Diamond Maps for Protein Binder Design: Inference-Time Scaling Survives Stochastic Flow Map Distillation

Anonymous Authors[1]

Abstract

Conditional protein generative models rely on inference-time scaling techniques such as Sequential Monte Carlo and reward-guided sampling to achieve competitive binder design performance. These techniques estimate a value function over candidate trajectories, which requires many evaluations of expensive pretrained models. Flow map distillation could relieve this bottleneck, but standard flow maps are deterministic and yield biased value-function estimates, collapsing the inference-time scaling that makes generative methods competitive with hallucination in the first place. We apply Posterior Diamond Maps to Proteina-Complexa, distilling its joint latent and Cα flow into a single-step stochastic sampler of the posterior $p_{1|t}(z \mid x_t)$ that admits consistent, low-cost value-function estimation. We pair the distilled model with Sequential Monte Carlo using AlphaFold-Multimer interface confidence as the reward, and use the PD-L1 binder design problem as proof of concept. On a 5K-step student, deterministic higher-order integration of the teacher saturates well short of FK-steering, while a (P, K) search-width sweep with the distilled SMC monotonically improves every tail metric tracked. These results are consistent with stochastic flow map distillation preserving the inference-time scaling behavior of the underlying generative model.

1. Introduction

De novo protein binder design has split into two paradigms. Hallucination methods such as BindCraft (Pacesa et al., 2025) and BoltzDesign1 (Cho et al., 2025) optimize a binder sequence against a structure prediction oracle. Generative methods such as RFDiffusion (Watson et al., 2023), Genie 3 (Lin et al., 2026), and Proteina-Complexa (Didi et al., 2026) learn explicit distributions over conditional structures. Generative methods compete with and increasingly surpass hallucination when paired with inference-time scaling where the underlying distribution is sampled many times under a reward such as AlphaFold-Multimer interface confidence, with Sequential Monte Carlo (SMC) or beam search concentrating compute on promising candidates.

The compute cost of this approach is dominated by the structure prediction oracle, not the generative model. AlphaFold-Multimer with five models and twenty recycles requires 200–300 seconds per evaluation, while a single generation step takes roughly 10 seconds. Speeding up the generator pays off only insofar as the savings are reinvested into more search rollouts under the oracle.

Flow map distillation (Boffi et al., 2025) compresses a multi-step probability flow ODE into a single network evaluation, with W_2 guarantees relative to the teacher. Naive application to a conditional protein generator would compress out a key capability. Inference-time scaling techniques require estimates of the value function

$$V_t(x_t) = \log \mathbb{E}_{z \sim p_{1|t}(\cdot|x_t)}\big[\exp\big(r(z)\big)\big], \tag{1}$$

which in turn requires sampling from the posterior $p_{1|t}(z \mid x_t)$. Standard flow maps are deterministic: given x_t, the predicted endpoint is fixed, and value function estimation collapses to a single-point denoiser approximation that is known to be biased (Chung et al., 2023).

Diamond Maps (Holderrieth et al., 2026a), built on GLASS Flows (Holderrieth et al., 2026b), are stochastic flow maps that produce posterior samples in a single forward pass. Posterior Diamond Maps in particular are distilled to sample $z \sim p_{1|t}(\cdot \mid x_t)$ given $(\bar{x}_0, x_t, t)$ for $\bar{x}_0 \sim \mathcal{N}(0, I)$, where the stochasticity in $\bar{x}_0$ traverses the posterior. This admits efficient and consistent value function estimation, restoring the inference-time scaling that direct flow map distillation collapses.

We adapt the above framework to distill Proteina-Complexa, the first application of stochastic flow maps to a protein generative model. Our contributions are:

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Submitted to the 2026 Workshop on Generative and Agentic AI for Biology (ICML 2026). Do not distribute.

- A Posterior Diamond Map distillation of Proteina-Complexa operating in the joint latent and Cα space, with an extension to the GLASS reparameterization
- An SMC inference procedure using AlphaFold-Multimer interface predicted aligned error as the reward, evaluated under the AF2M benchmark protocol.
- Empirical evidence on PD-L1 that (i) the GLASS reparameterization on the real teacher exhibits a non-zero Jensen gap of 0.117–0.162 relative L_2 between the deterministic denoiser and the GLASS posterior mean, (ii) deterministic higher-order integration (Heun) plateaus well below stochastic search at matched compute, and (iii) PDM-SMC tail metrics improve monotonically with search width even on a 5K-step student, surpassing single-pass at NFE $= 400$ on best reward.

