STAPLEBRIDGE: FINITE-SUPPORT INTERVENTION CONTROL FOR STAPLED PEPTIDE OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Hydrocarbon stapling covalently links two peptide side chains to constrain peptide conformation and can improve proteolytic stability and cellular uptake. Recent computational cyclic-peptide design has largely focused on *de novo* generation, yet redesigning a functional peptide from scratch can perturb sequence and structural context important for its function. Constrained local editing requires a framework to select chemically and geometrically valid modifications despite scarce paired linear-to-stapled supervision. To address this, we introduce **StapleBridge**, a chemistry-aware discrete Schrödinger-control framework for optimizing existing peptides through hydrocarbon stapling. For each peptide, StapleBridge constructs a finite set of executable interventions and maps each to a canonical minimal-edit product, making the optimal intervention distribution exactly computable from terminal utility. Distilling this distribution into an amortized controller provides plan-level supervision without paired examples and enables efficient intervention selection on unseen peptides. On test leads supported by the fixed, property-independent intervention catalog, StapleBridge improves predicted permeability for 81.13% of leads, with a mean permeability gain of 0.169 compared with 0.140 for the strongest matched plan-level baseline. Cross-predictor permeability evaluation and an orthogonal, post-selection 3D evaluation support our observed predicted gains and structural compatibility, while computational experiments on N-methylation and disulfide cyclization support extension to additional chemistries. Overall, StapleBridge enables property-guided refinement of existing peptide leads by learning which feasible chemical edits to apply without paired linear-to-stapled supervision. Code is available at `https://anonymous.4open.science/r/Staplebridge-A05A/`.

1 INTRODUCTION

Cyclic peptides are attractive therapeutic scaffolds because conformational constraint can improve structural preorganization and proteolytic stability while retaining the molecular recognition capacity of peptides (Zorzi et al., 2017; Vinogradov et al., 2019). Recent computational methods have made de novo cyclic-peptide design increasingly practical by incorporating geometric, cyclization-site, topological, and physicochemical constraints (Rettie et al., 2025; Jiang et al., 2025a; Zhang et al., 2026a; Zhao et al., 2026; Sun et al., 2026). These methods, however, primarily generate new molecules. When a functional peptide is already available, redesigning it from scratch may perturb sequence and structural context that contributes to its function, motivating approaches that introduce targeted chemical modifications while preserving the parent peptide.

Hydrocarbon stapling provides a localized and chemically interpretable way to modify an existing peptide while largely retaining its surrounding molecular scaffold (Schafmeister et al., 2000; Verdine & Hilinski, 2012; Walensky & Bird, 2014). Staple placement and sequence context affect cellular uptake and target engagement (Bird et al., 2016; Chandramohan et al., 2024; Reyes et al., 2025). Choosing an effective intervention therefore requires compatible staple chemistry, anchor spacing, and local geometry (Tan et al., 2016), while standardized paired data linking linear peptides to optimized stapled products remain scarce. Consequently, the desired stapled product is typically unknown during training, making direct supervised learning difficult. The key challenge is therefore to learn which feasible intervention to apply without direct supervision of the desired endpoint.

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.

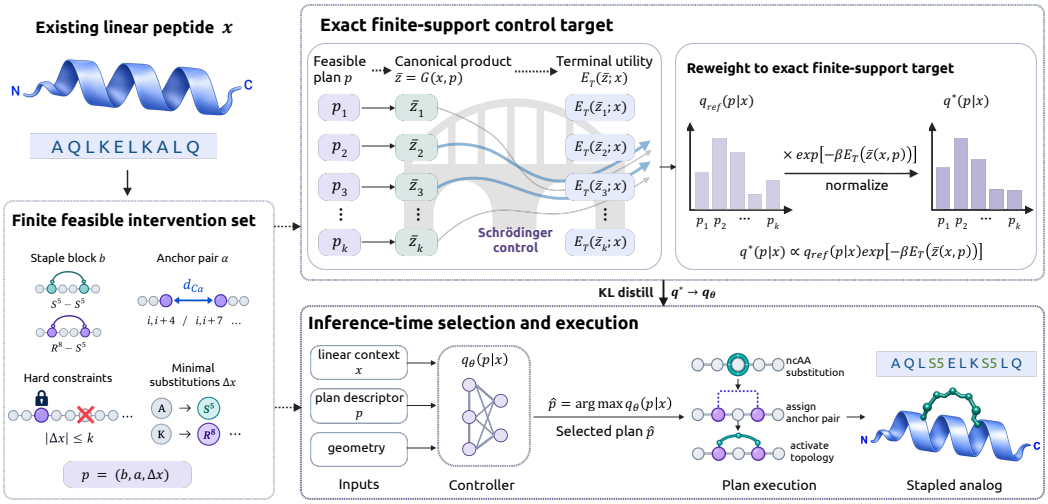

Figure 1: **Overview of StapleBridge.** Hydrocarbon stapling is represented by a finite, lead-specific space of chemically feasible interventions. The exact intervention distribution $q^\star(p \mid x)$ is computed during training and distilled into $q_\theta(p \mid x)$ for efficient plan selection and constraint-preserving execution on unseen peptide leads.

KL-regularized stochastic control provides a distribution over candidate modifications by balancing terminal utility against deviation from a specified reference process (Todorov, 2006; Chen et al., 2021; Garg et al., 2024). In this work, we introduce **StapleBridge**, a discrete Schrödinger-control framework for chemically constrained modification of existing peptides through hydrocarbon stapling (Figure 1). For each peptide, StapleBridge constructs a finite set of feasible stapling interventions and maps each intervention to a canonical minimal-edit product. This finite plan–product correspondence makes the soft-optimal intervention distribution exactly computable under the specified reference and terminal objective, which can then be distilled into a peptide-conditioned controller for inference.

On held-out leads, StapleBridge improves predicted permeability for 81.1% of supported peptides. Post-hoc permeability and structural analyses further show higher native-contact retention and lower interface deviation, while extensions to N-methylation and disulfide cyclization support the portability of the formulation. Our principal contributions are as follows.

1. We define a finite, peptide-specific intervention-control problem for localized hydrocarbon stapling.
2. We show that the canonical plan–product structure yields an exactly computable soft-terminal control target without requiring paired linear-to-stapled products.
3. We distill the training-time terminal-property target into a context-aware controller, enabling intervention selection on unseen peptides without terminal-property queries at inference.

2 Related Work

Peptide and cyclic-peptide design. Recent generative models support controllable biological sequence and therapeutic peptide design (Gruver et al., 2023; Stark et al., 2024; Tang et al., 2025b). Structure-aware and cyclic-peptide methods further incorporate sequence–structure compatibility, cyclization geometry, topology, and physicochemical objectives (Kong et al., 2024; Li et al., 2024; Jiang et al., 2025a; Yang et al., 2026; Zhang et al., 2026a; Wang et al., 2026; Zhao et al., 2026; Rettie et al., 2025; Sun et al., 2026). These approaches mainly address *de novo* generation. Our setting modifies an existing functional peptide through a localized hydrocarbon-stapling intervention.

Hydrocarbon stapling. Hydrocarbon stapling constrains peptide conformation through side-chain crosslinks and can improve helicity, proteolytic stability, and cellular uptake, with effects depending strongly on staple placement and sequence context (Schafmeister et al., 2000; Verdine & Hilinski,

2012; Walensky & Bird, 2014; Bird et al., 2016; Tan et al., 2016; Reyes et al., 2025). Tools such as StaPep support structural analysis and permeability prediction of stapled peptides (Wang et al., 2024; Chandramohan et al., 2024), but do not directly learn which feasible stapling intervention should be selected for a given peptide.

Schrödinger bridges and stochastic control. Schrödinger bridges and KL-regularized control define controlled dynamics relative to a reference process (Schrödinger, 1932; Léonard, 2013; Todorov, 2006; Chen et al., 2021). Recent methods learn diffusion bridges through iterative fitting, score and flow matching, or optimal bridge matching (De Bortoli et al., 2021; Shi et al., 2023; Tong et al., 2024; Gushchin et al., 2024). Extensions incorporate state costs, soft endpoint constraints, and reward-guided diffusion control (Liu et al., 2024; Garg et al., 2024; Han et al., 2025), while bridge distillation accelerates inference (Gushchin et al., 2025). Discrete formulations address graph transformation, categorical data, and peptide sequences (Kim et al., 2025; Ksenofontov & Korotin, 2025; Goel & Chatterjee, 2026). BranchSBM learns branching dynamics with branch-specific growth, and EntangledSBM learns interacting particle dynamics (Tang et al., 2026; 2025a). StapleBridge operates on a finite, peptide-specific intervention space, where a canonical plan–product mapping makes the soft-terminal optimum directly computable and suitable for amortized distillation.

3 METHOD

3.1 LEAD-SPECIFIC FEASIBLE INTERVENTION SPACE

Let $x = (x_1, \ldots, x_L)$ denote an observed linear peptide lead of length L. We restrict the search to chemically executable local interventions on x. A stapling intervention is $p = (b, a, \Delta x)$, where b specifies the staple block, $a = (i, j)$ specifies the anchors, and Δx is the minimal set of residue substitutions required to instantiate block b. The experiments include two hydrocarbon stapling modes, S5-S5 at $i, i{+}4$ and R8-S5 at $i, i{+}7$.

