PREFIX-REUSE EVOLUTIONARY SEARCH WITH NORMALIZED PROXY GUIDANCE FOR PROTEIN BINDER DESIGN

Anonymous authors
Paper under double-blind review

ABSTRACT

Test-time steering guides a pretrained diffusion model toward high-reward outputs without retraining. Effective guidance should be computed at the clean-sample estimate rather than at a noisy intermediate state. the key requirement is a useful local direction, the gradient of an objective the sampler explicitly optimizes. Moreover, inexpensive scores computed along unfinished trajectories can be unreliable, making the observed score of a completed candidate the more trustworthy basis for selection. In this work, we propose PRES-NPG, a framework that combines **P**refix-**R**euse **E**volutionary **S**earch (PRES) with **N**ormalized **P**roxy **G**uidance (NPG). NPG differentiates a proxy at the clean-sample estimate to obtain normalized local directions for inference-time steering. PRES evaluates completed candidates, selects their saved intermediate states based on these evaluations, and generates new continuations from the selected states across generations. We apply PRES-NPG to protein binder design using two models, PXDesign-d and Complexa. PRES-NPG improves binder quality through local directional guidance and produces a higher-quality retained set than both its registered random-parent control and independent best-of-N sampling, at a matched number of scored candidates. Code available at: The code can be found at this anonymous link: https://anonymous.4open.science/r/PRES-NPG-8E22.

1 INTRODUCTION

Test-time steering uses additional computation to shift the output distribution of a pretrained diffusion model toward a downstream criterion while keeping its parameters fixed (Singhal et al., 2025). Its design involves two coupled decisions: how to guide an individual denoising trajectory and how to allocate further sampling effort after observing completed outcomes. A common formulation encodes the downstream criterion as a terminal reward $r(x, \boldsymbol{c})$ and defines an exponentially tilted target distribution $\pi^\star(x \mid \boldsymbol{c}) \propto \mu(x \mid \boldsymbol{c}) e^{r(x,\boldsymbol{c})/\beta}$, where μ is the pretrained reference policy, x denotes a completed sample, $\boldsymbol{c}$ denotes conditioning information, and the fixed temperature $\beta > 0$ controls the tradeoff between reward and proximity to the pretrained distribution (Singhal et al., 2025; Li et al., 2024). The reward may be a score from a learned human-preference model (Xu et al., 2023) or a log-likelihood, such as $\log p(y \mid x, \boldsymbol{c})$ (Wu et al., 2023). We study this problem in computational protein design, where predictive scores guide selection and only a small subset of generated candidates is retained for further investigation.

Two common inference-time strategies are gradient-based guidance (Dhariwal & Nichol, 2021; Bansal et al., 2024) and best-of-N sampling (Singhal et al., 2025). Gradient-based guidance modifies denoising updates using a differentiable steering signal, whereas best-of-N sampling generates independent candidates and selects those with the highest terminal scores. Guidance can shape each trajectory, but it requires a differentiable reward or proxy during sampling; a reliable evaluator of completed candidates may not provide gradients (Singhal et al., 2025; Li et al., 2024). Independent best-of-N uses terminal scores to rank completed candidates, but does not use those outcomes to direct subsequent sampling from saved intermediate states.

These limitations motivate combining local guidance with further sampling informed by completed outcomes through two complementary scoring signals. A differentiable proxy provides a local

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.

direction for steering individual trajectories, while a terminal evaluator can score completed candidates even if it is nondifferentiable or unreliable on incomplete states (Bansal et al., 2024; Li et al., 2024). The terminal scores can then determine which trajectories receive further sampling.

In this work, we introduce PRES-NPG, a framework that combines **P**refix-**R**euse **E**volutionary **S**earch (PRES) with **N**ormalized **P**roxy **G**uidance (NPG). Within each trajectory, NPG differentiates a proxy at the denoiser's clean-sample estimate and applies the normalized direction to steer the noisy intermediate state, without differentiating through the denoiser. PRES extends this guided sampling across generations by selecting parents according to terminal fitness and resuming their saved intermediate states to generate new candidates. Final outputs are selected from the complete archive of candidates generated across generations.

Our methodological contributions are the following:

- **Objective and policy formulation.** We formulate the reward-tilted target $\pi^\star$ and construct the test-time policy π^{NPG} using normalized proxy directions, distinguishing exact continuation-based guidance from the implemented clean-estimate proxy.
- **Evolutionary search adaptation.** Building on evolutionary diffusion search, we formulate PRES over the guided continuation law. New candidates are generated by resuming saved intermediate states rather than re-noising completed samples. Parents are selected by observed terminal fitness from the current generation, and final outputs are selected from the complete archive.
- **Empirical evaluation.** We compare parent-selection rules and independent generation at the same number of scored candidates and a common final retention rule. We then re-score the same retained candidates with a second verifier to determine whether gains under the selection score persist under a different evaluation. We also apply the checkpoint mechanism to a second frozen sampler, Complexa's product-space flow matcher, and verify that restoring a checkpoint reproduces the original trajectory under the same random numbers.

2 PRELIMINARIES

2.1 REFERENCE POLICY AND REWARD TARGET

We seek a sampling policy that improves terminal reward while remaining close to a frozen generative model. We index denoising in sampling order, $x_0 \to \cdots \to x_T$, where x_0 is initial noise and x_T is the completed sample. Fixed conditioning is suppressed, and states include the variables required for Markov transitions. Let μ denote the pretrained reference policy, suppressing its fixed parameters. We use the same symbol for a policy and its induced distributions, with arguments indicating the relevant variables. The reference trajectory density factorizes as $\mu(x_{0:T}) = \mu(x_0)\prod_{t=0}^{T-1}\mu(x_{t+1} \mid x_t)$, where $\mu(x_0)$ is the initial noise density and $\mu(x_{t+1} \mid x_t)$ is the denoising transition. Integrating out $x_{0:T-1}$ gives the output density $\mu(x_T)$.

For a bounded terminal reward r and a fixed temperature $\beta > 0$, the desired output distribution solves

$$\pi^\star = \arg\max_{\pi} \left\{ \mathbb{E}_{x_T\sim\pi}[r(x_T)] - \beta D_{\mathrm{KL}}(\pi\|\mu) \right\}. \tag{1}$$

The reward term favors better outputs, while the KL term penalizes departure from the reference output distribution $\mu(x_T)$. Optimizing over normalized densities yields the exponential tilt (Li et al., 2024; Singhal et al., 2025):

$$\pi^\star(x_T) = \frac{\mu(x_T)e^{r(x_T)/\beta}}{Z_\beta}, \tag{2}$$

Here $Z_\beta = \mathbb{E}_{x_T\sim\mu}[e^{r(x_T)/\beta}]$ is the partition function, with the expectation over completed samples from μ. Smaller β gives greater weight to reward; larger β favors proximity to the reference policy. This target motivates test-time changes to the sampling policy while the pretrained parameters remain fixed.

2.2 DIFFUSION GUIDANCE

To construct a sampling policy with terminal density $\pi^\star$, consider a continuous-time reference policy μ given by $dX_t = b_t(X_t)\,dt + \sigma_t(X_t)\,dW_t$, where X_t denotes the random state corresponding to x_t,

with initial density $\mu(x_0)$ and terminal density $\mu(x_T)$. Here, b_t is the drift, W_t is standard Brownian motion, and $a_t(x) = \sigma_t(x)\sigma_t(x)^\top$ is the local noise covariance per unit time. Define the continuation moment

$$h_t(x_t) = \mathbb{E}_{x_T \sim \mu(\cdot|x_t)}\left[e^{r(x_T)/\beta}\right]. \tag{3}$$

The conditional density $\mu(x_T \mid x_t)$ describes completions from x_t under the reference policy. Thus, h_t averages their exponentiated terminal rewards, with $h_T(x_T) = e^{r(x_T)/\beta}$ and the same temperature β as in eq. (2).

A diffusion *score* is the gradient of a log state density (Song et al., 2021); the reference score at time t is $\nabla_{x_t} \log \mu(x_t)$. This vector-valued score is distinct from the scalar terminal reward r. Tilting the reference trajectories by $e^{r(X_T)/\beta}$ gives the time-t marginal density $\mu(x_t)h_t(x_t)/Z_\beta$, whose score satisfies

$$\nabla_{x_t} \log\left[\frac{\mu(x_t)h_t(x_t)}{Z_\beta}\right] = \nabla_{x_t} \log \mu(x_t) + \nabla_{x_t} \log h_t(x_t). \tag{4}$$

Thus, $\nabla_{x_t} \log h_t(x_t)$ is the exact score correction. Under the regularity conditions for the Doob h transform (Chopin et al., 2023), the corresponding guided dynamics are

$$dX_t = [b_t(X_t) + a_t(X_t)\nabla \log h_t(X_t)]\, dt + \sigma_t(X_t)\, dW_t^\star, \tag{5}$$

where $W_t^\star$ is Brownian motion under the tilted law. The score correction becomes a drift correction through multiplication by a_t; a numerical sampler then converts the drift into a coordinate update. With initial density $h_0(x_0)\mu(x_0)/Z_\beta$, this process has terminal density $\pi^\star$. Modifying only the drift while retaining the reference initialization does not generally recover that target. Computing h_t is generally intractable; section 3.1 develops a tractable proxy-based approximation using denoised predictions.

2.3 Trajectory search

Evolutionary algorithms update candidate populations through selection and variation (Beyer & Schwefel, 2002). For denoising trajectories, a *completed rollout* produces x_T, a *checkpoint* stores the sampler state at an intermediate time t, and a *continuation* resumes that state with fresh randomness under a specified sampling policy (He et al., 2025). A completed rollout provides an observed terminal reward, whereas a checkpoint's *continuation value* is the expected terminal reward of future completions under that policy. These quantities differ: a score that reliably ranks completed outputs need not reliably value intermediate states (Li et al., 2024). This motivates our division of labor: a differentiable proxy guides individual trajectories, and observed terminal rewards select saved prefixes for further sampling.

3 Method

Given a frozen conditional generator μ and a budget of B scored candidates, our goal is to return K high-reward protein binders without retraining. The conditioning $\boldsymbol{c}$ specifies the target structure, cropped epitope region, and hotspot residues. We use two complementary signals: a differentiable structural proxy $\widetilde{r}$ supplies local coordinate gradients, while a terminal selection fitness f ranks completed candidates using only metric values, without requiring gradients. The fitness f instantiates the generic terminal reward r from section 2.1 in our search. The proxy encodes tractable structural preferences and need not estimate f.

PRES-NPG uses these signals at different stages of sampling (Figure 1). NPG first turns μ into the guided policy π^{NPG} by modifying individual denoising updates (section 3.1). PRES then constructs π^{EVL} by selecting completed trajectories according to f, reusing their saved prefixes, and generating fresh guided continuations (section 3.2). The final output is the best K candidates from the complete search archive. Algorithm 1 in Appendix A summarizes the full procedure from μ to the final outputs of π^{EVL}. A held-out verifier e assesses these outputs but participates in neither guidance nor selection.

