PEPARD: FULL-ATOM PEPTIDE DESIGN VIA GEOMETRIC LATENT DIFFUSION WITH ADAPTIVE REWARD STEERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Target-aware peptide design requires simultaneously modeling molecular geometry, receptor interactions, and diverse biophysical constraints. However, existing methods often use coarse structural representations, separate peptide generation from pose prediction, or rely on fixed objectives during optimization. We propose PepARD, a unified framework for full-atom target-specific peptide design based on geometric latent diffusion. PepARD encodes peptide sequence and 3D structure into complementary invariant and $SE(3)$-equivariant latent spaces, while its Adaptive Multi-Channel $E(n)$-Equivariant Graph Neural Network (AMEGNN) captures detailed receptor-peptide interactions with variable heavy-atom configurations. Receptor-conditioned diffusion enables both *de novo* sequence-structure generation and fixed-sequence binding pose prediction within a shared framework. To further improve physical plausibility, an Adaptive Reward Weight Network dynamically balances structural, steric, geometric, and sequence objectives during latent refinement. Experiments on PepBench and PepBDB demonstrate that PepARD produces diverse and physically plausible peptides with favorable binding properties and improves recovery of native-like binding conformations.

1 INTRODUCTION

Target-specific peptide design is a fundamental challenge in computational drug discovery and therapeutics (Zhai et al., 2025). Compared with small molecules, peptides often provide high target affinity and specificity with reduced off-target toxicity (Ji et al., 2024; Torres et al., 2025), while offering advantages in tissue penetration and manufacturing cost over large protein biologics (Lai et al., 2025). The goal of target-conditioned peptide design is to generate amino acid sequences together with receptor-bound 3D conformations that are structurally compatible with a designated protein pocket (Zhang et al., 2025b; Chen et al., 2026). Recent diffusion and flow-matching models have substantially advanced *de novo* molecular generation (Zhang et al., 2025a; Morehead et al., 2026; Matsukiyo et al., 2024). In peptide design, RFdiffusion (Watson et al., 2023) enabled expressive backbone generation, followed by full-atom and sequence-structure co-design methods such as PepFlow (Li et al., 2024), PPDFlow (Lin et al., 2024), and PepGLAD (Kong et al., 2024), as well as cross-modal and guided approaches including UniMoMo (Kong et al., 2025b), PepTri (Nguyen et al., 2026), and PepFGLD (Niu et al., 2026). Despite this progress, three challenges remain.

A central limitation of existing peptide generation methods lies in their molecular representations. Many approaches rely on molecular graphs or residue-level C_α representations (Sayuti et al., 2025; Qin et al., 2026), which provide limited descriptions of side-chain geometry and fine-grained steric interactions. Equivariant architectures such as $E(n)$-EGNN (Satorras et al., 2021) and higher-order tensor networks (Brandstetter et al., 2022), together with equivariant diffusion models such as EDM (Hoogeboom et al., 2022), GeoDiff (Xu et al., 2022), and TargetDiff (Guan et al., 2023), have established strong geometric inductive biases for 3D molecular modeling. However, extending these approaches to irregular full-atom peptide representations with variable heavy-atom configurations remains challenging.

Another challenge arises from the separation between sequence generation and binding-pose prediction. Recent models such as AlphaFold3 (Abramson et al., 2024), DiffDock (Corso et al., 2023),

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.

and TorNet (Jing et al., 2022) have advanced structure and pose prediction, but generally operate on pre-existing molecular sequences or structures. This separation prevents the generated sequence and binding geometry from mutually adapting to the receptor and often necessitates downstream docking or structural optimization (Zhang et al., 2026).

Beyond representation and joint generation, balancing multiple and competing design objectives remains difficult. RL-based methods have been explored for sequence optimization (Wang et al., 2023; 2025) and coupled with diffusion sampling through RERD (Uehara et al., 2025) or heuristic guidance such as PepTri (Nguyen et al., 2026). Energy-guided generation has also been investigated for molecular optimization (Zhao et al., 2024). However, these approaches typically rely on predefined reward weights, which are difficult to adapt to changing steric, structural, and sequence constraints during generation.

To address these challenges, we propose **PepARD** (**Pep**tide **A**daptive **R**eward **D**iffusion), an end-to-end framework for target-specific full-atom peptide design. Our main contributions are threefold:

- **Unified Full-Atom Latent Diffusion.** We introduce a geometric latent diffusion framework that represents peptide sequence and full-atom geometry in complementary invariant and equivariant latent spaces. The receptor-conditioned generative process supports both *de novo* sequence-structure co-design and fixed-sequence binding conformation prediction within a shared framework, without requiring post-hoc docking.
- **Geometric Modeling with AMEGNN.** We develop AMEGNN with RollerPooling to model irregular full-atom representations with variable numbers of valid atoms across amino acids, while preserving geometric equivariance and capturing fine-grained receptor-peptide interactions.
- **Adaptive Post-Diffusion Reward Steering.** We introduce a lightweight Adaptive Reward Weight Network that dynamically predicts simplex-constrained weights over competing biophysical objectives during latent refinement. A gradient-guided steering alignment objective with an exponential moving baseline enables adaptive multi-objective refinement without relying on fixed reward weights.

2 METHODS

Problem Formulation. Given a target protein receptor with binding-pocket coordinates $X_r \in \mathbb{R}^{N_r \times 14 \times 3}$ and residue attributes $H_r \in \mathbb{R}^{N_r \times d_r}$, where N_r is the number of pocket residues and d_r is the receptor residue-feature dimension, the goal of target-specific peptide design is to generate a peptide defined by its amino acid sequence $S = (s_1, s_2, \ldots, s_L)$ with $s_i \in \{1, \ldots, 20\}$ and its 3D full-atom conformation $X \in \mathbb{R}^{L \times 14 \times 3}$, where L is the peptide length and each residue accommodates up to 14 heavy atoms. This formulation natively supports two task settings: *de novo* sequence-structure co-design, where both S and X are jointly generated, and fixed-sequence binding conformation prediction (*FixSeq*), where S is provided as a prior condition and only the receptor-bound conformation X is predicted.

Overall Architecture. PepARD is a two-stage geometric latent diffusion framework, illustrated in Figure 1. The first stage is a dual-stream equivariant autoencoder that compresses the full-atom complex into an invariant scalar latent $H \in \mathbb{R}^{L \times d}$ ($d = 8$) encoding residue chemical semantics and an $SE(3)$-equivariant coordinate latent $Z \in \mathbb{R}^{L \times C_z \times 3}$ ($C_z = 3$) capturing backbone and side-chain orientation topology. The second stage applies receptor-conditioned latent diffusion over joint trajectories $(H_t, Z_t)_{t=1}^{T}$ ($T = 100$), after which a differentiable multi-objective reward module scores the denoised latents and a post-diffusion refinement procedure updates them while training an Adaptive Reward Weight Network to balance competing biophysical objectives. Pre-trained decoders finally map the latents (H_K, Z_K) after $K = 10$ refinement steps back to the sequence $\hat{S}$ and full-atom coordinates $\hat{X}$, followed by Rosetta FastRelax for downstream evaluation, for FixSeq, $H_K = H_0$ by construction.

2.1 GEOMETRIC AUTOENCODER

Dual-Stream Latent Encoding. Given the peptide sequence S and full-atom coordinates X, a Separated Amino Acid Feature module first constructs sequence-only residue embeddings $H_{\text{init}} =$

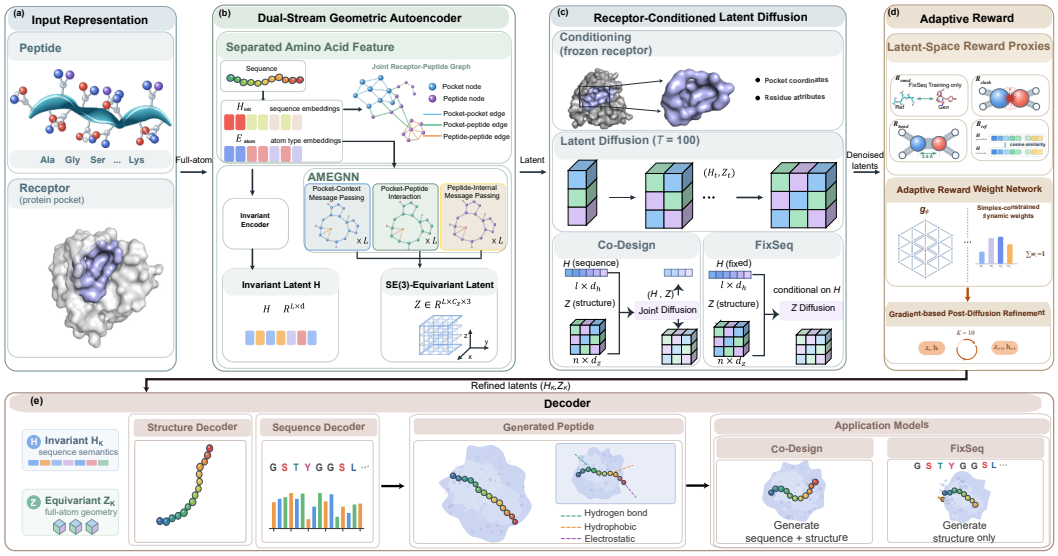


Figure 1: PepARD architecture. **(a)** Full-atom complex representation used for autoencoder training. **(b)** Dual-stream encoding into invariant and equivariant latents. **(c)** Receptor-conditioned diffusion over (H, Z) for co-design and over Z only for FixSeq, with H fixed. **(d)** Task-masked latent-space reward proxies and post-diffusion refinement, r_{rmsd} is active only during FixSeq training. **(e)** Decoding into full-atom peptide structures and task-specific outputs.