2. Background

2.1. Stochastic interpolants and flow matching

Following Albergo et al. (2023), the stochastic interpolant

$$I_t = \alpha_t z + \sigma_t \epsilon, \quad z \sim p_{\text{data}},\ \epsilon \sim \mathcal{N}(0, I_d), \tag{2}$$

with $\alpha_t, \sigma_t \in C^1([0,1])$, $\alpha_0 = \sigma_1 = 0$, $\alpha_1 = \sigma_0 = 1$, defines a probability path $p_t(x) = \text{Law}(I_t)$ between $\mathcal{N}(0, I_d)$ and p_{data}. The marginal velocity field

$$u_t(x) = \mathbb{E}[\dot{I}_t \mid I_t = x] \tag{3}$$

generates this path: integrating $\dot{x}_t = u_t(x_t)$ from $x_0 \sim \mathcal{N}(0, I)$ produces $x_1 \sim p_{\text{data}}$ (Lipman et al., 2023). The denoiser $D_t(x) = \mathbb{E}[z \mid I_t = x]$ is recoverable from u_t via

$$D_t(x) = \frac{\sigma_t u_t(x) - \dot{\sigma}_t x}{\dot{\alpha}_t \sigma_t - \alpha_t \dot{\sigma}_t}. \tag{4}$$

2.2. Flow maps

The flow map $X_{s,t} : \mathbb{R}^d \to \mathbb{R}^d$ for the ODE $\dot{x}_t = u_t(x_t)$ satisfies $X_{s,t}(x_s) = x_t$ along trajectories (Boffi et al., 2025). Equivalently it is the unique solution of the Lagrangian PDE

$$\partial_t X_{s,t}(x) = u_t(X_{s,t}(x)), \quad X_{s,s}(x) = x. \tag{5}$$

A neural flow map $\hat{X}^\theta_{s,t}$ can be distilled from a pretrained u_t via the Lagrangian distillation loss

$$\mathcal{L}_{\text{LMD}}(\theta) = \mathbb{E}\Big[\big\|\partial_t \hat{X}^\theta_{s,t}(I_s) - u_t\big(\hat{X}^\theta_{s,t}(I_s)\big)\big\|^2\Big], \tag{6}$$

which controls the Wasserstein distance between teacher and student endpoint distributions (Boffi et al., 2025). The Lagrangian formulation outperforms its Eulerian counterpart in practice when the optimal map is near-singular, as is typical in protein structure space.

2.3. Inference-time scaling and the role of stochasticity

Reward-aligned sampling targets the tilted distribution $p_r(z) \propto p_{\text{data}}(z)\exp(r(z))$. Sequential Monte Carlo evolves P particles along the probability path while reweighting by $\exp(V_t(x_t))$ (Singhal et al., 2025; Holderrieth et al., 2026a). Search and guidance similarly rely on V_t or its gradient. Estimating V_t accurately requires samples $z^k \sim p_{1|t}(\cdot \mid x_t)$, since

$$\hat{V}_t(x_t) = \log \tfrac{1}{K} \textstyle\sum_{k=1}^K \exp(r(z^k)) \tag{7}$$

is consistent only when z^k are drawn from the true posterior.

Three options for posterior sampling exist. Running the time-reversal SDE backward from t to 1 is unbiased but requires many simulation steps. Using the denoiser as a point estimate, $V_t(x_t) \approx r(D_t(x_t))$, is biased: the Jensen gap grows with the noise level (Chung et al., 2023). GLASS Flows (Holderrieth et al., 2026b) construct an inner ODE whose endpoint at inner time $s = 1$ is a sample from $p_{1|t}(\cdot \mid x_t)$, recovering the SDE marginal at lower cost. Diamond Maps (Holderrieth et al., 2026a) distill this inner flow into a single-step stochastic flow map. We adopt the third route.

3. Method

Figure 1 offers a sketch of the pipeline. The pretrained teacher is the Proteina-Complexa flow $u_t(x \mid c)$, where c encodes target structure, hotspot residues, and binder length, and $x = (z^{\text{lat}}, z^{C\alpha}) \in \mathcal{Z} = \mathbb{R}^{N\times d_z} \times \mathbb{R}^{N\times 3}$ is the joint partial-autoencoder latent and explicit Cα trace. The autoencoder decoder $\mathcal{D} : \mathcal{Z} \to \text{StructureSpace}$ maps z to all-atom coordinates.

3.1. GLASS reparameterization in joint latent space

For fixed (x_t, t), the posterior $p_{1|t}(z \mid x_t)$ admits an inner stochastic interpolant with state $\bar{x}_s$ and time s, transporting $\bar{x}_0 \sim \mathcal{N}(0, I)$ to $\bar{x}_1 \sim p_{1|t}(\cdot \mid x_t)$. The corresponding inner velocity field $\bar{u}_s(\bar{x}_s \mid x_t, t, c)$ admits a closed form as a linear function of the pretrained denoiser evaluated at a transformed input (Holderrieth et al., 2026a). Define the sufficient statistic

$$S_{s,t}(\bar{x}_s, x_t) = \frac{\alpha_s \sigma_t^2\, \bar{x}_s + \alpha_t \sigma_s^2\, x_t}{\sigma_t^2 \alpha_s^2 + \alpha_t^2 \sigma_s^2}, \tag{8}$$

the time reparameterization

$$t^*(s,t) = g^{-1}\!\left(\frac{\sigma_t^2 \sigma_s^2}{\sigma_t^2 \alpha_s^2 + \alpha_t^2 \sigma_s^2}\right), \quad g(t) = \frac{\sigma_t^2}{\alpha_t^2}, \tag{9}$$

and the inner velocity

$$\bar{u}_s(\bar{x}_s \mid x_t, t, c) = a_{s,t}\, \bar{x}_s + b_{s,t}\, D_{t^*}\big(\alpha_{t^*} S_{s,t}(\bar{x}_s, x_t) \,\big|\, c\big), \tag{10}$$

