DuET-MD: Training-Free Rare-Event Steering for Frozen Molecular Dynamics Surrogates

Anonymous authors
Paper under double-blind review

Abstract

Protein dynamics is central to understanding biological function, while atomistic molecular dynamics simulation, the gold-standard tool for characterizing protein dynamics, is too costly to observe rare transitions across high free-energy barriers. Recent MD surrogates accelerate trajectory generation by learning ps-to-ns finite-lag transitions, but faster propagation alone does not solve rare-event sampling: their default rollouts are concentrated around high-probability transitions learned from reference trajectories, leaving rare-event sampling unresolved. Applying conventional enhanced-sampling biases to frozen generative transition models is also non-trivial. We introduce DuET-MD, a training-free inference-time framework for steering frozen iterative MD surrogates. DuET-MD couples two particle processes across distinct clocks. Within each next-frame generation, inner particles use guidance and intermediate resampling to discover promising transitions before frame completion. Across physical time, an outer particle process uses the selected frame together with an estimated local normalizer to properly weight and resample competing trajectory prefixes. This enables path conditioning without retraining the surrogate. Across protein-folding and cryptic-pocket-opening tasks, DuET-MD increases the recovery of rare-transition candidates over default surrogate rollouts under matched finite sampling budgets. We further evaluate geometric validity and mechanistic ordering not encoded in the steering objective. DuET-MD thus extends accelerated MD surrogates from fast rollout to training-free rare-transition sampling.

1 Introduction

Molecular dynamics (MD) simulation generates atomistic trajectories by propagating molecular systems under physics-based force fields, providing a fundamental tool for characterizing protein motions underlying biological function (McCammon et al., 1977). Its practical timescale is nevertheless constrained by two fundamental difficulties. First, conventional atomistic MD typically requires 1–2 fs integration time steps to resolve fast atomic motions and maintain numerical stability (Hopkins et al., 2015). Second, many biologically meaningful processes, including conformational switching, loop rearrangement, cryptic-pocket opening, and ligand binding or unbinding, require transitions across high free-energy barriers and therefore occur rarely (Lindorff-Larsen et al., 2011; Bemelmans et al., 2026; Dror et al., 2011; Izrailev et al., 1997). Consequently, observing such events may require an impractically large number of integration steps.

Recent MD-surrogate models seek to address the first limitation by learning finite-lag molecular transitions from simulation trajectories. Instead of resolving every femtosecond-scale integration step, these models generate next conformations over substantially longer time intervals, enabling trajectory rollouts without explicitly simulating every intervening integration step (Klein et al., 2023; Schreiner et al., 2023; Yu et al., 2025; Antoniadis et al., 2026; Shen et al., 2025; Yu et al., 2026).

Longer transition intervals, however, do not by themselves eliminate the rare-event problem. A surrogate trained on reference trajectories learns the transition patterns represented in the training data. Biologically relevant paths may therefore remain unlikely to emerge under default rollout, even when the surrogate propagates molecular configurations much faster than conventional MD (Cheng et al., 2026). Enhanced-sampling methods address rare-event sampling by biasing

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.

molecular exploration toward otherwise rarely visited conformational regions and transitions. Representative approaches include umbrella sampling (Torrie & Valleau, 1977), steered molecular dynamics (Izrailev et al., 1997), and metadynamics (Laio & Parrinello, 2002). Related ideas have also been introduced into diffusion-based conformational and equilibrium-ensemble generators (Richman et al., 2025; Xie et al., 2026). Unlike MD surrogates, however, these ensemble generators do not directly yield temporally ordered molecular trajectories.

Applying enhanced sampling to pretrained diffusion-based MD surrogates is non-trivial. Unlike explicit MD integrators, these models typically do not provide an underlying force, potential energy, or tractable marginal transition density (Shen et al., 2025; Yu et al., 2026), making conventional biasing schemes difficult to apply directly. We therefore seek a training-free steering strategy that retains the frozen surrogate as an MD-derived transition prior while directing sampling toward low-probability transitions that are rarely exposed by default rollout. In this way, inference-time control leverages the transition knowledge already learned from MD trajectories without retraining the backbone or learning an auxiliary guidance model.

We introduce DUET-MD, which performs particle steering across two distinct timescales (Fig. 1). At the *internal generation clock*, gradient-based guidance and particle resampling redirect computation toward promising next-frame candidates before completion. At the *physical trajectory clock*, an outer particle process uses the selected next frame and its local normalizer estimate to update and resample competing trajectory prefixes. This coupling adapts Feynman–Kac steering and the proper-weighting construction of Nested SMC (Singhal et al., 2025; Naesseth et al., 2015; Chanchu et al., 2026) to frozen diffusion-based MD surrogates, enabling training-free rare-transition steering without additional training.