Each intervention uniquely determines a *canonical minimal-edit product*,

$$\bar{z}(x, p) = G(x, p), \tag{1}$$

obtained by applying exactly the substitutions in Δx, assigning the prescribed anchors and block, and activating the stapled topology.

We denote by $\mathcal{K}_{\mathrm{hard}}(x)$ the set of products satisfying the chemistry, geometry, edit-budget, identity, and protected-position constraints. The feasible intervention set is then

$$\mathcal{H}(x) = \{p : \bar{z}(x, p) \in \mathcal{K}_{\mathrm{hard}}(x)\}\,. \tag{2}$$

A lead is hard-supported when this feasible intervention set is non-empty,

$$h_{\mathrm{hard}}(x) = \mathbb{I}[\mathcal{H}(x) \neq \varnothing]\,. \tag{3}$$

Hard support depends on the predefined chemistry, geometry, and edit constraints. Property predictors neither construct $\mathcal{H}(x)$ nor determine $h_{\mathrm{hard}}(x)$, so the functional objective ranks interventions that are already executable. Unsupported leads have no feasible intervention under the fixed hydrocarbon design space. Exact constraint definitions and numerical thresholds appear in Appendix B.

Execution is represented after the intervention space has been fixed. A peptide design state is $z = (s, a, b, m)$, where s is the current monomer sequence, a is an anchor pair, b is the committed staple block, and $m \in \{\mathrm{linear}, \mathrm{stapled}\}$ is the topology. The initial state is $z_0(x) = (x, \varnothing, \varnothing, \mathrm{linear})$. A trajectory $\tau = (z_0, \ldots, z_T)$ follows a discrete edit graph containing ncAA substitution, anchor and block assignment, topology activation, and other admissible local edits. It completes intervention p when $z_T = \bar{z}(x, p)$. This canonical contract keeps intervention selection, terminal scoring, and execution tied to the same molecular state.

3.2 SOFT-TERMINAL CONTROL ON THE FEASIBLE INTERVENTION SPACE

The reference distribution covers both the intervention and its execution. Let $\widetilde{R}(\tau \mid x, p)$ denote the plan-aware reference edit process, and let

$$\mathcal{T}(x, p) = \{\tau : z_0 = z_0(x),\ z_T = \bar{z}(x, p)\}$$

be the set of canonical-completing trajectories. We define the completion-conditioned reference

$$R^{\mathrm{can}}(\tau \mid x, p) = \widetilde{R}\big(\tau \mid x, p,\, Z_T = \bar{z}(x, p)\big), \qquad \tau \in \mathcal{T}(x, p). \tag{4}$$

In practice, samples from this conditional law are obtained by drawing from the plan-aware reference process and retaining exact canonical completions.

The full reference path measure factorizes as

$$R_x(p, \tau) = q_{\mathrm{ref}}(p \mid x) R^{\mathrm{can}}(\tau \mid x, p). \tag{5}$$

The intervention-level reference incorporates empirical preferences over hydrocarbon stapling modes while correcting for the number of available anchor placements according to

$$q_{\mathrm{ref}}(p \mid x) \propto \frac{\pi_{\mathrm{emp}}(\mu(p))^{\alpha}}{n_{\mu(p)}(x)}, \qquad p \in \mathcal{H}(x), \tag{6}$$

where $\mu(p)$ denotes the hydrocarbon-stapling mode and $n_{\mu}(x)$ is the number of feasible interventions of mode μ for lead x. Here, π_{emp} is the empirical mode prior, and α controls its strength. The conditional edit reference assigns more probability to actions consistent with the committed intervention and less probability to unnecessary off-plan changes. Details of the reference parameterization are provided in Appendix B.

Lead optimization lacks a target endpoint distribution for each input peptide. We therefore use the soft-terminal Schrödinger-control objective

$$P_x^{\star} = \arg\min_{P \ll R_x} \left\{ \mathrm{KL}(P \| R_x) + \beta\, \mathbb{E}_P \left[E_T(Z_T; x) \right] \right\}. \tag{7}$$

Here, $P \ll R_x$ restricts the candidate path law P to the support of R_x, and $\beta > 0$ controls the terminal-energy penalty. The KL term measures the information required to deviate from the chemistry-aware reference process, while lower terminal energy corresponds to more desirable stapled products. Its optimal path law is the exponential tilt

$$P_x^{\star}(p, \tau) = \frac{R_x(p, \tau) \exp[-\beta E_T(Z_T; x)]}{\mathcal{Z}_x}. \tag{8}$$

Here, $\mathcal{Z}_x$ normalizes the tilted path law. The terminal energy is defined on the complete canonical product. Its general form is

$$E_T(\bar{z}; x) = \lambda_{\mathrm{prop}} E_{\mathrm{prop}}(\bar{z}) + \lambda_{\mathrm{edit}} d_{\mathrm{edit}}(\bar{z}, z_0) + \lambda_{\mathrm{geom}} \phi_{\mathrm{geom}}(\bar{z}) + \lambda_{\mathrm{block}} c(b). \tag{9}$$

The λ coefficients weight task-specific property energy, edit distance d_{edit}, geometric penalty ϕ_{geom}, and staple-block cost $c(b)$. Here, $E_{\mathrm{prop}}(\bar{z})$ denotes a task-specific terminal property energy defined on the canonical product, with lower values corresponding to more favorable terminal states. The concrete property scorer, coefficients, and evaluation protocol used in the reported experiments are described in Section 4 and Appendix B.

3.3 Exact intervention distribution

The canonical plan-product correspondence makes the key intervention marginal of equation 7 exactly tractable. For any $\tau \in \mathcal{T}(x, p)$,

$$Z_T = \bar{z}(x, p),$$

and hence the terminal energy is constant across all valid executions of the same intervention,

$$E_T(Z_T; x) = E_T(\bar{z}(x, p); x).$$

Marginalizing equation 8 over execution trajectories therefore yields

$$q^{\star}(p \mid x) = \frac{q_{\mathrm{ref}}(p \mid x) \exp[-\beta E_T(\bar{z}(x, p); x)]}{\sum_{p' \in \mathcal{H}(x)} q_{\mathrm{ref}}(p' \mid x) \exp[-\beta E_T(\bar{z}(x, p'); x)]}. \tag{10}$$

Moreover, the complete optimal law factorizes as

$$P_x^{\star}(p, \tau) = q^{\star}(p \mid x) R^{\mathrm{can}}(\tau \mid x, p). \tag{11}$$

Equation 10 is the central computational consequence of the formulation. The finite, lead-specific intervention space and deterministic plan–product mapping make $q^{\star}(\cdot \mid x)$ exactly computable without paired linear-to-stapled training examples. This exact intervention distribution provides the plan-level supervision signal used for amortized intervention control.

3.4 AMORTIZED INTERVENTION CONTROL

Direct evaluation of $q^\star$ requires constructing and terminal-scoring every feasible canonical product. Amortization removes these exhaustive terminal-property queries at deployment. We learn an amortized controller

$$q_\theta(p \mid x) = \mathrm{softmax}_{p \in \mathcal{H}(x)} \left[\log q_{\mathrm{ref}}(p \mid x) + s_\theta(x, p) \right], \tag{12}$$

where $s_\theta(x, p)$ combines sequence context, explicit intervention descriptors, and plan-specific geometric information. The exact target $q^\star$ incorporates terminal-property evaluations during training. We distill this supervision into q_θ, so that test-time selection requires the lead and feasible-plan features, without querying the terminal-property predictor. Training distills the exact intervention distribution by minimizing

$$\mathcal{L}_{\mathrm{plan}} = \mathrm{KL}\left(q^\star(\cdot \mid x) \| q_\theta(\cdot \mid x) \right). \tag{13}$$

This converts a per-lead exact stochastic-control solution into an amortized decision rule that generalizes to unseen peptide leads. Architecture and representation details, including the pretrained sequence encoder and geometric descriptors, are deferred to Appendix B.

3.5 CONSTRAINT-PRESERVING EXECUTION AND INFERENCE

Intervention selection and execution are separated. A chemical modification is selected from $q_\theta(p \mid x)$, after which a conditional transition policy assigns the edit sequence that realizes it. The deployed process factorizes as

$$P_\Theta(p, \tau \mid x) = q_\theta(p \mid x) P_\phi(\tau \mid x, p), \qquad \Theta = (\theta, \phi). \tag{14}$$

The transition policy is fitted by terminal-energy-weighted maximum likelihood on reference trajectories that reach the canonical product exactly. Detailed weights and training losses appear in Appendix B. Under the canonical contract, all successful executions of a fixed intervention share the same terminal product and utility. The exact intervention distribution $q^\star$ provides the terminal-property supervision, and the conditional policy learns valid realization of the selected plan. This distinction follows directly from equation 11.

