AMP-UGEN: PROMOTING ANTIMICROBIAL PEPTIDE GENERATION THROUGH PEPTIDE UNDERSTANDING

Anonymous authors
Paper under double-blind review

ABSTRACT

Existing antimicrobial peptide (AMP) generation models primarily learn the probability distribution of known AMP sequences. Although this paradigm can produce diverse AMP-like candidates, generation is commonly driven by sequence likelihoods, property labels, or predictor scores, without explicitly incorporating the functional and physicochemical knowledge captured by peptide understanding. We introduce AMP-UGen, an understanding-driven generation framework that turns such knowledge into learnable signals through two complementary mechanisms, Understanding-to-Design Transfer (UDT) and Understanding-Feedback Policy Optimization (UFPO). UDT transfers design-semantic states extracted by the peptide understanding model to a masked diffusion model, allowing peptide understanding to guide sequence construction. UFPO reuses the same understanding component to evaluate generated candidates and converts its assessments into feedback for updating the generation policy. Among the evaluated generation models, AMP-UGen achieves the highest predicted-positive rates under the frozen HydrMIC and AMPlify predictors, reaching 56.72% and 82.34%, respectively. During property optimization, UFPO reduces charge MAE from 1.183 to 0.311 and hydrophobic-moment MAE from 0.0273 to 0.0163. These results show that AMP-UGen improves predicted AMP quality and property control by integrating peptide understanding into both sequence construction and policy optimization, demonstrating its potential as a general framework for peptide generation.

1 INTRODUCTION

Antimicrobial peptides (AMPs) are promising therapeutic candidates for drug-resistant infections because their broad sequence space and diverse mechanisms of action offer opportunities beyond conventional antibiotics (Oliveira Júnior et al., 2025). Their practical discovery, however, remains difficult because viable candidates must jointly satisfy a desired antimicrobial function, physicochemical requirements, and biological plausibility. These requirements are coupled rather than independent, since a sequence alteration that improves one property can impair another aspect of peptide quality. Computational AMP design must therefore do more than sample unseen sequences. It must construct candidates that remain responsive to design intent while preserving the biological regularities of functional peptides.

Recent AMP generation models, ranging from variational and adversarial frameworks to diffusion-based approaches, have substantially expanded the computational peptide design space (Szymczak et al., 2023; Wang et al., 2024a; 2025b; Torres et al., 2025). These methods can learn broad sequence regularities from AMP corpora and generate diverse candidates, but their generation process is still largely driven by the probability distribution of observed sequences, external property labels, or predictor-derived scores. As a result, sequence construction is often weakly coupled with functional interpretation. The model may learn what AMP-like sequences look like, without explicitly turning peptide-function understanding into a design signal. This raises the central question of this work, namely how peptide understanding can be transformed from a passive predictive capability into a learnable signal that promotes AMP generation throughout candidate construction and generation-policy updates.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

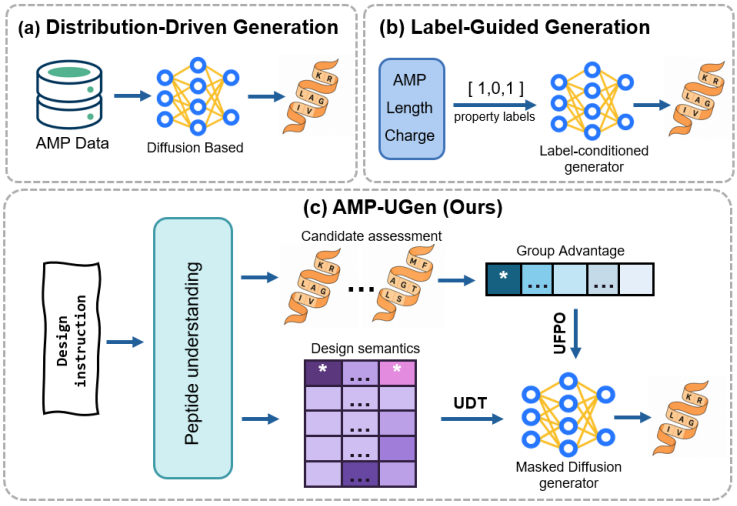

Figure 1: Comparison of three AMP generation paradigms. Existing approaches learn from AMP sequence distributions or simple property labels. AMP-UGen instead uses peptide understanding for forward design-semantic transfer through UDT and post-generation candidate feedback through UFPO.

We address this question with AMP-UGen, an understanding-driven AMP generation framework that couples an autoregressive peptide understanding model with a masked diffusion generation model. We first consider a direct autoregressive alternative in which the understanding model itself generates peptide sequences. Empirical comparisons indicate that this direct formulation is weaker than separating semantic interpretation from sequence construction. The autoregressive model is more effective as a source of design semantics, while masked diffusion provides a stronger sequence-construction mechanism. AMP-UGen therefore promotes generation through two named mechanisms. Understanding-to-Design Transfer (UDT) passes design-semantic states extracted by the understanding model to the masked diffusion model as explicit conditions, allowing peptide-understanding priors to guide candidate construction. Understanding-Feedback Policy Optimization (UFPO) then reuses the same understanding component to interpret generated candidates, combines its assessments with exact physicochemical feedback, and updates the generation policy. In this way, AMP-UGen incorporates peptide understanding into both sequence construction and subsequent policy optimization.

Experiments across peptide understanding, AMP generation, and property optimization support this design. AMP-domain fine-tuning of Prot2Text-V2 yields AMP-SFT and raises macro-F1 from 0.2029 to 0.5923. Among the evaluated generation models, AMP-UGen attains the highest predicted-positive rates under the frozen HydrMIC and AMPlify predictors, reaching 56.72% and 82.34%, respectively. During property optimization, UFPO reduces charge MAE from 1.183 to 0.311 and hydrophobic-moment MAE from 0.0273 to 0.0163. Ablations further show that removing understanding feedback reduces predicted AMP quality.

Our main contributions are listed below.

- We present AMP-UGen, a unified framework for AMP property prediction, peptide understanding, and de novo generation, supporting both peptide analysis and sequence design.
- We introduce Understanding-to-Design Transfer (UDT), which uses peptide-understanding states as explicit design conditions for masked diffusion generation and achieves the highest HydrMIC and AMPlify positive rates among the evaluated models.
- We propose Understanding-Feedback Policy Optimization (UFPO), which converts assessments from the same peptide understanding component into policy-optimization feedback and improves condition-level alignment with target physicochemical properties.
- We evaluate understanding-promoted generation across multiple frozen activity predictors, non-hemolysis screening, and property-control targets, demonstrating complementary benefits from pre-generation semantic transfer and post-generation understanding feedback.

2 Related Work

2.1 Distribution-Oriented AMP Generation

Sequence-based AMP discovery is commonly formulated as supervised recognition, which enables efficient screening but does not directly construct new candidates. Representative methods range from the convolutional and recurrent AMPScannerV2 and the short-peptide predictor DeepAmPEP30 to the attentive multi-scale AMPlify (Veltri et al., 2018; Yan et al., 2020; Li et al., 2022). Macrel extends prediction to AMP and hemolysis screening in genomic and metagenomic data, while ESCAPE provides a standardized multi-label benchmark for comparing AMP classifiers (Santos-Júnior et al., 2020; Ojeda et al., 2025). Collectively, these models extract useful sequence and activity patterns, yet their outputs remain task-specific labels or scores used to retrieve and rank existing sequences.

