SCIENCE CHINA
Information Sciences

• Supplementary File •

MafAFP: A mechanism-aware protein language model for antifungal peptide activity prediction

Han CHEN[1], Feiyi HUANG[1], Xuanbai REN[1], Hongxing XIANG[1], Zhonghui ZHANG[1], Daibo LIU[1], Xinyu YANG[1*], Xiangzheng FU[2*], Haowen CHEN[1*] & Xiangxiang ZENG[1]

[1]*College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China*
[2]*Computer Science and Control Engineering, Shenzhen University of Advanced Technology, Shenzhen 518107, China*

Appendix A Related work

Appendix A.1 Mechanism-aware antifungal peptide prediction

Fungal infections affect more than one billion people worldwide [1], yet clinical antifungal therapy remains confined to four major classes, polyenes, azoles, flucytosine, and echinocandins, whose utility is limited by toxicity and a narrow range of molecular targets [2]. Antifungal peptides (AFPs) offer an attractive alternative, typically providing lower toxicity, better biocompatibility, and strong activity against drug-resistant strains [3], which has motivated growing interest in computational AFP prediction [4].

Transformer-based protein language models (PLMs) such as ESM, TAPE, and ProtTrans [5–7] learn rich contextual representations from large unannotated sequence corpora, alleviating the data scarcity common in AFP research and outperforming traditional feature engineering in few-shot settings. Building on this, recent methods couple PLM embeddings with complementary information: AFP-MFL [8] integrates semantic, evolutionary, and physicochemical features; DeepAFP [9] combines a PLM with CNN and LSTM modules; a BiTCN-based multi-view model [10] fuses the same three feature views; DL-QSARES couples motif recombination with QSAR screening; and AMP-Designer [12] leverages a large language model to design antimicrobial peptides with desired properties.

Despite this progress, most AFP predictors still rely on post hoc fusion, concatenating PLM embeddings with physicochemical and evolutionary features before a conventional classifier, with only basic SHAP-based interpretation [13, 14]. Such pipelines tend to learn correlation-driven rules (length, charge, or composition biases) rather than the biophysical mechanisms underlying antifungal activity, and do not explicitly represent membrane-interaction processes such as membrane affinity, interfacial/transmembrane insertion propensity, and amphipathic helix formation. They are therefore difficult to interpret mechanistically and can behave unreliably under distribution shift. To address this gap, we introduce MafAFP, which integrates a PLM with a multi-attribute, biophysically motivated mechanism encoder that captures insertion propensity, amphipathic helicity, and membrane affinity across multiple sliding-window scales, and injects these cues into PLM embeddings via Feature-wise Linear Modulation (FiLM) and gated residual fusion for adaptive, mechanism-guided representation learning.

Appendix B Methods

We propose a physics-enhanced multi-attribute modulation framework, MafAFP. The overall architecture consists of three main components: (i) a Multi-Attribute Feature Extractor module, (ii) a protein language model fine-tuning module based on ESMC, and (iii) a feature modulation and fusion module. The fine-tuned ESMC model produces 1152-dimensional sequence semantic embeddings (Figure A1(a)), which are passed together with the global MAF representation into the downstream feature modulation and fusion module. This module primarily includes a FiLM component and a Gated Fusion Classifier. To facilitate effective interaction between semantic and mechanistic features, both the PLM embeddings and the global MAF vector are first projected into a shared 512-dimensional latent space. The global MAF representation then predicts FiLM modulation parameters γ and β via a small neural network to perform feature-wise modulation of the PLM embeddings. This design allows membrane-related biophysical signals to selectively emphasize or suppress specific semantic dimensions, rather than being treated as independent auxiliary features. In parallel, a gated residual fusion mechanism dynamically learns a gate weight that regulates the overall contribution of membrane-mechanism information to the backbone sequence features. This gating strategy prevents over-injection of physical priors and improves robustness under dataset heterogeneity and distribution shifts. Finally, the fused representation is fed to an MLP prediction head to produce the binary AFP activity output (Figure A1(c)).

Appendix B.1 The Multi-Attribute Feature Extractor

The Multi-Attribute Feature Extractor comprises five submodules: (i) multi-scale window-level MAF scoring; (ii) window-level insertion probability / ΔG_{if} regression; (iii) multi-scale global pooling to obtain the learned representation g_{learned}; (iv) biophysically motivated

* Corresponding author (email: xinyuyang21@hnu.edu.cn, fuxiangzheng@suat-sz.edu.cn, hwchen@hnu.edu.cn)

