

Classifier-driven generative adversarial networks for enhanced antimicrobial peptide design

Michaela Areti Zervou ^{1,2,*}, Effrosyni Doutsis ², Yannis Pantazis ³, Panagiotis Tsakalides ^{1,2}

¹Computer Science Department, University of Crete, University Campus, Voutes, 715 00, Heraklion, Greece

²Institute of Computer Science, Foundation for Research and Technology-Hellas, Nikolaou Plastira 100, 700 13, Heraklion, Greece

³Institute of Applied and Computational Mathematics, Foundation for Research and Technology-Hellas, Nikolaou Plastira 100, 700 13, Heraklion, Greece

*Corresponding author. Computer Science Department, University of Crete, University Campus, Voutes, 715 00, Heraklion, Greece. E-mail: zervou@ics.forth.gr

Abstract

The development of antimicrobial peptides (AMPs) presents a promising approach to addressing antibiotic-resistant pathogens. Computational methods, such as Feedback Generative Adversarial Networks (FBGANs), have demonstrated strong performance in optimizing AMP design. FBGAN operates as a classifier-guided Generative Adversarial Network (GAN), refining training data by replacing them with the classifier's most accurate predictions based on a predefined threshold. However, this method may introduce bias and constrain the diversity and quality of the generated peptides. To address these limitations, we propose a novel classifier-driven GAN (cdGAN) framework that seamlessly integrates classifier predictions into the generative model's loss function. This enables an adaptive, end-to-end learning process that enhances AMP generation without requiring explicit data modifications. By embedding classifier guidance within the loss computation, cdGAN dynamically optimizes both peptide diversity and functionality. Comparative studies indicate that cdGAN outperforms conventional guided-GAN architectures, such as Conditional GANs and Auxiliary Classifier GANs, while achieving performance comparable to or exceeding established AMP design methods. Additionally, cdGAN's flexible architecture allows for the simultaneous optimization of multiple peptide attributes. To demonstrate this capability, we introduce a multi-task classifier based on the Evolutionary Scale Modeling 2 (ESM2) model, enabling cdGAN to assess both antimicrobial activity and peptide structural properties in parallel. This enhancement improves the likelihood of generating viable therapeutic candidates with enhanced antimicrobial effectiveness and reduced toxicity.

Keywords: protein de novo design; antimicrobial peptides; generative adversarial networks; large protein language models; transfer learning; multi-task learning

Introduction

Antimicrobial peptides (AMPs) are a promising alternative to traditional antibiotics due to their broad-spectrum activity and ability to target bacterial membranes, which reduces the likelihood of resistance development [1]. One of the key factors in AMP functionality is the presence of α -helical structures, which enhance their amphiphilic nature that allows interaction with both hydrophobic and hydrophilic regions of bacterial membranes, leading to effective membrane disruption [2, 3]. However, therapeutic peptides must also exhibit minimal toxicity to human cells, particularly avoiding hemolysis, as excessive hemolytic activity can cause red blood cell destruction and severe side effects [4]. Many potent AMPs fail in clinical applications due to unintended cytotoxicity, limiting their therapeutic viability. These considerations highlight the importance of designing AMPs that balance antimicrobial potency, structural optimization, and safety [4].

To accelerate AMP discovery, deep learning approaches have been widely adopted, with Generative Adversarial Networks (GANs) emerging as a particularly promising tool [5–8]. GAN-based frameworks, such as Feedback GAN (FBGAN) [5], have

demonstrated their ability to generate peptide sequences that closely mimic natural AMPs. The FBGAN model employs an additional classifier to guide the generator toward producing more AMP-like sequences based on a predefined threshold score. More recently, FBGAN-ESM2 [8] incorporated a pre-trained Evolutionary Scale Modeling 2 (ESM2) classifier [9] to enhance peptide selection, improving the antimicrobial efficacy of generated peptides.

Although effective, the general FBGAN framework relies on selecting sequences based on predefined AMP score thresholds, which can introduce bias and limit peptide diversity. By prioritizing high-scoring sequences, this model risks overfitting to known patterns, thereby reducing its ability to identify novel AMPs with unique properties. Furthermore, because classifier predictions are not always accurate, misclassified sequences may be included in training, resulting in incorrect or misleading data. These challenges emphasize the need for a more adaptive and unbiased generative approach.

To address these issues, this work presents a classifier-driven GAN (cdGAN) framework that incorporates classifier predictions

Received: March 1, 2025. Revised: July 25, 2025. Accepted: August 15, 2025

© The Author(s) 2025. Published by Oxford University Press.

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 reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

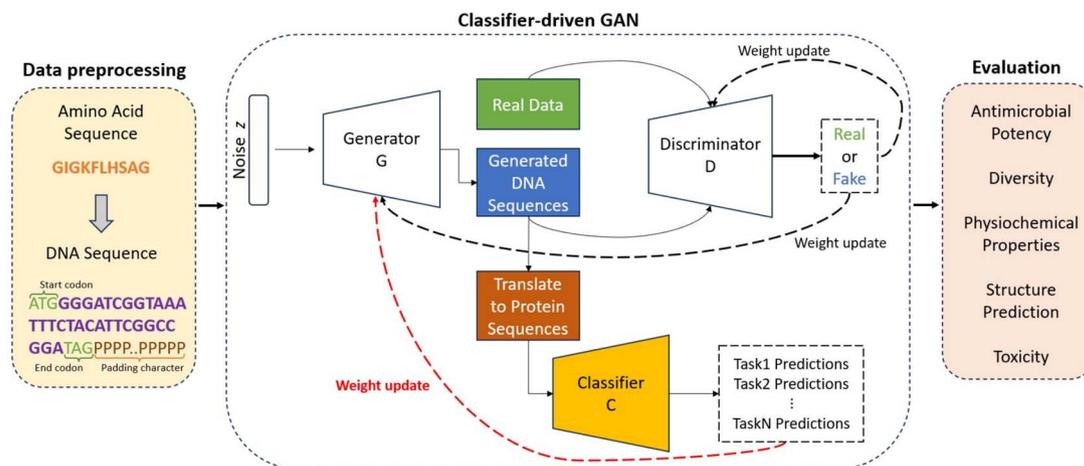


Figure 1. Architecture of the classifier-driven GAN (cdGAN) framework for AMP generation. The cdGAN framework integrates a generator (G), discriminator (D), and a multi-task classifier (C) to generate DNA-encoded antimicrobial peptide (AMP) sequences with biologically meaningful properties. In the preprocessing step, amino acid sequences are translated into DNA sequences using fixed-length encoding, including a start codon (ATG), a stop codon (TAG, TAA, or TGA), and a padding character P, which is constrained to appear only at the end of the sequence. The generator takes a noise vector z as input and produces synthetic DNA sequences. These sequences are evaluated by a discriminator, which attempts to distinguish between real and generated DNA sequences, providing feedback to improve the generator. Simultaneously, the DNA sequences are translated into protein sequences and passed through a classifier that predicts sequence-level properties relevant to antimicrobial peptide (AMP) functionality. The generator is trained using feedback from both the discriminator and the classifier, optimizing for both realistic sequence generation and the desired biological properties. Finally, evaluation metrics are applied to assess the quality of the generated sequences.

directly into the GAN loss function as provided in Fig. 1. Unlike traditional GANs, which rely on explicit selection of generated peptides, cdGAN allows for adaptive and continuous learning without data manipulation. This end-to-end approach reduces selection bias and improves model generalization, ultimately enhancing the diversity and novelty of generated AMPs. Moreover, by dynamically incorporating classifier feedback into the training process, cdGAN mitigates the risk of overfitting to predefined high-scoring peptides.

Beyond single-task antimicrobial design, cdGAN is extended to a multi-task framework that simultaneously optimizes antimicrobial activity and structural properties. Given the crucial role of α -helical structures in AMP function, we develop a novel multi-task classifier based on ESM2 that evaluates both antimicrobial potential and structural characteristics. This multi-task learning strategy enhances peptide design by ensuring that generated AMPs possess not only potent antimicrobial properties but also optimal structural features for membrane interaction [2, 3].