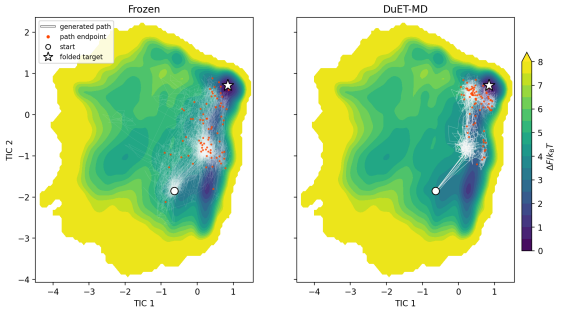


Figure 1: Representative BBA folding-directed trajectories in a fixed TICA landscape.

Our contributions are threefold:

- We formulate rare-transition steering of frozen iterative MD surrogates as a two-clock sequential inference problem. The internal generation clock controls how each next frame is proposed before completion, while the physical trajectory clock maintains competing trajectory prefixes over successive surrogate transitions. This separation exposes a proposal bottleneck that cannot be resolved by trajectory-level selection alone when useful completed proposals are rare.
- We demonstrate improved recovery of target-directed, geometrically valid transition candidates across fast-folding and cryptic-pocket steering tasks. DUET-MD consistently improves over unconditioned rollout and provides additional gains over complete-frame selection in regimes where useful next-frame proposals are difficult to obtain.
- Beyond task success, we show on cryptic-pocket systems that the generated trajectories recapitulate MD-observed conformational rearrangements that are not directly encoded in the steering objective, providing evidence that DUET-MD exploits the MD-derived transition structure learned by the frozen surrogate rather than merely satisfying the prescribed steering coordinate.

2 Related Work

Protein dynamics emulators and trajectory generation. Timewarp learns time-coarsened flow proposals, while implicit transfer operators learn stochastic transitions across temporal resolutions (Klein et al., 2023; Schreiner et al., 2023). EquiJump and UniSim use stochastic interpolants for molecular transitions (dos Santos Costa et al., 2024; Yu et al., 2025). Transferability is further studied through TITO and protein-language-model conditioning of implicit transfer operators

(Viguera Diez et al., 2026; Antoniadis et al., 2026). ConfRover combines temporal autoregression with an $\mathrm{SE}(3)$-aware diffusion decoder (Shen et al., 2025). PVB instead realizes finite-lag transitions through a variational encoder and an augmented stochastic bridge (Yu et al., 2026). MDGen treats molecular trajectories as generative objects and supports forward simulation and endpoint-conditioned transition-path generation (Jing et al., 2024). GeoTDM models geometric trajectories with equivariant spatiotemporal diffusion kernels (Han et al., 2024). Other trajectory generators combine reference and motion guidance or structure pretraining with temporal alignment (Cheng et al., 2025; Iyengar et al., 2026). These studies establish state-conditioned and trajectory-level generation. We build on that capability rather than learning a new dynamics model, and study particle control within and across successive calls to a frozen next-frame generator.

Inference-time control of biomolecular generators. ConforMix uses twisted diffusion sampling, structural filtering, and optional reweighting to expose conformations that are poorly explored by default sampling (Richman et al., 2025). Enhanced Diffusion Sampling develops biased ensemble generation and statistical reweighting for rare-state exploration and free-energy estimation (Xie et al., 2026). These approaches motivate the distinction between information represented by a learned distribution and information exposed by its default sampler. Our object is a state-conditioned physical-time path rather than an unordered conformational distribution. Dynamics-emulator control also includes learned history-aware biases (Cheng et al., 2026) and PVB's adjoint-matching finetuning (Yu et al., 2026). DuET-MD instead couples decoder-level steering with trajectory-level particle inference, without surrogate finetuning or auxiliary controller training.

Feynman–Kac steering and nested inference. Weighted-ensemble methods redistribute weighted trajectory replicas during stochastic dynamics (Zhang et al., 2010). Twisted Diffusion Sampling and general Feynman–Kac steering establish particle-based conditional generation with intermediate potentials (Wu et al., 2023; Singhal et al., 2025), while Feynman–Kac Correctors study weighting and resampling for diffusion-model compositions and guidance (Skreta et al., 2025). Nested SMC establishes the selected-sample–normalizer interface for embedding approximate inner sampling in an outer particle method (Naesseth et al., 2015). Discrete-diffusion NSMC further develops predictive-normalizer weighting for inference-time reward control (Chanchu et al., 2026). The locally adapted proposal, telescoping potentials, and proper-weighting identity used below follow this line of work.

3 Problem Formulation

3.1 Frozen State-Conditioned Dynamics Surrogate

Let $x_0 \in \mathcal{X}$ denote the fixed starting conformation and c_0 the backbone-specific initial context. Depending on the surrogate, c_0 may be determined solely by x_0 or may include additional preceding context frames. For $t \geq 1$, let $H_t = (x_0, x_1, \ldots, x_t)$ denote the generated physical trajectory prefix, and let c_{t-1} denote the model-specific context constructed from c_0 and H_{t-1}. Let $q(x_t \mid c_{t-1})$ the frozen finite-lag transition distribution of the surrogate. Repeated application of this transition induces the path distribution

$$Q(x_{1:T} \mid x_0, c_0) = \prod_{t=1}^{T} q(x_t \mid c_{t-1}), \tag{1}$$

where T is the physical trajectory horizon.