At inference, $\mathcal{H}(x)$ is constructed using the chemistry and geometry constraints. The distribution $q_\theta(p \mid x)$ is then evaluated over the feasible interventions, and the selected plan is

$$\hat{p} = \arg \max_{p \in \mathcal{H}(x)} q_\theta(p \mid x). \tag{15}$$

The conditional policy then decodes a trajectory terminating at $\bar{z}(x, \hat{p})$. The terminal property predictor used to construct the training target is not queried to exhaustively rerank feasible interventions at test time.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

We used PepBench (Kong et al., 2024) for model development and evaluated on the LNR benchmark (Tsaban et al., 2022). A fixed preprocessing pipeline produced 4,020 training, 111 validation, and 77 test leads of at most 25 residues. These datasets contain peptide leads without paired linear-to-hydrocarbon-stapled products. Under the property-independent hard-support criterion in equation 3, 53 test leads admit at least one valid stapling plan. We evaluated support-conditioned metrics on this fixed subset and report support coverage over the full test set. Appendix A provides complete preprocessing details.

We compared StapleBridge against four constrained baselines under the same property-independent feasible intervention space $\mathcal{H}(x)$. Under **Random Feasible**, plans are sampled uniformly from the strict feasible support. Under **Empirical Reference**, interventions are drawn from $q_{\mathrm{ref}}(p \mid x)$ without learned terminal control. **Generic-SB** is a matched flat action-space bridge with the same hydrocarbon-stapling task and surrogate objective but no explicit plan variable. **Direct Plan Ranker (DPR)** is a strong plan-level baseline with the same feasible plans, lead context, geometric features,

Table 1: **Hydrocarbon-stapling intervention performance.** All metrics are evaluated on the fixed 53-peptide subset of the 77-peptide test set that admits at least one valid stapling intervention. Property metrics use the frozen PeptiVerse surrogate. Random Feasible, Generic-SB, Direct Plan Ranker, and StapleBridge are reported as mean ± standard deviation over three independent runs; Empirical Reference is deterministic and reported once.

Method	Permeability Improved ↑	ΔPermeability ↑	Plan Regret ↓
Rand. Feasible	$69.81\% \pm 3.27\%$	0.127 ± 0.012	0.891 ± 0.160
Empirical Ref.	62.26%	0.062	1.104
Generic-SB	$68.55\% \pm 1.09\%$	0.123 ± 0.001	0.934 ± 0.026
Direct Plan Ranker	$73.58\% \pm 1.89\%$	0.140 ± 0.006	0.380 ± 0.031
StapleBridge	$\mathbf{81.13\% \pm 0.00\%}$	$\mathbf{0.169 \pm 0.004}$	$\mathbf{0.209 \pm 0.006}$

and intervention descriptors. Its training objective directly ranks terminal plans using their terminal scores. The Original Linear peptide is used as the parent reference for lead-relative property changes.

For the hydrocarbon-stapling experiments, we instantiate the task-specific terminal utility with a frozen PeptiVerse (Zhang et al., 2026b) property scorer. All methods use the same property-independent feasible support and frozen evaluation protocol. We report task-aligned property outcomes together with intervention-selection diagnostics; formal metric definitions are provided in Appendix D. Alongside property-based outcomes, we report Plan Regret as a common-objective diagnostic of intervention selection. For each lead, the intervention selected by every method is evaluated under the same fixed plan-scoring criterion, and regret measures its gap to the best feasible intervention within $\mathcal{H}(x)$ under that criterion. Lower values indicate closer alignment with the common intervention objective.

All orthogonal permeability and 3D analyses were performed post hoc on the same frozen products and did not affect training, checkpoint selection, or intervention ranking. The same fixed products were used throughout, with complete protocols in Appendices D and E.

4.2 StapleBridge improves intervention selection

Explicit plan-level modeling improved over the flat action-space baseline, as DPR reduced Plan Regret from 0.934 to 0.380 and increased mean ΔPermeability from 0.123 to 0.140 (Table 1). StapleBridge further improved both metrics, reducing Plan Regret to 0.209 and increasing mean ΔPermeability to 0.169, with 81.1% of supported test peptides showing improved predicted permeability. With the intervention space and input features matched, StapleBridge's lower Plan Regret and larger permeability gain relative to DPR support the value of distribution-level supervision beyond direct scalar plan ranking.

4.3 Plan-level supervision improves objective alignment

Each component contributed differently to intervention selection, with all variants retaining the same feasible intervention space and hard constraints (Table 2). Removing the functional tilt reduced the permeability-improvement rate from 81.1% to 69.2% and mean ΔPermeability from 0.169 to 0.120, consistent with its role in directing selection toward products with higher predicted permeability. Removing ESM2 local context produced a smaller degradation, reducing mean ΔPermeability to 0.151 and increasing Plan Regret to 0.268.

We further replaced the soft intervention target $q^\star$ with hard top-1 distillation. Hard top-1 supervision reduced mean ΔPermeability from 0.169 to 0.160 and increased Plan Regret from 0.209 to 0.247. It also weakened recovery of the broader intervention ordering, reducing Spearman agreement with $q^\star$ from 0.661 to 0.548. These results indicate that the full soft target provides supervision over the relative quality of feasible interventions beyond the identity of the highest-ranked plan.

Together, these ablations separate the roles of functional guidance, local sequence context, and distribution-level supervision in lead-specific intervention selection.

Table 2: **Ablation of intervention control.** All variants share the same feasible intervention space, hard constraints, and controller architecture. Hard Target Distillation replaces the soft target $q^\star$ with a one-hot top-1 target. Top-1 Agreement and Spearman quantify recovery of $q^\star$. Values are mean $\pm$ standard deviation over three runs.

Variant	Permeability Improved ↑	ΔPermeability ↑	Top-1 Agree. w/ $q^\star$ ↑	Spearman w/ $q^\star$ ↑	Plan Regret ↓
w/o Functional Tilt	$69.182 \pm 6.626\%$	0.120 ± 0.014	$26.415 \pm 6.536\%$	0.022 ± 0.154	0.822 ± 0.161
w/o ESM2 Local Context	$79.874 \pm 1.089\%$	0.151 ± 0.003	$49.686 \pm 1.089\%$	0.602 ± 0.008	0.268 ± 0.012
Hard Target Distillation	$80.503 \pm 1.089\%$	0.160 ± 0.005	$50.943 \pm 1.887\%$	0.548 ± 0.021	0.247 ± 0.015
StapleBridge (soft $q^\star$)	$\mathbf{81.132 \pm 0.000\%}$	$\mathbf{0.169 \pm 0.004}$	$\mathbf{51.572 \pm 1.089\%}$	$\mathbf{0.661 \pm 0.006}$	$\mathbf{0.209 \pm 0.006}$

Table 3: **Cross-predictor evaluation with CPMP.** Frozen products are evaluated using four assay-specific CPMP predictors. Values are mean $\pm$ standard deviation over three predefined seeds. Empirical Ref. is deterministic. Values are predicted $\log_{10} P$ with permeability expressed in cm/s; higher (less negative) values indicate higher predicted permeability.

Method	PAMPA ↑	Caco-2 ↑	RRCK ↑	MDCK ↑
Random Feasible	-7.047 ± 0.009	-6.702 ± 0.021	-6.192 ± 0.018	-6.047 ± 0.020
Empirical Ref.	-7.102	-6.785	-6.244	-6.103
Generic-SB	-6.986 ± 0.005	-6.643 ± 0.005	-6.159 ± 0.007	-5.996 ± 0.005
Direct Plan Ranker	-6.903 ± 0.009	-6.572 ± 0.021	-6.092 ± 0.009	-5.942 ± 0.017
StapleBridge	$\mathbf{-6.762 \pm 0.014}$	$\mathbf{-6.420 \pm 0.004}$	$\mathbf{-5.948 \pm 0.007}$	$\mathbf{-5.780 \pm 0.012}$

4.4 Cross-predictor evaluation with CPMP

We additionally evaluated the same frozen hydrocarbon-stapled products using the four assay-specific CPMP (Jiang et al., 2025b) predictors for PAMPA, Caco-2, RRCK, and MDCK. CPMP was applied after intervention selection and was not used for training, checkpoint selection, or plan ranking. StapleBridge achieved the strongest scores across the four endpoints (Table 3), providing consistent cross-predictor support that its selected interventions remain favorable under a separately developed predictive model.

4.5 Structural compatibility with the parent lead

We next evaluated whether the selected interventions remain compatible with the three-dimensional structure and receptor-bound context of the parent peptides. This is important because the starting peptide already encodes functional structural interactions that stapling should preserve as much as possible. For backbone evaluation, the frozen selected products were reconstructed from the deposited peptide structures and relaxed using the validation-frozen GFN-FF/ALPB-water protocol. Of the 53 supported peptides, 48 contain the complete N/CA/C/O backbone required for reconstruction. Both E and Z staple isomers were evaluated, yielding 96 molecule-isomer jobs per method. Receptor-interface preservation was evaluated separately on the common 31-peptide subset with complete measurements across all methods. This aggregate analysis uses the frozen reconstructed products in the deposited receptor coordinate frame, without the additional receptor-aware local refinement used in the case studies below. Backbone preservation is measured by Cα-RMSD, while receptor-interface preservation is quantified by native-contact retention (F_{nat}) and interface RMSD (iRMSD). Complete metric and reconstruction definitions are given in Appendices D and E.