Generative AMP design has progressed from latent-variable and adversarial models to diffusion models. PepVAE couples latent-space sampling with activity prediction, AMPGAN v2 introduces bidirectional attribute control, and HydrAMP supports both de novo and analogue generation with computational preselection (Dean et al., 2021; Van Oort et al., 2021; Szymczak et al., 2023). Recent diffusion systems further expand the design space. TG-CDDPM uses textual descriptions to guide conditional denoising, Diff-AMP organizes generation and reinforcement learning as separate modules, AMPGen preserves evolutionary information, and AMP-Diffusion and latent diffusion pipelines combine large-scale generation with downstream computational or experimental prioritization (Cao et al., 2025; Wang et al., 2024a; Jin et al., 2025; Torres et al., 2025; Wang et al., 2025b). These methods are effective at modeling AMP sequence distributions and producing diverse candidates. Their biological priors, however, are usually implicit in training data, pretrained representations, or post-generation filters, rather than expressed as design semantics that directly participate in sequence construction.

2.2 External-Signal-Guided Controllable Design

Another line of work introduces external signals to steer generation. CLaSS combines latent generation with classifier guidance and molecular simulation, while classifier guidance and computational screening are also used by AMPGAN v2, HydrAMP, DPLM, and Diff-AMP (Das et al., 2021; Van Oort et al., 2021; Szymczak et al., 2023; Wang et al., 2024b;a). AMP-Designer builds an AMP foundation model and combines prompt learning, knowledge distillation, activity prediction, and reinforcement learning (Wang et al., 2025a). More broadly, model-based reinforcement learning has been applied to biological sequence design, protein language models have served as proxy rewards for mutation policies, and CtrlProt uses preference optimization for single- and multi-attribute control (Angermueller et al., 2020; Subramanian et al., 2024; Liu et al., 2025b).

These methods demonstrate the importance of controllable generation and feedback optimization. Their control signals are often property labels, prompts, task-specific predictors, or reward models that are specified outside the process that interprets the full design instruction. This separation is not necessarily a flaw, since external predictors provide practical and measurable objectives. It does mean, however, that the generation process is commonly optimized toward externally supplied scores rather than toward a semantic state derived from a peptide understanding process. AMP-UGen targets this gap by making peptide understanding a source of both generation conditioning and feedback construction.

2.3 Peptide Understanding and Understanding-Promoted Generation

Large-scale protein models learn broad sequence, structural, and functional priors that can support interpretation and design. ESM-2 provides transferable representations from evolutionary-scale pretraining, while ProtST aligns proteins with biomedical text for representation learning and zero-shot prediction (Lin et al., 2023; Xu et al., 2023). InstructProtein supports instruction-based interaction between proteins and text, and Prot2Text-V2 generates open-ended functional descriptions from sequence representations (Wang et al., 2024c; Fei et al., 2025). These models encode biological context beyond a single AMP predictor, but they are typically used for understanding, annotation, or representation learning rather than direct AMP sequence construction.

Protein generation has also moved toward architectures that combine predictive and generative objectives. ProGen uses functional control tags for conditional autoregressive generation, ProtGPT2 learns natural-like protein sequences without task labels, ProteinDT aligns text and protein representations for guided generation and editing, DPLM combines predictive representations with diffusion generation, and xTrimoPGLM jointly accommodates understanding and generation objectives (Madani et al., 2023; Ferruz et al., 2022; Liu et al., 2025a; Wang et al., 2024b; Chen et al., 2024). Hyformer further shares parameters between autoregressive molecular generation and property prediction and evaluates this formulation on peptide design (Izdebski et al., 2025). These studies narrow the distance between prediction and generation at the level of objectives, representations, or shared parameters. AMP-UGen follows a different design principle for AMP generation. UDT converts peptide-understanding states into PDQ design semantics and uses them as direct conditions for masked diffusion, while UFPO uses understanding feedback together with exact physicochemical feedback to optimize the generation policy. The contribution is therefore not simply placing understanding and generation in one system, but using peptide understanding as an active mechanism that promotes generation before and during policy optimization.

3 METHOD

3.1 PROBLEM FORMULATION AND FRAMEWORK OVERVIEW

Let $c = (a, L, \mathbf{y})$ denote a peptide-design instruction, where a specifies the desired antimicrobial activity, L is the target length, and $\mathbf{y}$ contains optional physicochemical targets. We consider net charge, hydrophobicity measured by the grand average of hydropathy (GRAVY), and hydrophobic moment. The goal is to learn a conditional distribution over amino-acid sequences $\mathbf{x} = (x_1, \ldots, x_L)$,

$$p_\theta(\mathbf{x} \mid c), \qquad x_i \in \mathcal{V}_{\mathrm{aa}},$$

such that the generated sequence is both compatible with the functional intent in c and responsive to its explicit property constraints.

Our solution contains two understanding-promoted stages. As illustrated in Figure 2, the first stage, Understanding-to-Design Transfer (UDT), connects an AMP-oriented peptide understanding model to a masked diffusion generation model through a learnable semantic interface. This allows sequence construction to condition on biological knowledge represented by the understanding model rather than on isolated attribute labels. The second stage, Understanding-Feedback Policy Optimization (UFPO), closes the loop. Candidates sampled from AMP-UGen are interpreted by the same understanding model, and this understanding signal is combined with exact property feedback to optimize the generation policy. Thus, UDT turns understanding into pre-generation design semantics, whereas UFPO turns understanding into optimization feedback.

AMP-UGen achieves unification at the task and representation levels rather than by requiring understanding and generation to share an identical output space. Specifically, it connects an autoregressive peptide understanding model and a masked diffusion generator through an end-to-end learned semantic interface. Peptide Design Queries extract design-relevant states from the understanding model, allowing these states to participate directly in sequence denoising.

3.2 UNDERSTANDING-TO-DESIGN TRANSFER FOR AMP GENERATION

AMP-oriented peptide understanding. Prot2Text-V2 is a general protein-understanding model that generates functional descriptions from protein sequence representations (Fei et al., 2025). We initialize U_ϕ from its checkpoint and fine-tune it on AMP-domain instructions, yielding the AMP-SFT understanding component. This adaptation focuses its joint protein–text representation on antimicrobial function and physicochemical characteristics relevant to peptide design. Given a peptide sequence $\mathbf{s}$, an instruction q, and a target textual response $\mathbf{z}$, U_ϕ is trained with the autoregressive objective

$$\mathcal{L}_{\mathrm{und}} = -\sum_{j=1}^{|\mathbf{z}|} \log p_\phi(z_j \mid z_{<j}, \mathbf{s}, q).$$

The instruction data cover free-form peptide descriptions and questions about net charge, GRAVY, and hydrophobic moment. This adaptation is intended to concentrate the broad sequence and func-

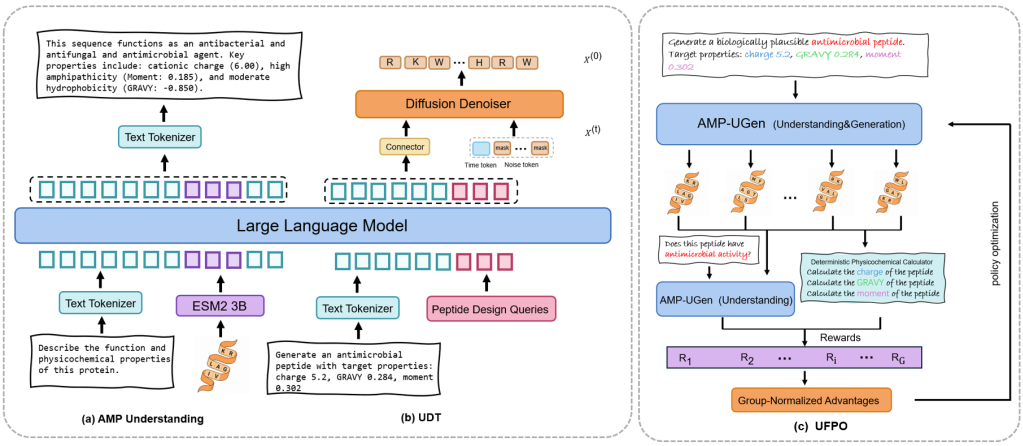

Figure 2: Overview of AMP-UGen. (a,b) Understanding-to-Design Transfer (UDT) uses Peptide Design Queries (PDQs) to extract generation-oriented design semantics from the autoregressive AMP understanding model and maps the resulting states into the conditioning space of a masked diffusion denoiser. (c) Understanding-Feedback Policy Optimization (UFPO) reuses the same peptide understanding component to evaluate generated candidates and convert understanding feedback into policy-optimization signals.

tional priors acquired during protein pretraining into a representation relevant to AMP interpretation, rather than to train a separate reward classifier from scratch.

