

Multiscale autoregressive modeling generates ordered protein conformational trajectories

Mingchen Chen

mingchenchen@cpl.ac.cn

Changping Laboratory

Yaoyao Xu

The Chinese University of Hong Kong, Shenzhen, China

Yu Liu

Changping Laboratory

Tianshu Yu

School of Data Science, The Chinese University of Hong Kong - Shenzhen <https://orcid.org/0000-0002-6537-1924>

Article

Keywords: protein dynamics, conformational pathways, transition intermediates, kinase regulation, molecular dynamics

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11163218/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Multiscale autoregressive modeling generates ordered protein conformational trajectories

Yaoyao Xu^{b,a}, Yu Liu^a, Tianshu Yu^{b,*}, Mingchen Chen^{a,*}

This manuscript was compiled on September 25, 2026

^aChangping Laboratory, Beijing

^bThe Chinese University of Hong Kong, Shenzhen, China

Corresponding authors: Tianshu Yu and Mingchen Chen. E-mail:

yutianshu@cuhk.edu.cn; mingchenchen@cpl.ac.cn

Lead author: Yaoyao Xu

Abstract

Protein function depends on transitions among conformational states, but generating long, ordered structural sequences by molecular dynamics (MD) remains computationally expensive at the scale of large protein panels. Many generative approaches emphasize conformational ensembles, whereas sequential trajectory generators must also control error accumulation over long rollouts. We present TEMPO, a hierarchical autoregressive model trained on MD trajectories to generate ordered protein-backbone conformations at two temporal resolutions. A slow-scale model advances anchor conformations at the 20-ns sampling stride used for training, and a fast-scale model generates 1-ns-spaced frames conditioned on the slow-scale context. On held-out mdCATH and ATLAS proteins, TEMPO reproduced MD-derived ensemble observables. On ATLAS, Pearson correlations were 0.91 for ensemble-averaged pairwise RMSD and 0.89 for residue-level RMSF; the leading principal-component direction agreed with MD in 76% of cases, compared with 44% for AlphaFlow. Relative to a single-scale autoregressive model, the hierarchy reduced long-horizon RMSD error from 8.62 to 1.78 Å, and a 400-frame rollout required less than 1 min. Across 10 endpoint-conditioned adenylate kinase rollouts, progression along a shared time-lagged coordinate was accompanied by progressive contact exchange during opening. Inference-time steering yielded conformations satisfying predefined DFG-in- and DFG-out-like local backbone criteria in 68 of 100 kinases. Across 252 kinases, fluctuation profiles showed variable agreement with crystallographic B-factors (mean Spearman $\rho = 0.358$; 22% with $\rho \geq 0.5$). TEMPO provides a framework for generating and comparing ordered protein conformational sequences.

Keywords: protein dynamics | conformational pathways | transition intermediates | kinase regulation | molecular dynamics

Main Text

Proteins are dynamic molecular machines whose functions depend on coordinated structural changes. Local fluctuations tune active-site geometry, whereas larger loop and domain rearrangements regulate binding, catalysis, and signaling [1–3]. Static structures can define the endpoints of such motions, and ensembles can describe populated conformations, but mechanisms also depend on how those conformations connect [4]. For many questions, the relevant description is therefore an ordered sequence of changes: which contacts exchange first, which regions move cooperatively, and which intermediates connect functional states.

Molecular dynamics (MD) simulations provide atomically resolved trajectories under physics-based models [5], but their cost limits repeated sampling and systematic comparisons across proteins, mutants, or ligand states. Deep generative models such as AlphaFlow and BioEmu can rapidly sample diverse protein conformational ensembles [6, 7], and recent methods such as MDGen have begun to model sequential trajectories [8]. Ensemble generation seeks to reproduce a distribution of structures, whereas trajectory generation must also preserve temporal dependence among frames. A smooth sequence may drift from MD-like conformational statistics, while an accurate ensemble need not provide meaningful ordering. Both ensemble-level and sequence-level evidence are therefore required.

Sequential generation becomes increasingly difficult with rollout length because each prediction becomes part of the context for the next, allowing geometric errors to accumulate. Protein motions are also organized across temporal scales, with local fluctuations embedded within slower collective rearrangements on a hierarchical conformational landscape [9]. A single update scale must therefore balance short-range variation against stable long-range evolution. A multiscale generator can separate these roles while retaining an explicitly ordered output.

Here we introduce TEMPO, a hierarchical autoregressive model trained directly on MD trajectories (Fig. 1). A slow-scale model advances anchor conformations at the 20-ns sampling stride represented during training, and a fast-scale model generates 1-ns-spaced frames conditioned on the slow-scale context. The slow anchors organize longer-range evolution, whereas the fast-scale model resolves structural variation between anchors. This decomposition is designed to limit long-rollout drift without reducing trajectory generation to independent sampling around each predicted state.

We benchmark TEMPO on held-out mdCATH and ATLAS proteins and use a single-scale ablation to assess the contribution of the multiscale hierarchy. We then examine

endpoint-conditioned adenylate kinase (ADK) opening, inference-time steering of local DFG geometry, family-scale kinase variation, and exploratory transfer to six G protein-coupled receptors (GPCRs). These analyses test ensemble fidelity, long-rollout stability, and whether frame ordering carries interpretable structural information.

