

RESEARCH ARTICLE

Fung-AI: An AI/ML-driven pipeline for antifungal peptide discovery

Daniel S. Berman¹, Libby M. Lewis¹, Tom D. Curtis¹, Olivia N. Tiburzi¹, Daniel F. Q. Smith², Arturo Casadevall², Laura J. Dunphy^{1*}

1 Johns Hopkins Applied Physics Laboratory, Laurel, Maryland, United States of America, **2** W. Harry Feinstone Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, Maryland, United States of America

* Laura.Dunphy@jhuapl.edu



Abstract

Emerging fungal pathogens represent a concerning threat to both global health and food security. In this study, we aimed to address our rising vulnerability to fungal pathogens through the development of the Fung-AI pipeline: an AI/ML-driven approach for antifungal discovery. A generative adversarial network (GAN) was trained to generate novel candidate antifungal peptide sequences. Next, *in silico* antifungal and hemolytic classifiers were built to further prioritize AI-generated peptides for experimental validation. From a pool of ~10,000 candidates, thirteen peptides were selected for testing over two-stages of experimentation. Five peptides were found to display mild antifungal activity against the wheat pathogen, *Fusarium graminearum*, with minimal inhibitory concentrations (MICs) ranging from 250 µg/mL to 500 µg/mL. Four of the five peptides also showed activity against the human pathogen, *Candida albicans* (MIC: 500 µg/mL). Two of our AI-generated antifungal peptides additionally demonstrated low cytotoxicity in HepG2 human liver carcinoma cells ($LC_{50} > 704.2$ µg/mL) indicating that they may be useful as scaffolds for future optimization for therapeutic applications. None of our peptides were found to considerably inhibit the emerging pathogen *C. auris*, suggesting the need for pathogen-specific down-selection of candidate peptides. Overall, we present a proof-of-principle, generative-AI-based approach for the rapid design of *de novo* antifungal peptides.

OPEN ACCESS

Citation: Berman DS, Lewis LM, Curtis TD, Tiburzi ON, Smith DFQ, Casadevall A, et al. (2026) Fung-AI: An AI/ML-driven pipeline for antifungal peptide discovery. PLoS Comput Biol 22(6): e1014105. <https://doi.org/10.1371/journal.pcbi.1014105>

Editor: Dirk Walther, Max Planck Institute of Molecular Plant Physiology: Max-Planck-Institut für molekulare Pflanzenphysiologie, GERMANY

Received: March 6, 2026

Accepted: May 29, 2026

Published: June 15, 2026

Copyright: © 2026 Berman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All data and code used in model development and figure generation can be found at <https://github.com/jhuapl-bio/fungai>.

Author summary

As humans, we are more biologically similar to fungi than we are to bacteria or viruses. This resemblance means that it is inherently difficult to find drugs that uniquely target fungal pathogens without causing side effects in people. Consequently, there are only five approved classes of antifungal drugs on the market today. As a result of limited therapeutic options as well as the emergence of drug resistance, fungal infections are notoriously difficult to treat. Antifungal peptides,

Funding: This research was supported by internal funding from the Johns Hopkins University Applied Physics Laboratory (JHU/APL). Authors DSB, LML, TDC, ONT, and LJD were employed at JHU/APL at the time of this work and their work on the study was performed within the scope of that employment. DFQS and AC were supported in part by the NIAID (U19-AI189183) which had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

short chains of ~2–50 amino acids, represent a promising alternative to traditional small molecule drugs, however, the possible search space for these drugs is intractably large. Here, we sought to leverage advancements in artificial intelligence and machine learning to accelerate the design of novel antifungal peptides. We first trained a generative AI model on publicly available examples of known antifungal and non-antifungal peptides. We then asked this model to suggest thousands of new “antifungal-like” peptides. Candidates were computationally down-selected, and promising peptides were synthesized and experimentally tested against both human and agricultural pathogens. Our main finding was that some, but not all, AI-generated peptides were active against fungal pathogens and were not notably harmful to human liver cells, highlighting that generative AI can be used to design novel antifungal drugs.

Introduction

A major threat of growing global concern, fungal pathogens have been estimated to cause at least 2.5 million infection-related deaths per year [1–3], while additionally driving the spoilage of between 10–23% of annual pre-harvest crop yields [4]. The negative impact of fungal pathogens has been exacerbated in recent years by both a general lack of known classes of antifungal drugs, as well as the emergence of multi- and pan-drug-resistant clinical isolates [5–8]. Compounding upon this issue, the discovery of novel antifungal drugs is slow and challenging, in part because fungi share much of their cell biology with mammalian cells [9,10].

Antifungal peptides, short polypeptides that are typically 2–50 amino acids in length, are thought to be less likely to cause fungal resistance, and therefore represent a promising alternative to traditional small molecule drugs [11,12]. Antifungal peptides, many of which are naturally produced by bacteria, plants, and animals, can kill fungal cells through a variety of mechanisms, including cell membrane disruption, lysis of the cell wall, and targeting of critical intracellular processes [12]. Published examples have demonstrated strong activity against relevant human pathogens, including *Candida albicans*, *Aspergillus fumigatus*, and *Cryptococcus neoformans* [10]. For example, the archaeal peptide, VLL-28 and a peptide isolated from the giant monkey frog, Skin-PYY, have been shown to have minimal inhibitory concentrations (MICs) of 88.5 µg/mL and 25 µg/mL against *C. albicans* [10,13,14], respectively. While some antifungal peptides can also exhibit high toxicity in mammals, this toxicity can sometimes be mitigated with the design of semi-synthetic or synthetic analogues [10,15].

Over the last decade, the rapid evolution of generative artificial intelligence (AI) has resulted in the explosion of AI-enabled research across the biological sciences. Generative modeling approaches, such as generative adversarial networks (GANs) [16], variational autoencoders (VAEs), and diffusion models [17], can be trained to turn random noise into realistic DNA, RNA, and protein sequences. For example, used in conjunction with other deep learning architectures, GANs have been applied

to predict viral evolution [18], discover new enzymes [19], and optimize DNA and protein sequences for specific functions [20,21]. More recently, deep learning has been used successfully to both detect (e.g., with discriminative classification models) [22–27] and design (e.g., with generative models) novel antimicrobial peptides [28–32]. However, while there has been some success in developing highly accurate AI/ML models for the classification of antifungal peptides [33–39], the direct application of generative AI to antifungal discovery has been much more limited, at least in part due to a lack of large, high-quality fungal datasets [40].

Given the recent success of others toward developing high-accuracy antifungal peptide classifiers, we hypothesized that, while limited, there could be sufficient data in the public domain to train an antifungal peptide generative model for the task of drug discovery. To this end, we present the Fung-AI pipeline, an *in silico* approach for antifungal discovery (Fig 1). The Fung-AI pipeline begins with a GAN, which has been trained on publicly available known antifungal and non-antifungal peptide sequences to rapidly generate diverse and novel candidate peptide sequences. Peptides are then down-selected based on predicted activity, toxicity, structure, as well as computed biophysical properties (e.g., net charge, hydrophobicity). To validate our approach, thirteen candidate peptides were synthesized and screened for activity against a panel of four fungal species, including both plant and human pathogens. Of the five peptides with antifungal activity, two were found to also have low cytotoxicity in a human liver cell line. Applied on a larger-scale, our computational, semi-automated approach has the potential to help fill the void in currently available antifungal countermeasures, addressing our rising vulnerability to fungal pathogens.

