# LaCPlow: Geometry-Preserving Partially Latent Flow Matching for Cyclic Peptide Generation

**Anonymous authors**
Paper under double-blind review

## Abstract

Cyclic peptides share much of their local stereochemistry with linear peptides, but cyclization alters their covalent topology and reshapes their global conformations through sparse, nonlocal chemical constraints. Satisfying these constraints depends on the relative placement and orientation of linked residues. We propose **LaCPlow**, a target-conditioned partially latent flow model that explicitly transports residue frames while encoding sequence and residue-local atom coordinates in a continuous latent representation. A representation-aligned joint vector field couples these components by predicting rotational, translational, and latent velocities. Because chemistry-specific specialization depends on the capabilities of its substrate, we first evaluate LaCPlow-Base without cyclization-specific supervision to assess its geometric modeling capacity. We then introduce LaCPlow-Cyc, which incorporates cyclization-specific supervision. LaCPlow-Cyc improves bond-forming geometry while largely preserving the stereochemical quality of the base model.

## 1 Introduction

Cyclic peptides retain the intrinsic properties of peptides while incorporating conformational restriction induced by cyclization, which can improve structural stability and target-binding specificity. These characteristics make cyclic peptides a promising therapeutic modality for targeting sites that are difficult to address with conventional small molecules or linear peptides (Fosgerau & Hoffmann, 2015; Ji et al., 2024; Thite et al., 2026).

Cyclic peptides impose chemical constraints on a small subset of residues, so cyclic peptide generative models must recognize these interactions either explicitly or implicitly. Existing cyclic peptide generative models address this requirement through divergent architectural paradigms. One line of work explicitly models geometric variables, keeping them directly accessible to the generative network (Wang et al., 2026; 2024; Li et al., 2024). However, the generative state spans heterogeneous spaces, so editing or optimizing it requires selecting the relevant modalities and defining operations appropriate to each domain. Conversely, latent-space methods compress part or all of the peptide into learned continuous representations (Jiang et al., 2025; Zhao et al., 2026; Zhang et al., 2026). Models designed for linear peptides such as PepGLAD (Kong et al., 2024) provide a continuous latent representation that can serve as a substrate for specialized tasks, including cyclic peptide generation. Such extensions depend on the generative capacity of the underlying substrate, while task-specific biases may trade off the specialized objective against other aspects of generated quality. In latent-based models, reconstruction fidelity is particularly relevant because stereochemically sensitive geometry must be recovered from the compressed representation. For cyclic peptide generation, a substrate intended for specialized extensions should therefore possess sufficient geometrical capabilities from the outset. Extensions addressing specific chemical objectives, whether through structural modification or optimization, should also preserve the substrate's geometric and stereochemical quality as far as possible. To meet these requirements, we propose LaCPlow, a target-conditioned partially latent flow model that jointly transports residue orientations $R$, $C_\alpha$ coordinates $x$, and a latent representation encoding sequence and residue-local atomistic coordinates. The explicit frames expose the global placement and orientation of each residue, while the latent variables encode the remaining residue-wise information. To capture dependencies among these components, we use a single representation-aligned, target-conditioned geometric denoiser to jointly predict



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

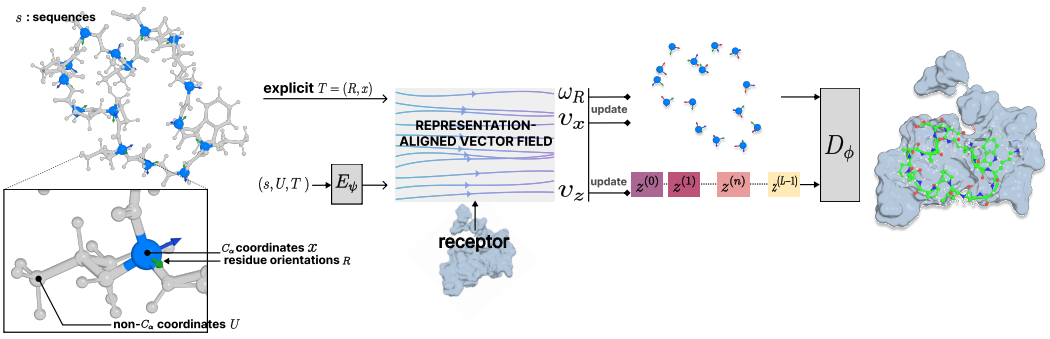



Figure 1: Overview of **LaCPlow**. The target-agnostic encoder maps $(s, U, T)$ to a residue-wise latent $z$, while the residue frames $T = (R, x)$ remain explicit generative variables. A target-conditioned representation-aligned joint vector field transports $(T, z)$ and jointly predicts rotational, translational, and latent velocities $(\omega_R, v_x, v_z)$. The frozen decoder reconstructs $(s, U) = \mathrm{D}_\phi(T, z)$.

translational, rotational, and latent velocities. We first assess how well LaCPlow models cyclic peptides without dedicated cyclization-specific supervision. We evaluate geometric quality before external energy minimization to separate model-generated structures from corrections introduced by subsequent relaxation. We then test whether LaCPlow can serve as a substrate when more precise local chemical objectives are required. To this end, we introduce LaCPlow-Cyc, which augments the latent representation with cyclization-specific information and adds a closure-specific reconstruction objective. This specialization substantially improves bond-forming geometry while largely preserving stereochemical quality. We also show that the backbone and torsional corrections during energy minimization are comparable to those of the base model.

**Main contributions.** **(i) A partially latent substrate for cyclic peptide generation, LaCPlow-Base.** We propose a generative substrate that explicitly transports residue frames while compressing sequence and residue-local atomistic information into a latent variable $z$. **(ii) Representation-aligned joint geometric flow.** We design a target-conditioned joint vector field that uses explicit residue frames for geometric interactions and predicts rotational, translational, and latent velocities according to their respective transformation laws. **(iii) Cyclization-specific specialization through LaCPlow-Cyc.** Using LaCPlow-Base as a substrate, we show that cyclization-specific supervision can substantially improve bond-forming geometry while largely preserving the stereochemical profile of the substrate.

## 2 Related Work

**Cyclic Generation under Limited Cyclic Peptide Data.** Early cyclic peptide modeling was developed under limited availability of cyclic peptide datasets, motivating approaches that leverage linear peptide data for cyclic peptide generation, such as a Harmonic-SDE-based approach (Zhou et al., 2025), zero-shot cyclic generation by composing geometric constraints (Jiang et al., 2025), and approaches that modify the relative positional encoding of AlphaFold2 (Jumper et al., 2021) for cyclic peptide modeling (Rettie et al., 2025).

**Cyclic-Peptide Generative Modeling and Property-Guided Modeling.** APCyc (Zhao et al., 2026) generates multi-objective-optimized cyclic peptides based on PepGLAD's latent diffusion model (LDM) architecture (Kong et al., 2024). Similarly, GeoCycler (Zhang et al., 2026) and CycDiff-DPO (Sun et al., 2026) apply preference or property optimization to CP-Composer (Jiang et al., 2025).

**Explicit, Latent, and Partially Latent Representations.** The recent release of large-scale cyclic peptide–protein complex datasets, such as CPSea (Yang et al., 2025), has enabled training target-binding cyclic peptide generative models from scratch. Explicit approaches directly model geometric states, including the flow-matching model PepFlow (Li et al., 2024) and the SE(3)-equivariant

diffusion model DiffPepBuilder (Wang et al., 2024), which follow the SE(3) principles of Yim et al. (2023; 2024). Latent approaches include PepGLAD (Kong et al., 2024), and partially latent flow-matching models include La-Proteina and Proteína-Complexa (Geffner et al., 2026; Didi et al., 2026).

# 3 METHOD

LaCPlow couples explicit residue-frame geometry with a continuous latent representation. Our design aligns compression, interactions, and transport with this representation. Local atomistic information and relative geometry are processed through rigid-motion-invariant features, and geometric velocities follow the transformation laws of the explicit frames. We describe LaCPlow-Base below; the cyclization-specific extension LaCPlow-Cyc is given in Appendix A.6.

## 3.1 PROBLEM FORMULATION AND REPRESENTATION

Given a target pocket $P$ and peptide length $L$, we model the conditional distribution $p(s, U, T \mid P)$ over amino-acid sequence $s$, residue-local non-$C_\alpha$ atom coordinates $U$, and residue frames $T$. Each residue has a frame $T_i = (R_i, x_i)$, where $x_i \in \mathbb{R}^3$ is its $C_\alpha$ position and $R_i \in SO(3)$ is defined by the $N$–$C_\alpha$–$C$ backbone. Using Atom37 (Jumper et al., 2021), we express an atom position $X_{i,a}$ in this frame as

$$U_{i,a} = R_i^\top (X_{i,a} - x_i), \qquad X_{i,a} = R_i U_{i,a} + x_i. \tag{1}$$

This decomposition separates residue placement and orientation from the atomistic configuration within each residue. Proper rotations preserve the internal geometry and chirality of that local configuration when placing it in space. Explicit frames also provide a coordinate system for orientation-sensitive interactions at every intermediate generative state, without transporting all peptide atoms.

Following partially latent modeling (Geffner et al., 2026; Didi et al., 2026), we introduce a continuous latent $z_i \in \mathbb{R}^{d_z}$ for reconstructing sequence and local atomistic structure. The generative state is

$$(T, z) \in SO(3)^L \times \mathbb{R}^{3L} \times \mathbb{R}^{Ld_z}, \tag{2}$$

and the model factorizes as

$$p_{\theta,\phi}(s, U, T \mid P) = \int p_\theta(T, z \mid P)\, p_\phi(s, U \mid T, z)\, \mathrm{d}z. \tag{3}$$

Here $p_\theta$ jointly generates frames and latents conditioned on the pocket, and $p_\phi$ reconstructs sequence and local atom coordinates.

## 3.2 VARIATIONAL AUTOENCODER

We learn the partially latent representation with a target-agnostic variational autoencoder (Kingma & Welling, 2014). The encoder defines $q_\psi(z \mid s, U, T)$, and the decoder reconstructs $(\hat{s}, \hat{U}) = D_\phi(T, z)$. Both networks use pair-biased Transformers (Jumper et al., 2021; Geffner et al., 2026; Didi et al., 2026), with frame geometry supplied through the relative quantities $\Delta x_{ij} = R_i^\top (x_j - x_i)$, $\Delta R_{ij} = R_i^\top R_j$. These quantities and $U$ are invariant under a common rigid transformation. They allow self-attention with pairwise attention biases to capture directional inter-residue relationships while preserving the latent posterior's invariance to a common rigid transformation of the input structure. The autoencoder is trained with

$$\mathcal{L}_{\mathrm{AE}} = \mathcal{L}_{\mathrm{seq}} + \mathcal{L}_{\mathrm{atom}} + \beta \mathcal{L}_{\mathrm{KL}}, \tag{4}$$

combining sequence reconstruction, local atom-coordinate reconstruction, and posterior regularization toward a standard Gaussian. The trained autoencoder is then frozen. Loss definitions and reconstruction quality are given in Appendix A.1.

## 3.3 JOINT FLOW MATCHING

We learn $p_\theta(T, z \mid P)$ by jointly transporting explicit frames and latent variables. We center peptide and target positions at the pocket $C_\alpha$ centroid $c_P$. For each training peptide, we sample

$z^1 \sim q_\psi(z \mid s, U, T^1)$ from the frozen encoder and sample the source state $(T^0, z^0)$ independently from equation 10. Orientations follow $SO(3)$ geodesics (Yim et al., 2024; Li et al., 2024), and positions and latents follow Euclidean interpolants (Geffner et al., 2026):

$$\begin{aligned} R_i^{t_T} &= R_i^0 \operatorname{Exp}\left[t_T \operatorname{Log}\left((R_i^0)^\top R_i^1\right)\right], \\ x_i^{t_T} &= (1-t_T)x_i^0 + t_T x_i^1, \\ z_i^{t_z} &= (1-t_z)z_i^0 + t_z z_i^1. \end{aligned} \tag{5}$$

Here Exp and Log are the exponential and logarithm maps on $SO(3)$. The geometric and latent time variables $t_T, t_z$ are sampled independently during training. Their conditional velocity targets are

$$\begin{aligned} \omega_{R,i}^* &= \operatorname{vee}\left[\operatorname{Log}\left((R_i^0)^\top R_i^1\right)\right], \\ v_{x,i}^* &= x_i^1 - x_i^0, \qquad v_{z,i}^* = z_i^1 - z_i^0. \end{aligned} \tag{6}$$

Here vee sends a skew-symmetric matrix to its rotation vector in $\mathbb{R}^3$.

A single network $V^\theta(T^{t_T}, z^{t_z}, P, t_T, t_z) = (\omega_R^\theta, v_x^\theta, v_z^\theta)$ predicts all three velocities from the joint state. Consequently, geometric motion can depend on the current latent representation, and latent evolution can depend on the current frames. The conditional flow-matching objective is

$$\mathcal{L}_{\mathrm{CFM}} = \mathbb{E}\left[\|\omega_R^\theta - \omega_R^*\|^2_{m\hat{m}} + \|v_x^\theta - v_x^*\|^2_m + \|v_z^\theta - v_z^*\|^2_m\right], \tag{7}$$

where $m$ masks valid residues and $\hat{m}$ valid frames. Clock distributions and loss normalizations are given in Appendix A.2.

### 3.4 Representation-aligned geometric vector field

The vector field uses the transported frames to express geometric interactions in query-local coordinates. Residue relations use the relative translations and rotations introduced for the autoencoder. For peptide-to-target conditioning, target atom $a$ of residue $m$ is expressed in the current frame of peptide residue $i$ as

$$r_{ima} = R_i^\top \left(X_{m,a}^P - x_i\right). \tag{8}$$

The vector $r_{ima}$ retains the direction and distance of the target atom relative to the peptide query while remaining invariant to a common rigid transformation of the peptide and target. Using these query-dependent features, LaCPlow constructs peptide–target and target–peptide relations in their respective query frames, unlike prior partially latent flow-matching models that obtain reverse-direction relations by transposing pair features (Didi et al., 2026).

We use gated Invariant Point Attention (IPA) (Jumper et al., 2021) to compare learned points expressed in the residue frames, returning invariant features $h_i$ in the query frame. During each vector-field evaluation, the residue frames are held fixed at $T^{t_T}$. Unlike iterative structure modules that update frames inside the network and predict a clean structure (Li et al., 2024), the numerical integrator updates the transported state. The readout follows the transformation law of each state component:

$$\omega_{R,i}^\theta = g_R(h_i), \qquad v_{x,i}^\theta = R_i g_x(h_i), \qquad v_{z,i}^\theta = g_z(h_i). \tag{9}$$

Here $g_R$, $g_x$, and $g_z$ are MLPs. The rotational velocity is represented in body-frame coordinates, the translational velocity transforms equivariantly in Cartesian space, and the latent velocity is invariant. Architecture details and transformation properties are given in Appendices A.3 and A.3.1.