$E_{\text{aa}}(S) \in \mathbb{R}^{L\times d}$ and per-channel atom-type embeddings $E_{\text{atom}} \in \mathbb{R}^{L\times 14\times d_c}$, where E_{aa} is learned and d_c is the atom-embedding dimension. The invariant stream encodes residue chemical semantics solely from the prescribed amino-acid sequence, parameterizing $q_{\text{enc}}(H \mid S) = \mathcal{N}(\boldsymbol{\mu}_H(S), \text{diag}(\boldsymbol{\sigma}_H^2(S)))$ over $H \in \mathbb{R}^{L\times d}$, with encoder-predicted mean $\boldsymbol{\mu}_H$ and standard deviation $\boldsymbol{\sigma}_H$. Its KL regularizer is $\mathcal{L}_{\text{KL}}^H = D_{\text{KL}}(q_{\text{enc}}(H \mid S) \,\|\, \mathcal{N}(\mathbf{0}, \mathbf{I}))$. In parallel, the equivariant geometric stream combines H_{init}, E_{atom}, and receptor-peptide geometry through AMEGNN (Section 2.2) operating on the joint graph of pocket and peptide residues. This stream parameterizes a Gaussian encoder posterior $q_{\text{enc}}(Z \mid X, S)$ over $Z \in \mathbb{R}^{L\times C_z\times 3}$, anchored at ground-truth backbone landmarks $Z^{\text{anchor}} \in \mathbb{R}^{L\times C_z\times 3}$ with backbone atom labels $(\text{N}, \text{C}_\alpha, \text{C})$, and prior standard deviation σ_p. For FixSeq, we use $H_0 = \boldsymbol{\mu}_H(S)$ as the fixed sequence condition, no quantity derived from the ground-truth conformation X_{target} is used to construct H_0:

$$\mathcal{L}_{\text{KL}}^Z = D_{\text{KL}}\big(q_{\text{enc}}(Z \mid X, S) \,\|\, \mathcal{N}(Z^{\text{anchor}}, \sigma_p^2\mathbf{I})\big)\,. \tag{1}$$

Because Z^{anchor} transforms equivariantly under rigid-body transformations ($SE(3)$ rotations and translations), the Euclidean distance metric and the resulting divergence $\mathcal{L}_{\text{KL}}^Z$ remain strictly $SE(3)$-invariant. The KL balancing factors are set to $\beta_H = 0.3$ and $\beta_Z = 0.5$.

Decoding and Training Objectives. Reconstruction is executed by two specialized sub-networks. The sequence decoder $\mathcal{D}_{\text{seq}}$ predicts residue emission logits $\hat{S} \in \mathbb{R}^{L\times 20}$ from H conditioned on Z and the pocket context, supervised by cross-entropy $\mathcal{L}_{\text{seq}} = \text{CE}(\hat{S}, S)$. Concurrently, the structural decoder $\mathcal{D}_{\text{struct}}$ reconstructs full heavy-atom coordinates $\hat{X} \in \mathbb{R}^{L\times 14\times 3}$ from Z conditioned on H, supervised by a composite loss $\mathcal{L}_{\text{struct}}$ comprising coordinate mean squared error $\mathcal{L}_{\text{MSE}}$, bond-length consistency, and side-chain torsional constraints. The overall autoencoder objective is:

$$\mathcal{L}_{\text{AE}} = (1-\gamma_{\text{AE}})\left(\mathcal{L}_{\text{seq}} + \beta_H\mathcal{L}_{\text{KL}}^H\right) + \gamma_{\text{AE}}\left(\mathcal{L}_{\text{struct}} + \beta_Z\mathcal{L}_{\text{KL}}^Z\right), \tag{2}$$

where $\gamma_{\text{AE}} = 0.5$ balances sequence and structural fidelity.

2.2 AMEGNN Architecture

AMEGNN generalizes geometric message passing to multi-channel coordinate tensors, capturing higher-order spatial interactions across residues with variable heavy-atom counts.

Multi-Channel Representation and Joint Graph. Each residue node i maintains a multi-channel coordinate tensor $Y_i \in \mathbb{R}^{C\times 3}$ ($C = 14$ for full-atom modeling in the autoencoder and $C = C_z = 3$ for coordinate latents during diffusion), scalar features H_i, and a validity mask $m_i \in \{0,1\}^C$ indicating active atoms or channels, thus $n_i = \sum_{c=1}^{C} m_i^c$ is the number of active channels at node i. The receptor pocket and the peptide are represented as a unified heterogeneous graph. Edges are partitioned into *pocket-context edges*, *pocket-peptide interaction edges*, and *peptide internal edges*. Pocket coordinates are frozen via an update mask $u_i \in \{0,1\}^{C\times 1}$ ($u_i = \mathbf{0}$ for receptor nodes), preserving the rigid receptor structure while allowing bidirectional contextual propagation.

Multi-Channel Message Passing. For each edge (i, j), pairwise inter-channel distances $D_{ij} \in \mathbb{R}^{C\times C}$ with entries $D_{ij}^{ab} = \|Y_i^a - Y_j^b\|_2$, where $a, b \in \{1, \dots, C\}$ index channels, are projected via radial basis functions and modulated by atom-type embeddings into geometric attributes R_{ij}. Raw edge messages are computed as:

$$M_{ij} = \mathrm{MLP}_{\mathrm{edge}}\left(\left[H_i \,\|\, H_j \,\|\, R_{ij} \,\|\, E_{ij}^{\mathrm{attr}}\right]\right), \tag{3}$$

where $\mathrm{MLP}_{\mathrm{edge}}$ produces the channel-aware edge message M_{ij}, E_{ij}^{attr} encodes chemical-bond and interaction edge-type indicators, and $[\cdot\|\cdot]$ denotes feature concatenation.

RollerPooling for Irregular Channel Masks. Because residues contain variable numbers of valid channels $n_i, n_j \le C$, naive spatial pooling over padded matrices introduces non-physical boundary artifacts. RollerPooling applies normalized pooling operators $\{P_{n_i,n_j}\}$ tailored to active channels:

$$\bar{M}_{ij} = \mathrm{RollerPool}(M_{ij}; P_{n_i,n_j}), \tag{4}$$

where P_{n_i,n_j} restricts aggregation strictly to the valid $n_i \times n_j$ sub-block of channel pairs. For intra-residue self-interaction edges ($i = j$), RollerPooling contracts over the non-redundant upper-triangular band to prevent double-counting internal atom pairs.

Equivariant Layer Propagation. Each layer updates node invariants and coordinate tensors via:

$$H_i^{(\ell+1)} = H_i^{(\ell)} + \mathrm{MLP}_{\mathrm{node}}\left(\left[H_i^{(\ell)} \,\|\, \sum_{j\in\mathcal{N}(i)} \bar{M}_{ij}\right]\right), \tag{5}$$

$$Y_i^{(\ell+1)} = Y_i^{(\ell)} + u_i \odot \sum_{j\in\mathcal{N}(i)} (Y_i^{(\ell)} - Y_j^{(\ell)}) \odot \mathrm{MLP}_{\mathrm{coord}}(\bar{M}_{ij}), \tag{6}$$

where ℓ indexes layers, $\mathcal{N}(i)$ is the neighbor set of node i, and $\odot$ denotes element-wise multiplication with broadcasting over Cartesian coordinates. Here, $\mathrm{MLP}_{\mathrm{node}}$ outputs an invariant feature update with the same dimension as H_i, while $\mathrm{MLP}_{\mathrm{coord}}$ maps the pooled edge message $\bar{M}_{ij}$ to channel-wise scalar coefficients that weight the relative coordinate displacements. Because updates rely purely on relative displacements and invariant messages, AMEGNN guarantees exact $E(3)$-equivariance for coordinates and invariance for scalar features.

