300, *S. aureus* clinical isolates (VISA panel; CDC AR Isolate
888 Bank), *Escherichia coli* BW25113, *Pseudomonas aeruginosa* PAO1, *Klebsiella pneumoniae*
889 ATCC 43816, *Enterococcus faecium* ATCC 19434, and *Acinetobacter baumannii* ATCC 17978
890 were grown overnight at 37°C in 3 ml of LB medium with shaking. Overnight cultures were then
891 diluted 1:10,000 into fresh LB. Cells were then introduced to two-fold serial dilutions of each
892 compound under investigation in a final volume of 100 µl in Costar 96-well flat-bottom plates.
893 Plates were incubated at 37°C without shaking (*E. coli* BW25113, *P. aeruginosa* PAO1, *K.*
894 *pneumoniae* ATCC 43816, *E. faecium* ATCC 19434, and *A. baumannii* ATCC 17978) or with
895 shaking at 900 rpm (*S. aureus* RN4220, *S. aureus* USA300, and *S. aureus* clinical isolates) until
896 untreated control cultures reached stationary phase. Plates were then read at 600 nm using a
897 BioTek Synergy Neo2 plate reader.

898

899 *C.R.E.A.M. assay*

900 *S. aureus* USA300 cells were grown overnight in 3 ml LB medium at 37°C with shaking. ~10⁶
901 □CFU in 100□µl liquid LB was deposited onto solid LB agar plates. ~100 mg of control or
902 SyntheMol-RL-generated molecule in Glaxal Base was deposited onto inoculated agar plates,
903 and plates were incubated for 18 h at 37°C. For chemical preparation, molecules were weighed
904 and solubilized in 10% DMSO and then added to 100 mg Glaxal Base to a final concentration of

905 2% w/v. The solution was mixed thoroughly to ensure an even distribution of the compound in
906 the Glaxal Base carrier. For the control, the same amount of vehicle (DMSO) was measured
907 and mixed through Glaxal Base.

908

909 *Bacterial Cell Killing*

910 *S. aureus* USA300 cells were grown overnight in 3 ml LB medium at 37°C with shaking and
911 diluted 1:10,000 into fresh LB. In 96-well flat-bottom plates, cells were grown to the required
912 density in a final volume of 100 µl at 37°C with shaking at 900 rpm, at which time compound
913 was added at the indicated concentration and cultures were incubated for the required duration.
914 Cells were then pelleted in 96-well flat-bottom plates by centrifugation at 4,000 x g for 15□ min
915 at 4□°C and washed in ice-cold sterile PBS. After washing, cells were ten-fold serially diluted in
916 PBS and plated on solid LB agar medium.

917

918 *Mouse Infection Model*

919 Mouse model experiments were conducted according to the guidelines set by the Canadian
920 Council on Animal Care, using protocols approved by the Animal Review Ethics Board and
921 McMaster University under Animal Use Protocol no. 22-04-10. Six- to eight-week-old female
922 C57BL/6N mice were purchased from Charles River Laboratories (#027, QC, CAN). Animals
923 were housed in a specific pathogen-free barrier facility under Containment Level 2 conditions
924 and maintained on a 12 h light: 12 h dark cycle, which was temperature-controlled (21°C) at 30-
925 50% humidity. No animals were excluded from the analysis, and blinding was considered
926 unnecessary.

927

928 Mice were pre-treated with 150□mg/kg (day T-4) and 100□mg/kg (day T-1) of
929 cyclophosphamide to render mice neutropenic. On day T-0, mice were anesthetized using
930 isoflurane and administered buprenorphine-sustained release as an analgesic at 0.6□mg/ml

931 subcutaneously. A 2 cm² abrasion on the dorsal surface of the mouse was inflicted through
932 tape-stripping to the basal layer of the epidermis using approximately 30-35 pieces of autoclave
933 tape. Mice were immediately infected with ~2.15 x 10⁷ CFU *S. aureus* USA300 directly pipetted
934 onto the wound bed. The infection was left to establish for 1 h before the first treatment with
935 Glaxal Base supplemented with vehicle (10% DMSO) or synthecin (2% wt/vol). Mice (n=5)
936 were treated 1 h, 4 h, 8 h, 12 h, and 20 h post-infection with 150 mg Glaxal Base with
937 synthecin (treatment) or DMSO (control). Mice were sacrificed at experimental endpoint (24 hpi)
938 and wound tissue was aseptically collected, homogenized in PBS and plated on solid LB agar
939 medium to quantify bacterial load. For chemical preparation, synthecin was weighed and
940 solubilized in 10% DMSO and then added to 150 mg Glaxal Base to a final concentration of 2%
941 w/v. The solution was mixed thoroughly to ensure an even distribution of the compound in the
942 carrier. For control groups, the same amount of vehicle (DMSO) was measured and mixed
943 through Glaxal Base.

944

945 **Data Availability**

946 Training data (for *S. aureus* and aqueous solubility), molecular building blocks and reactions,
947 generated molecules are in Supplementary Data 1-12. The purity spectra of the synthesized
948 molecules are in Supplementary Data 13. All of this data, along with trained property prediction
949 models and the molecules screened by VS-Chemprop, are available on Zenodo at
950 <https://doi.org/10.5281/zenodo.15391267> (ref.⁶³).

951

952 **Code Availability**

953 Code for data processing and SyntheMol-RL molecule generation is available on GitHub at
954 <https://github.com/swansonk14/SyntheMol> and on Zenodo at
955 <https://doi.org/10.5281/zenodo.15392449> (ref.⁶⁴). This code repository makes use of general
956 cheminformatics functions from https://github.com/swansonk14/chem_utils as well as Chemprop

957 model code from <https://github.com/chemprop/chemprop> and GFlowNet model code from
958 <https://github.com/swansonk14/gflownet/tree/antibiotics>.

959

960 **References**

- 961 1. Murray, C. J. L. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a
962 systematic analysis. *The Lancet* **399**, 629–655 (2022).
- 963 2. Rice, L. B. Federal Funding for the Study of Antimicrobial Resistance in Nosocomial
964 Pathogens: No ESKAPE. *J. Infect. Dis.* **197**, 1079–1081 (2008).
- 965 3. Enright, M. C. *et al.* The evolutionary history of methicillin-resistant *Staphylococcus aureus*
966 (MRSA). *Proc. Natl. Acad. Sci.* **99**, 7687–7692 (2002).
- 967 4. Miller, L. S. & Cho, J. S. Immunity against *Staphylococcus aureus* cutaneous infections.
968 *Nat. Rev. Immunol.* **11**, 505–518 (2011).
- 969 5. Tacconelli, E. *et al.* Discovery, research, and development of new antibiotics: the WHO
970 priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18**, 318–327
971 (2018).
- 972 6. Stokes, J. M. *et al.* A Deep Learning Approach to Antibiotic Discovery. *Cell* **180**, 688–
973 702.e13 (2020).
- 974 7. Wong, F. *et al.* Discovery of a structural class of antibiotics with explainable deep learning.
975 *Nature* **626**, 177–185 (2024).
- 976 8. Wan, F., Torres, M. D. T., Peng, J. & de la Fuente-Nunez, C. Deep-learning-enabled
977 antibiotic discovery through molecular de-extinction. *Nat. Biomed. Eng.* **8**, 854–871 (2024).
- 978 9. Melo, M. C. R., Maasch, J. R. M. A. & de la Fuente-Nunez, C. Accelerating antibiotic
979 discovery through artificial intelligence. *Commun. Biol.* **4**, 1050 (2021).
- 980 10. Yang, K. *et al.* Analyzing Learned Molecular Representations for Property Prediction. *J.*
981 *Chem. Inf. Model.* **59**, 3370–3388 (2019).