StapleBridge showed the lowest mean/median backbone deviation, with a Cα-RMSD of 1.106/0.660 Å, and the highest observed relaxation completion rate of 89.6%, compared with 88.5% for DPR and 87.5% for the other baselines (Table 4A). All reconstructed molecules retained their frozen product identity, and every successful relaxation completed under ALPB water without gas-phase fallback. Under the frozen structural protocol, StapleBridge-selected products preserved the deposited parent backbone more closely than the evaluated baselines while maintaining a high relaxation completion rate. Appendix F provides representative molecule-level examples showing how differences in anchor placement and staple chemistry translate into distinct relaxed peptide geometries.

Table 4: **Post-hoc structural compatibility. A**, parent-peptide backbone preservation after frozen product reconstruction and relaxation; relaxation success is reported over 96 E/Z molecule-isomer jobs per method. **B**, receptor-interface preservation on the common $n = 31$ subset. Brackets denote bootstrap 95% confidence intervals.

A. Parent-peptide backbone preservation		
Method	Cα-RMSD mean/med. (Å) ↓	Relax. success ↑
Random Feasible	1.394 / 1.412	84/96 (87.5%)
Empirical Ref.	1.373 / 1.230	84/96 (87.5%)
Generic-SB	1.370 / 1.130	84/96 (87.5%)
DPR	1.300 / 1.150	85/96 (88.5%)
StapleBridge	**1.106 / 0.660**	**86/96 (89.6%)**
B. Receptor-interface preservation ($n = 31$)		
Method	Contact retention ↑	iRMSD (Å) ↓
Random Feasible	0.686 [0.656, 0.710]	1.416 [1.237, 1.629]
Empirical Ref.	0.682 [0.661, 0.701]	1.386 [1.227, 1.550]
Generic-SB	0.701 [0.683, 0.739]	1.215 [1.084, 1.369]
DPR	0.730 [0.704, 0.753]	1.280 [1.153, 1.414]
StapleBridge	**0.795 [0.776, 0.823]**	**0.883 [0.741, 1.050]**

We next examine whether this structural compatibility extends to the receptor-bound context. On the common $n = 31$ subset, StapleBridge achieved the highest native-contact retention ($F_{\text{nat}} = 0.795$) and the lowest iRMSD (0.883 Å), compared with 0.730 and 1.280 Å for DPR. Under the common receptor-frame evaluation, StapleBridge preserved more native receptor contacts and produced less interface displacement than the matched baselines. Together with the backbone analysis, these results support structural preservation across both parent-peptide and receptor-interface views.

4.6 Molecular case studies under receptor-aware refinement

We further examined whether selected interventions remain compatible with the receptor-bound geometry after an additional receptor-aware local refinement. These case studies included the receptor during local Vina refinement and provide molecule-level illustrations (Figure 2).

For 6seo, StapleBridge retained more native contacts and a lower iRMSD than the three baselines after refinement ($F_{\text{nat}} = 0.677$, iRMSD$= 1.057$ Å), while increasing predicted permeability by $+0.365$. For 5onp, the selected intervention also remained close to the deposited receptor-bound geometry ($F_{\text{nat}} = 0.908$, iRMSD$= 1.000$ Å). Together, these examples illustrate that the selected stapling interventions can remain structurally compatible after an additional receptor-aware refinement step. Additional molecule-level comparisons of selected interventions and relaxed structures appear in Appendix G.

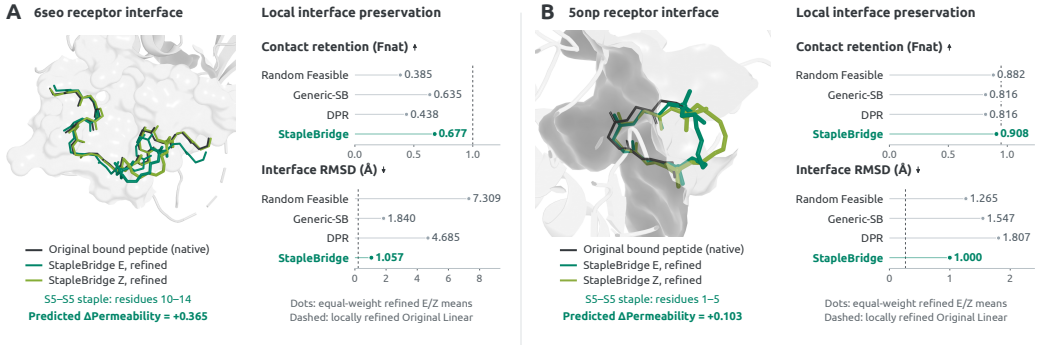

Figure 2: **Receptor-aware local-refinement case studies. A**, 6seo; **B**, 5onp. For each case, refined StapleBridge E/Z products are overlaid with the deposited receptor-bound peptide, alongside native-contact retention (F_{nat}) and iRMSD after local Vina refinement for all methods. Points denote equal-weight refined E/Z means; the dashed vertical reference line in each metric panel indicates the locally refined Original Linear value.

Table 5: **Portability beyond hydrocarbon stapling.** Results on disulfide cyclization and N-methylation using chemistry-specific intervention spaces. Values are mean ± standard deviation over three seeds. No Empirical Ref. baseline is reported for disulfide cyclization because no chemistry-specific empirical mode prior is defined.

Method	Permeability Improved (%) ↑	ΔPermeability ↑	Plan Regret ↓
Disulfide cyclization			
Rand. Feasible	56.02 ± 1.60	0.039 ± 0.003	0.942 ± 0.012
Generic-SB	71.76 ± 1.60	0.110 ± 0.028	0.578 ± 0.142
DPR	73.61 ± 1.39	0.143 ± 0.004	0.392 ± 0.025
StapleBridge	$\mathbf{80.56 \pm 0.00}$	$\mathbf{0.171 \pm 0.004}$	$\mathbf{0.274 \pm 0.023}$
N-methylation			
Rand. Feasible	59.72 ± 2.41	0.016 ± 0.007	0.340 ± 0.037
Empirical Ref.	61.11	0.018	0.330
Generic-SB	58.33 ± 7.22	0.019 ± 0.006	0.326 ± 0.028
DPR	61.11 ± 1.39	0.019 ± 0.003	0.336 ± 0.014
StapleBridge	$\mathbf{65.28 \pm 2.78}$	$\mathbf{0.026 \pm 0.003}$	$\mathbf{0.306 \pm 0.013}$

4.7 Portability beyond hydrocarbon stapling

Finally, we tested whether the structured intervention-control formulation can be instantiated for peptide chemistries beyond hydrocarbon stapling. The two settings are N-methylation, a local backbone chemical edit that does not alter peptide topology, and disulfide cyclization, a topology-changing covalent constraint. For each chemistry, we retained the same plan-based control formulation while replacing only the chemistry-specific intervention catalog and hard-feasibility rules. On the fixed 72-peptide evaluation sets, products selected with StapleBridge had the best predicted permeability improvement and the lowest Plan Regret for both chemistries (Table 5). These experiments used chemistry-specific models and intervention spaces to assess the portability of the control formulation. The results support extending plan-level control to chemically distinct peptide-editing problems.

5 Discussion and Conclusion

In this work, we introduced StapleBridge for optimizing existing peptide leads through chemically feasible local modifications. Grounded in KL-regularized stochastic control and soft-terminal Schrödinger formulations (Todorov, 2006; Chen et al., 2021; Garg et al., 2024), StapleBridge exploits the finite, lead-specific intervention space and canonical plan–product mapping to obtain an exactly computable intervention distribution under the specified reference and terminal objective. Distilling this exact target into an amortized controller enables intervention selection on unseen leads without exhaustive terminal-property reranking. Improved intervention selection, together with cross-predictor permeability evaluation and frozen structural analyses, supports the use of the formulation for constrained peptide refinement.

Experimental stapling studies have established that staple placement and chemistry influence cellular activity and target interactions (Chandramohan et al., 2024; Reyes et al., 2025). Our results provide a computational basis for prioritizing these modifications, with permeability improvements consistently supported by PeptiVerse and the separately trained CPMP predictors(Zhang et al., 2026b; Jiang et al., 2025b). Prospective paired measurements of parent and modified peptides will be required to determine whether these computational gains translate into improved measured permeability with retained target binding. Coverage also depends on the intervention catalog, which supported 53 of the 77 test leads. The N-methylation and disulfide-cyclization results support extending the same control formulation through chemistry-specific catalogs and feasibility rules. Overall, StapleBridge enables property-guided refinement of existing peptide leads by learning feasible chemical interventions from exact control targets without paired parent–product supervision.

AI USE STATEMENT

We used generative AI tools to refine the writing and assist with manuscript review. The authors take responsibility for the final manuscript.

REPRODUCIBILITY STATEMENT

Section 3 defines the feasible interventions, exact control target, and training objectives. Appendices A–D document data preprocessing, splits, implementation settings, baselines, and evaluation metrics. Appendix E describes the frozen structural protocol, and Appendix H specifies repeated-run aggregation and checkpoint selection. Code is available at `https://anonymous.4open.science/r/Staplebridge-A05A/`.