Results

In silico generation and down-selection of candidate antifungal peptides

A total of 9,994 unique peptide sequences ranging from 10–35 amino acids in length were computationally generated with our custom GAN (S1 Text). Following peptide generation, we applied a series of in-house and publicly available *in silico* tools to further screen for generated peptides that were the most likely to be antifungal, non-hemolytic (e.g., not obviously toxic to humans), and physiologically realistic. In addition, the novelty of promising candidates was assessed by querying generated peptides against known proteins in the NCBI Non-Redundant (nr) Protein Database. Cumulatively, in the first phase of this effort, the Fung-AI pipeline (Fig 1), was used to down-select five candidate peptides for experimental validation. More specifically, generated peptides were sequentially removed and assessed throughout the Fung-AI pipeline (Fig 2A). First, all 9,994 peptides were passed through three binary classifiers adapted from literature [36,37], which were trained to predict whether an individual peptide would have antifungal activity (Fig 2B–2D). Using a probability cutoff of 0.5 for each model, a total of 3,578 peptides were predicted to be antifungal by all three classifiers. Notably, the probability

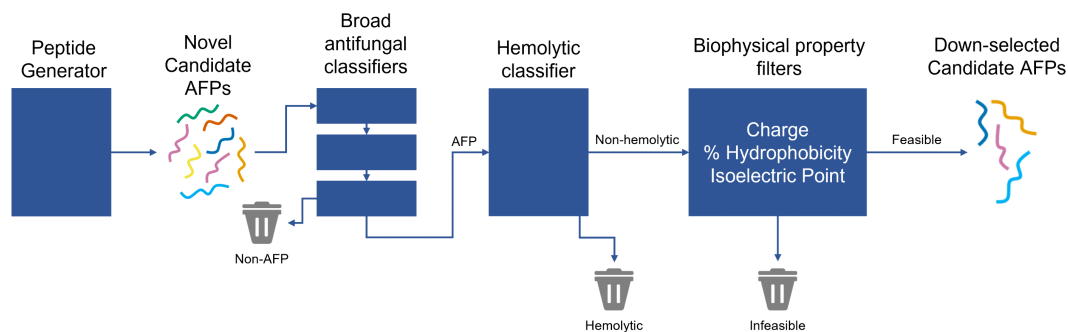


Fig 1. High-level overview of Fung-AI pipeline for antifungal discovery. Candidate antifungal peptides (AFPs) were generated and then down-selected by passing them through three antifungal classifiers and a hemolytic classifier. AFP-like, non-hemolytic candidates with reasonable biophysical properties and diverse secondary structures were retained for *in vitro* antifungal and cytotoxicity testing.

<https://doi.org/10.1371/journal.pcbi.1014105.g001>

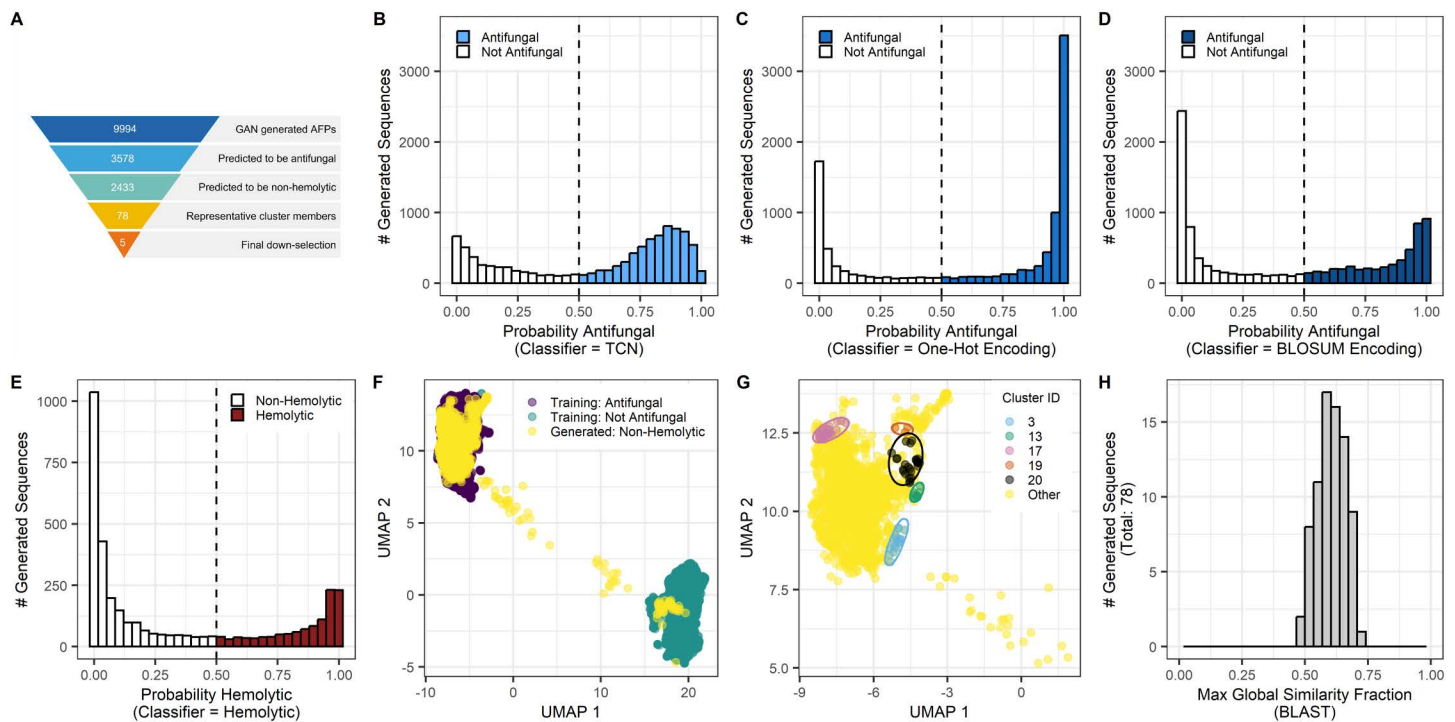


Fig 2. Down-selection of AI-generated candidate peptides through the *in silico* Fung-AI pipeline. **A)** Overview of the process employed to down-select candidate peptides for experimental testing. Probability distributions of 9,994 peptides being antifungal as predicted by the **B)** TCN-based, **C)** one-hot encoding-based, and **D)** BLOSUM encoding-based binary antifungal classifiers. **E)** Probability distributions of 3,578 peptides being hemolytic as predicted by the binary hemolytic classifier. Classifier cutoffs were set at 0.5, denoted by vertical dashed lines. **F)** UMAP dimensionality reduction and visualization of training data compared to predicted antifungal and non-hemolytic AI-generated peptides. **G)** UMAP of five down-selected HDBSCAN clusters, containing 78 peptides of interest. Yellow denotes the rest of the generated peptides that were “AFP-like”. **H)** A total of 78 peptides representative of clusters of interest were run through BLAST. The max global similarity fraction denotes the “best” match for each generated sequence from the NCBI Non-Redundant (nr) Protein Database. The maximum possible score is 1 and the minimum possible score is 0.

<https://doi.org/10.1371/journal.pcbi.1014105.g002>

distributions of each classifier were concentrated at the tails, indicating strong predictions for the majority of the generated peptides, as opposed to weak predictions around the classification cutoff.

A common challenge when developing antimicrobial peptides is that peptides active against microorganisms, especially fungi which are eukaryotic and thus share many characteristics with humans, may also cause cytotoxicity in mammalian cells [12]. Therefore, in an attempt to increase the chances of identifying relatively safe therapeutic candidates, the remaining 3,578 peptides that were predicted to be antifungal by all three classifiers were passed through a binary hemolytic classifier adapted from Yaseen *et al.* [41] to predict whether or not peptides would have hemolytic activity (Fig 2E). Hemolytic peptides were removed, and embeddings of the remaining 2,433 generated peptides were compared with training data using Uniform Manifold Approximation and Projection (UMAP) (Fig 2F). The majority of predicted antifungal and non-hemolytic generated peptides clustered with known antifungal peptides from the training data.

To identify groups of similar peptides that could be sampled for experimental validation, we used hierarchical density-based spatial clustering of applications with noise (HDBSCAN) to identify putative antifungal, non-hemolytic peptide clusters. An HDBSCAN model was trained on known antifungal sequences from the training dataset, omitting non-antifungal sequences. Using a minimum cluster size of 40 peptides to reduce noise, a total of 21 clusters were formed from the antifungal training data (Table 1). These 21 clusters were down-selected to nine clusters of interest, where at least 70% of members in each cluster were predicted to be non-hemolytic. Generated sequences were then mapped to these clusters, with

Table 1. Summary of peptide clusters identified with HDBSCAN. Clusters were determined based on peptides in the antifungal training dataset. Generated peptides were assigned to training clusters. Clusters were considered non-hemolytic if the percent of non-hemolytic peptides from training data was greater than 70%. Italicized rows denote non-hemolytic clusters. Bolded rows denote the final five down-selected clusters.