- 982 11. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual
983 screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **3**, 935–949
984 (2004).
- 985 12. Bilodeau, C., Jin, W., Jaakkola, T., Barzilay, R. & Jensen, K. F. Generative models for
986 molecular discovery: Recent advances and challenges. *WIREs Comput. Mol. Sci.* **12**, e1608
987 (2022).
- 988 13. Gangwal, A. & Lavecchia, A. Unleashing the power of generative AI in drug discovery. *Drug*
989 *Discov. Today* **29**, 103992 (2024).
- 990 14. Bian, Y. & Xie, X.-Q. Generative chemistry: drug discovery with deep learning generative
991 models. *J. Mol. Model.* **27**, 71 (2021).
- 992 15. Zeng, X. *et al.* Deep generative molecular design reshapes drug discovery. *Cell Rep. Med.*
993 **3**, 100794 (2022).
- 994 16. Gao, W. & Coley, C. W. The Synthesizability of Molecules Proposed by Generative Models.
995 *J. Chem. Inf. Model.* **60**, 5714–5723 (2020).
- 996 17. Swanson, K. *et al.* Generative AI for designing and validating easily synthesizable and
997 structurally novel antibiotics. *Nat. Mach. Intell.* **6**, 338–353 (2024).
- 998 18. Huang, K. *et al.* Therapeutics Data Commons: Machine Learning Datasets and Tasks for
999 Drug Discovery and Development. in *Thirty-fifth Conference on Neural Information*
1000 *Processing Systems Datasets and Benchmarks Track (Round 1)* (2021).
- 1001 19. Kreiswirth, B. N. *et al.* The toxic shock syndrome exotoxin structural gene is not detectably
1002 transmitted by a prophage. *Nature* **305**, 709–712 (1983).
- 1003 20. Zdrazil, B. *et al.* The ChEMBL Database in 2023: a drug discovery platform spanning
1004 multiple bioactivity data types and time periods. *Nucleic Acids Res.* **52**, D1180–D1192
1005 (2024).
- 1006 21. RDKit. <https://www.rdkit.org/>.

- 1007 22. Kocsis, L. & Szepesvári, C. Bandit Based Monte-Carlo Planning. in *Machine Learning:*
1008 *ECML 2006* (eds. Fürnkranz, J., Scheffer, T. & Spiliopoulou, M.) 282–293 (Springer, Berlin,
1009 Heidelberg, 2006). doi:10.1007/11871842_29.
- 1010 23. Coulom, R. Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search. in
1011 *Computers and Games* (eds. van den Herik, H. J., Ciancarini, P. & Donkers, H. H. L. M.
1012 (Jeroen)) 72–83 (Springer, Berlin, Heidelberg, 2007). doi:10.1007/978-3-540-75538-8_7.
- 1013 24. Zhou, Z., Kearnes, S., Li, L., Zare, R. N. & Riley, P. Optimization of Molecules via Deep
1014 Reinforcement Learning. *Sci. Rep.* **9**, 10752 (2019).
- 1015 25. Ståhl, N., Falkman, G., Karlsson, A., Mathiason, G. & Boström, J. Deep Reinforcement
1016 Learning for Multiparameter Optimization in de novo Drug Design. *J. Chem. Inf. Model.* **59**,
1017 3166–3176 (2019).
- 1018 26. Popova, M., Isayev, O. & Tropsha, A. Deep reinforcement learning for de novo drug design.
1019 *Sci. Adv.* **4**, eaap7885 (2018).
- 1020 27. Olivecrona, M., Blaschke, T., Engkvist, O. & Chen, H. Molecular de-novo design through
1021 deep reinforcement learning. *J. Cheminformatics* **9**, 48 (2017).
- 1022 28. Grygorenko, O. O. *et al.* Generating Multibillion Chemical Space of Readily Accessible
1023 Screening Compounds. *iScience* **23**, 101681 (2020).
- 1024 29. Chemical Spaces • Ultra-Large Compound Collections by BioSolveIT. *BioSolveIT*
1025 <https://www.biosolveit.de/chemical-spaces/>.
- 1026 30. Klug, D. M. *et al.* Open Source Antibiotics: Simple Diarylimidazoles Are Potent against
1027 Methicillin-Resistant *Staphylococcus aureus*. *ACS Infect. Dis.* **9**, 2423–2435 (2023).
- 1028 31. Tversky, A. Features of similarity. *Psychol. Rev.* **84**, 327–352 (1977).
- 1029 32. Maggiora, G., Vogt, M., Stumpfe, D. & Bajorath, J. Molecular Similarity in Medicinal
1030 Chemistry. *J. Med. Chem.* **57**, 3186–3204 (2014).
- 1031 33. Swanson, K. *et al.* ADMET-AI: a machine learning ADMET platform for evaluation of large-
1032 scale chemical libraries. *Bioinformatics* **40**, btae416 (2024).

- 1033 34. Mukherjee, S. *et al.* ADMET-AI Enables Interpretable Predictions of Drug-Induced
1034 Cardiotoxicity. *Circulation* **151**, 285–287 (2025).
- 1035 35. Bengio, E., Jain, M., Korablyov, M., Precup, D. & Bengio, Y. Flow Network based
1036 Generative Models for Non-Iterative Diverse Candidate Generation. in *Advances in Neural*
1037 *Information Processing Systems* vol. 34 27381–27394 (Curran Associates, Inc., 2021).
- 1038 36. Jain, M. *et al.* Multi-Objective GFlowNets. in *Proceedings of the 40th International*
1039 *Conference on Machine Learning* (eds. Krause, A. *et al.*) vol. 202 14631–14653 (PMLR,
1040 2023).
- 1041 37. Macielag, M. J. *et al.* Substituted Salicylanilides as Inhibitors of Two-Component Regulatory
1042 Systems in Bacteria. *J. Med. Chem.* **41**, 2939–2945 (1998).
- 1043 38. Rajamuthiah, R. *et al.* Repurposing Salicylanilide Anthelmintic Drugs to Combat Drug
1044 Resistant *Staphylococcus aureus*. *PLOS ONE* **10**, e0124595 (2015).
- 1045 39. Weiss, A. *et al.* Topical niclosamide (ATx201) reduces *Staphylococcus aureus* colonization
1046 and increases Shannon diversity of the skin microbiome in atopic dermatitis patients in a
1047 randomized, double-blind, placebo-controlled Phase 2 trial. *Clin. Transl. Med.* **12**, e790
1048 (2022).
- 1049 40. Lee, A. S. *et al.* Methicillin-resistant *Staphylococcus aureus*. *Nat. Rev. Dis. Primer* **4**, 1–23
1050 (2018).
- 1051 41. Lutgring, J. D. *et al.* FDA-CDC Antimicrobial Resistance Isolate Bank: a Publicly Available
1052 Resource To Support Research, Development, and Regulatory Requirements. *J. Clin.*
1053 *Microbiol.* **56**, 10.1128/jcm.01415-17 (2018).
- 1054 42. Moran, G. J. *et al.* Methicillin-Resistant *S. aureus* Infections among Patients in the
1055 Emergency Department. *N. Engl. J. Med.* **355**, 666–674 (2006).
- 1056 43. Siddiqui, A. H. & Koirala, J. Methicillin-Resistant *Staphylococcus aureus*. in *StatPearls*
1057 (StatPearls Publishing, Treasure Island (FL), 2025).
- 1058 44. Clemett, D. & Markham, A. Linezolid. *Drugs* **59**, 815–827 (2000).

- 1059 45. Du, Y. *et al.* Machine learning-aided generative molecular design. *Nat. Mach. Intell.* **6**, 589–
1060 604 (2024).
- 1061 46. Liu, C.-H. *et al.* RetroGNN: Approximating Retrosynthesis by Graph Neural Networks for De
1062 Novo Drug Design. Preprint at <https://doi.org/10.48550/arXiv.2011.13042> (2020).
- 1063 47. Gao, W. & Coley, C. W. The Synthesizability of Molecules Proposed by Generative Models.
1064 *J. Chem. Inf. Model.* **60**, 5714–5723 (2020).
- 1065 48. Cretu, M., Harris, C., Roy, J., Bengio, E. & Lio, P. SynFlowNet: Towards Molecule Design
1066 with Guaranteed Synthesis Pathways. in *ICLR 2024 Workshop on Generative and*
1067 *Experimental Perspectives for Biomolecular Design* (2024).
- 1068 49. Bradshaw, J., Paige, B., Kusner, M. J., Segler, M. H. S. & Hernández-Lobato, J. M. A Model
1069 to Search for Synthesizable Molecules. in *Proceedings of the 33rd International Conference*
1070 *on Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, USA,
1071 2019).
- 1072 50. Gottipati, S. K. *et al.* Learning to Navigate the Synthetically Accessible Chemical Space
1073 Using Reinforcement Learning. in *Proceedings of the 37th International Conference on*
1074 *Machine Learning* (JMLR.org, 2020).
- 1075 51. Horwood, J. & Noutahi, E. Molecular Design in Synthetically Accessible Chemical Space via
1076 Deep Reinforcement Learning. *ACS Omega* **5**, 32984–32994 (2020).
- 1077 52. Button, A., Merk, D., Hiss, J. A. & Schneider, G. Automated de novo molecular design by
1078 hybrid machine intelligence and rule-driven chemical synthesis. *Nat. Mach. Intell.* **1**, 307–
1079 315 (2019).
- 1080 53. Gao, W., Mercado, R. & Coley, C. W. Amortized Tree Generation for Bottom-up Synthesis
1081 Planning and Synthesizable Molecular Design. in *International Conference on Learning*
1082 *Representations* (2022).

