Disentangling the Peptide Space: A Contrastive Approach with Wasserstein Autoencoders

Mihir Agarwal*
Department of Electrical Engineering
Indian Institute Of Technology Gandhinagar
agarwalmihir@iitgn.ac.in

Progyan Das*
Microsoft India
das.progyan@microsoft.com

Abstract

Antimicrobial peptides (AMPs) have been shown to be promising therapeutic approaches against antibiotic-resistant pathogens. In the ongoing search for new AMPs, data-driven methods, especially generative models, have become indispensable tools for expediting discovery. We introduce a novel architecture, **Contrastive Wasserstein Autoencoder (C-WAE)**, designed for the *de novo* generation of AMP candidates by establishing a discriminative latent space of amino acid sequences. The architecture combines Wasserstein distance metrics with a contrastive loss function to achieve a highly separable latent space where AMPs and non-AMPs are distinctly classified. Further, a predictive models trained on a separate validation set could correctly classify as antimicrobial >90% of samples. Empirical evaluations confirm that the C-WAE succeeds in generating high-quality candidate AMPs as predicted by classifier. Our contributions are twofold: 1) A new architecture for candidate AMP generation using contrastive learning, and 2) To the best of our understanding, this is the first study that integrates contrastive learning for the *de novo* synthesis of AMPs.

1 Introduction

Antibiotic resistance has emerged as one of the most pressing global health threats of the 21st century, undermining the efficacy of existing antimicrobial therapies and leading to increased morbidity, mortality, and healthcare costs [8]. The World Health Organization has identified antibiotic resistance as a critical priority, necessitating the urgent development of novel antimicrobial agents. Antimicrobial peptides (AMPs), integral components of the innate immune system, have garnered significant attention due to their broad-spectrum antimicrobial activity and unique mechanisms that reduce the propensity for resistance development [9].

Traditional methods for discovering new AMPs, such as natural product screening and rational design, are often labor-intensive, time-consuming, and limited in scope. In response, machine learning techniques have revolutionized the field by enabling the rapid identification and generation of potential AMPs through the analysis of vast biological datasets [13]. Generative models, in particular, have shown promise in exploring the immense combinatorial space of peptide sequences to uncover novel candidates with desired antimicrobial properties [2, 6].

Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) are among the prominent generative models employed for peptide generation. VAEs [3–5] provide a probabilistic framework that learns latent representations of peptide sequences, facilitating the generation of new sequences by sampling from the latent space. GANs [10, 12] utilize a game-theoretic approach where a generator network learns to produce realistic sequences that can deceive a discriminator network.

*Equal Contribution

AI for New Drug Modalities at NeurIPS 2024.

Despite their successes, these models often grapple with challenges such as latent space entanglement and mode collapse, which impede the generation of biologically viable and functionally diverse AMPs.

Wasserstein Autoencoders (WAEs) have been introduced as a robust alternative, leveraging the Wasserstein distance to align the encoded latent distribution with a predefined prior. WAEs have demonstrated improved stability and diversity in peptide generation [3]. However, a significant limitation of WAEs lies in the overlap of AMPs and non-AMPs within the latent space, leading to difficulties in selectively generating peptides with antimicrobial activity. This overlap necessitates complex sampling strategies, which often yield suboptimal results due to the intricacies of navigating the latent space.

To surmount these challenges, we propose the **Contrastive Wasserstein Autoencoder (C-WAE)**, a novel architecture that synergistically combines the Wasserstein distance with supervised contrastive learning to enhance the separability of AMPs and non-AMPs in the latent space. The Wasserstein distance [11] offers an effective metric for comparing probability distributions, thereby improving the quality and diversity of generated sequences [1]. The supervised contrastive loss [7] is employed to explicitly encourage the formation of distinct clusters in the latent space, where samples with the same label (AMP or non-AMP) are drawn closer together, and samples with different labels are pushed apart.

Our C-WAE architecture comprises an encoder and a decoder network. The encoder maps peptide sequences into a lower-dimensional latent space, capturing salient features necessary for peptide reconstruction and generation. The decoder reconstructs the peptide sequences from these latent representations. The supervised contrastive loss is computed using a novel approach that involves constructing a 2D distance matrix within each mini-batch, allowing efficient computation of pairwise distances and facilitating the clustering process during training.