Finally, to ensure the safety of generated peptides, we implement a rigorous post-generation screening step using ToxinPred 3.0 [10] and HemoPI [11] to filter out toxic and hemolytic candidates. This additional screening step significantly enhances the therapeutic viability of generated AMPs by prioritizing sequences that are both effective and safe for clinical applications.

Materials and methods

Datasets and preprocessing

This study compares our proposed model with various guided-GAN variants and state-of-the-art approaches for AMP design. To train both the proposed cdGAN and baseline guided-GAN models, we constructed a dataset of 5,200 protein sequences ranging from 10 to 50 residues, including 2600 AMPs from APD3 [12] and 2600 non-AMPs from reviewed UniProt entries [13]. Non-AMPs were clustered using MMseqs2 [14] with a 50 percent sequence identity threshold, and one representative per cluster was selected to ensure diversity. Pairwise Needleman-Wunsch alignments [15]

showed an average sequence similarity of 0.2 between the AMP and non-AMP sets, limiting the usefulness of sequence alignment and encouraging the model to capture additional features. The diversity and physicochemical analyses (Tables 3, 5, and Fig. 6) support that the model learns biologically meaningful patterns beyond sequence similarity.

To enable multi-task classification, structural annotations were added using Porter 5 [16], which labels each sequence by secondary structure elements: helix (H), strand (E), and coil (C), while preserving the sequence length. Sequences with more than 10 consecutive helix (H) elements are classified as likely α -helical structures [5]. This structural information is combined with AMP annotations, forming a multi-task classification dataset capturing both antimicrobial activity and potential structural features of the peptides.

To standardize sequence encoding and introduce biological diversity, each amino acid sequence was converted to a complementary DNA (cDNA) sequence by assigning codons based on known genetic coding. For amino acids with multiple codons, one codon was randomly selected to introduce natural variability, following the approach in [5]. The process of translating the amino acid sequence into DNA through random codon assignments is reversible and does not affect the final protein's function [17, 18]. Finally, shorter sequences were padded with the character 'P' at the end to a uniform length of 156 nucleotides, ensuring input compatibility during training. An example is provided in Fig. 1.

Guided generative adversarial networks

In this work, we refer to the FBGAN-ESM2 model, as well as Conditional GANs (cGANs) and Auxiliary Classifier GANs (ACGANs), as guided-GANs, since they incorporate task-specific labels to steer generation toward desired outcomes.

For the implementation, we adopt the Wasserstein GAN with Gradient Penalty (WGAN-GP) framework [19], known for its stability during training and its ability to produce high-quality samples. The core of the WGAN-GP approach involves minimizing the Wasserstein distance [20] between the real data distribution P_r

and the distribution of generated samples P_g that is expressed as follows:

$$W(P_r, P_g) = \sup_{\|D\|_L \leq 1} \mathbb{E}_{x \sim P_r} [D(x)] - \mathbb{E}_{x \sim P_g} [D(x)],$$

where $\|D\|_L$ denotes the Lipschitz constraint on the discriminator D . Here, $\mathbb{E}_{x \sim P_r} [D(x)]$ and $\mathbb{E}_{x \sim P_g} [D(x)]$ represent the expected discriminator outputs on real and generated data, respectively.

To ensure the stability of the training process and uphold the Lipschitz constraint, a gradient penalty term is introduced, improving the quality of generated samples and promoting smoother training dynamics [19].

Conditional generative adversarial networks

Conditional GANs (cGANs) [21] extend the GAN framework by conditioning the generation process on auxiliary information, such as class labels or desired attributes. In cGANs, both the generator and the discriminator receive this conditional input, which guides the generation of samples in a targeted manner.

The objective function for the discriminator in a Wasserstein cGAN with gradient penalty can be written as:

$$L_D = \mathbb{E}_{x \sim P_r} [D(x|y)] - \mathbb{E}_{x \sim P_g} [D(x|y)] \\ + \lambda \mathbb{E}_{x \sim P_{\hat{x}}} [(\|\nabla_x D(x|y)\|_2 - 1)^2],$$

where $P_{\hat{x}}$ is the distribution of the interpolated samples between the real and generated data, and λ is the gradient penalty coefficient. The generator's objective function is:

$$L_G = -\mathbb{E}_{z \sim P_z} [D(G(z|y))].$$

Here, y represents the conditioning variable, and the generator aims to produce outputs consistent with this input.

Auxiliary classifier generative adversarial networks

Auxiliary Classifier GANs (ACGANs) [22] further enhance the conditional generation process by incorporating an auxiliary classifier within the discriminator architecture. This classifier is tasked with predicting the class labels associated with the generated samples, providing an additional layer of feedback during training.

The loss function for the discriminator in an Wasserstein ACGAN with gradient penalty comprises three components: the Wasserstein loss, the gradient penalty, and the classification loss:

$$L_D = \mathbb{E}_{x \sim P_r} [D(x|y)] - \mathbb{E}_{x \sim P_g} [D(x|y)] \\ + \lambda \mathbb{E}_{x \sim P_{\hat{x}}} [(\|\nabla_x D(x|y)\|_2 - 1)^2] \\ + \mathbb{E}_{x \sim P_r} [\log(Q(x|y))] + \mathbb{E}_{z \sim P_z} [\log(Q(G(z|y)|y))],$$

where Q is the auxiliary classifier predicting the label y . The generator's objective function is:

$$L_G = -\mathbb{E}_{z \sim P_z} [D(G(z|y)|y)] - \mathbb{E}_{z \sim P_z} [\log(Q(G(z|y)|y))].$$

This dual-loss structure encourages the generator to create samples that closely resemble real data while also meeting specific label characteristics.

Classifier-driven generative adversarial networks

This study introduces a novel classifier-driven Generative Adversarial Network (cdGAN) framework, designed to generate protein sequences with multiple desired properties. The cdGAN architecture extends the WGAN-GP model by incorporating a multi-task classifier that provides targeted feedback to guide the generative process.

The cdGAN consists of three core components: a generator G , a discriminator D , and a classifier C . The generator synthesizes protein sequences from latent noise vectors, while the discriminator evaluates these sequences to distinguish between real and generated data. The classifier predicts the likelihood that the generated sequences exhibit specific target properties, and this feedback is integrated into the generator's training, allowing it to produce sequences that meet defined biological criteria. To optimize multiple tasks simultaneously, the classifier output probabilities are multiplied, with a product closer to 1 indicating greater simultaneous satisfaction of all target properties. This product operator serves as a soft analogue to the logical AND operator, where all properties must be activated (i.e. set to ON) to have an effect. A logarithmic transformation is subsequently applied to enhance numerical stability and ensure the values are appropriately scaled within the overall loss function. The objective function of the cdGAN is expressed as:

$$\min_G \max_D \left(\mathbb{E}_{x \sim P_r} [D(x)] - \mathbb{E}_{z \sim P_z} [D(G(z))] \right. \\ \left. + \lambda \mathbb{E}_{x \sim P_{\hat{x}}} [(\|\nabla_x D(x)\|_2 - 1)^2] \right. \\ \left. - \sum_{t=1}^N a_t \cdot \mathbb{E}_{z \sim P_z} [\log C_t(G(z))] \right)$$

where λ controls the strength of the gradient penalty, $P_{\hat{x}}$ denotes the interpolated distribution between real and generated distributions, and $C_t(G(z)) = P(C_t = 1|G(z))$ represents the probability that the generated sample $G(z)$ belongs to the target class for task t , as assessed by the classifier C . Each classification task $t = 1, \dots, N$ is assigned a weight a_t , adjusting the influence of each task's feedback on the generator's objective.