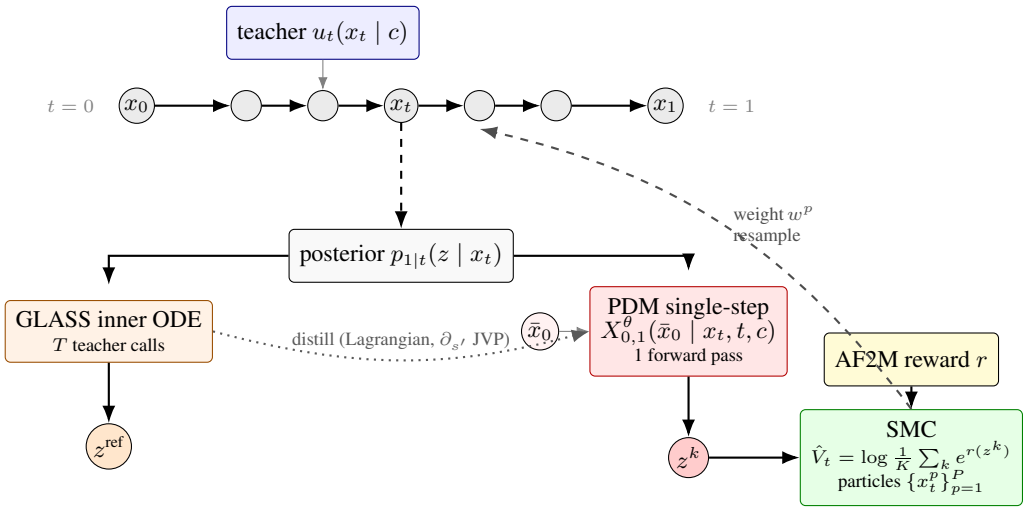

Figure 1. **Method overview.** The pretrained Proteina-Complexa teacher integrates the outer probability path u_t (top, blue). At any intermediate x_t, the posterior $p_{1|t}(\cdot \mid x_t)$ is what value-function estimation needs to query. *Two routes produce posterior samples.* Integrating the GLASS inner ODE on the teacher (orange) yields a reference posterior sample at the cost of T teacher forward passes per sample. The Posterior Diamond Map student (red) collapses this to a single forward pass via Lagrangian distillation against the GLASS inner velocity. Multiple cheap student samples drive the SMC value-function estimator (green), whose reweighting and resampling feed back to the outer particle ensemble.

where $a_{s,t}, b_{s,t}$ are scalar weights determined by the schedulers (Holderrieth et al., 2026a, App. A.4) and $D_t(\cdot \mid c)$ is the conditional denoiser obtained from the Proteina-Complexa velocity field via Equation (4). Equation (10) requires no retraining: $\bar{u}_s$ is a closed-form function of the frozen teacher.

Heterogeneous schedulers. Proteina-Complexa applies different scheduler scalings to the latent and Cα components. The GLASS construction extends componentwise: Equations (8) to (10) are computed independently for each component, with the denoiser query in Equation (10) dispatched as a single forward pass through the joint network at the (concatenated, heterogeneous) reparameterized time.

3.2. Posterior Diamond Map

The Posterior Diamond Map $X^\theta_{s,s'}(\bar{x}_s \mid x_t, t, c)$ is a flow map for the inner GLASS flow. We use the Boffi ansatz to enforce the boundary condition architecturally:

$$X^\theta_{s,s'}(\bar{x}_s \mid x_t, t, c) = \bar{x}_s + (s' - s)\, v^\theta_{s,s'}(\bar{x}_s, x_t, t, c). \quad (11)$$

The neural network v^θ reuses the Proteina-Complexa backbone: the time embedding is extended to three scalars (s, s', t), and the input projection accepts the concatenation of $\bar{x}_s$ and x_t. Weights are initialized from the teacher, and the conditioning encoder is frozen for the first 5K training steps to stabilize the new pathways before disturbing c-handling.