- 1083 54. Klarich, K., Goldman, B., Kramer, T., Riley, P. & Walters, W. P. Thompson Sampling—An
1084 Efficient Method for Searching Ultralarge Synthesis on Demand Databases. *J. Chem. Inf.
1085 Model.* **64**, 1158–1171 (2024).
- 1086 55. Gao, W., Luo, S. & Coley, C. W. Generative Artificial Intelligence for Navigating
1087 Synthesizable Chemical Space. Preprint at <https://doi.org/10.48550/arXiv.2410.03494>
1088 (2024).
- 1089 56. Koziarski, M. *et al.* RGFN: Synthesizable Molecular Generation Using GFlowNets. *Adv.
1090 Neural Inf. Process. Syst.* **37**, 46908–46955 (2024).
- 1091 57. Tong, W., Welsh, W. J., Shi, L., Fang, H. & Perkins, R. Structure-activity relationship
1092 approaches and applications. *Environ. Toxicol. Chem.* **22**, 1680–1695 (2003).
- 1093 58. J. Melander, R., V. Zurawski, D. & Melander, C. Narrow-spectrum antibacterial agents.
1094 *MedChemComm* **9**, 12–21 (2018).
- 1095 59. Rea, M. C. *et al.* Effect of broad- and narrow-spectrum antimicrobials on Clostridium difficile
1096 and microbial diversity in a model of the distal colon. *Proc. Natl. Acad. Sci.* **108**, 4639–4644
1097 (2011).
- 1098 60. Jernberg, C., Löfmark, S., Edlund, C. & Jansson, J. K. Long-term ecological impacts of
1099 antibiotic administration on the human intestinal microbiota. *ISME J.* **1**, 56–66 (2007).
- 1100 61. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J Mach Learn Res* **12**, 2825–
1101 2830 (2011).
- 1102 62. Hong, Z.-W. *et al.* Diversity-driven exploration strategy for deep reinforcement learning. in
1103 *Proceedings of the 32nd International Conference on Neural Information Processing
1104 Systems* 10510–10521 (Curran Associates Inc., Red Hook, NY, USA, 2018).
- 1105 63. Swanson, K. *et al.* SyntheMol-RL Data. Zenodo <https://doi.org/10.5281/zenodo.15391268>
1106 (2025).
- 1107 64. swansonk14, GaryLiu152 & Callon, K. swansonk14/SyntheMol: SyntheMol-RL Release.
1108 Zenodo <https://doi.org/10.5281/zenodo.15392449> (2025).

1109

1110 **Acknowledgments**

1111 This research was kindly supported by the Weston Family Foundation (J.M.S.); the David Braley
1112 Centre for Antibiotic Discovery (J.M.S.); the Canadian Institutes of Health Research (CIHR)
1113 (J.M.S.); the Natural Sciences and Engineering Research Council of Canada (NSERC) (J.M.S.);
1114 a generous gift from M. and M. Heersink (J.M.S.); and the Chan-Zuckerberg Biohub (J.Z.). In
1115 addition, this research was kindly supported by the Knight-Hennessy Scholarship (K.S.) and the
1116 Stanford Bio-X Fellowship (K.S.); a CIHR Canada Graduate Scholarship CGS-D (G.L., M.M.T.);
1117 a CIHR Canada Graduate Scholarship CGS-M (D.B.C., S.M., A.A.). E.D.B. was supported by a
1118 Tier 1 Canada Research Chair award and a Project Grant from CIHR (PJT 186281). We thank
1119 D. Richman for helping to inspire the name SyntheMol as well as for early discussions of the
1120 model.

1121

1122 **Author Contributions**

1123 Conceptualization was carried out by K.S., G.L., J.Z., and J.M.S. Model development was
1124 performed by K.S. and G.L. *In vitro* validation was carried out by D.B.C., S.M., and A.A. Murine
1125 studies were conducted by D.B.C. and M.M.T. K.S., G.L., D.B.C., S.M., J.Z., and J.M.S. wrote
1126 the paper. E.D.B., J.Z., and J.M.S. supervised the work.

1127

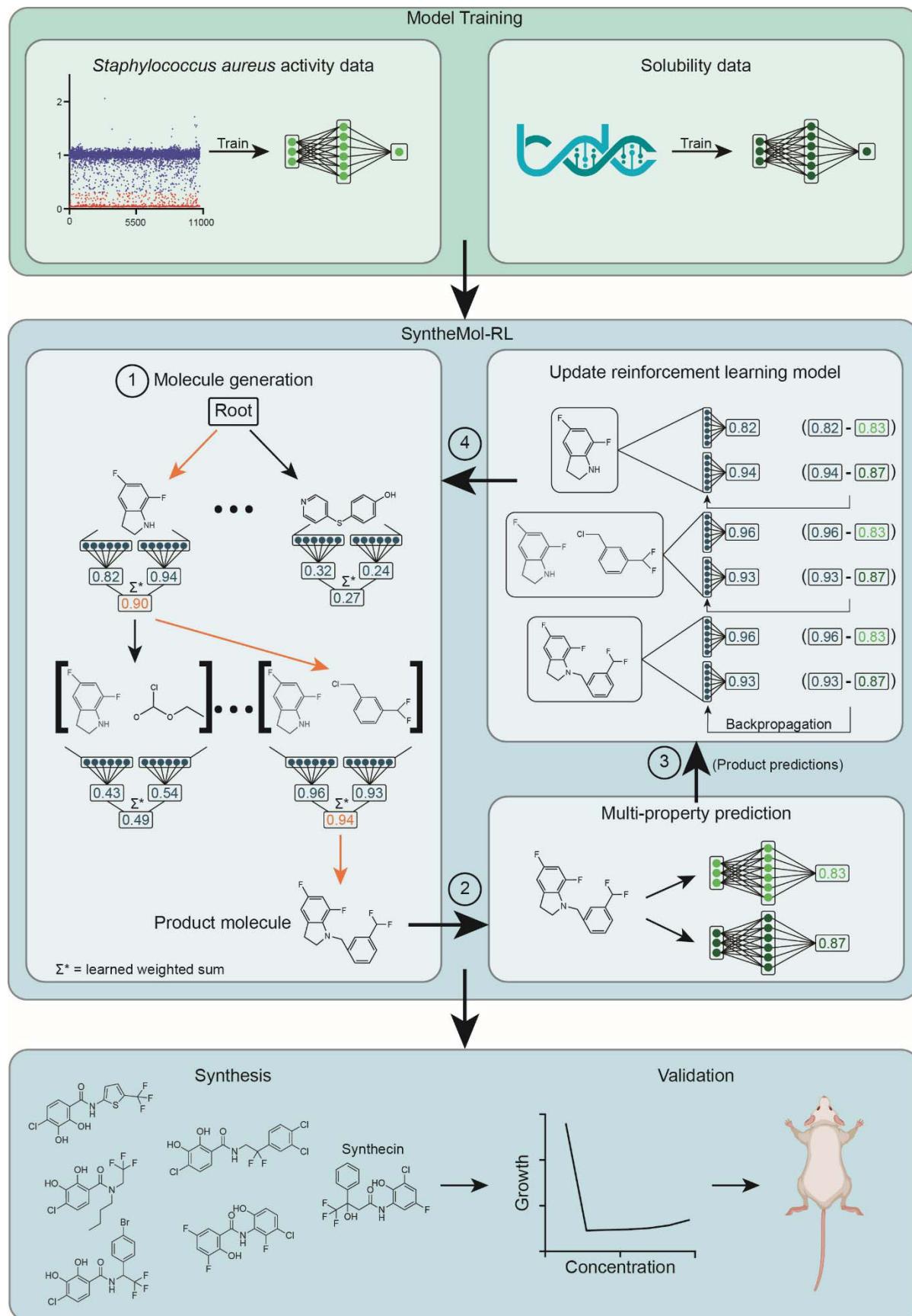
1128 **Competing Interests**

1129 K.S. is a part-time employee of Greenstone Biosciences and is a consultant at Merck & Co., Inc.
1130 G.L. & D.B.C. are consultants for Stoked Bio. J.M.S. is a founder of Stoked Bio.