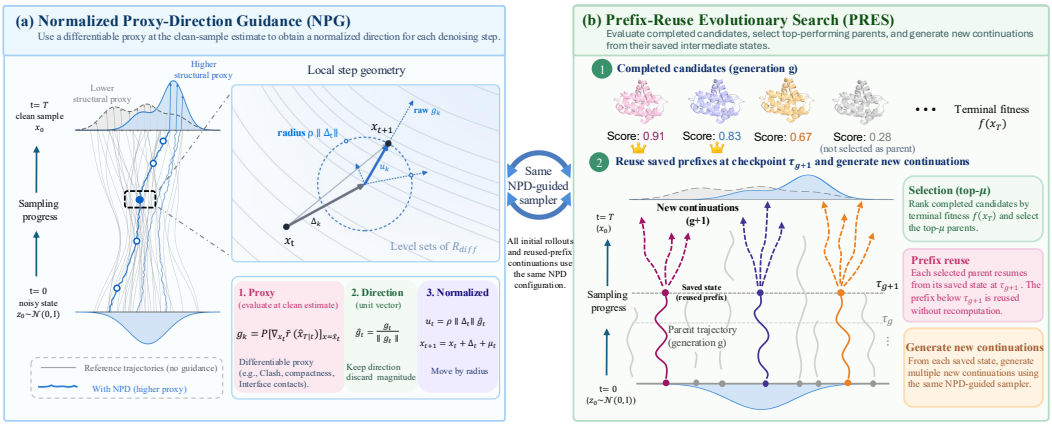


Figure 1: Overview of PRES-NPG. **(a) NPG:** At each denoising step, the frozen denoiser estimates a clean sample from the current noisy state. A differentiable structural proxy evaluated on this estimate supplies a local direction; after coordinate restriction and normalization, the direction is scaled by $\rho\|\Delta_t\|$ and added to the sampler update. The blue trajectories illustrate guidance toward higher proxy values relative to the unguided reference trajectories. **(b) PRES:** Completed trajectories in generation g are ranked by the fitness of their terminal samples, and the top M are selected as parents. Their saved prefixes through time $t(g+1)$ are reused without recomputation; each prefix is extended into N child trajectories with fresh randomness under the same NPG-guided sampler.

3.1 Proxy-Guided Sampling

At each denoising step, NPG takes the current state x_t, the frozen sampler μ, and the differentiable proxy $\widetilde{r}$, and returns an updated state x_{t+1}, as illustrated in Figure 1(a). Running the reference sampler provides an estimate $\widehat{x}_{T|t}$ of the clean terminal sample x_T and a base displacement Δ_t.

Guidance via Denoised Predictions. We evaluate the proxy on $\widehat{x}_{T|t}$ to assess the predicted structure rather than the noisy coordinates (Bansal et al., 2024). In our binder-design setting, $\widetilde{r}$ is a configured weighted sum of scale-normalized terms for clash avoidance, compactness, interface contacts, and hotspot proximity, oriented so that larger values are preferred; Appendix C gives their definitions. To connect this local signal to the continuation moment in eq. (3), we use the heuristic surrogate $\widetilde{h}_t(x_t) = e^{\widetilde{r}(\widehat{x}_{T|t})/\lambda}$, where $\lambda > 0$ is a proxy scale. This replaces the terminal reward with a structural proxy and the distribution of possible completions with a single denoised prediction. Where the functions are differentiable, the corresponding surrogate score correction is

$$\nabla_{x_t} \log \widetilde{h}_t(x_t) = \lambda^{-1} J_t^\top \nabla_{\widehat{x}_{T|t}} \widetilde{r}(\widehat{x}_{T|t}), \tag{6}$$

where $J_t = \partial \widehat{x}_{T|t} / \partial x_t$ is the denoiser Jacobian. The factor $J_t^\top$ maps the clean-coordinate gradient back to the noisy input. Computing this full correction differentiates through both the denoiser and the proxy.

Normalized Proxy Guidance. To obtain an inexpensive local direction, NPG uses the clean-coordinate derivative in eq. (6) but omits the denoiser Jacobian J_t. The guided positional variables in x_t and $\widehat{x}_{T|t}$ share the same Cartesian representation, so this derivative can be applied as a direction in the corresponding noisy-state coordinates. Let P denote the fixed projection onto binder coordinates, implemented as a mask that retains gradient components on binder atoms and sets all other components to zero. The guidance direction is

$$g_t = \mathsf{P}\left[\nabla_{\widehat{x}_{T|t}} \widetilde{r}(\widehat{x}_{T|t})\right]. \tag{7}$$

The mask prevents direct guidance updates to the conditioned target. Differentiation passes through the proxy alone, avoiding the denoiser backward pass and its activation storage. Omitting J_t is a directional approximation, not an exact score correction; Appendix B analyzes its local validity and limitations.

We next convert this direction into a coordinate displacement. For sampling steps $t = 0, \dots, T-1$, let $\mathcal{W} \subseteq \{0, \dots, T-1\}$ be the active guidance window, $\rho \geq 0$ the guidance-to-base step ratio, and $\delta_g \geq 0$ the numerical liveness threshold. We set

$$u_t = \begin{cases} \rho \|\Delta_t\| \, g_t / \|g_t\|, & t \in \mathcal{W} \text{ and } \|g_t\| > \delta_g, \\ 0, & \text{otherwise} \end{cases}. \tag{8}$$

The window is mapped through the sampler's native schedule, and norms and rescaling are computed separately for each candidate. On active steps, $\|u_t\| = \rho \|\Delta_t\|$: the proxy gradient determines the direction, while the reference displacement determines the scale. Thus, even a small gradient above the threshold receives the prescribed relative displacement. The intervention vanishes outside the window, for an inactive gradient, or when the base displacement is zero. Positive scalar factors such as $1/\lambda$ cancel under ideal normalization, so ρ controls the implemented step size rather than the target temperature β; its effect can still differ across samplers.

Adding this displacement to the reference update defines the guided policy:

$$\begin{aligned} x_t + \Delta_t &\sim \mu(\cdot \mid x_t), \\ x_{t+1} = x_t + \Delta_t + u_t &\sim \pi^{\mathrm{NPG}}(\cdot \mid x_t). \end{aligned} \tag{9}$$

Here u_t has the same units as the sampled coordinates; it is added to the state, not directly to the diffusion score. Any augmented solver or batch context follows its configured update, and coupled sampling is interpreted jointly over the population state. The reference initialization $x_0 \sim \mu$ is unchanged. This practical policy is motivated by $\pi^\star$ but does not guarantee exact sampling from it.

Iterating π^{NPG} produces a completed candidate x_T and saved intermediate states. Evaluating the terminal selection fitness $f(x_T)$ can now supply information unavailable to the local proxy, even when f has no usable gradient. The next stage uses these observed outcomes to select which saved states receive further guided continuations.

3.2 Prefix-Reuse Evolutionary Search

PRES takes completed trajectories and checkpoints from π^{NPG} and uses their terminal selection fitness to allocate further sampling, as illustrated in Figure 1 (b). We treat f as a black-box metric: only its values on completed candidates are required, so it may be non-differentiable or include an external evaluation pipeline. In our experiments, f is the AF2 ipTM score (Appendix D.2). Alternating fitness-based parent selection with fresh guided continuations constructs π^{EVL} (He et al., 2025; Yang et al., 2024a).

Let $g = 0, \dots, G$ index generations. Each branching generation $g \geq 1$ selects M parents and generates N children per parent, giving MN completed trajectories. We use $i = 1, \dots, M$ for the selected parent and $j = 1, \dots, N$ for its child. The branch time $t(g)$ follows sampling order, with $0 = t(0) < t(1) < \cdots < t(G) < T$; native schedule positions are mapped to these denoising steps.

Generation zero supplies an initial population $\mathcal{P}^0$ of MN full trajectories, drawn jointly as $(\Gamma^{0,\ell})_{\ell=1}^{MN} \sim \pi_0^{\mathrm{NPG}}$, where $\Gamma^{0,\ell} = (x_{0:T}^{0,\ell})$. Here π_0^{NPG} includes a fresh reference state for each candidate slot and guided denoising to T. Each rollout saves checkpoints at the later branch times.

Parent Selection and Prefix Reuse. To generate population g, select the M best completed trajectories from $\mathcal{P}^{g-1}$:

$$(\Gamma_{\mathrm{par}}^{g,1}, \dots, \Gamma_{\mathrm{par}}^{g,M}) = \mathrm{Top}_M(\mathcal{P}^{g-1}; f). \tag{10}$$

The operator ranks trajectories by f evaluated on their terminal samples at time T, not on their intermediate states. Selection is across all MN completed trajectories, rather than across averages of a parent's children. The selected trajectories are relabeled by rank i; ties are broken by candidate slot.

For selected parent i, let $\tau^{g,i} = \Gamma_{\mathrm{par}}^{g,i}\big|_{0:t(g)}$ denote its saved trajectory prefix through $t(g)$. Each of its N continuations receives a child-specific label, $\tau^{g,i,j} = \tau^{g,i}$. Thus, all children of a parent reuse the same prefix content, while j distinguishes their subsequent sampling runs. A saved prefix includes the checkpoint's solver and batch context needed to resume at $t(g)$, so no prefix denoising is repeated. Candidate slots are ordered as $(j-1)M + i$, preserving cyclic allocation across ranked parents.

Guided Continuation and Terminal Fitness. Starting from $\tau^{g,i,j}$, NPG generates a new suffix from $t(g)$ to T with fresh randomness. Denote its realized terminal sample by $x_{T|\tau^{g,i,j}}$. The full child trajectory and its fitness are

$$\Gamma^{g,i,j} = \big(\tau^{g,i,j},\, x_{t(g)+1:T|\tau^{g,i,j}}\big), \tag{11}$$

$$f^{g,i,j} = f\big(x_{T|\tau^{g,i,j}}\big)\,. \tag{12}$$

The first line concatenates the saved prefix and the newly sampled suffix. In particular, $x_{T|\tau^{g,i,j}}$ is an actual completed sample, not the denoiser prediction $\widehat{x}_{T|t}$ or a deterministic function of the prefix. Identical prefixes can therefore lead to different terminal samples and fitnesses. Fitness is evaluated only after denoising reaches T. PRES ranks these realized outcomes without separately estimating the expected fitness of future continuations from a prefix.

Let $\boldsymbol{\tau}^g = (\tau^{g,i,j})_{i,j}$ collect the labeled prefixes. Define $\pi^{\mathrm{NPG}}_{t(g)}(\cdot \mid \boldsymbol{\tau}^g)$ as the joint law of full trajectories formed by retaining these prefixes and sampling only their new suffixes. The conditional generation step of the evolutionary policy is

$$\pi^{\mathrm{EVL}}(\cdot \mid \mathcal{P}^{g-1}) := \pi^{\mathrm{NPG}}_{t(g)}(\cdot \mid \boldsymbol{\tau}^g). \tag{13}$$

Its output forms $\mathcal{P}^g = \{\Gamma^{g,i,j} : i = 1, \ldots, M,\ j = 1, \ldots, N\}$, which supplies parents for generation $g + 1$. This joint formulation preserves native batch coupling and does not assume independent children. Selection and shared prefixes also induce dependence across generations, while the generator remains frozen.