The student is trained against the GLASS velocity via the Lagrangian loss

$$\mathcal{L}(\theta) = \mathbb{E}\Big[\big\|\partial_{s'} X^\theta_{s,s'}(\bar{x}_s \mid x_t, t, c) - \mathrm{sg}\big(\bar{u}_{s'}(X^\theta_{s,s'}(\bar{x}_s \mid x_t, t, c) \mid x_t, t, c)\big)\big\|^2\Big], \quad (12)$$

with the expectation taken over c, $z \sim p_{\text{data}}(\cdot \mid c)$, $t \sim \mathrm{Uniform}[\epsilon, 1-\epsilon]$, $x_t \sim p_t(\cdot \mid z)$, $(s, s') \sim w_{s,s'}$, and $\bar{x}_s \sim p_s(\cdot \mid x_t)$. The partial derivative $\partial_{s'}$ is computed by forward-mode automatic differentiation through $v^\theta_{s,s'}$, comparable in cost to a single forward pass.

Training schedule. Following Boffi et al. (2025), training uses a strip-then-full schedule on (s, s'): a strip stage restricts $w_{s,s'} = \mathbb{1}[|s' - s| \le 0.1]$ (training a $K = 10$-step model), and a subsequent full stage extends to $w_{s,s'} = 1$ on $[0, 1]^2$. Direct one-step training without the strip warm-up is unstable in our experiments, consistent with the original observation on checkerboard data. Step counts and other hyperparameters are listed in Section A.

3.3. Sampling

Single-step sampling from the posterior at any (x_t, t, c) is

$$z^k = X^\theta_{0,1}(\bar{x}_0^k \mid x_t, t, c), \quad \bar{x}_0^k \sim \mathcal{N}(0, I_d). \quad (13)$$

The Monte Carlo estimator $\hat{V}_t(x_t)$ from Equation (7) requires K such samples and K reward evaluations.

3.4. SMC inference for binder design

Algorithm 1 SMC binder design with Posterior Diamond Map

Input: student $X^{\theta}_{0,1}$, teacher u_t, particles P, outer steps N, MC samples K, conditioning c, reward r
Initialize $\{x_0^p\}_{p=1}^P \sim \mathcal{N}(0, I)$, $\log w^p \leftarrow 0$
for $n = 0, \dots, N-1$ **do**
  $t \leftarrow n/N$
  if $n > 0$ **then**
    $\{p_{1|0} = p_{\text{data}}$ is independent of x_0, so the value-function update at $n = 0$ is uninformative$\}$
    for $p = 1, \dots, P$ **do**
      for $k = 1, \dots, K$ **do**
        $\bar{x}_0^{p,k} \sim \mathcal{N}(0, I)$
        $z^{p,k} \leftarrow X^{\theta}_{0,1}(\bar{x}_0^{p,k} \mid x_t^p, t, c)$
        $r^{p,k} \leftarrow r(\mathcal{D}(z^{p,k}), c)$
      end for
      $\hat{V}_t(x_t^p) \leftarrow \log \frac{1}{K} \sum_k \exp(r^{p,k})$
      $\log w^p \leftarrow \log w^p + \hat{V}_t(x_t^p)$
    end for
    $\text{ESS} \leftarrow \left(\sum_p \exp(\log w^p)\right)^2 / \sum_p \exp(2 \log w^p)$
    if $\text{ESS} < P/2$ **then**
      Systematically resample by $\text{softmax}(\log w)$; reset $\log w^p \leftarrow 0$
    end if
  end if
  $x^p_{t+1/N} \leftarrow x_t^p + \frac{1}{N} u_t(x_t^p \mid c)$ for all p
end for
Return: weighted samples $\{(x_1^p, w^p)\}$

Algorithm 1 plugs the Posterior Diamond Map into Sequential Monte Carlo over the Proteina-Complexa probability path. Particles advance along the deterministic outer ODE; weights accumulate value function estimates at each step; effective sample size triggers resampling.

4. Preliminary validation

We report a case study on PD-L1, a representative monomeric binder target from Zambaldi et al. (2024). We test three claims: a non-zero Jensen gap on the real teacher, a plateau of deterministic higher-order integration, and search-width scaling on tail metrics under the distilled SMC.

The student is initialized from the teacher checkpoint with a frozen backbone and 133K trainable adapter parameters and trained for 5K steps under the strip schedule of Section 3 ($\sim$7 GPU-h on 4×A100); full hyperparameters are in Section A. Reward is $r = -\text{ipAE}_{\text{interface}}$ from AlphaFold-Multimer with three recycles, where "pass" requires complex pTM > 0.8, minimum interface pAE < 1.5 Å, and complex self-consistency RMSD < 2.5 Å.

We report three main findings, more details in Section B:

(1) Jensen gap on the real teacher. The GLASS reparameterization (Equation (10)) requires no retraining and, on 20 different (x_t, c) initializations on the PD-L1 target at $t = 0.5$, with $K = 16$ posterior draws and $T = 50$ inner Euler steps, produces a posterior whose mean differs from the deterministic teacher denoiser $D_t(x_t \mid c)$ by relative L_2 of 0.117 (bb_ca) and 0.162 (latent), Table 2. This non-zero Jensen gap is critical for the inference-time-scaling argument where stochastic posterior estimation captures a signal the point estimate cannot.