### 3.5 Generation

We initialize the joint state with

$$R_i^0 \sim \mathrm{Haar}(SO(3)), \qquad x_i^0 = \varepsilon_i + d, \qquad z_i^0 \sim \mathcal{N}(0, I_{d_z}), \tag{10}$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma_x^2 I)$ and $d \sim \mathcal{N}(0, \sigma_d^2 I)$ is a shared positional offset. Geometry and latents advance along separate time schedules. At each step, velocities evaluated at the current joint state give

$$\begin{aligned} R_i^{k+1} &= R_i^k \operatorname{Exp}\left(\Delta t_T^k [\omega_{R,i}^\theta]_\times\right), \\ x_i^{k+1} &= x_i^k + \Delta t_T^k v_{x,i}^\theta, \\ z_i^{k+1} &= z_i^k + \Delta t_z^k v_{z,i}^\theta. \end{aligned} \tag{11}$$

Here $[\cdot]_\times$ is the skew-symmetric matrix representation of a rotation vector. The exponential-map update keeps orientations on $SO(3)$. The frozen decoder maps the endpoint to $(s, U)$, and $X_{i,a}^{\text{global}} = R_i^1 U_{i,a} + x_i^1 + c_P$ restores atom coordinates in the original target coordinate system. Sampling schedules and training settings are given in Appendices A.4 and A.5.

# 4 Experiments

## 4.1 Experimental setup

**Dataset.** We use CPCore for training and validation, a high-quality subset of CPSea (Yang et al., 2025) containing 71,867 cyclic peptide–protein complexes. For evaluation, we use CPSea_PDB, a set of high-resolution structures from the Protein Data Bank (PDB) processed following Wang et al. (2026), yielding 84 complexes after removing overlap with CPCore. Detailed preprocessing steps are given in Appendix D.1.

**Baselines.** PepGLAD (Kong et al., 2024), DiffPepBuilder (Wang et al., 2024), PepFlow (Li et al., 2024), APCyc (Zhao et al., 2026), and Proteína-Complexa (Didi et al., 2026) are retrained on the same CPCore split to generate cyclic peptides from scratch, while MuCO (Wang et al., 2026) is trained on peptide chains only. **PepGLAD (raw)** denotes the direct decoder output, whereas **PepGLAD (ideal.)** additionally applies its side-chain idealization procedure. For **DiffPepBuilder**, the model generates only backbone and $C_\beta$ atoms, so metrics that require side-chain atoms beyond $C_\beta$ are not reported. **APCyc (raw)** samples the cyclization class from the model, whereas **APCyc ($c^\star$)** supplies the ground-truth class at inference. For **Proteína-Complexa**, inference-time optimization is excluded so that comparisons reflect the model's own expressiveness. **MuCO** generates cyclic peptides conditioned on a supplied sequence and not on a target pocket, so only single-chain geometric metrics are reported. **CP-Composer** (Jiang et al., 2025) is designed to be trained on linear peptides, so we use the authors' released checkpoint and generate head-to-tail (H2T) cyclic peptides rather than internally cyclized peptides in order to match the terminal-linked cyclization types in CPSea_PDB as closely as possible. Parentheses report CP-Composer results restricted to targets whose ground-truth cyclization type is H2T. For each target, every model attempts 100 samples. Unless stated otherwise, reported values are first aggregated within target and then averaged uniformly over the 84 targets.

**Metrics** For energy minimization, we use a CHARMM36-based force-field engine (Huang & MacKerell Jr, 2013) modified by Yang et al. (2025); Wang et al. (2026). The cyclization type of a generated structure $X$ with sequence $s$ is $c(X, s)$:

$$c(X,s) = \begin{cases} \text{C2C}, & \text{if the termini are Cys–Cys and } d_{\text{C2C}}(X) < 2.5\,\text{Å}, \\ \text{K2DE}, & \text{if the termini are Lys–Asp/Glu and } d_{\text{K2DE}}(X) < d_{\text{H2T}}(X), \\ \text{H2T}, & \text{otherwise.} \end{cases} \tag{12}$$

The K2DE condition includes either terminal ordering.

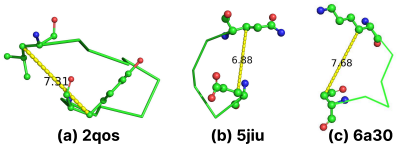



Figure 2: Linear peptides from LNR (Tsaban et al., 2022) dataset. Dashed lines indicate $d_{C_\beta}$.

**(1) $C_\beta$-distance window.** For comparison with the published benchmarks, we report $P_{C_\beta} = \mathbf{1}[3 \le d_{C_\beta} \le 8]$, where $d_{C_\beta}$ is the endpoint $C_\beta$–$C_\beta$ distance. Although this metric was proposed by Yang et al. (2025) and proved effective for constructing the CPSea dataset, we empirically observed that it occasionally misclassifies non-cyclic structures as successful closure (Figure 2). **(2) Bond-forming atom distance.** As a complement to $P_{C_\beta}$, we test whether the bond-forming atom distance for the assigned type $c(X, s)$ falls within its chemistry-specific window $P_{d_{\text{bond}}}(X, s) = \mathbf{1}\left[d_{\text{bond}}^{c(X,s)} \in W_{c(X,s)}\right]$. We use $W_{\text{C2C}} = [1.84, 2.2]$ Å and $W_{\text{H2T}} = W_{\text{K2DE}} = [1.2, 1.55]$ Å. The atom pairs and window calibration are detailed in Appendix B.1. **(3) Structural quality and affinity.** Chirality is 1 minus the fraction of D amino acids. We also report Ramachandran allowed and favored rates, median Rosetta $dG$ (Alford et al., 2017), energy success ($dG < 0$), and the median AutoDock Vina score (Morris et al., 2009; Trott & Olson, 2010). **(4) Physicochemical properties.** We report

the fraction of peptides satisfying GRAVY $< 0$ (Kyte & Doolittle, 1982). **(5) Diversity.** Per target, peptide $C_\alpha$ coordinates are clustered by single linkage with a 4 Å RMSD cutoff without superposition in the target-fixed coordinate frame. **(6) Sequence–structure self-consistency (scRMSD).** scRMSD is the Kabsch-superposed $C_\alpha$ RMSD between the minimized generated peptide and the relaxed rank-1 HighFold2 (Zhu et al., 2025) prediction. **(7) Novelty.** Novelty is 1 minus the within-target mean maximum query TM-score of the peptide against the CPCore training peptides. Closure, chirality, and Ramachandran in the tables are computed before minimization on every generated sample. The $dG$, Vina, GRAVY, diversity, novelty, and scRMSD use only post-minimization samples with $dG < 0$. Pre- and post-minimization displacements use every paired sample, without that filter. Detailed descriptions and evaluation of the metrics are given in Appendix D.

### 4.2 Evaluating Substrates for Cyclic Peptide Generation

Table 1: **Base generative substrate**. Pre-minimization geometry (top) and post-minimization quality (bottom); 84 targets × 100 samples. Best in **bold**, second best underlined.

| | | Closure | | Chirality | Ramachandran | |
|---|---|---|---|---|---|---|
| Representation | Model | $P_{d_{\text{bond}}}$ ↑ | $P_{C_\beta}$ ↑ | ↑ | Allowed↑ | Favored↑ |
| | *CPSea_PDB (ref)* | *100.0* | *98.8* | *100.0* | *97.3* | *89.8* |
| Explicit | PepFlow | 7.0 | 93.9 | **100.0** | 92.4 | 81.7 |
| | DiffPepBuilder | 1.4 | 95.0 | **100.0** | 91.0 | 78.2 |
| Fully latent | PepGLAD (raw) | 19.4 | 96.1 | 49.9 | 47.3 | 29.0 |
| | PepGLAD (ideal.) | 16.3 | 81.9 | **100.0** | 43.8 | 25.7 |
| Partially latent | Proteína-Complexa | 20.0 | **96.3** | 99.9 | 73.3 | 51.9 |
| Partially latent | **LaCPlow-Base** | **25.8** | 95.5 | **100.0** | **95.3** | **83.5** |

| | | Energy | Affinity | | GRAVY | div | nov | scRMSD | |
|---|---|---|---|---|---|---|---|---|---|
| Representation | Model | ↑ | $dG$ ↓ | Vina↓ | ↑ | ↑ | ↑ | H2T↓ | C2C↓ |
| | *CPSea_PDB (ref)* | *98.8* | −26.4 | −3.20 | *91.5* | — | *0.396* | *3.18* | *2.70* |
| Explicit | PepFlow | 87.3 | −15.9 | −3.98 | 59.3 | 0.50 | 0.572 | 3.52 | 3.50 |
| | DiffPepBuilder | 95.5 | −21.4 | −4.01 | 90.5 | 0.17 | 0.515 | 3.12 | — |
| Fully latent | PepGLAD (raw) | 91.5 | −18.1 | −3.55 | 85.9 | **0.62** | **0.612** | 3.69 | 3.66 |
| | PepGLAD (ideal.) | 90.9 | −17.8 | −3.56 | 86.1 | **0.62** | 0.608 | 3.73 | 3.65 |
| Partially latent | Proteína-Complexa | 85.3 | −18.4 | −3.70 | **91.5** | 0.12 | 0.508 | **2.52** | **2.68** |
| Partially latent | **LaCPlow-Base** | **97.2** | **−22.0** | **−4.11** | 91.1 | 0.20 | 0.541 | 3.52 | 3.63 |

PepFlow, DiffPepBuilder, and PepGLAD are designed to model linear peptides, whereas Proteína-Complexa is developed for modeling proteins. By combining explicit residue frames with a representation-aligned network, LaCPlow preserves stereochemically relevant geometry, even compared to fully explicit models. In other words, LaCPlow's explicit frames with a representation-aligned architecture capture cyclic-peptide geometric regularities directly from data. Post-minimization chirality and Ramachandran results are reported in Table 13, where LaCPlow achieves the highest chirality and Ramachandran pass rates following energy minimization. Taking both the energy pass rate and binding affinity into account, LaCPlow reliably generates peptides with favorable interfacial interactions with the target. Target-fixed diversity conflates peptide conformational diversity with variation in surface placement, whereas therapeutic campaigns typically prioritize diverse conformations at a specific binding site rather than variation in binding-site placement (Watson et al., 2023; Pacesa et al., 2025). We therefore also report sequence diversity and conformational diversity after Kabsch superposition. LaCPlow's sequence diversity is 0.955, and its aligned mean pairwise RMSD reaches 88% of the best baseline (Table 12).

### 4.3 Cyclization-specific specialization

We observed that cyclic peptide generative models built on existing substrates—such as APCyc, which builds upon PepGLAD—still struggle to effectively model cyclic peptides even when incorporating cyclization-specific architectural specialization. This limitation primarily stems from the fact that PepGLAD was originally designed to model linear peptides. Furthermore, as shown in Table 1, we empirically observed that fully latent models face inherent disadvantages in preserving stereochemical

Table 2: **Cyclization-specific specialization.** Best in **bold**, second best underlined.

| Model | Closure | | Chirality | Ramachandran | |
|---|---|---|---|---|---|
| | $P_{d_{\text{bond}}}$ ↑ | $P_{C_\beta}$ ↑ | ↑ | Allowed↑ | Favored↑ |
| *CPSea_PDB (ref)* | *100.0* | *98.8* | *100.0* | *97.3* | *89.8* |
| MuCO$^\star$ | 2.8 | 78.7 | 50.3 | **96.2** | **89.4** |
| APCyc (raw) | 9.0 | 74.4 | 50.1 | 43.3 | 24.9 |
| APCyc ($c^\star$) | 15.0 | 85.2 | 49.9 | 43.2 | 25.0 |
| CP-Composer | 0.4 (0.3) | 74.7 (75.4) | 50.2 (50.6) | 39.1 (39.2) | 18.6 (18.5) |
| **LaCPlow-Cyc** | **61.0** | **93.9** | **100.0** | 94.5 | 83.5 |

| Model | Energy | Affinity | | GRAVY | div | nov | scRMSD | |
|---|---|---|---|---|---|---|---|---|
| | ↑ | *dG* ↓ | Vina↓ | ↑ | ↑ | ↑ | H2T↓ | C2C↓ |
| *CPSea_PDB (ref)* | *98.8* | −26.4 | −3.20 | *91.5* | — | *0.396* | *3.18* | *2.70* |
| APCyc (raw) | 92.2 | −16.2 | −3.08 | 78.5 | 0.56 | **0.625** | 3.85 | 3.94 |
| APCyc ($c^\star$) | 92.0 | −16.1 | −3.12 | **86.9** | 0.56 | **0.625** | 3.88 | 3.87 |
| CP-Composer | 71.3 (73.6) | −13.4 (−11.5) | −2.49 (−2.31) | 69.2 (69.0) | **0.94** | 0.622 (0.602) | 3.83 (3.60) | — |
| **LaCPlow-Cyc** | **97.0** | **−20.6** | **−4.19** | 86.8 | 0.22 | 0.534 | **3.48** | **3.60** |

quality. Therefore, using LaCPlow-Base as our underlying substrate, we evaluate whether introducing more refined, cyclization-specific inductive biases enables the precise modeling of bond-forming geometries while largely preserving stereochemical quality. Table 2 compares alternative specializations using the same metrics as Table 1. The implementation of LaCPlow-Cyc is in Appendix A.6. Compared to existing generative models specifically tailored for cyclic peptides, LaCPlow-Cyc achieves high closure rates and balanced stereochemical quality. Notably, LaCPlow-Cyc improves bond-forming distance while largely preserving LaCPlow-Base's stereochemical quality and property metrics.