Each transition is generated through an iterative stochastic decoder,

$$Y_{t,L} \to \cdots \to Y_{t,\ell} \to Y_{t,\ell-1} \to \cdots \to Y_{t,0} \mapsto x_t. \tag{2}$$

Here, t indexes the *physical trajectory clock*, whereas ℓ indexes the *internal generation clock*.

3.2 User-Specified Steering Potential

We encode a user-specified steering condition $\mathcal{P}$ through a positive state-based potential

$$\psi_t(x_t) > 0, \tag{3}$$

which measures the compatibility of the generated conformation with the desired condition. The resulting terminal path target is

$$Q_{\mathcal{P}}(x_{1:T} \mid x_0, c_0) \propto Q(x_{1:T} \mid x_0, c_0)\, \psi_T(x_T), \tag{4}$$

with a finite, nonzero normalizing constant.

For sequential inference, positive intermediate potentials define the Feynman–Kac targets (Del Moral et al., 2006)

$$\gamma_t(H_t) = Q(x_{1:t} \mid x_0, c_0)\, \psi_t(x_t), \tag{5}$$

with $\psi_0 \equiv 1$. Intermediate potentials guide particle allocation during generation, while the terminal target remains specified by ψ_T. In our experiments, ψ_t measures either folded-state proximity or local-contact disruption.

4 METHOD

4.1 DUAL-CLOCK PARTICLE INFERENCE

DUET-MD maintains K outer particles along the *physical trajectory clock*. For each outer particle, M inner particles operate along the *internal generation clock* to steer the iterative generation of one next frame. The inner sampler returns a selected completed frame together with a local normalizer estimate, which are passed to the outer sampler through the Nested-SMC proper-weighting construction (Naesseth et al., 2015).

Without inner adaptation, a next frame is sampled directly as

$$x_t \sim q(\cdot \mid c_{t-1}), \tag{6}$$

with the standard physical-time incremental weight

$$G_t^{\text{outer}} = \frac{\psi_t(x_t)}{\psi_{t-1}(x_{t-1})}. \tag{7}$$

This Feynman–Kac update propagates compatibility information across physical time (Del Moral et al., 2006), but can only retain next frames that have already been proposed by q. Rare transitions may therefore remain inaccessible at finite sampling budgets even under strong selection.

For a parent context c, the ideal locally adapted proposal is

$$r_t(x \mid c) = \frac{q(x \mid c)\psi_t(x)}{Z_t(c)}, \tag{8}$$

where

$$Z_t(c) = \mathbb{E}_{X \sim q(\cdot \mid c)} \left[\psi_t(X)\right]. \tag{9}$$

Here, $Z_t(c)$ summarizes the one-step compatibility available from the current parent. DUET-MD approximates this adapted transition inside the frozen iterative decoder without evaluating the marginal next-frame density.

4.2 GUIDED INNER FEYNMAN–KAC SAMPLER

Within the iterative decoder, we use the predicted-clean structure to evaluate the steering potential. At selected stochastic decoder steps, the gradient of this potential biases the native score or drift toward higher compatibility, yielding a guided proposal (Chung et al., 2023; Wu et al., 2023). Let p_ℓ and g_ℓ denote the base and guided transition densities, respectively, at internal step ℓ. Because guidance changes the proposal but not the desired target, the importance correction over a decoder segment $\mathcal{S}$ is

$$\rho(\mathcal{S}) = \prod_{\ell \in \mathcal{S}} \frac{p_\ell(Y_{\ell-1} \mid Y_\ell, c)}{g_\ell(Y_{\ell-1} \mid Y_\ell, c)}. \tag{10}$$

Unguided steps have $g_\ell = p_\ell$ and therefore contribute no proposal correction. The specific guided transition depends on the surrogate architecture, while the same importance-correction principle applies in both backbones.

We first describe one intermediate checkpoint. Propagate M internal particles to checkpoint states Y_i, and let $\widehat{X}_i$ denote their predicted-clean frames. Define

$$\phi_i = \psi_t(\widehat{X}_i), \tag{11}$$

and let ρ_i denote the proposal correction accumulated from the start of the decoder to this checkpoint. The checkpoint weights are

$$u_i = \rho_i \phi_i. \tag{12}$$

The checkpoint states are resampled according to u_i and continued with independent continuation noise. For continuation i, let a_i be its checkpoint ancestor and $\bar{\rho}_i$ the proposal correction accumulated from that checkpoint to the completed frame X_i. The endpoint correction is

$$v_i = \bar{\rho}_i \frac{\psi_t(X_i)}{\phi_{a_i}}. \tag{13}$$

The potential ratio retains the intended clean-frame tilt, while ρ_i and $\bar{\rho}_i$ correct for the guided proposal.