Final Retention. Let $\mathcal{H}^0 = \{x_T^{0,\ell} : \ell = 1, \ldots, MN\}$ contain the initial terminal samples. The final archive $\mathcal{H} = \mathcal{H}^0 \uplus \biguplus_{g=1}^{G} \{x_{T|\tau^{g,i,j}}\}_{i,j}$ retains all $B = (G + 1)MN$ terminal samples. This multiset preserves candidate identities and multiplicities. Earlier samples remain eligible for final delivery even when their trajectories are no longer eligible parents.

Let $I_K(\mathcal{H}; f)$ index the best K archived terminal samples under f, for $1 \le K \le B$. Returning these samples completes π^{EVL}. Final ties follow the frozen readout's archive ordering and sorting convention, separately from the parent-selection tie-break. Thus, NPG uses differentiable proxy information to generate continuations, while PRES uses the fitness of their completed samples to select the next prefixes and the final outputs.

3.3 Evaluation statistics and controls

To distinguish final-output quality from whole-pool quality, enumerate archived terminal samples by $x_{T,1}, \ldots, x_{T,B}$. For reporting function u, the retained mean is $\widehat{V}_K(\mathcal{H}; u) = K^{-1} \sum_{\ell \in I_K(\mathcal{H};f)} u(x_{T,\ell})$, and the whole-pool mean is $\widehat{V}_{\mathrm{pool}}(\mathcal{H}; u) = B^{-1} \sum_{\ell=1}^{B} u(x_{T,\ell})$. Selection always uses f; u determines the reported endpoint. Setting $u = f$ reports selection fitness, whereas $u = e$ reports held-out evaluation of the same candidates. These are ordinary averages over different output sets, so an improvement in the whole-pool mean need not appear at a fixed retention depth K.

Matched Controls. PRES selects parents using eq. (10). The random-parent control instead samples M parents uniformly while keeping branching, scored-candidate budget, and final retention fixed, thereby isolating the effect of parent selection. Independent best-of-B applies the same retention rule to separately seeded full batches under an equal scored-candidate budget; independence is required across batches, not within them. This control does not match generation cost because PRES reuses prefixes. The controls therefore address different comparisons. Poorer random-parent performance affects interpretation, but does not bias the specified contrast.

Algorithm Overview. Algorithm 1 in section A summarizes PRES-NPG. At each generation, the method scores candidate trajectories, selects M parents using eq. (10), and uses NPG guidance to generate N continuations from each selected prefix. This process repeats until the scored-candidate budget is exhausted, after which the final retention rule selects the output designs.

4 EXPERIMENT

This section presents a comprehensive empirical evaluation of PRES-NPG. We ask three questions: (i) Can Normalized Proxy Guidance (NPG) steer a frozen sampler toward higher-quality binder candidates? (ii) At a matched scored-candidate budget, does score-based parent selection with prefix reuse improve the delivered set's selection score? (iii) Which guidance and search components matter most, and what quality–diversity trade-offs do the ablations reveal?

Experiment setup

Dataset. We evaluate on protein binder-design targets from ProtDBench (Liu et al., 2026) and an additional eight-target panel derived from BindCraft (Pacesa et al., 2025). Appendix D.1 details the target specifications and preparation.

Metrics. We evaluate generated binders using structural confidence, interface energy, sequence plausibility, design success, diversity, and novelty. AF2 ipTM guides selection, while Boltz-2 ipTM provides cross-verifier evaluation. Together, these metrics assess predicted design quality, structural variety, and similarity to known structures. Higher values are preferred except for Rosetta interface energy and mean maximum reference similarity, where lower values indicate more favourable energy and greater reference-set novelty, respectively. Detailed definitions are provided in Appendix D.2.

Baselines

We evaluate PRES-NPG against two registered controls: random-parent branching and independent best-of-64. Additional comparisons cover stock sampling, guidance-only sampling, best-of-N selection (N=128), and RFdiffusion (Watson et al., 2023). The broader baseline discussion also includes BoltzGen (Stark et al., 2025), search-only sampling, random selection, and stock sampling with Complexa (Didi et al., 2026). Method-specific descriptions, implementation choices, and the evaluation coverage of each baseline are provided in Appendix D.3.

4.1 RESULTS

Table 1 compares PRES-NPG with two registered controls and additional baselines on ProtDBench. In both the retained set and the full pool, PRES-NPG records higher AF2 and Boltz-2 i_{pTM}, higher AF2 success rates, and more favorable Rosetta energies than either control. The whole-pool gains show that the improvement is not limited to final selection. These gains come with lower structural diversity, particularly relative to the control without generational search. RFdiffusion covers only six targets and is therefore not directly comparable across the full panel.

Table 2 shows a similar pattern with Proteina-Complexa. In both the retained set and the full pool, PRES-NPD achieves higher Complexa and Boltz-2 i_{pTM}, higher success rates, and more favorable Rosetta energies than both controls. However, it produces fewer clusters and unique designs, again indicating a trade-off between predicted quality and structural diversity.

4.2 VISUALIZATION

We present five generated binders, one per target, in Figure 2. PRES-NPG produces binders with proper geometries and binding poses: the designed chain (wheat) packs against the specified hotspot residues (orange) and conforms to the receptor pocket (light blue), burying a substantial interface. The designs span folds, comprising helical bundles and one mixed α/β binder. We annotate every panel with three complementary metrics: the Rosetta total score of the relaxed complex and the AlphaFold2 confidence scores iPTM and pLDDT. Across the five examples our designs achieve high pLDDT and iPTM, indicating that the predicted binder–target complexes are structurally plausible.

4.3 ABLATION

Table 3 compares guidance formulations and retrospective selection rules. PRES-NPG achieves the strongest overall AF2 and Boltz-2 performance. Removing the hotspot term or replacing NPG with resampling reduces performance, while random and reversed guidance further support the value of the proxy gradient. Other variants offer isolated advantages in Rosetta energy, AF2-easy rate, pLDDT,

Table 1: PRES-NPG and two registered controls on ten ProtDBench targets, alongside additional baselines. For these three arms, the upper block reports the AF2-selected top eight designs; the lower block reports all 64 designs per cell. RFdiffusion covers six targets because its pipeline does not support the multichain receptor crops for VEGFA, IL17A, H1, and TNFa. Reference similarity is computed over the full pool. Lower Rosetta ΔG_{sep} is better.

Model + sampler	Rosetta ΔG_{sep}	i_{pTM} AF2	i_{pTM} Boltz-2	pLDDT	ESM-2 PLL	AF2-easy %	AF2-opt %	Clust. 6 Å	Uniq. %	Ref. sim.
Retained 8 of 64 — the delivered set										
PRES-NPG (`es`)	**−33.75**	**0.8061**	**0.8321**	**0.9192**	**−2.0101**	**77.1**	**45.4**	4.11	51.4	0.7583
− score-based selection (`rnd`)	−32.55	0.7255	0.7965	0.9044	−2.0067	67.1	37.1	4.53	56.6	0.7466
− generational search (`bon`)	−31.45	0.7508	0.8035	0.9093	−2.0112	71.7	44.2	**6.97**	**87.1**	0.7502
Stock / best-of-128	−29.62	0.7895	0.7801	0.9176	−1.9974	84.6	46.7	6.70	83.8	0.7652
Search-only resampling	−29.53	0.7267	0.7794	0.9097	−1.9827	71.7	36.7	6.30	78.8	0.7518
RFdiffusion + ProteinMPNN (6/10)	−29.72	0.7290	0.6842	0.9132	−1.9668	63.9	43.8	4.23	52.9	0.7598
Whole pool — no selection at readout										
PRES-NPG (`es`)	**−31.26**	**0.4932**	**0.7523**	**0.9004**	−1.9672	**39.2**	**18.8**	21.62	33.8	0.7591
− score-based selection (`rnd`)	−29.86	0.3420	0.6928	0.8999	−1.9639	20.1	8.1	21.74	34.0	0.7529
− generational search (`bon`)	−29.26	0.3584	0.7003	0.8956	**−1.9761**	21.8	9.3	**37.33**	**58.3**	0.7443
Stock / best-of-128	−26.82	0.3446	0.6928	0.8977	−1.9539	20.2	8.9	24.32	19.0	0.7584
Search-only resampling	−26.57	0.3272	0.6824	0.9001	−1.9627	19.0	7.7	25.37	39.6	0.7473
RFdiffusion + ProteinMPNN (6/10)	−28.61	0.4073	0.6930	0.8872	−1.9590	22.1	9.5	23.61	36.9	0.9582

Table 2: PRES-NPD and two controls using Proteina-Complexa. The upper block reports the top eight designs selected by Complexa's native i_{pTM} score after ProteinMPNN redesign and folding; the lower block reports all 64 designs per cell. Structural diversity uses Foldseek clustering at TM-score ≥ 0.5. Reference similarity is computed over the full pool.

Model + sampler	Rosetta ΔG_{sep}	i_{pTM} mpnn	i_{pTM} Boltz-2	pLDDT mpnn	ESM-2 PLL	Succ. %	Clust. 6 Å	Uniq. %	Ref. sim.
PRES-NPD, Complexa	**−32.06**	**0.6510**	**0.8102**	**0.8301**	**−1.8692**	**30.6**	3.22	40.25	0.7137
− Rollout-guided sel., Complexa	−30.77	0.5691	0.7937	0.8167	−1.8535	20.0	4.53	56.6	0.7285
− Generational search, Complexa	−29.72	0.6285	0.7832	0.8267	−1.8622	22.5	**5.85**	**73.1**	0.7255
Stock sampling, Complexa	−28.32	0.5813	0.7804	0.8072	−1.8316	11.8	4.12	51.5	0.7269
Complexa, whole pool of 64									
PRES-NPD, Complexa	**−31.71**	**0.4302**	**0.7306**	**0.7762**	**−1.8627**	**18.4**	10.75	16.8	0.7137
− Rollout-guided sel., Complexa	−26.35	0.3050	0.6833	0.7642	−1.8509	12.5	13.67	21.3	0.7285
− Generational search, Complexa	−29.99	0.3147	0.6692	0.7524	−1.8486	12.1	**27.65**	**43.2**	0.7255

or diversity. PRES-NPG yields a concentrated retained set but a diverse candidate pool. Because guidance variants were evaluated in separate, unpaired batches, these comparisons are descriptive rather than statistically conclusive. All selection rules improve AF2 and Boltz-2 scores over the unselected pool, but no rule dominates every metric. AF2-IG performs best on AF2 and Rosetta energy, Protenix-mini on Boltz-2 and pLDDT, and corrected Boltz-2 selection on the AF2-easy rate. Correcting the Boltz-2 complex specification also improves results over the legacy setup. Overall, selector rankings depend on the evaluation metric and do not indicate a general same-evaluator advantage. Cluster counts are not directly comparable because the selected sets contain eight designs, whereas the full pool contains 64.