### 4.4 Importance of precise modeling before energy minimization

Table 3: **Structural correction induced by energy minimization.**

| Model | Cartesian (Å) | | Angular RMSD (°) | | | | Closure | |
|---|---|---|---|---|---|---|---|---|
| | BB ↓ | SC ↓ | BB bond ↓ | SC bond ↓ | BB tor ↓ | SC tor ↓ | $\Delta d_{\text{bond}}$ (Å) ↓ | $\Delta C_\beta$ (Å) ↓ |
| PepGLAD (raw) | 2.61 | 3.10 | 22.9 | 41.8 | 66.2 | 72.9 | 0.64 | 0.95 |
| PepGLAD (ideal.) | 2.66 | 3.18 | 19.3 | **3.1** | 59.6 | 63.1 | 1.00 | 1.47 |
| DiffPepBuilder | 2.50 | — | 15.3 | — | 43.9 | — | 1.84 | 1.14 |
| PepFlow | 2.41 | 2.65 | 7.9 | 3.3 | 48.3 | 47.3 | 2.77 | 1.11 |
| APCyc (raw) | 3.28 | 3.88 | 27.8 | 41.2 | 74.8 | 76.2 | 2.51 | 1.89 |
| APCyc ($c^\star$) | 3.26 | 3.86 | 27.7 | 41.5 | 74.4 | 76.8 | 1.48 | 1.43 |
| Proteína-Complexa | 2.62 | 3.03 | 18.5 | 15.5 | 47.7 | 60.0 | 0.46 | 0.98 |
| MuCO$^\star$ | **1.26** | **1.53** | 12.1 | 31.9 | 45.4 | 57.3 | 3.70 | 1.29 |
| CP-Composer | 3.31 | 3.95 | 28.7 | 54.0 | 78.8 | 77.3 | 4.34 | 1.90 |
| LaCPlow-Base | 2.15 | 2.57 | **5.0** | 6.7 | 37.3 | 47.5 | 0.53 | **0.87** |
| LaCPlow-Cyc | 2.11 | 2.52 | 5.1 | 6.8 | **37.0** | **46.7** | **0.26** | 0.92 |

Energy minimization remains useful for correcting local geometric defects in generated cyclic peptides. However, if satisfying closure constraints requires substantial structural changes, the quality of the relaxed structure reflects the refinement procedure as well as the generative model. We therefore evaluate geometry before minimization and quantify how much the backbone, side chains, and closure site change during refinement. Table 3 quantifies geometrical changes during energy minimization, with metric definitions provided in Appendix D.6. Among the target-conditioned generators, both LaCPlow variants show the smallest refinement-induced backbone displacement and backbone torsional corrections. Although LaCPlow-Base does not attain the highest $P_{C_\beta}$ in Table 1, the $C_\beta$-distance distribution of LaCPlow-Base is concentrated near the accepted region (Figure 7), and energy minimization induces the smallest change in $C_\beta$ distance among target-conditioned generators. Relative to LaCPlow-Base, LaCPlow-Cyc has a markedly smaller $\Delta d_{\text{bond}}$, with only small changes

in the other refinement metrics, indicating that cyclization-specific information selectively improves bond-forming geometry.

### 4.5 Representation-aligned network analysis

Table 4: **Representation-aligned geometric network analysis.** The two $(R, x, z)$ rows at the bottom compare LaCPlow networks on the same representation state.

| Representation | Model / network | $P_{d_{\text{bond}}}$ ↑ | $P_{C_\beta}$ ↑ | Chirality↑ | Ramachandran | | $dG$ ↓ | Energy↑ |
|---|---|---|---|---|---|---|---|---|
| | | | | | Allowed↑ | Favored↑ | | |
| $(x, z)$ | Proteína-Complexa | <u>20.0</u> | **96.3** | <u>99.9</u> | 73.3 | 51.9 | −18.4 | 85.3 |
| $(R, x, z)$ | PairBiased Transformer | 17.7 | 90.0 | **100.0** | <u>89.4</u> | <u>75.2</u> | <u>−19.2</u> | <u>92.1</u> |
| $(R, x, z)$ | LaCPlow-Base | **25.8** | <u>95.5</u> | **100.0** | **95.3** | **83.5** | **−22.0** | **97.2** |

Proteína-Complexa (Didi et al., 2026) achieves high chirality using direct coordinate prediction and rotational augmentation. LaCPlow explicitly transports residue frames $R \in SO(3)$ and predicts atom coordinates in these local frames, allowing local handedness to be learned while preserving $SE(3)$-equivariance without rotational augmentation. Within the shared $(R, x, z)$ representation, the representation-aligned geometric network improves Ramachandran allowed rate from 89.4% to 95.3% and bond-distance success from 17.7% to 25.8% relative to the PairBiased Transformer.

Table 5: **Specialization ablations**

| Variant | $P_{d_{\text{bond}}}$ ↑ | $P_{C_\beta}$ ↑ | Chirality↑ | Ramachandran | | $dG$ ↓ | Energy↑ |
|---|---|---|---|---|---|---|---|
| | | | | Allowed↑ | Favored↑ | | |
| $(R, x, z)$ | 25.8 | **95.5** | **100.0** | **95.3** | **83.5** | **−22.0** | **97.2** |
| $(R, x, \mathcal{C}, z)$ | <u>49.1</u> | 86.6 | **100.0** | 91.6 | <u>78.1</u> | −19.5 | 93.2 |
| $(R, x, z) + \mathcal{C}\|z$ | **61.0** | <u>93.9</u> | **100.0** | <u>94.5</u> | **83.5** | <u>−20.6</u> | <u>97.0</u> |

When $\mathcal{C}$ appears in the variant column in Table 5, closure-aware loss terms are also added. Transporting $\mathcal{C}$ as a flow state raises $P_{d_{\text{bond}}}$, worsens $dG$, and reduces Ramachandran quality and energy success, showing that the choice of explicit representation matters. Adding closure-aware loss terms on $\mathcal{C}|z$ sharpens bond-forming geometry while largely preserving the substrate's stereochemistry. Refinement displacements for these rows are reported in Table 17. In Tables 4 and 5, each architectural variant of LaCPlow is instantiated and trained separately, rather than masked at inference. Implementation details for each row are given in Appendix A.6.

### 4.6 Latent–Geometry Coupling and Latent-Space Optimization

Latent generative models provide continuous variables for peptide optimization and have served as substrates for specialized design methods (Kong et al., 2024; Jiang et al., 2025; Sun et al., 2026; Zhao et al., 2026; Zhang et al., 2026). LaCPlow likewise includes a latent state, but its presence alone does not establish that latent edits influence the explicitly generated residue frames or provide a useful search direction. We therefore test whether perturbations to $z$ propagate to the frames through the joint flow and whether directed latent edits outperform matched random perturbations on an interface objective.

**Latent–geometry coupling.** On 128 randomly sampled CPCore validation pockets, we intervene only on $z$ at the midpoint of an otherwise fixed generation trajectory. We compare two conditions. In the *joint-propagation* condition, generation resumes from the perturbed midpoint state while both $T$ and $z$ continue to evolve. In the *decoder-only* condition, the endpoint frame $T^1$ is held fixed and only the endpoint latent is perturbed before decoding. Perturbation magnitude is controlled by $\rho \in \{0.05, 0.1, 0.25, 0.5\}$, the root-mean-square size of the perturbation in standardized latent coordinates (Appendix E.1). The induced response of the explicit state is measured as

$$\Delta x(\rho) = \sqrt{\frac{1}{L}\sum_i \left\| x_i^{1,\rho} - x_i^1 \right\|_2^2}, \qquad \Delta R(\rho) = \frac{1}{L}\sum_i \left\| \text{vee}\left[\text{Log}\left((R_i^1)^\top R_i^{1,\rho}\right)\right] \right\|_2 . \tag{13}$$

We report the median over the 128 pockets with IQR (25th–75th percentile) bands. The complete intervention, pairing, and aggregation protocol is given in Appendix E.1.

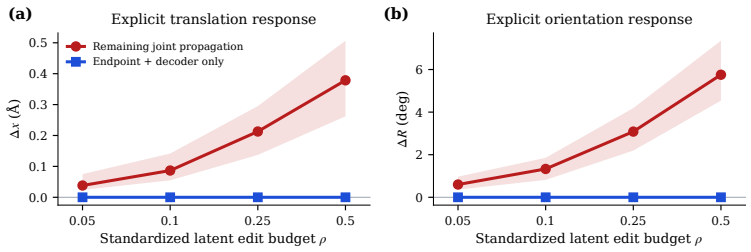



Figure 3: **Latent–geometry coupling.** Translation **(a)** and orientation **(b)** under joint propagation.

Figure 3 shows that perturbations applied only to $z$ produce increasing changes in both residue translations and orientations under joint propagation. Appendix E.2 examines sensitivity to larger perturbations.

**Latent editing.** We next test whether the same latent representation can serve as an optimization interface (Figure 4). Starting from the midpoint of an unguided trajectory, we edit only $z$ and resume joint propagation. As a proof of concept, we optimize a surrogate interface score $S_{\text{iface}} = S_{\text{match}} - S_{\text{rep}}$ with SPSA (Spall, 1992) in standardized latent coordinates, where $S_{\text{match}}$ scores peptide–pocket contacts and $S_{\text{rep}}$ penalizes steric overlap, and compare the resulting edits with random perturbations matched in latent displacement magnitude. Directed edits improve the surrogate score and reduce unrelaxed Rosetta $dG$ relative to the random control. The optimization protocol, control construction, and post-edit structural quality are given in Appendix E.3.

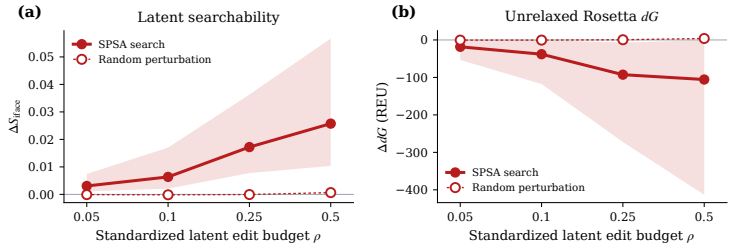



Figure 4: **Latent editing.** Interface score **(a)** and unrelaxed Rosetta $dG$ **(b)**.

Moreover, because residue frames are explicit, their evolution can be monitored throughout generation without intermediate decoding.

## 5 Conclusion

We presented LaCPlow, a target-conditioned partially latent flow model that combines explicit residue frames and latent atomistic information with a representation-aligned geometric network. LaCPlow-Base achieves strong geometric quality before energy minimization. LaCPlow-Cyc improves bond-forming geometry while largely preserving stereochemical quality and requiring comparable structural corrections during minimization. These results support LaCPlow as a substrate for cyclic peptide generation and chemistry-specific specialization. **Limitations.** The current model covers cyclization between terminal residues and inherits biases from predicted training structures. Moreover, scRMSD measures agreement with a structure predictor (Zhu et al., 2025), whose accuracy under cyclization constraints limits its use as a proxy for designability. We therefore interpret scRMSD alongside closure, stereochemical, and energetic metrics.

### AI USE STATEMENT

In this work, we used generative AI tools to assist with translation. We did not use generative AI tools to propose or refine hypotheses or to formulate mathematical claims. Additionally, we used generative AI tools to create or modify scientific figures or images, and to create or edit software code. We have reviewed all AI-assisted work. LLM-generated code was verified and tested for correctness by the authors. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

### REPRODUCIBILITY STATEMENT

Section 3 describes the model, training objectives, and generation procedure. Appendix A provides the architecture, optimization settings, and sampling schedules. Appendix D documents dataset construction, splits, evaluation stages, and metrics. Closure evaluation and latent intervention protocols are detailed in Appendices B and E, respectively. An anonymized implementation and configuration files are provided in the supplementary material.

## REFERENCES

Rebecca F. Alford, Andrew Leaver-Fay, Jeliazko R. Jeliazkov, Matthew J. O'Meara, Frank P. DiMaio, Hahnbeom Park, Maxim V. Shapovalov, P. Douglas Renfrew, Vikram K. Mulligan, Kalli Kappel, Jason W. Labonte, Michael S. Pacella, Richard Bonneau, Philip Bradley, Jr. Dunbrack, Roland L., Rhiju Das, David Baker, Brian Kuhlman, Tanja Kortemme, and Jeffrey J. Gray. The rosetta all-atom energy function for macromolecular modeling and design. *Journal of Chemical Theory and Computation*, 13(6):3031–3048, 04 2017. ISSN 1549-9618. doi: 10.1021/acs.jctc.7b00125. URL https://doi.org/10.1021/acs.jctc.7b00125.

Sidhartha Chaudhury, Sergey Lyskov, and Jeffrey J. Gray. Pyrosetta: a script-based interface for implementing molecular modeling algorithms using rosetta. *Bioinformatics*, 26(5):689–691, 03 2010. ISSN 1367-4803. doi: 10.1093/bioinformatics/btq007. URL https://doi.org/10.1093/bioinformatics/btq007.

J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N. P. King, and D. Baker. Robust deep learning–based protein sequence design using proteinmpnn. *Science*, 378(6615):49–56, 2022. doi: 10.1126/science.add2187. URL https://www.science.org/doi/abs/10.1126/science.add2187.

Kieran Didi, Zuobai Zhang, Guoqing Zhou, Danny Reidenbach, Zhonglin Cao, Sooyoung Cha, Tomas Geffner, Christian Dallago, Jian Tang, Michael M. Bronstein, Martin Steinegger, Emine Kucukbenli, Arash Vahdat, and Karsten Kreis. Scaling atomistic protein binder design with generative pretraining and test-time compute. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=qmCpJtFZra.

R. A. Engh and R. Huber. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallographica Section A*, 47(4):392–400, Jul 1991. doi: 10.1107/S0108767391001071. URL https://doi.org/10.1107/S0108767391001071.

Keld Fosgerau and Torsten Hoffmann. Peptide therapeutics: current status and future directions. *Drug Discovery Today*, 20(1):122–128, 2015. ISSN 1359-6446. doi: https://doi.org/10.1016/j.drudis.2014.10.003. URL https://www.sciencedirect.com/science/article/pii/S1359644614003997.

Tomas Geffner, Kieran Didi, Zhonglin Cao, Danny Reidenbach, Zuobai Zhang, Christian Dallago, Emine Kucukbenli, Karsten Kreis, and Arash Vahdat. La-proteina: Atomistic protein generation via partially latent flow matching. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=RDerF20JYT.

Bart Hazes and Bauke W. Dijkstra. Model building of disulfide bonds in proteins with known three-dimensional structure. *Protein Engineering*, 2(2):119–125, 07 1988. ISSN 0269-2139. doi: 10.1093/protein/2.2.119. URL https://doi.org/10.1093/protein/2.2.119.

Jing Huang and Alexander D. MacKerell Jr. Charmm36 all-atom additive protein force field: Validation based on comparison to nmr data. *Journal of Computational Chemistry*, 34(25): 2135–2145, 2013. doi: https://doi.org/10.1002/jcc.23354. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.23354.

Xinjian Ji, Alexander L. Nielsen, and Christian Heinis. Cyclic peptides for drug development. *Angewandte Chemie International Edition*, 63(3):e202308251, 2024. doi: https://doi.org/10.1002/anie.202308251. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202308251.

Dapeng Jiang, Xiangzhe Kong, Jiaqi Han, Mingyu Li, Rui Jiao, Wenbing Huang, Stefano Ermon, Jianzhu Ma, and Yang Liu. Zero-shot cyclic peptide design via composable geometric constraints. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 27553–27568. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/jiang25d.html.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, Aug 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03819-2. URL https://doi.org/10.1038/s41586-021-03819-2.

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In Yoshua Bengio and Yann LeCun (eds.), *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014. URL http://arxiv.org/abs/1312.6114.

Xiangzhe Kong, Yinjun Jia, Wenbing Huang, and Yang Liu. Full-atom peptide design with geometric latent diffusion. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 74808–74839. Curran Associates, Inc., 2024. doi: 10.52202/079017-2379. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/88ad9774ffcb7a272457e9396f793a07-Paper-Conference.pdf.

Jack Kyte and Russell F. Doolittle. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105–132, 1982. ISSN 0022-2836. doi: https://doi.org/10.1016/0022-2836(82)90515-0. URL https://www.sciencedirect.com/science/article/pii/0022283682905150.