Peptide Design Queries for semantic conditioning. For a design instruction c, we append K learnable *Peptide Design Queries* (PDQs), $\mathbf{Q}_D = (q_1^D, \ldots, q_K^D)$, to the language-model context. The hidden states at the PDQ positions form a fixed-length set of design semantic states,

$$\mathbf{H}_D = \mathrm{Select}_{\mathbf{Q}_D}\big(U_\phi([c; \mathbf{Q}_D])\big) \in \mathbb{R}^{K \times d}.$$

In our model, $K = 32$. The PDQs aggregate information about activity and property requirements from the understanding space and expose the resulting design semantics to the peptide generator without requiring the two modules to share their token vocabularies or output heads. This PDQ interface is the central connection in AMP-UGen. The design condition is first interpreted as protein-domain semantics and is then used to construct a sequence.

Understanding-conditioned masked diffusion. We instantiate the sequence generator with a discrete masked diffusion model following the masked denoising formulation of DPLM (Wang et al., 2024b). For a training peptide $\mathbf{x}_0$, a time step t determines a corruption process $q_t(\mathbf{x}_t \mid \mathbf{x}_0)$ that replaces a subset $\mathcal{M}_t$ of residues with mask tokens. Conditioned on $\mathbf{H}_D$, the denoiser predicts the original residues from the corrupted sequence.

$$p_\theta(\mathbf{x}_0 \mid \mathbf{x}_t, t, \mathbf{H}_D).$$

The diffusion objective is

$$\mathcal{L}_{\mathrm{diff}} = -\mathbb{E}_{\mathbf{x}_0, t, \mathbf{x}_t}\left[\sum_{i \in \mathcal{M}_t} \log p_\theta(x_{0,i} \mid \mathbf{x}_t, t, \mathbf{H}_D)\right].$$

We randomly replace a portion of semantic conditions with zero states during training, enabling conditional and unconditional predictions within the same denoiser. The AMP-adapted understanding backbone supplies a stable semantic prior, while the PDQs and masked diffusion generator are optimized end to end through $\mathcal{L}_{\mathrm{diff}}$. At inference, generation starts from a fully masked sequence and progressively reconstructs residues under $\mathbf{H}_D$ until a complete peptide is obtained.

Table 1: Peptide understanding and AMP generation results. Panel (a) reports functional and physicochemical understanding; AMP-SFT denotes Prot2Text-V2 fine-tuned on AMP instructions. AntiB., AntiF., AntiV., AntiP., and AntiM. denote the five activity labels. Panel (b) reports candidate-pool quality: AMP-SFT generator further tunes AMP-SFT on generation data; *w/o condition* omits design instructions; and *frozen PDQs* fixes the query embeddings. ESC-AB is the frozen ESCAPE antibacterial predictor, and Non-hem. is the predicted non-hemolytic rate. For every generation method, results are averaged over five independent runs with different random seeds, with 2,000 candidates generated per run (10,000 candidates in total per method). Best and second-best generated results are bold and underlined; ESCAPE train is a data reference.

(a) AMP understanding

Method	Functional classification						Charge		GRAVY		Moment	
	Macro-F1 ↑	AntiB. ↑	AntiF. ↑	AntiV. ↑	AntiP. ↑	AntiM. ↑	MAE ↓	r ↑	MAE ↓	r ↑	MAE ↓	r ↑
LLaMA-3.1-8B-Instruct	0.1997	0.3239	0.1478	0.1086	0.0093	0.4087	6.4717	−0.0859	2.9880	−0.0263	2.9225	0.1022
Prot2Text-V2	0.2029	0.3329	0.1487	0.1108	0.0104	0.4116	748.0599	−0.0221	891.4699	0.0467	55103.1673	0.0034
AMP-SFT (ours)	**0.5923**	**0.8710**	**0.5079**	**0.5460**	**0.1533**	**0.8831**	**2.1913**	**0.8589**	**0.2106**	**0.5589**	**0.0715**	**0.4425**

(b) AMP generation

Method	Unique ↑	Novel ↑	HydrAMP ↑	HydrMIC ↑	ESC-AB ↑	AMPlify ↑	Non-hem. ↑	Fitness ↑
ESCAPE train AMPs	100.00	0.00	70.27	42.75	77.33	77.10	44.78	0.1208
AMP-Diffusion	<u>99.95</u>	<u>99.98</u>	57.40	25.18	52.64	25.19	**64.68**	0.0622
Diff-AMP	**100.00**	**100.00**	48.76	17.03	7.24	7.72	2.61	0.0521
HydrAMP	**100.00**	**100.00**	**78.09**	<u>43.45</u>	<u>70.83</u>	<u>59.33</u>	36.36	0.0874
AMP-SFT generator	97.39	99.40	60.24	35.34	55.22	53.88	61.46	**0.0963**
AMP-UGen (w/o condition)	**100.00**	**100.00**	56.45	30.60	68.57	56.00	54.11	0.0614
AMP-UGen (frozen PDQs)	**100.00**	**100.00**	45.63	17.15	68.85	44.23	56.80	0.0717
AMP-UGen	**100.00**	**100.00**	<u>74.06</u>	**56.72**	**72.32**	**82.34**	<u>61.98</u>	<u>0.0917</u>

3.3 Understanding-Feedback Policy Optimization

Understanding-consistent reward. Although UDT transfers design semantics to the masked diffusion model, likelihood-based training alone does not directly optimize the quality of complete sampled peptides. For every condition c, UFPO samples a group of G candidates $\{\mathbf{x}^{(g)}\}_{g=1}^{G}$ from a frozen behavior policy $p_{\theta_{\text{old}}}$. Each candidate receives a composite reward

$$R(\mathbf{x}; c) = R_{\text{prop}}(\mathbf{x}; \mathbf{y}) + \lambda_u R_{\text{und}}(\mathbf{x}; c) + R_{\text{rep}}(\mathbf{x}).$$

The exact-property term measures normalized deviations from the requested targets.

$$R_{\text{prop}} = -\sum_{k \in \mathcal{P}} w_k \frac{|f_k(\mathbf{x}) - y_k|}{\tau_k},$$

where $\mathcal{P} = \{\text{charge}, \text{GRAVY}, \text{moment}\}$, f_k is a deterministic property calculator, and τ_k is the tolerance for property k. The repetition term penalizes excessive single-residue dominance and adjacent repeats. These explicit terms enforce measurable constraints, but they cannot fully characterize whether a sequence is globally compatible with antimicrobial function.

To supply this missing semantic signal, we reuse the understanding component of AMP-UGen to evaluate a generated candidate. Instead of treating the evaluator as a separate binary classifier, we score whether the candidate supports an affirmative AMP-related response under the same instruction-following interface.

$$S_\phi(\mathbf{x}, q, \mathbf{a}) = \frac{1}{|\mathbf{a}|} \sum_{j=1}^{|\mathbf{a}|} \log p_\phi(a_j \mid a_{<j}, \mathbf{x}, q),$$

$$R_{\text{und}}(\mathbf{x}; c) = S_\phi(\mathbf{x}, q_c, \mathbf{a}^+) - S_\phi(\mathbf{x}, q_c, \mathbf{a}^-),$$

where q_c is the AMP-function query associated with condition c, and $\mathbf{a}^+$ and $\mathbf{a}^-$ denote affirmative and negative textual responses. The same understanding model therefore participates on both sides of generation. It constructs the design semantic states $\mathbf{H}_D$ through PDQs before generation and

interprets the resulting peptide afterward. Compared with an independently trained discriminator, this reuse places condition interpretation and semantic feedback in a shared representation space, reducing potential mismatch between what the instruction expresses and what the optimizer rewards. Importantly, the understanding score complements rather than replaces the exact physicochemical constraints.