The resulting local normalizer estimate is

$$\widehat{Z}_t = \left(\frac{1}{M}\sum_{i=1}^{M} u_i\right)\left(\frac{1}{M}\sum_{i=1}^{M} v_i\right), \tag{14}$$

and one completed frame is selected according to

$$\Pr(J = i) = \frac{v_i}{\sum_j v_j}. \tag{15}$$

The inner sampler thus returns the properly weighted pair $(X_J, \widehat{Z}_t)$ to the outer particle system.

With multiple checkpoints, the same update is applied sequentially: each incremental weight combines the proposal correction accumulated since the previous checkpoint with the ratio between the current potential and that of its selected ancestor. When every intermediate checkpoint is followed by resampling, the local normalizer estimate is the product of the mean incremental weights. The full construction and proper-weighting derivation are given in Appendix A.

4.3 Outer Trajectory Weighting and Resampling

By the Nested-SMC proper-weighting construction (Naesseth et al., 2015), the selected frame X_t and local normalizer estimate $\widehat{Z}_t$ returned by the inner sampler can be used directly in the outer particle update. Under the ideal locally adapted proposal in Eq. (8), the physical-time incremental weight is

$$G_t = \frac{Z_t(c_{t-1})}{\psi_{t-1}(x_{t-1})}. \tag{16}$$

Replacing the intractable Z_t with its inner estimate gives

$$G_t = \frac{\widehat{Z}_t}{\psi_{t-1}(x_{t-1})}. \tag{17}$$

For each outer particle, we append the selected frame, update the backbone context, and multiply its previous weight by G_t. After normalization, trajectory–context pairs are resampled when the outer resampling criterion is met, and descendants continue with independent future randomness.

5 Experiments

5.1 Experimental Setup

Frozen Backbones and Baselines. We evaluate DuET-MD with two pretrained MD surrogates, PVB (Yu et al., 2026) and ConfRover (Shen et al., 2025), keeping all backbone parameters frozen and retaining each model's native transition lag and decoder discretization. We compare against

Frozen, the unconditioned rollout of the pretrained surrogate, and **Complete Nested (CN)**, a complete-frame Nested-SMC control (Naesseth et al., 2015; Chanchu et al., 2026) that performs selection only after multiple next-frame candidates have been fully generated. **DUET-MD** instead performs gradient-based guidance and particle resampling within the iterative decoder before frame completion, together with trajectory-level weighting and resampling across physical time.

Tasks. We consider two complementary conformational steering tasks. First, folding-directed generation shows whether a frozen MD surrogate can be steered from a fixed starting conformation toward a known folded region without retraining. We measure progress in a fixed two-dimensional time-lagged independent component analysis (TICA) space (Pérez-Hernández et al., 2013), using the distance to a folded reference as the steering coordinate. The TICA representation is fixed from reference trajectories and is not refitted to generated samples. We evaluate this setting on the fast-folding proteins BBA, WW-domain, and NTL9 (Lindorff-Larsen et al., 2011).

Moreover, we perform cryptic-pocket opening tasks to evaluate local conformational steering. Following the local-contact collective variable (CV) design of SLICE (Bemelmans et al., 2026), we define task-specific CVs from predefined side-chain contacts surrounding the region of interest and steer their separation, without specifying a target holo conformation. We study PRMT5, PRMT6, and SMARCA2 as known cryptic-pocket systems, with ubiquitin as a negative control. Task-specific CVs, potentials, evaluation criteria, and generation settings are described in the following sections and Appendix B.

5.2 Protein Folding

Setup and metrics. We evaluate folding-directed trajectory generation on three fast-folding proteins, BBA, WW-domain, and NTL9 (Lindorff-Larsen et al., 2011). For each protein, we define a fixed two-dimensional time-lagged independent component analysis (TICA) projection $z(\cdot)$ (Pérez-Hernández et al., 2013) and a folded reference $x^\star$. The steering potential is

$$d(x) = \|z(x) - z(x^\star)\|_2, \qquad \log \psi_t(x_t) = -a \left(\frac{d(x_t)}{d(x_0)} \right)^2 . \tag{18}$$

The TICA transforms and folded references are fixed before generation and are not fitted to generated trajectories.

We report valid-hit fractions, final native-contact recovery Q, and whole-path geometric validity. Native-contact recovery Q measures the fraction of contacts in the folded reference that are recovered in the generated structure. A valid hit requires satisfying the protein-specific TICA criterion together with whole-path validity. A trajectory is considered valid only if every generated frame satisfies (i) non-adjacent Cα distances ≥ 1.0 Å and (ii) adjacent Cα distances ≤ 5.50 Å. These criteria assess coarse geometric integrity rather than full atomistic physical validity. Backbone-specific horizons, TICA features, hit criteria, potential coefficients, and detailed sampling and evaluation settings are provided in Appendix B.2.