2.3 Receptor-Conditioned Latent Diffusion

Forward Perturbation. Let (H_0, Z_0) denote the clean encoded latents. Forward diffusion corrupts them across $T = 100$ steps using a cosine noise schedule $\bar{\alpha}_t = f(t)/f(0)$, where $f(t) = \cos^2\left(\frac{\pi}{2} \cdot \frac{t/T+s}{1+s}\right)$ with offset $s = 0.01$:

$$q_{\mathrm{diff}}(H_t, Z_t \mid H_0, Z_0) = \mathcal{N}\big(H_t; \sqrt{\bar{\alpha}_t}H_0, (1-\bar{\alpha}_t)\mathbf{I}\big)\,\mathcal{N}\big(Z_t; \sqrt{\bar{\alpha}_t}Z_0, (1-\bar{\alpha}_t)\mathbf{I}\big). \tag{7}$$

Conditioning and Denoising Objective. Before diffusion, coordinates are normalized relative to the pocket C_α centroid: $Z_{\mathrm{norm}} = (Z - \bar{X}_r^{C_\alpha})/\sigma_{\mathrm{coord}}$ with $\sigma_{\mathrm{coord}} = 10.0$ Å. The denoiser $\boldsymbol{\epsilon}_\theta$ utilizes an AMEGNN backbone parameterized by time embedding $\boldsymbol{\tau}_t$ and sequence position embeddings $\mathbf{p}_{\mathrm{seq}}$:

$$(\hat{\boldsymbol{\epsilon}}_H, \hat{\boldsymbol{\epsilon}}_Z) = \boldsymbol{\epsilon}_\theta\left(H_t, Z_t, \boldsymbol{\tau}_t, \mathbf{p}_{\mathrm{seq}}\,;\, X_r, H_r\right). \tag{8}$$

Here, θ denotes the learnable denoiser parameters, and $\hat{\boldsymbol{\epsilon}}_H$ and $\hat{\boldsymbol{\epsilon}}_Z$ are the predicted noises for the invariant and equivariant latent streams, respectively. The model is trained by minimizing the joint noise estimation objective:

$$\mathcal{L}_{\mathrm{LDM}} = \mathbb{E}_{t,\boldsymbol{\epsilon}}\left[\|\hat{\boldsymbol{\epsilon}}_Z - \boldsymbol{\epsilon}_Z\|_F^2 + \lambda_H\|\hat{\boldsymbol{\epsilon}}_H - \boldsymbol{\epsilon}_H\|_F^2\right], \tag{9}$$

where t is uniform on $\{1, \ldots, T\}$, $\boldsymbol{\epsilon}_H$ and $\boldsymbol{\epsilon}_Z$ are independent standard Gaussian noises, and $\lambda_H = 3C_z/d$ balances the two streams. Let $p_t = (H_t, Z_t)$ denote the joint latent state. The reverse sampling follows the ancestral DDPM transition, where $\alpha_t = \bar{\alpha}_t/\bar{\alpha}_{t-1}$ is the per-step retention factor:

$$p_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(p_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \hat{\boldsymbol{\epsilon}}_\theta(p_t) \right) + \sigma_t \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \text{ for } t > 1. \tag{10}$$

Here, conditioning is suppressed in $\hat{\boldsymbol{\epsilon}}_\theta(p_t)$, and $\sigma_t^2 = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}(1-\alpha_t)$ is the DDPM posterior variance, with $\sigma_1 = 0$.

Inference-Time Latent Geometry Guidance. During sampling, generation is stabilized using an empirical adjacent-residue distance prior $E(Z_t) = \frac{1}{L-1}\sum_{i=1}^{L-1} \max(0, |\|z_{i+1}^{C_\alpha} - z_i^{C_\alpha}\|_2 - \mu_d| - 3\sigma_d)^2$. Here, $z_i^{C_\alpha}$ is the C_α latent-channel coordinate of residue i, and μ_d and σ_d are the training-set mean and standard deviation of adjacent C_α distances. The gradient of this geometric energy penalizes non-physical chain breaks by perturbing the predicted noise:

$$\hat{\boldsymbol{\epsilon}}'_Z = \hat{\boldsymbol{\epsilon}}_Z + \lambda_E \sqrt{1-\bar{\alpha}_t}\, \nabla_{Z_t} E(Z_t), \tag{11}$$

where λ_E denotes the guidance strength.

Task Configuration and Target-Blind FixSeq Protocol. In *FixSeq*, the prescribed sequence is encoded as $H_0 = \boldsymbol{\mu}_H(S)$ using the sequence-only invariant encoder and is clamped throughout denoising and refinement, only Z is generated. The ground-truth conformation X_{target} is unavailable to all generative components, and r_{rmsd} is masked at inference. It is used only after generation for RMSD and DockQ evaluation. Full protocol details are given in Appendix A.2. For *de novo* co-design, both H and Z are jointly generated and refined.

2.4 Differentiable Multi-Objective Reward Modeling

Following reverse diffusion, denoised latents (Z_0, H_0) are scored by four differentiable reward functions $\mathbf{r} = [r_{\text{rmsd}}, r_{\text{clash}}, r_{\text{bond}}, r_{\text{seq}}]^\top$. These objectives operate on the compact latent variables rather than directly on decoded all-heavy-atom coordinates, in particular, r_{clash} and r_{bond} are latent-space geometric proxies that regularize refinement. Prior to reward computation, Z_0 is mapped back to the physical scale via $Z_0 \leftarrow Z_0 \cdot \sigma_{\text{coord}} + \bar{X}_r^{C_\alpha}$. Full-atom structures are produced only after decoding the refined latents and are subsequently evaluated after Rosetta FastRelax using all-heavy-atom and interface metrics.

Structure Fidelity Reward (r_{rmsd}, FixSeq training only). During training, this auxiliary reward evaluates alignment against the ground-truth coordinates X_{target} on the C_α latent channel:

$$r_{\text{rmsd}} = -\sqrt{\frac{1}{L}\sum_{i=1}^{L} \left\| z_i^{C_\alpha} - z_{\text{target},i}^{C_\alpha} \right\|_2^2 + \delta}, \tag{12}$$

where $z_{\text{target},i}^{C_\alpha}$ is the target C_α coordinate for residue i and $\delta = 10^{-6}$ ensures numerical stability. This term is omitted in co-design and during target-blind FixSeq inference, where its associated reward weight w_{rmsd} is set to zero, at test time, the remaining active reward weights are renormalized to sum to one.

Steric Clash Proxy (r_{clash}). Penalizes excessive proximity across active coordinate-latent channel pairs:

$$r_{\text{clash}} = -\frac{\eta}{N_{\text{ch}}} \sum_{1 \le a < b \le N_{\text{ch}}} \max(0, \tau_{\text{clash}} - \|z_a - z_b\|_2), \tag{13}$$

where z_a and z_b are coordinates in the flattened set of active peptide latent channels, $\tau_{\text{clash}} = 0.8\,\text{Å}$ is the proximity threshold, $\eta = 5.0$ is the penalty scale, and $N_{\text{ch}} = \sum_{i=1}^{L} n_i$ is the number of active channels.

Bond Length Regularization (r_{bond}). Penalizes distortion between consecutive C_α channels relative to standard equilibrium distance $d_{\text{target}} = 3.8\,\text{Å}$ with tolerance $\epsilon_{\text{bond}} = 0.5\,\text{Å}$:

$$r_{\text{bond}} = -\frac{1}{L-1}\sum_{i=1}^{L-1} \max\Big(0, \Big| \|z_{i+1}^{C_\alpha} - z_i^{C_\alpha}\|_2 - d_{\text{target}} \Big| - \epsilon_{\text{bond}} \Big). \tag{14}$$

Sequence Semantic Alignment (r_{seq}). Encourages scalar latents H to align with the learned embedding prototypes $\{\mathbf{e}_m\}_{m=1}^{20}$ of the 20 canonical amino-acid types using cosine similarity:

$$r_{\text{seq}} = \frac{1}{L}\sum_{i=1}^{L} \max_{m\in\{1,\dots,20\}} \left(\frac{H_i \cdot \mathbf{e}_m}{\|H_i\|_2 \|\mathbf{e}_m\|_2}\right). \tag{15}$$

2.5 Adaptive Reward Weight Network and Post-Diffusion Refinement

The Adaptive Reward Weight Network g_ϕ predicts four reward weights $\mathbf{w} \in \Delta^3 = \{\mathbf{w} \in \mathbb{R}^4_{\geq 0} : \sum_{j=1}^4 w_j = 1\}$ with a task mask:

$$\mathbf{w} = g_\phi(Z, H, \boldsymbol{\tau}_k, L) = \text{softmax}(f_\phi(\boldsymbol{\tau}_k, \mathbf{s}_{Z,H}) + \mathbf{M}_{\text{task}}), \tag{16}$$

Here, f_ϕ (parameters ϕ) outputs reward logits, $\boldsymbol{\tau}_k$ encodes refinement step k, and $\mathbf{s}_{Z,H}$ summarizes the latent state. The mask $\mathbf{M}_{\text{task}}$ is 0 for active rewards and $-\infty$ otherwise, so $\mathcal{A}_{\text{task}} = \{j : M_{\text{task},j} = 0\}$ is the active index set. Details are in Appendix A.1.

Post-Diffusion Latent Refinement. After obtaining denoised candidates (Z_0, H_0), we execute $K = 10$ steps of gradient-based refinement:

$$\mathbf{w}_k = g_\phi(Z_k, H_k, \boldsymbol{\tau}_k, L), \tag{17}$$

$$Z_{k+1} = Z_k + \gamma_{\text{refine}} \nabla_Z \mathcal{R}(Z_k, H_k; \mathbf{w}_k), \tag{18}$$

$$H_{k+1} = H_k + \gamma_{\text{refine}} \nabla_H \mathcal{R}(Z_k, H_k; \mathbf{w}_k), \tag{19}$$

where $\mathcal{R}(Z, H; \mathbf{w}) = \sum_{j=1}^4 w_j r_j(Z, H)$, with r_j and w_j denoting the j-th entries of the reward vector $\mathbf{r}$ and weight vector $\mathbf{w}$, respectively, $\gamma_{\text{refine}} = 10^{-3}$ is the refinement step size. Gradient clipping optionally stabilizes the updates. The update of H in Eq. equation 19 is applied only in co-design. In FixSeq, $H_{k+1} = H_k = H_0$ for every refinement step, so the provided amino-acid sequence remains fixed and only Z is optimized. Non-generated receptor contexts are clamped after each step.