Clipped pseudo-likelihood surrogate for masked diffusion. UFPO normalizes rewards within the candidates generated for the same condition,

$$A^{(g)} = \frac{R^{(g)} - \mu_R}{\sigma_R + \epsilon},$$

which makes optimization depend on relative candidate quality under a common design target. Because a masked diffusion model does not expose an autoregressive sequence likelihood, we construct a position-wise masked pseudo-log-likelihood,

$$\ell_\theta(\mathbf{x}, c) = \frac{1}{L} \sum_{i=1}^{L} \log p_\theta \left(x_i \mid \mathbf{x}_{\setminus i}^{[\mathrm{MASK}]}, \mathbf{H}_D \right).$$

Each valid residue is masked in turn, and its log-probability is averaged over valid positions at a fixed evaluation timestep. Applying the same construction to the current and frozen behavior denoisers gives the sequence-level surrogate ratio

$$\rho^{(g)} = \exp\Big(\ell_\theta(\mathbf{x}^{(g)}, c) - \ell_{\theta_{\mathrm{old}}}(\mathbf{x}^{(g)}, c) \Big).$$

We optimize the clipped UFPO surrogate

$$\mathcal{L}_{\mathrm{policy}} = -\frac{1}{G} \sum_{g=1}^{G} \min\big(\rho^{(g)} A^{(g)}, \; \mathrm{clip}(\rho^{(g)}, 1-\epsilon_c, 1+\epsilon_c) A^{(g)} \big).$$

To limit abrupt policy changes, we add a ratio-based drift penalty

$$\mathcal{L}_{\mathrm{drift}} = \frac{1}{G} \sum_{g=1}^{G} (\rho^{(g)} - 1)^2.$$

The final optimization objective is

$$\mathcal{L}_{\mathrm{UFPO}} = \mathcal{L}_{\mathrm{policy}} + \beta \mathcal{L}_{\mathrm{drift}}.$$

This local pseudo-likelihood ratio is not an importance ratio for the full diffusion trajectory; clipping and drift regularization serve as practical update stabilizers. During UFPO, the understanding component remains fixed while the masked diffusion generator is updated, preventing evaluator co-adaptation.

4 Experiments

4.1 Experimental Setup

Data construction. We construct AMP-UGen training data and held-out understanding evaluation data from ESCAPE, a standardized multilabel AMP benchmark integrating more than 80,000 peptides from 27 validated repositories (Ojeda et al., 2025). ESCAPE harmonizes functional annotations into antibacterial, antifungal, antiviral, antiparasitic, and general antimicrobial labels, and builds non-AMP examples from filtered unrelated protein sources. We follow the official ESCAPE data partition: Fold 1 and Fold 2 are combined for training (65,870 sequences), while the Test partition is used for held-out evaluation (16,489 sequences). The training and test partitions contain 17,174 and 4,235 AMP-positive sequences, respectively. For each sequence, we deterministically calculate length, net charge, Kyte–Doolittle GRAVY, and the Eisenberg hydrophobic moment at a 100° residue angle. We then convert the labeled sequences into instruction–response pairs for functional classification, physicochemical regression, sequence description, and conditional generation. The resulting instruction training set contains 80,276 examples. UFPO uses 9,966 AMP-derived design instructions with target length, charge, GRAVY, and hydrophobic moment. Their lengths range from 5 to 20 residues.

Table 2: Conditional-control and understanding-feedback comparisons. In Panel (a), MAEs are averaged over three random-seed runs per target and reported as mean ± standard deviation across the four primary targets; success uses charge/moment tolerances of $2.0/0.15$ and reports mean ± standard deviation across four equal-sized candidate subsets. Panel (b) reports exact candidate-level results pooled over the balanced and cationic settings shown in Figure 6.

(a) Conditional control and candidate-level success

Method	Charge MAE ↓	Moment MAE ↓	r_Q ↑	r_M ↑	Charge Success ↑	Moment Success ↑	Joint Success ↑
Base AMP-UGen	1.183 ± 1.084	0.0273 ± 0.0253	0.499	-0.300	26.00 ± 4.32	66.25 ± 5.12	19.50 ± 2.89
UFPO	0.311 ± 0.071	0.0163 ± 0.0095	**0.988**	**0.833**	61.75 ± 4.57	71.00 ± 5.16	43.75 ± 5.74

(b) Understanding-feedback ablation

Method	Charge Success ↑	Moment Success ↑	Joint Success ↑	HydrAMP ↑	HydrMIC ↑	ESC-AB ↑	AMPlify ↑
Base AMP-UGen	6.50	47.00	5.75	**74.06**	**56.72**	**72.32**	**82.34**
UFPO + understanding	**73.75**	73.25	**54.75**	65.52	37.35	**72.41**	71.45
UFPO w/o understanding	12.50	**74.00**	10.00	47.22	20.86	56.91	50.75

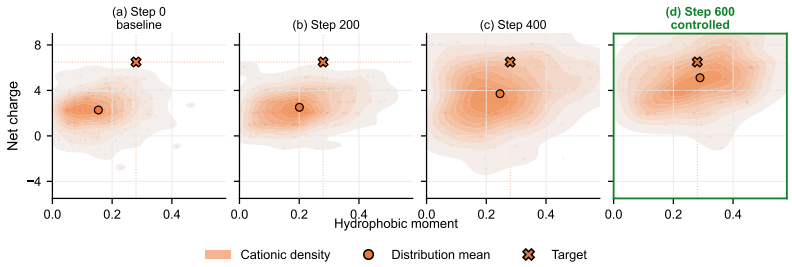

Figure 3: Conditional-control evolution for the representative cationic target from the base generator to UFPO step 600. Circles mark distribution means; crosses and dotted lines mark targets.

Baselines and metrics. For AMP understanding, we compare AMP-SFT, obtained by continuing supervised fine-tuning from Prot2Text-V2 on ESCAPE-derived AMP instructions, against LLaMA-8B and the unadapted Prot2Text-V2 checkpoint. For generation, we compare AMP-UGen with AMP-Diffusion (Torres et al., 2025), Diff-AMP (Wang et al., 2024a), HydrAMP (Szymczak et al., 2023), and AMP-SFT generator, an autoregressive baseline obtained by further fine-tuning AMP-SFT on peptide-generation data. The *w/o condition* variant retains trainable PDQs but omits design instructions, whereas *frozen PDQs* retains instructions but fixes the PDQ embeddings. Each generation method is evaluated over five independent runs with different random seeds. Each run generates 2,000 candidates, yielding 10,000 candidates in total per method. Generation-quality metrics are computed separately for each run and then averaged across the five runs. The same amino-acid validity checks are applied to all methods. Uniqueness measures non-duplication, and novelty is computed against the ESCAPE training sequences. Predicted-positive rates use the frozen HydrAMP, HydrMIC, AMPlify, and ESCAPE-AB evaluators (Szymczak et al., 2023; Li et al., 2022; Ojeda et al., 2025); HemoPI2 supplies predicted non-hemolytic rates (Rathore et al., 2025). These evaluators provide complementary rather than fully independent evidence: ESC-AB supplies an ESCAPE in-benchmark consistency check, the other activity predictors provide cross-model assessments, and HemoPI2 adds a safety-oriented endpoint. Their architectures span convolutional–recurrent, attentive recurrent, dual-branch Transformer, and random-forest models; full provenance is provided in Appendix B.4. All remain fixed and are used only post hoc, not in UDT training or UFPO rewards. Fitness measures amphipathic sequence quality.