5 Related work

Gradient-Based Guidance. Classifier guidance adds a scaled log-classifier gradient to the transition mean, with the classifier trained on noised states (Dhariwal & Nichol, 2021). Chroma applies that construction to protein structure, using classifiers trained to predict labels from noisy examples (Ingraham et al., 2023). A second line evaluates the objective at the model's clean estimate, as in diffusion posterior sampling and the forward-guidance component of Universal Guidance (Chung et al., 2022; Bansal et al., 2024). These updates backpropagate through the denoiser, retaining the Jacobian factor. Q-Guided Flow (QGF) guides a flow policy over actions by evaluating a learned critic at a one-step estimate of the denoised action. It replaces the corresponding Jacobian with the identity and reports improved performance over the variant that retains the Jacobian in its evaluated settings (Zhou et al., 2026). Guided-TTS scales the classifier gradient to a fixed multiple of the

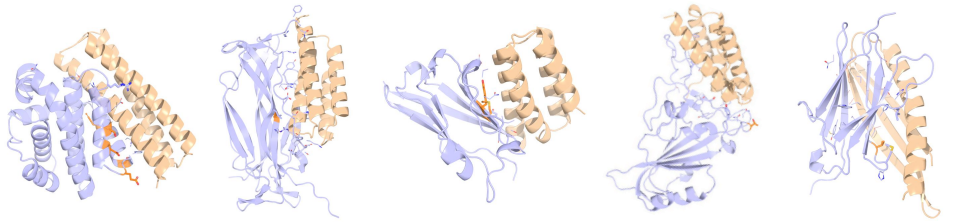


Figure 2: Binders (wheat) designed by PRES-NPG, where targets are in light blue with hotspot residues as orange sticks. Left to right: BHRF1, IL17A, PD-L1, SARS-CoV-2 RBD, TrkA.

Table 3: Ablation axis. Each cell reads **retained 8 (whole pool)** from one population: the parenthesised entry is reduced over the cell's whole pool and the entry before it over the same cell's top eight by AF2 i_{pTM}; the two are never pooled, and are headed accordingly.

	Rosetta ΔG_{sep}	i_{pTM} AF2	i_{pTM} Boltz-2	pLDDT	AF2-easy %	Clust. 6 Å
B. Guidance form						
PRES-NPG	**−33.75 (−31.26)**	**0.8061 (0.4932)**	**0.8321 (0.7523)**	**0.9192 (0.9004)**	77.1 (**39.2**)	4.11 (51.4)
w/o hotspot term	−33.53 (−31.09)	0.7728 (0.3641)	0.8094 (0.7031)	0.9114 (0.8939)	77.5 (23.1)	7.00 (38.63)
Burial-only reward	−33.78 (−33.13)	0.7724 (0.3583)	0.7904(0.6823)	0.9087 (0.8900)	71.3 (20.7)	7.23 (41.40)
Geometry reward	−33.10 (−30.17)	0.7539 (0.3481)	0.7792(0.6716)	0.9074 (0.8975)	64.6 (19.4)	6.93 (38.77)
Allocation	−32.34 (−29.50)	0.7555 (0.3510)	0.8048 (0.6947)	0.9096 (0.8941)	71.7 (20.8)	7.10 (39.30)
Guidance + resampling	−31.57 (−28.48)	0.7475 (0.3415)	0.8009 (0.6894)	0.9072 (0.8986)	67.5 (19.5)	6.40 (26.53)
Multi-objective	−30.30 (−27.25)	0.7415 (0.3614)	0.8027 (0.6899)	0.9122 (0.8980)	75.8 (22.8)	6.63 (35.27)
Random-direction	−30.14 (−26.91)	0.7398 (0.3508)	0.7912(0.6631)	0.9108 (0.8978)	78.3 (21.9)	6.83 (36.10)
Resampling only	−29.53 (−26.57)	0.7267 (0.3272)	0.7794 (0.6824)	0.9097 (0.9001)	71.7 (19.0)	6.30 (25.37)
Smooth-min	−28.36 (−26.46)	0.7021 (0.2517)	0.7288 (0.6498)	0.9152 (0.9150)	68.8 (9.9)	**7.67** (**49.17**)
Reversed reward	−25.43 (−21.92)	0.6417 (0.2815)	0.7431(0.6427)	0.9010 (0.8974)	57.1 (14.7)	6.90 (35.80)
C. Selection rule						
AF2-IG	−29.85	0.4618	0.6635	0.9084	47.2	6.94
Protenix-mini	−28.21	0.4543	**0.6731**	**0.9134**	46.4	**7.19**
Boltz-2, corrected complex	**−29.72**	**0.4614**	0.6412	0.9114	**48.1**	7.00
Boltz-2, legacy (mis-spec.)	−29.45	0.4135	0.6631	0.9083	40.2	6.96
No selection (pool of 64)	−28.40	0.3333	0.6355	0.9043	27.6	36.69

unconditional score norm (Kim et al., 2022), and DSG constrains guided updates to a sphere of prescribed radius centered at the unconditional transition mean (Yang et al., 2024b).

Evolutionary Search and Refinement. Twisted diffusion sampling and Feynman–Kac steering carry weighted particles against specified targets (Wu et al., 2023; Singhal et al., 2025). Each particle carries a weight built from a potential, together with a model–proposal transition ratio where the proposal departs from the model. Controlled sequential Monte Carlo learns proposal adjustments in the same framework (Heng et al., 2020). Diffusion-ES draws parents with probability proportional to an exponentiated terminal reward and re-noises selected completed outputs (Yang et al., 2024a). Partial diffusion with subsequent filtering is an established refinement approach in protein design (Vázquez Torres et al., 2024). PRES-NPG selects parents using completed-rollout fitness and resumes their saved intermediate states under normalized proxy guidance, without re-noising completed designs.

6 Conclusion

We introduced PRES-NPG, a test-time framework that combines normalized proxy guidance with evolutionary search over saved diffusion states. By steering individual trajectories and reusing prefixes selected from completed outcomes, it improves predicted binder quality without retraining or backpropagating through the denoiser. Experiments with PXDesign-d and Proteina-Complexa demonstrate gains over independent sampling at matched candidate budgets, alongside reduced structural diversity. These results support combining local guidance with outcome-driven trajectory reuse for protein binder design.

AI USE STATEMENT

AI-assisted tools were used to improve the clarity and presentation of the manuscript, support literature retrieval and the identification of related work, and assist in monitoring the execution of the experiment.

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, et al. The rosetta all-atom energy function for macromolecular modeling and design. *Journal of chemical theory and computation*, 13(6):3031, 2017.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Roni Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. In *International Conference on Learning Representations*, volume 2024, pp. 51304–51323, 2024.

Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1): 235–242, 2000.

Hans-Georg Beyer and Hans-Paul Schwefel. Evolution strategies–a comprehensive introduction. *Natural computing*, 1(1):3–52, 2002.

Léon Bottou, Frank E Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *SIAM review*, 60(2):223–311, 2018.

Nicolas Chopin, Andras Fulop, Jeremy Heng, and Alexandre H Thiery. Computational doob h-transforms for online filtering of discretely observed diffusions. In *International Conference on Machine Learning*, pp. 5904–5923. PMLR, 2023.

Hyungjin Chung, Jeongsol Kim, Michael T Mccann, Marc L Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. *arXiv preprint arXiv:2209.14687*, 2022.

Justas Dauparas, Ivan Anishchenko, Nathaniel Bennett, Hua Bai, Robert J Ragotte, Lukas F Milles, Basile IM Wicky, Alexis Courbet, Rob J de Haas, Neville Bethel, et al. Robust deep learning–based protein sequence design using proteinmpnn. *Science*, 378(6615):49–56, 2022.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

Kieran Didi, Zuobai Zhang, Guoqing Zhou, Danny Reidenbach, Zhonglin Cao, Sooyoung Cha, Tomas Geffner, Christian Dallago, Jian Tang, Michael Bronstein, et al. Scaling atomistic protein binder design with generative pretraining and test-time compute. In *International Conference on Learning Representations*, volume 2026, pp. 61056–61109, 2026.

Alex Dytso, H Vincent Poor, and Shlomo Shamai Shitz. Conditional mean estimation in gaussian noise: A meta derivative identity with applications. *IEEE Transactions on Information Theory*, 69 (3):1883–1898, 2022.

Bradley Efron. Tweedie's formula and selection bias. *Journal of the American Statistical Association*, 106(496):1602–1614, 2011.

Karl Gustafson. The angle of an operator and positive operator products. *Bulletin of the American Mathematical Society*, 74(3):488–492, 1968.

Haoran He, Jiajun Liang, Xintao Wang, Pengfei Wan, Di Zhang, Kun Gai, and Ling Pan. Scaling image and video generation via test-time evolutionary search. *arXiv preprint arXiv:2505.17618*, 2025.

Jeremy Heng, Adrian N Bishop, George Deligiannidis, and Arnaud Doucet. Controlled sequential monte carlo. *The Annals of Statistics*, 48(5):2904–2929, 2020.

John B Ingraham, Max Baranov, Zak Costello, Karl W Barber, Wujie Wang, Ahmed Ismail, Vincent Frappier, Dana M Lord, Christopher Ng-Thow-Hing, Erik R Van Vlack, et al. Illuminating protein space with a programmable generative model. *Nature*, 623(7989):1070–1078, 2023.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.

Wolfgang Kabsch. A solution for the best rotation to relate two sets of vectors. *Foundations of Crystallography*, 32(5):922–923, 1976.

Heeseung Kim, Sungwon Kim, and Sungroh Yoon. Guided-tts: A diffusion model for text-to-speech via classifier guidance. In *International Conference on Machine Learning*, pp. 11119–11133. PMLR, 2022.

Xiner Li, Yulai Zhao, Chenyu Wang, Gabriele Scalia, Gokcen Eraslan, Surag Nair, Tommaso Biancalani, Shuiwang Ji, Aviv Regev, Sergey Levine, et al. Derivative-free guidance in continuous and discrete diffusion models with soft value-based decoding. *arXiv preprint arXiv:2408.08252*, 2024.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.

Cong Liu, Milong Ren, Jiaqi Guan, Chengyue Gong, Jinyuan Sun, Xinshi Chen, and Wenzhi Xiao. Protdbench: A unified benchmark of protein binder design and evaluation. *arXiv preprint arXiv:2605.04118*, 2026.

Martin Pacesa, Lennart Nickel, Christian Schellhaas, Joseph Schmidt, Ekaterina Pyatova, Lucas Kissling, Patrick Barendse, Jagrity Choudhury, Srajan Kapoor, Ana Alcaraz-Serna, et al. One-shot design of functional protein binders with bindcraft. *Nature*, 646(8084):483–492, 2025.

Saro Passaro, Gabriele Corso, Jeremy Wohlwend, Mateo Reveiz, Stephan Thaler, Vignesh Ram Somnath, Noah Getz, Tally Portnoi, Julien Roy, Hannes Stark, et al. Boltz-2: Towards accurate and efficient binding affinity prediction. *BioRxiv*, 2025.

Raghav Singhal, Zachary Horvitz, Ryan Teehan, Mengye Ren, Zhou Yu, Kathleen McKeown, and Rajesh Ranganath. A general framework for inference-time scaling and steering of diffusion models. *arXiv preprint arXiv:2501.06848*, 2025.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Hannes Stark, Felix Faltings, MinGyu Choi, Yuxin Xie, Eunsu Hur, Timothy O'Donnell, Anton Bushuiev, Talip Uçar, Saro Passaro, Weian Mao, et al. Boltzgen: Toward universal binder design. *BioRxiv*, 2025.