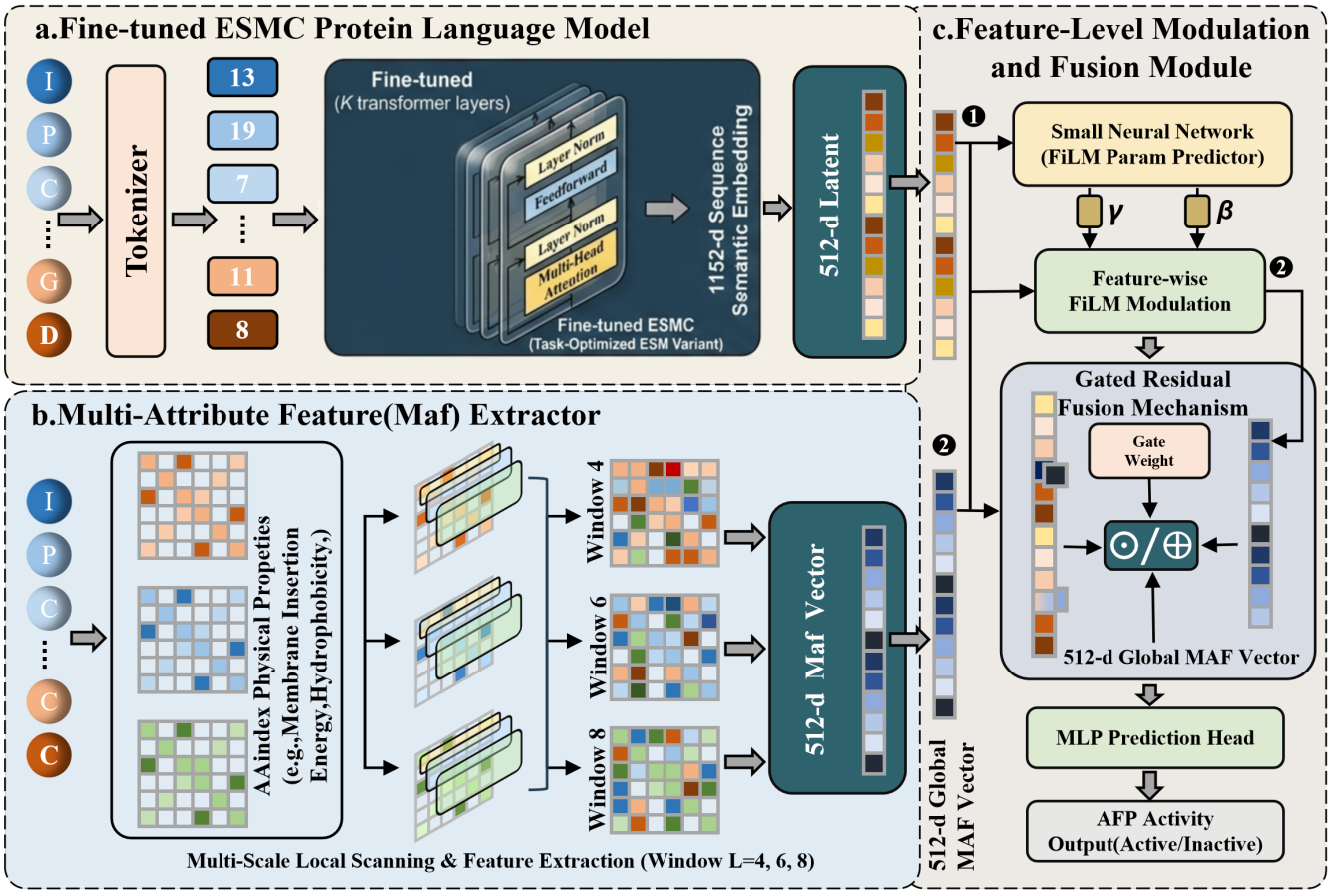

Figure A1 The MafAFP Multi-Attribute Modulation Framework. (a) The fine-tuned ESMC model. (b) The MafExtractor. (c) The feature-level modulation and fusion module.

ΔG_{if} and HM convolutions to derive the physical representation g_{phys}; and (v) an MLP-based fusion layer that integrates these into the final sequence-level representation g_{global}.

The local ΔG_{if}/HM approximation (with convolution kernels initialized as window averages over amino acid properties, i.e., each kernel element is normalized by $1/w$) is formulated as follows:

$$z_{b,s}(t) = (\text{conv}_{w_s} * X_b)(t), \tag{B1}$$

The window-level insertion probability and the ΔG_{if} proxy are defined as follows:

$$p_{b,s}(t) = \sigma\left(\omega_s^{\text{cls}} z_{b,s}(t) + b_s^{\text{cls}}\right), \quad g_{b,s}(t) = \omega_s^{\text{reg}} z_{b,s}(t) + b_s^{\text{reg}}. \tag{B2}$$

The multi-scale global MAF feature is defined as follows, where $c_{b,s}(t)$ denotes the pre-sigmoid classification logit (i.e., $c_{b,s}(t) = \sigma^{-1}(p_{b,s}(t))$):

$$c_{b,s}(t) = \omega_s^{\text{cls}} z_{b,s}(t) + b_s^{\text{cls}}, \tag{B3}$$

$$g_{\text{learned}}(b) = \text{MLP}_{\text{learn}}\left(\left\{\min_t c_{b,s}(t),\ \text{E}_t\left[c_{b,s}(t)\right],\ \text{TopK}_3\left(c_{b,s}(t)\right)\right\}_{s=1}^{S}\right). \tag{B4}$$

The biophysically motivated ΔG_{if}/HM feature is defined as follows ($\Delta G_{b,s}^{\text{WW}}$ and $\mu_{b,s}$ are produced by convolutions with kernels that are fixed or initialized from physical kernels):

$$g_{\text{phys}}(b) = \text{MLP}_{\text{phys}}\left(\left[\left\{\min_t \Delta G_{b,s}^{\text{WW}}(t),\ \max_t \mu_{b,s}(t)\right\}_{s=1}^{S},\ \text{best_pos}(b),\ \text{best_wprop}(b)\right]\right). \tag{B5}$$

The sequence-level MAF representation is defined as follows:

$$g_{\text{global}}(b) = \text{MLP}_{\text{fusion}}\left(\left[g_{\text{learned}}(b);\ g_{\text{phys}}(b)\right]\right). \tag{B6}$$

To enable the model to simultaneously learn both transmembrane-region patterns and global membrane-affinity characteristics of AFPs, we collected all transmembrane protein sequences and their annotated transmembrane segments from the TOPDB database as strong labels for transmembrane-fragment recognition. In parallel, experimentally validated AFP sequences were retrieved from DRAMP, APD3, and dbAMP as weak labels, and were used to jointly train the MafExtractor.

Algorithm B1 Multi-Attribute Feature Extraction (MafExtractor)

Require: Peptide sequence $\mathbf{x} = (x_1, x_2, \ldots, x_L)$; window sizes $\mathcal{W} = \{4, 6, 8\}$
Ensure: Global MAF representation $g_{\text{global}} \in \mathbb{R}^H$

// Part I: Learned pathway
$X \leftarrow \text{AminoAcidEmbed}(\mathbf{x})$ ▷ $X \in \mathbb{R}^{D \times L}$
for each $w \in \mathcal{W}$ **do**
  $z_w \leftarrow \text{Conv1D}(X;\ K_w)$ ▷ $z_w \in \mathbb{R}^{L-w+1}$, trainable kernel
  for each position $t = 1, \ldots, L - w + 1$ **do**
    $c_w(t) \leftarrow \omega_w^{\text{cls}} \cdot z_w(t) + b_w^{\text{cls}}$
    $p_w(t) \leftarrow \sigma(c_w(t))$ ▷ Insertion probability
    $\hat{g}_w(t) \leftarrow \omega_w^{\text{reg}} \cdot z_w(t) + b_w^{\text{reg}}$ ▷ ΔG_{if} proxy
  end for
  $f_w \leftarrow [\min_t c_w(t),\ \text{mean}_t\, c_w(t),\ \text{TopK}_3(c_w(t))]$
end for
$g_{\text{learned}} \leftarrow \text{MLP}_{\text{learn}}(\text{Concat}(f_4, f_6, f_8))$