Greg Landrum, Paolo Tosco, Ricardo Rodriguez, Brian Kelley, David Cosgrove, Riccardo Vianello, sriniker, Peter Gedeck, Gareth Jones, Dan Nealschneider, Eisuke Kawashima, NadineSchneider, tadhurst cdd, Andrew Dalke, Niels Maeder, Matt Swain, Yakov Pechersky, Brian Cole, Kevin Boyd, Samo Turk, Aleksandr Savelev, Rachel Walker, Alain Vaucher, Maciej Wójcikowski, Hussein Faara, Ichiru Take, Vincent F. Scalfani, Steven Kearnes, Kazuya Ujihara, and Daniel Probst. rdkit/rdkit: 2026_03_6 (q1 2026) release, August 2026. URL https://doi.org/10.5281/zenodo.22140358.

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 Ruslan Salakhutdinov, Zico Kolter, Katherine A. Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*, volume 235 of *Proceedings of Machine Learning Research*, pp. 27615–27640. PMLR / OpenReview.net, 2024. URL https://proceedings.mlr.press/v235/li24o.html.

Simon C. Lovell, Ian W. Davis, W. Bryan Arendall III, Paul I. W. de Bakker, J. Michael Word, Michael G. Prisant, Jane S. Richardson, and David C. Richardson. Structure validation by C$\alpha$ geometry: $\phi$, $\psi$ and C$\beta$ deviation. *Proteins: Structure, Function, and Bioinformatics*, 50(3): 437–450, 2003. doi: https://doi.org/10.1002/prot.10286. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.10286.

Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16):2785–2791, 2009. doi: https://doi.org/10.1002/jcc.21256. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21256.

Martin Pacesa, Lennart Nickel, Christian Schellhaas, Joseph Schmidt, Ekaterina Pyatova, Lucas Kissling, Patrick Barendse, Jagrity Choudhury, Srajan Kapoor, Ana Alcaraz-Serna, Yehlin Cho, Kourosh H. Ghamary, Laura Vinué, Brahm J. Yachnin, Andrew M. Wollacott, Stephen Buckley, Adrie H. Westphal, Simon Lindhoud, Sandrine Georgeon, Casper A. Goverde, Georgios N. Hatzopoulos, Pierre Gönczy, Yannick D. Muller, Gerald Schwank, Daan C. Swarts, Alex J. Vecchio, Bernard L. Schneider, Sergey Ovchinnikov, and Bruno E. Correia. One-shot design of functional protein binders with bindcraft. *Nature*, 646(8084):483–492, Oct 2025. ISSN 1476-4687. doi: 10.1038/s41586-025-09429-6. URL https://doi.org/10.1038/s41586-025-09429-6.

Stephen A. Rettie, Katelyn V. Campbell, Asim K. Bera, Alex Kang, Simon Kozlov, Yensi Flores Bueso, Joshmyn De La Cruz, Maggie Ahlrichs, Suna Cheng, Stacey R. Gerben, Mila Lamb, Analisa Murray, Victor Adebomi, Guangfeng Zhou, Frank DiMaio, Sergey Ovchinnikov, and Gaurav Bhardwaj. Cyclic peptide structure prediction and design using alphafold2. *Nature Communications*, 16(1):4730, May 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-59940-7. URL https://doi.org/10.1038/s41467-025-59940-7.

Philipp Schake, Sarah Naomi Bolz, Katja Linnemann, and Michael Schroeder. PLIP 2025: introducing protein-protein interactions to the protein-ligand interaction profiler. *Nucleic acids research*, 53 (W1):W463–W465, 2025. ISSN 1362-4962. doi: 10.1093/nar/gkaf361.

J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automatic Control*, 37(3):332–341, 1992. doi: 10.1109/9.119632.

Martin Steinegger and Johannes Söding. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35(11):1026–1028, Nov 2017. ISSN 1546-1696. doi: 10.1038/nbt.3988. URL https://doi.org/10.1038/nbt.3988.

Heqi Sun, Hong Tan, Yanyi Chu, Jiayi Li, Ruixuan Wang, and Dong-Qing Wei. Design of permeability-optimized target-binding macrocycles via direct preference optimization. *Chemical Science*, 17(20):10223–10236, 05 2026. ISSN 2041-6520. doi: 10.1039/d6sc01722c. URL https://doi.org/10.1039/d6sc01722c.

Trupti Thite, Kavita A Iyer, Preeti Jain, Janet M Sasso, Rumiana Tenchov, Sudarshan Murmu, Mahaveer Surendranath, Ankush Maind, Umesh T Balande, and Qiongqiong Angela Zhou. Cyclic peptides in modern drug discovery: Trends and therapeutic directions. *Journal of Medicinal Chemistry*, 69(15):17770–17799, 07 2026. ISSN 0022-2623. doi: 10.1021/acs.jmedchem.6c01522. URL https://doi.org/10.1021/acs.jmedchem.6c01522.

Oleg Trott and Arthur J. Olson. Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2):455–461, 2010. ISSN 1096-987X. doi: 10.1002/jcc.21334. URL http://dx.doi.org/10.1002/jcc.21334.

Tomer Tsaban, Julia K. Varga, Orly Avraham, Ziv Ben-Aharon, Alisa Khramushin, and Ora Schueler-Furman. Harnessing protein folding neural networks for peptide–protein docking. *Nature Communications*, 13(1):176, Jan 2022. ISSN 2041-1723. doi: 10.1038/s41467-021-27838-9. URL https://doi.org/10.1038/s41467-021-27838-9.

Michel van Kempen, Stephanie S. Kim, Charlotte Tumescheit, Milot Mirdita, Jeongjae Lee, Cameron L. M. Gilchrist, Johannes Söding, and Martin Steinegger. Fast and accurate protein structure search

with foldseek. *Nature Biotechnology*, 42(2):243–246, Feb 2024. ISSN 1546-1696. doi: 10.1038/s41587-023-01773-0. URL https://doi.org/10.1038/s41587-023-01773-0.

Fanhao Wang, Yuzhe Wang, Laiyi Feng, Changsheng Zhang, and Luhua Lai. Target-specific de novo peptide binder design with diffpepbuilder. *Journal of Chemical Information and Modeling*, 64(24):9135–9149, 09 2024. ISSN 1549-9596. doi: 10.1021/acs.jcim.4c00975. URL https://doi.org/10.1021/acs.jcim.4c00975.

Yitian Wang, Fanmeng Wang, Angxiao Yue, Wentao Guo, Yaning Cui, and Hongteng Xu. MuCO: Generative peptide cyclization empowered by multi-stage conformation optimization. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=6l8bvRoJZN.

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, Basile I. M. Wicky, Nikita Hanikel, Samuel J. Pellock, Alexis Courbet, William Sheffler, Jue Wang, Preetham Venkatesh, Isaac Sappington, Susana Vázquez Torres, Anna Lauko, Valentin De Bortoli, Emile Mathieu, Sergey Ovchinnikov, Regina Barzilay, Tommi S. Jaakkola, Frank DiMaio, Minkyung Baek, and David Baker. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, Aug 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06415-8. URL https://doi.org/10.1038/s41586-023-06415-8.

Ziyi Yang, Hanyuan Xie, Yinjun Jia, Xiangzhe Kong, Jiqing Zheng, Ziting Zhang, Yang Liu, Lei Liu, and Yanyan Lan. Cpsea: Large-scale cyclic peptide-protein complex dataset for machine learning in cyclic peptide design. In Danielle Belgrave, Cheng Zhang, Laura N. Montoya, Hsuan-Tien Lin, Razvan Pascanu, Piotr Koniusz, Marzyeh Ghassemi, Nancy Chen, Iván Vladimir Meza Ruíz, and Arturo Loaiza-Bonilla (eds.), *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025, NeurIPS 2025, San Diego, CA, USA, December 2-7, 2025 / Mexico City, Mexico, November 30 - December 5, 2025*, 2025. URL http://papers.nips.cc/paper_files/paper/2025/hash/3eec719ab86712d32b065c5977f94ad0-Abstract-Datasets_and_Benchmarks_Track.html.

Jason Yim, Brian L. Trippe, Valentin De Bortoli, Emile Mathieu, Arnaud Doucet, Regina Barzilay, and Tommi Jaakkola. SE(3) diffusion model with application to protein backbone generation. 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. 40001–40039. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/yim23a.html.

Jason Yim, Andrew Campbell, Emile Mathieu, Andrew Y. K. Foong, Michael Gastegger, Jose Jimenez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S. Veeling, Frank Noe, Regina Barzilay, and Tommi Jaakkola. Improved motif-scaffolding with SE(3) flow matching. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856. URL https://openreview.net/forum?id=fa1ne8xDGn.

Jingjie Zhang, Hanqun Cao, Haosen Shi, He Mutian, Yu Wang, Zijun Gao, Fang Wu, Xiaojun Yao, Chang-Yu Hsieh, Sinno Jialin Pan, Pranam Chatterjee, Chunbin Gu, and Pheng-Ann Heng. Geocycler: Reward-aligned 3d diffusion for constraint-conditioned cyclic peptide design, 2026. URL https://arxiv.org/abs/2605.23407.

Huifeng Zhao, Dejun Jiang, Chao Shen, Jintu Zhang, Xujun Zhang, Xiaorui Wang, Dou Nie, Tingjun Hou, and Yu Kang. Comprehensive evaluation of 10 docking programs on a diverse set of protein–cyclic peptide complexes. *Journal of Chemical Information and Modeling*, 64(6):2112–2124, 03 2024. ISSN 1549-9596. doi: 10.1021/acs.jcim.3c01921. URL https://doi.org/10.1021/acs.jcim.3c01921.

Yifan Zhao, Lang Qin, and Jintai Chen. Apcyc: Property-informed design of cyclic peptides via automated cyclization. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, KDD '26, pp. 13020–13031, New York, NY, USA, 2026. Association for Computing Machinery. ISBN 9798400722592. doi: 10.1145/3770855.3818908. URL https://doi.org/10.1145/3770855.3818908.

Xiangxin Zhou, Mingyu Li, Yi Xiao, Jiahan Li, Dongyu Xue, Zaixiang Zheng, Jianzhu Ma, and Quanquan Gu. Designing cyclic peptides via harmonic SDE with atom-bond modeling. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=ERu2ZiAnR7.

Yi Zhou, Connelly Barnes, Jingwan Lu, Jimei Yang, and Hao Li. On the continuity of rotation representations in neural networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5738–5746, 2019. doi: 10.1109/CVPR.2019.00589.

Cheng Zhu, Sen Cao, Tianfeng Shang, Jingjing Guo, An Su, Chengxi Li, and Hongliang Duan. Predicting the structures of cyclic peptides containing unnatural amino acids by highfold2. *Briefings in Bioinformatics*, 26(3):bbaf202, 05 2025. ISSN 1477-4054. doi: 10.1093/bib/bbaf202. URL https://doi.org/10.1093/bib/bbaf202.

# A Model Details

## A.1 Autoencoder architecture and objective

**Residue and pair representations.** Encoder residue features combine amino-acid embeddings, pooled atom embeddings computed from $U$ by an MLP, and sinusoidal residue-index embeddings. Decoder residue features combine $z$ with the same form of index embedding. Both networks receive the explicit frames through directed pair features. For residues $i$ and $j$, define

$$\varphi_{ij} = \left[R_i^\top (x_j - x_i);\ \rho(R_i^\top R_j);\ \mathrm{RBF}(\|x_j - x_i\|)\right], \tag{14}$$

where $\rho$ concatenates the first two columns of a rotation matrix into a six-dimensional vector (Zhou et al., 2019). The autoencoder pair representation is

$$p_{ij} = F_{\mathrm{pair}}\left([F_{\mathrm{geom}}(\varphi_{ij});\ E_{\mathrm{rel}}(\mathrm{clip}(i-j,-32,32))]\right), \tag{15}$$

where $F_{\mathrm{geom}}$ and $F_{\mathrm{pair}}$ are MLPs. Each network constructs these features from its supplied frames. The relative-index embedding preserves the sign of sequence separation.

**Reconstruction losses.** Let $m_i$ indicate a valid peptide residue, and let $m_{i,a}^{\mathrm{atom}}$ select valid non-$C_\alpha$ atoms for coordinate reconstruction. The two reconstruction terms in equation 4 are

$$\mathcal{L}_{\mathrm{seq}} = -\frac{\sum_i m_i \log p_\phi(s_i \mid T, z)}{\sum_i m_i}, \tag{16}$$

$$\mathcal{L}_{\mathrm{atom}} = \frac{\sum_{i,a} m_{i,a}^{\mathrm{atom}} \|\hat{U}_{i,a} - U_{i,a}\|_2^2}{\sum_{i,a} m_{i,a}^{\mathrm{atom}}}. \tag{17}$$

The decoder sets $\hat{U}_{i,\mathrm{CA}} = 0$.

**Latent regularization.** The posterior is a residue-wise diagonal Gaussian, $q_\psi(z \mid s, U, T) = \prod_i \mathcal{N}(z_i; \mu_i, \mathrm{diag}(\sigma_i^2))$. We regularize it toward a standard Gaussian using a per-residue, per-dimension KL floor $\eta = 0.05$ nats:

$$\mathcal{L}_{\mathrm{KL}} = \frac{1}{\sum_i m_i} \sum_i m_i \sum_{k=1}^{d_z} \max\left\{\eta,\ \frac{\mu_{ik}^2 + \sigma_{ik}^2 - 1 - 2\log\sigma_{ik}}{2}\right\}. \tag{18}$$

The coefficient $\beta$ increases linearly from zero to $10^{-4}$ over the first $10^4$ optimization steps.

**Reconstruction evaluation.** Table 6 reports reconstruction on the 1,583 CPCore validation complexes using posterior means and ground-truth frames. Recovery is evaluated per residue, and all-atom RMSD is averaged over complexes.

Table 6: Reconstruction performance of the LaCPlow-Base autoencoder on the CPCore validation set.

| Metric | Value |
|---|---|
| Amino-acid recovery (%) | 100.0 |
| All-atom RMSD (Å) | 0.23 |

### A.2 Flow-matching objective

The translational prior in equation 10 uses $\sigma_x = 0.375\,\mathrm{nm}$ and $\sigma_d = 0.686\,\mathrm{nm}$. For each training example, geometric and latent times are sampled independently from the mixtures used in partially latent flow matching (Geffner et al., 2026; Didi et al., 2026):

$$\begin{aligned} t_T &\sim 0.02\,\mathrm{Unif}(0,1) + 0.98\,\mathrm{Beta}(1.9, 1.0), \\ t_z &\sim 0.02\,\mathrm{Unif}(0,1) + 0.98\,\mathrm{Beta}(1.0, 1.5). \end{aligned} \tag{19}$$

Both times are clipped to $[\varepsilon_{\mathrm{grid}}, 1-\varepsilon_{\mathrm{grid}}]$, with $\varepsilon_{\mathrm{grid}} = 10^{-3}$.