Protenix Team, Milong Ren, Jinyuan Sun, Jiaqi Guan, Cong Liu, Chengyue Gong, Yuzhe Wang, Lan Wang, Qixu Cai, Wenzhi Ma, et al. Pxdesign: Fast, modular, and accurate de novo design of protein binders. *BioRxiv*, pp. 2025–08, 2025.

Michel Van Kempen, Stephanie S Kim, Charlotte Tumescheit, Milot Mirdita, Jeongjae Lee, Cameron LM Gilchrist, Johannes Söding, and Martin Steinegger. Fast and accurate protein structure search with foldseek. *Nature biotechnology*, 42(2):243–246, 2024.

Susana Vázquez Torres, Philip JY Leung, Preetham Venkatesh, Isaac D Lutz, Fabian Hink, Huu-Hien Huynh, Jessica Becker, Andy Hsien-Wei Yeh, David Juergens, Nathaniel R Bennett, et al. De novo design of high-affinity binders of bioactive helical peptides. *Nature*, 626(7998):435–442, 2024.

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

Luhuan Wu, Brian Trippe, Christian Naesseth, David Blei, and John P Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. *Advances in Neural Information Processing Systems*, 36:31372–31403, 2023.

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. Imagereward: Learning and evaluating human preferences for text-to-image generation. *Advances in Neural Information Processing Systems*, 36:15903–15935, 2023.

Brian Yang, Huangyuan Su, Nikolaos Gkanatsios, Tsung-Wei Ke, Ayush Jain, Jeff Schneider, and Katerina Fragkiadaki. Diffusion-es: Gradient-free planning with diffusion for autonomous and instruction-guided driving. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15342–15353. IEEE, 2024a.

Lingxiao Yang, Shutong Ding, Yifan Cai, Jingyi Yu, Jingya Wang, and Ye Shi. Guidance with spherical gaussian constraint for conditional diffusion. *arXiv preprint arXiv:2402.03201*, 2024b.

Yang Zhang and Jeffrey Skolnick. Tm-align: a protein structure alignment algorithm based on the tm-score. *Nucleic acids research*, 33(7):2302–2309, 2005.

Zhiyuan Zhou, Andy Peng, Charles Xu, Qiyang Li, Tobias Springenberg, Kevin Frans, and Sergey Levine. Test-time gradient guidance of flow policies in reinforcement learning. *arXiv preprint arXiv:2606.11087*, 2026.

A PRES-NPG ALGORITHM

Algorithm 1 gives the complete sampling procedure described in sections 3.1 and 3.2.

Algorithm 1 PRES-NPG: Sampling with π^{EVL}

Input: Frozen sampler μ, conditioning $\boldsymbol{c}$, proxy $\widetilde{r}$, terminal fitness f; M parents, N children per parent, G branching generations, K outputs; branch times $0 < t(1) < \cdots < t(G) < T$; guidance settings $\rho, \mathcal{W}, \delta_g, \mathsf{P}$.

Output: K terminal samples returned by π^{EVL} from $B = (G+1)MN$ scored candidates.

Initialize MN states $(x_0^{0,\ell})_{\ell=1}^{MN}$ from μ, with their native solver/batch context.
$(\Gamma^{0,\ell})_\ell \leftarrow \textsc{Continue}((x_0^{0,\ell})_\ell, 0)$; set $\mathcal{P}^0 \leftarrow \{\Gamma^{0,\ell}\}_\ell$.
Evaluate $f(x_T^{0,\ell})$ for all ℓ; initialize $\mathcal{H} \leftarrow \{x_T^{0,\ell}\}_\ell$ with these fitnesses.
for $g = 1, \ldots, G$ **do**
  $(\Gamma_{\mathrm{par}}^{g,1}, \ldots, \Gamma_{\mathrm{par}}^{g,M}) \leftarrow \mathrm{Top}_M(\mathcal{P}^{g-1}; f)$.
  Retrieve $\tau^{g,i} \leftarrow \Gamma_{\mathrm{par}}^{g,i}\big|_{0:t(g)}$ for $i = 1, \ldots, M$.
  Set $\tau^{g,i,j} \leftarrow \tau^{g,i}$ for $i = 1, \ldots, M$ and $j = 1, \ldots, N$.
  $(\Gamma^{g,i,j})_{i,j} \leftarrow \textsc{Continue}((\tau^{g,i,j})_{i,j}, t(g))$ (generation under π^{EVL}).
  Read each terminal sample $x_{T|\tau^{g,i,j}}$ from $\Gamma^{g,i,j}$; evaluate $f^{g,i,j} \leftarrow f(x_{T|\tau^{g,i,j}})$.
  $\mathcal{P}^g \leftarrow \{\Gamma^{g,i,j}\}_{i,j}$; $\mathcal{H} \leftarrow \mathcal{H} \uplus \{x_{T|\tau^{g,i,j}}\}_{i,j}$, retaining fitnesses.
end for
return the K highest-fitness terminal samples in $\mathcal{H}$.
Subroutine $\textsc{Continue}$(saved prefixes, t_{start}):
Restore prefix states and context (an initial x_0 is a one-state prefix); use fresh randomness.
for $t = t_{\mathrm{start}}, \ldots, T-1$ **do**
  if $t \in \{t(1), \ldots, t(G)\}$ **then**
    Save each trajectory's prefix and resumable solver/batch context at time t.
  end if
  Run μ jointly to obtain each candidate's denoised prediction $\widehat{x}_{T|t}$ and displacement Δ_t.
  for each candidate slot (indices suppressed) **do**
    $u_t \leftarrow 0$.
    if $t \in \mathcal{W}$ **then**
      $g_t \leftarrow \mathsf{P}[\nabla_{\widehat{x}_{T|t}} \widetilde{r}(\widehat{x}_{T|t})]$; differentiate only the proxy.
      if $\|g_t\| > \delta_g$ **then**
        $u_t \leftarrow \rho \|\Delta_t\| \, g_t / \|g_t\|$.
      end if
    end if
    $x_{t+1} \leftarrow x_t + \Delta_t + u_t$ (update under π^{NPG}).
  end for
  Apply native auxiliary-state updates and append the new states to their trajectories.
end for
return the completed trajectories, retaining their saved checkpoints.

B LOCAL DIRECTIONAL VALIDITY AND STABILITY OF NORMALIZED PROXY GUIDANCE

We analyze when the NPG direction in section 3.1 locally increases the composed proxy and how the step size affects this guarantee.

Setting. Fix the conditioning and sampling time t with corruption noise level $s_t > 0$. Under Gaussian corruption $X_t = X_T + s_t\xi$, with $\xi \sim \mathcal{N}(0, I)$ independent of X_T, write $\widehat{x}_{T|t}(y) = \mathbb{E}[X_T \mid X_t = y]$ to make the exact denoiser's dependence on its noisy input y explicit. Here s_t is the corruption standard deviation, distinct from the SDE diffusion coefficient $\sigma_t(x)$. Tweedie's formula and its derivative give (Efron, 2011)

$$\begin{aligned} \widehat{x}_{T|t}(y) &= y + s_t^2 \nabla_y \log p_t(y), \\ J := \frac{\partial \widehat{x}_{T|t}(y)}{\partial y} &= I + s_t^2 \nabla_y^2 \log p_t(y) \\ &= s_t^{-2} \operatorname{Cov}(X_T \mid X_t = y). \end{aligned} \tag{14}$$

Here p_t is the noisy marginal density in this Gaussian model. Within this idealized model, $J \approx I$ requires posterior covariance approximately $s_t^2 I$; isotropy alone gives only a scalar multiple of I. A positive scalar multiple preserves the normalized direction, but general anisotropy can change it. These identities do not establish that a learned denoiser satisfies the approximation.

The input is taken after augmentation and noise injection. Define

$$F(y) = \widetilde{r}(\widehat{x}_{T|t}(y)), \qquad \tilde{g}(y) = \nabla_x \widetilde{r}(x)|_{x=\widehat{x}_{T|t}(y)}. \tag{15}$$

Here $\tilde{g}$ is the unrestricted proxy gradient. The Jacobian J in eq. (14) is symmetric and positive semidefinite (Dytso et al., 2022, Proposition 1).

Directional Validity. For $\tilde{g} \neq 0$, the chain rule gives $\nabla F(y) = J\tilde{g}$, hence

$$\left\langle \nabla F(y), \frac{\tilde{g}}{\|\tilde{g}\|} \right\rangle = \frac{\tilde{g}^\top \operatorname{Cov}(X_T \mid X_t = y)\tilde{g}}{s_t^2 \|\tilde{g}\|} \geq 0. \tag{16}$$

The inequality is strict when $J\tilde{g} \neq 0$. Thus, the unrestricted direction is locally ascending under this condition, without requiring $J \approx I$. It agrees with the normalized gradient of F exactly when $J\tilde{g} = c\tilde{g}$ for some $c > 0$. When $J\tilde{g} \neq 0$, the classical operator-angle bound (Gustafson, 1968) gives

$$\cos \angle(J\tilde{g}, \tilde{g}) \geq \frac{2\sqrt{\lambda_{\min}\lambda_{\max}}}{\lambda_{\min} + \lambda_{\max}}, \tag{17}$$

where $\lambda_{\min}, \lambda_{\max}$ are the extreme eigenvalues of J. This lower bound is positive for positive-definite J and zero for singular nonzero J.

Guidance-Only Step. Define

$$\eta(y) = \frac{\tilde{g}^\top J\tilde{g}}{\|\tilde{g}\|^2}, \qquad \left\langle \nabla F(y), \frac{\tilde{g}}{\|\tilde{g}\|} \right\rangle = \eta(y)\|\tilde{g}\|. \tag{18}$$

Assume ∇F is L_F-Lipschitz, with $L_F > 0$, on the segment from y to $y + u$, where $u = \alpha\tilde{g}/\|\tilde{g}\|$ and $\alpha \geq 0$. The smoothness bound (Bottou et al., 2018, Eq. (4.3)), applied to $-F$, yields

$$F(y + u) - F(y) \geq \alpha\eta\|\tilde{g}\| - \frac{L_F}{2}\alpha^2. \tag{19}$$

Consequently, a sufficient condition for strict increase is

$$0 < \alpha < \frac{2\eta\|\tilde{g}\|}{L_F}. \tag{20}$$

This requires $\eta > 0$ along the chosen direction, rather than a uniform positive lower bound on every eigenvalue of J.