At the core of this framework is the multi-task classifier, which is developed using transfer learning to evaluate multiple characteristics of protein sequences. To achieve this, we chose the ESM2 model [9], which has demonstrated strong performance in antimicrobial peptide classification, with the ESM2-t12 variant selected specifically for its computational efficiency and effective balance of representational power [23]. The classifier is trained using a combined loss function to simultaneously optimize multiple tasks:

$$\mathcal{L}_{mit} = \sum_{t=1}^N \lambda_t \mathcal{L}_t,$$

where $t = 1, \dots, N$ is the number of tasks, \mathcal{L}_t is the loss for each task, and λ_t is a hyperparameter that determines the weight of each task's loss, and it must satisfy $\lambda_k \geq 0$. In addition, the sum of all task-specific weights λ_t is constrained to $\sum_t \lambda_t = 1$, ensuring a balanced contribution of each task to the total loss. By adjusting the values of λ_t , the model can prioritize certain tasks, allowing for

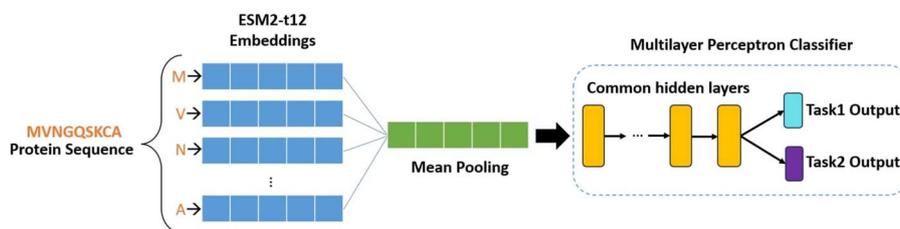


Figure 2. Overview of the proposed transfer learning-based architecture utilizing the ESM2 protein language model and a Multilayer Perceptron (MLP) classifier.

flexible optimization and achieving balanced performance across all tasks.

Results

In this section, we present the computational performance of each model. Technical details on model parameters, experimental settings, and hardware specifications are provided in the [supplementary material](#).

Classifier performance analysis

The classification of AMPs is important in drug design and the development of targeted therapies [5, 8, 24]. AMPs often adopt α -helical structures, which are particularly advantageous due to their ability to disrupt cellular protective layers. Consequently, detecting the structural fold of an AMP provides valuable insights into its functional mechanisms.

Numerous classifiers have been designed to identify AMP properties [25–28], while others focus on classifying structural folds [16, 29–31]. A recent study proposed a multi-task learning framework leveraging Transformer networks and k -mers technique to classify both AMP properties and α -helical structure [32] simultaneously. The k -mers method is applied on DNA sequences to provide meaningful representations, however, it requires careful optimization of k to achieve a balance between sequence detail and computational efficiency.

To eliminate the need for manual sequence preprocessing and to work directly at the protein level, we propose a novel multi-task classifier, illustrated in Fig. 2. Our model utilizes embeddings derived from the pre-trained ESM2 protein language model, which processes amino acid sequences directly, and feeds them into a lightweight Multilayer Perceptron (MLP) for efficient and scalable classification. The joint loss function for the two prediction tasks is defined as follows:

$$\mathcal{L}_{\text{mlt}} = \lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2 \quad \lambda_2 = 1 - \lambda_1, \quad \lambda_1, \lambda_2 \geq 0$$

Table 1 compares the single-task classification performance of our proposed MLP-ESM2 classifier against the k -mers Transformer model for both AMP (Task 1) and α -helical folding prediction (Task 2) independently. As demonstrated, the MLP-ESM2 classifier achieves significantly better performance compared to the k -mers-based model for both tasks as evaluated by three different metrics: the accuracy (ACC), the area under the curve (AUC), and the F1-score. The results indicate that the proposed MLP-ESM2 classifier provides a robust and computationally efficient approach to protein property prediction, achieving high accuracy, AUC, and F1 scores across both tasks. This suggests that leveraging transfer learning with ESM2 embeddings can significantly enhance classification performance in bioinformatics tasks while

Table 1. Single-task classification results for AMP (Task 1) and α -helical folding prediction (Task 2) using the proposed MLP-ESM2 classifier and a k -mers Transformer model. Standard deviations are reported in parentheses

	Task 1 ($\lambda_1 = 1$)		Task 2 ($\lambda_2 = 1$)	
	MLP-ESM2	k -mers Transformer	MLP-ESM2	k -mers Transformer
ACC	91.6 (1.0)	81.1 (1.4)	92.8 (0.8)	74.7 (1.2)
AUC	97.1 (0.3)	88.2 (0.6)	98.1 (0.3)	80.7 (1.3)
F1	91.7 (1.0)	80.9 (1.7)	92.8 (0.8)	80.2 (1.1)

Table 2. Performance comparison of the optimal k -mers Transformer ($\lambda_1 = 0.7$) and MLP-ESM2 ($\lambda_1 = 0.4$) multi-task classifiers in AMP (Task 1) and α -helical folding prediction (Task 2)

	k -mers Transformer		MLP-ESM2	
	Task 1	Task 2	Task 1	Task 2
ACC	79.9 (2.1)	75.9 (0.4)	91.3 (0.7)	92.6 (0.8)
AUC	88.3 (0.8)	82.0 (1.7)	96.9 (0.5)	97.9 (0.2)
F1	79.0 (3.5)	82.0 (0.5)	91.4 (0.8)	93.3 (0.8)

mitigating the complexity associated with training larger networks.

The performance of the proposed multi-task MLP-ESM2 classifier across various λ_1 and λ_2 values is presented in the [Supplementary Table 1](#). The optimal values were selected based on the Pareto front shown in [Supplementary Figure 1](#). This figure illustrates that variations in λ_1 have minimal impact on overall classification performance, enabling robust classification despite shifts in task prioritization. Table 2 highlights the performance of the multi-task classifiers, demonstrating that our proposed model significantly outperforms the k -mers Transformer on both tasks.

Comparison of cdGAN with other guided-GANs

In this section, we analyse the proposed cdGAN model against baseline guided-GAN architectures, including cGANs, ACGANs, and FBGAN-ESM2. We generated 5000 sequences per model using the final trained generator, with random noise as input and, for conditional models, the appropriate AMP class label. These generated sequences were then used for all subsequent comparisons against real AMP sequences from reference databases. In particular, we evaluated their diversity and predicted antimicrobial potency. High diversity is crucial to ensure the generative model is not simply reproducing known sequences. For antimicrobial potency, we employed predictive models to estimate the biological activity of the generated sequences. This evaluation ensures that