Algorithm B2 Multi-Attribute Feature Extraction (MafExtractor) — continued

// Part II: Biophysical pathway
$X_{\text{oh}} \leftarrow \text{OneHot}_{20}(\mathbf{x})$ ▷ $X_{\text{oh}} \in \mathbb{R}^{20 \times L}$
for each $w \in \mathcal{W}$ **do**
  for each position t **do**
    $\Delta G_w^{\text{WW}}(t) \leftarrow \frac{1}{w} \sum_{j=0}^{w-1} K_{\text{WW}}(j) \cdot X_{\text{oh}}(t+j)$ ▷ Fixed WW kernel
    $\text{HM}_w(t) \leftarrow \frac{1}{w} \sqrt{\left(\sum_{j=0}^{w-1} H_j \cos(j\delta) \cdot X_{\text{oh}}(t+j)\right)^2 + \left(\sum_{j=0}^{w-1} H_j \sin(j\delta) \cdot X_{\text{oh}}(t+j)\right)^2}$
  end for
end for
best_val $\leftarrow +\infty$; best_pos $\leftarrow 0$; best_wprop $\leftarrow 0$
for each $w \in \mathcal{W}$ **do**
  $t^* \leftarrow \arg\min_t \Delta G_w^{\text{WW}}(t)$
  if $\Delta G_w^{\text{WW}}(t^*) <$ best_val **then**
    best_val $\leftarrow \Delta G_w^{\text{WW}}(t^*)$
    best_pos $\leftarrow (t^* + (w-1)/2)\ /\ (L-1)$
    best_wprop $\leftarrow w\ /\ \max(\mathcal{W})$
  end if
end for
$g_{\text{phys}} \leftarrow \text{MLP}_{\text{phys}}\Big([\{\min_t \Delta G_w^{\text{WW}}(t),\ \max_t \text{HM}_w(t)\}_{w \in \mathcal{W}},\ \text{best_pos},\ \text{best_wprop}]\Big)$
// Part III: Fusion
$g_{\text{global}} \leftarrow \text{MLP}_{\text{fusion}}(\text{Concat}(g_{\text{learned}},\ g_{\text{phys}}))$
return g_{global}

Appendix B.2 The protein language model fine-tuning module based on ESMC

We adopt the ESMC-600M protein language model which is specifically designed for sequence-level embedding and therefore aligns well with the needs of our framework. ESMC-600M comprises 30 transformer layers. To better adapt its output embeddings to the AFP recognition task, we employed a gradual unfreezing strategy on the last several transformer blocks. After extensive empirical evaluation, we fixed the number of trainable layers to two and jointly fine-tuned these layers during training of the main task, thereby avoiding a separate pre-finetuning stage and reducing additional computational cost while maintaining competitive performance. To achieve more refined learning rate scheduling during joint fine-tuning, we introduce a custom function that allows flexible adjustment of learning rates across different layers of the model. The learning rate for Transformer layers decays progressively with the layer index, with the base learning rate applied to the final layer. This layer-wise decay enables different learning rates at various stages of training. Additionally, the classification head's learning rate is scaled by a factor, relative to the base learning rate, and is not subject to weight decay.

Appendix B.3 The feature modulation and fusion module

We feed the sequence-level representation h_{seq} together with the global MAF feature m into the feature fusion module. First, both h_{seq} and m are projected into a shared latent space to obtain the PLM query q and the MAF representation. A FiLM modulation network then predicts dimension-wise scaling and bias parameters (γ, β) from m, and applies a residual affine transformation to the PLM query:

$$h' = (1 + \gamma) \odot q + \beta, \tag{B7}$$

where $\gamma = \tanh(\tilde{\gamma}) \cdot s$ with s being a learnable scale parameter, and q denotes the projected PLM query vector. The residual form $(1 + \gamma)$ ensures identity mapping when γ=0, thereby explicitly injecting membrane-affinity-related physicochemical attributes into the representation while maintaining training stability. Finally, we employ a temperature-scaled sigmoid gate g_{maf} together with a learnable global weight α_{maf} to perform residual fusion:

$$z = h' + \alpha_{\text{maf}}\, g_{\text{maf}} \odot z_{\text{maf}}. \tag{B8}$$

Algorithm B3 FiLM Modulation and Gated Residual Fusion

Require: PLM embedding $h_{\text{seq}} \in \mathbb{R}^{L\times D}$; MAF global feature $h_{\text{maf}} \in \mathbb{R}^{Dm}$
Ensure: AFP prediction logit $\hat{y} \in \mathbb{R}$
// Step 1: Project to shared latent space
$z_{\text{seq}} \leftarrow W_{\text{seq}} \cdot h_{\text{seq}}$ ▷ $z_{\text{seq}} \in \mathbb{R}^{L\times H}$, $H = 512$
$z_{\text{maf}} \leftarrow W_{\text{maf}} \cdot h_{\text{maf}}$ ▷ $z_{\text{maf}} \in \mathbb{R}^{H}$
$q \leftarrow z_{\text{seq}}[1, :]$ ▷ Query from first token
// Step 2: FiLM modulation
$[\tilde{\gamma},\ \beta] \leftarrow \text{MLP}_{\text{film}}(h_{\text{maf}})$ ▷ $\tilde{\gamma}, \beta \in \mathbb{R}^{H}$
$\gamma \leftarrow \tanh(\tilde{\gamma}) \cdot s$ ▷ s: learnable scale
$h' \leftarrow (1+\gamma) \odot q + \beta$ ▷ Feature-wise affine transform
// Step 3: Gated residual fusion
$g_{\text{maf}} \leftarrow \sigma(W_g \cdot h_{\text{maf}} \,/\, \tau)$ ▷ $\tau = 2.0$
$\alpha_{\text{maf}} \leftarrow \sigma(\text{logit}_{\alpha})$ ▷ Learnable global weight
$z \leftarrow h' + \alpha_{\text{maf}} \cdot (g_{\text{maf}} \odot z_{\text{maf}})$
// Step 4: Prediction
$z \leftarrow \text{Dropout}(\text{LayerNorm}(z))$
$\hat{y} \leftarrow \text{MLP}_{\text{head}}(z)$ ▷ $512 \to 256 \to 1$
return $\hat{y}$

Appendix C Experiments

Appendix C.1 Datasets

High-quality datasets are essential for developing reliable machine-learning models. Following previous AFP prediction studies, we collected four publicly available benchmark datasets, namely Antifp_Main [15], Antifp_DS1 [15], Antifp_DS2 [15], and Deep-AFPpred [16]. These datasets were used exactly as released by the original authors. Redundancy reduction and sequence de-duplication had already been conducted during dataset construction to minimize overlap and sequence similarity between training and testing subsets. To ensure direct comparability with previously reported results, no additional filtering, redundancy removal, or data re-partitioning was performed in this study. A summary of the datasets and their statistical characteristics is provided in Table C1.

Table C1 Summary of AFP benchmark datasets.

Dataset	Training Set	Test Set	Description
Antifp_Main	1168/1168	291/291	The non-AFPs were randomly selected from Swiss-Prot and antimicrobial peptides other than AFPs obtained from DRAMP.
Antifp_DS1	1168/1168	291/291	The non-AFPs consisted of antimicrobial peptides with biological activities other than anti-fungal activity.
Antifp_DS2	1168/1168	291/291	The non-AFPs were randomly selected from Swiss-Prot.
Deep-AFPpred	2062/2062	1379/1379	AFPs of length 5–30 amino acids were collected from CAMP, DRAMP and StarPep. Non-AFPs of the same length range were randomly generated from Swiss-Prot.