Conditional-control protocol. We evaluate conditional control on four primary target configurations, balanced, cationic, low-charge, and low-moment. High-charge and high-moment are treated as stress cases. Each method–condition pair is generated three times with different random seeds and the same candidate budget. Let Q^* and M^* denote the requested net charge and hydrophobic moment. Condition-level errors $|\bar{Q} - Q^*|$ and $|\bar{M} - M^*|$ are first averaged over runs and then summarized across the four target configurations. We also report r_Q and r_M, the Pearson correlations between requested values and condition-level generated means. Candidate-level success is evalu-

ated on the balanced and cationic pools under tolerances of 2.0 for charge and 0.15 for moment, and reports mean $\pm$ within-pool standard deviation across four equal-sized candidate subsets.

4.2 Peptide Understanding and Generation Quality

Table 1 reports the understanding capabilities that supply AMP-oriented semantics and the resulting generation quality. Panel (a) summarizes functional and physicochemical understanding on the held-out ESCAPE test set, while Panel (b) compares generated candidate pools.

Understanding quality. Fine-tuning the original Prot2Text-V2 protein-understanding checkpoint raises macro-F1 from 0.2029 to 0.5923 in AMP-SFT; LLaMA-3.1-8B-Instruct obtains 0.1997. AMP-SFT also yields the lowest charge, GRAVY, and hydrophobic-moment MAEs, supporting AMP-specific adaptation of general protein understanding.

Computational candidate evaluation. Table 1 shows that AMP-UGen generates valid and novel candidates while maintaining strong predicted antimicrobial quality. Under exact-match criteria, its candidates remain fully unique and novel. Among the compared generation methods, AMP-UGen obtains the best HydrMIC and AMPlify positive rates, and also improves ESC-AB over the baselines. Its predicted non-hemolytic rate reaches 61.98%, ranking second to AMP-Diffusion (64.68%) and slightly exceeding AMP-SFT generator (61.46%). The frozen-PDQ variant reaches 56.80% but substantially underperforms the full model on HydrMIC and AMPlify. HydrAMP gives a different ranking, where HydrAMP itself remains the strongest method under its own predictor, and AMP-UGen obtains the second-highest fitness score. These differences expose trade-offs that would be hidden by a single aggregate score. Agreement across heterogeneous frozen efficacy evaluators, together with the HemoPI2 non-hemolysis screen, provides a broader computational assessment than any individual predictor.

4.3 Conditional Control and Understanding Feedback

Role of Understanding-Derived Conditions and Trainable PDQs. Removing the design-condition input while retaining trainable PDQs preserves uniqueness and novelty, but reduces HydrAMP from 74.06% to 56.45%, HydrMIC from 56.72% to 30.60%, ESC-AB from 72.32% to 68.57%, and AMPlify from 82.34% to 56.00%; fitness also decreases from 0.0917 to 0.0614. Its pooled predicted non-hemolytic rate is 54.11%. Because this variant retains the query pathway and its trainable parameters, the degradation isolates the contribution of condition-dependent semantic input rather than simply removing the conditioning interface. Complementarily, freezing the PDQ query embeddings while retaining condition input reduces HydrMIC to 17.15% and AMPlify to 44.23%. Together, the two controls indicate that both understanding-derived conditions and adaptation of the query interface contribute to transferring useful design semantics to the generator.

We evaluate whether understanding feedback improves property-directed control after AMP-UGen has learned the base generation distribution. Table 2 shows that UFPO substantially improves property alignment, increasing Charge, Moment, and Joint Success from 26.00%, 66.25%, and 19.50% to 61.75%, 71.00%, and 43.75%, respectively. In the direct ablation, removing understanding feedback markedly reduces charge and joint success as well as all four frozen-predictor positive rates. Together with the distributional shifts in Figure 3, these results indicate that understanding feedback improves both target satisfaction and predicted antimicrobial quality.

5 Conclusion

AMP-UGen promotes generation by transferring understanding-derived design semantics to masked diffusion through UDT and reusing the understanding component as UFPO feedback. It achieves the highest HydrMIC and AMPlify rates together with the second-highest predicted non-hemolytic rate among the evaluated generation models. UFPO improves property alignment, while its understanding feedback increases joint constraint satisfaction and consistently raises the positive rates of all four frozen activity predictors relative to UFPO without understanding. These results support peptide understanding as an active signal for both sequence construction and generation-policy optimization.

AI USE STATEMENT

Generative AI tools were used to assist with English-language editing, manuscript organization, LaTeX formatting, and the presentation of author-provided experimental results. They were not used to generate experimental data, conduct biological experiments, or autonomously determine the scientific conclusions of this work. The authors reviewed and revised all AI-assisted content and take full responsibility for the final text, equations, tables, citations, claims, and presentation.

ETHICS STATEMENT

This study is computational and uses publicly available peptide datasets and released prediction models. It does not involve human participants, animal subjects, private information, or personally identifiable data. Because antimicrobial-peptide design may have dual-use implications, the generated sequences are reported only as computational candidates and should not be interpreted as experimentally validated therapeutics. Any future synthesis or biological testing should follow institutional biosafety procedures and include appropriate antimicrobial-activity, cytotoxicity, hemolysis, stability, and off-target assessments. We also acknowledge that biases and coverage limitations in public AMP datasets and computational predictors can affect which sequence families are favored by the model.

REPRODUCIBILITY STATEMENT

The main paper and appendix document the ESCAPE data partitions and sample counts, instruction construction, deterministic physicochemical calculations, model objectives, candidate budgets, random-seed protocol, success thresholds, and evaluation metrics used in this study. The appendix additionally provides the training stages and hyperparameters, formulas for sequence- and property-level metrics, and the architectures, development data, and decision protocols of the frozen evaluators. All external datasets, baseline methods, and released predictors are identified through citations. These details are intended to support independent reconstruction of the training and evaluation procedures and to distinguish exact calculations from predictor-based in-silico estimates.

REFERENCES

Christof Angermueller, David Dohan, David Belanger, Ramya Deshpande, Kevin Murphy, and Lucy J. Colwell. Model-based reinforcement learning for biological sequence design. In *International Conference on Learning Representations*, 2020.

Junhang Cao, Jun Zhang, Qiyuan Yu, Junkai Ji, Jianqiang Li, Shan He, and Zexuan Zhu. TG-CDDPM: Text-guided antimicrobial peptides generation based on conditional denoising diffusion probabilistic model. *Briefings in Bioinformatics*, 26(1):bbae644, 2025. doi: 10.1093/bib/bbae644.

Bo Chen, Xingyi Cheng, Pan Li, Yangli-ao Geng, Jing Gong, Shen Li, Zhilei Bei, Xu Tan, Boyan Wang, Xin Zeng, Chiming Liu, Aohan Zeng, Yuxiao Dong, Jie Tang, and Le Song. xTrimoPGLM: Unified 100b-scale pre-trained transformer for deciphering the language of protein. *arXiv preprint arXiv:2401.06199*, 2024. URL https://arxiv.org/abs/2401.06199.

Payel Das, Tom Sercu, Kahini Wadhawan, Inkit Padhi, Sebastian Gehrmann, Flaviu Cipcigan, Vijil Chenthamarakshan, Hendrik Strobelt, Cicero Dos Santos, Pin-Yu Chen, Yi Yan Yang, Jeremy P. K. Tan, James Hedrick, Jason Crain, and Aleksandra Mojsilovic. Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. *Nature Biomedical Engineering*, 5(6):613–623, 2021. doi: 10.1038/s41551-021-00689-x.

Scott N. Dean, Jerome Anthony E. Alvarez, Dan Zabetakis, Scott A. Walper, and Anthony P. Malanoski. PepVAE: Variational autoencoder framework for antimicrobial peptide generation and activity prediction. *Frontiers in Microbiology*, 12:725727, 2021. doi: 10.3389/fmicb.2021.725727.