Cluster ID	Number of Cluster Members from Antifungal Training Data	Percent Non-Hemolytic from Antifungal Training Data	Number of Cluster Members from Generated Sequences
0	40	0.90	0
1	46	0.89	0
2	77	0.79	0
3	66	0.82	17
4	41	0.63	53
5	72	0.47	178
6	204	0.38	165
7	151	0.52	226
8	125	0.35	9
9	53	0.53	172
10	41	0.20	24
11	44	0.16	59
12	151	0.37	265
13	102	0.84	10
14	50	0.84	0
15	71	0.31	1
16	88	0.41	22
17	97	0.88	23
18	210	0.47	6
19	61	0.70	3
20	148	0.95	25
Unassigned	675	0.61	1175

<https://doi.org/10.1371/journal.pcbi.1014105.t001>

a total of 78 generated sequences mapping to five out of the nine clusters, and none mapping to the remaining four clusters (**Fig 2G and Table 1**). When queried against short known proteins (< 250 amino acids), no considerable similarities (> 90% query cover and > 80% percent identity) were found between the top BLAST hits and these 78 sequences, with all max global similarity fractions below 0.75 for hits with at least 50% query cover. This suggests that our AI-generated sequences are truly novel and not simply copies or near-relatives of known natural or synthetic proteins (**Fig 2H**).

Finally, having down-selected to five candidate clusters of interest, one peptide from each cluster was selected based on a combination of biophysical properties, predicted structure, and synthesizability. Following published literature on synthetic therapeutic peptide design [11], we filtered for mildly cationic peptides (e.g., positive net charge below +6) composed of 30–60% amino acids with hydrophobic side chains (**Table 2**). The PEP-FOLD4 server [42] was used to predict the structures of the remaining peptides. For each cluster, one peptide with strong cluster mapping (probability > 0.9) and a qualitatively representative linear structure (e.g., alpha-helical, random coil, etc.) was selected for synthesis (**Fig 3**). Thus altogether, the Fung-AI pipeline is a semi-automated, fully computational approach to generating and priority ranking novel antifungal peptide candidates for experimental testing.

Antifungal activity and cytotoxicity of AI-generated peptides

To evaluate the antifungal potential of candidate peptides, we first measured MICs of our five down-selected AI-generated peptides against the relevant wheat pathogen, *Fusarium graminearum*, and the model organism

Table 2. Biophysical properties of peptides selected for experimental validation. AA=amino acid.

Peptide ID	Cluster ID	Sequence	Isoelectric Point	Net Charge	Fraction Hydrophobic AAs
40	17	KILKRQRLLLGGLLKYLKAGK	11.22	5.76	0.5
50	13	KRRWPPKPVFSPNAAAYKCNP	10.46	4.75	0.33
27	3	GYGTTFIAPPGGLRGALPQFS	8.75	0.76	0.38
63	19	GLVLIKAARAGNLLKR	12	3.76	0.56
23	20	KAALKPKFTGAVTAFGLPQ	10.3	2.76	0.47
12	17	QLLRPLFTKLLGGLSKILKKF	11.33	4.76	0.48
17	17	RLLQALLNLRKLLKGLLSRLS	12	5.76	0.45
48	17	KKTYTKLKKLNGAIKYLVGHA	10.22	5.84	0.43
65	17	KAKLRQAVLLKLLAGVKSSL	11.39	5.76	0.52
20	17	RLKKLWWKTCLLKGCLTLKA	10.21	5.74	0.45
29	17	GLRHYARVKLKLALLNLNKSC	10.46	4.83	0.48
44	17	KLAKLLALRLKSLVFGLRYRT	11.74	5.76	0.57
54	17	KAIPLLRRRSIVLLLRGFGV	12	4.76	0.57

<https://doi.org/10.1371/journal.pcbi.1014105.t002>

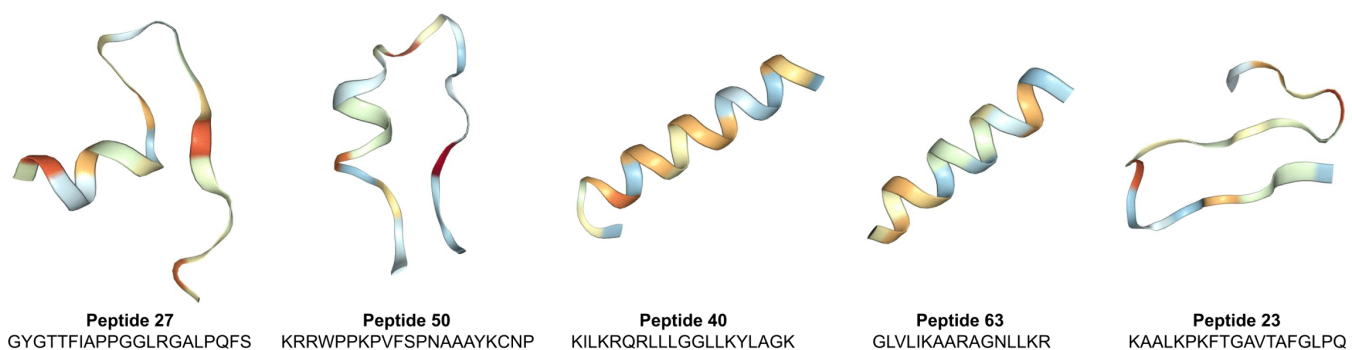


Fig 3. Predicted structures of down-selected AI-generated peptides. Structures predicted with the PEP-FOLD4 server. Residues are colored by hydrophobicity (red = hydrophobic, blue = hydrophilic).

<https://doi.org/10.1371/journal.pcbi.1014105.g003>

Saccharomyces cerevisiae. Peptide 40 from cluster 17 showed inhibitory activity against both organisms, prompting a second phase of testing, during which time the remaining eight peptides from cluster 17 with strong cluster membership and which met down-selection criteria were also synthesized and tested (S1 Fig). A total of five out of nine peptides from cluster 17 showed inhibitory activity against *F. graminearum*. These five peptides were then further tested against an expanded fungal panel of human-relevant pathogens: *Candida albicans* and *Candida auris*. We finally assessed cytotoxicity of the down-selected peptides by determining the LC₅₀ values in HepG2 human liver carcinoma cells, where the LC₅₀ is defined as the lethal concentration that kills 50% of cells.

Peptides 12 and 40 exhibited the broadest and strongest antifungal activity with MICs of 250 µg/mL against *F. graminearum* and 500 µg/mL against both *S. cerevisiae* and *C. albicans* (Table 3). Peptide 48 showed weaker, but equally broad activity, inhibiting *F. graminearum*, *S. cerevisiae*, and *C. albicans* at 500 µg/mL. In contrast, Peptide 17 displayed activity only against *F. graminearum* at 500 µg/mL. Peptide 65 was active against *F. graminearum* and *C. albicans* at 500 µg/mL but had no effect on *S. cerevisiae*. Notably, none of the peptides inhibited *C. auris* at concentrations up to 500 µg/mL (S2 Fig).

Cytotoxicity varied substantially among the peptides (Table 3 and Fig 4). Peptide 12 had the lowest LC₅₀ value (66.06 µg/mL), indicating higher cytotoxic potential in HepG2 cells. In contrast, no toxicity was observed for Peptide 48 at the

Table 3. MIC ($\mu\text{g/mL}$) and LC_{50} ($\mu\text{g/mL}$) activity against a panel of fungi and HepG2 cells, respectively. R^2 denotes the fit of the dose-response curve used to calculate LC_{50} values. NT: No toxicity observed. AmB: Amphotericin B. N/A: Not available.

Peptide ID	Sequence	<i>F. graminearum</i>	<i>S. cerevisiae</i>	<i>C. albicans</i>	<i>C. auris</i>	HepG2 (LC_{50})	Fit (R^2)
12	QLLRPLFTKLLGGLSKILKKF	250	500	500	>500	66.06	0.985
17	RLLQALLNLRKCLKGLLSRLS	500	>500	>500	>500	85.88	0.998
40	KILKRQRLLLGGLLKYLAGK	250*	500	500*	>500*	315.5	0.994
48	KKTYTKLKKLNGAIKYLVGHA	500	500	500	>500	NT	N/A
65	KAKLRQAVLLKLLAGVKSSL	500	>500	500	>500	704.2	0.987
AmB	N/A	1	0.5	0.5	0.5	N/A	N/A

*MIC value determined on both crude purity and >95% purity peptide.