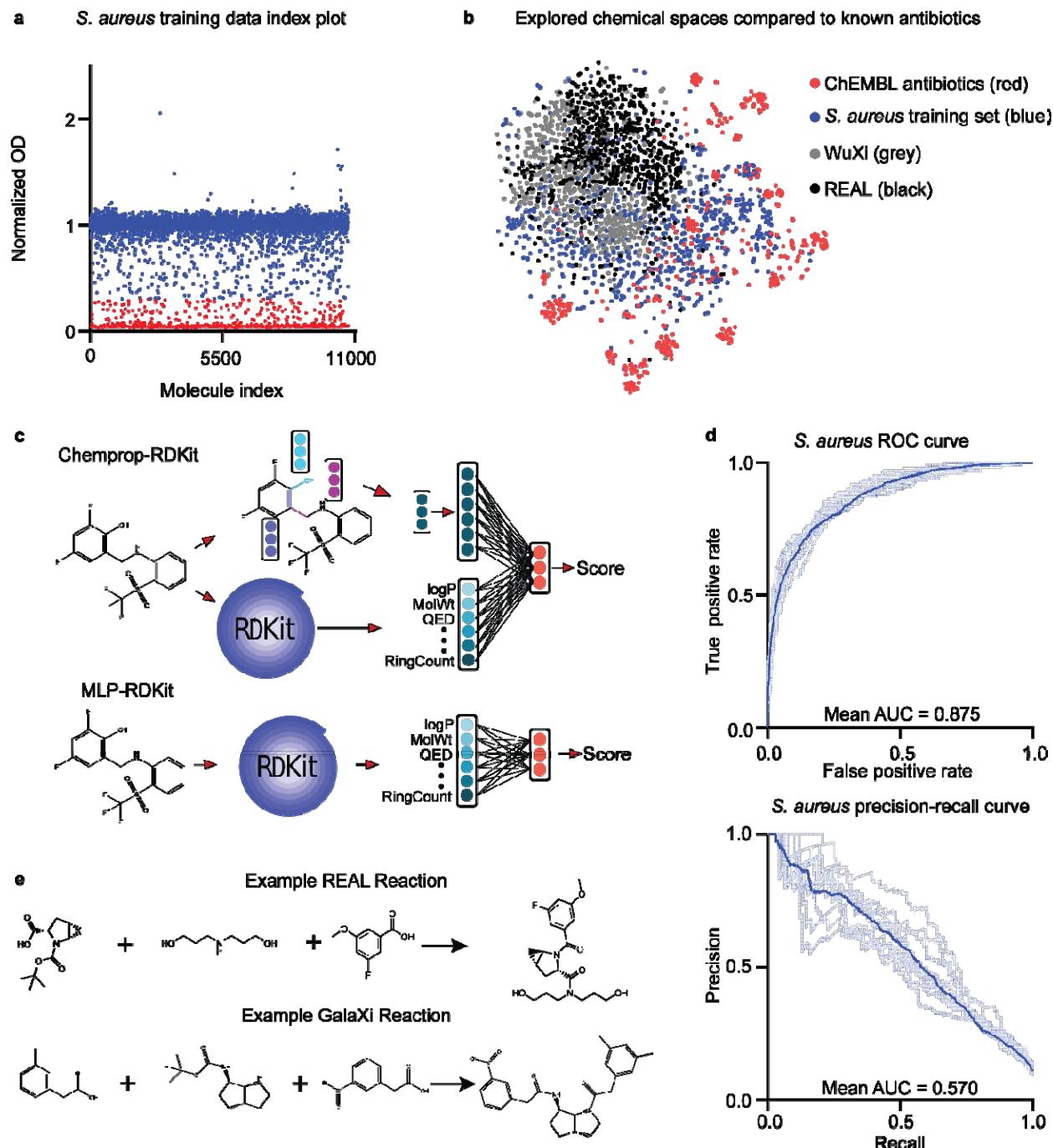
1131

1132 **Figures**

1133 *Main Figures*



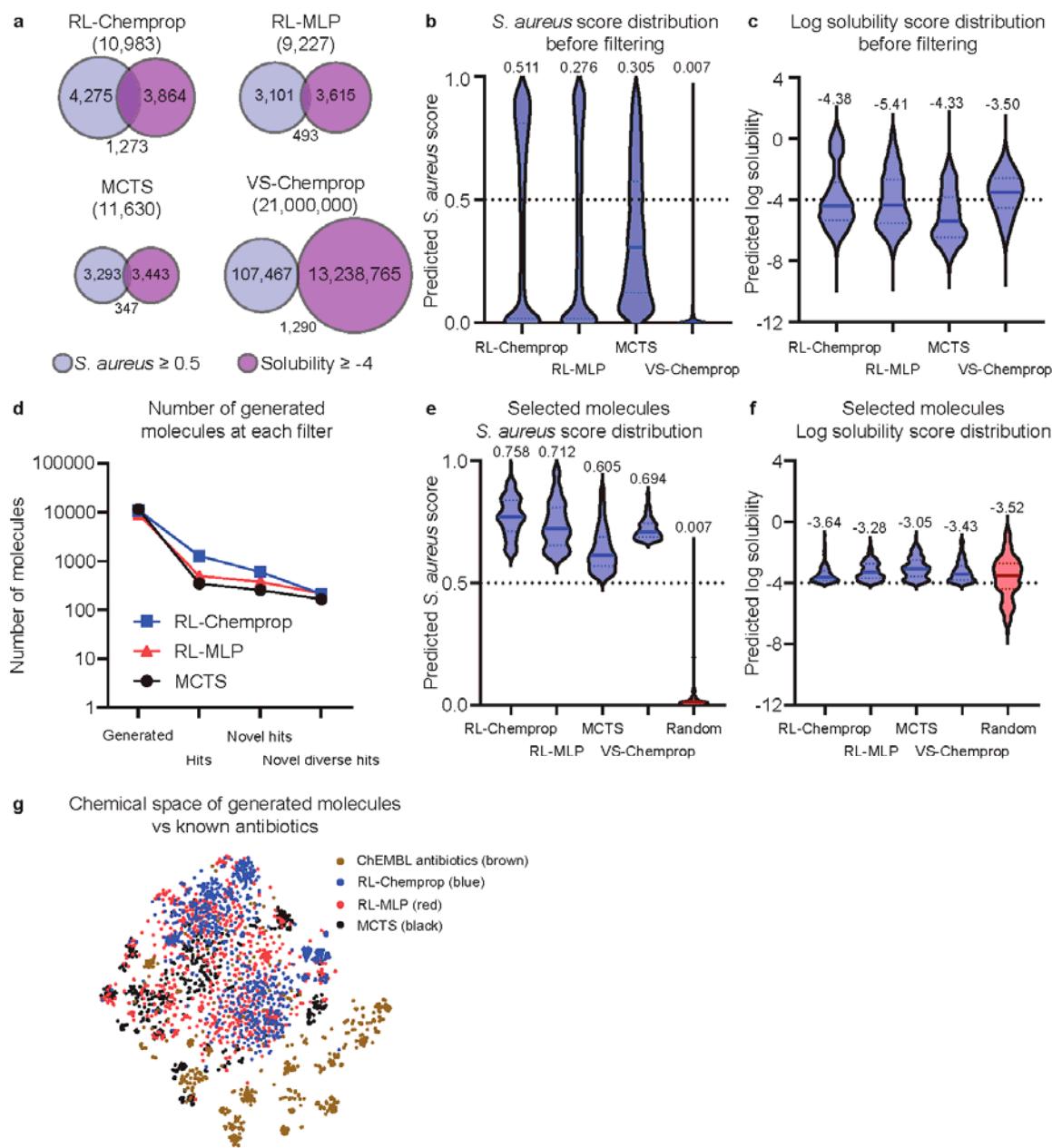
1135 **Figure 1. Overview of the SyntheMol-RL pipeline.** The SyntheMol-RL pipeline can be broken
1136 into three primary steps. (1) Property prediction models are trained based on a combination of
1137 in-house chemical screening data against *S. aureus* as well as aqueous solubility data obtained
1138 from the Therapeutic Data Commons. (2) SyntheMol-RL generates molecules guided by these
1139 property prediction models. A reinforcement learning (RL) model is used to iteratively select
1140 molecular building blocks based on their potential for generating molecules with promising *S.*
1141 *aureus* antibacterial activity and aqueous solubility (RL scores shown in blue). The generated
1142 molecule is then scored using the property prediction models, and these scores are used to
1143 update the RL model to improve its generative capacity (property prediction scores shown in
1144 green). (3) A prioritized subset of molecules is selected for synthesis and laboratory testing
1145 based on predicted properties as well as novelty and diversity. Notably, this SyntheMol-RL
1146 pipeline generated a novel compound that can effectively treat a MRSA wound infection in a
1147 mouse model, showing the utility of SyntheMol-RL for real-world drug discovery applications.