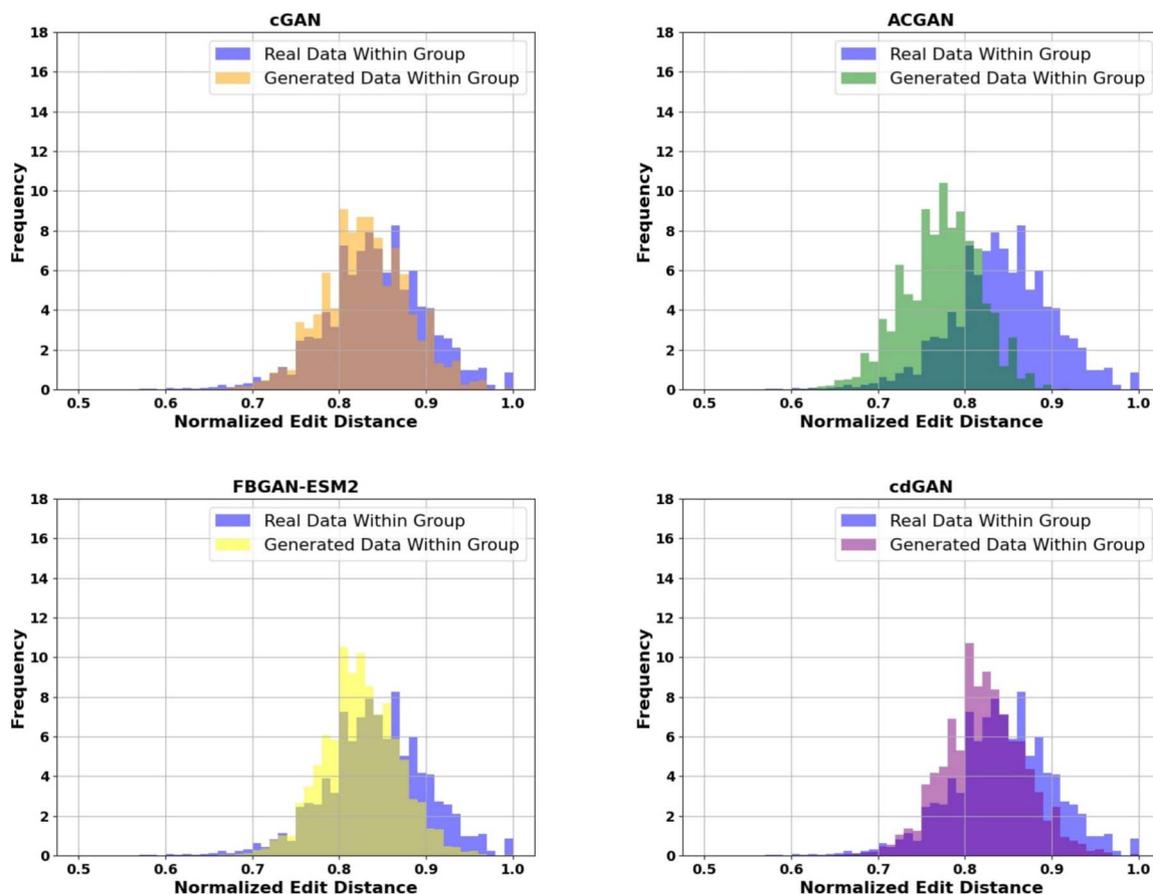


Figure 3. Normalized within group edit distance distribution for real and generated data, highlighting the fidelity and diversity of different models.

the generated peptides are not only diverse but also functionally relevant, aligning to discover new antimicrobial peptides.

The generated data were compared to real AMPs sourced from six widely recognized datasets: XUAMP [33], CAMP [34], dbAMP [35], DRAMP [36], LAMP [37], and YADAMP [38]. Duplicate sequences were removed, and only peptides with lengths between 10 to 50 amino acids were selected for analysis.

Diversity of generated peptides

The edit distance analysis, shown in Fig. 3, provides insights into the diversity of the generated sequences. The normalized edit distance measures the difference between real and generated sequences by counting the minimum number of insertions, deletions, or substitutions required to transform one sequence into another, offering a way to assess how closely models capture the distribution of real peptides. As provided, the distribution of the generated data for cdGAN, FBGAN-ESM2, and cGAN aligns closely with the distribution of real data. In contrast, ACGAN shows the weakest alignment, with its generated distributions deviating significantly from the real data, highlighting challenges in capturing diversity.

Table 3 shows both training data similarity and in-between sequence similarity for each model. To highlight the classifier's role in the GAN framework, we include a plain GAN, which follows the cdGAN architecture but is trained only on AMP sequences without classifier guidance. More details on the architectures can be found in the [supplementary material](#).

For in-between sequence similarity, ACGAN shows the highest redundancy (40.3%), meaning its generated peptides are highly

Table 3. Percentage of sequence similarity and diversity within the generated data for each model, and novelty relative to the training data

	In-between Sequence Similarity	Diversity	Novelty
Real data	22.9%	100%	100%
cGAN	28.2%	100%	100%
ACGAN	40.3%	100%	100%
FBGAN-ESM2	27.6%	100%	100%
cdGAN (This work)	27.4%	100%	100%

similar to each other. In contrast, cdGAN achieves the lowest in-between similarity (27.4%), followed closely by FBGAN-ESM2 (27.6%) and the plain GAN (27.6%), highlighting their ability to generate a more diverse set of peptides. Finally, all models exhibited 100% novelty, confirming that none of the generated sequences overlapped with the training set.

Antimicrobial property prediction

While the previous metrics provide valuable insights into the fidelity and diversity of generated sequences, they do not capture functional relevance. To address this, we evaluated the generated sequences using the CAMP_{R4} platform [34], which predicts AMP activity through external classifiers such as Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN). The classification accuracy for RF is 86.5%, 84.1% for SVM, and 82.2% for ANN. This step is essential for assessing how well

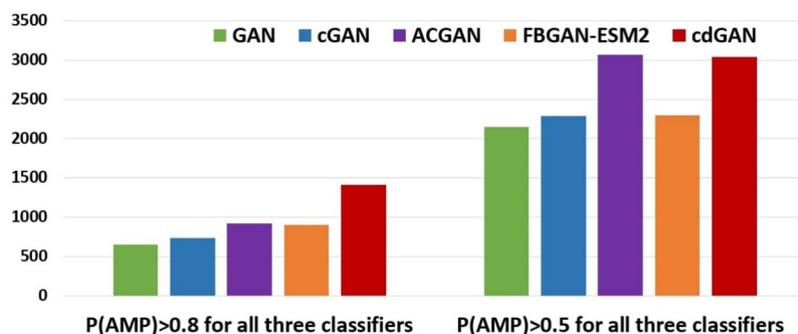


Figure 4. Fraction of generated peptides classified as positive by every classifier of $CAMP_{R4}$ (RF, SVM, and ANN) with probabilities $P(AMP) > 0.5$ (right) and $P(AMP) > 0.8$ (left), respectively.

Table 4. Percentage of predicted AMPs with $CAMP_{R4}$ platform equipped with Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) models for AMP prediction. The best performance per classifier is highlighted in bold, and the standard deviation is reported in parentheses

	CAMP _{R4} predictions			
	RF	SVM	ANN	Average (std)
GAN	58.0%	51.5%	52.8%	54.1% (2.8)
cGAN	59.2%	56.1%	57.2%	57.5% (1.2)
ACGAN	77.6%	72.2%	65.6%	71.8% (4.9)
FBGAN-ESM2	59.1%	55.5%	58.9%	57.8% (1.6)
cdGAN (This work)	75.8%	68.6%	70.5%	71.6% (3.0)

the generated sequences generalize to unseen data and whether they retain key functional characteristics. By leveraging external classifiers trained on diverse datasets, we minimize biases introduced by our specific training data, ensuring a more objective and robust evaluation of the generated peptides' applicability in real-world contexts.

The results summarized in Table 4, indicate that the RF classifier consistently identifies a higher proportion of sequences as AMPs across all models. Overall, ACGAN and cdGAN achieve the strongest performance; however, the high performance of ACGAN is likely attributed to the high sequence similarity of its generated peptides, which as shown in Table 3, is 40.3%. This increased similarity allows the generated sequences to align closely with patterns recognized by the classifiers, reflecting a trade-off between reduced diversity and enhanced classifier-predicted functionality. Moreover, the performance of ACGAN is not consistent across all classifiers, with the ANN-based classifier achieving approximately 10% lower accuracy since it relies on learned features, unlike RF and SVM, which benefit from direct sequence similarity. In contrast, cdGAN performs better with ANN classifiers, and given the flexibility of neural networks, there is significant potential for further improvements, which is not as easily achievable with RF and SVM. Overall, cdGAN demonstrates a more balanced performance across classifiers, maintaining high diversity while generating sequences with sufficient similarity to real AMPs. Lastly, GAN, cGAN, and FBGAN-ESM2 exhibit comparable performance, though they underperform compared to cdGAN and ACGAN.

Figure 4 illustrates the number of sequences predicted as AMPs by all three classifiers, categorized by the thresholds $P(AMP) > 0.5$ and $P(AMP) > 0.8$. At the lower threshold, ACGAN and cdGAN emerge as top performers, generating the largest number

Table 5. Percentage of diversity and sequence similarity per model

	In-between Sequence Similarity	Diversity
Real data	22.9%	100%
AMPGAN	26.8%	100%
HydrAMP	21.7%	98.9%
RLGen	36.9%	100%
cdGAN (This work)	27.4%	100%

of predicted AMPs, followed by FBGAN-ESM2, cGAN, and GAN. When considering the stricter threshold, cdGAN retains strong performance, while the rest of the models experience notable declines. Overall, these results demonstrate that cdGAN achieves the best trade-off between diversity, sequence similarity, and classifier-predicted functionality, making it an effective model for AMP design.