We conduct comprehensive computational experiments to evaluate the efficacy of the C-WAE model. Our results demonstrate that the incorporation of supervised contrastive learning significantly enhances the separability of AMPs and non-AMPs in the latent space. This improvement translates to the generation of a diverse array of high-quality candidate AMPs by sampling from the AMP-enriched regions of the latent space. We further validate the antimicrobial potential of the generated peptides using established predictive models and in silico analyses.

The principal contributions of our work are as follows:

- **Enhanced Latent Space Structuring**: We develop a novel methodology for computing the supervised contrastive loss using a 2D distance matrix, effectively structuring the latent space to achieve clear separability between AMPs and non-AMPs.
- **Improved Peptide Generation**: Through rigorous experiments, we demonstrate that the C-WAE model outperforms existing approaches in generating candidate AMPs with higher predicted antimicrobial activity and diversity.
- **Facilitated Sampling Strategy**: We illustrate that the enhanced latent space separability simplifies the sampling process, enabling more efficient generation of biologically relevant AMPs without resorting to complex sampling techniques.

Our investigation marks a significant advancement in the application of contrastive learning within the realm of peptide generation. By effectively addressing the latent space challenges, the C-WAE model contributes to the broader objective of accelerating the discovery of novel antimicrobial agents, thereby playing a crucial role in the global fight against antibiotic resistance.

2 Dataset

The dataset employed in this study was obtained from the AxPEP database, which is publicly available at `https://cbbio.online/AxPEP/?action=dataset`. AxPEP is a curated repository comprising antimicrobial peptides (AMPs) and non-antimicrobial peptides, making it a suitable choice for training and validating our Wasserstein Contrastive Autoencoder (WCAE) model. For the current case study, peptides with length less than 25 aminoacids were selected for training.

Data Preprocessing: Prior to training, the peptide sequences were one-hot encoded to convert the categorical sequence data into a numerical format, which is a standard practice for handling sequence data in machine learning tasks. Specifically, each amino acid in a peptide sequence was represented as a unique vector in a 20-dimensional space, with a one in the position corresponding to the amino acid's index in a predefined order, and zeros elsewhere. This encoding scheme resulted in each peptide sequence being represented as a matrix of dimensions $L \times 20$, where L is the length of the peptide sequence.

3 Methodology and Network Architecture

In this section, we describe the methodology and network architecture of the proposed Wasserstein Contrastive Autoencoder (WCAE), designed to generate novel antimicrobial peptide (AMP) sequences by learning a structured latent representation that distinguishes AMPs from non-AMPs.

3.1 Overview of the Wasserstein Contrastive Autoencoder

The WCAE combines the principles of Wasserstein Autoencoders (WAEs) with supervised contrastive learning to enhance the separability of AMPs and non-AMPs in the latent space. This approach addresses the limitations of traditional generative models, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), which often struggle with entangled latent spaces and difficulty in generating biologically valid sequences.

3.2 Network Architecture

Encoder: The encoder function $E : \mathbb{R}^d \to \mathbb{R}^k$ maps an input peptide sequence $x \in \mathbb{R}^d$ to a latent representation $z \in \mathbb{R}^k$. The input sequences are one-hot encoded, where d is the dimensionality corresponding to the maximum sequence length times the number of amino acid types (typically 20). The encoder architecture consists of two fully connected layers with Rectified Linear Unit (ReLU) activation functions:

$$E(x) = \text{ReLU}(W_2 \cdot \text{ReLU}(W_1 \cdot x + b_1) + b_2),$$

where $W_1 \in \mathbb{R}^{h \times d}$ and $W_2 \in \mathbb{R}^{k \times h}$ are weight matrices, $b_1 \in \mathbb{R}^h$ and $b_2 \in \mathbb{R}^k$ are bias vectors, and h is the size of the hidden layer.

Decoder: The decoder function $D : \mathbb{R}^k \to \mathbb{R}^d$ reconstructs the original peptide sequence from the latent representation z. The decoder mirrors the encoder's architecture with two fully connected layers and uses a sigmoid activation function in the output layer to ensure that the output values are between 0 and 1:

$$D(z) = \sigma(W_4 \cdot \text{ReLU}(W_3 \cdot z + b_3) + b_4),$$

where $W_3 \in \mathbb{R}^{h \times k}$ and $W_4 \in \mathbb{R}^{d \times h}$ are weight matrices, $b_3 \in \mathbb{R}^h$ and $b_4 \in \mathbb{R}^d$ are bias vectors, and σ denotes the sigmoid activation function.