Results. Table 1 shows a consistent benefit of particle-based conditioning over frozen rollout, with the benefit of pre-completion steering becoming most apparent when successful next-frame proposals are scarce. On PVB, DUET-MD provides the clearest improvement over complete-frame selection for WW-domain and a smaller but consistent gain for NTL9. BBA exhibits a different regime: frozen rollout already reaches the prescribed target with non-negligible frequency, and CN captures most of the remaining gain, leaving little additional benefit for pre-completion intervention. This pattern is consistent with the inner clock acting primarily as a proposal-discovery mechanism, with its benefit depending on the difficulty of obtaining useful completed proposals.

The same distinction is visible on ConfRover BBA, whereas raw WW-domain and NTL9 trajectories fail the coarse whole-path geometry gate and are therefore not assigned target-reaching scores. Native-contact recovery Q, although not optimized by the steering potential, generally improves over frozen rollout, providing an independent indication of increased native-like structural agreement. However, its improvement does not consistently track the gain in TICA target attainment. This highlights a limitation inherent to CV-based enhanced sampling: steering toward a selected collective variable does not necessarily recover structural features that are not represented by that

Table 1: Folding-directed trajectory generation on fast-folding proteins. Values are run-level means ± 95% CI half-widths. Best target-reaching and structural-agreement results are shown in bold.

Protein	Method	Valid hit (%) ↑	Native-contact recovery Q ↑	Path validity (%) ↑ Non-adj. Cα ≥ 1.0 Å	Adj. Cα ≤ 5.50 Å	Joint
(a) PVB						
BBA	Frozen	28.28 ± 2.04	0.3364 ± 0.0031	100.00 ± 0.00	99.84 ± 0.35	99.84 ± 0.35
	CN	70.42 ± 12.10	0.3781 ± 0.0175	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
	DUET-MD	$\mathbf{75.42 \pm 12.10}$	$\mathbf{0.3839 \pm 0.0183}$	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
WW-domain	Frozen	0.00 ± 0.00	0.1347 ± 0.0018	100.00 ± 0.00	99.87 ± 0.20	99.87 ± 0.20
	CN	23.21 ± 11.70	0.1878 ± 0.0173	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
	DUET-MD	$\mathbf{50.00 \pm 15.41}$	$\mathbf{0.2250 \pm 0.0214}$	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
NTL9	Frozen	0.04 ± 0.09	0.1327 ± 0.0022	100.00 ± 0.00	99.60 ± 0.24	99.60 ± 0.24
	CN	27.86 ± 11.03	$\mathbf{0.2212 \pm 0.0216}$	100.00 ± 0.00	99.64 ± 0.73	99.64 ± 0.73
	DUET-MD	$\mathbf{35.36 \pm 13.40}$	0.2053 ± 0.0198	100.00 ± 0.00	97.14 ± 5.81	97.14 ± 5.81
(b) ConfRover						
BBA	Frozen	14.58 ± 1.53	0.4187 ± 0.0055	34.95 ± 1.94	98.02 ± 0.59	34.32 ± 1.87
	CN	39.17 ± 11.15	0.4663 ± 0.0228	41.67 ± 11.84	97.08 ± 3.61	40.42 ± 11.80
	DUET-MD	$\mathbf{40.42 \pm 14.33}$	$\mathbf{0.4675 \pm 0.0227}$	42.50 ± 14.48	100.00 ± 0.00	42.50 ± 14.48
WW-domain	Frozen	–	–	0.00 ± 0.00	99.61 ± 1.24	0.00 ± 0.00
	CN	–	–	0.00 ± 0.00	100.00 ± 0.00	0.00 ± 0.00
	DUET-MD	–	–	0.00 ± 0.00	100.00 ± 0.00	0.00 ± 0.00
NTL9	Frozen	–	–	0.00 ± 0.00	99.22 ± 2.49	0.00 ± 0.00
	CN	–	–	0.00 ± 0.00	100.00 ± 0.00	0.00 ± 0.00
	DUET-MD	–	–	0.00 ± 0.00	100.00 ± 0.00	0.00 ± 0.00

variable. We therefore interpret these results as recovery of target-directed, geometrically valid transition candidates rather than evidence of complete native folding.

5.3 CRYPTIC-POCKET OPENING

Setup and metrics. We evaluate cryptic-pocket opening with frozen PVB on PRMT5, PRMT6, and SMARCA2, with ubiquitin as a negative control. Following SLICE (Bemelmans et al., 2026), DUET-MD steers the separation of predefined local intramolecular contacts surrounding each region of interest, without conditioning on a target holo conformation. Task-specific regions, contact sets, steering potentials, and generation settings are provided in Appendix B.3.

We report stable contact-open mass for trajectories whose steered contacts remain at least 6.0 Å over the final two frames. We evaluate every generated frame with fpocket (Le Guilloux et al., 2009). For the three positive systems, the evaluation site corresponds to the holo-ligand binding site. For ubiquitin, we use the steered contact region as the evaluation site. We report path-max Druggability Score (DS) (Schmidtke & Barril, 2010), defined as the trajectory-weighted mean of each trajectory's maximum target-site DS over trajectories with at least one target-site pocket detection, together with the fraction of frames reaching DS≥ 0.5 and DS≥ 0.7. We use the same whole-path geometric-validity criteria as in the folding experiments (Section 5.2).