1148

1149 **Figure 2. Property prediction models and chemical spaces.** a, A diverse set of small
1150 molecules were screened against *S. aureus* RN4220 at 50 μ M final concentration. The mean
1151 normalized growth across two biological replicates is plotted. The hit cutoff is calculated as:
1152 (the mean minus two standard deviations). Points below the cutoff are considered “hits”
1153 and are colored red ($n = 1,137$), and points above the cutoff are “non-hits” and are colored blue

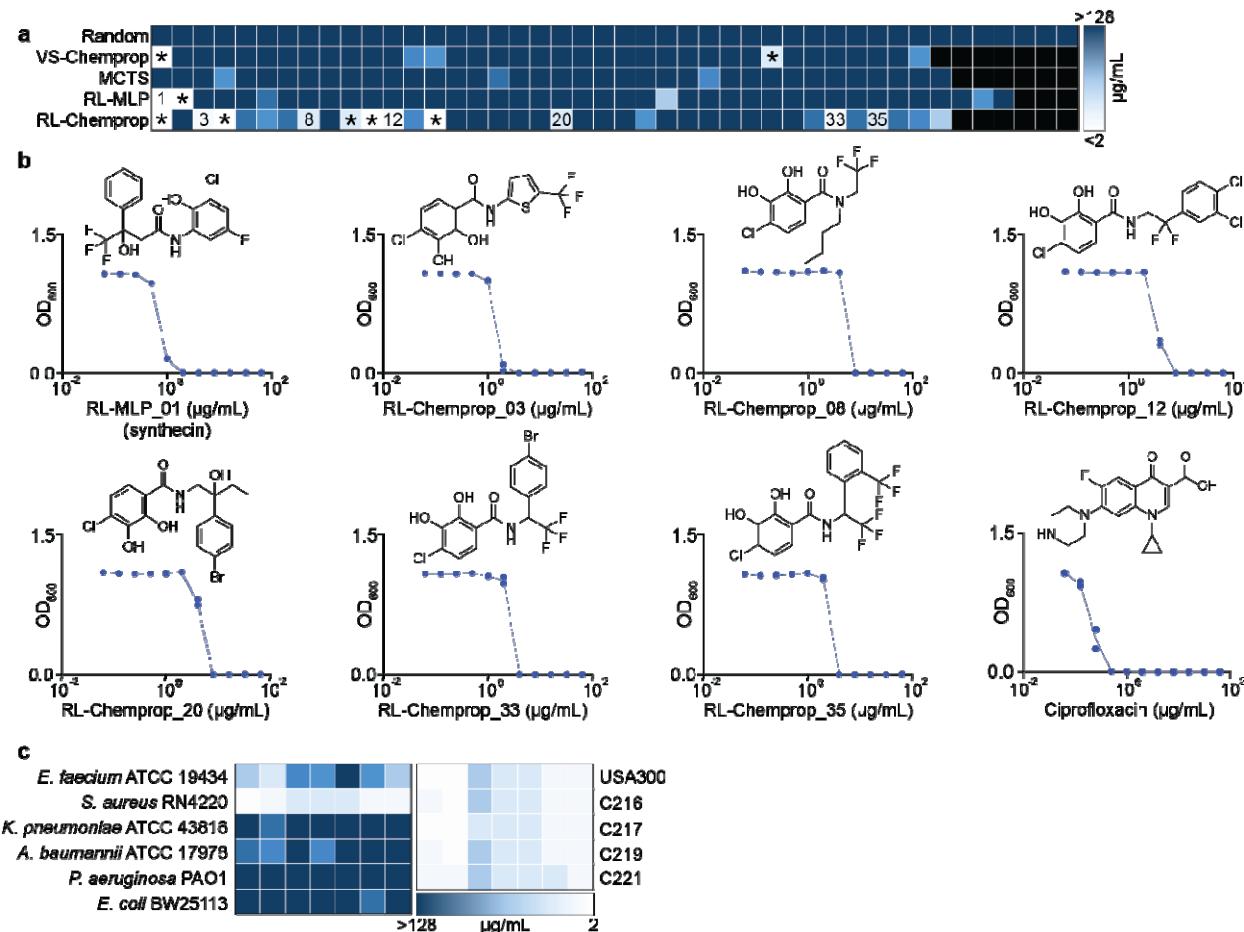
1154 ($n = 9,521$), totaling 10,658 data points used for training. b, A t-SNE visualization of chemical
1155 spaces represented by ChEMBL antibiotics and the *S. aureus* training set in comparison to the
1156 chemical spaces explored by SyntheMol-RL. c, Depictions of the two types of models used in
1157 this paper. Chemprop-RDKit (top) was consistently used as the property predictor in each
1158 experiment. Both architectures were used for the reinforcement learning model. d, ROC and
1159 precision-recall curves for each model in an ensemble of 10 models used as the *S. aureus*
1160 activity property prediction model. Dark curves represent the average of all 10 models in the
1161 ensemble. Area under the curve (AUC) is indicated on the respective plots. e, Example
1162 reactions are shown for the two ultra-large chemical spaces used for SyntheMol-RL: the
1163 Enamine REAL Space (top) and the WuXi GalaXi (bottom).



1164

1165 **Figure 3. Generating compounds with SyntheMol-RL.** a, Venn diagrams showing the
1166 number of molecules from each SyntheMol method (RL-Chemprop, RL-MLP, MCTS) or virtual
1167 screening method (VS-Chemprop) that pass the respective thresholds for predicted *S. aureus*
1168 activity and solubility, respectively. b, Violin plots displaying the distribution of *S. aureus* scores
1169 across the four methods. The dotted line indicates the threshold used to define “hit” (≥ 0.5). c,
1170 Violin plots displaying the distribution of predicted log solubility scores across four methods. The

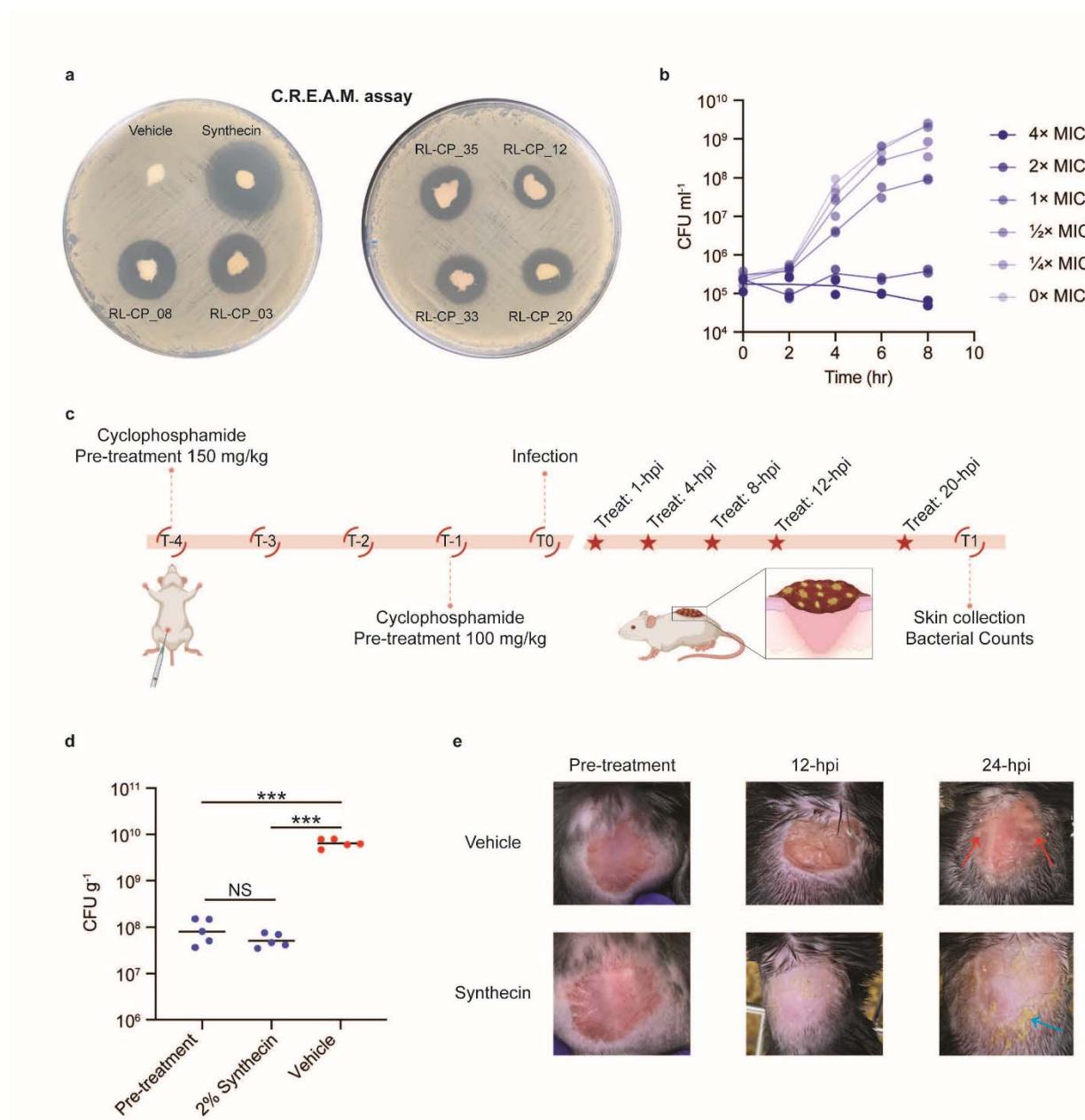
1171 dotted line indicates the threshold used to define “hit” (≥ -4). d, Line graph representing the
1172 number of molecules present at each stage of post-hoc filtering for each version of SyntheMol.
1173 e, Violin plots displaying the distribution of *S. aureus* scores across the four methods after
1174 applying filters outlined in Methods. Random molecules (red) were randomly selected from the
1175 chemical space and serve as a baseline for comparison. The dotted line indicates the threshold
1176 used to define “hit” (≥ 0.5). f, Violin plots displaying the distribution of predicted log solubility
1177 scores across the four methods after applying filters outlined in Methods as well as random
1178 molecules (red). The dotted line indicates the threshold used to define “hit” (≥ -4). g, A t-SNE
1179 visualization of the training set along with all generated molecules from each of the three
1180 SyntheMol models compared to known antibiotics from ChEMBL.