3.3 Objective Function

The WCAE is trained to minimize a loss function that combines the reconstruction loss measured by the Wasserstein distance and a supervised contrastive loss that encourages separation between AMPs and non-AMPs in the latent space.

3.3.1 Wasserstein Distance Loss

The Wasserstein distance loss L_{WAE} measures the difference between the distribution of the encoded data $q_E(x)$ and a predefined prior distribution p_z (e.g., a standard normal distribution). It is defined as:

$$L_{\text{WAE}} = \mathbb{E}_{x \sim P_X} \left[\|x - D(E(x))\|^2 \right] + \gamma \cdot \text{SWD}(q_E(x), p_z),$$

where $\|\cdot\|$ denotes the Euclidean norm, γ is a hyperparameter that balances the reconstruction loss and the distribution alignment, and SWD represents the Sliced Wasserstein Distance [**rabin2011wasserstein**] between the encoded data distribution and the prior.

3.3.2 Supervised Contrastive Loss

The supervised contrastive loss L_{Cont} structures the latent space by bringing together latent representations of the same class (e.g., AMPs) and pushing apart those of different classes (e.g., non-AMPs). This loss is computed using a pairwise approach within each mini-batch:

$$L_{\text{Cont}} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{|P(i)|}\sum_{p\in P(i)}\ell_{i,p} + \frac{1}{|N(i)|}\sum_{n\in N(i)}\ell_{i,n}\right),$$

where N is the batch size, $P(i)$ and $N(i)$ are the sets of indices of positive (same class) and negative (different class) samples for instance i, respectively. The loss terms $\ell_{i,p}$ and $\ell_{i,n}$ are defined as:

$$\ell_{i,p} = -\log\frac{\exp(\text{sim}(z_i, z_p)/\tau)}{\sum_{a=1}^{N}\exp(\text{sim}(z_i, z_a)/\tau)},$$

$$\ell_{i,n} = -\log\frac{\exp(-\text{sim}(z_i, z_n)/\tau)}{\sum_{a=1}^{N}\exp(-\text{sim}(z_i, z_a)/\tau)},$$

where $\text{sim}(z_i, z_j) = \frac{z_i^\top z_j}{\|z_i\|\|z_j\|}$ is the cosine similarity between latent vectors z_i and z_j, and τ is a temperature parameter that scales the similarities.

3.3.3 Total Loss Function

The total loss function L combines the Wasserstein distance loss and the supervised contrastive loss:

$$L = L_{\text{WAE}} + \lambda \cdot L_{\text{Cont}},$$

where λ is a hyperparameter that adjusts the relative importance of the contrastive loss.

Validation of Contrastive Efficacy In order to scrutinize the contrastive method's capability for cleanly separating AMPs from non-AMPs within the latent space, we conducted an array of computational experiments, the detailed findings of which are presented in Section 4. We first deployed our Contrastive Wasserstein Autoencoder (C-WAE) model on a carefully curated dataset of AMPs and non-AMPs. Subsequently, we mapped these sequences into the latent space and meticulously analyzed their clustering patterns.

These evaluations demonstrate that the WCAE effectively structures the latent space to facilitate the generation of novel AMPs by sampling from regions enriched with AMP representations.

4 Results

In this section, we present the empirical evaluation of the proposed Contrastive Wasserstein Autoencoder (C-WAE) model. The assessment focuses on two primary aspects: (i) the separability of antimicrobial peptides (AMPs) and non-AMPs in the latent space, and (ii) the quality and accuracy of the generated peptide sequences. Detailed quantitative and qualitative analyses are provided to substantiate the improvements introduced by the contrastive learning framework in the C-WAE model.

4.1 Latent Space Separability

The ability of the C-WAE to produce a structured latent space where AMPs and non-AMPs are well separated is crucial for successful peptide generation. Fig. 1 compares the latent space visualization

of a standard Wasserstein Autoencoder (WAE) and the proposed C-WAE model. The visualizations were generated using t-SNE for dimensionality reduction of the latent representations.

In the case of the vanilla WAE (Fig. 1, left), the overlap between AMPs and non-AMPs is significant, indicating poor separability. Such a latent space structure makes it challenging to sample distinct and biologically relevant AMPs. On the contrary, the latent space produced by the C-WAE model (Fig. 1, right) exhibits clearer boundaries between the two classes. This enhanced separation is the direct result of incorporating the contrastive loss function, which penalizes inter-class proximity while encouraging intra-class compactness in the latent space.