The component losses in equation 7 are normalized by the number of valid scalar components:

$$\begin{aligned} \mathcal{L}_R &= \|\omega_R^\theta - \omega_R^*\|^2_{m\hat{m}} = \frac{\sum_i m_i \hat{m}_i \|\omega_{R,i}^\theta - \omega_{R,i}^*\|_2^2}{3\sum_i m_i \hat{m}_i}, \\ \mathcal{L}_x &= \|v_x^\theta - v_x^*\|^2_m = \frac{\sum_i m_i \|v_{x,i}^\theta - v_{x,i}^*\|_2^2}{3\sum_i m_i}, \\ \mathcal{L}_z &= \|v_z^\theta - v_z^*\|^2_m = \frac{\sum_i m_i \|v_{z,i}^\theta - v_{z,i}^*\|_2^2}{d_z \sum_i m_i}. \end{aligned} \tag{20}$$

The frame-validity mask $\hat{m}_i$ applies only to the rotational loss.

### A.3 Target-conditioned geometric denoiser

**Directed residue relations.** Peptide and target residue tokens are concatenated into the sequence $[\mathrm{peptide} \,\|\, \mathrm{target}]$. For every query–key pair, the geometric features follow equation 14, using the corresponding frames in Table 7. Here $T_i$ denotes a current peptide frame and $T_m^P = (R_m^P, x_m^P)$ a fixed target frame. Signed relative-index embeddings are included for PP pairs. Four learned relation embeddings distinguish PP, PT, TP, and TT:

$$p_{ij} = m_i m_j \left[\widetilde{p}_{ij} + e_{\mathrm{rel}}(\ell_{ij})\right], \qquad \ell_{ij} \in \{\mathrm{PP}, \mathrm{PT}, \mathrm{TP}, \mathrm{TT}\}, \tag{21}$$

where $\widetilde{p}_{ij}$ is the pair representation before adding the relation embedding, and the masks refer to the concatenated sequence of peptide and target residue tokens.

Table 7: Frames and atom-level vectors for the four directed relations. All rows use the same form of residue geometry in equation 14.

| Relation | Query frame | Key frame | Atom-level vector |
|---|---|---|---|
| PP | $T_i$ | $T_j$ | — |
| PT | $T_i$ | $T_m^P$ | $R_i^\top (X_{m,a}^P - x_i)$ |
| TP | $T_m^P$ | $T_i$ | $(R_m^P)^\top (x_i - X_{m,a}^P)$ |
| TT | $T_m^P$ | $T_n^P$ | — |

**Atom-level target conditioning.** For a peptide query $i$ and target residue $m$, let $r_{ima}$ be the target-atom displacement expressed in the query residue frame as defined in equation 8, and let $d_{ima} = \|r_{ima}\|_2$. Atom features are

$$f_{ima} = m_{ma}\, F_{\mathrm{atom}}\left([m_{ma} r_{ima};\ \mathrm{RBF}(d_{ima});\ e_a;\ m_{ma}]\right), \tag{22}$$

where $F_{\text{atom}}$ is an MLP, $e_a$ is an atom-identity embedding for the Atom37 representation, and $m_{ma}$ is the target atom mask. Attention pooling conditioned on residue-pair features normalizes only over valid atoms,

$$\alpha_{ima} = \frac{m_{ma} \exp(\ell_{ima})}{\sum_{a'} m_{ma'} \exp(\ell_{ima'})}, \qquad \ell_{ima} = \frac{q_{im}^\top k_{ima}}{\sqrt{d_q}} - \lambda_d d_{ima}, \tag{23}$$

with learned query and key projections of the pair and atom features. Atoms with $m_{ma} = 0$ are omitted from the denominator. If residue $m$ has no valid atom, the weights and the pooled feature vector are zero. The distance coefficient $\lambda_d$ is constrained to be positive using the softplus function and initialized to $0.1$. The corresponding TP vectors are given in Table 7.

The atom vectors combine relative frame geometry with target-local coordinates $U^P_{m,a}$:

$$r_{ima} = R_i^\top (x^P_m - x_i) + (R_i^\top R^P_m) U^P_{m,a}. \tag{24}$$

Their pooled features are concatenated with the residue-pair geometric features and mapped by an MLP to the peptide–target and target–peptide pair representations.

**Self-conditioning.** The denoiser receives auxiliary endpoint estimates $\mathcal{S} = (\hat{T}^1, \hat{z}^1)$, omitted from the vector-field notation in the main text. These are computed from predicted velocities as

$$\begin{aligned} \hat{R}^1_i &= R_i^{t_T} \operatorname{Exp}\left((1-t_T)[\omega^\theta_{R,i}]_\times\right), \\ \hat{x}^1_i &= x_i^{t_T} + (1-t_T) v^\theta_{x,i}, \\ \hat{z}^1_i &= z_i^{t_z} + (1-t_z) v^\theta_{z,i}. \end{aligned} \tag{25}$$

For latent self-conditioning, $\hat{z}^1$ is passed through a zero-initialized linear projection and concatenated with the peptide residue features. For geometric self-conditioning, frame-pair features constructed from $\hat{T}^1$ are passed through an MLP and concatenated with the PP features; only the final linear layer of this MLP is zero-initialized.

With probability $1/2$ during training, a preliminary evaluation with zero self-conditioning features provides detached estimates for a second evaluation of the same noisy state. Otherwise, training uses zero self-conditioning features. During sampling, these features are zero at the first step and are obtained from the previous step's endpoint estimates thereafter.

### A.3.1 Rigid-motion equivariance

Residue orientations are constructed by Gram–Schmidt orthogonalization: the first axis follows $C - C_\alpha$, the second follows the component of $N - C_\alpha$ orthogonal to the first, and the third is their cross product.

To track the effect of a common rigid motion $(Q, b) \in SE(3)$ before pocket centering, denote original Cartesian coordinates by a tilde. The peptide and target transform as

$$\begin{aligned} \widetilde{x}'_i &= Q\widetilde{x}_i + b, & R'_i &= QR_i, \\ \widetilde{X}^{P\prime}_{m,a} &= Q\widetilde{X}^P_{m,a} + b, & R^{P\prime}_m &= QR^P_m, \qquad c'_P = Qc_P + b. \end{aligned} \tag{26}$$

Thus, centered coordinates satisfy

$$x'_i = (Q\widetilde{x}_i + b) - (Qc_P + b) = Q(\widetilde{x}_i - c_P) = Qx_i, \tag{27}$$

with the same relation for target coordinates.

The pair features are invariant because

$$\begin{aligned} \Delta x'_{ij} &= (QR_i)^\top (Qx_j - Qx_i) = \Delta x_{ij}, \\ \Delta R'_{ij} &= (QR_i)^\top (QR_j) = \Delta R_{ij}, \\ r'_{ima} &= (QR_i)^\top (QX^P_{m,a} - Qx_i) = r_{ima}. \end{aligned} \tag{28}$$

Local atom coordinates $U$ are invariant under the same transformation. The autoencoder posterior is therefore invariant, since its residue and pair inputs are invariant.

For IPA, a learned local point $q_i^\ell$ is transformed into the pocket-centered coordinate system by $y_i^\ell = R_i q_i^\ell + x_i$. In centered coordinates,

$$y_i^{\ell\prime} = Q y_i^\ell. \tag{29}$$

Point distances are unchanged, and mapping aggregated points back into the query frame yields invariant hidden features. The readout in equation 9 consequently satisfies $(\omega_R', v_x', v_z') = (\omega_R, Q v_x, v_z)$.

The body-frame rotational target has the same invariance:

$$\mathrm{Log}\left((QR_i^0)^\top (QR_i^1)\right) = \mathrm{Log}\left((R_i^0)^\top R_i^1\right). \tag{30}$$

The translational target transforms by $Q$, while the latent target is invariant. The Haar rotation prior and isotropic translational prior preserve these transformation laws in distribution. Endpoint estimates in equation 25 transform consistently as well, so self-conditioning preserves the same symmetry.

### A.4 Sampling and time schedules

We use the updates in equation 11 with $N = 100$ steps. A uniform integration grid $\tau_k = k/N$ is mapped to modality-specific times:

$$\begin{aligned} t_M(\tau) &= \varepsilon_{\mathrm{grid}} + (1 - \varepsilon_{\mathrm{grid}}) f_M(\tau), \qquad M \in \{T, z\}, \\ f_T(\tau) &= \frac{1 - 10^{-p\tau}}{1 - 10^{-p}}, \qquad f_z(\tau) = \tau^p. \end{aligned} \tag{31}$$

We set $p = 2$ and $\varepsilon_{\mathrm{grid}} = 10^{-3}$. This default schedule, denoted by $\mathrm{exponential/quadratic}$, uses the exponential and quadratic discretization schedules of Proteína-Complexa (Didi et al., 2026) for geometry and latent variables, respectively. The integration increments are $\Delta t_M^k = t_M(\tau_{k+1}) - t_M(\tau_k)$.

#### A.4.1 Sampling-schedule ablation

We compare $\mathrm{exponential/quadratic}$ with two alternatives:

$$\begin{aligned} \mathrm{uniform/uniform}: &\quad f_T(\tau) = \tau, \quad f_z(\tau) = \tau, \\ \mathrm{uniform/quadratic}: &\quad f_T(\tau) = \tau, \quad f_z(\tau) = \tau^2. \end{aligned}$$

All schedules use the same LaCPlow-Base checkpoint, integration step count, and $\varepsilon_{\mathrm{grid}}$. We generate 100 samples for each of 128 CPCore validation pockets and report all metrics before energy minimization in Table 8. For the LaCPlow variants in Tables 4 and 5, we instead fix the default schedule and vary the model.

Table 8: Sampling-schedule ablation using a fixed LaCPlow-Base checkpoint.

| Schedule | $P_{d_{\mathrm{bond}}}$ ↑ | $P_{C_\beta}$ ↑ | Chirality↑ | Ramachandran | |
|---|---|---|---|---|---|
| | | | | Allowed↑ | Favored↑ |
| exponential/quadratic | **29.2** | <u>97.0</u> | **100.0** | **95.7** | **84.1** |
| uniform/uniform | 26.6 | **97.6** | **100.0** | 94.7 | 82.7 |
| uniform/quadratic | <u>28.5</u> | **97.6** | **100.0** | <u>95.3</u> | <u>84.0</u> |

### A.5 Architecture Hyperparameters and Sampling Time

Both training stages use AdamW with $(\beta_1, \beta_2) = (0.9, 0.95)$, weight decay $10^{-2}$, global gradient-norm clipping at 1.0, and bf16 mixed precision. The autoencoder is trained with a learning rate of $3 \times 10^{-4}$ and a batch size of 32. Its selected checkpoint is frozen for flow training, which uses posterior samples, a learning rate of $2 \times 10^{-4}$, a micro-batch size of 8, and gradient accumulation over four micro-batches. For each stage, validation is performed once per epoch, and training stops after 20 epochs without a decrease in the monitored CPCore validation loss. Table 9 gives the architecture settings.

Table 10 reports trainable parameter counts and average wall-clock sampling time per peptide on a single NVIDIA A6000, measured by generating one peptide at batch size 1 for each of the 84 CPSea_PDB targets. Checkpoint loading is excluded.

Table 9: Architecture hyperparameters.

| | VAE encoder | VAE decoder | Flow denoiser |
|---|---|---|---|
| Attention | pair-biased | pair-biased | gated IPA |
| Token width | 256 | 256 | 256 |
| Pair width | 128 | 128 | 256 |
| Conditioning width | — | — | 256 |
| Blocks | 8 | 8 | 10 |
| Heads | 8 | 8 | 8 |
| Scalar query/key dimension | — | — | 32 |
| Scalar value dimension | — | — | 32 |
| Query/key points | — | — | 6 |
| Value points | — | — | 8 |
| Pair value dimension | — | — | 16 |
| Transition | SwiGLU, $\times 4$ | SwiGLU, $\times 4$ | SwiGLU, $\times 4$ |
| QK LayerNorm | yes | yes | yes |
| Adaptive LN / output scale | none | none | yes |
| Dropout | 0 | 0 | 0 |
| Attention gating | yes | yes | yes |
| Pair / triangular updates | none | none | none |
| Encoder $\log \sigma$ clamp | $[-7, 5]$ | — | — |
| Latent dimension $d_z$ | | 8 | |
| Radial bases | | 32 over $[0, 3.2]$ nm | |
| Relative-index clipping | | $[-32, 32]$ | |

Table 10: Trainable parameter counts and sampling time per peptide. Parameter counts are in millions.

| Model | AE / VAE | Flow / generator | Total | Time (s) |
|---|---|---|---|---|
| PepGLAD | 3.59 | 1.11 | 4.70 | 1.31 |
| APCyc | 3.59 | 1.73 | 5.32 | 0.88 |
| CP-Composer | 3.59 | 1.94 | 5.52 | 1.55 |
| PepFlow | — | 6.88 | 6.88 | 3.13 |
| LaCPlow-Base | 21.32 | 20.73 | 42.05 | 2.01 |
| LaCPlow-Cyc | 21.38 | 20.73 | 42.11 | 2.01 |
| DiffPepBuilder | — | 103.66 | 103.66 | 7.27 |
| Proteína-Complexa | 256.18 | 159.04 | 415.22 | 10.30 |
| MuCO | — | $672.18 + 18.10$ | 690.28 | 3.34 |

### A.6 CYCLIZATION SPECIALIZATION AND NETWORK VARIANTS

Each variant in Tables 4 and 5 has a separately trained autoencoder and flow. All use the sequence, coordinate, and KL terms defined in Appendix A.1, and flow losses are included only for transported modalities. Table 11 summarizes the additional supervision.

**Cyclization supervision.** For a training example $b$ with cyclization type $c_b$, the endpoint distance loss is

$$\mathcal{L}_{\text{close}} = \frac{1}{B}\sum_{b=1}^{B}\left[d^{c_b}_{\text{bond}}(\hat{X}^{(b)}) - d^{c_b}_{\text{bond}}(X^{(b)})\right]^2. \tag{32}$$

Here $d^{c}_{\text{bond}}$ measures the separation of the chemistry-specific bond-forming atoms in residues 1 and $L$ (Appendix B.1). Decoded coordinates are placed in the ground-truth frames before evaluating this loss.

In LaCPlow-Cyc, the encoder receives the ground-truth cyclization type when forming the posterior, $q_\psi(z \mid s, U, T, \mathcal{C})$. A classification head outputs logits $e_{\mathcal{C}}(z) = h_{\mathcal{C}}([\bar{z};\, z_1;\, z_L])$, where $\bar{z}$ is the mean of the latent vectors over valid residues. These logits are supervised by

$$\mathcal{L}_{\text{cls}} = \text{CE}\big(e_{\mathcal{C}}(z), c^\star\big), \tag{33}$$

where $c^\star$ is the ground-truth cyclization type and CE is cross-entropy on the logits. The decoder is conditioned on the probabilities $\text{softmax}(e_{\mathcal{C}}(z))$, $(s, U) = D_\phi\big(T, z, \text{softmax}(e_{\mathcal{C}}(z))\big)$. The autoencoder objective adds $\mathcal{L}_{\text{close}} + \mathcal{L}_{\text{cls}}$ to equation 4.