Results. Table 2 reports contact-open success alongside fpocket DS that are not used for steering. Across all four systems, DUET-MD increases stable contact-opening over complete-frame selection, showing that pre-completion steering improves access to the requested local conformational transition. In the known cryptic-pocket systems, the downstream pocket response is protein-dependent: DUET-MD increases the occurrence of high-DS target-site pockets over CN for PRMT5 and SMARCA2, whereas PRMT6 favors CN despite stronger contact opening under DUET-MD. Because neither fpocket scores nor holo-ligand geometry enters the steering objective, this separation shows that improved control of the chosen CV does not trivially determine pocket quality.

The ubiquitin negative control tests whether DUET-MD simply drives trajectories toward the prescribed steering objective irrespective of the physical prior encoded by the frozen surrogate. Table 2 shows that high-DS target-site pockets remain rare and no DS≥ 0.7 frame is observed under CN or

Table 2: Contact-program realization and independent target-site pocket evaluation. fpocket scores are not used for steering. Best results among the known cryptic-pocket systems are shown in bold.

			Target-site fpocket DS		
System	Method	Stable contact-open mass ↑	≥ 0.5 frames (%) ↑	≥ 0.7 frames (%) ↑	Path-max DS ↑
PRMT5	Frozen	0.000	0.99	0.67	0.549
	CN	0.219	2.21	1.91	0.558
	DUET-MD	**0.829**	**3.17**	**2.43**	**0.610**
PRMT6	Frozen	0.094	8.35	5.58	0.861
	CN	0.617	**10.07**	**6.94**	0.859
	DUET-MD	**0.753**	8.42	4.99	**0.876**
SMARCA2	Frozen	0.016	2.86	1.76	0.634
	CN	0.373	4.60	1.87	0.607
	DUET-MD	**0.763**	**7.20**	**4.04**	**0.800**
Ubiquitin	Frozen	0.016	0.23	0.11	0.289
	CN	0.104	0.09	0.00	0.312
	DUET-MD	0.489	0.26	0.00	0.320

DUET-MD, consistent with steering remaining constrained by the MD-derived transition prior of the frozen surrogate rather than forcing arbitrary bias-satisfying structures.

Mechanistic consistency. Beyond pocket scores, we further examine whether DUET-MD recovers the sequential structural rearrangements reported by SLICE (Bemelmans et al., 2026). SLICE identified a hierarchical opening mechanism in PRMT6 in which H163–M373 separation is followed by EE-loop displacement and L161 clearance. In our setup, H163–M373 is included among the steering contacts, but neither the temporal ordering nor the downstream rearrangements are encoded in the steering objective. A representative DUET-MD trajectory recovers this reported progression (Fig. 2), suggesting that DUET-MD does not merely generate arbitrary bias-satisfying paths but can preserve mechanistic transition structure learned from MD-derived trajectories.

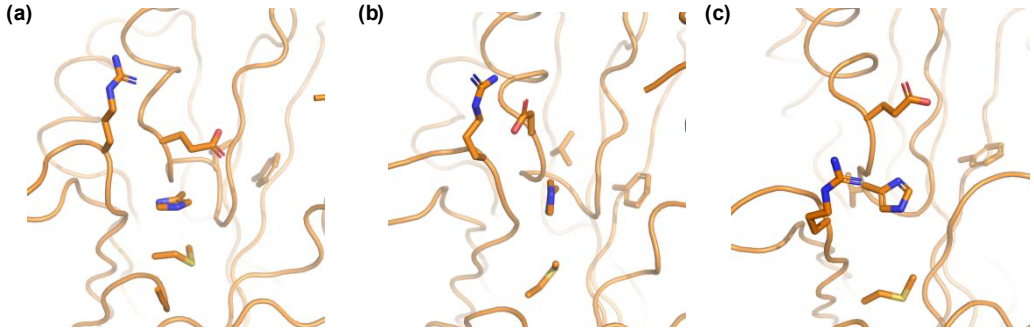


Figure 2: Representative structural snapshots from a PRMT6 pocket-opening trajectory generated by DUET-MD. (a) H163–M373 in contact. (b) Separation of H163–M373. (c) L161 vacates the pocket-facing region.

6 LIMITATIONS

DUET-MD remains constrained by both the frozen surrogate and the user-specified steering potential. It cannot recover transitions that are not practically accessible under the surrogate at a finite sampling budget, and successful steering toward a chosen CV does not guarantee recovery of structural features omitted from that CV. Moreover, our whole-path validity criteria assess only coarse geometric integrity rather than full atomistic physical validity, and the nominal transition intervals of the surrogates should not be interpreted as calibrated physical kinetics. We therefore view the generated paths as target-directed transition candidates for downstream mechanistic analysis, rather than estimates of unbiased pathway probabilities or transition rates.