ETHICS STATEMENT

This work is entirely computational and does not involve human participants, identifiable personal data, animal experiments, or prospective biological experiments. All evaluations are performed using previously available peptide–protein structural benchmarks and computational property predictors. StapleBridge is intended to support peptide lead optimization by prioritizing chemically feasible interventions for downstream validation; the predicted permeability and structural outcomes reported in this work should therefore not be interpreted as experimentally established therapeutic efficacy or safety. As with other computational molecular-design methods, any candidates proposed by the framework require appropriate experimental validation and domain-specific safety assessment before practical use.

REFERENCES

Gregory H Bird, Emanuele Mazzola, Kwadwo Opoku-Nsiah, Margaret A Lammert, Marina Godes, Donna S Neuberg, and Loren D Walensky. Biophysical determinants for cellular uptake of hydrocarbon-stapled peptide helices. *Nature chemical biology*, 12(10):845–852, 2016.

Arun Chandramohan, Hubert Josien, Tsz Ying Yuen, Ruchia Duggal, Diana Spiegelberg, Lin Yan, Yu-Chi Angela Juang, Lan Ge, Pietro G Aronica, Hung Yi Kristal Kaan, et al. Design-rules for stapled peptides with in vivo activity and their application to mdm2/x antagonists. *Nature communications*, 15(1):489, 2024.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Optimal transport in systems and control. *Annual Review of Control, Robotics, and Autonomous Systems*, 4(1):89–113, 2021.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. *Advances in neural information processing systems*, 34:17695–17709, 2021.

Jhanvi Garg, Xianyang Zhang, and Quan Zhou. Soft-constrained schrödinger bridge: a stochastic control approach. In *International Conference on Artificial Intelligence and Statistics*, pp. 4429–4437. PMLR, 2024.

Shrey Goel and Pranam Chatterjee. Minimal-action discrete schrödinger bridge matching for peptide sequence design. In *ICLR 2026 2nd Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy*, 2026. URL `https://openreview.net/forum?id=94KB26ww3D`.

Nate Gruver, Samuel Stanton, Nathan Frey, Tim GJ Rudner, Isidro Hotzel, Julien Lafrance-Vanasse, Arvind Rajpal, Kyunghyun Cho, and Andrew G Wilson. Protein design with guided discrete diffusion. *Advances in neural information processing systems*, 36:12489–12517, 2023.

Nikita Gushchin, Sergei Kholkin, Evgeny Burnaev, and Alexander Korotin. Light and optimal Schrödinger bridge matching. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 17100–17122. PMLR, 2024. URL `https://proceedings.mlr.press/v235/gushchin24a.html`.

Nikita Gushchin, David Li, Daniil Selikhanovych, Evgeny Burnaev, Dmitry Baranchuk, and Alexander Korotin. Inverse bridge matching distillation. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 21471–21496. PMLR, 2025. URL https://proceedings.mlr.press/v267/gushchin25b.html.

Yinbin Han, Meisam Razaviyayn, and Renyuan Xu. Stochastic control for fine-tuning diffusion models: Optimality, regularity, and convergence. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 21844–21870. PMLR, 2025. URL https://proceedings.mlr.press/v267/han25g.html.

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 2025a. URL https://proceedings.mlr.press/v267/jiang25d.html.

Dawei Jiang, Zixi Chen, and Hongli Du. Cyclic peptide membrane permeability prediction using deep learning model based on molecular attention transformer. *Frontiers in Bioinformatics*, 5: 1566174, 2025b.

Jun Hyeong Kim, Seonghwan Kim, Seokhyun Moon, Hyeongwoo Kim, Jeheon Woo, and Woo Youn Kim. Discrete diffusion schrödinger bridge matching for graph transformation. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 13831–13877, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/2438d634f0ed1640934d31376c110a92-Paper-Conference.pdf.

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

Grigoriy Ksenofontov and Alexander Korotin. Categorical Schrödinger bridge matching. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 31727–31751. PMLR, 2025. URL https://proceedings.mlr.press/v267/ksenofontov25a.html.

Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *arXiv: Probability*, 2013. URL https://api.semanticscholar.org/CorpusID:14387241.

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. *arXiv preprint arXiv:2406.00735*, 2024.

Guan-Horng Liu, Yaron Lipman, Maximilian Nickel, Brian Karrer, Evangelos Theodorou, and Ricky TQ Chen. Generalized schrödinger bridge matching. In *International Conference on Learning Representations*, volume 2024, pp. 14527–14552, 2024.

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, et al. Cyclic peptide structure prediction and design using alphafold2. *Nature Communications*, 16(1):4730, 2025.

Lauren Reyes, Lena Naser, Warren S Weiner, Darren Thifault, Erik Stahl, Liam McCreary, Rohini Nott, Colton Quick, Alex Buchberger, Carlos Alvarado, et al. Structure-based discovery of hydrocarbon-stapled paxillin peptides that block fak scaffolding in cancer. *Nature communications*, 16(1):2060, 2025.

Christian E Schafmeister, Julia Po, and Gregory L Verdine. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *Journal of the American Chemical Society*, 122(24):5891–5892, 2000.

Erwin Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. In *Annales de l'institut Henri Poincaré*, volume 2, pp. 269–310, 1932.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion schrödinger bridge matching. *Advances in neural information processing systems*, 36:62183–62223, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/c428adf74782c2092d254329b6b02482-Abstract-Conference.html.

Hannes Stark, Bowen Jing, Chenyu Wang, Gabriele Corso, Bonnie Berger, Regina Barzilay, and Tommi Jaakkola. Dirichlet flow matching with applications to DNA sequence design. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 46495–46513, 2024. URL https://proceedings.mlr.press/v235/stark24b.html.

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, 2026.

Yaw Sing Tan, David P Lane, and Chandra S Verma. Stapled peptide design: principles and roles of computation. *Drug Discovery Today*, 21(10):1642–1653, 2016.

Sophia Tang, Yinuo Zhang, and Pranam Chatterjee. Entangled Schrödinger bridge matching. *arXiv preprint arXiv:2511.07406*, 2025a. URL https://arxiv.org/abs/2511.07406.

Sophia Tang, Yinuo Zhang, and Pranam Chatterjee. PepTune: De novo generation of therapeutic peptides with multi-objective-guided discrete diffusion. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 59017–59065, 2025b. URL https://arxiv.org/abs/2412.17780.

Sophia Tang, Yinuo Zhang, Alexander Tong, and Pranam Chatterjee. Branched Schrödinger bridge matching. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=ctq8BfUXWz.

Emanuel Todorov. Linearly-solvable markov decision problems. *Advances in neural information processing systems*, 19, 2006.

Alexander Y. Tong, Nikolay Malkin, Kilian Fatras, Lazar Atanackovic, Yanlei Zhang, Guillaume Huguet, Guy Wolf, and Yoshua Bengio. Simulation-free Schrödinger bridges via score and flow matching. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238, pp. 1279–1287. PMLR, 2024. URL https://proceedings.mlr.press/v238/tong24a.html.

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, 2022.

Gregory L Verdine and Gerard J Hilinski. Stapled peptides for intracellular drug targets. In *Methods in enzymology*, volume 503, pp. 3–33. Elsevier, 2012.

Alexander A Vinogradov, Yizhen Yin, and Hiroaki Suga. Macrocyclic peptides as drug candidates: recent progress and remaining challenges. *Journal of the American Chemical Society*, 141(10): 4167–4181, 2019.

Loren D Walensky and Gregory H Bird. Hydrocarbon-stapled peptides: principles, practice, and progress: miniperspective. *Journal of medicinal chemistry*, 57(15):6275, 2014.

Yitian Wang, Fanmeng Wang, Angxiao Yue, Wentao Guo, Yaning Cui, and Hongteng Xu. Muco: Generative peptide cyclization empowered by multi-stage conformation optimization. *arXiv preprint arXiv:2602.11189*, 2026.

Zhe Wang, Jianping Wu, Mengjun Zheng, Chenchen Geng, Borui Zhen, Wei Zhang, Hui Wu, Zhengyang Xu, Gang Xu, Si Chen, et al. Stapep: an open-source toolkit for structure prediction, feature extraction, and rational design of hydrocarbon-stapled peptides. *Journal of Chemical Information and Modeling*, 64(24):9361–9373, 2024.

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. *Advances in Neural Information Processing Systems*, 38, 2026.

Jingjie Zhang, Hanqun Cao, Haosen Shi, He Mutian, Yu Wang, Zijun Gao, Fang Wu, Xiaojun Yao, Chang-Yu Hsieh, Sinno Jialin Pan, et al. Geocycler: Reward-aligned 3d diffusion for constraint-conditioned cyclic peptide design. *arXiv preprint arXiv:2605.23407*, 2026a.

Yinuo Zhang, Sophia Tang, Tong Chen, Elizabeth Mahood, Sophia Vincoff, and Pranam Chatterjee. PeptiVerse: A unified platform for therapeutic peptide property prediction. *Nature Communications*, 17:6819, 2026b. doi: 10.1038/s41467-026-74167-w. URL `https://www.nature.com/articles/s41467-026-74167-w`.

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*, pp. 13020–13031, 2026.

Alessandro Zorzi, Kaycie Deyle, and Christian Heinis. Cyclic peptide therapeutics: past, present and future. *Current opinion in chemical biology*, 38:24–29, 2017.