<https://doi.org/10.1371/journal.pcbi.1014105.t003>

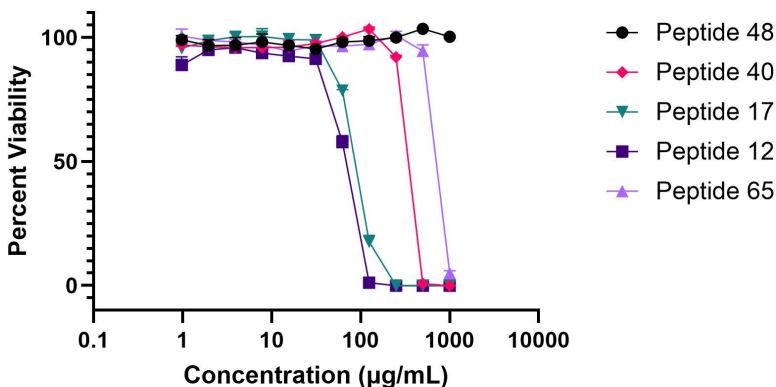


Fig 4. Cytotoxicity of AI-generated peptides in human liver cells. Percent viability of HepG2 cells at increasing concentrations of peptide. Each point denotes the mean and error bars denote the standard deviation across three replicate wells.

<https://doi.org/10.1371/journal.pcbi.1014105.g004>

concentrations tested, and Peptide 65 showed only mild cytotoxicity with an LC_{50} of 704.2 $\mu\text{g/mL}$. Peptide 40 and Peptide 17 exhibited intermediate cytotoxicity with LC_{50} values of 315.5 $\mu\text{g/mL}$ and 85.88 $\mu\text{g/mL}$, respectively.

Overall, peptides 12 and 40 demonstrated modest antifungal activity, particularly against the wheat pathogen, *F. graminearum*. While these peptides were also active against *C. albicans*, their therapeutic window against human pathogens may be limited by cytotoxicity. Peptides 48 and 65, despite showing weaker antifungal activity and being predicted to form cationic alpha-helical structures (Fig 5), may be safer scaffolds for future optimization due to their low cytotoxicity profile.

Discussion

Here, we have developed and experimentally validated the Fung-AI pipeline, an AI/ML supported approach to design *de novo* antifungal peptides. We generated ~10,000 candidate peptides *in silico* with a custom GAN, computationally down-selected hits, and evaluated peptide antifungal activity across a panel of fungal pathogens of relevance in agriculture and human health. Testing fewer than 20 peptides across two phases of experimental validation, we identified one cluster of peptides from which five out of nine synthesized peptides (55%) displayed activity against at least one fungal species. Of these five peptides, two displayed both antifungal activity in *F. graminearum* and *C. albicans* as well as low cytotoxicity in a human liver cell line (e.g., LC_{50} was less than the MIC), suggesting there may be the potential to further optimize these peptides for therapeutic applications (Table 3). Notably, while the sequences of these peptides are distinct from known proteins, they are cationic and predicted to form alpha-helical structures, meaning that they may act in a

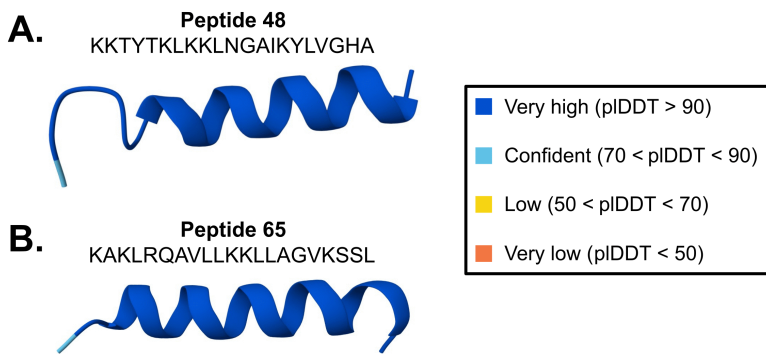


Fig 5. Predicted structures of antifungal and non-cytotoxic AI-generated peptides. Structures were predicted for (A) peptide 48 and (B) peptide 65 with the publicly available AlphaFold3 server. Residues are colored by the predicted local distance difference test (pIDDT) confidence score. Dark blue: very high confidence (pIDDT > 90). Light blue: confident (70 < pIDDT < 90).

<https://doi.org/10.1371/journal.pcbi.1014105.g005>

functionally similar way to known antimicrobial peptides. Overall, we have demonstrated the utility of our peptide generator and the Fung-AI pipeline by discovering novel antifungal peptides with minimal experimental screening.

Through building the Fung-AI pipeline, we have identified four key limitations of our approach as well as opportunities to improve the usefulness of AI-guided approaches for antifungal discovery. First, publicly available antifungal peptide data is limited. For example, while there are over ~35,000 antimicrobial peptides in the data repository of antimicrobial peptides (DRAMP) database, only ~5% of the peptides in DRAMP are labeled as antifungal [43]. Therefore, in order to maximize the size of our training dataset, our GAN was trained on any known antifungal peptides, regardless of target species. While it is interesting that such a heterogeneous training set resulted in the generation of peptides that were broadly active against both *F. graminearum* and *C. albicans*, our peptides had relatively high MICs (Table 3) compared to known antifungal peptides such as VLL-28 (MIC *C. albicans*: 88.5 µg/mL) [13]. Using smaller pathogen-specific datasets to fine-tune our GAN or develop additional antifungal classifiers could potentially increase the chances of generating targeted, high-potency therapeutics. In addition, it may be possible to modify these peptides to increase their activity. Alternatively, a recent study by Wang *et al.* found that training a diffusion model on a broad collection of antimicrobial peptides (antibacterial and antifungal) was sufficient to discover peptides with activity against *C. glabrata* and *Cryptococcus neoformans* [44].

Second, while based on their broad-spectrum activity, positive charge, and predicted linear alpha-helical structures, we hypothesize that peptides 48 and 65 may be membrane disrupting, we did not experimentally determine the mechanism of action (MOA) of any of our AI-generated peptides. Those looking to leverage our pipeline or similar methods should aim to incorporate MOA prediction into the down-selection process. Additionally, newer generative-AI methods such as protein diffusion models [17] or small molecule diffusion models [45,46] could be explored to generate peptides or small molecule drugs against specific fungal targets.

Third, none of the peptides evaluated in this work were found to considerably inhibit the growth of *C. auris*, a fungal pathogen of growing clinical concern that often shows a multi-drug resistance phenotype. *C. auris* was first described in 2009 [47] and consequently, publicly available data sets on this pathogen are fairly scarce. For example, only 78 peptides out of ~24,000 total entries contained in the Database of Antimicrobial Activity and Structure of Peptides (DBAASP) [48] are associated with any strain of *C. auris*. Future studies should aim to collect high-throughput, high-quality multi-modal data (*e.g.*, small molecule antifungal susceptibility, antifungal peptide screens, whole-genome sequencing data, clinical metadata, *etc.*) for this pathogen to enable the development of targeted and effective therapeutic strategies.

Finally, in this study we implemented an *in silico* screening strategy to down-select AI-generated peptides for experimental testing. This strategy prioritized peptides based on their predicted antifungal and hemolytic activities in addition to

their calculated biophysical properties with the goal of finding safe and effective hits. Future efforts should explore both the hit rate of the Fung-AI GAN in the absence of any down-selection method as well as the impact of employing alternative down-selection approaches to prioritize peptides based on additional desired properties of interest such as host specificity or peptide structural diversity.

In summary, the Fung-AI pipeline demonstrates the value of applying generative AI toward the challenge of antifungal drug discovery. Our results motivate the continued exploration, screening, and optimization of AI-generated peptides against relevant agricultural and human fungal pathogens.

Materials and Methods

Construction of the *in silico* Fung-AI pipeline

Dataset curation. Two datasets were created for this work: one capturing antifungal peptides and non-antifungal peptides, and another of hemolytic peptides and non-hemolytic peptides.

