UKRI Centre for Doctoral Training in
Artificial Intelligence for Healthcare

Imperial College London

Tutorial Report

A Deep Learning approach to antibiotic discovery

Stokes et al., 2020

Anisia Talianu

April 1, 2025

1 Paper details

Title: A deep learning approach to antibiotic discovery [7]

Authors: Jonathan Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M Donghia, Craig R MacNair, Shawn French, Lindsey A Carfrae, Zohar Bloom-Ackermann, Victoria M Tran, Anush Chiappino-Pepe, Ahmed H Badran, Ian W Andrews, Emma J Chory, George M Church, Eric D Brown, Tommi S Jaakkola, Regina Barzilay, James J Collins.

Publication details: *Cell*, 20th February 2020

2 Background

Antibiotic agents are the cornerstone of modern medicine. Their discovery and employment has revolutionised the management of infectious diseases, significantly improved quality of life and prolonged life expectancy. However, the inappropriate use of antimicrobial agents has been an important driver of antimicrobial resistance, the process by which bacteria mutate to become resistant to antibiotic agents. In order to keep up with the increasing number of resistant bacteria there are increased efforts to discover new antibiotic agents. However, the process of antibiotic discovery is very long, costly and has mostly failed thus far.

In the paper "A deep learning approach to antibiotic discovery", the authors employ a pioneering method of using deep learning (DL) to identify novel antibiotic compounds by predicting antibacterial properties based on chemical structures. The aim is to optimise the process of antibiotic discovery by reducing the time and cost of discovery using DL methods. The aim of the study is to develop a DL model that would be applied during early drug discovery to score compounds based on the likelihood of exhibiting antimicrobial properties. These predictions would be made *in silico* based on the chemical structure of the drug molecules, which would then be tested *in vitro* and *in vivo* to confirm efficacy. The paper was published in the high-impact journal *Cell* in February 2020 and has led to the discovery of a novel antibiotic, Halicin.

3 Paper Content

3.1 Context and Motivation

Antimicrobial resistance (AMR) is the process by which bacteria develop or acquire mechanisms of resistance to antimicrobial agents. AMR is now a major global health emergency projected to be responsible for 10 million deaths a year by 2050 [4]. The main driver of resistance is the use of antibiotics to treat human infections, whether that use is appropriate or inappropriate. Once a bacterial strain develops resistance to an antibiotic agent, the compound essentially becomes obsolete for the treatment of infections caused by that pathogen. Currently, the rate at which antibiotic resistance develops surpasses that of antibiotic discovery, which brings about the daunting prospect of eventually running out of therapeutically effective antibiotics.

The impact of AMR and the antibiotic crisis is many fold. Firstly, we will no longer be able to treat common infections, increasing risk of mortality and morbidity. These infections caused by resistant strains will be more likely to spread, posing major risks for population health. Fi-

nally, the development of AMR has strong socioeconomic implications due to an increased cost of care, in-hospital length of stay, and an increase in the workload of healthcare professionals.

The most productive era of antibiotic discovery was from 1940s to mid 1960s [8](Figure 1.) and was known as the golden era of antibiotic discovery. These classes of antibiotics were discovered through screening of natural products, particularly secondary metabolites synthesised by soil-dwelling microbes. However, by the mid 1960s our ability to discover novel antibiotics using this method was significantly diminished as it relied on the diversity of naturally available compounds which had already been exhausted. Therefore, antibiotic discovery approaches turned to high throughput screening of semi-synthetic derivatives, a process by which the scaffold of a natural compound is modified to increase efficacy and potency, and decrease toxicity. However, the chemical space, or the vast theoretical set of all possible chemical compounds that could be screened for therapeutic activity is so enormous, making it difficult to predict which modifications will work. This often leads to more failures than successes when engineering next-generation versions of existing natural antibiotics. Finally, discovering entirely synthetic antibiotics involves screening very large chemical libraries to find compounds that limit bacterial growth. This is a trial-and-error process using high-throughput screening (HTS) techniques where thousands to millions of chemical compounds are screened for antibacterial properties through an automated method. HTS is very expensive and slow, and chemical libraries need to be specifically curated. Unfortunately, since the implementation of HTS in the 1980s, no new clinical antibiotics have been discovered using this method.