7 Conclusion

We introduced DuET-MD, a training-free framework for rare-event steering of frozen iterative MD surrogates. By coupling pre-completion particle search with trajectory-level weighting and resampling across physical time, DuET-MD enables inference-time control without retraining the underlying dynamics model. Across folding and cryptic-pocket opening, the dual-clock construction improves access to low-probability conformational transitions, particularly when useful next-frame proposals are scarce. Importantly, DuET-MD does not simply force arbitrary bias-driven structures: pocket responses remain protein-dependent, the negative control does not automatically form a high-DS cavity, and generated trajectories exhibit additional rearrangements beyond the steered contacts. These results position inference-time particle control as a practical route toward controllable rare-transition sampling with pretrained MD surrogates.

AI Use Statement

Generative AI tools were used to assist with literature retrieval and discovery, research ideation and execution, and the development and verification of mathematical arguments. All AI-assisted outputs, including cited sources, methodological decisions, and mathematical claims, were independently reviewed and verified by the authors. The authors take full responsibility for the final manuscript.

Reproducibility Statement

The main text specifies the target distribution and DuET-MD inference procedure. Appendix A provides the proper-weighting derivation, and Appendix B provides the experimental configurations and evaluation definitions used in our study.

References

Panagiotis Antoniadis, Beatrice Pavesi, Simon Olsson, and Ole Winther. Protein language model embeddings improve generalization of implicit transfer operators. In *Proceedings of the International Conference on Machine Learning*, 2026.

Martijn P. Bemelmans, Alberto Borsatto, Simone Marsili, Francesco L. Gervasio, and Vineet Pande. Exploring the structural basis of cryptic pocket formation driven by extensive protein conformational changes in drug targets. *Journal of Chemical Theory and Computation*, 22(6):2714–2724, 2026. doi: 10.1021/acs.jctc.5c02016.

Robert B. Best, Gerhard Hummer, and William A. Eaton. Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences*, 110(44): 17874–17879, 2013. doi: 10.1073/pnas.1311599110.

Lohithsai Yadala Chanchu, Hany Abdulsamad, and Christian A. Naesseth. Discrete diffusion inference-time control with nested sequential Monte Carlo. arXiv preprint arXiv:2608.20123, 2026. URL https://arxiv.org/abs/2608.20123.

Kaihui Cheng, Ce Liu, Qingkun Su, Jun Wang, Liwei Zhang, Yining Tang, Yao Yao, Siyu Zhu, and Yuan Qi. 4D diffusion for dynamic protein structure prediction with reference and motion guidance. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 93–101, 2025. doi: 10.1609/aaai.v39i1.31984.

Kaihui Cheng, Zhiqiang Cai, Wenkai Xiang, Zhihang Hu, Siyu Zhu, Tzuhsiung Yang, and Yuan Qi. Learning implicit bias in generative spaces for accelerating protein dynamics emulation. arXiv preprint arXiv:2606.01833, 2026. URL https://arxiv.org/abs/2606.01833.

Hyungjin Chung, Jeongsol Kim, Michael T. McCann, Marc L. Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023.

Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436, 2006. doi: 10.1111/j.1467-9868.2006.00553.x.

Allan dos Santos Costa, Ilan Mitnikov, Franco Pellegrini, Ameya Daigavane, Mario Geiger, Zhonglin Cao, Karsten Kreis, Tess Smidt, Emine Kucukbenli, and Joseph Jacobson. EquiJump: Protein dynamics simulation via SO(3)-equivariant stochastic interpolants. arXiv preprint arXiv:2410.09667, 2024. URL https://arxiv.org/abs/2410.09667.

Randal Douc, Olivier Cappé, and Eric Moulines. Comparison of resampling schemes for particle filtering. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, 2005.

Ron O. Dror, Albert C. Pan, Daniel H. Arlow, David W. Borhani, Paul Maragakis, Yibing Shan, Huafeng Xu, and David E. Shaw. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proceedings of the National Academy of Sciences*, 108(32):13118–13123, 2011. doi: 10.1073/pnas.1104614108.

Jiaqi Han, Minkai Xu, Aaron Lou, Haotian Ye, and Stefano Ermon. Geometric trajectory diffusion models. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Chad W. Hopkins, Scott Le Grand, Ross C. Walker, and Adrian E. Roitberg. Long-time-step molecular dynamics through hydrogen mass repartitioning. *Journal of Chemical Theory and Computation*, 11(4):1864–1874, 2015. doi: 10.1021/ct5010406.

Aniketh Iyengar, Jiaqi Han, Pengwei Sun, Mingjian Jiang, Jianwen Xie, and Stefano Ermon. Align your structures: Generating trajectories with structure pretraining for molecular dynamics. In *The Fourteenth International Conference on Learning Representations*, 2026.

S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophysical Journal*, 72(4):1568–1581, 1997. doi: 10.1016/S0006-3495(97)78804-0.