APPENDIX

A DATASET CONSTRUCTION AND PREPROCESSING

Data provenance. The primary development dataset, **PepBench**, is constructed from all dimers extracted from the Protein Data Bank (PDB). We filter for complexes with receptors longer than 30 residues and ligands between 4 and 25 residues. To ensure non-redundancy, we remove duplicated complexes where both the receptor and peptide share over 90% sequence identity, resulting in 6,105 non-redundant complexes.

For held-out testing, we use the **LNR benchmark**, comprising 93 expert-curated protein–peptide complexes (Tsaban et al., 2022). To minimize target leakage, receptor sequences are clustered at 40% sequence identity, and development examples sharing a cluster with an LNR receptor are removed. This leakage-control step removes 1,834 of the 6,105 non-redundant complexes, leaving 4,271 development complexes, which are split into 4,157 training and 114 validation complexes. The resulting PepBench split contains 4,157 training and 114 validation complexes, while LNR is reserved exclusively for testing.

PDB-to-lead preprocessing. We apply a fixed preprocessing pipeline to all splits. For each complex, we extract the specified receptor and peptide chains from the corresponding PDB structure, convert residues to their one-letter sequence representation, and retain one Cα coordinate per residue. Examples are discarded if either the receptor or peptide chain cannot be parsed into a valid sequence or if the peptide falls outside the preprocessing length range of 5-25 residues. This produces 4,020 training leads, 111 validation leads, and 78 LNR test leads from the original 4,157/114/93 entries. Among the 78 processed LNR examples, one peptide exceeds this limit, yielding the final 77-lead test benchmark used in all reported experiments. Table 6 summarizes the resulting splits.

Table 6: **Dataset construction.** The train and validation sets are derived from the PDB-based benchmark splits, whereas the held-out test set is taken from LNR. The final hydrocarbon test benchmark additionally restricts peptide length to at most 25 residues.

Split	Upstream entries	After PDB preprocessing	Final benchmark
Train	4,157	4,020	4,020
Validation	114	111	111
LNR Test	93	78	77

No outcome-dependent filtering. The preprocessing procedure is independent of StapleBridge predictions and optimization outcomes. We do not filter or reweight examples according to PeptiVerse properties, hard-support status, generated candidates, or model success. The processed train and validation files are direct outputs of the PDB-to-lead extraction pipeline, with no additional normalization, data augmentation, or feature-engineering stage. Thus, the reduction from the upstream splits reflects only fixed structural parsing and sequence-length criteria.

Structural information. For every retained example, the preprocessing pipeline stores both the peptide and receptor sequences together with their Cα coordinates from the input PDB complex. These coordinates are used for predefined geometry-aware feasibility checks and plan-specific geometric descriptors; they do not provide paired stapled-product supervision. The benchmark itself contains linear peptide leads and does not provide experimentally paired hydrocarbon-stapled products.

Hard-support coverage. Hard support is determined independently of the property predictors using the criterion in equation 3. Of the 77 held-out LNR leads, 53 admit at least one valid hydrocarbon-stapling plan and 24 do not. We fix this 53-lead set before property evaluation and use it consistently for all support-conditioned comparisons. We retained unsupported leads when reporting test-set coverage.

B Implementation Details

The primary configuration uses 4,020 training leads, 111 validation leads, 10 epochs, horizon 8, at most 128 neighboring actions, chunk size 32, and four committed plans per lead for execution-policy training. Main results for StapleBridge and stochastic/learned baselines are reported over three independent runs using predefined random seeds. The repeated runs use seeds 42, 101, and 202. Seed 42 is used as the predefined representative run for post-hoc structural analysis and case visualization. The control coefficient is $\beta = 1$ and the plan-loss weight is 1. The plan scorer uses hidden dimension 128. The ESM2-650M reference is frozen and evaluated offline with a persistent cache. PeptiVerse is used in strict SMILES mode with fallback disabled and predictions cached. The hydrocarbon property scorer constructs neutral canonical E/Z product SMILES and averages their permeability-penetrance predictions.

The reported catalog enables S5-S5/$i, i{+}4$ and R8-S5/$i, i{+}7$. The Cα windows stored with the enabled blocks are 4.5-7.5 Å and 8.5-13.0 Å. The hard-support predicate uses geometry tolerance 2.5, edit budget 6, minimum sequence identity 0.6, and disallows protected-position edits. The same property-free predicate is applied in training, validation, and test evaluation. The terminal-energy coefficients are $\lambda_{\text{prop}} = 5.0$, $\lambda_{\text{edit}} = 0.2$, $\lambda_{\text{geom}} = 1.0$, and $\lambda_{\text{block}} = 0.2$, with $\beta = 1$.

Execution-policy objective. For the sampled batch $\mathcal{B}_x$ of canonical-completing reference trajectories for lead x, normalized terminal-energy weights are

$$w_x(\tau) = \frac{\exp[-\beta E_T(Z_T; x)]}{\sum_{\tau' \in \mathcal{B}_x} \exp[-\beta E_T(Z'_T; x)]}. \tag{16}$$

The weighted likelihood objective is

$$\mathcal{L}_{\text{exec}} = -\sum_{\tau} w_x(\tau) \sum_{t} \log P_\phi(z_{t+1} \mid z_t, x, p, t), \tag{17}$$

with non-completing and non-canonical trajectories receiving no positive weight. The full learning objective is

$$\mathcal{L} = \mathcal{L}_{\text{exec}} + \lambda_{\text{plan}} \mathcal{L}_{\text{plan}}. \tag{18}$$

C Baseline Definitions

All baselines operate on the same hydrocarbon-stapling task, use the same property-independent hard-feasible intervention support, and are evaluated with the same frozen PeptiVerse and structural protocols as StapleBridge. They differ in how an intervention is selected.

Random Feasible. For each supported lead, an intervention is sampled uniformly from the common strict feasible set $\mathcal{H}(x)$. This baseline controls for gains that can arise solely from restricting generation to chemically and geometrically valid hydrocarbon staples, without empirical staple preferences or learned lead-specific control. Results are averaged over three predefined independent random seeds.

Empirical Reference. Under Empirical Reference, interventions are drawn directly from the intervention-level reference distribution $q_{\mathrm{ref}}(p \mid x)$ without learned terminal control. This comparison tests whether population-level staple preferences encoded by the reference process are sufficient for lead optimization. Under the frozen evaluation protocol this baseline is deterministic and is therefore evaluated once.

Generic-SB. Generic-SB is the primary matched learned baseline. It shares the peptide leads, hydrocarbon chemistry, hard-feasibility constraints, surrogate property objective, and evaluation pipeline with StapleBridge, while replacing the explicit plan-level representation and exact intervention target $q^\star(p \mid x)$ with a flat action-space bridge. This comparison isolates the contribution of structured plan-level supervision from generic bridge-based optimization. Results are reported as mean $\pm$ standard deviation over the same three predefined seeds used for StapleBridge.

Direct Plan Ranker. Direct Plan Ranker (DPR) is a matched plan-level learned baseline designed to separate the benefit of explicit plan representation from exact plan-level supervision. For each lead, the same property-independent feasible intervention set $\mathcal{H}(x)$, lead representation, plan descriptors, and geometry features are used as in the StapleBridge controller. Training fits a scalar score from the terminal objective and ranks feasible plans by that score. Inference returns the highest-ranked feasible intervention without access to $q^\star$ or test-time property reranking. This baseline therefore controls for explicit plan-level modeling and tests the additional value of exact plan-level supervision. Results are reported as mean $\pm$ standard deviation over the same three predefined seeds used for StapleBridge.

Original Linear. The unmodified linear peptide is the parent reference for computing lead-relative property changes such as ΔPermeability.

D Evaluation Metrics

Evaluation covers three complementary perspectives, namely property improvement relative to the input lead, intervention-selection quality, and post-hoc structural compatibility. Unless otherwise stated, all support-conditioned metrics are computed on the fixed 53-lead hard-supported subset of the 77-lead test set.

Predicted property improvement. For each lead x, let $\hat{p}$ denote the intervention selected by a method and $\bar{z}(x, \hat{p})$ its corresponding canonical stapled product. Let $f_{\mathrm{perm}}(\cdot)$ denote the PeptiVerse-predicted permeability. For stapled products, predictions are averaged over the corresponding E/Z hydrocarbon products using the same frozen evaluation protocol for all methods.

We define the lead-relative permeability change as

$$\Delta\mathrm{Perm}(x) = f_{\mathrm{perm}}(\bar{z}(x, \hat{p})) - f_{\mathrm{perm}}(x). \tag{19}$$

The reported ΔPermeability is the mean of these lead-relative changes over the fixed hard-supported test set.

Permeability Improved measures the fraction of supported leads for which the selected stapled product has higher predicted permeability than its linear parent,

$$\mathrm{PermImproved} = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} \mathbb{I}\left[\Delta\mathrm{Perm}(x) > 0\right], \tag{20}$$

where $\mathcal{S}$ denotes the fixed 53-lead hard-supported subset. These quantities report surrogate-model predictions.