The variant with an explicit cyclization-class state instead transports $\mathcal{C}$ as an additional Euclidean variable and conditions the decoder on that state. Its class logits occupy a zero-sum subspace of dimension $K_{\mathcal{C}} - 1$, where $K_{\mathcal{C}}$ is the number of classes. The additional flow loss is

$$\mathcal{L}_{\mathcal{C}} = \frac{1}{B}\sum_{b=1}^{B}\frac{\|v^{(b)}_{\mathcal{C}} - u^{(b)}_{\mathcal{C}}\|_2^2}{K_{\mathcal{C}} - 1}, \tag{34}$$

where $v_{\mathcal{C}}$ is the predicted velocity of the cyclization-class logits and $u_{\mathcal{C}}$ is the interpolant velocity. The generated class is $\arg\max \ell_{\mathcal{C}}$, with $\ell_{\mathcal{C}}$ the class logits. This variant adds only $\mathcal{L}_{\text{close}}$ to the shared autoencoder objective.

**Representation with network variants.** The $(R, x, z)$ PairBias variant reconstructs atom coordinates directly in the residue frame. This variant retains the explicit-frame representation while replacing the representation-aligned geometric network with pair-biased attention. The $(R, x, z)$ representation-aligned variant is LaCPlow-Base.

Table 11: Objectives for the representation variants. All rows include $\mathcal{L}_{\text{seq}}$, $\mathcal{L}_{\text{atom}}$, and $\beta\mathcal{L}_{\text{KL}}$.

| Representation | Network | Atom reconstruction target | Additional AE losses | Flow losses |
|---|---|---|---|---|
| $(R, x, z)$ | PairBias | $U$ | — | $\mathcal{L}_R + \mathcal{L}_x + \mathcal{L}_z$ |
| $(R, x, z)$ | Representation-aligned | $U$ | — | $\mathcal{L}_R + \mathcal{L}_x + \mathcal{L}_z$ |
| $(R, x, \mathcal{C}, z)$ | Representation-aligned | $U$ | $\mathcal{L}_{\text{close}}$ | $\mathcal{L}_R + \mathcal{L}_x + \mathcal{L}_z + \mathcal{L}_{\mathcal{C}}$ |
| $(R, x, z) + \mathcal{C}\vert z$ | Representation-aligned | $U$ | $\mathcal{L}_{\text{close}}, \mathcal{L}_{\text{cls}}$ | $\mathcal{L}_R + \mathcal{L}_x + \mathcal{L}_z$ |

Additionally, for LaCPlow-Cyc, the predicted cyclization type is compatible with the decoded endpoint residue identities in 99.9% of samples.

## B CHEMISTRY-SPECIFIC CLOSURE EVALUATION

### B.1 BOND-FORMING ATOM DEFINITIONS AND WINDOW CALIBRATION

The CPSea $C_\beta$ window was derived from $C_\beta$–$C_\beta$ distances in 1,823 disulfide, 263 isopeptide, and 173 main-chain cyclization structures, with the final $[3, 8]$ Å interval selected from kernel-density estimates (Yang et al., 2025). For glycine, a virtual $C_\beta$ constructed using the idealized-geometry formula used in ProteinMPNN (Dauparas et al., 2022) is used. The atom pairs and reported windows are

| Cyclization type | Required residues | Bond-forming atoms | Window |
|---|---|---|---|
| H2T | none | $N_1$–$C_L$ | $[1.2, 1.55]$ Å |
| C2C | Cys–Cys | SG–SG | $[1.84, 2.2]$ Å |
| K2DE | Lys–Asp | NZ–CG | $[1.2, 1.55]$ Å |
| K2DE | Lys–Glu | NZ–CD | $[1.2, 1.55]$ Å |

H2T and K2DE form amide C–N bonds, whose standard crystallographic length is approximately 1.33 Å (Engh & Huber, 1991); disulfide S–S bonds are centered near 2.03 Å (Hazes & Dijkstra, 1988). We broaden these bond-length ranges to cover the variation in CPCore and CPSet (Zhao et al., 2024). Reversed Asp–Lys and Glu–Lys orderings are handled symmetrically. As with $P_{C_\beta}$, $P_{d_{\text{bond}}}$ is a geometric proximity criterion and does not by itself establish a covalent bond.

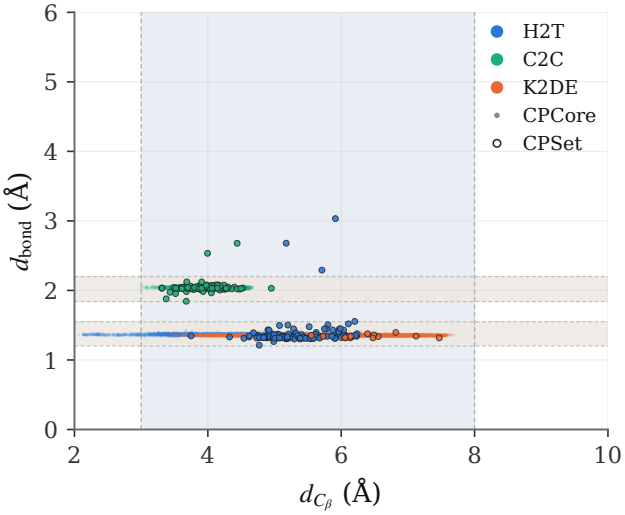


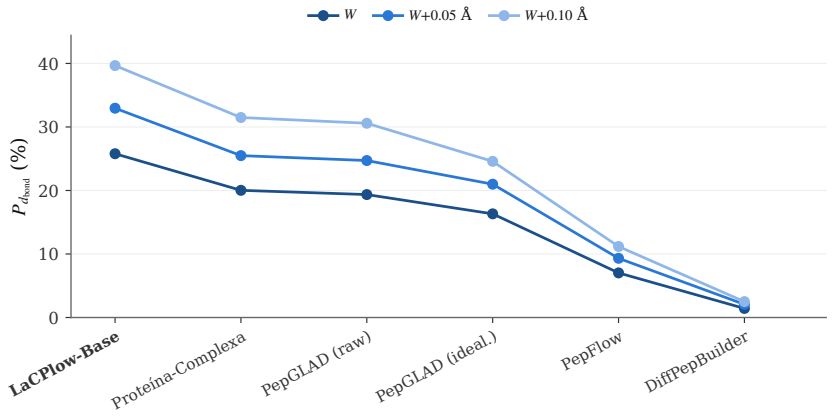

Figure 5: Reference closure geometry in CPCore (71,867 structures; small points) and CPSet (large points). The $d_{\text{bond}}$ window contains 100.0% of CPCore and 97.9% of CPSet structures.

### B.2 Complementarity and window sensitivity

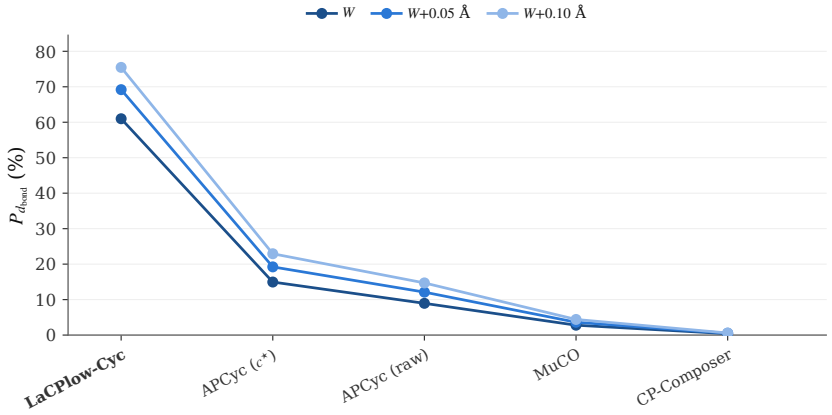
To test whether the bond-forming-atom criterion is more selective than the $C_\beta$-distance criterion on generated linear-peptide structures, we evaluate the 85,469 nominally linear peptides introduced by Wang et al. (2026) from sequences retained after filtering by the backbone mass fraction of the full CPSea dataset (approximately 2.71M structures) (Yang et al., 2025). On this control set, 40.62% of structures satisfy $P_{C_\beta}$, 6.03% satisfy $P_{d_{\text{bond}}}$, and 5.90% satisfy both. The conditional pass rates are

$$P(P_{d_{\text{bond}}}{=}1 \mid P_{C_\beta}{=}1) = \frac{5043}{34717} = 14.53\%,$$
$$P(P_{C_\beta}{=}1 \mid P_{d_{\text{bond}}}{=}1) = \frac{5043}{5151} = 97.90\%.$$

Thus, the chemistry-specific criterion is substantially more selective, while almost every structure that passes it also satisfies the $C_\beta$ window.




(a) Models without dedicated cyclization-specific components and LaCPlow-Base.




(b) Cyclization-aware generators and LaCPlow-Cyc.

Figure 6: Sensitivity of $P_{d_{\text{bond}}}$ to the window boundaries. Curves show the reported window $W$ and intervals obtained by widening both bounds by $0.05$ Å and $0.10$ Å.

Figure 6 shows that the model ranking is stable under modest widening of the chemistry-specific intervals. Figures 7 and 8 provide the corresponding joint distributions of $C_\beta$ separation and $d_{\text{bond}}^{c(X,s)}$ for all evaluated generators.

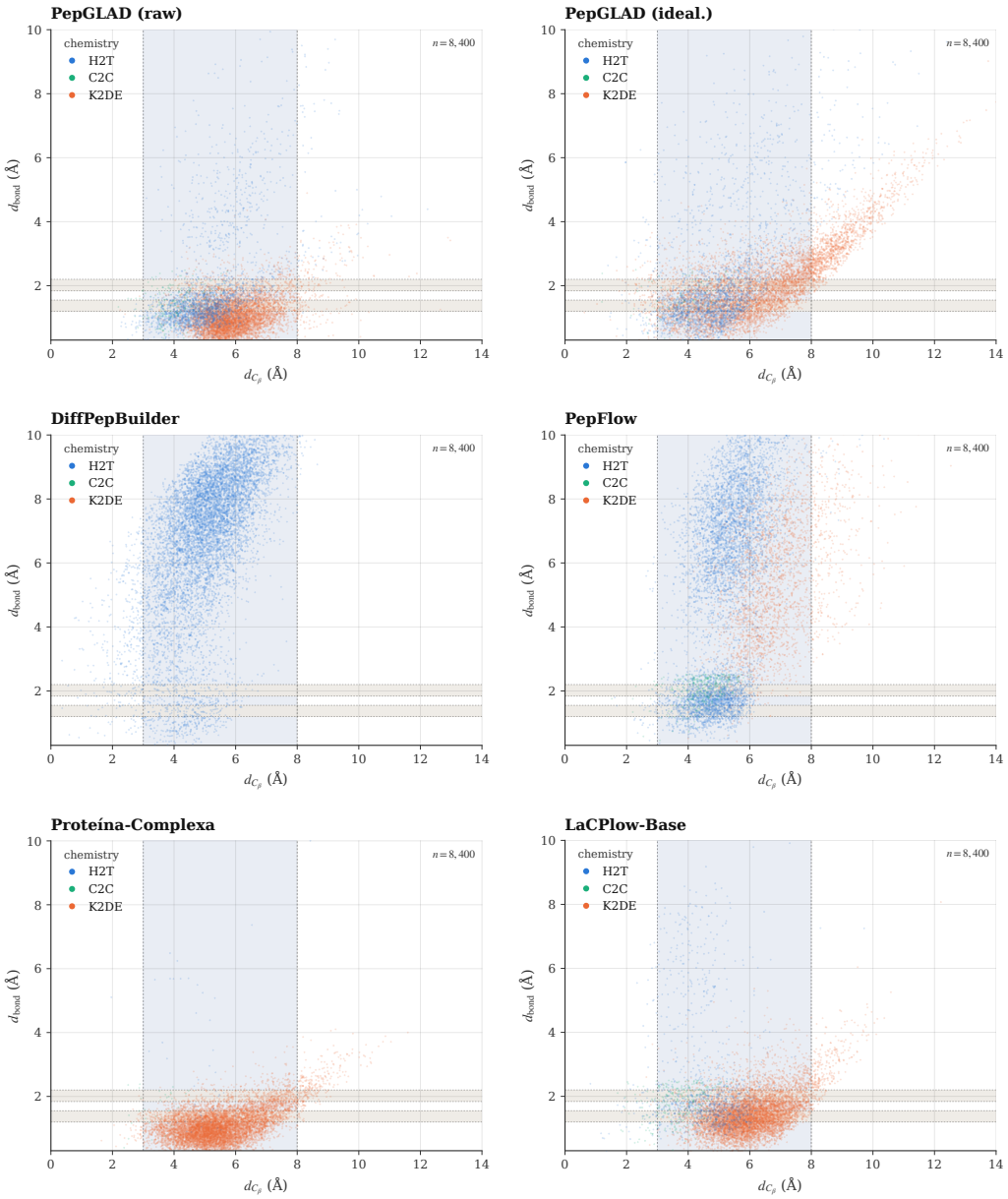



Figure 7: Unrelaxed outputs of models without dedicated cyclization-specific components and LaCPlow-Base on the axes of Figure 5. Each panel contains 8,400 samples.