Comparison with state-of-the-art AMP design models

To further assess the potential of the proposed cdGAN model we compare it against state-of-the-art AMP design models, namely, AMPGAN [39], HydrAMP [6], and RLGen [7]. AMPGAN uses bidirectional conditional GANs [40], conditioning on diverse AMP attributes, such as pathogen specificity, mechanisms of action, and minimum inhibitory concentration (MIC), to tailor generated peptides. RLGen is a reinforcement learning-guided generative adversarial network (GAN) framework that leverages a thermodynamic diffusion process to better simulate AMP conformations, while HydrAMP leverages conditional variational auto-encoders [41] to generate peptides based on properties like antimicrobial activity and MIC.

The sequence similarity metric presented in Table 5 reveals that HydrAMP achieves the lowest similarity (21.7%), outperforming even the real data (22.9%) and indicating its superior ability to generate highly diverse sequences. AMPGAN (26.8%) and cdGAN (27.4%) exhibit moderate levels of similarity, whereas RLGen performs the worst (36.9%), indicating reduced diversity in its generated sequences. In terms of the percentage of unique sequences, all models besides HydrAMP achieve no duplicated outputs.

The $CAMP_{R4}$ predictive results provided in Table 6 show that cdGAN consistently outperforms all other models across all classifiers (RF, SVM, and ANN) highlighting the model's ability to generate sequences that are both diverse and biologically relevant. Figure 5 demonstrates the number of sequences predicted as

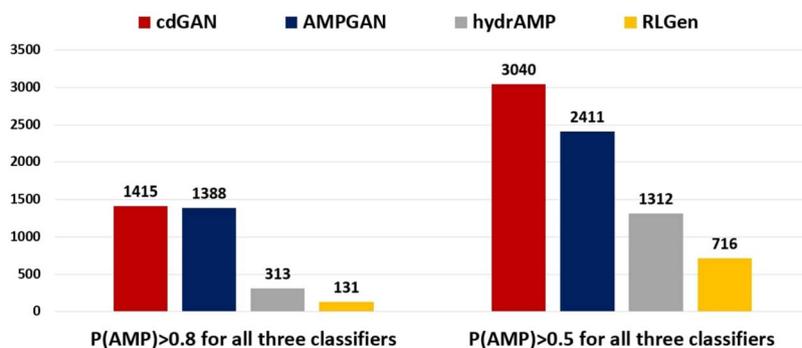


Figure 5. Fraction of generated peptides classified as positive by every classifier of $CAMP_{R4}$ (RF, SVM, and ANN) with probabilities $P(AMP) > 0.5$ (right) and $P(AMP) > 0.8$ (left), respectively.

Table 6. Percentage of predicted AMPs with $CAMP_{R4}$ platform equipped with Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) models for AMP prediction. The best performance per classifier is highlighted in bold, and the standard deviation is reported in parentheses

	CAMP _{R4} predictions			
	RF	SVM	ANN	Average (std)
AMPGAN	64.6%	53.8%	56.6%	58.3% (4.5)
HydrAMP	58.8%	52.6%	32.9%	48.1% (11.0)
RLGen	34.7%	21.1%	21.7%	25.8% (6.2)
cdGAN (This work)	75.8%	68.6%	70.5%	71.6% (3.0)

AMPs by all three classifiers, categorized by the thresholds $P(AMP) > 0.5$ and $P(AMP) > 0.8$. At both thresholds, cdGAN outperforms the other models by generating the highest number of predicted AMPs. However, under the stricter threshold, AMPGAN closely follows in performance.

This analysis highlights that cdGAN and AMPGAN are the leading models in generating predicted AMPs with high confidence. cdGAN demonstrates an advantage in generating sequences under broader confidence criteria, while AMPGAN remains competitive at stricter thresholds. HydrAMP and RLGen, in contrast, face challenges in producing sequences that consistently meet high-confidence criteria.

Physicochemical properties and amino acid composition

To evaluate the quality and functional relevance of the generated peptides, we analysed four physicochemical properties [42] of AMP sequences with $P(AMP) > 0.8$ across all classifiers, for each model. Particularly, we evaluate the distribution of charge, isoelectric point (pI), aromaticity, and hydrophobicity ratio. The charge of AMPs plays a crucial role in their interaction with microbial membranes, and the distribution of charge among the generated sequences provides important insights into how well the models replicate this fundamental characteristic. The isoelectric point (pI) distribution, at physiological pH, for the real data has a wide distribution of pI values, with a median around 10, reflecting the prevalence of cationic peptides among AMPs, a key feature for membrane-targeting functionality. Aromaticity is also a crucial feature in AMPs as it can influence their interaction with microbial membranes and overall efficacy. Finally, hydrophobicity is a key factor in shaping the functional properties of AMPs, as it affects their interaction with microbial membranes.

The violin plots in Fig 6 present the distribution of these four physicochemical properties for AMP sequences with a

predicted confidence of $P(AMP) > 0.8$. The properties of sequences generated by the cdGAN model are compared against real data, random data, AMPGAN, HydrAMP, and RLGen. Statistical significance was evaluated using the two-sided Mann-Whitney-Wilcoxon test [43], a non-parametric method suitable for comparing independent samples without assuming normality. No multiple testing corrections were applied, as the number of statistical comparisons was limited.

As presented, the distribution of physicochemical features reveals that cdGAN-generated sequences most closely resemble real AMPs. For charge and pI, cdGAN produces distributions that, while statistically different from real data ($p \leq 0.05$ and $p \leq 0.001$, respectively), are more balanced compared to other models.

In terms of aromaticity, cdGAN-generated peptides show a significant but moderate deviation from real data ($p \leq 0.01$), while AMPGAN's and RLGen's results are not significantly different (ns), suggesting better alignment in this specific feature.

For hydrophobicity, cdGAN again demonstrates moderate deviations ($p \leq 0.01$), whereas HydrAMP shows no significant difference from real data, and models like AMPGAN and RLGen exhibit more pronounced differences ($p \leq 0.0001$), indicating less biologically relevant outputs.

Overall, while AMPGAN and HydrAMP demonstrate partial success in replicating specific physicochemical properties, cdGAN offers more balanced results across all features.

Multi-task cdGAN performance

To demonstrate the effectiveness of the proposed cdGAN model in multi-task design, we present a case study focused on the design of AMPs with specific α -helical structure. The ability to generate AMPs with desired structural properties is crucial for enhancing their efficacy and stability, contributing to the development of new therapeutic agents.

Table 7 provides the percentage of predicted AMPs with $CAMP_{R4}$ platform, as well as the percentage of the helical structures predicted with the proposed classifier. The results show that both the single-task and multi-task cdGAN models yield relatively high AMP scores across all classifiers from the $CAMP_{R4}$ platform. In particular, the multi-task cdGAN model consistently produces AMP sequences with an average score of 71.1%, with a standard deviation of 1.7%.