Gradient-Guided Steering Alignment. Backpropagating through the unrolled refinement trajectory ($K = 10$) is memory intensive and can destabilize gradients. We therefore optimize the weight network ϕ through an implicit steering signal. Let $\chi_H = \mathbb{I}[\text{CoDesign}]$ indicate whether the scalar latent is actively refined. The aggregate directional pressure is computed only over task-active latent variables:

$$s_{\text{steer}} = \frac{1}{K}\sum_{k=0}^{K-1} \frac{\|\nabla_{Z_k} \mathcal{R}(Z_k, H_k; \mathbf{w}_k)\|_2 + \chi_H \|\nabla_{H_k} \mathcal{R}(Z_k, H_k; \mathbf{w}_k)\|_2}{1 + \chi_H}. \tag{20}$$

Here, each $\|\cdot\|_2$ is the Euclidean norm of the corresponding vectorized gradient tensor. Accordingly, both Z- and H-directional pressures contribute in co-design, whereas FixSeq uses only the Z-gradient because $H_k \equiv H_0$ is clamped. To evaluate outcomes independently of the dynamic weight scale, we compute the intrinsic reward over the same task-active reward set:

$$\mathcal{R}_{\text{eval}} = \frac{1}{|\mathcal{A}_{\text{task}}|} \sum_{j\in\mathcal{A}_{\text{task}}} \tilde{r}_j(Z_K, H_K), \tag{21}$$

where $\tilde{r}_j$ is the z-score-normalized reward using training-set statistics. Thus, r_{rmsd} is excluded whenever it is task-masked. With the scalar exponential moving baseline $b \leftarrow 0.99b + 0.01\mathcal{R}_{\text{eval}}$, the steering loss is

$$\mathcal{L}_{\text{steer}}(\phi) = -\lambda_{\text{steer}}\, s_{\text{steer}}\, (\mathcal{R}_{\text{eval}} - b), \tag{22}$$

where $\lambda_{\text{steer}} = 0.01$. Gradients reach ϕ through s_{steer} and the predicted reward weights, while $\mathcal{R}_{\text{eval}}$ and b are detached. This rewards strong steering when the refined candidate exceeds the moving baseline and suppresses over-steering otherwise. The joint objective is:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{LDM}} + \mathcal{L}_{\text{steer}}(\phi). \tag{23}$$

The alignment update is applied every $F = 200$ diffusion iterations. At inference, the trained g_ϕ performs the same task-aware refinement without access to the ground-truth conformation, optimization details are provided in Appendix A.1.

3 EXPERIMENT

3.1 PREPROCESSING OF DATASET

We evaluate PepARD on two widely recognized benchmark datasets for peptide-protein complexes: PepBench (Kong et al., 2024) and PepBDB (Wen et al., 2019). PepBench provides 6,105 curated high-quality structures, from which the independent LNR (Low Non-Redundant) test set, consisting of 93 complexes sharing $< 30\%$ sequence identity with any training sample, evaluates out-of-distribution generalizability under low sequence redundancy. PepBDB is a comprehensive structural repository comprising 13,301 experimentally determined crystal structures, partitioned into training, validation, and test subsets with a strict 30% sequence identity cutoff to eliminate data leakage. Full details of pocket extraction, atom padding, and non-standard residue handling are given in Appendix C.

3.2 EVALUATION METRICS

To systematically evaluate generative quality, physical plausibility, and structural fidelity, we employ a comprehensive suite of metrics across both tasks. Exploration capacity is quantified by sequence diversity (SeqDiv, ↑) and structural diversity (StrDiv, ↑), interface interaction fidelity by contact consistency (Con, ↑), thermodynamic stability by predicted binding free energy (ΔG, ↓) under Rosetta scoring, and overall design quality by the success rate (Success, ↑). For binding conformation generation, we report backbone RMSD ($\text{RMSD}_{\text{C}_\alpha}$, ↓), all-heavy-atom RMSD ($\text{RMSD}_{\text{atom}}$, ↓), and DockQ (↑) against crystallized ground-truth structures. Precise definitions of all eight metrics are provided in Appendix C.

3.3 SEQUENCE-STRUCTURE CO-DESIGN

3.3.1 ABLATION STUDY

AMEGNN is PepARD's geometric backbone, coupling residue features and full-atom coordinates through $E(3)$-equivariant multi-channel message passing. Because its geometry supports both diffusion and refinement, we begin with component ablations (Table 1). Replacing the equivariant updates sharply reduces structural diversity and PepBDB success, confirming AMEGNN's role in geometric generalization. Removing refinement preserves competitive sequence diversity and contact consistency but weakens structural diversity, ΔG, and success, indicating that diffusion alone is insufficient. Uniform reward weights likewise degrade structural and energetic metrics, showing the benefit of adaptive objective balancing. Overall, the complete model provides the most consistent diversity-fidelity trade-off and the highest success rate on both datasets.

Table 1: Ablation study results for the co-design task, evaluated on the PepBench and PepBDB datasets. (**Best**, <u>Second best</u>)

Model	PepBench					PepBDB				
	SeqDiv(↑)	StrDiv(↑)	Con(↑)	ΔG(↓)	Success(↑)	SeqDiv(↑)	StrDiv(↑)	Con(↑)	ΔG(↓)	Success(↑)
w/o $E(3)$-Equiv. AMEGNN	0.879	0.336	0.722	-11.14	0.347	0.921	0.698	0.879	-13.38	0.218
w/o Refinement	0.902	0.388	<u>0.778</u>	<u>-14.67</u>	0.390	**0.960**	0.724	**0.935**	<u>-14.65</u>	<u>0.320</u>
w/o Adaptive Weights	<u>0.910</u>	<u>0.405</u>	0.762	-11.82	<u>0.399</u>	0.941	<u>0.744</u>	<u>0.934</u>	-12.63	0.276
PepARD	**0.933**	**0.637**	**0.791**	**-19.90**	**0.409**	<u>0.947</u>	**0.821**	0.928	**-17.87**	**0.502**

3.3.2 COMPARISON

We benchmark PepARD against leading generative models in Table 2, averaging results over three runs. PepFGLD did not report diversity metrics in its original *de novo* evaluation (−). Across both benchmarks, PepARD achieves the best balance of sequence and structural diversity with strong contact consistency. UniMoMo reports a lower nominal ΔG on PepBench (-34.35 kcal/mol), but this coincides with severe atomic clashes ($> 45\%$ vs. $< 5\%$ for PepARD) and low contact fidelity (0.439). Although PepFGLD attains a higher PepBench success rate (0.594) with more conservative, low-diversity templates, PepARD generalizes better to the larger PepBDB benchmark, achieving the highest success rate (0.502, a 61.9% gain over PepTri) while preserving conformational novelty.

Table 2: Performance comparison of our model against baselines on the PepBench and PepBDB datasets for the co-design task. (**Best**, Second best)

Model	PepBench					PepBDB				
	SeqDiv(↑)	StrDiv(↑)	Con(↑)	ΔG(↓)	Success(↑)	SeqDiv(↑)	StrDiv(↑)	Con(↑)	ΔG(↓)	Success(↑)
PepFlow (Li et al., 2024)	0.830	0.440	0.670	-17.05	0.310	0.670	0.580	0.740	-17.05	0.300
PepGLAD (Kong et al., 2024)	0.901	0.342	0.789	-15.63	0.290	0.921	0.813	0.789	-14.48	0.150
UniMoMo (Kong et al., 2025b)	0.920	0.620	0.439	**-34.35**	0.340	0.900	0.810	0.920	**-18.89**	0.440
PepTri (Nguyen et al., 2026)	0.930	0.620	0.790	-19.39	0.400	0.930	0.630	0.890	-18.15	0.310
PepFGLD (Niu et al., 2026)	-	-	**0.816**	-19.86	**0.594**	-	-	**0.948**	-18.74	0.497
PepARD	**0.933**	**0.637**	0.791	-19.90	0.409	**0.947**	**0.821**	0.928	-17.87	**0.502**

3.4 Binding Conformation Generation

3.4.1 Ablation Study

Table 3 tests the same components for FixSeq. Replacing AMEGNN's equivariant updates increases backbone RMSD and lowers DockQ on both datasets, confirming its importance for receptor-bound geometry. Removing refinement also worsens backbone RMSD and DockQ, with smaller changes in all-atom RMSD because refinement acts on the compact coordinate latent. Uniform reward weights further reduce DockQ, supporting task-adaptive objective balancing. The complete model therefore achieves the best backbone accuracy and DockQ on both datasets, although its PepBDB all-atom RMSD is slightly higher than the static-weight variant, reflecting a trade-off between global pose quality and decoder-reconstructed side-chain detail.

Table 3: Ablation study results for the FixSeq task, evaluated on the PepBench and PepBDB datasets. (**Best**, Second best)

Model	PepBench			PepBDB		
	RMSD_{C_α} (↓)	$\text{RMSD}_{\text{atom}}$ (↓)	DockQ (↑)	RMSD_{C_α} (↓)	$\text{RMSD}_{\text{atom}}$ (↓)	DockQ (↑)
w/o $E(3)$-Equiv. AMEGNN	5.20	6.34	0.524	10.79	9.96	0.272
w/o Refinement	4.96	5.63	0.526	8.92	9.80	0.352
w/o Adaptive Weights	4.79	5.95	0.590	8.57	**9.53**	0.366
PepARD	**4.11**	**5.49**	**0.643**	**8.49**	9.62	**0.386**

3.4.2 Comparison

We compare PepARD with complex modeling and docking baselines in Table 4. On PepBench, PepARD achieves the lowest backbone deviation ($\text{RMSD}_{C_\alpha} = 4.11$ Å) and highest DockQ (0.643), outperforming PepFGLD (0.607) and HSRN (0.508). On PepBDB, it obtains the second-lowest backbone error (8.49 Å vs. PepMimic's 8.41 Å) and a competitive DockQ of 0.386 versus HSRN's 0.394. PepFGLD retains a slightly lower all-atom RMSD (9.45 Å vs. 9.62 Å). Overall, PepARD achieves competitive binding-mode recovery across diverse pocket geometries without post-hoc docking.