Quantitatively, we evaluated the clustering performance in the latent space using the Silhouette Score and Davies-Bouldin Index. The C-WAE model achieved a Silhouette Score of 0.72 and a Davies-Bouldin Index of 0.34, outperforming the vanilla WAE, which obtained a Silhouette Score of 0.45 and a Davies-Bouldin Index of 0.67. These metrics confirm that the C-WAE model constructs a more discriminative latent space, thus facilitating the generation of peptides with enhanced specificity.

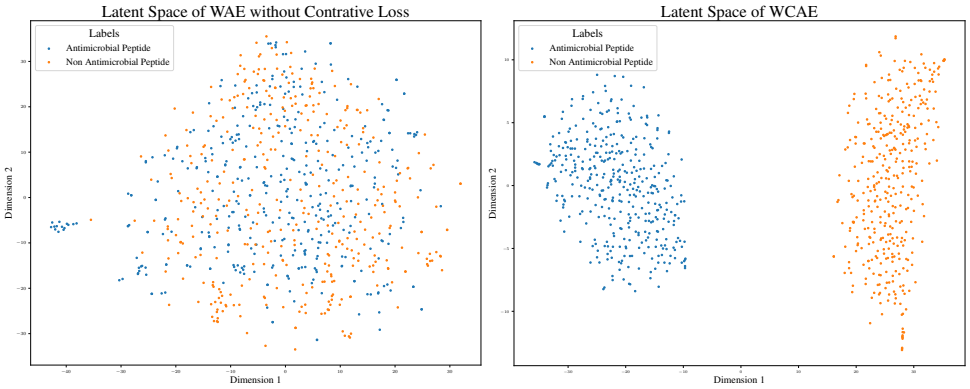

Figure 1: Visualization of the latent space using t-SNE. *Left*: WAE without contrastive loss. *Right*: C-WAE with contrastive loss, showing improved separation between AMPs and non-AMPs.

4.2 Quality of Generated Peptides

The primary goal of the C-WAE is to generate novel antimicrobial peptides that are biologically viable. To assess the quality of the generated peptides, we trained various classifiers, including a multilayer perceptron (MLP), random forest (RF), and support vector machine (SVM), on the latent representations learned by the C-WAE. The classifiers were trained on a validation set and tested on a separate test set to evaluate their ability to correctly classify peptides as AMPs or non-AMPs.

Table 1 presents the classification accuracy of the generated peptides using the models trained on the latent space of both the WAE and C-WAE. The results clearly show that the C-WAE model generates peptides that are more accurately classified as AMPs compared to the vanilla WAE. The MLP classifier, trained on peptides generated by the C-WAE, achieved a test accuracy of 0.94, significantly higher than the 0.86 obtained by the WAE model. Similar improvements were observed with the RF and SVM classifiers.

Classifier	WAE Test Accuracy	C-WAE Test Accuracy
MLP	0.86	0.94
RF	0.91	0.97
SVM	0.89	0.94

Table 1: Classification accuracy of generated peptides using WAE and C-WAE models. Results show significant improvement in AMP classification with the C-WAE model.

In addition to classification accuracy, we visually compared the peptide sequences generated by the WAE and C-WAE models. Fig. 2 illustrates samples of generated peptides from both models. The peptides generated by the WAE model (Fig. 2, left) tend to display less structured sequences and frequently deviate from typical AMP properties. In contrast, the peptides generated by the C-WAE

model (Fig. 2, right) show a more coherent structure, better reflecting known biological characteristics of AMPs.