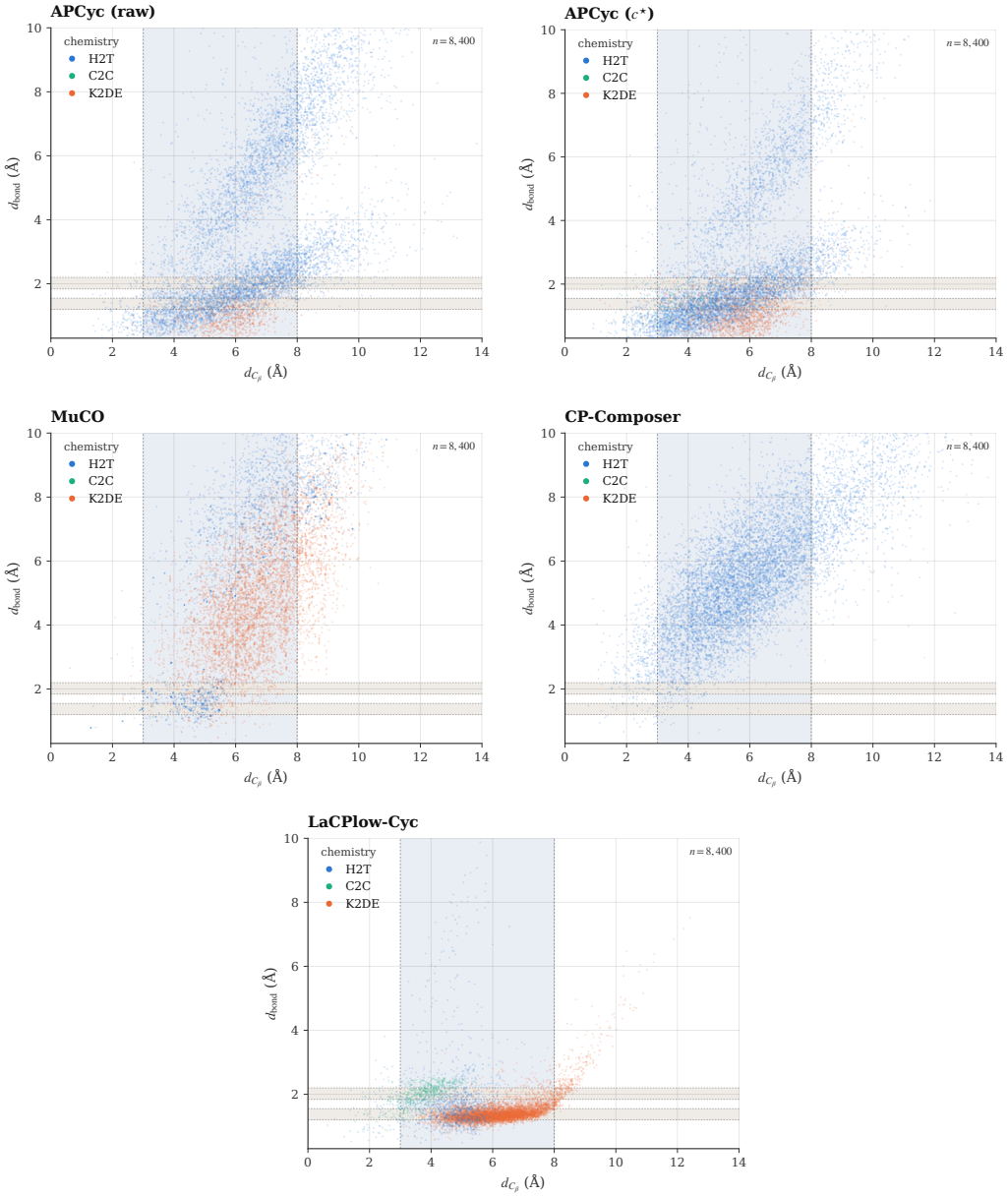



Figure 8: Unrelaxed outputs of cyclization-aware generators and LaCPlow-Cyc on the axes of Figure 5. Each panel contains 8,400 samples.

### B.3 Assigned Cyclization Types

All samples are scored using the assigned cyclization type $c(X, s)$ in equation 12. This assignment depends only on the generated sequence and coordinates, not on a target annotation or the predicted cyclization type. Figure 9 summarizes these assignments before energy minimization.

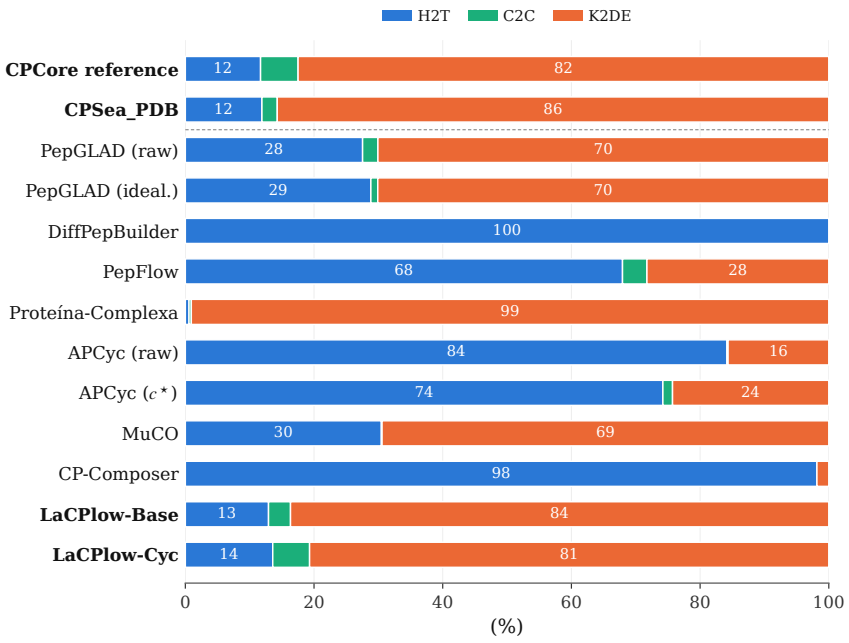



Figure 9: Distribution of cyclization types assigned by $c(X, s)$ before energy minimization. LaCPlow-Base and LaCPlow-Cyc are the bottom two rows.

## C Additional Diversity Analysis

The diversity score combines conformational variation with target-relative placement. For each target, peptide $C_\alpha$ coordinates are clustered by single linkage at $4$ Å using RMSD without superposition in the target-fixed coordinate frame. To isolate conformation, we additionally report the mean pairwise RMSD before and after optimal rigid superposition:

$$\overline{\mathrm{RMSD}}_{\mathrm{unaligned}} = \frac{2}{N(N-1)} \sum_{i<j} \mathrm{RMSD}(X_i, X_j),$$

$$\overline{\mathrm{RMSD}}_{\mathrm{aligned}} = \frac{2}{N(N-1)} \sum_{i<j} \min_{Q \in SO(3),\, t} \mathrm{RMSD}(QX_i + t, X_j).$$

The aligned quantity removes global rigid motion. Because RMSD does not admit an additive decomposition, we report the unaligned-minus-aligned RMSD only as a descriptive difference, not as an independent pose component. Percentages are relative to the best row in each column, over the $84$ targets in CPSea_PDB.

Sequence diversity is computed independently by single-linkage clustering at a BLOSUM62 global-alignment identity distance threshold of $0.4$. We report the number of clusters divided by the number of generated sequences using the same target-wise aggregation. MuCO is omitted because its sequence is supplied as input.

Table 12: Diversity on the samples of Tables 1 and 2. The structural- and sequence-diversity columns report single-linkage clustering scores. The unaligned and aligned columns report mean pairwise $C_\alpha$ RMSD before and after rigid superposition.

| | structural diversity | | Sequence | | Mean pairwise $C_\alpha$ RMSD (Å) | | | |
|---|---|---|---|---|---|---|---|---|
| Model | div | % best | Seq. | % best | unaligned | aligned | % best | unaligned − aligned |
| PepGLAD (raw) | 0.62 | 66 | 0.997 | **100** | 8.600 | 3.938 | 97 | 4.662 |
| PepGLAD (ideal.) | 0.62 | 66 | 0.997 | **100** | 8.506 | 3.926 | 96 | 4.580 |
| DiffPepBuilder | 0.17 | 18 | 0.425 | 43 | 6.706 | 3.290 | 81 | 3.416 |
| PepFlow | 0.50 | 53 | 0.945 | 94 | 8.785 | **4.074** | **100** | 4.711 |
| APCyc (raw) | 0.56 | 60 | 0.998 | **100** | 7.533 | 4.017 | 99 | 3.516 |
| APCyc ($c^\star$) | 0.56 | 60 | 0.999 | **100** | 7.590 | 4.027 | 99 | 3.563 |
| Proteína-Complexa | 0.12 | 13 | 0.170 | 17 | 7.306 | 3.185 | 78 | 4.121 |
| MuCO$^\star$ | 0.82 | 87 | — | — | 9.517 | 3.635 | 89 | 5.882 |
| CP-Composer | **0.94** | **100** | **1.000** | **100** | **10.098** | 3.950 | 97 | 6.148 |
| LaCPlow-Base | 0.20 | 21 | 0.955 | 96 | 5.818 | 3.583 | 88 | **2.235** |
| LaCPlow-Cyc | 0.22 | 23 | 0.937 | 94 | 6.011 | 3.581 | 88 | 2.430 |

# D Evaluation Protocol

## D.1 Datasets and splits

Training uses CPCore, the 71,867-complex high-quality subset of CPSea (Yang et al., 2025). We cluster complexes with Foldseek `easy-multimercluster` (van Kempen et al., 2024) using multimer TM-score 0.65, chain TM-score 0.5, interface lDDT 0.65, and `alignment-type` 2, obtaining 30,819 clusters. We randomly select 616 clusters (1,583 complexes) to form the validation set; the remaining 70,284 complexes are used for training.

Evaluation uses the high-resolution CPSea_PDB subset processed following Wang et al. (2026). Retaining complexes where the cyclic peptide backbone exceeds 20% of the total molecular mass leaves 85 complexes. We then filter CPSea_PDB receptors against CPCore with MMseqs2 (Steinegger & Söding, 2017) at 40% identity and 95% coverage, following Yang et al. (2025). Removing one overlapping receptor leaves 84 held-out targets.

## D.2 Evaluation Stages, Filtering, and Aggregation

Chain L denotes the peptide and chain R the target pocket. Unless stated otherwise, metrics are averaged within each target and then averaged uniformly over the 84 targets. The $dG < 0$ filter removes 716 (PepGLAD raw), 762 (PepGLAD idealized), 379 (DiffPepBuilder), 1,069 (PepFlow), 655 (APCyc raw), 672 (APCyc $c^\star$), 1,237 (Proteína-Complexa), 2,410 (CP-Composer), 238 (LaCPlow-Base), and 255 (LaCPlow-Cyc). MuCO has no receptor, so none of its samples are removed.

## D.3 Stereochemical and structural metrics

**Chirality.** For each residue with N, CA, C, and CB atoms, we compute

$$\chi = \big((\mathrm{N} - \mathrm{CA}) \times (\mathrm{C} - \mathrm{CA})\big) \cdot (\mathrm{CB} - \mathrm{CA}),$$

where $\chi < 0$ denotes a D residue. The reported score is $(1 - \text{number of D residues}/\text{number of evaluated residues})$, averaged within target and then over targets. Glycine and residues missing any required atom are omitted.

**Ramachandran.** We obtain $\phi/\psi$ angles of the peptide and evaluate allowed and favored regions using PyRAMA (Lovell et al., 2003). The allowed and favored rates are calculated as the fractions of valid $(\phi, \psi)$ pairs falling within the allowed and favored regions.

**Post-minimization comparison.** Table 13 reports chirality and Ramachandran quality after energy minimization.

Table 13: Chirality and Ramachandran quality after energy minimization for the samples in Tables 1 and 2.

| Model | Chirality↑ | Ramachandran | |
|---|---|---|---|
| | | Allowed↑ | Favored↑ |
| *CPSea_PDB (ref)* | *100.0* | *97.4* | *90.2* |
| PepGLAD (raw) | 49.8 | 62.9 | 36.5 |
| PepGLAD (ideal.) | 99.5 | 75.8 | 44.6 |
| DiffPepBuilder | 97.2 | 89.5 | 70.7 |
| PepFlow | 98.8 | 88.3 | 67.6 |
| APCyc (raw) | 49.8 | 59.7 | 33.4 |
| APCyc ($c^\star$) | 49.5 | 59.2 | 33.6 |
| Proteína-Complexa | 99.3 | 89.7 | 71.2 |
| MuCO$^\star$ | 50.6 | 79.8 | 59.2 |
| CP-Composer | 50.0 | 63.5 | 34.4 |
| LaCPlow-Base | **99.8** | **92.6** | 74.6 |
| LaCPlow-Cyc | **99.8** | 92.5 | **75.0** |

### D.4 Affinity, interactions, and molecular properties

**Rosetta $dG$ and energy success.** We compute $dG$ with the Rosetta all-atom energy function (Alford et al., 2017) through PyRosetta (Chaudhury et al., 2010). Energy success is $dG < 0$, and $dG$ is summarized by its median to limit the effect of structures with severe steric clashes.

**Vina score.** AutoDock Vina 1.1.2 (Trott & Olson, 2010) is run in `score_only` mode without redocking. Scores are summarized by their median.

**Interaction composition.** PLIP (Schake et al., 2025) is run as `plip -f FILE --peptides L -x`. For each structure, the count of an interaction type is divided by the total detected interaction count, and these per-structure fractions are averaged. Structures with no detected interaction are omitted.

Table 14: Protein–peptide interaction composition and LogP for the samples in Tables 1 and 2.

| Model | Hydrophobic | H-bond | Salt bridge | LogP |
|---|---|---|---|---|
| *CPSea_PDB (ref)* | *52.0* | *37.8* | *8.7* | *24.4* |
| PepGLAD (raw) | 55.6 | 35.3 | 6.3 | 50.8 |
| PepGLAD (ideal.) | 55.8 | 34.9 | 6.2 | 50.4 |
| DiffPepBuilder | 58.0 | 30.7 | 9.2 | 64.4 |
| PepFlow | 45.4 | 48.2 | 4.8 | 18.7 |
| APCyc (raw) | 57.6 | 32.5 | 7.2 | 62.7 |
| APCyc ($c^\star$) | 56.1 | 34.1 | 7.2 | 58.5 |
| Proteína-Complexa | 58.5 | 29.7 | 8.8 | 91.0 |
| CP-Composer | 47.1 | 40.7 | 10.7 | 19.8 (24.2) |
| LaCPlow-Base | 53.2 | 34.8 | 9.7 | 37.8 |
| LaCPlow-Cyc | 53.9 | 35.9 | 8.0 | 44.9 |

**GRAVY and logP.** GRAVY (Kyte & Doolittle, 1982) is the mean Kyte–Doolittle hydropathy of the chain L sequence. LogP is estimated by RDKit (Landrum et al., 2026). We report the fractions satisfying GRAVY $< 0$ and LogP $> -6$, respectively.

### D.5 Sequence–structure self-consistency

Using HighFold2 (Zhu et al., 2025), scRMSD is the Kabsch-superposed $C_\alpha$ RMSD between the minimized generated peptide and the relaxed rank-1 HighFold2 prediction. Samples are grouped

as H2T or C2C according to the structure-assigned cyclization type, and K2DE is omitted because HighFold2 has no isopeptide mode. Scores are first averaged within each target and then averaged across targets using an unweighted mean.

### D.6 Structural Changes During Energy Minimization

For all pre- and post-minimization comparisons, atoms in chain L are matched in the shared target frame. Cartesian backbone and side-chain RMSDs are root-mean-square distances of those matched atoms. The assigned cyclization type is read from the pre-minimization structure and is not recomputed afterward. $\Delta C_\beta$ and $\Delta d_{\mathrm{bond}}$ are the absolute changes in the endpoint $C_\beta$ distance and in the distance between that same bond-forming atom pair.

Backbone bond-angle RMSD pools N–$C_\alpha$–C, $C_\alpha$–C–$N_{\mathrm{next}}$, and C–$N_{\mathrm{next}}$–$C_{\alpha,\mathrm{next}}$. Side-chain bond-angle RMSD pools N–$C_\alpha$–$C_\beta$ and consecutive heavy-atom triples along the side-chain atom sequences used to define $\chi$. Backbone torsion RMSD pools $\phi$, $\psi$, and $\omega$; side-chain torsion RMSD pools $\chi_1$–$\chi_4$. Angles across the cyclization junction are omitted so that the same angle set is compared before and after minimization. For a periodic angle $\alpha$, we use

$$\delta(\alpha) = \mathrm{wrap}_{(-180^\circ,180^\circ]}\big(\alpha_{\mathrm{post}} - \alpha_{\mathrm{pre}}\big), \mathrm{TorRMSD} = \sqrt{\frac{\sum_i \delta(\phi_i)^2 + \sum_j \delta(\psi_j)^2 + \sum_k \delta(\omega_k)^2}{N_\phi + N_\psi + N_\omega}} \tag{35}$$

with the analogous pooled expression for $\chi_1$–$\chi_4$. We pool the squared angular differences over all valid observations before taking the square root.