Table 4: Performance comparison of our model against baselines on the PepBench and PepBDB datasets for the FixSeq task. (**Best**, Second best)

Model	PepBench			PepBDB		
	RMSD_{C_α} (↓)	$\text{RMSD}_{\text{atom}}$ (↓)	DockQ(↑)	RMSD_{C_α} (↓)	$\text{RMSD}_{\text{atom}}$(↓)	DockQ(↑)
HSRN (Jin et al., 2022)	6.02	7.59	0.508	9.28	9.72	**0.394**
dyMEAN (Kong et al., 2023)	7.96	8.35	0.374	17.64	17.56	0.142
PepGLAD (Kong et al., 2024)	4.45	5.63	0.558	11.02	11.99	0.290
PepMimic (Kong et al., 2025a)	5.12	6.33	0.520	**8.41**	9.49	0.371
PepFGLD (Niu et al., 2026)	4.14	**5.41**	0.607	8.86	**9.45**	0.345
PepARD	**4.11**	5.49	**0.643**	8.49	9.62	0.386

3.5 CASE STUDY

While benchmark metrics quantify global generative fidelity, evaluating whether designed peptides are biophysically viable requires granular, property-level characterization. We therefore examine PepARD on a demanding disease-relevant target: the X-linked lymphoproliferative (XLP) protein SAP bound to the 11-residue cytoplasmic peptide of SLAM (PDB: 1d4t, KSLTIYAQVQK), whose shallow, solvent-exposed interface presents a stringent test for surface complementarity.

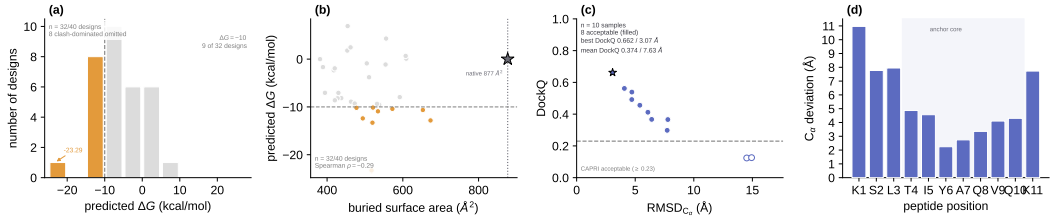


Figure 2: SLAM/SAP case study. **(a)** Predicted ΔG for the 32 non-clashing candidates among 40 generated Co-Design samples, eight clash-dominated candidates are omitted. **(b)** Buried surface area vs. predicted ΔG. **(c)** FixSeq RMSD$_{C\alpha}$ vs. DockQ. **(d)** Median FixSeq per-residue deviation.

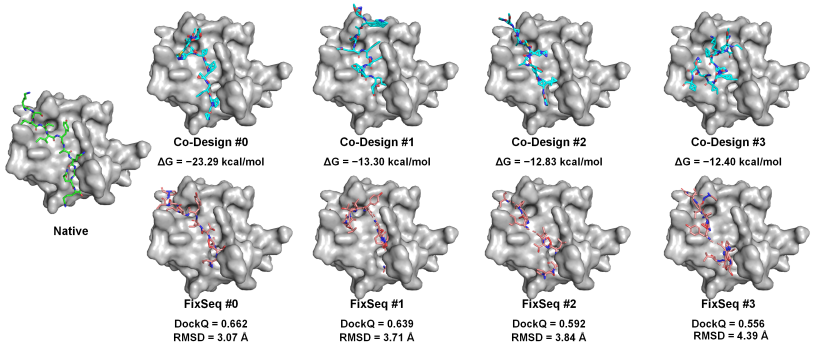


Figure 3: Structural visualizations of the four labeled candidates from each design mode, together with the native complex. Co-Design panels report predicted ΔG (kcal/mol), FixSeq panels report DockQ and RMSD (Å). Candidate indices #0-#3 follow the order displayed.

In unconstrained co-design ($n = 40$), PepARD explores diverse chemical spaces with low sequence identity to SLAM. After relaxation and clash filtering, 32 candidates remain, 9 of these have predicted $\Delta G < -10\,\mathrm{kcal/mol}$ (Figure 2 (a)). Buried surface area shows a modest negative correlation with predicted ΔG (Spearman $\rho = -0.29$, Figure 2 (b)). In the FixSeq setting ($n = 10$), 8 samples meet the CAPRI-acceptable threshold of DockQ $\geq$ 0.23 (Figure 2 (c)). The central Thr4-Gln10 region has lower per-residue deviation than the peptide termini in the median pose (Figure 2 (d)). Representative structures and their annotated scores are shown in Figure 3, further analyses appear in Appendix D.

4 CONCLUSION

We proposed PepARD, a geometric latent diffusion framework for target-specific full-atom peptide design and binding conformation prediction. PepARD combines dual-stream latents, AMEGNN, and adaptive reward-steered refinement to model variable heavy-atom geometry and balance competing objectives. Its receptor-conditioned generative process supports both *de novo* co-design and fixed-sequence conformation generation within a shared framework. Results on PepBench and PepBDB demonstrate diverse, physically plausible co-designs and improved fixed-sequence pose recovery. Ablation studies further show that equivariant modeling and adaptive refinement provide complementary gains across evaluation metrics. Future work will incorporate receptor flexibility and non-canonical amino acids.

AI USE STATEMENT

As declared in the submission, artificial intelligence tools were utilized in this work to aid and polish the writing. Specifically, these tools were employed to refine grammar, enhance sentence flow, and improve the overall readability of the manuscript. The authors meticulously reviewed, edited, and verified all AI-assisted text. All scientific concepts, experimental designs, and conclusions remain entirely the original work of the human authors, who take full responsibility for the final content and integrity of this paper.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.

Johannes Brandstetter, Rob Hesselink, Elise van der Pol, Erik J Bekkers, and Max Welling. Geometric and physical quantities improve e(3) equivariant message passing. In *International Conference on Learning Representations*, 2022.

Yaoyu Chen, Xiaoli Lin, Ziyi Gong, and Jun Pang. TSMGen: Target-specific molecule generation via higher-order structural dependencies and context-aware bidirectional fusion. In *Forty-third International Conference on Machine Learning*, 2026.

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi S. Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. In *The Eleventh International Conference on Learning Representations*, 2023.

Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, and Jianzhu Ma. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. In *The Eleventh International Conference on Learning Representations*, 2023.

Emiel Hoogeboom, Vıctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887. PMLR, 2022.

Xinjian Ji, Alexander L Nielsen, and Christian Heinis. Cyclic peptides for drug development. *Angewandte Chemie International Edition*, 63(3):e202308251, 2024.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Antibody-antigen docking and design via hierarchical structure refinement. In *International Conference on Machine Learning*, pp. 10217–10227. PMLR, 2022.

Bowen Jing, Gabriele Corso, Jeffrey Chang, Regina Barzilay, and Tommi Jaakkola. Torsional diffusion for molecular conformer generation. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, pp. 24240–24253, 2022.

Xiangzhe Kong, Wenbing Huang, and Yang Liu. End-to-end full-atom antibody design. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 17409–17429. PMLR, 23–29 Jul 2023.

Xiangzhe Kong, Yinjun Jia, Wenbing Huang, and Yang Liu. Full-atom peptide design with geometric latent diffusion. *Advances in Neural Information Processing Systems*, 37:74808–74839, 2024.

Xiangzhe Kong, Rui Jiao, Haowei Lin, Ruihan Guo, Wenbing Huang, Wei-Ying Ma, Zihua Wang, Yang Liu, and Jianzhu Ma. Peptide design through binding interface mimicry with pepmimic. *Nature biomedical engineering*, pp. 1–16, 2025a.

Xiangzhe Kong, Zishen Zhang, Ziting Zhang, Rui Jiao, Jianzhu Ma, Wenbing Huang, Kai Liu, and Yang Liu. Unimomo: Unified generative modeling of 3d molecules for de novo binder design. In *International Conference on Machine Learning*, pp. 31397–31418. PMLR, 2025b.

Leshan Lai, Yuansheng Liu, Bosheng Song, Keqin Li, and Xiangxiang Zeng. Deep generative models for therapeutic peptide discovery: a comprehensive review. *ACM Computing Surveys*, 57 (6):1–29, 2025.

Jiahan Li, Chaoran Cheng, Zuofan Wu, Ruihan Guo, Shitong Luo, Zhizhou Ren, Jian Peng, and Jianzhu Ma. Full-atom peptide design based on multi-modal flow matching. In *International Conference on Machine Learning*, pp. 27615–27640. PMLR, 2024.

Haitao Lin, Odin Zhang, Huifeng Zhao, Dejun Jiang, Lirong Wu, Zicheng Liu, Yufei Huang, and Stan Z Li. Ppflow: Target-aware peptide design with torsional flow matching. In *International Conference on Machine Learning*, pp. 30510–30528. PMLR, 2024.