Overall, the process of antibiotic discovery is slow, expensive, and inefficient. The most common problem encountered in antibiotic discovery is that of dereplication, where the same molecules or structurally similar molecules are being repeatedly discovered rather than molecules with truly novel molecular structures. There is therefore an urgent need to develop novel approaches that increase the rate of drug discovery and decrease the cost.

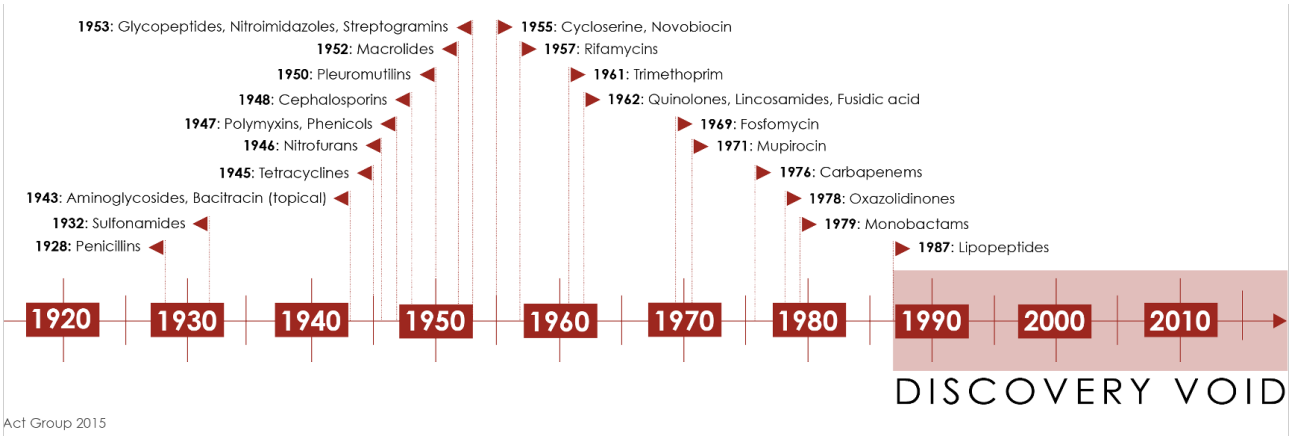

Figure 1: **Antibiotic discovery timeline.** Time-line of the discovery of antibiotic classes that are in clinical use. The discovery void is the period starting 1987 marked by the discovery of the last antibiotic class that was successfully approved for therapeutic use.[8]

3.2 Deep Learning Methodology

This paper explores the application of a DL model, Chemprop, to overcome the current limitations in the antibiotic discovery pipeline. The aim would be to shift early drug discovery from HTS to computational simulations. Performing early drug discovery stages *in silico* would allow

the exploration of vast chemical spaces that are beyond the reach of experimental approaches.

In order to find algorithmic solutions for molecular property prediction, molecules need to be represented as vectors. In chemoinformatics, molecules were represented using fingerprints and descriptors. Fingerprint representations are usually binary vectors used to log the presence or absence of molecular substructures at a given location in the molecule. On the other hand, descriptors are vectors that contain molecular properties thought to be relevant to the task of interest.

Fingerprint and descriptors rely on expert knowledge and need to be designed manually. This approach is time-consuming, and it is very difficult to predict which molecular properties actually hold predictive value. As a result, these models did not achieve enough accuracy to replace traditional drug discovery methods.

The innovative approach applied in this article consists in not only automatically learning the mapping of the vectorised molecular representation to the target property using AI approaches, but also learning the vectorised representations of the molecules itself (Figure 2.) This shift from the deterministic approach to molecular representations to learned molecular representations holds the potential to gain in prediction accuracy. The authors employed message-passing graph neural network (MP-GNN) techniques to derive the learned molecular representations from their string representations, a method discussed in future sections.