The antifungal peptide dataset consisted of 9,388 peptides from published literature [35], DRAMP 4.0 [43], and CAMPR4 [49]. Antifungal peptides from CAMPR4 and DRAMP 4.0 could be naturally occurring or synthetic, but had to have experimentally validated antifungal properties. Peptides only hypothesized to be antifungal were not included in the dataset. A total of 2,204 peptides were gathered from the CAMPR4 database, of which 982 (44.6%) were antifungal and 1,222 (55.4%) were non-antifungal. An additional 1,811 peptides, of which all were antifungal, were pulled from the DRAMP 4.0 database. Finally, 5,373 peptides from Sharma et al. [37] were included, of which 1,932 (36.0%) were antifungal and 3,441 (64.0%) were non-antifungal. In all, the peptides ranged from 2 to 726 amino acids in length, with a mean length of 26.0 ($\sigma=31.5$) and a median length of 20.0 amino acids. As this distribution is heavily skewed with a very long tail, we set a minimum peptide length of 10 and a maximum of 35 amino acids. After removing sequences that were too long or too short, the final dataset contained 7,335 peptides, of which 3,423 were antifungal peptides (46.7%) and 3,912 were non-antifungal peptides (53.3%). The data was split into training and test sets using a random 80/20 split. Training and test sets had similar class balances (46.9% and 45.6% antifungal, respectively), amino acid frequencies (p -value > 0.05, Kolmogorov-Smirnov), and sequence length distributions (p -value > 0.05, Kolmogorov-Smirnov).

A dataset of hemolytic and non-hemolytic peptides published by Yaseen *et al.* [41] was used in the development of our hemolytic classifier. This dataset contained 3,804 peptides, of which 1,576 (41.4%) were hemolytic and 2,228 (58.6%) were non-hemolytic. The mean peptide sequence length was 18.3 ($\sigma=6.2$) amino acids with a maximum length of 35 and a minimum of 10 amino acids. A five-fold cross-validation strategy was used. For each fold, data was split into training and test sets using a random 80/20 split. On average across the five folds, 41.4% of peptides in both the training and test sets were labeled as hemolytic. Kolmogorov-Smirnov tests of the average amino acid frequencies and sequences length distributions showed no statistically significant difference between the training and test datasets (p -value > 0.05).

Antifungal peptide classifiers. Three antifungal peptide classifiers were identified from literature, retrained using our antifungal peptide dataset, and implemented within the Fung-AI pipeline. All three classifiers were used to down-select AI-generated candidate peptides, while one model was additionally used to monitor the behavior of the GAN throughout training. Detailed descriptions of the architectures and training regimens for each model are provided in the **Supporting Information (S1 Text)**.

The first model, based on temporal convolutional networks (TCNs) [50] leveraged an architecture described by Singh et al [36]. Following training, the TCN model achieved an accuracy of 86.8%, area under the curve (AUC) of 92.5%, and F1-score of 86.1% on our antifungal peptide datasets.

The second and third models used the architecture described in Sharma et al. [37], which is based on a one-dimensional convolutional neural network bidirectional long short-term memory (1DCNN-BiLSTM) architecture. The two models were differentiated by the input and the presence of an embedding layer. The second model took as input the one-hot encoded versions of the peptides, forgoing the need for an embedding layer. The third model, more similar to Sharma

et al. [37], used an embedding layer with a dimension of 20. Our embedding layer was seeded with pretrained BLOSUM weights [51], as the pretrained model Sharma et al. [37] used for their embeddings was not available at the time of this work. After training on our antifungal peptide dataset, the second and third models achieved accuracies of 85.5% and 86.3%, AUCs of 92.1% and 92.4%, and F1-scores of 84.6% and 84.7% respectively.

Hemolytic classifier. The hemolytic peptide classifier is based on published work by Yaseen et al. [41], but modified to use the architecture of the third antifungal peptide classifier (1DCNN-BiLSTM, using the input based on BLOSUM weights) [51]. We were able to achieve comparable results to those referenced in the paper using this architecture. The model was trained on a random subset of 80% of the hemolytic peptide dataset for 20 epochs with a batch size of 16, using the Adam optimizer with a learning rate of 0.03. The remaining 20% of the data was used for testing. This process was repeated for five folds, from which we used the best performing model which achieved an accuracy of 81.1%, AUC of 89.1%, and F1-score of 78.3%, as compared to the 88% AUC reported in Yaseen et al. [41].

Generative adversarial network. A GAN is a generative model framework composed of two neural networks, a generator G , and a discriminator D . The two models are trained alternately, such that they are in a zero-sum game in which the generator tries to create realistic looking outputs and the discriminator tries to identify real and fake data. The generator creates fake data, the discriminator is trained on real and fake data, then the discriminator weights are frozen and the generator is trained to fool the discriminator. This process repeats until the model reaches a point where the loss function is not improving with additional training.

Unlike standard GANs, our model has an additional encoder that pairs with the generator, which in turn acts as both a generator and a decoder. As the dataset was somewhat small, which resulted in poor performance when trained solely on antifungal peptides, we applied two strategies to create a larger dataset for training the GAN. First, we trained on both antifungal and non-antifungal peptides, accounting for this in the discriminator loss function. For the purpose of training the discriminator, the non-antifungal peptides were labeled as “fake”. The purpose of this was to push the generator to produce peptides more likely to be antifungal. Second, we trained an autoencoder while simultaneously training the generator, which had the benefit of reducing mode collapse. As a result, the generator, which is also the decoder of the autoencoder, was trained using both autoencoder reconstruction loss and GAN adversarial loss. The model was forced to still be able to reconstruct real antifungal and non-antifungal peptides, thereby enforcing meaning in the embedding space. Thus, there are three models that compose the GAN: the discriminator, the generator/decoder, and the encoder. This setup prevented mode collapse and enabled the generation of diverse output sequences.

Following model training, the GAN was used to generate 10,000 peptides using standard noise as input. A stochastic approach was implemented during this process to increase peptide diversity. All 10,000 peptides were within the desired range of 10–35 amino acids, with no repeated sequences and no overlap with training or test data. Six peptides did not end on an end token and were removed, resulting in a final count of 9,994 generated peptide sequences.

Detailed descriptions of the GAN architecture, loss function, training methods, and peptide generation process are provided in [S1 Text](#).

Unsupervised clustering of peptide sequences. The BLOSUM embedding layer from the 1DCNN-BiLSTM antifungal classifier was used to generate embeddings for both the known and GAN-output peptide sequences. UMAP was then used to reduce the peptide embeddings down to two dimensions, fitting the UMAP model to classifier training data [52]. Following feature compression with UMAP, the hierarchical density-based spatial clustering of applications with noise (HDBSCAN) algorithm was used to cluster the training data with the following parameter settings: `min_cluster_size=40`, `min_samples=2`, `gen_min_span_tree=True`, and `prediction_data=True`. Generated peptides were then assigned to the existing clusters for visualization and down-selection for experimental testing using the `approximate_predict` method. The `probabilities_` attribute was used to report the probability of down-selected peptides belonging to their assigned clusters. HDBSCAN version 0.8.33 [53], UMAP-learn version 0.5.5 [52] were implemented in Python version 3.9.13.

Peptide alignment to known proteins. In order to characterize the relative novelty of the AI-generated peptides, the Basic Local Alignment Search Tool (BLAST) version 2.15.0 [54] was used to compare generated peptides to documented protein sequences found in nature. A maximum of 100 BLAST sequences were returned for each query, and reference hits greater than 250 amino acids or with less than 50% query cover were removed.

An additional metric was developed to quantitatively determine the “best” match for each generated sequence when queried against the standard NCBI Non-Redundant (nr) Protein Database used by BLAST. This metric, referred to here as the “Global Similarity Fraction”, is the scaled product of two BLAST alignment metrics: query cover and percent identify:

$$\text{Global Similarity Fraction} = \frac{\text{Query Cover Percent}}{100.0} * \frac{\text{Percent Identity}}{100.0}$$

where the query cover is defined as the percent of the generated sequence that was found in the returned BLAST sequence and the percent identity is defined as the percent of the returned BLAST sequence that identically matches the generated sequence. For each generated sequence, the BLAST result with the maximum Global Similarity Fraction was assigned as the “best”, or most similar, reference protein sequence.

Biophysical property calculations. Biophysical properties including net charge, isoelectric point, and the percent of hydrophobic residues were calculated for a subset of peptide sequences during down-selection for experimental validation. The percent of hydrophobic residues was calculated from peptide sequence, based on the number of amino acids containing aliphatic or aromatic hydrophobic side chains (A, I, L, M, V, F, W, and Y) [55]. Net charges and isoelectric points were calculated with the Biopython package version 1.83 [56]. Positively charged (*e.g.*, 0 < net charge < +6) peptides composed of 30–60% hydrophobic amino acids were further considered for experimental testing.

Prediction of peptide structures. The publicly available server-based tool for peptide structural prediction, PEPFOLD-4 [42] was used to predict the structures of candidate peptides of interest during initial down-selection for experimental testing. The AlphaFold3 server [57] was used to predict and visualize the structures of peptides following experimental validation.