Yuki Matsukiyo, Chikashige Yamanaka, and Yoshihiro Yamanishi. De novo generation of chemical structures of inhibitor and activator candidates for therapeutic target proteins by a transformer-based variational autoencoder and bayesian optimization. *Journal of Chemical Information and Modeling*, 64(7):2345–2355, 2024.

Alex Morehead, Lazar Atanackovic, Akshata Hegde, Yanli Wang, Frimpong Boadu, Joel Selvaraj, Alexander Tong, Aditi Krishnapriyan, and Jianlin Cheng. Flow matching for generative modelling in bioinformatics and computational biology. *Nature Machine Intelligence*, pp. 1–18, 2026.

Ngoc-Quang Nguyen, Jaeyoon Jung, Seijung Kim, Sunkyu Kim, and Jaewoo Kang. Peptri: Tri-guided all-atom diffusion for peptide design via physics, evolution, and mutual information. In *The Fourteenth International Conference on Learning Representations*, 2026.

Dongjiang Niu, Xiaofeng Wang, Zhiqiang Wei, and Zhen Li. Flexibility-aware geometric latent diffusion for full-atom peptide design. In *Forty-third International Conference on Machine Learning*, 2026.

Rui Qin, Zijie Chen, Yurong Li, Meijing Fang, Longji Shen, Yilong Su, Odin Zhang, Qinghan Wang, Qun Su, Jike Wang, et al. Revisiting target-aware de novo molecular generation with tarpass: Between rational design and texas sharpshooter. *Advanced Science*, pp. e75411, 2026.

Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Tiara Natasha Binte Sayuti, Kakuly Mittal, Tan Lai Heng, and Jagath C Rajapakse. Target-aware latent diffusion model for design of apoptosis-inducing anticancer peptides. *Computers in Biology and Medicine*, 199:111286, 2025.

Marcelo DT Torres, Leo 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), 2025.

Masatoshi Uehara, Xingyu Su, Yulai Zhao, Xiner Li, Aviv Regev, Shuiwang Ji, Sergey Levine, and Tommaso Biancalani. Reward-guided iterative refinement in diffusion models at test-time with applications to protein and dna design. In *International Conference on Machine Learning*, pp. 60515–60529. PMLR, 2025.

Yi Wang, Hui Tang, Lichao Huang, Lulu Pan, Lixiang Yang, Huanming Yang, Feng Mu, and Meng Yang. Self-play reinforcement learning guides protein engineering. *Nature Machine Intelligence*, 5(8):845–860, 2023.

Ziwen Wang, Jiajun Fan, Ruihan Guo, Thao Nguyen, Heng Ji, and Ge Liu. Proteinzero: Self-improving protein generation via online reinforcement learning. *arXiv preprint arXiv:2506.07459*, 2025.

Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

Zeyu Wen, Jiahua He, Huanyu Tao, and Sheng-You Huang. Pepbdb: a comprehensive structural database of biological peptide–protein interactions. *Bioinformatics*, 35(1):175–177, 2019.

Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. In *International Conference on Learning Representations*, 2022.

Silong Zhai, Tiantao Liu, Shaolong Lin, Dan Li, Huanxiang Liu, Xiaojun Yao, and Tingjun Hou. Artificial intelligence in peptide-based drug design. *Drug Discovery Today*, 30(2):104300, 2025.

Buke Zhang, Junjie Zhu, Hai-Feng Chen, and Lu-Ning Liu. Alphafold-based peptide structure prediction: Opportunities, limitations, and future directions. *Biotechnology Advances*, pp. 108977, 2026.

Peining Zhang, Daniel Baker, Minghu Song, and Jinbo Bi. Unraveling the potential of diffusion models in small-molecule generation. *Drug Discovery Today*, 30(7):104413, 2025a.

Renxin Zhang, Zhiyuan Chen, Shuhan Li, Haohao Lv, Jinjun Li, Naixue Yang, and Shaoxing Dai. Proteome-wide identification and comparison of drug pockets for discovering new drug indications and side effects. *Molecules*, 30(2):260, 2025b.

Dengwei Zhao, Jingyuan Zhou, Shikui Tu, and Lei Xu. De novo drug design by multi-objective path consistency learning with beam a* search. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 21(6):2459–2470, 2024.

A Additional Method Details

A.1 Adaptive Reward Network Details

The Adaptive Reward Weight Network conditions its prediction on the refinement iteration and a compact summary of the current latent state,

$$\mathbf{s}_{Z,H} = \left[\overline{\|Z\|},\ \mathrm{std}(\|Z\|),\ \hat{c}_{\mathrm{clash}},\ \overline{\|H\|},\ \mathrm{std}(\|H\|),\ \mathrm{std}_{\mathrm{flatten}}(H),\ L/100\right] \in \mathbb{R}^7, \tag{24}$$

Here, $\overline{\|Z\|}$ and $\mathrm{std}(\|Z\|)$ are the mean and standard deviation of active-channel coordinate norms, $\overline{\|H\|}$ and $\mathrm{std}(\|H\|)$ are the corresponding statistics of per-residue feature norms, $\mathrm{std}_{\mathrm{flatten}}(H)$ is the standard deviation over all scalar entries of H, and $L/100$ is the peptide length normalized by the fixed scale 100. The conflict summary $\hat{c}_{\mathrm{clash}}$ estimates the fraction of sampled active-channel pairs separated by less than τ_{clash}. The network maps the seven statistics to a 128-dimensional state encoding, concatenates it with a 128-dimensional step encoding, and applies a $256 \to 128 \to 64$ MLP and projection head. The task mask is $\mathbf{M}_{\mathrm{task}} = [-\infty, 0, 0, 0]^\top$ whenever no reference conformation is available, including co-design and target-blind FixSeq inference, and $\mathbf{0}$ during FixSeq training. Projection-head biases are initialized to $[0.0, 0.4, 0.3, 0.3]^\top$ for co-design, where the first dimension is masked, and $[0.5, 0.2, 0.15, 0.15]^\top$ for FixSeq training.

For steering alignment, $\mathcal{R}_{\mathrm{eval}}$ and the exponential moving baseline are treated as detached scalars, so gradients reach ϕ only through the steering magnitude s_{steer} and the predicted reward weights. The objective therefore assigns credit when stronger steering coincides with above-baseline refined outcomes and suppresses excessive steering otherwise. The alignment update is executed every $F = 200$ diffusion iterations. In co-design, both latent streams contribute to the directional-pressure term, in FixSeq, the sequence latent is fixed and only the Z-gradient contributes.

A.2 Target-Blind FixSeq Inference

At test time, FixSeq receives only the receptor pocket and prescribed peptide sequence. The crystallized peptide conformation X_{target} is not provided to the sequence encoder, denoiser, Adaptive Reward Weight Network, or latent refinement procedure. The sequence-only invariant branch produces $H_0 = \boldsymbol{\mu}_H(S)$, so constructing the fixed sequence condition requires neither X_{target} nor generated peptide coordinates. This latent remains clamped as $H_t = H_k = H_0$ throughout denoising and all refinement steps, while only Z is generated and optimized. The structure-fidelity reward r_{rmsd} is used only as an auxiliary signal during FixSeq training, at inference it is masked, and the remaining active reward weights are renormalized to sum to one. After refinement, the decoder maps (H_0, Z_K) to the full-atom peptide structure. Only then is X_{target} accessed to compute RMSD and DockQ. By contrast, in de novo co-design both H and Z are generated and refined before decoding.

A.3 Latent-Space Reward Proxies and Full-Atom Decoding

The four refinement objectives in Section 2.4 are computed on the compact latent variables (H, Z) rather than directly on the decoded all-heavy-atom coordinates. In particular, the clash and bond terms operate on the $C_z = 3$ coordinate-latent channels and should therefore be interpreted as latent-space geometric proxies that regularize the refinement trajectory. After K refinement steps, the pretrained decoders map (H_K, Z_K) to the peptide sequence and full-atom coordinates. Full-atom structural and interface quality is then evaluated after decoding and Rosetta FastRelax using all-heavy-atom RMSD, clash ratio, contact consistency, and Rosetta/FoldX scores. Thus, "full-atom" refers to the reconstructed model output and its downstream evaluation, whereas adaptive reward steering is performed in the compact latent space.

B Algorithm Summaries

This section summarizes the algorithmic workflows for the two core components of PepARD: the equivariant graph neural layer (Section B.1) and the reward-guided training procedure (Section B.2). Notation follows Section 2.

B.1 AMEGNN Layer with RollerPooling

Algorithm 1 details one propagation step of the $E(3)$-equivariant AMEGNN layer, which is shared by both the autoencoder and the denoiser.