Plan Regret. We additionally evaluate each selected intervention using a fixed common criterion that combines terminal product utility with the predefined intervention reference. For each peptide, Plan Regret is the difference between the score of the selected intervention and the best score attainable within the same feasible intervention set. We report the mean across supported leads, with lower values indicating closer alignment with this common intervention objective.

Recovery of the exact intervention distribution. For analyses of the amortized controller, we additionally compare $q_\theta(p \mid x)$ with the exact intervention distribution $q^\star(p \mid x)$. **Top-1 Agreement** is the fraction of supported leads for which the highest-probability intervention under q_θ matches that under $q^\star$,

$$\mathrm{Top1} = \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} \mathbb{I}\left[\arg\max_p q_\theta(p \mid x) = \arg\max_p q^\star(p \mid x) \right]. \tag{21}$$

Spearman correlation measures agreement between the complete rankings of feasible interventions induced by q_θ and $q^\star$ for each lead, and is averaged across supported leads. Top-1 Agreement therefore measures recovery of the best intervention, whereas Spearman correlation captures agreement over the broader feasible-plan ordering.

Structural compatibility. Structural validation is performed after intervention selection and does not affect training, checkpoint selection, or test-time plan ranking. Each frozen hydrocarbon-stapled product is explicitly reconstructed and relaxed using the predefined 3D protocol. We use lead-backbone Cα-RMSD as the primary Structural compatibility metric,

$$\text{CA-RMSD} = \sqrt{\frac{1}{N_{\mathrm{CA}}} \sum_{i=1}^{N_{\mathrm{CA}}} \left\| \mathbf{r}_i^{\mathrm{stapled}} - \mathbf{r}_i^{\mathrm{lead}} \right\|_2^2}, \tag{22}$$

where N_{CA} is the number of aligned Cα atoms and $\mathbf{r}_i^{\mathrm{stapled}}$ and $\mathbf{r}_i^{\mathrm{lead}}$ are their coordinates after backbone alignment. Lower values indicate that the relaxed stapled product remains closer to the deposited backbone geometry of the input lead.

We additionally report the change in helicity relative to the corresponding linear lead and the fraction of reconstructed molecule-isomers that successfully complete the frozen relaxation protocol. Relaxation success is a 3D processing and geometry-sanity measure.

Receptor-interface preservation. We evaluate whether the selected stapling intervention preserves the receptor-bound geometry of the parent peptide using two complementary metrics. Native-contact retention (F_{nat}) is the fraction of receptor–peptide contacts present in the deposited complex that remain present after reconstruction of the frozen stapled product in the deposited receptor coordinate frame. Interface RMSD (iRMSD) measures the RMSD of the peptide–receptor interface relative to the deposited complex after receptor alignment. Higher F_{nat} indicates greater preservation of native intermolecular contacts, whereas lower iRMSD indicates smaller displacement of the binding interface. This aggregate evaluation does not include the additional receptor-aware local refinement used in the molecular case studies. For hydrocarbon staples, E and Z stereoisomers are evaluated independently and combined with equal weight.

Repeated-run aggregation. Random Feasible, Generic-SB, Direct Plan Ranker, and StapleBridge are evaluated over the same three predefined independent seeds where applicable. For each run, metrics are first computed over the fixed supported lead set, after which we report the mean and standard deviation across runs. Empirical Reference is deterministic under the frozen evaluation protocol and is therefore evaluated once.

E Orthogonal 3D Validation of Frozen StapleBridge Designs

Although StapleBridge is trained entirely in the discrete intervention space, each predicted stapling plan corresponds to a concrete hydrocarbon-stapled molecule. We therefore performed a post-hoc 3D analysis to examine whether the frozen designs remain structurally compatible with the input linear leads. All relaxation-derived structures, descriptors, and scores were computed after training, checkpoint selection, and intervention selection. The analysis evaluates the frozen selected products.

Table 7: **Post-hoc structural compatibility of frozen products.** Lead Cα-RMSD measures deviation from the deposited linear-lead structure. ΔHelicity is computed relative to the corresponding deposited lead. Relaxation success denotes molecule-isomers that successfully completed the frozen ALPB-water relaxation protocol with numerical convergence. It measures structural-processing success.

Method	Cα-RMSD mean / median ↓	ΔHelicity ↑	Relaxation success ↑
Random Feasible	1.394 / 1.412	0.028	84/96 (87.5%)
Empirical Ref.	1.373 / 1.230	0.041	84/96 (87.5%)
Generic-SB	1.370 / 1.130	0.019	84/96 (87.5%)
Direct Plan Ranker	1.300 / 1.150	0.067	85/96 (88.5%)
StapleBridge	**1.106 / 0.660**	**0.112**	**86/96 (89.6%)**

Frozen structural protocol. For each structurally eligible hard-supported test lead, we started from the deposited linear-peptide structure and reconstructed the frozen S5/R8 hydrocarbon-stapled product. Of the 53 hard-supported test leads, 48 contained the complete deposited N/CA/C/O backbone required by the frozen structural protocol. The remaining five were fixed input-level abstentions. Newly introduced side-chain and linker coordinates were initialized locally while retaining the input peptide backbone. Structures were subsequently relaxed using the protocol selected exclusively on the validation split. The protocol used GFN-FF with ALPB water, tight optimization, and restrained peptide-backbone atoms. The E and Z staple isomers were reconstructed and analyzed independently and aggregated with equal weight. The same frozen protocol was applied to Random Feasible, Empirical Reference Only, Generic-SB, Direct Plan Ranker, and StapleBridge. No test-dependent structural reranking, product reselection, global refolding, or post-hoc protocol adjustment was performed.

Lead-structure preservation. To assess preservation of the input lead, we use the Cα RMSD between the relaxed stapled product and the deposited linear-lead structure as the primary structural readout. Lower values indicate that the stapling intervention perturbs the original lead backbone less strongly. Products selected with StapleBridge had the lowest mean and median backbone deviation under the frozen structural protocol among all evaluated stapling strategies (Table 7). Their mean/median Cα-RMSD is 1.106/0.660 Å, compared with 1.300/1.150 Å for Direct Plan Ranker (DPR) and 1.370/1.130 Å for Generic-SB. The mean Cα-RMSD under Random Feasible and Empirical Reference are 1.394 and 1.373 Å, respectively. This trend is notable because backbone preservation is not included in the optimization objective.

Products selected with StapleBridge also showed the largest mean increase in helical character after relaxation. All evaluated stapling strategies exhibited positive mean ΔHelicity relative to their corresponding deposited leads, with StapleBridge achieving a ΔHelicity of $+0.112$, compared with $+0.067$ for Direct Plan Ranker, $+0.041$ for Empirical Reference, $+0.028$ for Random Feasible, and $+0.019$ for Generic-SB. Together with the lower backbone deviation, these results indicate that the interventions selected by StapleBridge remain compatible with the parent-peptide conformation while showing a stronger tendency to preserve or increase helical character under the frozen relaxation protocol.

3D reconstruction and geometry sanity. The explicit 3D reconstruction adds a chemistry-level sanity check on the frozen designs. Across the five stapling strategies, Random Feasible, Empirical Reference, Generic-SB, Direct Plan Ranker, and StapleBridge, all 480 reconstructed molecule-isomers retained canonical molecular identity with their frozen product SMILES. Every successfully completed structural optimization was explicitly verified under ALPB-water and reached the requested convergence criterion, with no gas-phase fallback. We observed no systematic violations of macrocycle size or S5/R8 stereochemical configuration. A small number of molecule-isomers did not satisfy the predefined geometry-QC criteria and were included in the reported protocol-level statistics. The products selected with StapleBridge have the highest numerical solvent-aware relaxation completion rate at 89.6%, compared with 88.5% for DPR and 87.5% for Random Feasible, Empirical Reference, and Generic-SB.

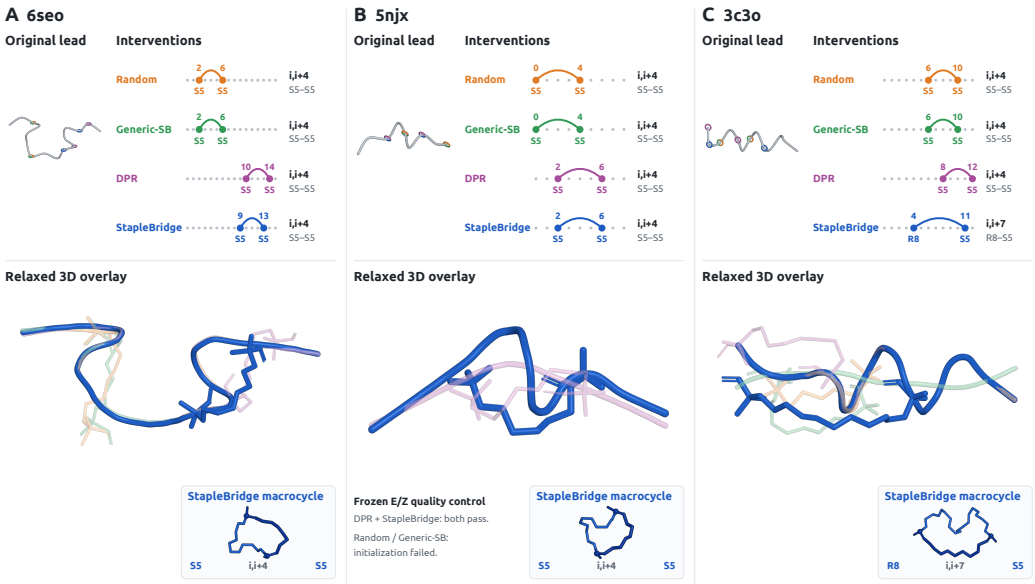

Figure 3: **Frozen case studies of lead-specific staple interventions.** Three test-set examples compare the frozen interventions and solvent-relaxed 3D geometries of Random Feasible, Generic-SB, Direct Plan Ranker (DPR), and StapleBridge. Insets show the corresponding StapleBridge macrocycles.