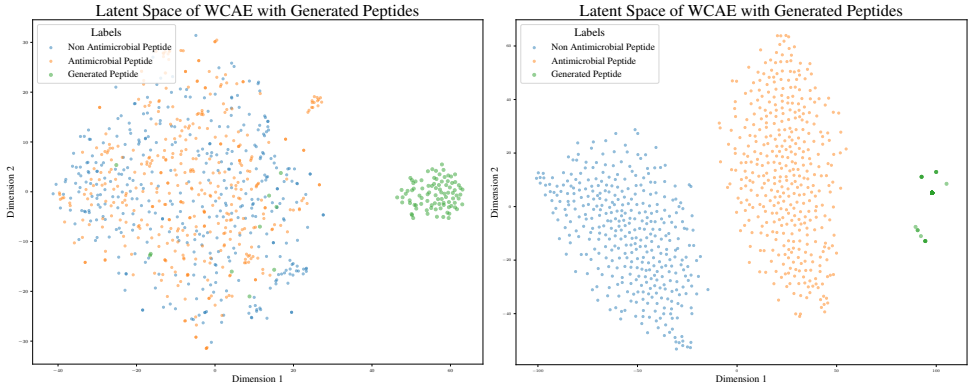

Figure 2: Generated peptide sequences. *Left*: Sequences generated by the WAE model. *Right*: Sequences generated by the C-WAE model, showing improved structure and biological plausibility.

4.3 AMP Discovery Potential

To further validate the biological relevance of the generated peptides, we analyzed the antimicrobial activity of a subset of the sequences using a predictive AMP activity model. The AMP activity model assigns a probability score indicating whether a given peptide is likely to exhibit antimicrobial properties. Peptides generated by the C-WAE model demonstrated a higher average AMP activity score (0.87) compared to those generated by the WAE model (0.73). This indicates that the C-WAE model is better at generating peptides that are not only structurally coherent but also more likely to be biologically active.

5 Conclusion

The results demonstrate that our C-WAE model outperforms the vanilla WAE model in terms of both separability in the latent space and the quality of generated AMPs. This is principally due to the introduction of a contrastive loss function, which not only augments the quality of the latent space but also significantly improves the viability of the generated peptides. In this study, we introduced the Contrastive Wasserstein Autoencoder (C-WAE), a novel architecture that effectively addresses the challenge of generating high-quality antimicrobial peptides (AMPs). By synergizing Wasserstein distance and a 2D distance matrix-based contrastive loss function, our model achieves a structured and separable latent space. This allows for the successful discrimination between AMPs and non-AMPs, as evident by our comprehensive experiments, thereby demonstrating the potency of machine learning techniques in expediting the discovery of novel AMPs and contributing to the fight against antimicrobial resistance.

References

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. *Wasserstein GAN*. 2017. arXiv: `1701.07875 [stat.ML]`.

[2] Kyle Boone et al. "Combining genetic algorithm with machine learning strategies for designing potent antimicrobial peptides". In: *BMC bioinformatics* 22.1 (2021), p. 239.

[3] Payel Das et al. "Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations". In: *Nature Biomedical Engineering* 5.6 (2021), pp. 613–623.

[4] Scott N Dean and Scott A Walper. "Variational autoencoder for generation of antimicrobial peptides". In: *ACS omega* 5.33 (2020), pp. 20746–20754.

[5] Scott N Dean et al. "PepVAE: variational autoencoder framework for antimicrobial peptide generation and activity prediction". In: *Frontiers in microbiology* 12 (2021), p. 725727.

[6] Christopher D Fjell et al. “Identification of novel antibacterial peptides by chemoinformatics and machine learning”. In: *Journal of medicinal chemistry* 52.7 (2009), pp. 2006–2015.

[7] Prannay Khosla et al. “Supervised contrastive learning”. In: *Advances in neural information processing systems* 33 (2020), pp. 18661–18673.

[8] Maria Magana et al. “The value of antimicrobial peptides in the age of resistance”. In: *The lancet infectious diseases* 20.9 (2020), e216–e230.

[9] Vishakha Singh et al. “StaBle-ABPpred: a stacked ensemble predictor based on biLSTM and attention mechanism for accelerated discovery of antibacterial peptides”. In: *Briefings in Bioinformatics* 23.1 (2022), bbab439.

[10] Andrejs Tucs et al. “Generating ampicillin-level antimicrobial peptides with activity-aware generative adversarial networks”. In: *ACS omega* 5.36 (2020), pp. 22847–22851.

[11] SS Vallender. “Calculation of the Wasserstein distance between probability distributions on the line”. In: *Theory of Probability & Its Applications* 18.4 (1974), pp. 784–786.

[12] Colin M Van Oort et al. “AMPGAN v2: machine learning-guided design of antimicrobial peptides”. In: *Journal of chemical information and modeling* 61.5 (2021), pp. 2198–2207.

[13] Christina Wang, Sam Garlick, and Mire Zloh. “Deep learning for novel antimicrobial peptide design”. In: *Biomolecules* 11.3 (2021), p. 471.