Appendix C.2 Predictive performance and cross-dataset generalization of MafAFP on benchmark datasets

To comprehensively evaluate MafAFP, we conducted systematic experiments on benchmark datasets and, under the same data partitions, compared its performance with four representative existing methods, Deep-AFPpred [16], iAFP-fLRM [17], AFP-MFL [8], and ESM2-AFPpred [11]. We employed five standard evaluation metrics: accuracy (ACC), sensitivity (Sn), area under the ROC curve (AUC), Matthews correlation coefficient (MCC), and F1 score [18].

As shown in Figure C1(a), MafAFP achieved performance comparable to or better than existing methods across all datasets. In particular, MafAFP obtained the highest average AUC of 0.987 (Figure C1(b)), while exhibiting relatively small performance variation among datasets, indicating strong predictive stability.

To further assess cross-dataset generalization, we adopted a leave-one-dataset-out (LODO) evaluation strategy, in which the model was trained on all but one dataset and tested on the held-out dataset. In Figure C1(c), MafAFP consistently maintained high AUC values across all held-out datasets. Although a moderate decrease relative to within-dataset evaluation was observed, overall performance remained competitive, demonstrating robust transferability across different data distributions.

Furthermore, the ROC and precision–recall (PR) curves (Figure C1(d)) remained smooth and stable across datasets, further supporting the robustness of MafAFP under varying decision thresholds and benchmark settings.

Appendix C.3 Ablation study on the MAF module

To dissect the specific contribution of the MAF module to the overall architecture, we performed ablation studies comparing various baseline models: a baseline model without MAF (NoMafBaseline), a baseline model with simple concatenation of MAF features (SCMafBaseline) and the full MafAFP model incorporating MAF-based physical mechanism modeling. We examined the corresponding ROC and PR curves (Figure C2(a)) and the bar plots of the five evaluation metrics on the benchmark datasets (Figure C2(b)). The results show that, after introducing the MafExtractor, the ROC and PR curves shift noticeably towards the upper left and upper right, respectively, with the AUC improving by up to approximately 5%. These findings indicate that explicit physical mechanism

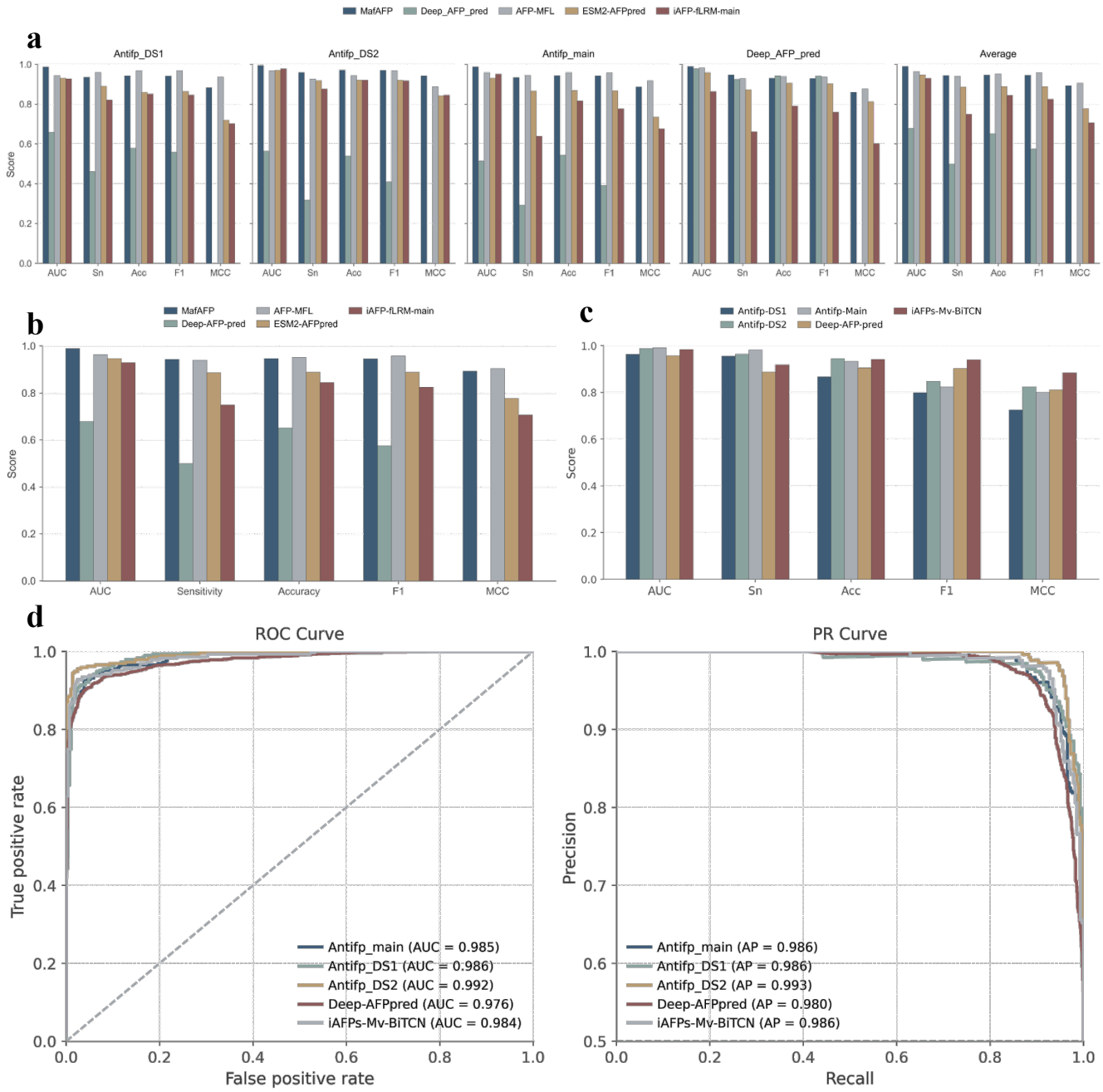

Figure C1 Performance of MafAFP. (a) Comparison of performance with four representative existing methods across datasets. (b) The average performance comparison with four representative existing methods across datasets. (c) Performance of MafAFP under the leave-one-dataset-out setting. (d) The ROC and PR curves of MafAFP on each dataset.

modeling can substantially enhance the discriminative power of the predictor. Importantly, while introducing this mechanism-aware component and markedly improving interpretability, MafAFP still maintains overall predictive performance that is competitive with current state-of-the-art methods and demonstrates strong generalization across heterogeneous AFP datasets.