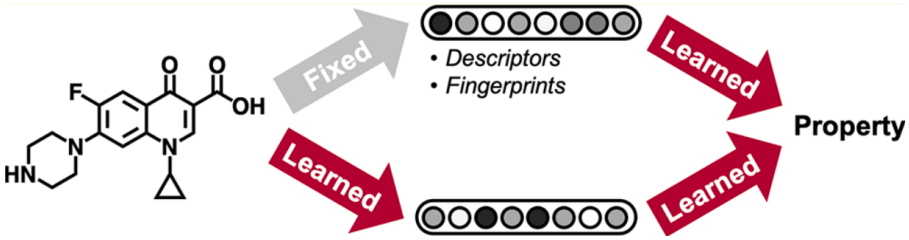

Figure 2: Deterministic vs. Learned molecular representations for property prediction [10]

3.2.1 Message-Passing Graph Neural Networks

Graph neural networks (GNNs) are a very useful way of representing data that is associated with an underlying structure, such as molecules and chemical structures. GNNs consist of nodes and edges. In molecular graphs, the nodes correspond to the atoms whilst the edges represent the bonds formed between the atoms of a molecule. Each node is associated with an initial feature vector containing atom and bond features, derived using the RDKit Python library [5]. In order for the GNNs to learn from the data, feature attributes and connectivity information of a graph is represented using feature matrices. In brief, node feature matrices contain the features for all the nodes in the graph, whilst an adjacency matrix keeps track of which nodes are connected in the graph to define how the GNN will aggregate information from surrounding neighbours.

Once the feature and adjacency matrices are initialised, a series of message-passing or graph convolution steps occurs, during which information is passed between atoms (nodes) through their bonds (edges) allowing the model to learn molecular structures directly. Each node collects information from its directly connected neighbours by summing, averaging or applying a weighted function to combine neighbour features. At each message-passing step, which represents a layer of the GNN, each node collects information from further nodes to create a learned

molecular embedding. At the end of the convolution operations, the MP-GNN vectorised representations of local chemical regions are summed into a single continuous vector to capture the complexity of an entire molecule thereby providing a wholistic representation.

3.2.2 Chemprop model architecture

The Chemprop DL framework to predict molecular properties consists of two sub-parts. The MP-GNN encoder learns task-specific molecular representations from the molecular graph. Molecular representation are then fed into a feed forward neural network (FFNN) decoder that predicts the target chemical property by learning the mapping from molecular representation to antibiotic activity (Figure 3.). The use of MP-GNN in the Chemprop architecture provides several benefits. By learning molecular features automatically, handcrafted descriptors and fingerprints are no longer needed. Additionally, model generalisation is improved as the model now captures complex molecular relationships.

Furthermore, there are several reasons why MP-GNNs are a well-suited architectural choice for molecular property prediction tasks. First, GNNs inherently capture the underlying structural information in molecules by representing atoms as nodes, and atomic bonds as the edges connecting nodes. In addition, the message-passing mechanism is very effective in modeling the atomic interactions within the molecule. Message passing allows the model to capture both local chemistry as well as global information with regards to the structure, geometry, and bond types present within the molecule. These molecular features are crucial for determining chemical properties and biological activity. Finally, MP-GNNs are interpretable and scalable to large molecules, making them powerful tools for molecular property prediction and drug discovery.