Bowen Jing, Hannes Stärk, Tommi Jaakkola, and Bonnie Berger. Generative modeling of molecular dynamics trajectories. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Leon Klein, Andrew Y. K. Foong, Tor Erlend Fjelde, Bruno Mlodozeniec, Marc Brockschmidt, Sebastian Nowozin, Frank Noé, and Ryota Tomioka. Timewarp: Transferable acceleration of molecular dynamics by learning time-coarsened dynamics. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, 2002. doi: 10.1073/pnas.202427399.

Vincent Le Guilloux, Peter Schmidtke, and Pierre Tuffery. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics*, 10:168, 2009. doi: 10.1186/1471-2105-10-168.

Kresten Lindorff-Larsen, Stefano Piana, Ron O. Dror, and David E. Shaw. How fast-folding proteins fold. *Science*, 334(6055):517–520, 2011. doi: 10.1126/science.1208351.

J. Andrew McCammon, Bruce R. Gelin, and Martin Karplus. Dynamics of folded proteins. *Nature*, 267:585–590, 1977. doi: 10.1038/267585a0.

Christian A. Naesseth, Fredrik Lindsten, and Thomas B. Schön. Nested sequential Monte Carlo methods. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 1292–1301. PMLR, 2015.

Guillermo Pérez-Hernández, Fabian Paul, Toni Giorgino, Gianni De Fabritiis, and Frank Noé. Identification of slow molecular order parameters for Markov model construction. *The Journal of Chemical Physics*, 139(1):015102, 2013. doi: 10.1063/1.4811489.

Daniel D. Richman, Jessica Karaguesian, Carl-Mikael Suomivuori, and Ron O. Dror. Unlocking hidden biomolecular conformational landscapes in diffusion models at inference time. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Peter Schmidtke and Xavier Barril. Understanding and predicting druggability. a high-throughput method for detection of drug binding sites. *Journal of Medicinal Chemistry*, 53(15):5858–5867, 2010. doi: 10.1021/jm100574m.

Mathias Schreiner, Ole Winther, and Simon Olsson. Implicit transfer operator learning: Multiple time-resolution models for molecular dynamics. In *Advances in Neural Information Processing Systems*, volume 36, pp. 36449–36462, 2023.

Kiyoung Seong, Seonghyun Park, Seonghwan Kim, Woo Youn Kim, and Sungsoo Ahn. Transition path sampling with improved off-policy training of diffusion path samplers. In *The Thirteenth International Conference on Learning Representations*, 2025.

Yuning Shen, Lihao Wang, Huizhuo Yuan, Yan Wang, Bangji Yang, and Quanquan Gu. Simultaneous modeling of protein conformation and dynamics via autoregression. In *Advances in Neural Information Processing Systems*, volume 38, 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. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 55810–55827. PMLR, 2025.

Marta Skreta, Tara Akhound-Sadegh, Viktor Ohanesian, Roberto Bondesan, Alán Aspuru-Guzik, Arnaud Doucet, Rob Brekelmans, Alexander Tong, and Kirill Neklyudov. Feynman–Kac correctors in diffusion: Annealing, guidance, and product of experts. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 55906–55949. PMLR, 2025.

G. M. Torrie and J. P. Valleau. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, 1977. doi: 10.1016/0021-9991(77)90121-8.

Juan Viguera Diez, Mathias Schreiner, and Simon Olsson. Transferable generative models bridge femtosecond to nanosecond time-step molecular dynamics. *Science Advances*, 12(15):eaed2333, 2026. doi: 10.1126/sciadv.aed2333.

Paul C. Whitford, Jeffrey K. Noel, Shachi Gosavi, Alexander Schug, Kevin Y. Sanbonmatsu, and José N. Onuchic. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. *Proteins: Structure, Function, and Bioinformatics*, 75(2): 430–441, 2009. doi: 10.1002/prot.22253.

Luhuan Wu, Brian L. Trippe, Christian A. Naesseth, David M. Blei, and John P. Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Yu Xie, Ludwig Winkler, Lixin Sun, Sarah Lewis, Adam E. Foster, José Jiménez Luna, Tim Hempel, Michael Gastegger, Yaoyi Chen, Iryna Zaporozhets, Cecilia Clementi, Christopher M. Bishop, and Frank Noé. Enhanced diffusion sampling: Efficient rare event sampling and free energy calculation with diffusion models. arXiv preprint arXiv:2602.16634, 2026. URL https://arxiv.org/abs/2602.16634.

Ziyang Yu, Wenbing Huang, and Yang Liu. UniSim: A unified simulator for time-coarsened dynamics of biomolecules. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 73097–73114. PMLR, 2025.

Ziyang Yu, Wenbing Huang, and Yang Liu. Unified biomolecular trajectory generation via pretrained variational bridge. In *The Fourteenth International Conference on Learning Representations*, 2026.

Bin W. Zhang, David Jasnow, and Daniel M. Zuckerman. The "weighted ensemble" path sampling method is statistically exact for a broad class of stochastic processes and binning procedures. *The Journal of Chemical Physics*, 132(5):054107, 2010. doi: 10.1063/1.3306345.