Moreover, the multi-task cdGAN outperforms the single-task cdGAN, achieving 27.0% helical content compared to 21.1% for the single-task version. This suggests that multi-task learning, by optimizing for both AMP activity and helical structure, leads

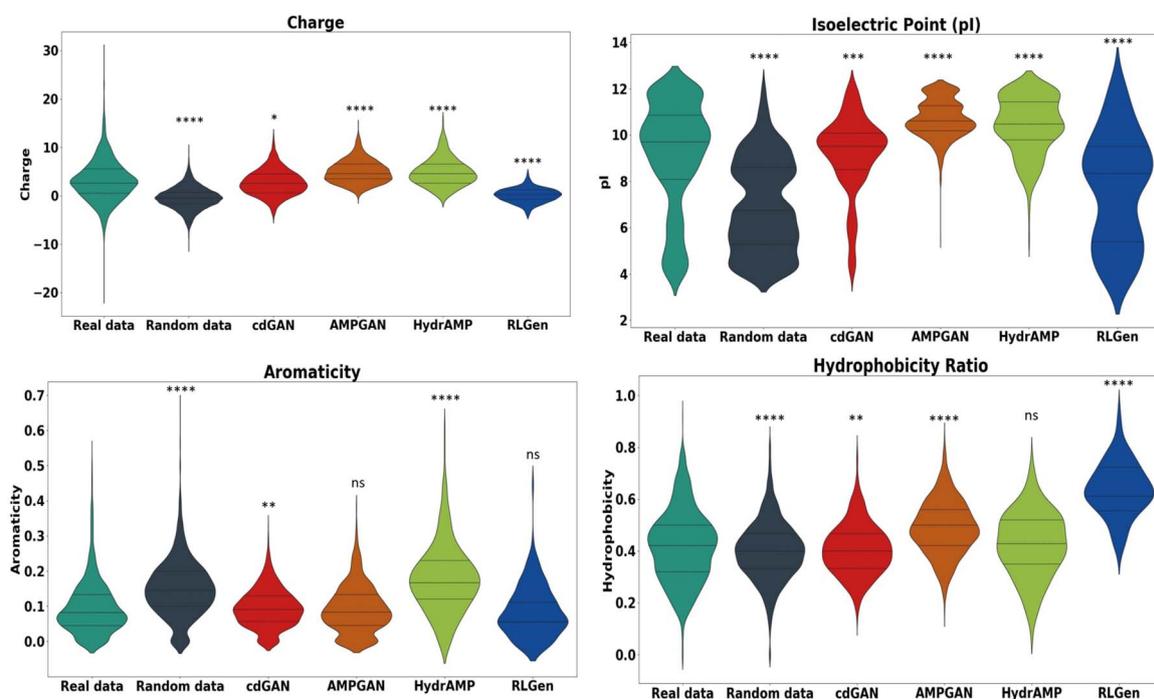


Figure 6. Distribution of physicochemical features such as charge, isoelectric point, aromaticity, and hydrophobicity, of real, random, and generated data per model. Statistical significance was evaluated using the two-sided Mann–Whitney–Wilcoxon test, with p-values denoted as: ns for non-significant values ($p > 0.05$), * ($p \leq 0.05$) indicating a modest level of significance, ** ($p \leq 0.01$) indicating stronger significance, *** ($p \leq 0.001$) indicating very strong significance, and **** ($p \leq 0.0001$) indicating the highest level of significance.

Table 7. Percentage of predicted AMPs and α -helical structure. The standard deviation is reported in parentheses

	CAMP _{R4} predictions				Helical structure prediction
	RF	SVM	ANN	Average	
cdGAN (single-task)	75.8%	68.6%	70.5%	71.6% (3.0)	21.1%
cdGAN (multi-task)	73.4%	69.3%	70.6%	71.1% (1.7)	27.0%

to better generation of helical peptides. These results support the notion that multi-task models can better capture complex relationships between multiple properties.

To evaluate the dual-objective capability of each generative model considering a more realistic peptide design scenario, we first selected sequences classified as AMPs with a probability greater than 95% across all three classifiers of CAMP_{R4}. These sequences were subsequently evaluated for α -helical content using our proprietary helix prediction classifier, retaining only those with a helix probability exceeding 95% as well. This stringent threshold ensures that only the most reliable and biologically meaningful candidates are selected for analysis, aligning with the ultimate goal of designing effective therapeutic agents.

The reason we rely on our helix prediction classifier, rather than other methods, is its higher accuracy compared to other related works that classify the structure of AMP sequences [32]. Other existing generic classifiers for secondary structures lack the precision required for AMP validation and achieve lower performance, in general, [16, 29–31]. By employing our classifier, we ensure that the validation of generated sequences aligns with the specialized requirements of AMP research, offering a robust and accurate assessment of their α -helical content. The liability of our classifier is also validated with AlphaFold 3 [44] server with the results depicted in Fig. 8.

Table 8 offers insights into each model's capability to generate sequences that are both AMP and possess a helical structure with high confidence ($P(\text{AMP} > 0.95)$ and $P(\text{helix} > 0.95)$). As shown, AMPGAN demonstrates the highest count of sequences meeting both objectives. This notable performance likely stems from the fact that AMPGAN incorporates more biological knowledge into its design, enabling it to generate sequences that are not only likely to be AMPs but also possess the necessary structural characteristics, such as α -helical content. In contrast, other models, including cGAN, ACGAN, FBGAN-ESM2, and HydrAMP, show a remarkably lower number of dual-objective sequences, with HydrAMP failing to generate any qualifying sequences. This underperformance could be attributed to differences in training objectives or architectural limitations that prevent these models from effectively capturing both AMP activity and α -helical structural features. RLGen performs slightly better, as its training on thermodynamic principles inherently considers the interdependence between amino acids, which indirectly incorporates structural information, but still falls short of expectations. The relatively limited performance of these models highlights the inherent challenges in designing generative frameworks capable of optimizing multiple complex properties simultaneously.

Considering the proposed cdGAN model, in its multi-task configuration, generates 21 sequences meeting the dual objectives, doubling the count achieved by its single-task counterpart that produces 10 sequences.

The toxicity of these top-scored sequences is also evaluated with ToxinPred 3.0 [10] to assess their potential for therapeutic applications. Ideally, AMPs that are effective at fighting infections should also exhibit low toxicity to human cells, making them more suitable for development into safe and effective treatments [45]. The results, presented in Fig. 7, demonstrate that 95.2% of the sequences generated by the cdGAN model are non-toxic, compared to 80.7% for AMPGAN, indicating a significant

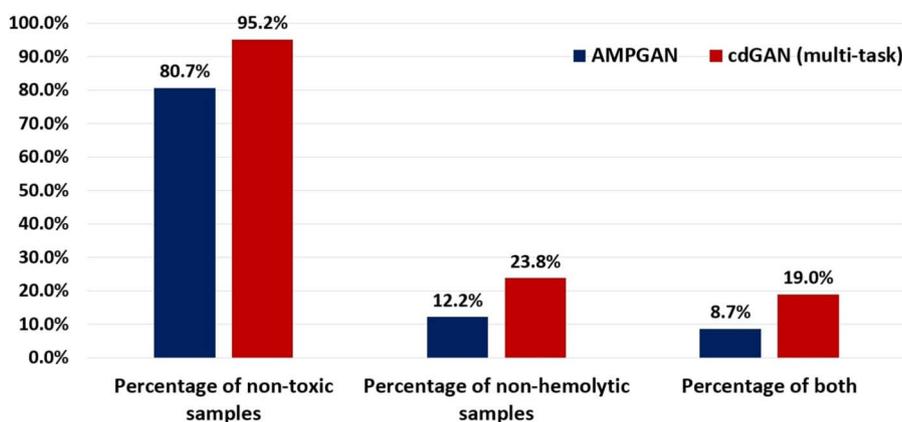


Figure 7. Performance of AMPGAN and cdGAN (multi-task) models based on the percentage of generated AMPs with $P(\text{AMP}) > 0.95$ and $P(\text{helix}) > 0.95$, which are non-toxic, non-hemolytic, and satisfy both the previous conditions.

Table 8. Number of sequences generated by each model that meet dual objectives of high AMP probability ($P(\text{AMP}) > 0.95$) and high α -helical content probability ($P(\text{helix}) > 0.95$)

	Number of sequences with $P(\text{AMP}) > 0.95$ and $P(\text{helix}) > 0.95$
GAN	1
cGAN	1
ACGAN	1
FBGAN-ESM2	2
AMPGAN	57
HydrAMP	0
RLGen	6
cdGAN (single-task)	10
cdGAN (multi-task)	21

improvement in safety. Similarly, the hemolytic potential of these sequences is assessed using the HemoPI tool [11]. Hemolytic AMPs are less desirable for therapeutic use as they can damage red blood cells, posing a risk of hemolysis. The results show that 23.8% of the cdGAN-generated sequences are non-hemolytic, nearly twice as many as the 12.2% achieved by AMPGAN. Furthermore, 19.0% of the cdGAN sequences meet both criteria of being non-toxic and non-hemolytic, a marked improvement over the 8.7% achieved by AMPGAN. These results highlight the ability of cdGAN's multi-task learning framework in generating sequences with lower predicted toxicity and hemolysis, making them better suited for therapeutic applications.