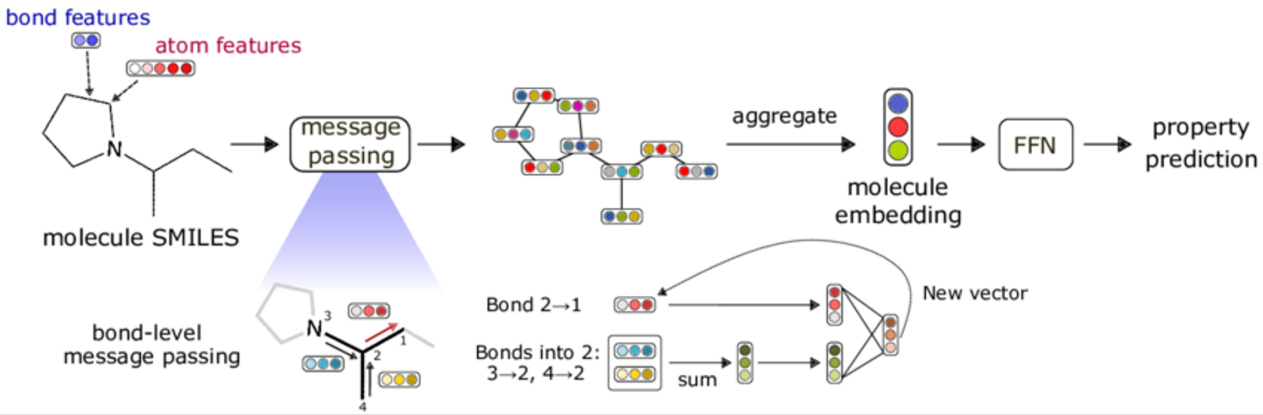

Figure 3: Deep learning framework of the Chemprop model [3]

3.2.3 Experimental approach

The researchers adopted a 3-stage approach for tackling the antibiotic discovery problem. Firstly, they trained the Chemprop model to predict growth inhibition of E.Coli using a training set of 2,335 known molecules. These molecules were curated from a library of FDA-approved drugs as well as purified natural products. The graph structure of every molecule was tied to an activity score (0/1) denoting whether the compound exhibited antibiotic properties, which was determined empirically in the lab (Figure 4). The trained model was then applied to the ZINC15 drug repurposing library which contains over 107 million molecules to identify lead compounds [2]. The compounds were ranked based on the model's predicted score. Lastly, lead compounds were selected for empirical testing *in vitro* and *in vivo*.

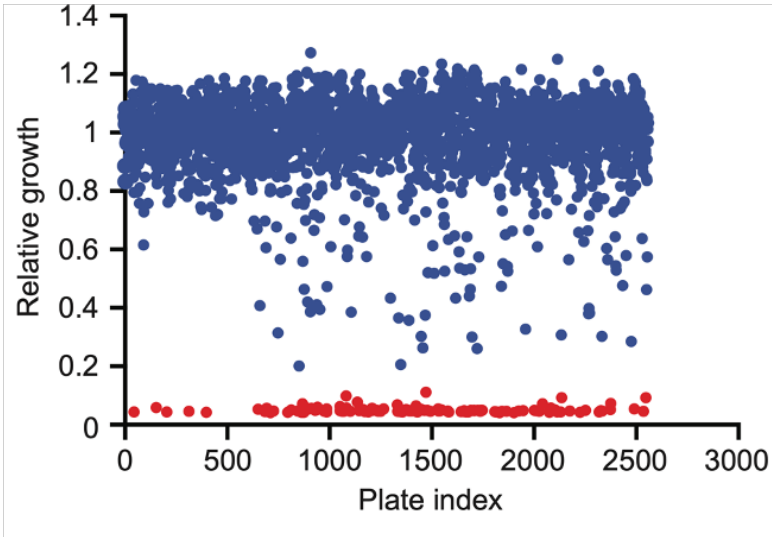

Figure 4: E.Coli growth assay of 2,335 compounds FDA-approved molecules and natural products used to derive the Chemprop training set

3.3 Empirical Validation

3.3.1 Empirical validation of model predictions

The 99 molecules assigned the highest prediction scores by the Chemprop algorithm were selected and tested empirically in the lab to determine whether they indeed displayed growth inhibitory activity against E.Coli. Bacterial growth inhibition assays revealed that 51 out of the 99 tested compounds exhibited antibiotic properties (Figure 5). This is a necessary yet not sufficient molecular feature for candidate antibiotics. As previously discussed, the ideal molecular candidate would also need to be structurally different to conventional antibiotics in order to overcome the challenge of dereplication. The molecular candidate that fulfilled both criteria and was selected for further empirical analysis was Halicin.