Experimental validation of the Fung-AI pipeline

Fungal strains. *Saccharomyces cerevisiae* S288C (ATCC; 204508), *Candida albicans* 23Q (BEI Resources; NR-29341) and *Candida auris* CDC 381 (Casadevall Lab, CDC & FDA Antibiotic Resistance Isolate Bank) were cultured with Yeast Peptone Dextrose (YPD) media at 30°C. *Fusarium graminearum* PH-1 (ATCC; MYA-4620) was maintained on Potato Dextrose Agar (PDA) at 30°C.

Peptide synthesis preparation. The tested peptides in Table 4 were synthesized (ABI Scientific) with either crude purity and desalting for the initial assays, or with 95% purity for hit validation assays. A 10 mg/mL stock solution of each peptide was prepared in filter sterilized Milli-Q water. Stock solutions were subsequently diluted in YPD media to obtain 1 mg/mL working solutions.

Minimum inhibitory concentration (MIC) assays. Inoculums for *S. cerevisiae*, *C. albicans*, and *C. auris* were prepared by diluting an overnight culture 1:100 in YPD media and incubating until cultures reached an OD600 of 0.2. MIC assays for *F. graminearum* PH-1 were adapted from Al-Hatmi, et al [58]. Briefly, *F. graminearum* PH-1 spores were prepared by streaking *Fusarium* on a PDA plate and incubating at room temperature for five days. Spores were scraped into a solution of Phosphate Buffered Saline (PBS) with 0.005% Tween 20. The solution was allowed to settle for three minutes. The upper solution was carefully transferred to a new tube and vortexed for 15 seconds. The spore suspension was adjusted to an OD530 of 0.15 in YPD media.

For *S. cerevisiae*, *C. albicans*, and *F. graminearum*, 100 µL of the diluted fungal culture or spore solution was added to each well of a microtiter plate. A 100 µL aliquot of the peptide working solution (1 mg/mL) or amphotericin B positive control (32 µg/mL) was

Table 4. Selection of peptides tested for inhibitory activity and/or cytotoxicity. Testing phase denotes whether peptides were from original screen or deep dive into promising candidate clusters. Sc=Saccharomyces cerevisiae, Fg=Fusarium graminearum, Cal=Candida albicans, Cau=Candida auris, H=HepG2 cells.

Peptide ID	Cluster ID	Sequence	Testing Phase	Tested strains and cell lines
40	17	KILKRQRLLLGGLLKYLKAGK	Phase 1	Sc, Fg
50	13	KRRWPPKPVFSPNAAAYKCNP		Sc, Fg
27	3	GYGTTFIAPPGGLRGALPQFS		Sc, Fg
63	19	GLVLIKAARAGNLLKR		Sc, Fg
23	20	KAALKPKFTGAVTAFGLPQ		Sc, Fg
40	17	KILKRQRLLLGGLLKYLKAGK	Phase 2	Sc, Fg*, Cal*, Cau*, H
12	17	QLLRPLFTKLLGGLSKILKFF		Sc, Fg, Cal, Cau, H
17	17	RLLQALLNLRKCLKLGLLSRLS		Sc, Fg, Cal, Cau, H
48	17	KKTYTKLKKLNGAIKYLVGHA		Sc, Fg, Cal, Cau, H
65	17	KAKLRQAVLLKLLAGVKSSL		Sc, Fg, Cal, Cau, H
20	17	RLKKLWWKTCLLKGLCTLKA		Sc, Fg
29	17	GLRHYARVCLKLALLNLNKSC		Sc, Fg
44	17	KLAKLLALRLKSLVFGRLRYRT		Sc, Fg
54	17	KAIPLLRRRSIVLLLRGFGV		Sc, Fg

*In Phase II, activity against *F. graminearum*, *C. albicans*, and *C. auris* was measured with both crude and >95% purity formulations of peptide 40. No differences in the MICs were found between the two purity levels. All other testing for all other peptides was performed on crude purity peptides.

<https://doi.org/10.1371/journal.pcbi.1014105.t004>

added to the first row, with each condition tested in triplicate. For *C. auris*, 50 µL of diluted fungal culture and peptide working solution were added to the wells for a total assay volume of 100 µL. A two-fold serial dilution was performed down to the 7th row, leaving the last row as a no-treatment control, giving a tested range of 500-7.8 µg/mL for the peptides and 16-0.25 µg/mL for amphotericin B. *S. cerevisiae*, *C. albicans* and *C. auris* microtiter plates were incubated at 28°C for 48 hours and visually inspected for growth, while *F. graminearum* was incubated at 34°C. MIC is reported as the lowest concentration of antifungal that inhibited growth.

Cell cytotoxicity assays. To assess the toxicity of peptides in mammalian cell culture, a cytotoxicity assay was performed. Briefly, HepG2 cells (ATCC; HB-8065) were seeded in a 96-well plate at a density of 25,000 cells per well and allowed to grow overnight at 37 °C, 5% CO₂. Eagle's minimum Essential Medium (EMEM) with 10% fetal bovine serum (FBS) was used for cell growth. Peptides of interest were diluted serially from 1000 to 0.98 µg/mL with a two-fold dilution scheme. For testing, 100 µL of the diluted peptide solution was added to each well, with each concentration tested in triplicate. Each assay plate included the appropriate negative, positive and vehicle controls. The samples and cells were allowed to incubate for 24 hours at 37°C prior to a viability assessment with CellTiter Blue (Promega). CellTiter Blue was used in accordance with manufacturer's recommendations with a four-hour incubation time.

Percent viability was calculated by comparing the fluorescence value of the test well to that of the negative control. Each concentration was tested in triplicate (n=3 wells). Dose-response curves were fit on averaged viability data for each peptide using a pre-defined non-linear regression model in Prism (Model: [Inhibitor] vs. normalized response – Variable slope). The LD₅₀ (µg/mL) and model fit (R²) were calculated from each model.

Supporting information

S1 Fig. Example of the Fusarium graminearum MIC assay for a representative selection of active and inactive peptides. Images of the MIC assay for four peptides. Amphotericin B (AmB) was used as a positive control (blue) and water was used as a negative control. The bottom row of each plate was a no-treatment control. The MIC was reported as the lowest concentration of antifungal that inhibited growth.

(TIF)

S2 Fig. MIC assay of peptides against *Candida auris* strain CDC 381. *C. auris* is an emerging pathogen that is difficult to treat clinically. We tested peptides that were active against *F. graminearum* against *C. auris*. Amphotericin B (AmB) was used as a positive control. The bottom row of each plate was a no-treatment control. None of the tested peptides were found to inhibit *C. auris* strain CDC 381.

(TIF)

S1 Text. Document providing detailed descriptions of AI-based methods used in the Fung-AI pipeline.

(PDF)

Author contributions

Conceptualization: Daniel S Berman, Tom D Curtis, Arturo Casadevall, Laura J. Dunphy.

Data curation: Daniel S Berman, Libby M Lewis.

Formal analysis: Daniel S Berman, Libby M Lewis, Tom D Curtis, Olivia N Tiburzi, Daniel FQ Smith.

Funding acquisition: Daniel S Berman, Laura J. Dunphy.

Investigation: Daniel S Berman, Tom D Curtis, Olivia N Tiburzi, Daniel FQ Smith.

Methodology: Daniel S Berman, Libby M Lewis.

Project administration: Laura J. Dunphy.

Resources: Arturo Casadevall.

Software: Daniel S Berman, Libby M Lewis.

Supervision: Daniel S Berman, Arturo Casadevall, Laura J. Dunphy.

Validation: Daniel S Berman, Libby M Lewis.

Visualization: Daniel S Berman, Libby M Lewis, Tom D Curtis, Olivia N Tiburzi, Daniel FQ Smith, Laura J. Dunphy.

Writing – original draft: Daniel S Berman, Libby M Lewis, Tom D Curtis, Olivia N Tiburzi, Daniel FQ Smith, Laura J. Dunphy.

Writing – review & editing: Daniel S Berman, Libby M Lewis, Tom D Curtis, Olivia N Tiburzi, Daniel FQ Smith, Arturo Casadevall, Laura J. Dunphy.