Algorithm 1 AMEGNN Layer with RollerPooling (one propagation step)

Require: node scalars $\{H_i\}$, coordinate tensors $\{Y_i \in \mathbb{R}^{C\times 3}\}$, validity masks $\{m_i \in \{0,1\}^C\}$ with $n_i = \sum_c m_i^c$, update masks $\{u_i\}$ ($u_i = \mathbf{0}$ for receptor nodes), edge set $\mathcal{E}$, edge-type attributes E_{ij}^{attr}

Ensure: updated $\{H_i\}$, $\{Y_i\}$

for all $(i,j) \in \mathcal{E}$ **do**
  $D_{ij}^{ab} \leftarrow \|Y_i^a - Y_j^b\|_2$ for all channel pairs $a, b \in \{1,\dots,C\}$
  $R_{ij} \leftarrow \mathrm{RBF}(D_{ij}) \oplus E_{\text{atom}}$
  $M_{ij} \leftarrow \mathrm{MLP}_{\text{edge}}\big([\, H_i \,\|\, H_j \,\|\, R_{ij} \,\|\, E_{ij}^{\text{attr}} \,]\big)$
  $m_{ij} \leftarrow \mathrm{RollerPool}(M_{ij}; P_{n_i,n_j})$
end for
$H_i \leftarrow H_i + \mathrm{MLP}_{\text{node}}\big([\, H_i \,\|\, \sum_{j\in\mathcal{N}(i)} m_{ij} \,]\big)$
$Y_i \leftarrow Y_i + u_i \odot \sum_{j\in\mathcal{N}(i)} (Y_i - Y_j) \odot \mathrm{MLP}_{\text{coord}}(m_{ij})$
return $\{H_i\}$, $\{Y_i\}$

B.2 Adaptive Reward-Steered Training

Algorithm 2 outlines the full training procedure of PepARD. Here, H and Z denote the invariant and $SE(3)$-equivariant latents ($d = 8$, $C_z = 3$), (X_r, H_r) the receptor pocket, $T = 100$ the number of diffusion steps, $K = 10$ the number of refinement steps, and $\mathcal{A}_{\text{task}}$ the set of task-active rewards.

C Datasets, Experimental Setup, and Evaluation Metrics

C.1 Data preprocessing.

Raw PDB files are parsed using Biopython, and receptor binding pockets are extracted via a two-stage spatial criterion: interface residues with any heavy atom within 6.0 Å of the peptide are first identified, and the pocket boundary is subsequently expanded to include all receptor residues within 10.0 Å of the peptide C_β (or C_α for glycine) atoms. In structural tensor modeling, each residue is parameterized as a structured block containing up to 14 heavy atoms, missing atoms are padded with the centroid coordinate of valid atoms and masked out during loss calculation, while non-standard amino acids are tokenized as `UNK` with their native 3D coordinates retained for geometric modeling. PepBench is complemented by ProtFrag for structural data augmentation.

C.2 Evaluation metrics.

For the sequence-structure co-design task, exploration capacity is quantified by sequence diversity (SeqDiv, ↑), defined as the mean pairwise normalized Levenshtein distance, and structural diversity (StrDiv, ↑), computed as the average pairwise C_α-RMSD across generated candidates. Interface interaction fidelity is measured by contact consistency (Con, ↑), which computes the fraction of native receptor-peptide contact pairs (heavy-atom distance < 5.0 Å) preserved in the generated complex. Binding thermodynamic stability is assessed via predicted binding free energy (ΔG, ↓) using Rosetta and FoldX force-field scoring functions, alongside the overall design success rate (Success, ↑), denoting the proportion of candidates that simultaneously surpass native binding affinity ($\Delta G \le \Delta G_{\text{native}}$) and maintain low steric conflict (clash ratio < 0.15). For the binding conformation generation task, structural prediction precision is evaluated by backbone root-mean-square deviation (RMSD_{C_α}, ↓) and all-heavy-atom RMSD ($\text{RMSD}_{\text{atom}}$, ↓) against crystallized ground-truth structures, with DockQ (↑) serving as the standardized benchmark score measuring overall protein-peptide docking pose quality.

Algorithm 2 Adaptive Reward-Steered Training of PepARD

Require: frozen pre-trained autoencoder; denoiser $\boldsymbol{\epsilon}_\theta$; reward weight network g_ϕ; task $\in \{\text{CoDesign}, \text{FixSeq}\}$; $T = 100$, $K = 10$, λ_E, $\lambda_{\text{steer}} = 0.01$, $\gamma_{\text{refine}} = 10^{-3}$, $F = 200$

Ensure: trained parameters θ, ϕ

while not converged **do**
  encode a mini-batch with the frozen autoencoder to obtain (H_0, Z_0), with $H_0 = \boldsymbol{\mu}_H(S)$ in FixSeq
  $t \sim \mathcal{U}\{1, \dots, T\}$, $\boldsymbol{\epsilon}_H, \boldsymbol{\epsilon}_Z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$; form the perturbed state (H_t, Z_t) (Eq. 6)
  $(\hat{\boldsymbol{\epsilon}}_H, \hat{\boldsymbol{\epsilon}}_Z) \leftarrow \boldsymbol{\epsilon}_\theta(H_t, Z_t, \boldsymbol{\tau}_t, \mathbf{p}_{\text{seq}}; X_r, H_r)$
  $\mathcal{L}_{\text{LDM}} \leftarrow \|\hat{\boldsymbol{\epsilon}}_Z - \boldsymbol{\epsilon}_Z\|_F^2 + \lambda_H \|\hat{\boldsymbol{\epsilon}}_H - \boldsymbol{\epsilon}_H\|_F^2$
  if iteration $\bmod F = 0$ **then**
    $Z_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$; $H_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in CoDesign, $H_T \leftarrow H_0$ in FixSeq
    for $t = T$ down to 1 **do**
      $(\hat{\boldsymbol{\epsilon}}_H, \hat{\boldsymbol{\epsilon}}_Z) \leftarrow \boldsymbol{\epsilon}_\theta(H_t, Z_t, \boldsymbol{\tau}_t, \mathbf{p}_{\text{seq}}; X_r, H_r)$
      $\hat{\boldsymbol{\epsilon}}_Z \leftarrow \hat{\boldsymbol{\epsilon}}_Z + \lambda_E \sqrt{1 - \bar{\alpha}_t}\, \nabla_{Z_t} E(Z_t)$
      $(H_{t-1}, Z_{t-1}) \leftarrow$ ancestral DDPM step (Eq. 9)
      if FixSeq **then**
        $H_{t-1} \leftarrow H_0$
      end if
    end for
    $Z_0 \leftarrow Z_0 \cdot \sigma_{\text{coord}} + \bar{X}_r^{C_\alpha}$; $\chi_H \leftarrow \mathbb{I}[\text{CoDesign}]$; $s_{\text{steer}} \leftarrow 0$
    for $k = 0$ to $K - 1$ **do**
      $\mathbf{w}_k \leftarrow \operatorname{softmax}\big(f_\phi(\boldsymbol{\tau}_k, \mathbf{s}_{Z_k, H_k}) + \mathbf{M}_{\text{task}}\big)$
      $g_k^Z \leftarrow \nabla_{Z_k} \mathcal{R}(Z_k, H_k; \mathbf{w}_k)$; if $\chi_H = 1$ also $g_k^H \leftarrow \nabla_{H_k} \mathcal{R}(Z_k, H_k; \mathbf{w}_k)$
      $s_{\text{steer}} \leftarrow s_{\text{steer}} + \dfrac{\|g_k^Z\|_2 + \chi_H \|g_k^H\|_2}{1 + \chi_H}$
      $Z_{k+1} \leftarrow Z_k + \gamma_{\text{refine}} g_k^Z$; if $\chi_H = 1$ also $H_{k+1} \leftarrow H_k + \gamma_{\text{refine}} g_k^H$
    end for
    $s_{\text{steer}} \leftarrow s_{\text{steer}} / K$
    $\mathcal{R}_{\text{eval}} \leftarrow \frac{1}{|\mathcal{A}_{\text{task}}|} \sum_{j \in \mathcal{A}_{\text{task}}} \tilde{r}_j(Z_K, H_K)$
    $b \leftarrow 0.99\, b + 0.01\, \mathcal{R}_{\text{eval}}$
    $\mathcal{L}_{\text{steer}} \leftarrow -\lambda_{\text{steer}}\, s_{\text{steer}} \big[\, \text{sg}(\mathcal{R}_{\text{eval}}) - \text{sg}(b) \,\big]$
  else
    $\mathcal{L}_{\text{steer}} \leftarrow 0$
  end if
  $\mathcal{L}_{\text{total}} \leftarrow \mathcal{L}_{\text{LDM}} + \mathcal{L}_{\text{steer}}$
  update θ, ϕ by one optimizer step on $\mathcal{L}_{\text{total}}$
end while
return θ, ϕ

C.3 Implementation Details

Environment. The model was implemented in Python 3.9 using PyTorch 1.13.1 with CUDA 11.7. All experiments were conducted on an Ubuntu 24.04 server equipped with two NVIDIA RTX A6000 GPUs (48GB VRAM each).

Training. The LDM is trained with AdamW ($\mathrm{lr} = 10^{-4}$, default weight decay 0.01) and a ReduceLROnPlateau scheduler (factor 0.6, patience 3, $\min_\mathrm{lr} = 5 \times 10^{-6}$) for at most 1000 (FixSeq) / 500 (Co-Design) epochs, with validation every 10 epochs, early stopping after 10 non-improving validations, and the top-10 checkpoints retained (selection criterion: validation RMSD for FixSeq, validation loss for Co-Design). No gradient clipping, learning-rate warmup, or weight EMA is used; batches are composed dynamically under a token-complexity budget of 6×10^4.

D Extended Case Study: In-Depth Physicochemical Analysis (1d4t)

In this section, we expand upon the case study on the SLAM/SAP complex (PDB: 1d4t) presented in Section 3.5, providing a comprehensive characterization of the physicochemical properties, contact metrics, and radar profile comparisons between the designed peptides and the wild-type ligand.