Appendix C.4 Effects of PLM fine-tuning on MafAFP embedding distributions and predictive performance

To investigate how task-specific fine-tuning affects the representation space of antifungal peptides (AFPs), we compared the distributions of sequence-level ESMC embeddings before and after fine-tuning. Peptide sequences from all benchmark datasets were pooled, and residue-level embeddings were averaged to obtain a single representation for each peptide. The resulting embeddings were projected into two dimensions using principal component analysis (PCA) (Figure C3(a)) and t-distributed stochastic neighbor embedding (t-SNE) (Figure C3(b)). These visualizations provide a direct assessment of how fine-tuning reshapes the embedding space with respect to AFP

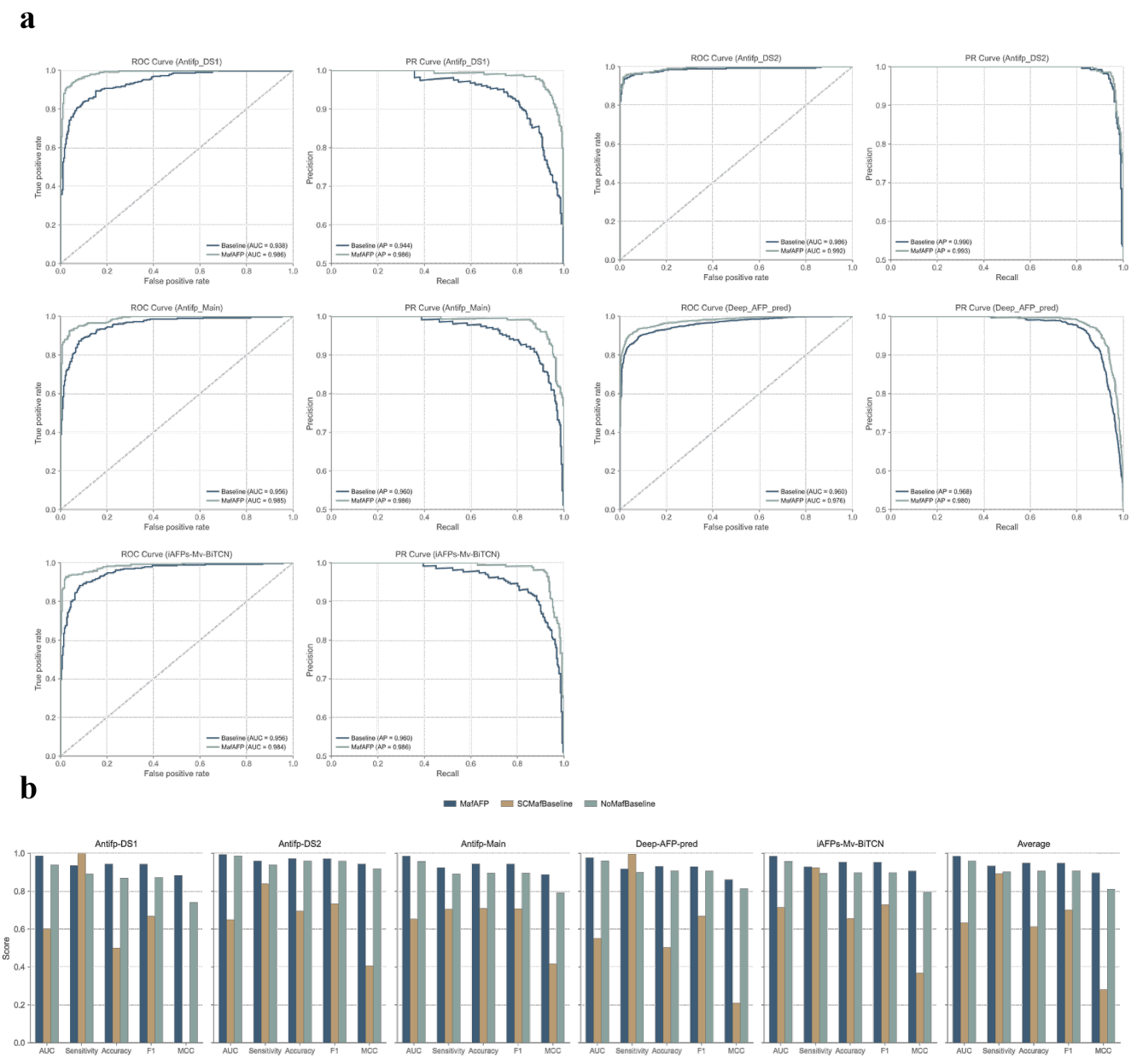

Figure C2 Performance for NoMafBaseline model and MafAFP. (a) The ROC and PR curves on benchmark datasets for NoMafBaseline model and MafAFP. (b) The evaluation metrics on benchmark datasets for all baseline models and MafAFP.

and non-AFP discrimination.

In Figure C3, fine-tuning substantially improves the separation between AFP and non-AFP samples, indicating that the model adapts the pretrained representation toward AFP-relevant sequence patterns. We further evaluated the effect of fine-tuning different numbers of ESMC encoder layers using ROC and precision–recall (PR) curves (Figure C3(c)), together with aggregate performance metrics (Figure C3(d)). Compared with the frozen-backbone model, fine-tuned variants consistently achieved improved predictive performance, confirming the importance of task-specific adaptation. However, increasing the number of trainable encoder layers beyond the final few layers produced only marginal improvements. This observation suggests diminishing returns from deeper fine-tuning and is consistent with the relatively limited size of current AFP datasets. Under such data-constrained conditions, extensive fine-tuning increases the risk of overfitting while providing little additional benefit to generalization performance.

Appendix C.5 Biophysically motivated interpretability of MafAFP: Datasets and Sequence Level Visualization of membrane-related features

To evaluate whether MafAFP captures membrane-related biophysical characteristics of antifungal peptides (AFPs), we analyzed the internal multi-scale MAF features learned from all AFP and non-AFP sequences across the benchmark datasets. Figure C4(a)-(b) show the positional distributions of the average multi-scale classification scores. AFP sequences exhibit pronounced and cross-scale-consistent