Binder-Coordinate Restriction. We use the same fixed binder-coordinate mask P as in eq. (7). For $\mathsf{P}\tilde{g} \neq 0$, define the applied unit direction and its directional derivative by

$$d_{\mathsf{P}} = \frac{\mathsf{P}\tilde{g}}{\|\mathsf{P}\tilde{g}\|}, \qquad a_{\mathsf{P}}(y) = \langle J\tilde{g}, d_{\mathsf{P}} \rangle. \tag{21}$$

Unlike the unrestricted derivative, a_{P} can be negative. Partitioning the gradient and Jacobian into binder ($\mathcal{B}$) and target ($\mathcal{T}$) coordinates gives

$$\langle J\tilde{g}, \mathsf{P}\tilde{g} \rangle = \tilde{g}_{\mathcal{B}}^{\top} J_{\mathcal{B}\mathcal{B}} \tilde{g}_{\mathcal{B}} + \tilde{g}_{\mathcal{T}}^{\top} J_{\mathcal{T}\mathcal{B}} \tilde{g}_{\mathcal{B}}. \tag{22}$$

By Cauchy–Schwarz, a sufficient condition for $a_{\mathsf{P}} \geq 0$ is

$$\|J_{\mathcal{T}\mathcal{B}} \tilde{g}_{\mathcal{B}}\| \|\tilde{g}_{\mathcal{T}}\| \leq \tilde{g}_{\mathcal{B}}^{\top} J_{\mathcal{B}\mathcal{B}} \tilde{g}_{\mathcal{B}}.$$

Nonnegativity for every gradient holds exactly when $\mathsf{P}J = J\mathsf{P}$: the symmetric part of $\mathsf{P}J$ can be positive semidefinite only if its binder–target block vanishes. This stronger condition holds when conditioning fixes target coordinates exactly, but is not assumed here.

Base Displacement and Normalization. When guidance is active, eq. (8) sets $u_t = \alpha d_{\mathsf{P}}$ with $\alpha = \rho \|\Delta_t\|$. Keep the same noise level and conditioning in F, and assume ∇F is L_F-Lipschitz, with $L_F > 0$, on a neighborhood containing both segments $y \to y + \Delta_t$ and $y + \Delta_t \to y + \Delta_t + \alpha d_{\mathsf{P}}$. Applying the descent lemma at $y + \Delta_t$ and bounding the change in ∇F from y gives

$$F(y + \Delta_t + \alpha d_{\mathsf{P}}) - F(y + \Delta_t) \geq \alpha \left(a_{\mathsf{P}} - L_F \|\Delta_t\| \right) - \frac{L_F}{2} \alpha^2. \tag{23}$$

Thus, strict improvement over the base update is guaranteed when

$$0 < \alpha < \frac{2 \left(a_{\mathsf{P}} - L_F \|\Delta_t\| \right)}{L_F}.$$

For $\rho > 0$ and $\|\Delta_t\| > 0$, this is equivalent to $a_{\mathsf{P}} > L_F(1 + \rho/2)\|\Delta_t\|$. The base displacement therefore tightens the guidance-only bound $0 < \alpha < 2a_{\mathsf{P}}/L_F$ along the same projected direction.

The implementation tests liveness after masking and the configured positive rescaling, which defines an effective threshold on $\|\mathsf{P}\tilde{g}\|$. Inactive steps have $u_t = 0$. For active steps, positive reward scaling multiplies both a_{P} and L_F by the same factor, leaving the sufficient step-size window unchanged. Across states, that window depends on a_{P}, L_F, and $\|\Delta_t\|$; the processed-gradient norm alone does not determine it.

C Structural proxy

The proxy is a scale-normalized weighted sum of coordinate-based rewards, oriented so that larger values are preferred:

$$\widetilde{r}(x) = \sum_j \omega_j \frac{R_j(x)}{q_j}, \qquad q_j > 0. \tag{24}$$

Weights $\omega_j \geq 0$ and reference scales q_j are fixed by the configuration; a zero weight disables a term. These component scales differ from the subsequent normalization of the combined gradient. Let $\mathcal{B}$ be the binder atoms, $\mathcal{A}_{\mathrm{tar}}$ the target atoms, $\mathcal{A} = \mathcal{B} \cup \mathcal{A}_{\mathrm{tar}}$, $\mathcal{B}_\alpha$ the binder C_α atoms, and $\mathcal{H}_{\mathrm{hot}} \subseteq \mathcal{A}_{\mathrm{tar}}$ the hotspot atoms.

Clash Avoidance. Let $m_{ij} = 0$ for self-pairs and exempt near-neighbor pairs along the binder chain, and $m_{ij} = 1$ otherwise. The clash reward is

$$R_{\mathrm{clash}}(x) = -\sum_{i \in \mathcal{B}} \sum_{j \in \mathcal{A}} m_{ij} \left[d_0 - \|x_i - x_j\| \right]_+^2, \tag{25}$$

where $[v]_+ = \max(v, 0)$ and d_0 is the minimum-distance threshold. Binder–binder pairs appear in both orders, while binder–target pairs appear once.

Compactness. For binder radius of gyration $R_g(x)$, a length-dependent reference R_g^{ref}, and tolerance $\tau_{R_g} > 0$, define

$$R_{\mathrm{rg}}(x) = -\left(\frac{R_g(x) - R_g^{\mathrm{ref}}}{\tau_{R_g}}\right)^2. \tag{26}$$

The penalty discourages both excessive collapse and excessive extension.

Interface Contacts. Define a soft contact by

$$c_{ij}(d, s) = \mathrm{sigmoid}\left(\frac{d - \|x_i - x_j\|}{s}\right), \qquad s > 0. \tag{27}$$

The configured burial term encourages interface formation while softening the reward for further contacts. With hotspot weights v_j, positive saturation scales s_1, s_2, and $\phi_s(v) = s\tanh(v/s)$, it takes the form

$$R_{\mathrm{burial}}(x) = \phi_{s_2}\left[\frac{1}{|\mathcal{B}_\alpha|}\sum_{i\in\mathcal{B}_\alpha}\phi_{s_1}\left(\sum_{j\in\mathcal{A}_{\mathrm{tar}}} v_j c_{ij}(d_b, s_b)\right)\right]. \tag{28}$$

Here v_j is the configured hotspot multiplier on hotspot atoms and one otherwise.

Hotspot Proximity. The hotspot term measures the fraction of interface contacts involving hotspot atoms:

$$R_{\mathrm{hotspot}}(x) = \frac{\displaystyle\sum_{i\in\mathcal{B}_\alpha}\sum_{j\in\mathcal{H}_{\mathrm{hot}}} c_{ij}(d_h, s_h)}{\max\left\{\displaystyle\sum_{i\in\mathcal{B}_\alpha}\sum_{j\in\mathcal{A}_{\mathrm{tar}}} c_{ij}(d_h, s_h),\ 10^{-6}\right\}}. \tag{29}$$

Using the same nonnegative contact function in numerator and denominator ensures $R_{\mathrm{hotspot}} \in [0, 1]$.

D Experiment Detail

D.1 Dataset

ProtDBench. ProtDBench (Liu et al., 2026) defines ten binder-design tasks, each specifying a receptor structure, crop, hotspot residues, and binder-length range. We obtain receptor structures from the Protein Data Bank (Berman et al., 2000) and convert them to the generator's input format. Table 4 lists all ten task specifications.

BindCraft. To evaluate performance on additional targets, we construct an eight-target panel from BindCraft (Pacesa et al., 2025): PD1, PD-L1, IFNAR2, CLDN1, DerF7, DerF21, BetV1, and CD45. We retain targets with residue numbering compatible with our input conversion and specified hotspots, and fix binder length to 80 residues. PD-L1 overlaps with ProtDBench and serves as a positive control; we therefore report pooled results both including and excluding it. DerF7, DerF21, BetV1, and CD45 each have one hotspot residue, and all eight targets use target templates. Results are reported in Tables 7 and 8.

D.2 Evaluation Metrics Details

We group the evaluation metrics into confidence, energy, design quality, and diversity and novelty.

Confidence. AF2 ipTM and Boltz-2 ipTM measure predicted confidence in the receptor–binder interface using AlphaFold2 and Boltz-2, respectively (Jumper et al., 2021; Passaro et al., 2025). AF2 ipTM guides selection, while Boltz-2 ipTM is used only for cross-verifier evaluation. pLDDT measures local structural confidence, averaged over binder residues from the AF2 prediction and reported on a 0–1 scale. Higher ipTM and pLDDT values indicate greater predicted interface and local structural confidence, respectively.

Table 4: The ProtDBench binder-design targets (Liu et al., 2026), converted to the generator's input format. Chains use mmCIF-assigned identifiers, and residue indices are 1-based sequential positions rather than author numbering. Comma-separated intervals indicate discontinuous crops. Each receptor chain occupies one row; dashes indicate no hotspots.

Target	PDB	Chain	Crop	Hotspots	Binder	Panel range
BHRF1	2wh6	A	15–171	78, 87, 90, 95, 98, 106	100	80–120
SC2RBD	6m0j	B	15–208	167, 171, 176, 182, 187	100	80–120
IL7RA	3di3	B	21–213	62, 84, 143	80	50–120
PDL1	5o45	A	1–116	40, 99, 107	80	50–120
TrkA	1www	C	1–101	13, 15, 52	80	50–120
Insulin	4zxb	E	6–155	64, 88, 96	80	40–120
VEGFA	1bj1	C	7–100	—	90	50–140
		D	7–100	74, 76, 84		
IL17A	4hsa	A	7–19, 31–121	84, 106	90	50–140
		B	9–117	57		
H1	5vli	A	1–45, 47–50, 76–80, 107–111, 258–322	—	80	40–120
		B	1–68, 80–170	21, 45, 52		
TNFa	1tnf	A	12–157	113	80	50–120
		B	12–157	—		
		C	12–157	73		

Table 5: BindCraft targets used in the additional evaluation panel (Pacesa et al., 2025). All targets use the full supplied chain A, sequential 1-based residue numbering, target templates, and an 80-residue binder. The final column gives the original BindCraft binder-length range.

Target	Crop	Hotspots	Length range
PD1	1–115	33, 95, 97, 102	80–150
PDL1	1–115	37, 39, 49, 98	64–155
IFNAR2	1–103	45, 73, 75, 77, 89, 91	60–175
CLDN1	1–188	31, 46, 55, 152	80–175
DerF7	1–196	115	70–185
DerF21	1–118	10	70–185
BetV1	1–159	24	70–185
CD45	1–88	1	80–200

Energy. Rosetta interface energy measures the energy difference between the complex and its separated partners under the Rosetta all-atom energy function (Alford et al., 2017), reported in Rosetta energy units (REU). Both states are repacked with backbones fixed. Lower values indicate more favourable interface energies. The calculation uses the generated glycine-placeholder binder backbone and therefore does not assess the designed sequence.

Design quality. ESM-2 PLL measures the plausibility of the ProteinMPNN-designed binder sequence (Dauparas et al., 2022) under ESM-2 650M (Lin et al., 2023). It averages the log-probability of each observed residue when that position is masked; higher, less negative values indicate more plausible sequences. *Success rate* is the percentage of designs passing all thresholds within a specified criterion. AF2-easy requires binder pLDDT > 0.80, ipTM > 0.50, scaled interface PAE < 0.35, and bound–unbound binder C_α RMSD < 3.5 Å. AF2-optimistic requires binder pLDDT > 0.90, unscaled interface PAE < 7.0, and monomer-to-design RMSD < 1.5 Å. Protenix-mini requires binder-restricted ipTM > 0.85, binder pTM > 0.88, and predicted-to-design RMSD < 2.5 Å. Higher success rates are preferred. The pooled column uses AF2-easy, with per-target rates reported to show variation across targets. AF2-based criteria share the selector's model family.