D.1 Energetics and contact recovery.

Of the 40 generated Co-Design candidates, 32 remain after relaxation and clash filtering, and 9 of these 32 have predicted $\Delta G < -10\,\mathrm{kcal/mol}$ (Figure 2a). The four Co-Design examples displayed in Figure 3 are annotated with predicted ΔG values of -23.29, -13.30, -12.83, and $-12.40\,\mathrm{kcal/mol}$ for candidates #0-#3, respectively. These are individual figure annotations, not a mean over the 40 generated designs. At a 4.5 Å heavy-atom cutoff, the top-scoring design retains only 6 of the 50 native pocket contacts, forms 23 de novo contacts, and buries $517\,\text{Å}^2$, indicating an interaction pattern distinct from the native complex. Among the 32 evaluated candidates, buried surface area has a modest negative association with predicted binding score (Spearman $\rho = -0.29$, Figure 2b). The Co-Design mean of $499 \pm 73\,\text{Å}^2$ is below the native buried area of $877\,\text{Å}^2$ (Table 5), polar contacts are likewise lower (8.5 ± 5.3 vs. 12). The score and interface measurements describe predicted structures and do not establish experimental affinity.

D.2 FixSeq pose recovery.

For the ten FixSeq samples, Figure 2c reports 8 CAPRI-acceptable poses (DockQ ≥ 0.23), a mean DockQ of 0.374, and a mean RMSD_{C_α} of 7.63 Å. Figure 3 separately labels its four displayed FixSeq examples #0-#3 with DockQ values 0.662, 0.639, 0.592, and 0.556 and corresponding RMSD values 3.07, 3.71, 3.84, and 4.39 Å, respectively. Table 5 preserves their order and keeps these individual labels separate from the ten-sample summary. The median-pose per-residue profile in Figure 2d places the smallest deviation near the central anchor region (Tyr6: 2.24 Å) and the largest at Lys1 (10.96 Å), indicating that terminal flexibility contributes substantially to the observed pose error.

D.3 Macroscopic and interface property summary.

Table 5 combines cohort-level physicochemical and interface properties with the individual scores printed in Figure 3. Because FixSeq holds the native sequence fixed, its sequence-derived properties match the native values. Co-Design candidates are similar in mean molecular weight (1251 ± 93 vs. 1278.5 Da) and hydrophobic moment (0.284 ± 0.181 vs. 0.315) but differ in isoelectric point (6.86 ± 3.18 vs. 10.22) and net charge (-0.08 ± 1.57 vs. $+2.00$). Both generated groups have smaller reported buried surface areas than the native complex (FixSeq: $443\,\text{Å}^2$, Co-Design: $499 \pm 73\,\text{Å}^2$, native: $877\,\text{Å}^2$). The Figure 3 rows list four examples per task and should not be interpreted as cohort averages or experimental binding measurements.

D.4 HYDROPHOBIC AND ELECTROSTATIC RE-BALANCING.

The native SLAM peptide is short and strongly basic, its $+2.00$ charge carried by the two lysines that flank the anchor core. The Co-Design arm does not reproduce this electrostatic identity: it approximately halves both the polar ($0.36 \rightarrow 0.17$) and the basic-residue ($0.18 \rightarrow 0.11$) fractions, introduces acidic residues where the native has none ($0.00 \rightarrow 0.12$), and devotes 21% of the sequence to proline and glycine, yielding a net-neutral average charge (Figure 4a,b). The re-balancing is not uniform across designs: candidates with $\Delta G \leq -10$ kcal/mol have more native-like mean hydropathy than the remaining candidates (GRAVY -0.01 vs. -0.43, aliphatic index 95.7 vs. 76.6). No single physicochemical descriptor strongly predicts the score (all $|\rho| \leq 0.29$, $p > 0.1$), with buried surface area showing the largest observed association ($\rho = -0.29$).

D.5 PHYSICOCHEMICAL PROFILES ACROSS THE TWO DESIGN ARMS.

Figure 5 compares eight physicochemical descriptors of the four best FixSeq and Co-Design samples predicted to bind ($\Delta G < 0$) against the native peptide. Because FixSeq poses a frozen sequence, its six sequence-derived descriptors (molecular weight, net charge, pI, GRAVY, hydrophobic moment and aliphatic index) coincide exactly with the native profile, and the two arms separate only on the interface descriptors: FixSeq loses 18-34% of the native buried surface area (580-716 Å^2 vs. 877 Å^2) and up to two thirds of the polar contacts. Co-Design departs from the native profile along the sequence axes as well: its hydropathy spans 1.9 units (-1.77 to $+0.15$, against the native -0.30), whereas the four designs stay close to the native isoelectric point (9.75-10.13 vs. 10.22), suggesting that charge character appears to be far more constrained than hydropathy once binding is enforced. Notably, the buried surface area stays below the native value in every panel, identifying interface size, rather than sequence-level polarity, as the descriptor on which both arms fall short of the native peptide.

Table 5: **Properties of PepARD-designed peptides and the native SLAM peptide (1d4t).** The upper block summarizes the indicated cohorts, the lower block transcribes the four examples labeled in Figure 3 in candidate order #0-#3. Predicted ΔG values are Rosetta scores for relative ranking, not experimental affinities. A dash denotes a metric not reported for that group.

Property	**Native**	**FixSeq** ($n = 10$)	**Co-Design** ($n = 40$)
Peptide length	11	11	11
Molecular weight (Da)	1278.5	1278.5	1251 ± 93
Isoelectric point (pI)	10.22	10.22	6.86 ± 3.18
Net charge (pH 7.0)	$+2.00$	$+2.00$	-0.08 ± 1.57
GRAVY	-0.30	-0.30	-0.34 ± 1.03
Hydrophobic moment (μH)	0.315	0.315	0.284 ± 0.181
Buried surface area (Å^2)	877	443	499 ± 73
Polar contacts	12	11	8.5 ± 5.3
Figure 3 annotations for candidates #0-#3			
Predicted ΔG (kcal/mol)	-	-	$-23.29,\ -13.30,\ -12.83,\ -12.40$
DockQ	-	$0.662,\ 0.639,\ 0.592,\ 0.556$	-
RMSD (Å)	-	$3.07,\ 3.71,\ 3.84,\ 4.39$	-

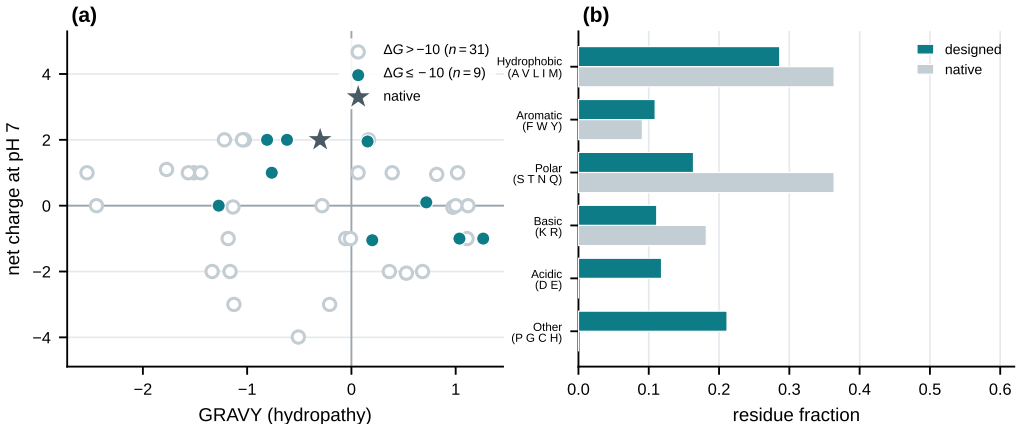


Figure 4: **Physicochemical properties of the** 40 **Co-Design sequences on 1D4T. (a)** Hydropathy versus net charge at pH 7, coloured by predicted ΔG, the native SLAM peptide is starred. **(b)** Amino-acid class composition of the Co-Design sequences versus the native peptide.

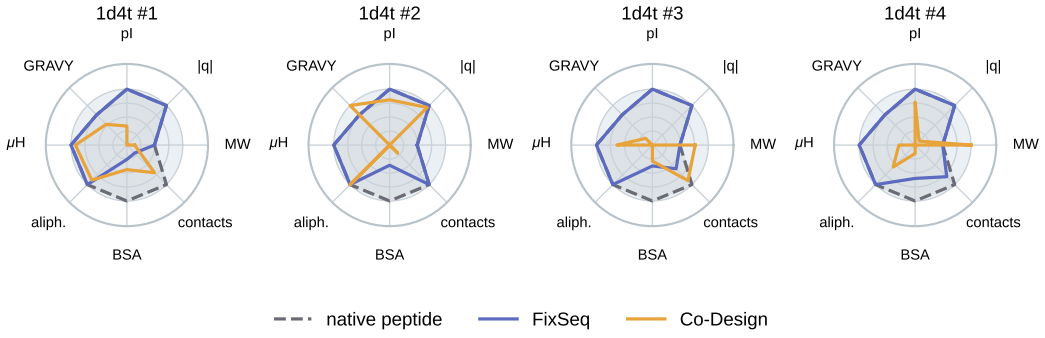


Figure 5: **Physicochemical profiles of design arms on 1D4T.** Eight properties of top-ranked FixSeq and Co-Design peptides compared to native SLAM (dashed line).