Xiao Fei, Michail Chatzianastasis, Sarah Carneiro, Hadi Abdine, Lawrence Petalidis, and Michalis Vazirgiannis. Prot2Text-V2: Protein function prediction with multimodal contrastive alignment. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Noelia Ferruz, Steffen Schmidt, and Birte Höcker. ProtGPT2 is a deep unsupervised language model for protein design. *Nature Communications*, 13:4348, 2022. doi: 10.1038/s41467-022-32007-7.

Adam Izdebski, Jan Olszewski, Pankhil Gawade, Krzysztof Koras, Serra Korkmaz, Valentin Rauscher, Jakub M. Tomczak, and Ewa Szczurek. Unified molecule generation and property prediction. *arXiv preprint arXiv:2504.16559*, 2025. URL https://arxiv.org/abs/2504.16559.

Shuwen Jin, Zihan Zeng, Xiyan Xiong, Baicheng Huang, Li Tang, Hongsheng Wang, Xiao Ma, Xiaochun Tang, Guoqing Shao, Xingxu Huang, et al. AMPGen: An evolutionary information-reserved and diffusion-driven generative model for de novo design of antimicrobial peptides. *Communications Biology*, 8:839, 2025. doi: 10.1038/s42003-025-08282-7.

Catherine Li, Dan Sutherland, Spencer A. Hammond, Chen Yang, Fabrice Taho, Nicholas H. Bergman, Steven Houston, René L. Warren, Andrew Wong, Linda M. N. Hoang, and Fiona S. L. Brinkman. AMPlify: Attentive deep learning model for discovery of novel antimicrobial peptides effective against WHO priority pathogens. *BMC Genomics*, 23:77, 2022. doi: 10.1186/s12864-022-08310-4.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023. doi: 10.1126/science.ade2574.

Shengchao Liu, Yanjing Li, Zhuoxinran Li, Anthony Gitter, Yutao Zhu, Jiarui Lu, Zhao Xu, Weili Nie, Arvind Ramanathan, Chaowei Xiao, Jian Tang, Hongyu Guo, and Anima Anandkumar. A text-guided protein design framework. *Nature Machine Intelligence*, 7:580–591, 2025a. doi: 10.1038/s42256-025-01011-z.

Xiangyu Liu, Yi Liu, Silei Chen, and Wei Hu. Controllable protein sequence generation with LLM preference optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 505–513, 2025b. doi: 10.1609/aaai.v39i1.32030.

Ali Madani, Ben Krause, Eric R. Greene, Subu Subramanian, Benjamin P. Mohr, James M. Holton, Jr. Olmos, Jose Luis, Caiming Xiong, Zachary Z. Sun, Richard Socher, James S. Fraser, and Nikhil Naik. Large language models generate functional protein sequences across diverse families. *Nature Biotechnology*, 41(8):1099–1106, 2023. doi: 10.1038/s41587-022-01618-2.

Sebastian Ojeda, Rafael Velasquez, Nicolás Aparicio, Juanita Puentes, Paula Cárdenas, Nicolás Andrade, Gabriel González, Sergio Rincón, Carolina Muñoz-Camargo, and Pablo Arbeláez. A standardized benchmark for multilabel antimicrobial peptide classification. In *Advances in Neural Information Processing Systems*, volume 38, 2025. URL https://papers.nips.cc/paper_files/paper/2025/hash/bbd28018d38fc0a51021139206af9542-Abstract-Datasets_and_Benchmarks_Track.html.

Nelson G. Oliveira Júnior, Camila M. Souza, Danieli F. Buccini, Marlon H. Cardoso, and Octávio L. Franco. Antimicrobial peptides: Structure, functions and translational applications. *Nature Reviews Microbiology*, 23:687–700, 2025. doi: 10.1038/s41579-025-01200-y.

Anand Singh Rathore, Nishant Kumar, Shubham Choudhury, Naman Kumar Mehta, and Gajendra P. S. Raghava. Prediction of hemolytic peptides and their hemolytic concentration. *Communications Biology*, 8:176, 2025. doi: 10.1038/s42003-025-07615-w.

Célio D. Santos-Júnior, Shaojun Pan, Xing-Ming Zhao, and Luis Pedro Coelho. Macrel: Antimicrobial peptide screening in genomes and metagenomes. *PeerJ*, 8:e10555, 2020. doi: 10.7717/peerj.10555.

Jithendaraa Subramanian, Shivakanth Sujit, Niloy Irtisam, Umong Sain, Riashat Islam, Derek Nowrouzezahrai, and Samira Ebrahimi Kahou. Reinforcement learning for sequence design leveraging protein language models. *arXiv preprint arXiv:2407.03154*, 2024. URL https://arxiv.org/abs/2407.03154.

Paulina Szymczak, Marcin Możejko, Tomasz Grzegorzek, Radosław Jurczak, Marta Bauer, Damian Neubauer, Karol Sikora, Michał Michalski, Jacek Sroka, Piotr Setny, Wojciech Kamysz, and Ewa Szczurek. Discovering highly potent antimicrobial peptides with deep generative model HydrAMP. *Nature Communications*, 14:1453, 2023. doi: 10.1038/s41467-023-36994-z.

Marcelo D. T. Torres, Tianlai Chen, Fangping Wan, Pranam Chatterjee, and Cesar de la Fuente-Nunez. Generative latent diffusion language modeling yields anti-infective synthetic peptides. *Cell Biomaterials*, 1(9):100183, 2025. doi: 10.1016/j.celbio.2025.100183.

Colin M. Van Oort, Jonathon B. Ferrell, Jacob M. Remington, Safwan Wshah, and Jianing Li. AMPGAN v2: Machine learning-guided design of antimicrobial peptides. *Journal of Chemical Information and Modeling*, 61(5):2198–2207, 2021. doi: 10.1021/acs.jcim.0c01441.

Daniel Veltri, Uday Kamath, and Amarda Shehu. Deep learning improves antimicrobial peptide recognition. *Bioinformatics*, 34(16):2740–2747, 2018. doi: 10.1093/bioinformatics/bty179.

Jike Wang, Jianwen Feng, Yu Kang, Peichen Pan, Jingxuan Ge, Yan Wang, Mingyang Wang, Zhenxing Wu, Xingcai Zhang, Jiameng Yu, Xujun Zhang, Tianyue Wang, Lirong Wen, Guangning Yan, Yafeng Deng, Hui Shi, Chang-Yu Hsieh, Zhihui Jiang, and Tingjun Hou. Discovery of antimicrobial peptides with notable antibacterial potency by an LLM-based foundation model. *Science Advances*, 11(10):eads8932, 2025a. doi: 10.1126/sciadv.ads8932.

Rui Wang, Tao Wang, Linlin Zhuo, Jinhang Wei, Xiangzheng Fu, Quan Zou, and Xiaojun Yao. Diff-AMP: Tailored designed antimicrobial peptide framework with all-in-one generation, identification, prediction and optimization. *Briefings in Bioinformatics*, 25(2):bbae078, 2024a. doi: 10.1093/bib/bbae078.

Xinyou Wang, Zaixiang Zheng, Fei Ye, Dongyu Xue, Shujian Huang, and Quanquan Gu. Diffusion language models are versatile protein learners. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 52309–52333, 2024b.

Yeji Wang, Minghui Song, Fujing Liu, Zhen Liang, Rui Hong, Yuemei Dong, Huaizu Luan, Xiaojie Fu, Wenchang Yuan, Wenjie Fang, Gang Li, Hongxiang Lou, and Wenqiang Chang. Artificial intelligence using a latent diffusion model enables the generation of diverse and potent antimicrobial peptides. *Science Advances*, 11(6):eadp7171, 2025b. doi: 10.1126/sciadv.adp7171.

Zeyuan Wang, Qiang Zhang, Keyan Ding, Ming Qin, Xiang Zhuang, Xiaotong Li, and Huajun Chen. InstructProtein: Aligning human and protein language via knowledge instruction. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pp. 1114–1136, 2024c. doi: 10.18653/v1/2024.acl-long.62.