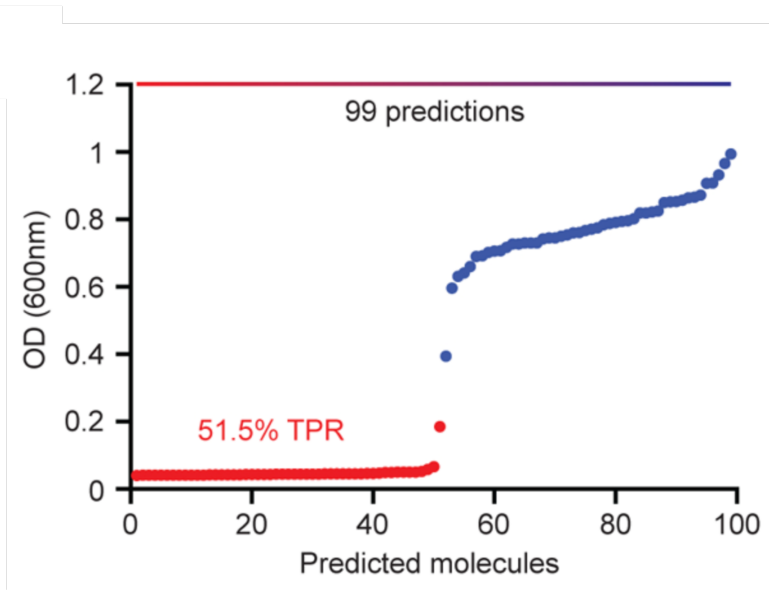

Figure 5: Screening of the 99 candidate compounds with the highest predicted algorithmic score for E.Coli growth inhibition. True Postive Rate (TPR), Optical Density(OD)

3.3.2 Halicin Is a Broad-Spectrum Bactericidal Antibiotic

Dose response curves for growth inhibition of E.Coli revealed that Halicin was a potent antibiotic, and that it not only had bacteriostatic properties in E.Coli cultures but it was in fact bactericidal. The authors then set to verify whether Halicin would retain activity against drug-resistant pathogens. They incubated clinical isolates of carbapenem-resistant *Enterobacteriaceae* (CRE), *A. baumannii*, and *Pseudomonas aeruginosa* with Halicin. These are clinically problematic pathogens labeled as needing urgent antibiotic treatment by the World Health Organization. Halicin showed strong inhibitory activity against both *Enterobacteriaceae* and *A. baumannii*, making Halicin a broad-spectrum agent that retains activity in multiple bacterial strains.

3.3.3 Halicin Dissipates the pH Component of the Proton Motive Force

Next, the authors conducted a set of experiments aimed at elucidating the mechanims of action (MOA) underlying Halicin's antiboitc activity. For this, they deliberately attempted to induce resistance in E.Coli bacterial strains by treating cells with sub-therapeutic concentrations of Halicin for a prolonged period of time in order to study the differences in gene expression in resistant and non-resistant bacteria. However, during a 30-day experiment in liquid culture the team were unable to evolve resistance to Halicin, whilst they were readily able to develop resistance using a conventional antibiotic, ciprofloxacin.
The researchers therefore turned to RNA-sequencing methods to look at the gene expression difference in E.Coli cells before and after Halicin treatment. They observed a downregulation in genes involved in locomotion, flagellar biosynthesis and membrane protein complexes, which is often associated with the dissipation of the pH component of the proton motive force, the electrochemical gradient across the extracellular membrane of E.Coli, which cells must maintain for viability.

3.3.4 Halicin Displays Efficacy in Murine Models of Infection

Given that Halicin displayed broad-spectrum bactericidal activity and was not susceptible to developing resistance, the authors next asked whether the compound might have utility as an antibiotic *in vivo*. For this, they tested Halicin in two murine models of infection, an *A Baumanni* model of skin infection, and a *C. difficile* model of gut infection. Halicin was topically administered to the skin infection model which resolved after 24h of treatment. Similarly, systemic Halicin administration was able to eradicate the *A.Baumannii* infection within 3-4 days of treatment (Figure 6). These results therefore demonstrated the *in vivo* efficacy of the drug molecule.