**Per-angle breakdown.** Tables 15 and 16 report the individual components of the pooled angular columns in Table 3.

Table 15: Backbone bond-angle and torsion RMSDs for each component of the pooled backbone columns in Table 3. Values are in degrees.

| | Backbone bond-angle RMSD (°) | | | Backbone torsion RMSD (°) | | |
|---|---|---|---|---|---|---|
| Model | N–CA–C | CA–C–N | C–N–CA | $\phi$ | $\psi$ | $\omega$ |
| PepGLAD (raw) | 24.3 | 19.1 | 24.2 | 71.5 | 71.0 | 51.9 |
| PepGLAD (ideal.) | 5.1 | 24.4 | 22.5 | 66.1 | 65.2 | 42.1 |
| DiffPepBuilder | <u>4.6</u> | 18.5 | 18.7 | 49.6 | 49.6 | 24.4 |
| PepFlow | 4.7 | 9.1 | 8.7 | 55.0 | 58.5 | 19.3 |
| APCyc (raw) | 31.3 | 24.1 | 26.6 | 78.0 | 78.2 | 65.2 |
| APCyc ($c^\star$) | 31.0 | 24.0 | 26.6 | 77.5 | 78.3 | 64.5 |
| Proteína-Complexa | 7.5 | 23.0 | 21.2 | 53.7 | 55.4 | 26.4 |
| MuCO$^\star$ | 11.2 | 15.8 | 7.9 | 54.6 | 51.6 | 20.0 |
| CP-Composer | 31.6 | 25.2 | 28.2 | 86.4 | 82.7 | 63.2 |
| LaCPlow-Base | **3.5** | <u>5.3</u> | **5.8** | <u>41.6</u> | <u>45.2</u> | **16.7** |
| LaCPlow-Cyc | **3.5** | **5.2** | <u>5.9</u> | **41.3** | **44.7** | <u>17.0</u> |

Table 16: Side-chain torsion RMSDs for each $\chi$ component of the pooled side-chain torsion column in Table 3. Values are in degrees. DiffPepBuilder does not generate side-chain atoms beyond $C_\beta$.

| Model | $\chi_1$ | $\chi_2$ | $\chi_3$ | $\chi_4$ |
|---|---|---|---|---|
| PepGLAD (raw) | 68.2 | 76.1 | 65.5 | 57.0 |
| PepGLAD (ideal.) | 51.4 | 66.1 | 76.1 | 80.0 |
| DiffPepBuilder | — | — | — | — |
| PepFlow | 38.6 | 50.2 | 77.8 | 71.5 |
| APCyc (raw) | 73.9 | 77.7 | 68.0 | 56.8 |
| APCyc ($c^\star$) | 73.5 | 78.0 | 72.8 | 59.4 |
| Proteína-Complexa | 46.4 | 62.0 | 73.6 | 89.2 |
| MuCO$^\star$ | 51.7 | 64.5 | **44.1** | **48.5** |
| CP-Composer | 77.3 | 77.3 | 74.9 | 50.6 |
| LaCPlow-Base | 35.3 | 50.1 | 56.7 | 57.4 |
| LaCPlow-Cyc | **34.0** | **48.3** | 59.2 | 59.9 |

Table 17: Structural displacement during energy minimization for the representations of Tables 4 and 5.

| | Cartesian (Å) | | Angular RMSD (°) | | | | Closure | |
|---|---|---|---|---|---|---|---|---|
| Representation | BB ↓ | SC ↓ | BB bond ↓ | SC bond ↓ | BB tor ↓ | SC tor ↓ | $\Delta d_{\text{bond}}$ (Å) ↓ | $\Delta C_\beta$ (Å) ↓ |
| $(R, x, z)$ | 2.15 | 2.57 | **5.0** | **6.7** | 37.3 | 47.5 | 0.53 | **0.87** |
| $(R, x, \mathcal{C}, z)$ | 2.38 | 2.82 | 7.2 | 7.9 | 41.8 | 50.0 | 0.44 | 1.13 |
| $(R, x, z) + \mathcal{C}\|z$ | **2.11** | **2.52** | 5.1 | 6.8 | **37.0** | **46.7** | **0.26** | 0.92 |

# E Latent–Geometry Coupling and Optimization

## E.1 Midpoint latent-intervention protocol

We evaluate latent–geometry coupling on 128 randomly sampled CPCore validation pockets. For each pocket, we first generate one unguided trajectory

$$(T^0, z^0) \to \cdots \to (T^q, z^q) \to \cdots \to (T^1, z^1),$$

using $N = 100$ Euler steps and record the state at the midpoint $q = 0.5$.

For this diagnostic, geometry and latent time are aligned as $t_T = t_z = \tau$. The shared uniform schedule has generation quality comparable to the default sampling schedule (Table 8) and leaves substantial remaining integration time for both the explicit and latent states after the intervention.

The unperturbed endpoint is decoded as

$$(s^{\text{base}}, U^{\text{base}}) = D_\phi(T^1, z^1), \qquad X^{\text{base}}_{i,a} = R^1_i U^{\text{base}}_{i,a} + x^1_i + c_P. \tag{36}$$

All intervention quantities are evaluated before energy minimization.

At the midpoint, $z^q$ is standardized using the per-latent-dimension mean and standard deviation of the latent state at that integration time. Let $\delta$ be the perturbation in these standardized coordinates and let $m_i$ mark a valid residue. The budget $\rho$ is the root-mean-square norm

$$\text{RMS}(\delta) = \sqrt{\frac{\sum_{i,k} m_i \delta^2_{ik}}{d_z \sum_i m_i}}, \qquad \text{RMS}(\delta) = \rho. \tag{37}$$

We sample a Rademacher direction $u$ with $\text{RMS}(u) = 1$ and set

$$\delta = \rho u, \qquad \rho \in \{0.05, 0.1, 0.25, 0.5\}.$$

The explicit state $T^q$ is never directly modified. For each pocket and each $\rho$, we use $M = 2$ independent directions and evaluate both signs $\{\pm u\}$.

We compare two intervention conditions using the same perturbation direction and budget. In the *joint-propagation* condition, integration resumes from $(T^q, z^{q,\rho})$, yielding $(T^{1,\rho}_{\text{joint}}, z^{1,\rho}_{\text{joint}})$. In the *decoder-only* condition, the unperturbed endpoint frame $T^1$ is held fixed and $z^1$ is perturbed using the corresponding per-channel mean and standard deviation at the endpoint, after which decoding is performed without further integration. Consequently, $\Delta x = \Delta R = 0$ in the decoder-only condition by construction.

Explicit geometric response is measured in the target-fixed coordinate frame using equation 13. For each pocket and perturbation budget, statistics are first aggregated as the median over the two directions and both signs, and the reported curve is then the median over the 128 pockets with IQR (25th–75th percentile) bands.

### E.2 Large-Magnitude Latent Perturbation with Isotropic Latent Budgets

The main latent-intervention analysis uses $\rho \leq 0.5$. We repeat the protocol of Appendix E.1 on the same 128 pockets with the extended budgets

$$\rho \in \{0.05, 0.1, 0.25, 0.5, 1, 2, 4\}.$$

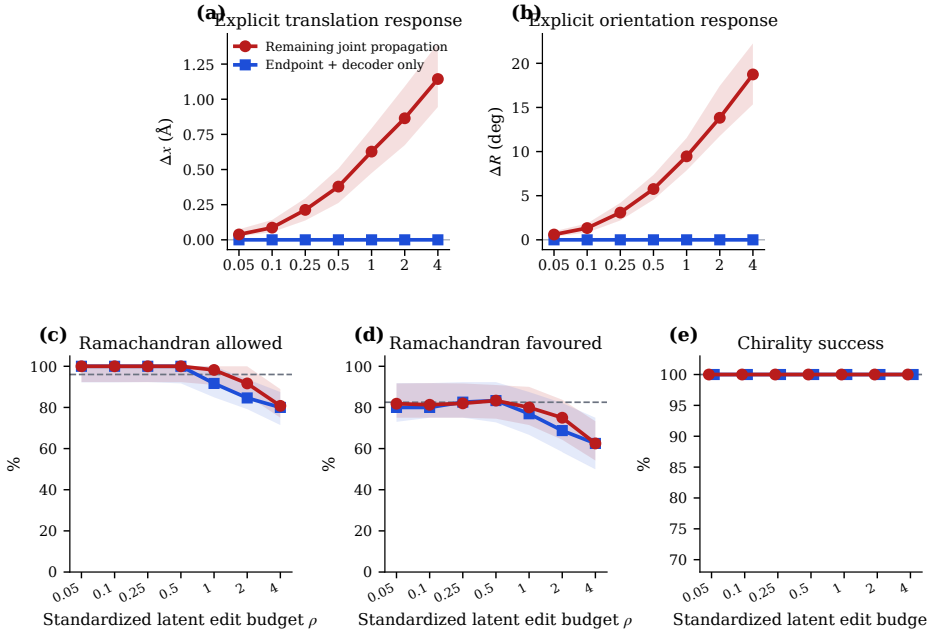



Figure 10: **Large-magnitude latent perturbation stress test.**

The additional panels report Ramachandran allowed and favored rates and the peptide L-residue fraction; dashed lines denote the unguided baseline. Chirality remains at $100\%$ across all budgets, whereas Ramachandran allowed and favored fractions decrease for $\rho \geq 1$. We therefore restrict the main coupling analysis to $\rho \leq 0.5$, where the perturbation probes local latent–geometry coupling.

### E.3 Latent-Editing Protocol and Post-Edit Quality

We use the midpoint intervention setting of Appendix E.1 and optimize only the latent state before resuming joint propagation. The surrogate objective is

$$S_{\text{iface}} = S_{\text{match}} - S_{\text{rep}}. \tag{38}$$

Residue distance $d_{ij}$ is the minimum heavy-atom distance between peptide residue $i$ and pocket residue $j$. The contact weight is $\sigma\big((4.5 - d_{ij})/0.3\big)$, with $\sigma$ the logistic function. $S_{\text{match}}$ is the mean

weight of a maximum-weight bipartite matching of cardinality $k$, in which each peptide residue and each pocket residue is used at most once. The integer $k$ is the training-set median number of distinct 4.5 Å contacts at the generated peptide length. $S_{\text{rep}}$ is a van der Waals overlap penalty.

We optimize the latent with direction-only SPSA (Spall, 1992) for $K = 6$ steps in standardized coordinates. Each step draws a Rademacher direction $u$ with $\text{RMS}(u) = 1$ and queries the score at displacements $\pm cu$ with probe radius $c = 0.05$. The estimated gradient is $\hat{g} = \big(S_+ - S_-\big)u/(2c)$. The update direction is $\hat{g}/(\text{RMS}(\hat{g}) + 10^{-12})$, and the step length is $\alpha = \rho/K$. The displacement from the captured latent is then projected onto $\text{RMS}(\delta) \le \rho$. Each step uses two score queries. The experiments use $\rho \in \{0.05, 0.1, 0.25, 0.5\}$.

For each optimized sample, the random-control perturbation is matched to the displacement magnitude actually realized by SPSA on that sample.

After the edited latent is propagated to the endpoint, we evaluate both the surrogate interface score and unrelaxed Rosetta $dG$ (Figure 4). Rosetta $dG$ is not used as the optimization objective. Because all scores in that figure are computed before energy minimization, decreases in $dG$ indicate lower Rosetta interface energies before minimization rather than experimentally validated improvements in binding affinity.

**Geometric Quality During Latent Optimization.** Figure 11 reports the closure-distance and stereochemical metrics for the structures obtained by integrating each edited midpoint state to the endpoint and decoding after every SPSA step. Panels (a) and (b) report $P_{d_{\text{bond}}}$ and $P_{C_\beta}$, respectively, for the assigned cyclization type $c(X, s)$. Panels (c)–(e) report chirality and Ramachandran allowed and favored rates, all evaluated before energy minimization. The dashed line is the unedited baseline ($k = 0$).

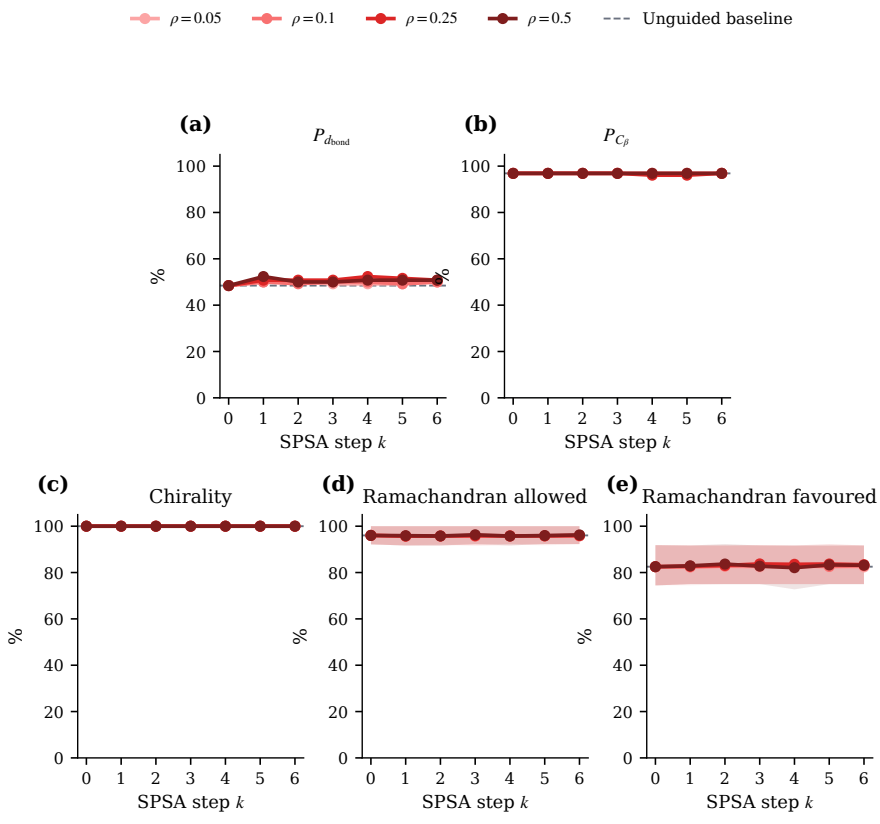



Figure 11: **SPSA step-wise geometry.** Closure and stereochemical quality during latent optimization on 128 CPCore validation pockets. Metrics are evaluated on unrelaxed endpoints after each SPSA step. Dashed lines indicate the unedited baseline.