(2) Deterministic integration plateaus. Heun's method on the teacher at matched NFE $\in \{32, 100, 400\}$ yields -0.758, -0.756, -0.751 mean reward and never crosses -0.7, while FK-steering (Singhal et al., 2025) at NFE $= 400$ achieves -0.366 mean / 0.000 best and single-pass Euler at the same NFE achieves -0.480 mean. Better integration accuracy alone does not close the gap to stochastic search (Table 1 and Figure 2).

(3) PDM-SMC tail metrics scale with search width. Sweeping $(P, K) \in \{(4, 1), (8, 2), (16, 4)\}$ on the distilled student, best reward improves $-0.257 \rightarrow -0.203 \rightarrow -0.186$ and the count of AF2M passes (under the reduced-recycle reward of Section A) climbs $1 \rightarrow 3 \rightarrow 7$ out of $\{16, 32, 64\}$ samples, while mean reward stays flat at ~ -0.55 (Table 1). At $(P, K) = (16, 4)$ the best reward already exceeds single-pass Euler at NFE $= 400$. This monotone improvement on the tail is the inference-time-scaling property the paper claims survives distillation; absolute counts rather than percentages are reported because the sample sizes are too small for stable rate estimates.

5. Discussion

Our result reframes a question that has shadowed protein generative modeling: does flow map distillation help or hurt? Direct distillation accelerates each forward pass at the cost of the inference-time scaling that makes generative methods competitive in the first place. Stochastic flow map distillation breaks this tradeoff. The compute saved by the faster generator is reinvested into more search rollouts under the structure prediction oracle, and the value function estimates that drive the search remain unbiased.

The framing also speaks to a broader observation that scaling laws in language modeling have not transferred cleanly to biological generative models. Inference-time scaling has been a partial substitute, but it requires expensive structure prediction oracles in the loop. Methods that make inference-time scaling cheaper, without compromising its statistical foundations, may be a more productive direction than continued scaling of base models that increasingly bottleneck on data quality and architectural inductive biases.

Impact Statement

This paper presents work whose goal is to advance the field of computational protein design. We identify no immediate ethical concerns specific to the methodological contribution. Generative models for protein binder design have dual-use potential, and the benchmark protocol used here is consistent with community norms for in silico evaluation prior to wet-lab validation. We do not release model weights or sequences for designs targeting human pathogens or human proteins of biosecurity concern.

References

Albergo, M. S., Boffi, N. M., and Vanden-Eijnden, E. Stochastic interpolants: A unifying framework for flows and diffusions. In *International Conference on Learning Representations*, 2023.

Boffi, N. M., Albergo, M. S., and Vanden-Eijnden, E. How to build a consistency model: Learning flow maps via self-distillation. In *Advances in Neural Information Processing Systems*, 2025.

Cho, Y., Pacesa, M., Zhang, Z., Correia, B. E., and Ovchinnikov, S. BoltzDesign1: Inverting all-atom structure prediction for generalizable biomolecular design. *bioRxiv*, 2025. doi: 10.1101/2025.04.06.647261. preprint.

Chung, H., Kim, J., Mccann, M. T., Klasky, M. L., and Ye, J. C. Diffusion posterior sampling for general noisy inverse problems. In *International Conference on Learning Representations*, 2023.

Didi, K., Zhang, Z., Zhou, G., Reidenbach, D., Cao, Z., Cha, S., Geffner, T., Dallago, C., Tang, J., Bronstein, M. M., Steinegger, M., Kucukbenli, E., Vahdat, A., and Kreis, K. Scaling atomistic protein binder design with generative pretraining and test-time compute. In *International Conference on Learning Representations*, 2026. URL https://research.nvidia.com/labs/genair/proteina-complexa/.

Holderrieth, P., Chen, D., Eyring, L., Shah, I., Anantharaman, G., He, Y., Akata, Z., Jaakkola, T., Boffi, N. M., and Simchowitz, M. Diamond maps: Efficient reward alignment via stochastic flow maps, 2026a. URL https://arxiv.org/abs/2602.05993.

Holderrieth, P., Singer, U., Jaakkola, T., Chen, R. T. Q., Lipman, Y., and Karrer, B. Glass flows: Transition sampling for alignment of flow and diffusion models, 2026b. URL https://arxiv.org/abs/2509.25170.

Lewis, S., Hempel, T., Jiménez-Luna, J., Gastegger, M., Xie, Y., Foong, A. Y. K., Satorras, V. G., Abdin, O., Veeling, B. S., Zaporozhets, I., Chen, Y., Yang, S., et al. Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science*, 2025. doi: 10.1126/science.adv9817.

Lin, Y., Lee, M., Vermani, A., Jiang, E., De Cooman, S., Spetko, M., and AlQuraishi, M. Fast and ultra-capable protein design: Advancing the frontier through atomistic se(3)-equivariance with genie 3. *bioRxiv*, 2026. doi: 10.64898/2026.05.01.722168. URL https://www.biorxiv.org/content/early/2026/05/05/2026.05.01.722168.

Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nickel, M., and Le, M. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

Pacesa, M., Nickel, L., Schellhaas, C., Schmidt, J., Pyatova, E., Kissling, L., Barendse, P., Choudhury, J., Kapoor, S., Alcaraz-Serna, A., Cho, Y., Ghamary, K. H., Vinué, L., Yachnin, B. J., Wollacott, A. M., Buckley, S., Westphal, A. H., Lindhoud, S., Georgeon, S., Goverde, C. A., Hatzopoulos, G. N., Gönczy, P., Muller, Y. D., Schwank, G., Swarts, D. C., Vecchio, A. J., Schneider, B. L., Ovchinnikov, S., and Correia, B. E. One-shot design of functional protein binders with bindcraft. *Nature*, 646(8084):483–492, Oct 2025. ISSN 1476-4687. doi: 10.1038/s41586-025-09429-6. URL https://doi.org/10.1038/s41586-025-09429-6.

Singhal, R., Horvitz, Z., Teehan, R., Ren, M., Yu, Z., McKeown, K., and Ranganath, R. A general framework for inference-time scaling and steering of diffusion models. *arXiv preprint arXiv:2501.06848*, 2025.

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., et al. De novo design of protein structure and function with RFdiffusion. *Nature*, 620(7976):1089–1100, 2023. doi: 10.1038/s41586-023-06415-8.

Zambaldi, V., La, D., Chu, A. E., Patani, H., Danson, A. E., Kwan, T. O. C., Frerix, T., Schneider, R. G., Saxton, D., Thillaisundaram, A., Wu, Z., Moraes, I., Lange, O., Papa, E., Stanton, G., Martin, V., Singh, S., Wong, L. H., Bates, R., Kohl, S. A., Abramson, J., Senior, A. W., Alguel, Y., Wu, M. Y., Aspalter, I. M., Bentley, K., Bauer, D. L. V., Cherepanov, P., Hassabis, D., Kohli, P., Fergus, R., and Wang, J. De novo design of high-affinity protein binders with alphaproteo, 2024. URL https://arxiv.org/abs/2409.08022.

A. Implementation details

A.1. Network architecture

The Posterior Diamond Map student $v^{\theta}_{s,s'}(\bar{x}_s, x_t, t, c)$ reuses the Proteina-Complexa denoising backbone. Modifications relative to the teacher: (i) the time embedding is replaced by an MLP $[3, 256, 512]$ taking the concatenation of (s, s', t) as input; (ii) the input projection accepts $2d$ channels (concatenation of $\bar{x}_s$ and x_t) instead of d. All other parameters are initialized from the teacher checkpoint.

A.2. Training hyperparameters

Optimizer	AdamW, $\beta = (0.9, 0.999)$, weight decay 0
Learning rate	1×10^{-4}
Batch size	32 (8 per GPU $\times$ 4 GPUs)
Mixed precision	bfloat16
EMA decay (sampling)	0.9999
Gradient clip norm	1.0
Outer / inner time clamp ϵ	10^{-3}
Cond. encoder freeze	first 5,000 steps
Strip schedule	$w_{s,s'} = \mathbb{1}[\lvert s' - s\rvert \leq 0.1]$
Training duration	5,000 steps, $\sim$ 7 GPU-h on 4 $\times$A100 (DDP)

Direct one-step training without the strip warm-up is unstable in our experiments, consistent with Boffi et al. (2025)'s observation on synthetic data.

A.3. Reward computation

We use ColabFold with AlphaFold-Multimer models. "Pass" designs require binder pTM > 0.8, minimum interface pAE < 1.5 Å, complex RMSD < 2.5 Å, and $\geq 80\%$ of hotspot residues within 5 Å of the binder.

For the SMC reward we use one AF2M model with three recycles and initial-guess templating from the design; this keeps per-step cost tractable so the search-width sweep stays inside the methods-paper compute budget. The reduced-recycle reward is a noisier proxy of the canonical multi-model AF2M score and the SMC numbers in this paper should be interpreted under that proxy rather than as direct benchmark numbers.

B. Additional verification

B.1. Integration-baseline calibration

Table 1 reports the calibration runs referenced in Section 4. All runs use the same teacher checkpoint and the PD-L1 target with hotspots and binder length matched to Zambaldi et al. (2024). The single-pass rows sweep Euler step count $\mathrm{NFE} \in \{8, 32, 100, 400\}$ at fixed sampler with $n = 8$ draws per NFE. The FK-steering reference uses Proteina-Complexa's beam search at $\mathrm{NFE} = 400$. Reward is $-\mathrm{ipAE}_{\text{interface}}$ from AlphaFold-Multimer under the AF2M protocol, so larger is better.

The Heun rows use deterministic predictor-corrector integration on the teacher's outer ODE (no stochastic noise injection, no self-conditioning). Single-pass Euler retains both, which contributes to the gap at high NFE; Thus deterministic higher-order integration cannot substitute for stochastic posterior estimation.