Summary. Taken together, these results provide an orthogonal structural validation of the permeability-oriented designs. Products selected with StapleBridge combine the strongest functional improvement with the lowest mean and median deviation from the deposited lead backbone, the largest average increase in helical character, and a high rate of successful solvent-aware 3D relaxation. Because these structural quantities were computed after product selection, the results indicate that the learned stapling interventions remain compatible with the structural identity and explicit three-dimensional realization of the input peptide leads.

F Case studies of lead-specific staple selection

The selected interventions illustrate lead-specific changes in anchor placement and staple chemistry (Figure 3). The three examples cover improved anchor selection under matched chemistry, convergence with a strong plan-level baseline when the preferred intervention is clear, and adaptation of staple chemistry and spacing to the local lead context. All displayed candidates were frozen before visualization, and the relaxed 3D structures were computed for post-hoc analysis. Anchor positions in this appendix are reported using zero-based indexing to match the implementation; residue numbers in the main-text figures use one-based indexing.

For 6seo, all four selections used S5–S5/$i, i+4$ chemistry but differ in anchor placement. The selected positions were 2:6 under Random Feasible and Generic-SB, 10:14 under DPR, and 9:13 under StapleBridge. Predicted permeability under DPR was 0.730, with $\Delta\mathrm{Perm.} = +0.275$ and Plan Regret 0.868. The corresponding StapleBridge values were 0.820, $+0.365$, and 0.417. Thus, even under matched staple chemistry, exact plan-level supervision yields a more favorable lead-specific intervention.

The 5njx case showed agreement between DPR and StapleBridge. The selected 0:4 S5–S5/$i, i+4$ plan under Random Feasible and Generic-SB failed E/Z structural initialization, whereas the shared 2:6 intervention under DPR and StapleBridge passed quality control. This shared intervention had permeability 0.639, $\Delta\mathrm{Perm.} = +0.310$, and zero Plan Regret, illustrating that DPR and StapleBridge can converge to the same intervention when both rank the same feasible plan highest.

The 3c3o case illustrated adaptation in both anchor placement and staple chemistry. The selection under Random Feasible and Generic-SB was S5–S5/$i, i+4$ at positions 6:10. Under DPR, the

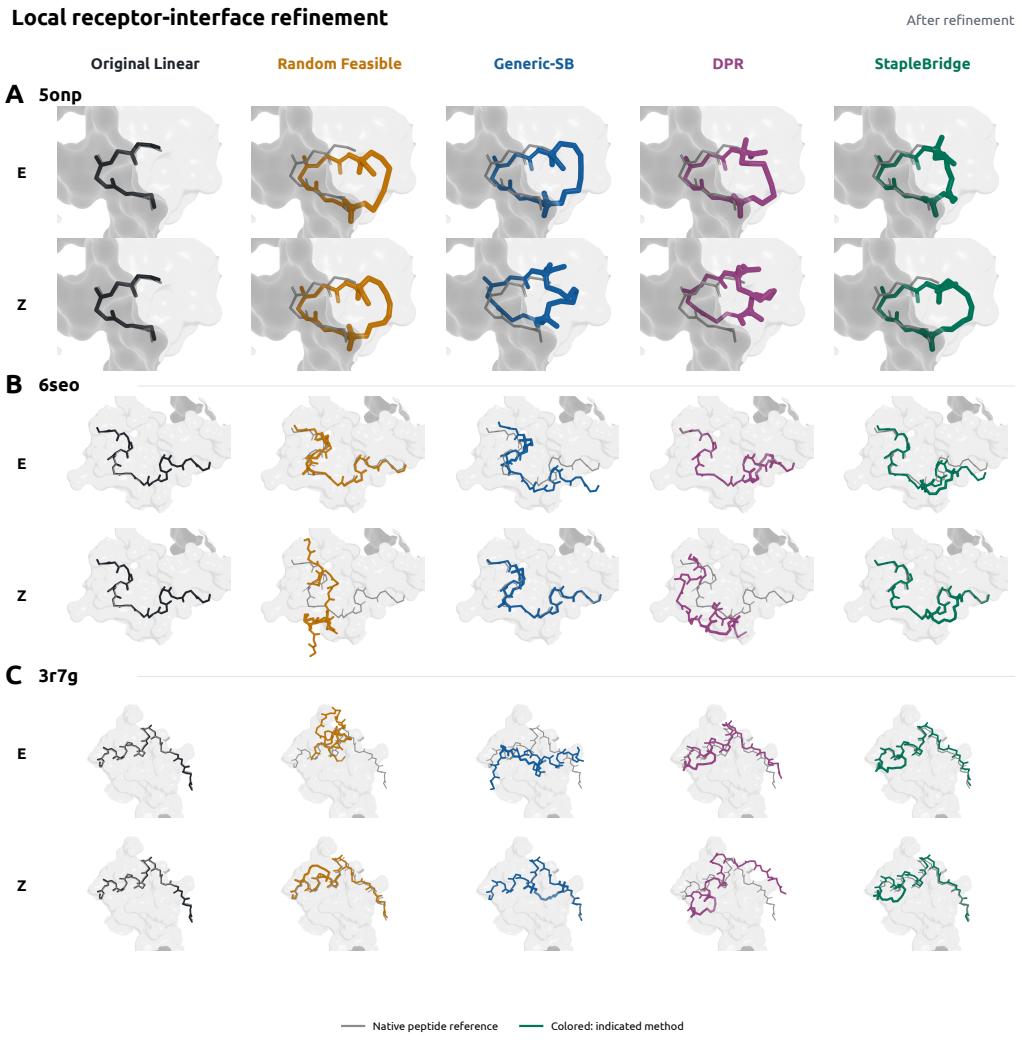

Figure 4: **Expanded receptor-interface refinement examples.** Three representative complexes are shown after an additional receptor-aware local refinement step: **A**, 5onp; **B**, 6seo; and **C**, 3r7g. Columns compare the deposited Original Linear peptide with products selected by Random Feasible, Generic-SB, Direct Plan Ranker (DPR), and StapleBridge. E and Z hydrocarbon-staple stereoisomers are shown separately. The native bound peptide is shown as the reference, while colored structures denote the corresponding method-specific refined products.

same chemistry shifted to positions 8:12, with $\Delta\mathrm{Perm.} = +0.450$ and Plan Regret 0.783. Under StapleBridge, an R8–S5/$i, i+7$ intervention at 4:11 gave permeability 0.782, $\Delta\mathrm{Perm.} = +0.583$, and Plan Regret 0.352. The example illustrates adaptation of both anchor location and staple mode to the lead context.

These cases complement the aggregate results with molecule-level comparisons of anchor placement, staple chemistry, and structural compatibility.

G Expanded receptor-interface refinement cases

Figure 4 provides an expanded molecular view of the receptor-aware local refinement analysis accompanying Figure 2 in the main text. We show three representative complexes (5onp, 6seo, and 3r7g) and compare the frozen interventions selected by Random Feasible, Generic-SB, Direct Plan

Ranker (DPR), and StapleBridge. For each hydrocarbon-stapled product, the E and Z stereoisomers are refined separately with the receptor present and visualized against the deposited bound peptide. This analysis is intentionally qualitative and is distinct from the aggregate $n = 31$ receptor-interface evaluation reported in Table 4. The aggregate analysis evaluates frozen reconstructed products in a common receptor coordinate frame without this additional receptor-aware refinement, whereas the examples here undergo a separate local refinement step. The examples illustrate how intervention choice can lead to different receptor-bound geometries even when all methods start from the same parent complex. In 5onp, the methods produce relatively compact refined conformations around the deposited binding region, while the E/Z stereoisomers retain method-specific local geometries. The 6seo and 3r7g examples exhibit larger differences across interventions, highlighting the sensitivity of the refined receptor-bound configuration to staple placement and chemistry. Across all panels, the StapleBridge products are shown without any post-hoc product reselection; the refinement is applied after the intervention has been fixed.

H Reproducibility Statement

We fixed the splits, hard-support definition, control coefficient, PeptiVerse mode, E/Z aggregation, and checkpoint-selection rule across runs. Random Feasible, Generic-SB, DPR, and StapleBridge were evaluated over three predefined seeds, with results reported as mean $\pm$ standard deviation. Empirical Reference was deterministic and evaluated once. Each learned run used its predefined validation criterion for checkpoint selection. StapleBridge used validation $\mathrm{KL}(q^\star \| q_\theta)$. Test metrics were excluded from model selection.