Minghao Xu, Xinyu Yuan, Santiago Miret, and Jian Tang. ProtST: Multi-modality learning of protein sequences and biomedical texts. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 38749–38767, 2023.

Jielu Yan, Pratiti Bhadra, Ang Li, Pooja Sethiya, Longguang Qin, Hio Kuan Tai, Koon Ho Wong, and Shirley W. I. Siu. Deep-AmPEP30: Improve short antimicrobial peptides prediction with deep learning. *Molecular Therapy–Nucleic Acids*, 20:882–894, 2020. doi: 10.1016/j.omtn.2020.05.006.

A Training and Implementation Details

AMP-UGen is trained in three stages. First, peptide-understanding adaptation continues supervised fine-tuning from Prot2Text-V2 on ESCAPE-derived functional, physicochemical, description, and generation instructions using LoRA. Second, UDT trains the peptide design queries (PDQs) and masked diffusion denoiser to convert understanding-derived design states into peptide sequences. Third, UFPO updates the masked diffusion generator using grouped candidates, exact-property rewards, and feedback from the fixed understanding component. Table 3 lists the settings used in these trainable stages.

Table 3: Training settings used for peptide-understanding adaptation, UDT-based generation training, and UFPO optimization. Micro batch denotes the per-device batch size before gradient accumulation.

Stage	Optimization	Batch and duration	Stage-specific settings
Peptide-understanding adaptation	Learning rate 1×10^{-5}	Micro batch 1; gradient accumulation 16; 5 epochs	LoRA rank 16
UDT generation training	Denoiser learning rate 1×10^{-4}; PDQ learning rate 1×10^{-5}	Micro batch 16; gradient accumulation 4; 80 epochs	50 diffusion timesteps; 32 PDQs
UFPO optimization	Learning rate 5×10^{-6}; clip threshold 0.2; ratio-drift coefficient 0.03	4 candidates per instruction; 50 generation steps	Reparameterized masked decoding with Gumbel argmax; reward weights: charge 0.7, GRAVY 1.0, moment 1.0, understanding 0.8

For the no-understanding UFPO ablation, the understanding-feedback weight is set from 0.8 to 0 while all other UFPO settings remain unchanged. The sequence-level likelihood ratio used for clipping is the masked pseudo-likelihood surrogate defined in Section 3.3, not a full-trajectory importance ratio for the multi-step diffusion process.

B Evaluation Protocol and Supporting Results

This appendix specifies the metrics used for computational candidate evaluation. All predictor-based quantities are in-silico estimates and are not substitutes for antibacterial, MIC, cytotoxicity, or hemolysis assays.

B.1 Candidate-level sequence metrics

Let $\mathcal{G}$ be the generated candidate multiset after amino-acid validity checking, $N = |\mathcal{G}|, \mathcal{U} = \mathrm{set}(\mathcal{G})$ its set of distinct sequences, and $\mathcal{T}$ the ESCAPE training-sequence set. We compute

$$\mathrm{Uniqueness} = \frac{|\mathcal{U}|}{N}, \qquad \mathrm{Novelty} = \frac{|\{\mathbf{x} \in \mathcal{U} : \mathbf{x} \notin \mathcal{T}\}|}{|\mathcal{U}|}.$$

Novelty therefore measures exact sequence non-overlap rather than scaffold- or homology-level novelty.

For a frozen binary evaluator e with score $s_e(\mathbf{x})$ and its fixed decision threshold τ_e, the reported positive rate is

$$\mathrm{PosRate}_e = \frac{1}{N}\sum_{\mathbf{x}\in\mathcal{G}} \mathbb{I}[s_e(\mathbf{x}) \geq \tau_e].$$

The same definition is used for the HemoPI2 non-hemolytic rate after reversing the positive event to "Non-Hemolytic." No evaluator in Table 1 is updated on generated candidates or used as a UDT training target or UFPO reward.

B.2 PHYSICOCHEMICAL DESCRIPTORS

For a peptide $\mathbf{x} = (x_1, \ldots, x_L)$, net charge follows the deterministic implementation

$$Q(\mathbf{x}) = n_K + n_R + 0.1n_H - n_D - n_E + 1 - \mathbb{I}_{\text{nonamidated}},$$

where n_a is the count of amino acid a. The last two terms are the N- and C-terminal contributions and cancel under the default non-amidated setting. Hydrophobicity is the Kyte–Doolittle grand average of hydropathy,

$$\text{GRAVY}(\mathbf{x}) = \frac{1}{L}\sum_{i=1}^{L} h_{\text{KD}}(x_i),$$

and hydrophobic moment uses the Eisenberg consensus scale with $\theta = 100°$,

$$M(\mathbf{x}) = \frac{1}{L}\sqrt{\left(\sum_{i=1}^{L} h_{\text{E}}(x_i)\cos((i-1)\theta)\right)^2 + \left(\sum_{i=1}^{L} h_{\text{E}}(x_i)\sin((i-1)\theta)\right)^2}.$$

The fitness score in Table 1 measures an amphipathic vector magnitude normalized by an exponential residue penalty,

$$\text{Fitness}(\mathbf{x}) = \frac{\sqrt{\left(\sum_{i=1}^{L} h_F(x_i)\cos(i\theta)\right)^2 + \left(\sum_{i=1}^{L} h_F(x_i)\sin(i\theta)\right)^2}}{\sum_{i=1}^{L}\exp(h_X(x_i))}, \qquad \theta = 100°,$$

where h_F and h_X are the residue scales used by the released evaluation implementation. Higher values favor a strong amphipathic arrangement while penalizing unfavorable residue contributions; the score is a sequence-level computational proxy rather than an experimental fitness measurement.

B.3 CONDITIONAL-CONTROL METRICS

For requested charge and hydrophobic-moment targets Q^* and M^*, and generated condition-level means $\bar{Q}$ and $\bar{M}$, we define

$$\text{MAE}_Q = |\bar{Q} - Q^*|, \qquad \text{MAE}_M = |\bar{M} - M^*|.$$

Charge Success and Moment Success are the fractions of candidates satisfying $|Q(\mathbf{x}) - Q^*| \leq 2.0$ and $|M(\mathbf{x}) - M^*| \leq 0.15$, respectively; Joint Success requires both inequalities. The correlations r_Q and r_M are Pearson correlations between requested values and generated condition-level means.

B.4 FROZEN PREDICTORS: ARCHITECTURES, DEVELOPMENT DATA, AND USE

Table 4 summarizes the provenance of the frozen evaluators. The predictors differ in target definition, model family, and corpus construction, so agreement across them is less dependent on one particular decision boundary. This heterogeneity should not, however, be interpreted as complete statistical independence: HydrAMP and HydrMIC originate from the same study, and ESC-AB is developed on the ESCAPE benchmark that also supplies AMP-UGen's domain-adaptation data.

The ESCAPE connection makes ESC-AB especially relevant as an in-benchmark consistency check: it tests whether candidates remain compatible with the functional labeling framework used for AMP-domain adaptation. It does not, by itself, establish out-of-distribution generalization, which is why the differently constructed HydrAMP, HydrMIC, and AMPlify predictors are reported alongside it.

B.5 WHY MULTIPLE FROZEN EVALUATORS ARE USED

The four activity outputs probe related but non-identical notions. HydrAMP measures generic AMP likeness, HydrMIC targets strong activity against *E. coli*, AMPlify supplies an independently developed attentive classifier, and ESC-AB targets the antibacterial label in a multilabel benchmark. HemoPI2 adds a distinct safety-oriented endpoint. Their rankings are not forced to agree: for example, HydrAMP ranks its own generator highest under the HydrAMP classifier, whereas AMP-UGen

Table 4: Frozen evaluator provenance. Dataset descriptions and architectures follow the corresponding original works. "Use here" denotes the output summarized in Table 1.