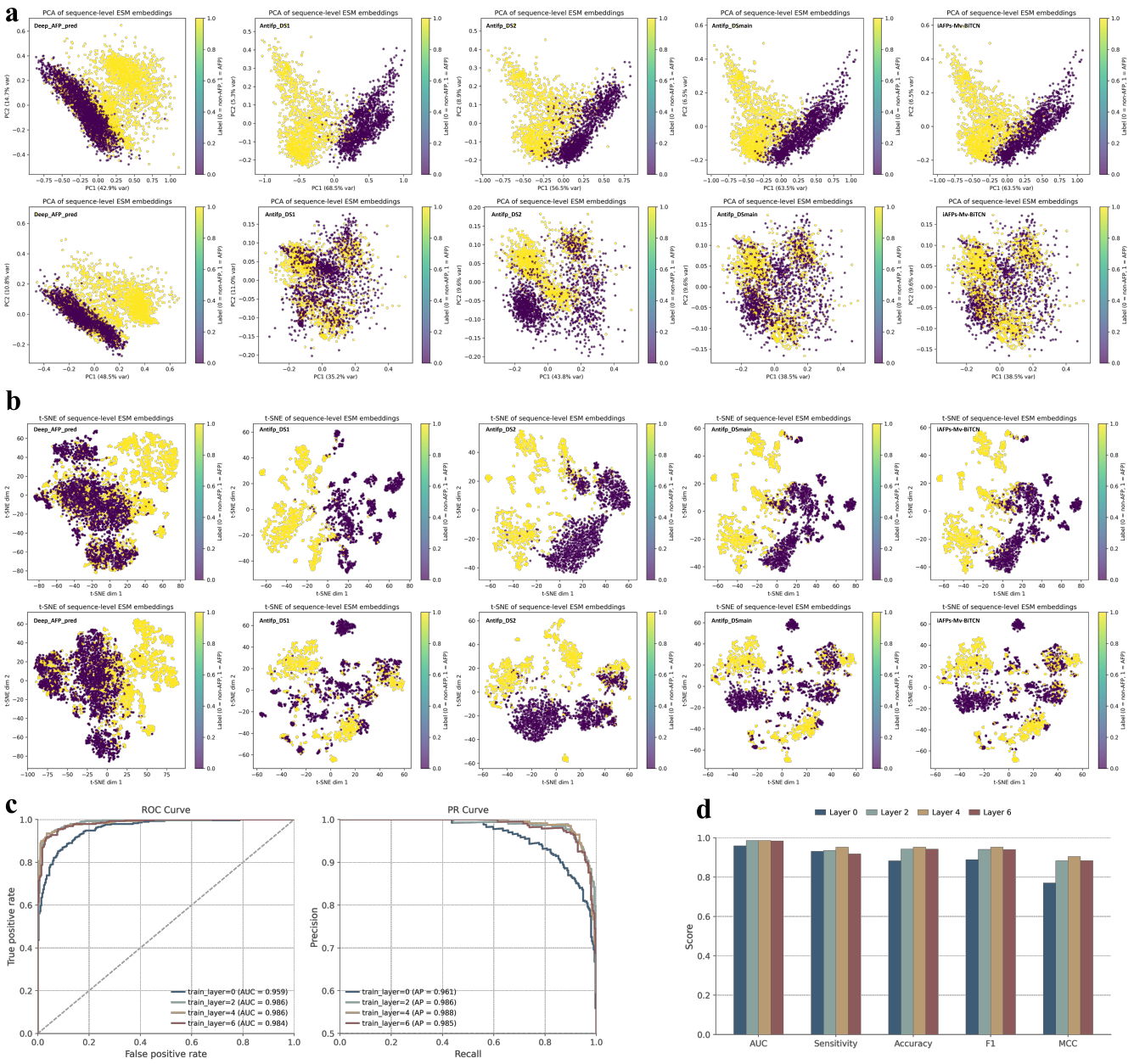

Figure C3 Fine-tuned PLM learned the biological properties of AFPs. (a) The PCA plot visualizing the embeddings from the fine-tuned model and the untrained model. (b) The t-SNE plot visualizing the embeddings from the fine-tuned model and the untrained model. (c) The ROC and PR curves for fine-tuned encoder layers. (d) The predictive performance for fine-tuned encoder layers.

Table C2 Effect of fine-tuning different numbers of ESMC encoder layers on MafAFP performance evaluated on the Antifp_DS1 dataset.

Model	AUC	Sn	ACC	F1	MCC
MafAFP (0 layers)	0.959	0.931	0.883	0.889	0.770
MafAFP (2 layers)	0.986	0.935	0.942	0.941	0.883
MafAFP (4 layers)	0.986	0.952	0.952	0.952	0.904
MafAFP (6 layers)	0.984	0.918	0.942	0.940	0.884

peaks, whereas non-AFP sequences display substantially flatter profiles. These results suggest that MafAFP identifies structured local regions that are preferentially associated with AFP activity.

We further compared the distributions of interfacial insertion free energy (ΔG_{if}) and hydrophobic moment (HM) derived from the physical pathway of MafAFP. As shown in Figure C4(c)-(d), AFP sequences exhibit stronger fluctuations in both quantities, with local minima in ΔG_{if} frequently coinciding with local maxima in HM. Such regions correspond to segments possessing both favorable