1181

1182 **Figure 4. Generated molecules show potent *in vitro* activity.** a, A heat map showing the
1183 minimum inhibitory concentrations (MICs) of up to 44 compounds from each generative
1184 SyntheMol method (RL-Chemprop, RL-MLP, MCTS), the virtual screening method (VS-
1185 Chemprop), or the random molecules. Molecules were tested against *S. aureus* RN4220 in LB
1186 medium, with compounds tested at concentrations ranging from 128 $\mu\text{g/ml}$ to 2 $\mu\text{g/ml}$ in two-fold
1187 serial dilutions. Lighter colors indicate lower MIC values. Experiments were performed in
1188 biological duplicate. Annotated squares denote “hit” molecules with $\text{MIC} \leq 8 \mu\text{g/ml}$. Those
1189 denoted with an asterisk (*) indicate molecules that did not pass a thorough manual literature
1190 search for structural novelty. b, Growth inhibition of *S. aureus* RN4220 by the seven structurally
1191 novel molecules that passed the manual literature search. Ciprofloxacin is shown as a positive
1192 control. Structures of compounds are shown. Experiments were performed using two-fold serial

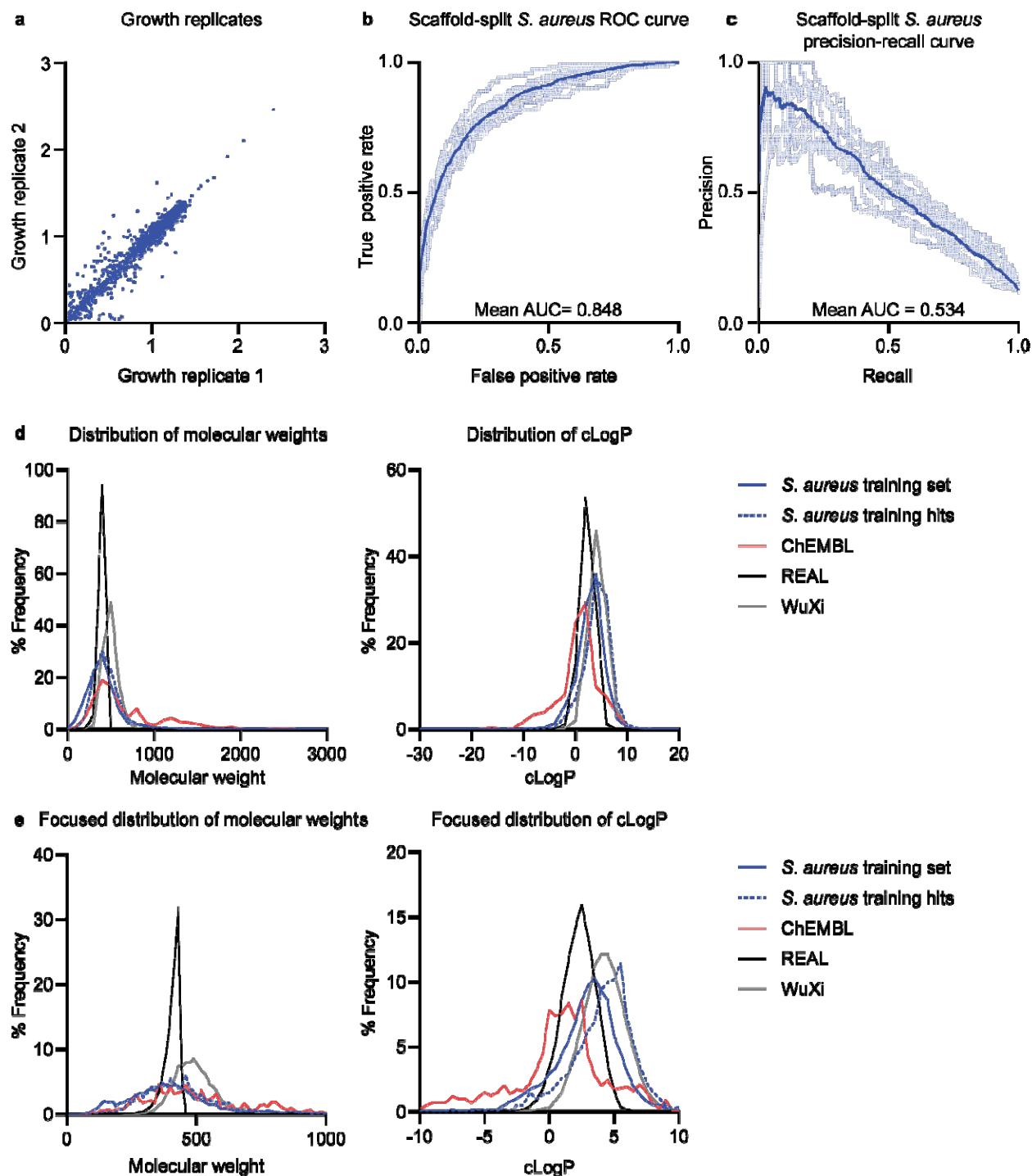
1193 dilution series. Experiments were performed in biological duplicate, with each point representing
1194 one replicate. c, Heat maps depicting the MICs of the seven novel compounds against
1195 representative isolates of the ESKAPE pathogens (left) and against a representative sample of
1196 *S. aureus* clinical isolates from the CDC AR Isolate Bank (right). Molecules are presented in the
1197 same order as b from left to right. Molecules were tested in LB medium, with compound
1198 concentrations ranging from 128 µg/ml to 2 µg/ml in two-fold serial dilutions. Lighter colors
1199 indicate lower MIC values. Experiments were performed in biological duplicate.



1207 Experiments were conducted in biological duplicate. Individual replicates with means connected
1208 are plotted. c, Overview of the infection model used to assess the *in vivo* antibacterial efficacy of
1209 synthecin against *S. aureus* USA300. In a wound infection model, mice were infected with *S.*
1210 *aureus* USA300 ($\sim 2.15 \times 10^7$ CFU). At 1-hour post-infection (hpi), mice were treated with
1211 vehicle (10% DMSO) (n = 5) or 2% synthecin (n = 5). Mice received treatments 4-, 8-, 12-, and
1212 20-hpi. d, Bacterial load of *S. aureus* USA300 from wound tissue prior to treatment (n = 5) and
1213 of vehicle (n = 5) or 2% synthecin-treated (n = 5) mice 24-hpi. Pre-treatment represents the
1214 bacterial load at the time of initial treatment (1-hpi). Black lines represent the geometric mean of
1215 the bacterial load for each group. NS means non statistically significant. ***P using unpaired
1216 two-sided t-test with Welch's correction (Pre-treatment vs 2% synthecin, P = 0.1774; Pre-
1217 treatment vs Vehicle, P = 0.0005; Vehicle vs 2% synthecin, P = 0.0004). e, Representative
1218 images of the dorsal surface of mice pre-treatment (1-hpi), after 12 h of treatment with vehicle or
1219 synthecin, and after 24 h of treatment with vehicle or synthecin. Note the inflammation seen in
1220 the vehicle control (red arrows) that is absent in the synthecin treated mouse. Note the yellow
1221 plaques, which is molecule precipitate, seen in the synthecin treated-mouse (blue arrow) that is
1222 absent in the vehicle-treated mouse.