Evaluator	Prediction target	Architecture	Development data	Use here
HydrAMP (M_{AMP})	Generic antimicrobial-peptide probability	AMPScanner-style convolutional–recurrent classifier	The HydrAMP study assembled peptides of at most 25 residues from curated AMP databases and UniProt. Its balanced AMP-classifier data contain 11,131 AMP positives from dbAMP, AMP Scanner, and DRAMP and 11,131 length-matched UniProt negatives after filtering and sequence clustering (Szymczak et al., 2023).	AMP-positive rate
HydrMIC (M_{MIC})	Probability of low MIC against *E. coli* (MIC $\leq 32\ \mu$g/mL)	128-dimensional residue embedding, 64-unit LSTM, one-dimensional pooling, 100-unit LSTM, and sigmoid output	GRAMPA-derived MIC measurements restricted to peptides of at most 25 residues and tested against *E. coli*; the released development set contains 2,126 positives and 1,318 negatives under the 32 μg/mL threshold (Szymczak et al., 2023).	Low-MIC positive rate
AMPlify	Generic antimicrobial-peptide probability	Bidirectional LSTM followed by multi-head scaled dot-product attention and context attention; five-model ensemble	A nonredundant positive corpus of 4,173 mature AMP sequences from APD3 and DADP, with non-AMP sequences selected from UniProt; the released procedure constructs matched positive and negative training partitions (Li et al., 2022).	AMP-positive rate
ESC-AB	Antibacterial label within multilabel AMP classification	Dual-branch Transformer encoding sequence and structural representations with bidirectional cross-attention	ESCAPE integrates more than 80,000 peptides from 27 repositories, harmonizes five functional labels, and supplies official train/test partitions; ESC-AB denotes its antibacterial output (Ojeda et al., 2025).	Antibacterial-positive rate
HemoPI2 RF	Hemolytic versus non-hemolytic peptide	Random forest over composition and physicochemical descriptors	HemoPI2 contains 1,926 unique peptides with experimentally reported HC_{50} against mammalian red blood cells: 891 are labeled hemolytic at $\mathrm{HC}_{50} \leq 100\ \mu$M and 1,035 non-hemolytic above that threshold (Rathore et al., 2025). We use the released random-forest classifier with a fixed decision threshold of 0.46.	Non-hemolytic rate

is strongest under HydrMIC and AMPlify. Reporting the individual outputs rather than averaging them exposes this evaluator dependence.

All five predictors remain fixed and are applied only after generation. Consequently, AMP-UGen cannot improve these scores through evaluator co-adaptation during UDT or UFPO training. Results that improve across predictors with different architectures and target definitions provide stronger computational evidence than a gain under one evaluator alone. Nevertheless, shared public AMP sources, domain shift to artificially generated sequences, calibration error, and reduced reliability on unusual or very short peptides remain possible. The reported rates therefore support comparative in-silico evaluation, not claims of experimentally verified antimicrobial activity or safety.

B.6 Generated-sequence amino-acid composition

Figure 4 compares residue frequencies in all 17,174 ESCAPE training AMPs and a randomly sampled subset of 5,000 AMP-UGen used for the full-model evaluation. AMP-UGen retains broad coverage of the 20 standard residues while reweighting the composition toward Lys, Arg, His, and Trp and away from acidic residues, consistent with the cationic design targets represented in the generated pool. The base-2 Jensen–Shannon divergence between the two normalized residue distributions is 0.0670. This comparison characterizes a controlled compositional shift rather than exact reproduction of the training distribution.

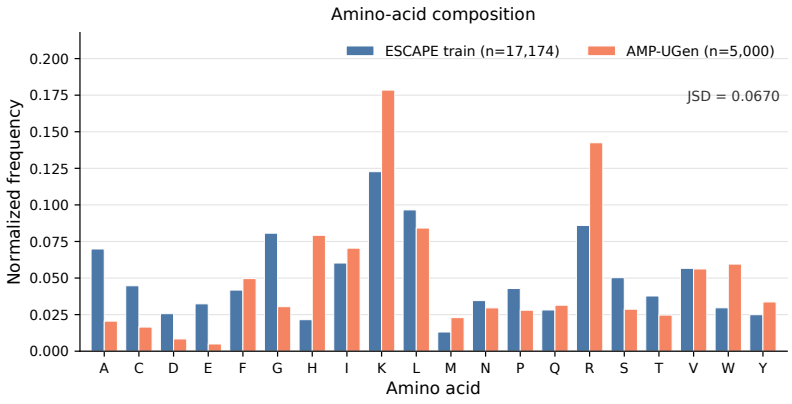

Figure 4: Normalized amino-acid composition of the ESCAPE training AMPs and AMP-UGen candidates. Frequencies are computed over all standard residues in each complete sequence pool; JSD denotes the base-2 Jensen–Shannon divergence.

C ADDITIONAL UFPO VISUALIZATIONS

Figure 5 complements the representative cationic trajectory in the main paper with the balanced target. The distribution mean moves toward the requested charge–moment region over optimization, while the density contours show that this change involves the candidate distribution rather than only a few isolated sequences.

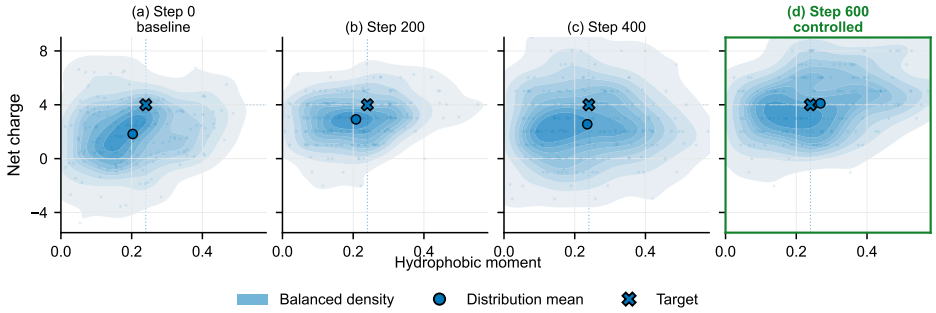

Figure 5: Representative UFPO trajectory for the balanced target from the base generator to step 600. Blue contours denote the generated-candidate density, circles denote distribution means, and crosses and dotted lines denote the requested target.

Figure 6 provides a representative distributional comparison of the base generator, UFPO with understanding feedback, and UFPO without understanding feedback. The exact-property calculator remains active in both UFPO variants, whereas the understanding-reward weight is set to zero in the no-understanding variant. The pools are illustrative rather than seed-aggregated; the quantitative predictor comparisons are reported in Table 2.

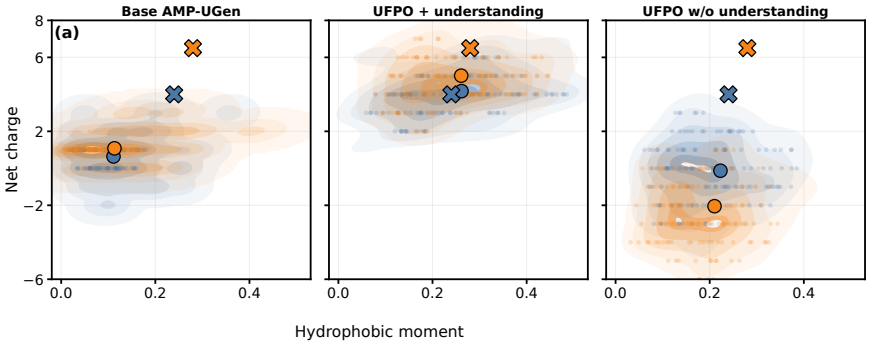

Figure 6: Representative UFPO candidate distributions with and without understanding feedback. Blue and orange contours denote the balanced and cationic requests, respectively. Only the distribution panels are shown; frozen-predictor rates are reported separately in Table 2.