Diversity and novelty. *Structural cluster count* measures the number of single-linkage clusters among generated binder backbones, using a 6 Å cutoff on pairwise C_α RMSD after Kabsch superposition (Kabsch, 1976). *Unique fraction* is the cluster count divided by the number of designs. Higher

values indicate greater structural diversity when comparing equally sized sets. *Mean maximum reference similarity* averages each binder backbone's highest query-normalized TM-score against a staged PDB reference database (Berman et al., 2000), using Foldseek in TM-align mode (Van Kempen et al., 2024; Zhang & Skolnick, 2005). Lower values indicate greater structural novelty relative to that database. Diversity and novelty describe the generated population and should be interpreted alongside design quality.

D.3 BASELINES

- − **Rollout-guided selection** (`rnd`): Selects parents uniformly from the current generation while preserving prefix reuse, continuation sampling, normalized proxy guidance, and final AF2 selection. This isolates the contribution of score-based parent selection.
- − **Generational search** (`bon`): Generates 64 candidates in independently seeded batches of full guided trajectories, without prefix reuse or generational parent selection. It retains the best eight by AF2 ipTM, matching the candidate budget and retention rule of PRES-NPG but not generation cost.
- **Stock sampling**: Uses the frozen PXDesign-d model's native sampling procedure (Team et al., 2025) without added guidance or resampling. Retained-set results use AF2-based top-eight selection.
- **Guidance only**: Adds normalized proxy guidance to the frozen model without resampling or generational search. These separately collected results provide context rather than a paired guidance ablation.
- **Best-of-N** ($N = 128$): Generates 128 candidates with stock sampling and ranks them by AF2 ipTM. Retained-set and pool results average the best eight and all 128 candidates, respectively. Its larger, unguided pool distinguishes it from the guided best-of-64 control.
- **RFdiffusion**: An external diffusion model for target-conditioned backbone generation, followed by sequence design (Watson et al., 2023). This comparison evaluates the complete design pipeline.
- **BoltzGen**: An external all-atom diffusion model combining protein design and structure prediction (Stark et al., 2025). It provides a complete-system reference, but no numerical result is reported in the present table.
- **Search-only sampling**: Applies score-based particle resampling at intermediate checkpoints without gradient guidance. Unlike PRES, it selects using intermediate scores rather than completed-rollout evaluations of saved prefixes.
- **Random selection**: Uniformly selects completed designs from a stock-generated pool. Its expected subset mean equals the full-pool mean. This conceptual control is not separately evaluated and differs from `rnd`, which randomizes parents but retains final AF2 ranking.
- **Stock sampling, Complexa**: Uses native Proteina-Complexa sampling (Didi et al., 2026) without PRES-NPG steering, providing a reference from a second design framework. Comparisons require common evaluation procedures and overlapping tasks.

D.4 ABLATION SETTINGS

We evaluate ablations on ten ProtDBench targets with three seeds each. The full method, random-parent control, and independent guided control form the matched comparison. Guidance variants share the targets, seeds, and frozen nominal dose but use different batching; their results are descriptive levels without paired contrasts against PRES-NPG.

- **PRES-NPG (full)**: Combines normalized proxy guidance with prefix-reuse evolutionary search. Across four generations of 16 candidates, completed-candidate AF2 scores select four parents whose saved intermediate states seed fresh continuations. Final selection operates on the complete 64-candidate archive.
- − **Rollout-guided selection** (`rnd`): Selects parents uniformly from the current generation while preserving prefix reuse, continuation sampling, normalized proxy guidance, and final AF2 selection. This isolates the contribution of score-based parent selection.

Table 6: Generator-specific settings and shared search configuration. Guidance windows and branch positions follow each sampler's native schedule. Guided-step counts refer to full trajectories.

	PXDesign-d (Team et al., 2025)	**Proteina-Complexa** (Didi et al., 2026)
Checkpoint	Released PXDesign-d checkpoint (v0.1.0)	Released Proteina-Complexa checkpoint
Sampler	400 denoising steps; diffusion chunk size 16; original sampler with an added guidance hook	Product-space flow sampler, 400 steps; branch positions mapped through the coordinate schedule (native coordinate-schedule value 0.70 corresponds to step 103)
Guidance (NPG)	$\rho = 0.04$, window $[0.70, 0.85]$; 60 guided steps	$\rho = 0.07$, window $[0.05, 0.85]$; 155 guided steps; term scales fixed from prior calibration
Selection score	ipTM from AlphaFold2 with an initial guess (AF2-IG)	AlphaFold2-Multimer ipTM after ProteinMPNN sequence redesign
Shared configuration		
Arms	PRES, random-parent branching, and independent best-of-64; each scores 64 candidates per run	
Search (PRES)	$M = 4$ parents and $N = 4$ children per parent, giving $MN = 16$ candidates per generation, with $G = 3$ branch positions at $(0.70, 0.80, 0.90)$	
Archive and delivery	$B = 64$ archived candidates; $K = 8$ retained designs	

- $-$ **Generational search** (`bon`): Generates 64 candidates in four independently seeded batches of full guided trajectories, without prefix reuse or generational parent selection. It matches the candidate budget and retention rule of PRES-NPG, but not generation cost.
- $-$ **Guidance** ($N = 128$): Uses stock sampling without added guidance, resampling, or generational search, then retains the best eight of 128 candidates by AF2 i_{pTM}. Because both the search procedure and candidate budget differ, this row provides unguided reference levels rather than an isolated guidance ablation.
- **No hotspot term**: Removes the explicit hotspot reward and the additional hotspot weighting inside the burial reward. Clash avoidance, compactness, and generic interface burial remain active. This examines hotspot-specific guidance while preserving the generator's conditioning.
- **Burial-only reward**: Retains only the generic binder–target contact reward, disabling the clash, compactness, and hotspot terms. This examines whether interface burial alone provides an effective guidance signal.
- **Allocation**: Redistributes guidance strength across candidates using a reward-dependent Pearson-χ^2 allocation rule while preserving their proxy-gradient directions. This examines how allocating the steering budget across candidates affects generation.
- **Geometry**: Preconditions the proxy gradient using a graph of binder backbone connectivity, then normalizes the resulting direction to the prescribed Euclidean displacement magnitude. This examines guidance directions that favor coordinated molecular motion.
- **Hybrid**: Combines normalized proxy guidance with intermediate particle resampling based on structural-stability scores. It includes both interventions within generation, without the outer PRES search loop.
- **Multi-objective**: Replaces the fixed weighted-sum gradient with an adaptive convex combination of individual objective gradients using the multiple-gradient descent algorithm. This seeks a common improvement direction across objectives rather than a fixed trade-off.
- **Random direction**: Replaces the proxy-gradient direction with an isotropic random direction on binder coordinates while preserving the corresponding displacement magnitude and activation conditions. This controls for perturbation magnitude when examining the contribution of the guidance direction.
- **Particle resampling**: Resamples particles at intermediate checkpoints using changes in structural-stability scores, without gradient guidance. Unlike PRES, its internal selection uses intermediate trajectory information rather than completed-candidate AF2 evaluations.

Table 7: **AF2 interface confidence on the BindCraft panel.** Eight targets, three arms, and two seeds yield 48 chains of 64 designs. We report mean AF2 ipTM for the AF2-selected top eight and all 64 designs. `rnd` randomizes parent selection; `bon` removes generational search. Target-level results average two seeds; pooled results weight targets equally. Bracketed values are 95% confidence intervals for paired arm differences from 20,000 bootstrap draws over eight target clusters. The panel includes bcPDL1, which overlaps with the main benchmark.

	AF2 ipTM			Arm contrast	
Target	PRES-NPG (`es`)	w/o score-based parent selection (`rnd`)	w/o generational search (`bon`)	`es` − `rnd`	`es` − `bon`
AF2-selected top 8 of 64					
bcBetV1	0.6113	0.4587	0.3700	+0.1525	+0.2413
bcCD45	0.3263	0.2344	0.1944	+0.0919	+0.1319
bcCLDN1	0.5719	0.6631	0.5575	−0.0913	+0.0144
bcDerF21	0.8600	0.8319	0.8512	+0.0281	+0.0087
bcDerF7	0.8825	0.8806	0.8819	+0.0019	+0.0006
bcIFNAR2	0.8644	0.8431	0.8431	+0.0212	+0.0212
bcPD1	0.7706	0.7525	0.7562	+0.0181	+0.0144
bcPDL1	0.8356	0.8237	0.8287	+0.0119	+0.0069
All 8 targets	0.7153	0.6860	0.6604	+0.0293	+0.0549
		`es`−`rnd`: [−0.0145, +0.0762]		`es`−`bon`: [+0.0093, +0.1149]	
All 64 designs					
bcBetV1	0.1957	0.1644	0.1536	+0.0313	+0.0421
bcCD45	0.1245	0.1014	0.0917	+0.0231	+0.0328
bcCLDN1	0.2351	0.2270	0.2049	+0.0081	+0.0302
bcDerF21	0.6033	0.4877	0.5141	+0.1156	+0.0892
bcDerF7	0.6798	0.5438	0.5770	+0.1360	+0.1029
bcIFNAR2	0.5917	0.4690	0.4461	+0.1227	+0.1456
bcPD1	0.4099	0.3215	0.3216	+0.0884	+0.0883
bcPDL1	0.5036	0.4465	0.4208	+0.0571	+0.0828
All 8 targets	0.4179	0.3452	0.3412	+0.0728	+0.0767
		`es`−`rnd`: [+0.0410, +0.1053]		`es`−`bon`: [+0.0524, +0.1036]	

- **Smooth-min**: Replaces the weighted-sum reward with a smooth minimum of calibrated objective margins, emphasizing the least-satisfied criterion.
- **Reversed reward**: Negates all active guidance-reward weights while preserving the nominal guidance dose and final AF2 ranking. This negative control steers against the original proxy to examine whether its orientation affects candidate quality.

D.5 Hyperparameters

Table 6 summarizes the fixed configurations. Both campaigns use the same search parameters and a budget of 64 scored candidates per run, with eight designs retained. Guidance settings are generator-specific. Selection uses interface predicted template modelling (ipTM) scores from the respective evaluation pipelines.

E Additional Results

We further evaluate PRES-NPG on eight BindCraft targets, comparing it with random parent selection and independent generation. Table 7 reports AF2 interface confidence and paired arm differences, including a separate summary excluding the overlapping bcPDL1 control. Table 8 complements this analysis with predictor scores, Rosetta interface energy, structural diversity, and reference similarity. Both tables distinguish the AF2-selected top eight designs from the full pool of 64, separating delivered-set quality from overall generation performance.

F Sensitivity to the selection score

We examine how the scoring rule affects final candidate retention and parent selection in PRES-NPG using fixed candidate archives and random-parent search runs. We compare the AF2-IG rule used by PRES-NPG with Protenix-mini and Boltz-2, including a reanalysis that corrects the target-complex inputs supplied to Boltz-2. As shown in table 9, the correction improves the retained-set AF2 score by 0.0480 (95% CI $[0.0058, 0.0971]$), while the change in the retrospective offspring contrast remains uncertain. Paired comparisons with Protenix-mini do not resolve a difference in either

Table 8: **Predictor scores and structural diversity on the BindCraft panel.** Results average target means across all eight targets in table 7. The upper block reports the AF2-selected top eight; the lower block reports all 64 designs. Protenix-mini reports held-out interface confidence (ipTM). Rosetta interface energy is the mean across targets of seed-level medians after constrained FastRelax; three designs with positive post-relaxation total energy are excluded. Lower energy is preferable. Clusters use a 6 Å Cα RMSD threshold; uniqueness is 100 times the cluster count divided by the number of designs. Compare these metrics within each block. Reference similarity averages each design's maximum Foldseek `qtmscore` against the PDB reference; lower values indicate greater structural novelty. * denotes a paired difference whose target-clustered 95% bootstrap interval excludes zero (20,000 draws).