The three PDM-SMC rows use the 5K-step student described in Section A and sweep $(P, K) \in \{(4, 1), (8, 2), (16, 4)\}$. Doubling search width at each step leaves the mean reward unchanged at ~ -0.55 but improves the best-reward tail monotonically ($-0.257 \to -0.203 \to -0.186$), surpassing single-pass at $\mathrm{NFE} = 400$ (-0.229) by $P = 8$. Absolute counts of designs passing the reduced-recycle AF2M check climb $1 \to 3 \to 7$ out of $\{16, 32, 64\}$ samples; we report counts rather than rates because the sample sizes are too small for the rate estimates to be stable, and because the reduced-recycle reward used during SMC (Section A) is a noisier proxy of canonical AF2M scoring. The mean-reward gap to FK-steering (-0.366) on PD-L1 quantifies the headroom available for a fully-trained student under the same SMC procedure.

Table 1. Integration-baseline calibration on PD-L1. Single-pass Euler and deterministic Heun on the teacher at varying NFE, against the same teacher with FK-steering at NFE $= 400$, and a PDM-SMC sweep with the distilled student. Heun's method beats single-pass Euler at NFE $= 32$ but plateaus at ~ -0.75 for all higher NFE, trailing FK-steering at NFE $= 400$ by 0.39 mean. PDM-SMC at minimal configuration ($P = 4$, $K = 1$) already matches single-pass-400 on the best-reward tail metric.

Run	Mean reward	Best reward	n
FK-steering, NFE $= 400$	-0.366	0.000	8704
Single-pass, NFE $= 400$	-0.480	-0.229	8
Single-pass, NFE $= 100$	-0.664	-0.346	8
Single-pass, NFE $= 32$	-0.846	-0.818	8
Single-pass, NFE $= 8$	-0.912	-0.852	8
Heun, NFE $= 400$	-0.751	-0.692	8
Heun, NFE $= 100$	-0.756	-0.677	8
Heun, NFE $= 32$	-0.758	-0.691	8
PDM-SMC, P=4, K=1	-0.548	-0.257	16
PDM-SMC, P=8, K=2	-0.555	-0.203	32
PDM-SMC, P=16, K=4	-0.546	-0.186	64

B.2. GLASS reparameterization verification on the teacher

The verification suite consumes a teacher checkpoint and an optional student checkpoint and reports (i) marginal preservation of $X^{\theta}_{0,1}$ against a T-step GLASS Flow ODE integration of the teacher, (ii) moment drift between the student's posterior samples and the teacher's, and (iii) ProteinMPNN $\rightarrow$ ESMFold self-consistency RMSD distributions for both.

Table 2. GLASS marginal preservation on the real Proteina-Complexa teacher (PD-L1, 20 different (x_t, c) initializations, $K = 16$ posterior draws per initialization, $T = 50$ inner Euler steps, fixed outer time $t = 0.5$). "Jensen gap" is the relative L_2 between the deterministic teacher denoiser $D_t(x_t \mid c)$ and the GLASS posterior mean over K draws.

Modality	GLASS std	GLASS $\|\cdot\|_2$	Jensen gap	GLASS-vs-data z
bb_ca	1.12	77.4	0.117	0.275
local_latents	1.13	126.3	0.162	0.573

The GLASS endpoint distribution at $s = 1$ has unit-scale moments (Table 2) consistent with sampling the conditional outer endpoint $p_{1|t}(\cdot \mid x_t)$. The Jensen gap of 0.117 (bb_ca) and 0.162 (latent) quantifies the bias incurred by replacing a true posterior expectation $\mathbb{E}_{z\sim p_{1|t}}[r(z)]$ with the single-point denoiser estimate $r(D_t(x_t))$: even at $t = 0.5$ on a fine-tuned teacher, the gap is $\sim$ 12–16% of the posterior-mean magnitude. This non-zero Jensen gap is the entry point for the inference-time-scaling argument: stochastic posterior estimation captures a signal that point estimates do not.

The GLASS-vs-data column reports the relative L_2 between a posterior draw and the original ground-truth z used to construct x_t. The latent gap (0.573) substantially exceeds the bb_ca gap (0.275), consistent with the latent space encoding higher-frequency information (sidechain conformations) whose posterior support is broader than the Cα trace at the same outer noise level.

B.3. PDM-SMC search-width scaling on PD-L1

Figure 2 plots the data of Table 1 on a single best-reward axis to make the qualitative argument visual: deterministic baselines (Heun, single-pass Euler) plateau or saturate as samples grow, while PDM-SMC tracks the FK-steering reference along the per-sample efficiency frontier and continues to improve with search width.