1223

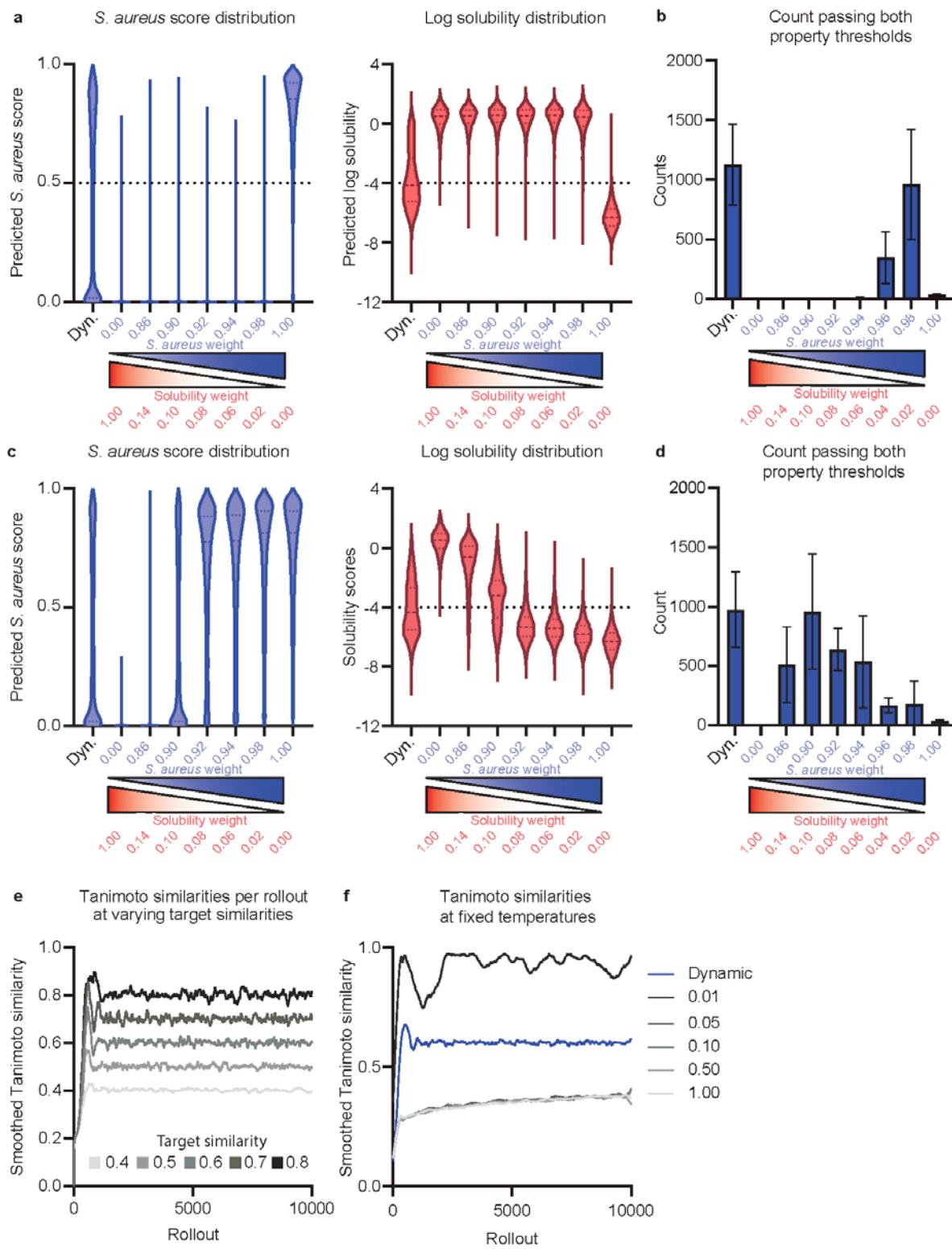
1224 *Extended Data Figures*



1225

1226 **Extended Data Figure 1. Analysis of *S. aureus* property predictor model and chemical**
1227 **spaces.** a, Normalized growth of *S. aureus* RN4220 in duplicate experiments used as the
1228 training data. b, ROC and c, precision-recall curves for each *S. aureus* activity property

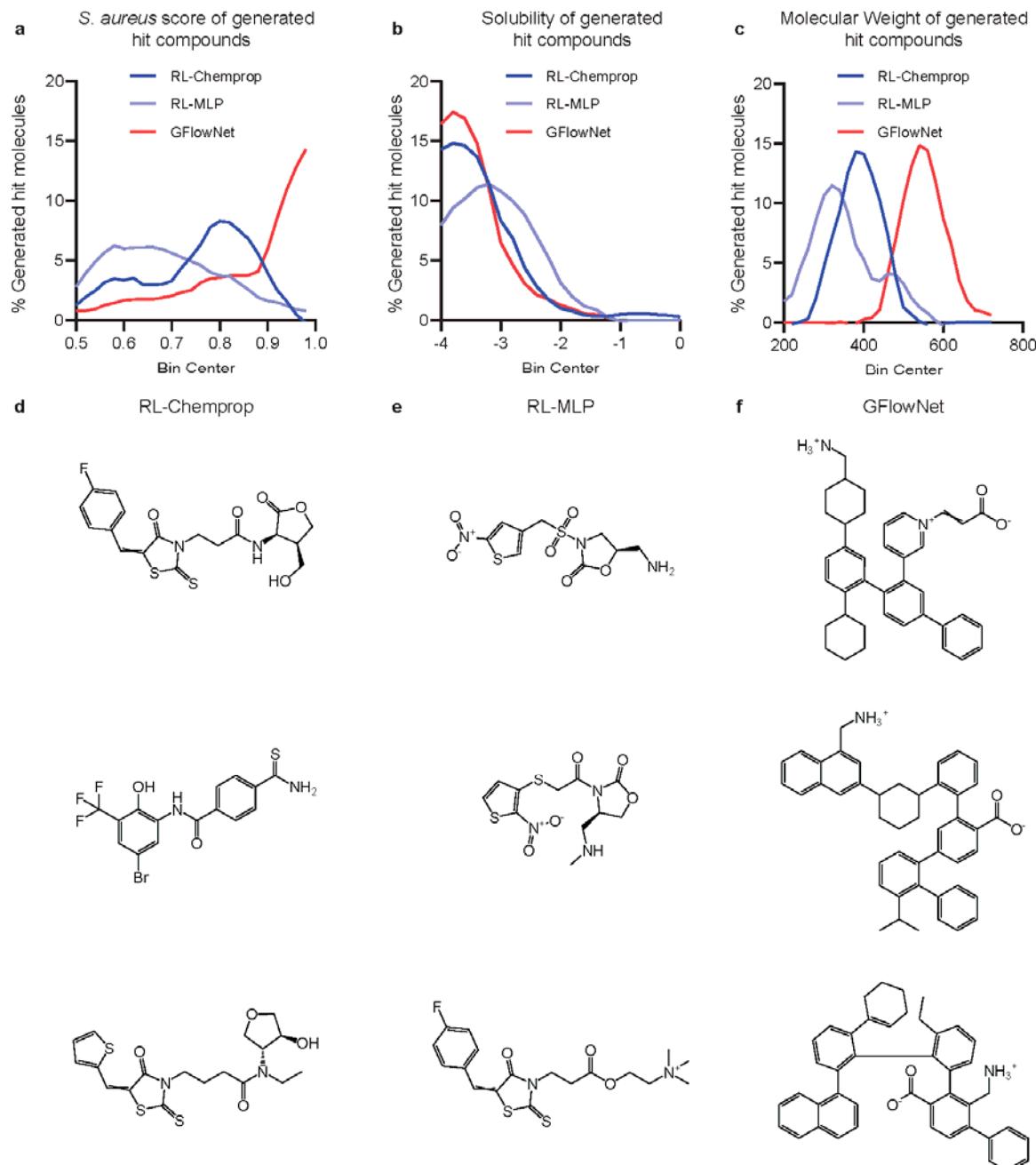
1229 prediction model in an ensemble of 10 split by scaffold. Dark curves represent the average
1230 across all models. Area under the curve is indicated on respective graphs. d, Density plots
1231 showing the distribution of molecular weight and cLogP across 1) the *S. aureus* model training
1232 set 2) a collection of known antibiotics from ChEMBL and 3) the REAL and WuXi GalaXi
1233 chemical spaces explored by SyntheMol-RL. Not visualized are two molecules with molecular
1234 weights exceeding >3,000 g/mol. e, Same as d, but focused on the region of the distribution
1235 containing the majority of molecules for visual clarity.