Arm / paired difference	AF2 i_{pTM}	Rosetta ΔG_{sep}	pLDDT	AF2 pass %	Protenix-mini	Clust. 6 Å	Uniq. %	Ref. sim.
(a) AF2-selected top 8 of 64								
PRES-NPG (`es`)	**0.7153**	−27.70	**0.9287**	67.97	**0.4335**	2.25	28.12	0.7614
w/o score-based parent selection (`rnd`)	0.6860	−26.98	0.9243	**72.66**	0.4199	3.94	49.22	0.7545
w/o generational search (`bon`)	0.6604	**−28.67**	0.9223	64.06	0.4330	**6.38**	**79.69**	0.7643
`es` − `rnd`	+0.0293	−0.72	+0.0044	−4.69	+0.0136	−1.69*	−21.09*	+0.0069
`es` − `bon`	+0.0549*	+0.97	+0.0063	+3.91	+0.0005	−4.12*	−51.56*	−0.0029
(b) All 64 designs								
PRES-NPG (`es`)	**0.4180**	**−26.64**	**0.9154**	**35.84**	**0.3705**	11.44	17.87	0.7661
w/o score-based parent selection (`rnd`)	0.3451	−25.65	0.9130	27.73	0.3466	11.19	17.48	0.7652
w/o generational search (`bon`)	0.3412	−25.52	0.9125	26.07	0.3446	**27.56**	**43.07**	0.7725
`es` − `rnd`	+0.0728	−0.99	+0.0024	+8.11	+0.0240	+0.25	+0.39	+0.0009
`es` − `bon`	+0.0767	−1.11	+0.0029	+9.77	+0.0259	−16.12	−25.20	−0.0064

setting. Table 10 shows higher AF2 confidence after the correction, with small observed differences in Rosetta energy and similar retained-set cluster counts. These results assess the selection component of PRES-NPG; they do not establish the performance of a full search run with an alternative scorer or experimental binding activity.

Table 9: **Selection performance under alternative scoring rules.** All outcomes are evaluated using AF2 i_{pTM} on the same nine targets. In (a), Δ is the mean score of the selected eight designs minus the 64-design pool mean. In (b), Δ compares offspring of would-be top-four parents with offspring of other parents in random-parent runs. Paired differences compare target means between rules. Intervals are 95% target-clustered bootstrap CIs from 20,000 draws.

Selection rule / paired difference	AF2 Δi_{pTM}	95% CI	Relative effect κ / $\Delta\kappa$	Negative targets
(a) Final retention: eight of 64 designs, 52 archives				
AF2-IG (PRES-NPG reference)	+0.3965	[+0.3283, +0.4598]	1.000	0/9
Protenix-mini	+0.1209	[+0.0640, +0.1765]	0.305	1/9
Boltz-2 (original input)	+0.0801	[+0.0386, +0.1228]	0.202	0/9
Boltz-2 (corrected input)	+0.1281	[+0.0787, +0.1764]	0.323	0/9
Corrected − original Boltz-2	+0.0480	[+0.0058, +0.0971]	$\Delta\kappa$ +0.121	2/9
Corrected Boltz-2 − Protenix-mini	+0.0072	[−0.0190, +0.0357]	$\Delta\kappa$ +0.018	5/9
(b) Offspring response: retrospective analysis of 52 random-parent runs				
AF2-IG (PRES-NPG reference)	+0.1180	[+0.0780, +0.1589]	1.000	0/9
Protenix-mini	+0.0600	[+0.0155, +0.1116]	0.509	1/9
Boltz-2 (original input)	+0.0458	[+0.0034, +0.0932]	0.388	3/9
Boltz-2 (corrected input)	+0.0749	[+0.0342, +0.1176]	0.635	1/9
Corrected − original Boltz-2	+0.0292	[−0.0164, +0.0790]	$\Delta\kappa$ +0.247	3/9
Corrected Boltz-2 − Protenix-mini	+0.0149	[−0.0161, +0.0504]	$\Delta\kappa$ +0.126	4/9

G Analysis of compute budget

PRES-NPG is test-time search over a frozen generator, so its cost is the cost of running that generator and its scorers, not of training anything. Two quantities are worth separating, and the paper's claims depend on which one is meant. *Scored candidates* are matched across the three registered arms by

Table 10: **Predicted quality and structural diversity after final selection.** Each rule retains eight of 64 designs from the same 52 archives across nine targets. Every metric is recomputed for the designs selected by that rule. Values are means over targets. AF2 pass rate is the percentage with $i_{\mathrm{pTM}} > 0.55$; clusters use a Cα RMSD threshold of 6 Å. Arrows indicate preferred directions for quality metrics.

Selection rule	AF2 $i_{\mathrm{pTM}}\uparrow$	Rosetta $\Delta G_{\mathrm{sep}}\downarrow$	pLDDT↑	$i_{\mathrm{pAE}}\downarrow$	AF2 pass (%)↑	Clusters
Retained eight designs						
AF2-IG (reference)	0.7298	−29.85	0.9084	0.2805	84.0	6.94
Protenix-mini	0.4543	−28.21	0.9134	0.5314	46.4	7.19
Boltz-2 (original input)	0.4135	−29.45	0.9083	0.5681	40.2	6.96
Boltz-2 (corrected input)	0.4614	−29.72	0.9114	0.5244	48.1	7.00
No selection (64 designs)	0.3333	−28.40	0.9043	0.6377	27.6	36.69

construction: every chain of every arm produces and scores exactly 64 designs. *Generation work* is not matched, and PRES-NPG is the cheapest of the three: prefix reuse means a continuation traverses only the suffix below its branch position, so a 64-design archive costs 1.6 full 16-candidate batches against independent best-of-64's 4.0. Table 11 gives the budget of each row in the main table; Table 12 decomposes what that budget actually cost on the clock.

The Step Ratio Is Not the Speedup. These two tables disagree by design, and the disagreement is the finding. Counting denoiser steps, PRES-NPG does 2.50× less generation work than its best-of-64 control (10,240 candidate-steps against 25,600). Measuring the clock, a PRES-NPG chain finishes only 1.20× faster (2,369 s against 2,848 s median). The step-count model is not wrong about the part it describes — the generation phase alone separates by 2.49× (345.6 s against 861.5 s), which reproduces the 2.50× prediction almost exactly — but generation is 15% of a PRES-NPG chain and 32% of a best-of-64 chain. The matched verifier cost dilutes the rest. Any speedup quoted from Table 11 alone overstates the realized saving by roughly a factor of two.

The Verifier Does Not Dominate by Orders of Magnitude. The accounting this campaign inherited assumed structure prediction would outweigh generation by two to three orders of magnitude, and budgeted on that basis. The campaign's own logs refute it. Scoring costs 3.83× generation for PRES-NPG and 1.49× for best-of-64 — and 1.49× is the relevant number for whether search is affordable, because an arm that generates more pays that cost against a fixed verifier bill. The revision matters beyond bookkeeping: under the inherited assumption, prefix reuse would have been invisible in end-to-end time, and the 1.20× we measure would have been 1.00×.

Table 12: **Where a chain's wall-clock actually goes.** Median seconds per task–seed cell over the 40 chains of each registered arm, from the phase timestamps of the 480 generation logs and the per-chain ledgers. All entries are *Meas.* The verifier columns are matched by construction — every arm scores the same 64 designs with the same instruments — so the arms separate only in the generation row, where they separate by 2.49×, against the 2.50× the step count predicts. End to end that becomes 1.20×, because generation is 15% of a PRES-NPG chain and 32% of a best-of-64 chain. Component medians are taken independently and do not sum to the chain median.

Phase (median s per cell)	**PRES-NPG**	`rnd`	`bon`
Generation, 4 invocations	345.6	344.4	861.5
AF2 selector stage, 128 predictions	894.6	871.9	854.2
Protenix-mini scoring, 64 designs	429.8	431.1	426.3
Protenix-base target re-fold, 4 calls[a]	435.6	438.2	436.3
Process start and checkpoint load, 4 invocations	180.3	179.7	183.0
Chain total	2,369	2,377	2,848
Generation share of the chain	15%	15%	32%
Verifier-to-generation ratio[b]	3.83×	3.78×	1.49×

Table 11: **Compute budget per task–seed cell, by row of Table 1.** The upper block is the three registered arms, which are matched on scored candidates and on nothing else; the lower block is the four rows that are levels rather than contrasts, and share with the upper block only the scoring instrument and the retention depth $K = 8$. *No arm of any campaign was compute-matched to any other*, and none is claimed to be. Denoiser steps are per scored candidate; batch equivalents count full 400-step trajectories of the 16-candidate batch that is the sampler's actual unit, so 1.6 means a branching chain does 1.6 batches' worth of sampling to fill a 64-design archive. GPU-hours are reported only where the jobs ran one-per-GPU and are therefore comparable; the lower block ran under heavy oversubscription and its recorded hours are given in note f rather than printed here, because they are not commensurable with the upper block's. Provenance by column: scored candidates and AF2 predictions *Meas.*; denoiser steps, batch equivalents and guided steps *Der.* from the per-generation step counts recorded in every run; median cell and GPU-hours *Meas.* from the run ledgers.

Model + sampler	Scored cand.	Steps per cand.	Batch equiv.	Guided steps	AF2 pred.	Median cell (s)	Campaign GPU-h
Registered arms, 40 paired cells each, one chain per GPU							
PRES-NPG	64	160	1.6	141/640	128	2,369	27.55
− Rollout-guided sel. (`rnd`)	64	160	1.6	141/640	128	2,377	27.28
− Generational search (`bon`)	64	400	4.0	240/1,600	128	2,848	32.82
Levels, not contrasts: different cells, seeds and campaigns							
Stock sampling	32	400	2.0	0	64		
Guidance only	32	400	2.0		64		
Best-of-N (N=128)	128	400	8.0		256		
RFdiffusion	192	50		0			

All results are in silico and require in vitro validation. Moreover, AF2-IG is used for parent selection, final retention, and a primary reported endpoint, so part of the observed gain may reflect optimization of the selector. Cross-verifier and Rosetta analyses provide complementary evidence but cannot eliminate shared modeling biases or establish experimental binding.

Prefix reuse concentrates search around a small number of high-scoring lineages and may reduce diversity in the delivered set. The guidance analysis is local and relies on idealized denoiser assumptions, while the empirical comparisons match scored candidates rather than end-to-end compute. The observed quality and speed gains may therefore not transfer to other generators, scoring pipelines, or hardware.

Future work could estimate prefix continuation values from multiple descendants or learned value models, adapt the NPG strength and guidance window to the current sampling state, and choose branch locations and child allocations dynamically. Diversity-aware or multi-objective parent selection could further balance exploitation of promising prefixes with exploration of distinct structural modes.