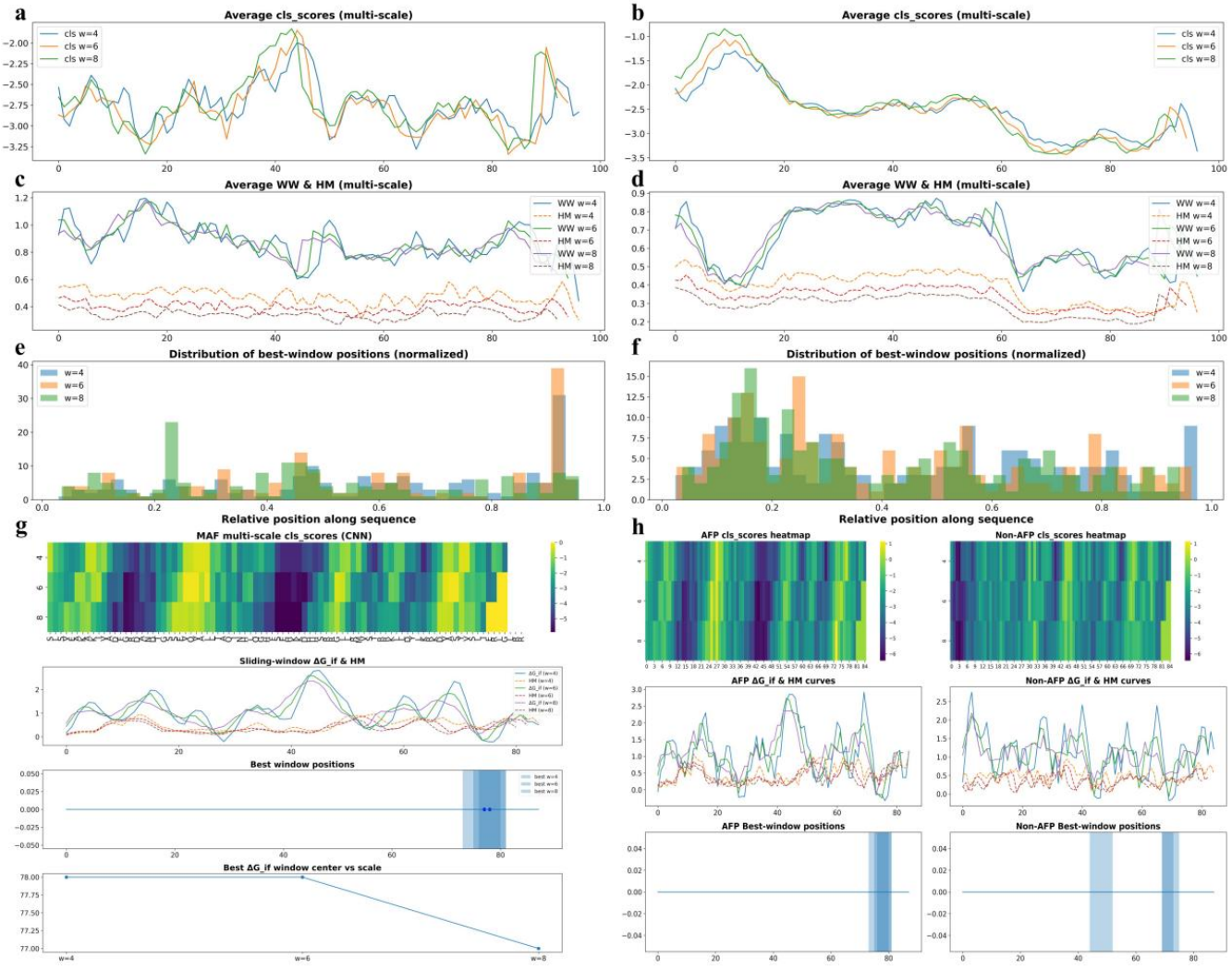

Figure C4 Datasets and Sequence Level Visualization of Raw Biophysical Features. (a) The distribution of multi-scale average classification scores along the sequence for AFPs. (b) The distribution of multi-scale average classification scores along the sequence for non-AFPs. (c) The distribution of sliding-window interfacial insertion free energy and hydrophobic moment for AFPs. (d) The distribution of sliding-window interfacial insertion free energy and hydrophobic moment for non-AFPs. (e) The positional distribution of the best window for each AFP. (f) The positional distribution of the best window for each non-AFP. (g) The multi-scale feature maps for a single AFP sequence. (h) The direct comparison between an AFP and a non-AFP sequence.

membrane-insertion propensity and amphipathic character, consistent with the classical membrane-active mechanism of AFPs. In contrast, non-AFP sequences lack similarly coordinated patterns. Analysis of the highest-scoring windows further revealed that AFPs exhibit clear positional enrichment across different window sizes, whereas non-AFPs show more diffuse distributions (Figure C4(e)-(f)), indicating that the learned local signals are biologically structured rather than randomly distributed.

To investigate these features at the sequence level, we visualized representative MAF maps for individual peptides (Figure C4(g)-(h)). For AFP sequences, the multi-scale cls_score heatmaps consistently identify a localized high-response region across different window sizes, suggesting the presence of a stable membrane-active segment. The corresponding ΔG_{if}/HM profiles reveal coincident insertion-favorable and amphipathic signatures within the same region, providing complementary biophysical evidence. Moreover, the highest-scoring windows remain concentrated within a common sequence segment across scales, indicating that MafAFP consistently localizes a single functional hotspot. By contrast, non-AFP sequences exhibit weaker and more diffuse responses without a clearly defined insertion motif.

Taken together, the dataset-level analyses (Figure C4(a)-(f)) demonstrate that MafAFP learns discriminative local patterns associated with membrane insertion and amphipathicity, while the sequence-level visualizations (Figure C4(g)-(h)) show that these patterns correspond to biologically meaningful membrane-active segments. These results provide mechanistic support for the predictive capability of MafAFP and highlight the interpretability of its multi-scale membrane-aware representations.

Appendix C.6 Multi-attribute modulation and feature fusion in MafAFP

Unlike conventional feature-concatenation strategies, MafAFP incorporates membrane-related biophysical information through a dedicated modulation pathway that directly reshapes protein language model (PLM) representations. To characterize this mechanism, we