Figure 8 presents AlphaFold3 [44] structural predictions for cdGAN-generated AMPs with high predicted antimicrobial activity ($P(\text{AMP}) > 0.95$) and strong α -helical likelihood ($P(\text{helix}) > 0.95$). The per-residue local distance difference test (pLDDT) scores are exceptionally high across all cases (ranging from 80 to 95), demonstrating a high degree of confidence in the overall structural quality. The predicted Template Modeling (pTM) scores for the peptides range from 0.27 to 0.39. Although these are below the 0.5 threshold typically associated with high-confidence global folds, this is expected in de novo design, where the sequences are novel and not derived from known structural templates [46]. These moderate scores suggest the peptides may adopt coherent, protein-like structures despite lacking homology to known proteins. However, the low pTM scores also reflect limited confidence

in the predicted global folds, and these structural predictions should therefore be interpreted with caution.

Discussion

The findings of this study highlight the multifaceted nature of AMP generation, emphasizing the balance between productivity, functional relevance, and physicochemical fidelity. Among the evaluated models, cdGAN emerges as a robust performer, excelling across multiple dimensions and achieving a balance that other models struggle to replicate.

The analysis of physicochemical properties further emphasizes cdGAN's strengths. Its ability to replicate the charge and pI distributions of real AMPs, while maintaining hydrophobicity levels consistent with natural variability, reflects a nuanced understanding of AMP characteristics. This is statistically supported by the results of the Mann–Whitney–Wilcoxon test, where cdGAN-generated peptides demonstrated significant similarity to real data in terms of charge ($p \leq 0.05$) and hydrophobicity ($p \leq 0.01$). In contrast, AMPGAN and HydrAMP generate peptides with elevated charge profiles, suggesting a potential bias toward membrane-targeting functionality, which, while useful for specific applications, may limit the overall diversity. Similarly, RLGen showed statistically significant deviations ($p \leq 0.0001$) from real AMPs, exhibiting overly broad, misaligned property ranges. These findings underscore the importance of balancing functional optimization with fidelity to real AMP profiles, a challenge that cdGAN addresses effectively.

The multi-task learning framework employed by cdGAN demonstrates a clear advantage in optimizing for multiple objectives simultaneously. In its multi-task configuration, cdGAN produces a higher proportion of sequences exhibiting both AMP activity and α -helical structure, outperforming its single-task counterpart doubling the count of sequences meeting these dual objectives.

The toxicity and hemolytic analysis further underscores the therapeutic potential of multi-task cdGAN-generated peptides. Among high-confidence AMP sequences ($P(\text{AMP}) > 0.8$), 95.2% of the sequences generated by cdGAN are classified as non-toxic, compared to 80.7% for the competitive AMPGAN model. This represents a significant improvement in toxicity reduction, highlighting cdGAN's superior ability to generate therapeutically viable peptides. Similarly, in the hemolysis analysis conducted with the HemoPI tool, nearly double of the cdGAN sequences

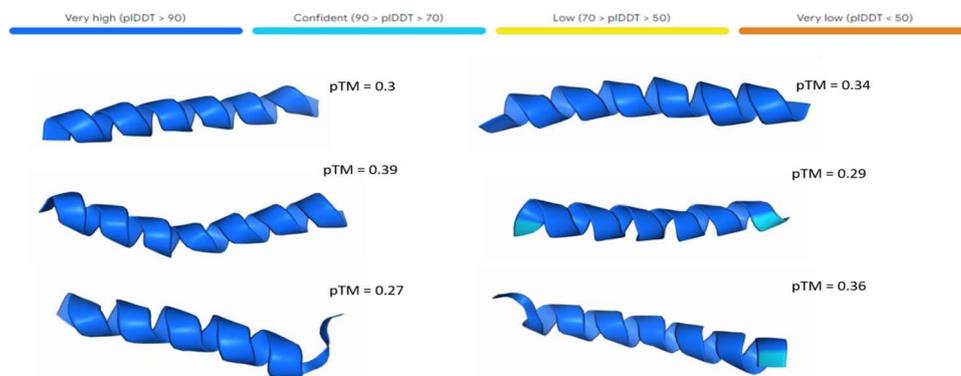


Figure 8. AlphaFold3 structural prediction of peptides generated by the cdGAN model, with probabilities $P(\text{AMP}) > 0.95$ and $P(\text{helix}) > 0.95$. The predicted Template Modeling (pTM) and per-residue local distance difference test (pLDDT) scores are also provided.

were classified as non-hemolytic, compared to AMPGAN's and more the double satisfied both criteria (non-toxic and non-hemolytic). These findings showcase cdGAN's ability to balance structural functionality, therapeutic safety, and biological relevance.

The structural predictions utilizing AlphaFold3 further emphasize the capabilities of the multi-task cdGAN model in generating de novo AMPs. Specifically, the exceptionally high per-residue local distance difference test (pLDDT) scores (> 90) indicate a high degree of confidence in the structural predictions. Additionally, the predicted Template Modeling (pTM) scores, ranging from 0.27 to 0.39, are consistent with the expectations for de novo design, where the sequences are not based on existing structural templates. These results validate the model's ability to generate peptides with α -helical structure, which is essential for membrane disruption and antimicrobial efficacy.

In summary, while models like AMPGAN, HydrAMP, and RLGen demonstrate strengths in specific aspects, they fail to replicate the comprehensive balance achieved by cdGAN. The ability of cdGAN to simultaneously optimize for antimicrobial activity, structural relevance, physicochemical fidelity, toxicity, and hemolytic potential underscores its versatility and effectiveness as a design tool for therapeutic peptide discovery.

Key Points

- A classifier-driven GAN framework (cdGAN), which integrates classifier predictions directly into the loss function, allowing for dynamic and adaptive AMP generation.
- Extension of cdGAN to a multi-task setting, incorporating both antimicrobial activity and structural optimization using a novel ESM2-based multi-task classifier.
- Enhance peptide safety by performing post-generation toxicity and hemolysis screening, ensuring that generated peptides are both effective and viable for therapeutic applications.

Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Conflict of interest: We declare that we have no conflicts of interest related to this publication.

Funding

This work has been supported by the TITAN ERA Chair project (contract no. 101086741) within the Horizon Europe Framework Program of the European Commission and the Hellenic Foundation for Research and Innovation (HFRI) Ph.D. Fellowship grant no. 5647.

Data availability

The datasets and code used in this study are publicly available at https://github.com/aretiz/amp_de_novo_design_cdGAN.

References

1. Somase V, Desai SA, Patel VP. et al. Antimicrobial peptides: Potential alternative to antibiotics and overcoming limitations for future therapeutic applications. *Int J Pept Res Ther* 2024; **30**:45.
2. Kumar P, Kizhakkedathu JN, Straus SK. Antimicrobial peptides: diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo. *Biomolecules* 2018; **8**:4.
3. Giuliani A, Pirri G, Nicoletto S. Antimicrobial peptides: an overview of a promising class of therapeutics. *Open Life Sci* 2007; **2**:1–33. <https://doi.org/10.2478/s11535-007-0010-5>
4. Dijksteel GS, Ulrich MMW, Middelkoop E. et al. Lessons learned from clinical trials using antimicrobial peptides (amps). *Front Microbiol* 2021; **12**:616979. <https://doi.org/10.3389/fmicb.2021.616979>
5. Gupta A, Zou J. Feedback GAN for DNA optimizes protein functions. *Nat Mach Intell* 2019; **1**:105–11. <https://doi.org/10.1038/s42256-019-0017-4>
6. Szymczak P, Możejko M, Grzegorzec T. et al. Discovering highly potent antimicrobial peptides with deep generative model hydramp. *Nat Commun* 2023; **14**:1453. <https://doi.org/10.1038/s41467-023-36994-z>
7. Wang R, Wang T, Zhuo L. et al. Diff-amp: tailored designed antimicrobial peptide framework with all-in-one generation, identification, prediction and optimization. *Brief Bioinform* 2024; **25**:bbae078.
8. Zervou MA, Doutsis E, Pantazis Y. et al. De novo antimicrobial peptide design with feedback generative adversarial networks. *Int J Mol Sci* 2024; **25**:5506. <https://doi.org/10.3390/ijms25105506>