4 Strength, Weaknesses and Future Perspectives

4.1 Strengths

The paper clearly demonstrates the potential of AI in drug discovery and provides the first compelling study showing that DL can be used to identify structurally new antibiotics thereby accelerating the drug discovery pipieline. By relying on the Chemprop model's predictions to find initial compound leads that are then empirically validated using biomedical methodologies, the authors highlight the potential of multi-disciplinary strategies to solving complex medical problems.
A further strength of this study resides in the methodical, step-wise approach to conducting research through the formulation of sequential hypotheses driven by experimental findings.

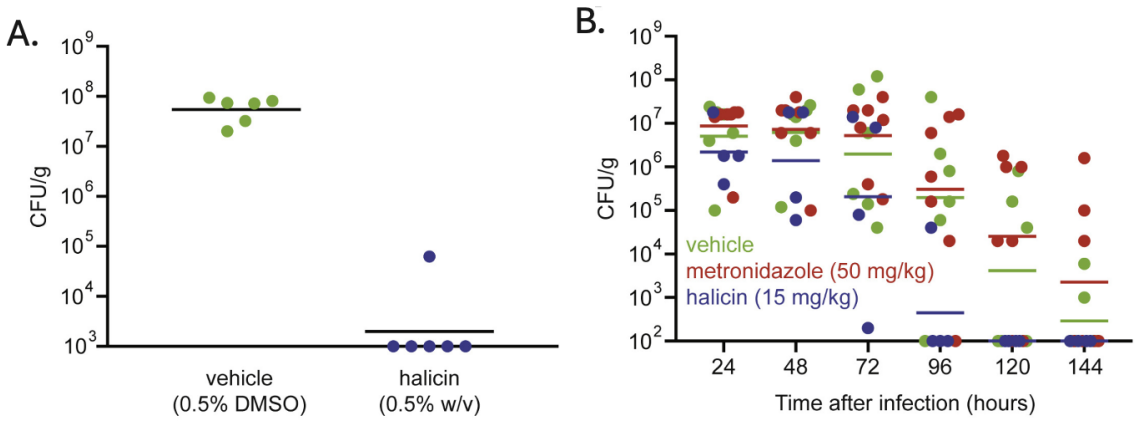

Figure 6: **A. Wound infection murine model** Mice were infected with A. baumannii and treated with either vehicle (green) or halicin (blue) **B. Gut infection murine model** Mice were infected with *C. difficile* and treated with Metronidazole (red), vehicle (green) and halicin (blue)

Finally, the research presented in this paper led to the discovery of a structurally and functionally novel antibiotic, Halicin, the first antibiotic to be discovered through the application of AI methods.

4.2 Weaknesses

Despite indisputable strengths, it is worth addressing some of the limitation of the discussed study. Firstly, the Chemprop model was trained on a relatively small dataset of only 2,335 molecules which is unlikely to be sufficient in encompassing the full chemical diversity of potential antibiotics. These data limitations may restrict model generalisation to entirely new chemical spaces. Further limitations are related to the lack of explainability in model outputs. The model identifies compounds that are likely to have growth inhibitory properties against E.Coli, however, Chemprop does not provide any insights into the potential mechanism of action of candidate molecules. This means that more extensive follow-up experiments are required to elucidate molecular mechanisms.

Non-inclusion of pharmacokinetic and toxicity profiles of molecular compounds during the training phase poses additional challenges. Indeed, training examples were derived from *in vitro* experiments on bacterial cultures and do which does not directly translate to clinical efficacy *in vivo*. While Halicin was shown to be effective in mouse models, broader *in vivo* studies were not conducted. This raises questions about safety, pharmacokinetics, and efficacy in more complex systems or against a broader range of infections. For example, some compounds may exhibit bactericidal properties but have limited bioavailability, or high toxicity, therefore limiting their utility for clinical applications. Pharmacokinetic profiling and safety studies remain a bottleneck in pharmacological research, further delaying antibiotic discovery. To address these challenges, incorporating pharmacokinetic and toxicity information along with developing explainability methods would further allow the Chemprop model to increase the rate of antibiotic discovery.