1236

1237 **Extended Data Figure 2. SyntheMol-RL ablation experiments.** a, Distribution of predicted *S.*
1238 *aureus* activity and log solubility scores for the final dynamic weighting system (Dyn) and fixed

1239 weightings using RL-Chemprop. The sum of weights is equal to one. b, Number of molecules
1240 generated that pass property “hit” thresholds for *S. aureus* activity and log solubility (>0.5 and >
1241 -4 respectively). Dynamic weighting achieves optimal or near-optimal performance without need
1242 for additional fine-tuning, shown by generating the most “hit” compounds. c, Distribution of
1243 predicted *S. aureus* activity and log solubility scores for the final dynamic weighting system and
1244 fixed weightings using RL-MLP. The sum of weights is equal to one. d, Number of molecules
1245 generated that pass property “hit” thresholds for *S. aureus* activity and log solubility (> 0.5 and >
1246 -4 respectively). e, Tanimoto similarities of each rollout compared to previous rollouts for RL-
1247 Chemprop set at desired target similarities in range [0.4, 0.8]. The model can effectively
1248 generate compounds at a user-defined diversity. f, Tanimoto similarities of each rollout
1249 compared to previous rollouts for RL-Chemprop at fixed temperatures and dynamic temperature
1250 for reference. Similarity is highly sensitive to changes in temperature, further necessitating the
1251 use of dynamic changes to output desired diversity.

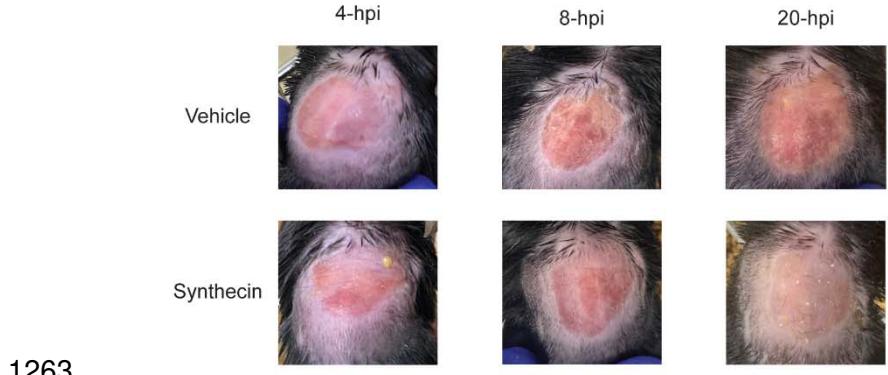


1252

1253 Extended Data Figure 3. Comparisons of hits generated by SyntheMol-RL and GFlowNet.

1254 SyntheMol-RL and GFlowNet generated “hit” compounds that passed filters for predicted
1255 antibacterial activity (>0.5), log solubility (>-4), novelty, and diversity (Methods). Additionally,
1256 GFlowNet molecules were filtered for synthesizability (Synthetic Accessibility score <4) to
1257 ensure a fair comparison to SyntheMol-RL compounds, which are intrinsically synthetically

1258 accessible. RL-Chemprop generated 495 hits (total = 10,983), RL-MLP generated 356 hits (total
1259 = 9,227), and GFlowNet generated 313 hits (total = 10,304). These hits were compared based
1260 on three properties of interest: a, *S. aureus* score, b, log solubility score, and c, molecular
1261 weight. Representative samples are also visualized for d, RL-Chemprop, e, RL-MLP, and f,
1262 GFlowNet.

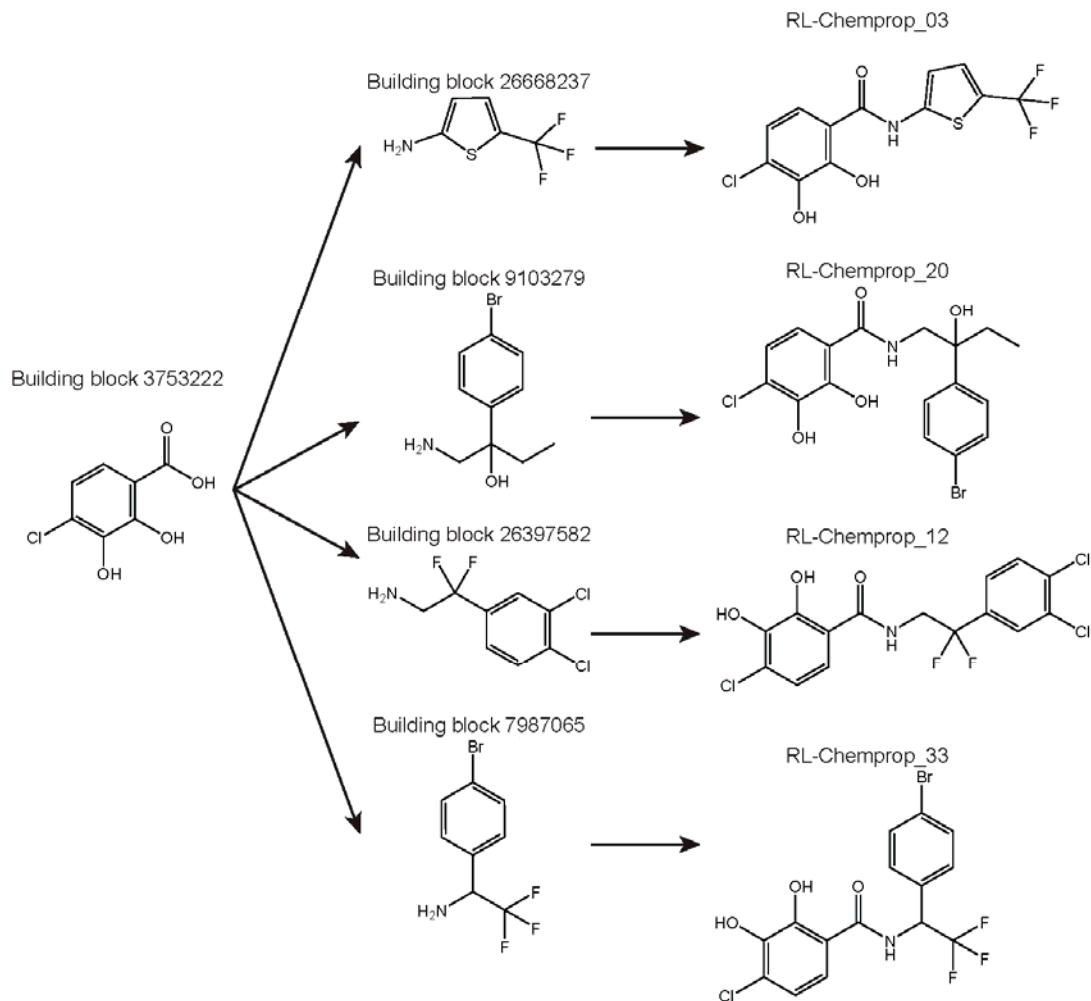


1263

1264 **Extended Data Figure 4. Dorsal surface of mice with an *S. aureus* wound infection model.**

1265 Representative images of the dorsal surface of mice after 4 h, 8 h, and 20 h of treatment with

1266 vehicle or synthecin. Note the marked inflammation on the vehicle treated animals.



1267

1268 **Extended Data Figure 5. Example of molecular generations centered around particular**
1269 **building blocks.** Depicted is building block 3753222, which SyntheMol-RL used multiple times
1270 when constructing hit compounds. Notably, all four compounds shown were validated to have
1271 an MIC $\leq 8 \mu\text{g/ml}$ in *S. aureus* RN4220. SyntheMol-RL seemingly employs a strategy of finding a
1272 particularly promising building block, then to satisfy the diversity objective it expands on that
1273 building block with a diverse set of “secondary” building blocks.