9. Lin Z, Akin H, Rao R. et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 2023;**379**:1123–30. <https://doi.org/10.1126/science.ade2574>
10. Rathore AS, Choudhury S, Arora A. et al. Toxinpred 3.0: an improved method for predicting the toxicity of peptides. *Comput Biol Med* 2024;**179**:108926. <https://doi.org/10.1016/j.combiomed.2024.108926>
11. Chaudhary K, Kumar R, Singh S. et al. A web server and mobile app for computing hemolytic potency of peptides. *Sci Rep* 2016;**6**:22843. <https://doi.org/10.1038/srep22843>
12. Wang G, Li X, Wang Z. Apd3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res* 2016;**44**:D1087–93. <https://doi.org/10.1093/nar/gkv1278>
13. UniProt Consortium. Uniprot: a worldwide hub of protein knowledge. *Nucleic Acids Res* 2019;**47**:D506–15. <https://doi.org/10.1093/nar/gky1049>
14. Steinegger M, Söding J. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 2017;**35**:1026–8. <https://doi.org/10.1038/nbt.3988>
15. Likic V. The Needleman–Wunsch algorithm for sequence alignment. In: *Lecture Given at the 7th Melbourne Bioinformatics Course, BiO21 Molecular Science and Biotechnology Institute*, pp. 1–46. University of Melbourne, 2008.
16. Torrisi M, Kaleel M, Pollastri G. Deeper profiles and cascaded recurrent and convolutional neural networks for state-of-the-art protein secondary structure prediction. *Sci Rep* 2019;**9**:12374. <https://doi.org/10.1038/s41598-019-48786-x>
17. Parvathy ST, Udayasuriyan V, Bhadana V. Codon usage bias. *Mol Biol Rep* 2022;**49**:539–65. <https://doi.org/10.1007/s11033-021-06749-4>
18. Chaney JL, Steele A, Carmichael R. et al. Widespread position-specific conservation of synonymous rare codons within coding sequences. *PLoS Comput Biol* 2017;**13**:e1005531. <https://doi.org/10.1371/journal.pcbi.1005531>
19. Gulrajani I, Ahmed F, Arjovsky M. et al. Improved training of Wasserstein GANs. *Adv Neural Inf Process Syst* 2017;**30**.
20. Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: *International Conference on Machine Learning*, pp. 214–23. PMLR, 2017.
21. Mirza M. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784. 2014.
22. Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier GANs. In: *International Conference on Machine Learning*, pp. 2642–51. PMLR, 2017.
23. Georgoulis E, Zervou MA, Pantazis Y. Transfer Learning on Protein Language Models Improves Antimicrobial Peptide Classification, Preprint, Research Square, <https://doi.org/10.21203/rs.3.rs-5768912/v1>
24. Mookherjee N, Anderson MA, Haagsman HP. et al. Antimicrobial host defence peptides: functions and clinical potential. *Nat Rev Drug Discov* 2020;**19**:311–32. <https://doi.org/10.1038/s41573-019-0058-8>
25. Bhadra P, Yan J, Li J. et al. Ampep: sequence-based prediction of antimicrobial peptides using distribution patterns of amino acid properties and random forest. *Sci Rep* 2018;**8**:1697.
26. Chung C-R, Kuo T-R, Li-Ching W. et al. Characterization and identification of antimicrobial peptides with different functional activities. *Brief Bioinform* 2020;**21**:1098–114. <https://doi.org/10.1093/bib/bbz043>
27. Lawrence TJ, Carper DL, Spangler MK. et al. Ampeppy 1.0: a portable and accurate antimicrobial peptide prediction tool. *Bioinformatics* 2021;**37**:2058–60. <https://doi.org/10.1093/bioinformatics/btaa917>
28. Yan K, Lv H, Guo Y. et al. Sampred-gat: prediction of antimicrobial peptide by graph attention network and predicted peptide structure. *Bioinformatics* 2023;**39**:btac715.
29. Zervou MA, Doutsis E, Tsakalides P. Efficient protein structural class prediction via chaos game representation and recurrent neural networks. In: *ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5. IEEE, 2023.
30. Zervou MA, Doutsis E, Pavlidis P. et al. Structural classification of proteins based on the computationally efficient recurrence quantification analysis and horizontal visibility graphs. *Bioinformatics* 2021;**37**:1796–804. <https://doi.org/10.1093/bioinformatics/btab407>
31. Panda B, Majhi B. A novel improved prediction of protein structural class using deep recurrent neural network. *Evol Intell* 2021;**14**:253–60. <https://doi.org/10.1007/s12065-018-0171-3>
32. Zervou MA, Doutsis E, Pantazis Y. et al. Multitask classification of antimicrobial peptides for simultaneous assessment of antimicrobial property and structural fold. In: *ICASSP 2024–2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1836–40. IEEE, 2024.
33. Jing X, Li F, Leier A. et al. Comprehensive assessment of machine learning-based methods for predicting antimicrobial peptides. *Brief Bioinform* 2021;**22**:bbab083. <https://doi.org/10.1093/bib/bbab083>
34. Gawde U, Chakraborty S, Waghu FH. et al. Campr4: a database of natural and synthetic antimicrobial peptides. *Nucleic Acids Res* 2023;**51**:D377–83. <https://doi.org/10.1093/nar/gkac933>
35. Jhong J-H, Chi Y-H, Li W-C. et al. Dbamp: an integrated resource for exploring antimicrobial peptides with functional activities and physicochemical properties on transcriptome and proteome data. *Nucleic Acids Res* 2019;**47**:D285–97. <https://doi.org/10.1093/nar/gky1030>
36. Kang X, Dong F, Shi C. et al. Dramp 2.0, an updated data repository of antimicrobial peptides. *Sci Data* 2019;**6**:148. <https://doi.org/10.1038/s41597-019-0154-y>
37. Ye G, Wu H, Huang J. et al. Lamp2: a major update of the database linking antimicrobial peptides. *Database* 2020;**2020**:baaa061.
38. Poggio SP, Sessa L, Concilio S. et al. Yadamp: yet another database of antimicrobial peptides. *Int J Antimicrob Agents* 2012;**39**:346–51. <https://doi.org/10.1016/j.ijantimicag.2011.12.003>
39. Van Oort CM, Ferrell JB, Remington JM. et al. Ampgan v2: machine learning-guided design of antimicrobial peptides. *J Chem Inf Model* 2021;**61**:2198–207. <https://doi.org/10.1021/acs.jcim.0c01441>
40. Jaiswal A, AbdAlmageed W, Wu Y. et al. Bidirectional conditional generative adversarial networks. In Jawahar CV, Li H, Mori G, Schindler K (Eds.), *Asian Conference on Computer Vision* (pp. 216–232). Cham: Springer International Publishing.
41. Kingma DP, Mohamed S, Rezende DJ. et al. Semi-supervised learning with deep generative models. *Adv Neural Inf Process Syst* 2014;**27**.
42. Kang S-J, Kim D-H, Mishig-Ochir T. et al. Antimicrobial peptides: their physicochemical properties and therapeutic application. *Arch Pharm Res* 2012;**35**:409–13. <https://doi.org/10.1007/s12272-012-0302-9>
43. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* 1947;**18**:50–60. <https://doi.org/10.1214/aoms/1177730491>
44. Abramson J, Adler J, Dunger J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500.

45. Zhang Q-Y, Yan Z-B, Meng Y-M. *et al.* Antimicrobial peptides: mechanism of action, activity and clinical potential. *Mil Med Res* 2021;**8**:1–25.
46. Jumper J, Evans R, Pritzel A. *et al.* Highly accurate protein structure prediction with alphafold. *Nature* 2021;**596**:583–9. <https://doi.org/10.1038/s41586-021-03819-2>