References

1. Rayens E, Norris KA. Prevalence and healthcare burden of fungal infections in the United States, 2018. *Open Forum Infect Dis.* 2022;9(1):ofab593. <https://doi.org/10.1093/ofid/ofab593>
2. Bongomin F, Gago S, Oladele RO, Denning DW. Global and multi-national prevalence of fungal diseases-estimate precision. *J Fungi (Basel).* 2017;3(4):57. <https://doi.org/10.3390/jof3040057> PMID: [29371573](https://pubmed.ncbi.nlm.nih.gov/29371573/)
3. Denning DW. Global incidence and mortality of severe fungal disease. *Lancet Infect Dis.* 2024;24(7):e428–38. [https://doi.org/10.1016/S1473-3099\(23\)00692-8](https://doi.org/10.1016/S1473-3099(23)00692-8) PMID: [38224705](https://pubmed.ncbi.nlm.nih.gov/38224705/)
4. Steinberg G, Gurr SJ. Fungi, fungicide discovery and global food security. *Fungal Genet Biol.* 2020;144:103476. <https://doi.org/10.1016/j.fgb.2020.103476> PMID: [33053432](https://pubmed.ncbi.nlm.nih.gov/33053432/)
5. Jacobs SE, Jacobs JL, Dennis EK, Taimur S, Rana M, Patel D, et al. *Candida auris* pan-drug-resistant to four classes of antifungal agents. *Antimicrob Agents Chemother.* 2022;66(7):e0005322. <https://doi.org/10.1128/aac.00053-22> PMID: [35770999](https://pubmed.ncbi.nlm.nih.gov/35770999/)
6. Casadevall A. Global warming could drive the emergence of new fungal pathogens. *Nat Microbiol.* 2023;8(12):2217–9. <https://doi.org/10.1038/s41564-023-01512-w> PMID: [38030896](https://pubmed.ncbi.nlm.nih.gov/38030896/)
7. Huang J, Hu P, Ye L, Shen Z, Chen X, Liu F, et al. Pan-drug resistance and hypervirulence in a human fungal pathogen are enabled by mutagenesis induced by mammalian body temperature. *Nat Microbiol.* 2024;9(7):1686–99. <https://doi.org/10.1038/s41564-024-01720-y> PMID: [38898217](https://pubmed.ncbi.nlm.nih.gov/38898217/)
8. Casadevall A, Kontoyiannis DP, Robert V. On the emergence of *Candida auris*: climate change, azoles, swamps, and birds. *mBio.* 2019;10(4):e01397-19. <https://doi.org/10.1128/mBio.01397-19> PMID: [31337723](https://pubmed.ncbi.nlm.nih.gov/31337723/)

9. Ostrosky-Zeichner L, Casadevall A, Galgiani JN, Odds FC, Rex JH. An insight into the antifungal pipeline: selected new molecules and beyond. *Nat Rev Drug Discov.* 2010;9(9):719–27. <https://doi.org/10.1038/nrd3074> PMID: [20725094](https://pubmed.ncbi.nlm.nih.gov/20725094/)
10. Buda De Cesare G, Cristy SA, Garsin DA, Lorenz MC. Antimicrobial peptides: a new frontier in antifungal therapy. *mBio.* 2020;11(6):10.1128/mbio.02123-20. <https://doi.org/10.1128/mbio.02123-20>
11. Fernández de Ullivarri M, Arbulu S, Garcia-Gutierrez E, Cotter PD. Antifungal Peptides as Therapeutic Agents. *Front Cell Infect Microbiol.* 2020;10:105. <https://doi.org/10.3389/fcimb.2020.00105> PMID: [32257965](https://pubmed.ncbi.nlm.nih.gov/32257965/)
12. Sharma A, Singh G, Bhatti JS, Gill SK, Arya SK. Antifungal peptides: therapeutic potential and challenges before their commercial success. *Int J Biol Macromol.* 2025;284:137957. <https://doi.org/10.1016/j.ijbiomac.2024.137957>
13. Roscetto E, Contursi P, Vollaro A, Fusco S, Notomista E, Catania MR. Antifungal and anti-biofilm activity of the first cryptic antimicrobial peptide from an archaeal protein against *Candida* spp. clinical isolates. *Sci Rep.* 2018;8:17570. <https://doi.org/10.1038/s41598-018-35530-0>
14. Vouldoukis I, Shai Y, Nicolas P, Mor A. Broad spectrum antibiotic activity of the skin-PYY. *FEBS Lett.* 1996;380(3):237–40. [https://doi.org/10.1016/0014-5793\(96\)00050-6](https://doi.org/10.1016/0014-5793(96)00050-6) PMID: [8601432](https://pubmed.ncbi.nlm.nih.gov/8601432/)
15. Debono M, Abbott BJ, Turner JR, Howard LC, Gordee RS, Hunt AS, et al. Synthesis and evaluation of LY121019, a member of a series of semi-synthetic analogues of the antifungal lipopeptide echinocandin B. *Ann N Y Acad Sci.* 1988;544:152–67. <https://doi.org/10.1111/j.1749-6632.1988.tb40398.x> PMID: [3063167](https://pubmed.ncbi.nlm.nih.gov/3063167/)
16. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S. Generative Adversarial Networks. *arXiv.* 2014. Accessed 2025 August 5. <http://arxiv.org/abs/1406.2661>
17. Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, et al. De novo design of protein structure and function with RFdiffusion. *Nature.* 2023;620(7976):1089–100. <https://doi.org/10.1038/s41586-023-06415-8> PMID: [37433327](https://pubmed.ncbi.nlm.nih.gov/37433327/)
18. Berman DS, Howser C, Mehoke T, Ertlund AW, Evans JD. MutaGAN: a sequence-to-sequence GAN framework to predict mutations of evolving protein populations. *Virus Evol.* 2023;9(1):vead022. <https://doi.org/10.1093/ve/vead022> PMID: [37066021](https://pubmed.ncbi.nlm.nih.gov/37066021/)
19. Repecka D, Jauniskis V, Karpus L, Rembeza E, Rokaitis I, Zrimec J. Expanding functional protein sequence spaces using generative adversarial networks. *Nat Mach Intell.* 2021;3(4):324–33. <https://doi.org/10.1038/s42256-021-00310-5>
20. Gupta A, Zou J. Feedback GAN (FBGAN) for DNA: a novel feedback-loop architecture for optimizing protein functions. *arXiv.* 2018. <https://doi.org/10.48550/arXiv.1804.01694>
21. Mardikoraem M, Wang Z, Pascual N, Woldring D. Generative models for protein sequence modeling: recent advances and future directions. *Brief Bioinform.* 2023;24(6):bbad358. <https://doi.org/10.1093/bib/bbad358> PMID: [37864295](https://pubmed.ncbi.nlm.nih.gov/37864295/)
22. Veltri D, Kamath U, Shehu A. Deep learning improves antimicrobial peptide recognition. *Bioinformatics.* 2018;34(16):2740–7. <https://doi.org/10.1093/bioinformatics/bty179> PMID: [29590297](https://pubmed.ncbi.nlm.nih.gov/29590297/)
23. Witten J, Witten Z. Deep learning regression model for antimicrobial peptide design. 2019. Accessed 2021 November 23.
24. Yan J, Bhadra P, Li A, Sethiya P, Qin L, Tai HK, et al. Deep-AmPEP30: improve short antimicrobial peptides prediction with deep learning. *Mol Ther Nucleic Acids.* 2020;20:882–94. <https://doi.org/10.1016/j.omtn.2020.05.006> PMID: [32464552](https://pubmed.ncbi.nlm.nih.gov/32464552/)
25. Szymczak P, Możejko M, Grzegorzek T, Jurczak R, Bauer M, Neubauer D, et al. Discovering highly potent antimicrobial peptides with deep generative model HydrAMP. *Nat Commun.* 2023;14(1):1453. <https://doi.org/10.1038/s41467-023-36994-z> PMID: [36922490](https://pubmed.ncbi.nlm.nih.gov/36922490/)
26. Ruiz Puentes P, Henao MC, Cifuentes J, Muñoz-Camargo C, Reyes LH, Cruz JC. Rational discovery of antimicrobial peptides by means of artificial intelligence. *Membranes.* 2022;12(7):708. <https://doi.org/10.3390/membranes12070708> PMID: [35877911](https://pubmed.ncbi.nlm.nih.gov/35877911/)
27. Li C, Sutherland D, Hammond SA, Yang C, Taho F, Bergman L, et al. AMPlify: attentive deep learning model for discovery of novel antimicrobial peptides effective against WHO priority pathogens. *BMC Genomics.* 2022;23(1):77. <https://doi.org/10.1186/s12864-022-08310-4> PMID: [35078402](https://pubmed.ncbi.nlm.nih.gov/35078402/)
28. Cao Q, Ge C, Wang X, Harvey PJ, Zhang Z, Ma Y, et al. Designing antimicrobial peptides using deep learning and molecular dynamic simulations. *Brief Bioinform.* 2023;24(2):bbad058. <https://doi.org/10.1093/bib/bbad058> PMID: [36857616](https://pubmed.ncbi.nlm.nih.gov/36857616/)
29. Wang C, Garlick S, Zloh M. Deep learning for novel antimicrobial peptide design. *Biomolecules.* 2021;11(3):471. <https://doi.org/10.3390/biom11030471> PMID: [33810011](https://pubmed.ncbi.nlm.nih.gov/33810011/)
30. Yan J, Cai J, Zhang B, Wang Y, Wong DF, Siu SWI. Recent progress in the discovery and design of antimicrobial peptides using traditional machine learning and deep learning. *Antibiotics (Basel).* 2022;11(10):1451. <https://doi.org/10.3390/antibiotics11101451> PMID: [36290108](https://pubmed.ncbi.nlm.nih.gov/36290108/)
31. Yu H, Wang R, Qiao J, Wei L. Multi-CGAN: deep generative model-based multiproperty antimicrobial peptide design. *J Chem Inf Model.* 2024;64(1):316–26. <https://doi.org/10.1021/acs.jcim.3c01881>
32. Dean SN, Walper SA. Variational autoencoder for generation of antimicrobial peptides. *ACS Omega.* 2020;5(33):20746–54. <https://doi.org/10.1021/acsomega.0c00442> PMID: [32875208](https://pubmed.ncbi.nlm.nih.gov/32875208/)
33. Mousavizadegan M, Mohabatkar H. An evaluation on different machine learning algorithms for classification and prediction of antifungal peptides. *Med Chem.* 2016;12(8):795–800. <https://doi.org/10.2174/1573406412666160229150823> PMID: [26924627](https://pubmed.ncbi.nlm.nih.gov/26924627/)
34. Agrawal P, Bhalla S, Chaudhary K, Kumar R, Sharma M, Raghava GPS. In silico approach for prediction of antifungal peptides. *Frontiers in Microbiology.* 2018;9. <https://doi.org/10.3389/fmicb.2018.00323>
35. Ahmad A, Akbar S, Khan S, Hayat M, Ali F, Ahmed A, et al. Deep-AntiFP: prediction of antifungal peptides using distant multi-informative features incorporating with deep neural networks. *Chemom Intell Lab Syst.* 2021;208:104214. <https://doi.org/10.1016/j.chemolab.2020.104214>