4.3 Future perspectives

This paper is a milestone in the integration of AI into drug discovery. The authors published a follow-up study in 2023 in which the Chemprop model was used on a different bacterial strain, *A.Baumannii*. This study led to the discovery of Abaucin, a narrow-spectrum antibiotic, and

showed the continued utility of Chemprop in identifying potent antibacterial agents [6].

In addition, the research team published a paper exploring the integration of explainable AI (XAI) methodology to address some of the limitation of the initial study. They hypothesised that model predictions could be explained on the level of chemical substructures. They used tree searching to determine which part of the molecule had the highest positive predictive value and therefore was likely to be responsible for the therapeutic activity of the compound. This improves transparency of model predictions and allows for more efficient exploration of chemical spaces based on the newly identified subgroups. [9]
Of note, in 2024, a few years following the above-mentioned publications, the U.S. Department of Health and Human Services announced funding for projects leveraging AI to accelerate antibiotic discovery [1], reflecting a growing recognition of AI's role in addressing antimicrobial resistance.

5 Acknowledgements

I would like to acknowledge the support of UK Research and Innovation (UKRI Centre for Doctoral Training in Digital Healthcare grant number EP/Y030974/1) for funding this work and would like to thank Dr. Ahmen Fetit for his guidance and support throughout the AI for Health research tutorial.

References

[1] ARPA-H. Arpa-h project to accelerate discovery and development of new antibiotics, 2025. Accessed: 2025-04-01.

[2] David K. Duvenaud, Dougal Maclaurin, J. Adam L. Donnelly, D. D. W. Ong, and R. P. Adams. Convolutional networks on graphs for learning molecular fingerprints. *Journal of Chemical Information and Modeling*, 52(11):3149–3157, 2012. Accessed: 2025-04-01.

[3] Xiaodong Liu, Yi Wang, Cheng Wang, Hui Wang, Shanshan Li, Hao Li, and Qianqian Zhang. Deepchem: A deep learning framework for chemistry. *Journal of Chemical Information and Modeling*, 63(12):4090–4099, 2023. Accessed: 2025-04-01.

[4] Mohsen Naghavi, Stein Emil Vollset, Kevin S Ikuta, Lucien R Swetschinski, Authia P Gray, Eve E Wool, Gisela Robles Aguilar, Tomislav Mestrovic, Georgia Smith, Chieh Han, et al. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *The Lancet*, 404(10459):1199–1226, 2024.

[5] RDKit. Getting started with rdkit in python, 2021. Accessed: 2025-04-01.

[6] Jonathan M. Stokes, Gary Liu, Kevin Yang, Alexia M. Swanson, Joshua S. Collins, Paul W. Brown, and James J. Collins. Deep learning-guided discovery of an antibiotic targeting *Acinetobacter baumannii*. *Nature Chemical Biology*, 2023.

[7] Jonathan M. Stokes, Kevin Yang, Kyle Swanson, Marius S. Jin, Wendy Cubillos-Ruiz, Anna C. Donghia, Craig R. MacNair, Sebastian French, Hamza Carfrae, Emma M. Bloom-Ackermann, et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4):688–702.e13, 2020.

[8] The React Group. The world needs new antibiotics â so why aren't they developed?, 2021. Accessed: 2025-04-01.

[9] Felix J. Wong, Hongzhuo Zheng, Jonathan M. Stokes, Kevin Yang, Kyle Swanson, Gary Liu, Wengong Jin, James J. Collins, Regina Barzilay, and Tommi S. Jaakkola. Discovery of a structural class of antibiotics with explainable deep learning. *Nature*, 615:186–190, 2023.

[10] Kevin Yang, Kyle Swanson, Wengong Jin, Connor Coley, Philipp Eiden, Hua Gao, Angel Guzman-Perez, Timothy Hopper, Brian Kelley, Miriam Mathea, Andrew Palmer, Volker Settels, Tommi Jaakkola, Klavs Jensen, and Regina Barzilay. Analyzing learned molecular representations for property prediction. *Journal of Chemical Information and Modeling*, 59(8):3370–3388, 2019.