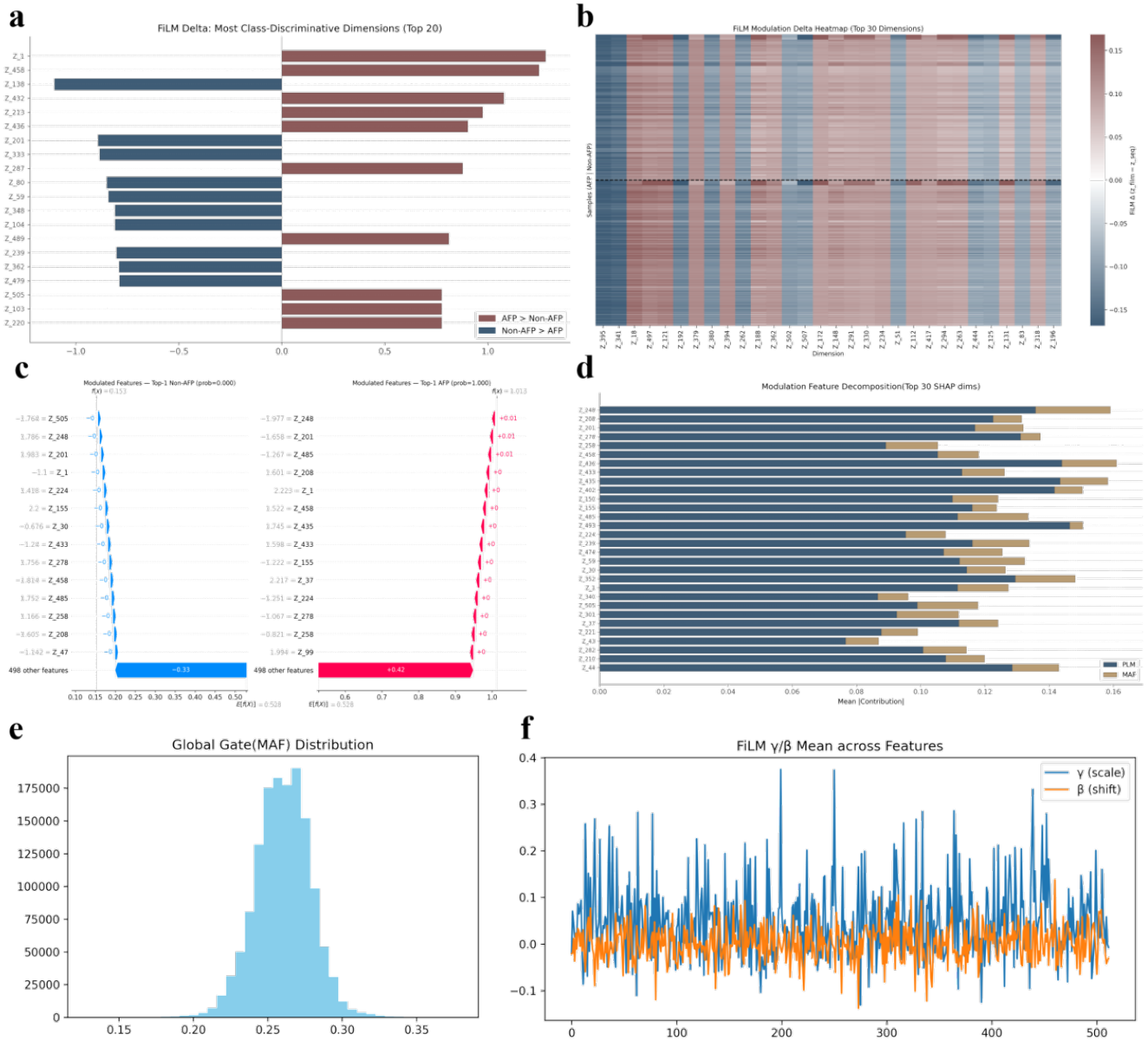

Figure C5 Multi-attribute modulation and feature fusion in MafAFP. (a) The Most Class-Discriminative Dimensions. (b) The FiLM Modulation Delta Heatmap. (c) The Waterfall Chart for Non-AFP. (d) The PLM/MAF Branch Contribution Decomposition. (e) The Global Gate (MAF) distribution. (f) The FiLM parameter analysis.

first examined the class-separability of all 512 FiLM-modulated latent dimensions. Only a small subset of channels exhibited strong class discrimination, with dimensions such as Z_1 and Z_458 showing the largest effect sizes (Cohen's $d > 1$), while most channels displayed substantially weaker shifts (Figure C5(a)). Consistent with this observation, the FiLM modulation delta heatmap revealed that the most strongly modulated channels were preferentially activated in AFP samples but remained comparatively weak in non-AFP samples (Figure C5(b)). These results indicate that biophysical information is incorporated through sparse and selective modulation rather than uniformly distributed across the latent space.

To investigate how these modulated features influence model predictions, we performed SHAP analysis on representative AFP and non-AFP samples. Non-AFP predictions were primarily driven by a small number of strongly inhibitory dimensions, whereas AFP predictions were supported by a limited set of strongly positive dimensions (Figure C5(c)). This opposite attribution pattern suggests that the same latent space encodes class-discriminative evidence in different directions, with final decisions determined by a relatively small number of influential channels.

We next examined the origin of these predictive signals by comparing the contributions of the PLM and MAF branches across the top 30 SHAP-ranked dimensions. PLM-derived features accounted for the largest effects overall, particularly in dimensions such as Z_248, Z_505, and Z_201, indicating that the pretrained language model provides the primary discriminative representation for AFP recognition (Figure C5(d)). In contrast, MAF-derived contributions were distributed more evenly across multiple channels. This

complementary pattern suggests that the MAF branch does not replace sequence representations, but instead refines the decision boundary by introducing mechanistically relevant physicochemical constraints.

Analysis of the fusion process further supported this interpretation. The distribution of global MAF gate values peaked at approximately 0.25–0.30, indicating that the physicochemical branch contributes roughly one-quarter to one-third of the fused representation across samples (Figure C5(e)). Likewise, the FiLM parameters γ and β remained within a relatively narrow range (approximately -0.10 to $+0.15$), consistent with mild but persistent feature-wise modulation (Figure C5(f)). Positive γ values correspond to channel amplification, whereas negative values indicate attenuation, while β introduces small baseline shifts. The similar but non-identical profiles of γ and β suggest that multiplicative and additive modulation operate cooperatively rather than redundantly.

Taken together, these results demonstrate that MafAFP integrates membrane-related biophysical information through sparse, moderate, and interpretable modulation of PLM representations. This mechanism preserves the strong sequence modeling capability of the language model while introducing biologically meaningful constraints that improve AFP discrimination. The code for this work is available for download at https://github.com/Hanszera/MAF_AFP.

References

1 Brown GD, Denning DW, Gow NA, et al. Hidden killers: human fungal infections. Sci Transl Med, 2012, 4: 165rv113
2 Perfect JR. The antifungal pipeline: a reality check. Nat Rev Drug Discov, 2017, 16: 603–616
3 Singh S, Chaudhary K, Dhanda SK, et al. SATPdb: a database of structurally annotated therapeutic peptides. Nucleic Acids Res, 2016, 44(D1): D1119–D1126
4 Fang C, Moriwaki Y, Li C, Shimizu K. Prediction of antifungal peptides by deep learning with character embedding. IPSJ Trans Bioinform, 2019, 12: 21–29
5 Lin Z, Akin H, Rao R, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science, 2023, 379(6637): 1123–1130
6 Rao R, Bhattacharya N, Thomas N, et al. Evaluating protein transfer learning with TAPE. Adv Neural Inf Process Syst, 2019, 32: 9689–9701
7 Elnaggar A, Heinzinger M, Dallago C, et al. ProtTrans: towards cracking the language of life's code through self-supervised deep learning and high performance computing. IEEE Trans Pattern Anal Mach Intell, 2021, 44: 7112–7127
8 Fang Y, Xu F, Wei L, et al. AFP-MFL: accurate identification of antifungal peptides using multi-view feature learning. Brief Bioinform, 2023, 24(1): bbac606
9 Yao L, Zhang Y, Li W, et al. DeepAFP: an effective computational framework for identifying antifungal peptides based on deep learning. Protein Sci, 2023, 32(10): e4758
10 Akbar S, Zou Q, Raza A, et al. iAFPs-Mv-BiTCN: predicting antifungal peptides using self-attention transformer embedding and transform evolutionary-based multi-view features with bidirectional temporal convolutional networks. Artif Intell Med, 2024, 151: 102860
11 Yin K, Li R, Zhang S, et al. Deep learning combined with quantitative structure–activity relationship accelerates de novo design of antifungal peptides. Adv Sci, 2025, 12(13): 2412488
12 Wang J, Feng J, Kang Y, et al. Discovery of antimicrobial peptides with notable antibacterial potency by an LLM-based foundation model. Sci Adv, 2025, 11(10): eads8932
13 Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell, 2020, 2: 56–67
14 Ding W, Xu Q, Liu S, et al. SAMF: a self-adaptive protein modeling framework. Bioinformatics, 2021, 37(22): 4075–4082
15 Agrawal P, Bhalla S, Chaudhary K, et al. In silico approach for prediction of antifungal peptides. Front Microbiol, 2018, 9: 323
16 Sharma R, Shrivastava S, Kumar Singh S, et al. Deep-AFPpred: identifying novel antifungal peptides using pretrained embeddings from seq2vec with 1DCNN-BiLSTM. Brief Bioinform, 2022, 23(1): bbab422
17 Zhang S, Cheng J, Zhou G, et al. iAFP-fLRM: accurate identification of antifungal peptides via hybrid deep learning architecture and multi-modal feature fusion. Biophys Chem, 2026: 107591
18 Dwivedi AK. Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Comput Appl, 2018, 29: 685–693