36. Singh V, Shrivastava S, Kumar Singh S, Kumar A, Saxena S. Accelerating the discovery of antifungal peptides using deep temporal convolutional networks. *Brief Bioinform.* 2022;23(2):bbac008. <https://doi.org/10.1093/bib/bbac008> PMID: 35152278
37. Sharma R, Shrivastava S, Kumar Singh S, Kumar A, Saxena S, Kumar Singh R. Deep-AFPred: identifying novel antifungal peptides using pretrained embeddings from seq2vec with 1DCNN-BiLSTM. *Brief Bioinform.* 2022;23(1):bbab422. <https://doi.org/10.1093/bib/bbab422> PMID: 34670278
38. Yao L, Zhang Y, Li W, Chung C-R, Guan J, Zhang W, et al. DeepAFP: An effective computational framework for identifying antifungal peptides based on deep learning. *Protein Sci.* 2023;32(10):e4758. <https://doi.org/10.1002/pro.4758> PMID: 37595093
39. Fang C, Moriwaki Y, Li C, Shimizu K. Prediction of antifungal peptides by deep learning with character embedding. *IPSI Transactions on Bioinformatics.* 2019;12(0):21–9. <https://doi.org/10.2197/ipsitbio.12.21>
40. Li Y, Qiao Y, Ma Y, Xue P, Ding C. AI in fungal drug development: opportunities, challenges, and future outlook. *Front Cell Infect Microbiol.* 2025;15:1610743. <https://doi.org/10.3389/fcimb.2025.1610743> PMID: 40470259
41. Yaseen A, Gull S, Akhtar N, Amin I, Minhas F. HemoNet: Predicting hemolytic activity of peptides with integrated feature learning. *J Bioinform Comput Biol.* 2021;19(5):2150021. <https://doi.org/10.1142/S0219720021500219> PMID: 34353244
42. Rey J, Murail S, de Vries S, Derreumaux P, Tuffery P. PEP-FOLD4: a pH-dependent force field for peptide structure prediction in aqueous solution. *Nucleic Acids Res.* 2023;51(W1):W432–7. <https://doi.org/10.1093/nar/gkad376> PMID: 37166962
43. Ma T, Liu Y, Yu B, Sun X, Yao H, Hao C. DRAMP 4.0: an open-access data repository dedicated to the clinical translation of antimicrobial peptides. *Nucleic Acids Research.* 2025;53(D1):D403–10. <https://doi.org/10.1093/nar/gkae1046>
44. Wang Y, Song M, Liu F, Liang Z, Hong R, Dong Y. Artificial intelligence using a latent diffusion model enables the generation of diverse and potent antimicrobial peptides. *Sci Adv.* 2025;11(6):eadp7171. <https://doi.org/10.1126/sciadv.adp7171>
45. Oestreich M, Merdivan E, Lee M, Schultze JL, Piraud M, Becker M. DrugDiff: small molecule diffusion model with flexible guidance towards molecular properties. *J Cheminform.* 2025;17(1):23. <https://doi.org/10.1186/s13321-025-00965-x> PMID: 40001177
46. Krishnan A, Anahtar MN, Valeri JA, Jin W, Donghia NM, Sieben L, et al. A generative deep learning approach to de novo antibiotic design. *Cell.* 2025;188(21):5962–5979.e22. <https://doi.org/10.1016/j.cell.2025.07.033> PMID: 40816267
47. Satoh K, Makimura K, Hasumi Y, Nishiyama Y, Uchida K, Yamaguchi H. *Candida auris* sp. nov., a novel ascomycetous yeast isolated from the external ear canal of an inpatient in a Japanese hospital. *Microbiol Immunol.* 2009;53(1):41–4. <https://doi.org/10.1111/j.1348-0421.2008.00083.x> PMID: 19161556
48. Pirtskhalava M, Amstrong AA, Grigolava M, Chubinidze M, Alimbarashvili E, Vishnepolsky B, et al. DBAASP v3: database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics. *Nucleic Acids Res.* 2021;49(D1):D288–97. <https://doi.org/10.1093/nar/gkaa991> PMID: 33151284
49. Gawde U, Chakraborty S, Waghv FH, Barai RS, Khanderkar A, Indraguru R, et al. CAMPR4: a database of natural and synthetic antimicrobial peptides. *Nucleic Acids Res.* 2023;51(D1):D377–83. <https://doi.org/10.1093/nar/gkac933> PMID: 36370097
50. Bai S, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv.* 2018. <https://doi.org/10.48550/arXiv.1803.01271>
51. ElAbd H, Bromberg Y, Hoarfrost A, Lenz T, Franke A, Wendorff M. Amino acid encoding for deep learning applications. *BMC Bioinformatics.* 2020;21(1):235. <https://doi.org/10.1186/s12859-020-03546-x> PMID: 32517697
52. McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. *arXiv.* 2020. <https://doi.org/10.48550/arXiv.1802.03426>
53. McInnes L, Healy J, Astels S. hdbSCAN: hierarchical density based clustering. *JOSS.* 2017;2(11):205. <https://doi.org/10.21105/joss.00205>
54. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421> PMID: 20003500
55. Amino Acids Reference Chart. Accessed 2023 October 1. <https://www.sigmaaldrich.com/US/en/technical-documents/technical-article/protein-biology/protein-structural-analysis/amino-acid-reference-chart>
56. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics.* 2009;25(11):1422–3. <https://doi.org/10.1093/bioinformatics/btp163>
57. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature.* 2024;630(8016):493–500. <https://doi.org/10.1038/s41586-024-07487-w>
58. Al-Hatmi AMS, Curfs-Breuker I, De Hoog GS, Meis JF, Verweij PE. Antifungal susceptibility testing of fusarium: a practical approach. *J Fungi.* 2017;3(2):2. <https://doi.org/10.3390/jof3020019>