B.4. Pilot student posterior fidelity

The 5K-step student does not yet match the teacher's GLASS posterior pointwise (`rel_mean_l2` $\approx$ 0.81–0.99 at 4 conditions $\times$ $K = 4$), consistent with the small trainable parameter count (133K) relative to the joint state dimensionality. The search-width scaling reported in Table 1 therefore reflects the inference-time-scaling property of the framework rather than a fully-converged student: tail-metric improvement under (P, K) growth holds even when single-sample posterior

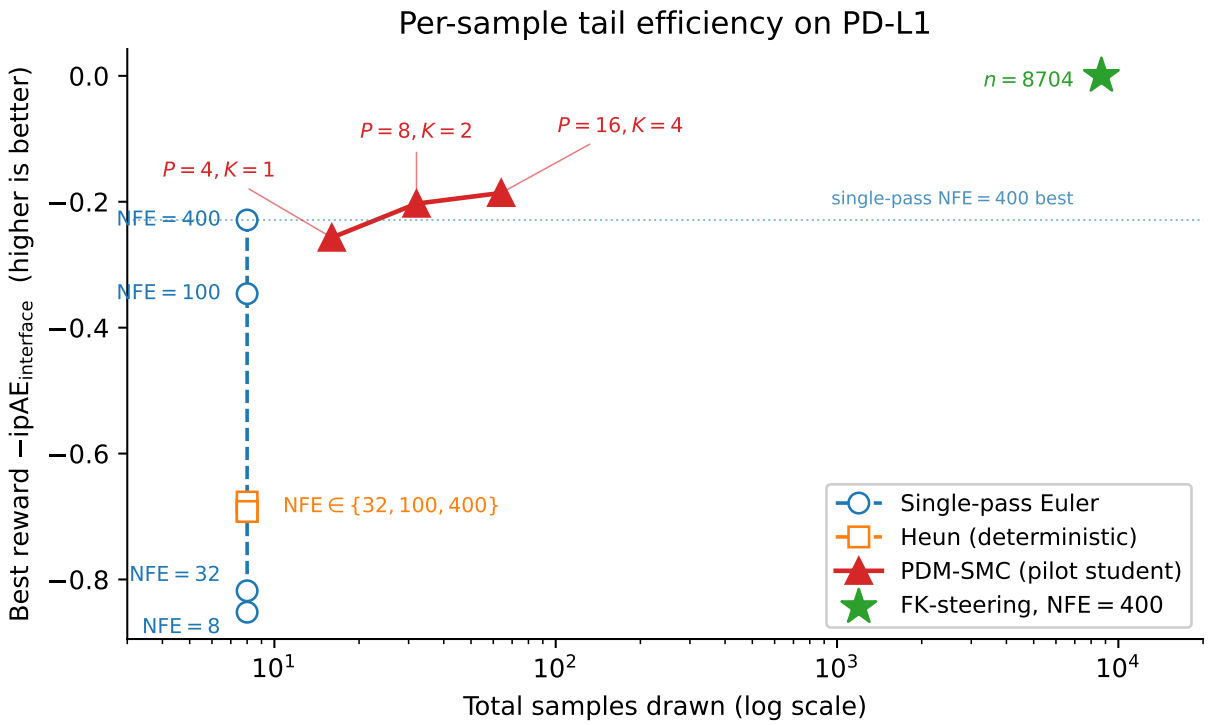

Figure 2. **Per-sample tail efficiency on PD-L1.** Best $-$ipAE reward as a function of total samples drawn (x-axis, log scale). Single-pass Euler at NFE $= 400$ peaks at -0.229 on 8 samples; deterministic Heun saturates near -0.69 at all NFE. PDM-SMC on the 5K-step student improves monotonically with (P, K) (red triangles), surpassing single-pass at $(P, K) = (8, 2)$ and reaching -0.186 at $(P, K) = (16, 4)$. FK-steering on the teacher requires $\sim$8700 samples to reach 0.

fidelity is partial.

C. Limitations

The reward, AlphaFold-Multimer interface pAE, is an in silico oracle: "passing" designs are those the oracle scores well, not necessarily ones that bind in vitro. The results should be read as evidence that stochastic flow map distillation *does not collapse* the inference-time scaling property of Proteina-Complexa rather than as a blanket claim that the distilled model advances the experimental frontier of binder design. Additionally, our Heun baseline is deterministic and disables self-conditioning, while single-pass Euler in our reference implementation retains both stochastic noise injection and self-conditioning; this contributes to the gap at high NFE.

D. Future work

Two natural directions extend the present work. First, a multi-target study under a longer-trained student would test the scaling claim across binder design problems of varying difficulty (monomeric / multimeric, conformationally rigid / flexible). Wet-lab validation in the manner of the Adaptyv Bio competitions would be the experimental next step beyond the in silico setting. Second, the Posterior Diamond Map framework extends naturally to other reward signals where SMC is currently expensive: motif scaffolding under multi-state RMSD constraints (Watson et al., 2023), multistate design via composite objectives, and binder design against unknown conformations using ensemble-aware predictors (Lewis et al., 2025). The latent space of partial autoencoders is already exploited for compression; combining stochastic flow maps with conformational sampling in the same latent space is a natural next step.