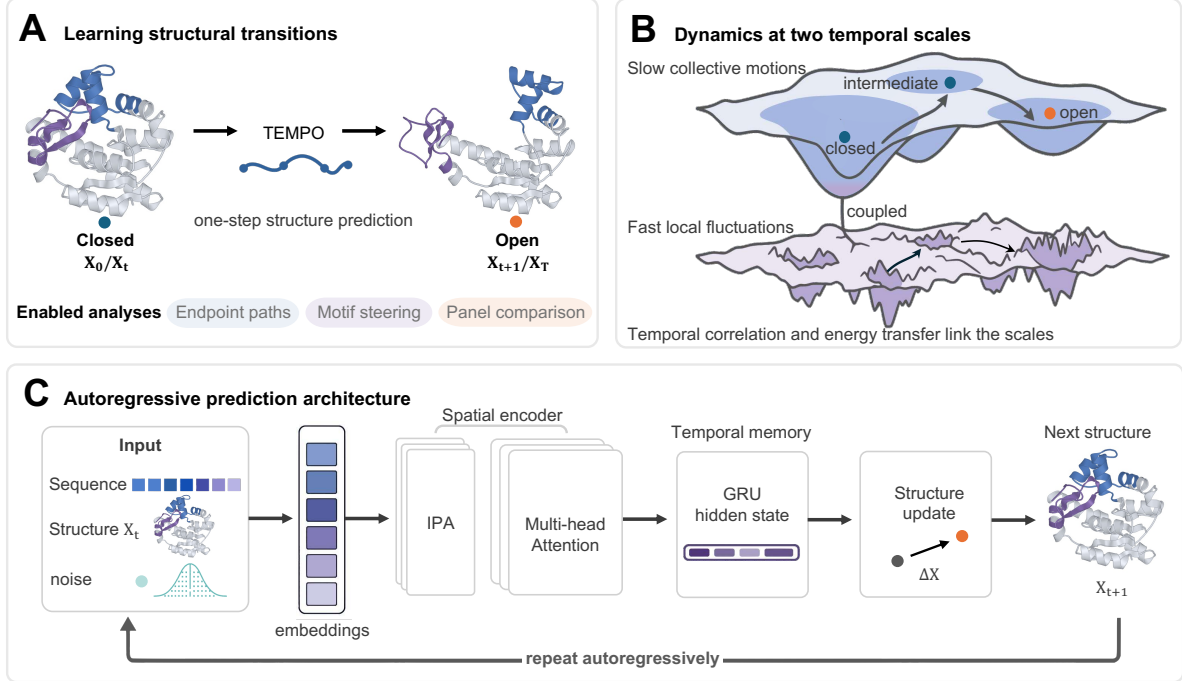


Figure 1: TEMPO generates protein conformational trajectories hierarchically. (A) TEMPO advances conformations autoregressively to generate an ordered structural trajectory. The schematics illustrate the three applications examined in this study: endpoint-conditioned generation toward a specified conformation, inference-time steering of local-motif geometry, and trajectory comparison across protein panels. (B) TEMPO resolves motion hierarchically: a slow autoregressive model advances anchor conformations at the 20-ns sampling stride used during training, whereas a fast autoregressive model generates intermediate frames at the 1-ns stride while conditioning on the slow-scale context. During inference, slow-scale anchors are generated first, after which the corresponding fast-scale segments are assembled into an ordered trajectory. The lower landscape enlarges the highlighted region above, illustrating how fast local fluctuations are resolved within slower collective conformational changes. (C) Network architecture shared by slow- and fast-scale models. Sequence features, structural context, and stochastic noise are embedded and processed using invariant point attention (IPA) and multi-head attention. Further details of the representation, architecture, training procedure, and loss functions are provided in Materials and Methods and SI Appendix.

Results

TEMPO was trained on mdCATH and evaluated on held-out mdCATH and ATLAS proteins [10, 11]. The ADK analysis used the endpoint-conditioned model; the kinase and

GPCR analyses used the mdCATH-trained base model, except for the explicit restraint applied during DFG steering.

A multiscale hierarchy stabilizes long rollouts while preserving MD-like conformational variation

We compared TEMPO with BioEmu [7], AlphaFlow, ESMFlow [6], and MDGen [8] using measures of ensemble breadth, residue mobility, collective coordinates [12, 13], and weak or transient contacts [14]. On ATLAS, TEMPO achieved Pearson correlations of 0.91 for ensemble-averaged pairwise RMSD and 0.89 for residue-level RMSF. It yielded the lowest PCA-space 2-Wasserstein distance on both ATLAS and mdCATH, with values of 0.60 and 2.33, respectively (Fig. 2B and C). The leading principal-component direction met the prespecified agreement criterion in 76% of ATLAS cases, compared with 44% for AlphaFlow. Jaccard similarities for weak and transient contacts were 0.43 and 0.20 on mdCATH and 0.74 and 0.30 on ATLAS, respectively. Metric definitions are provided in SI Appendix, Section S3.

Relative to a single-scale model rolled out for 400 frames, the multiscale model reduced long-horizon RMSD error from 8.62 to 1.78 Å, increased pairwise-RMSD correlation from 0.14 to 0.77, and increased global RMSF correlation from 0.15 to 0.67 (SI Appendix, Table S2). A 400-frame TEMPO rollout required less than 1 min, approximately 750-fold less time than AlphaFlow under the reported evaluation settings (Fig. 2A).

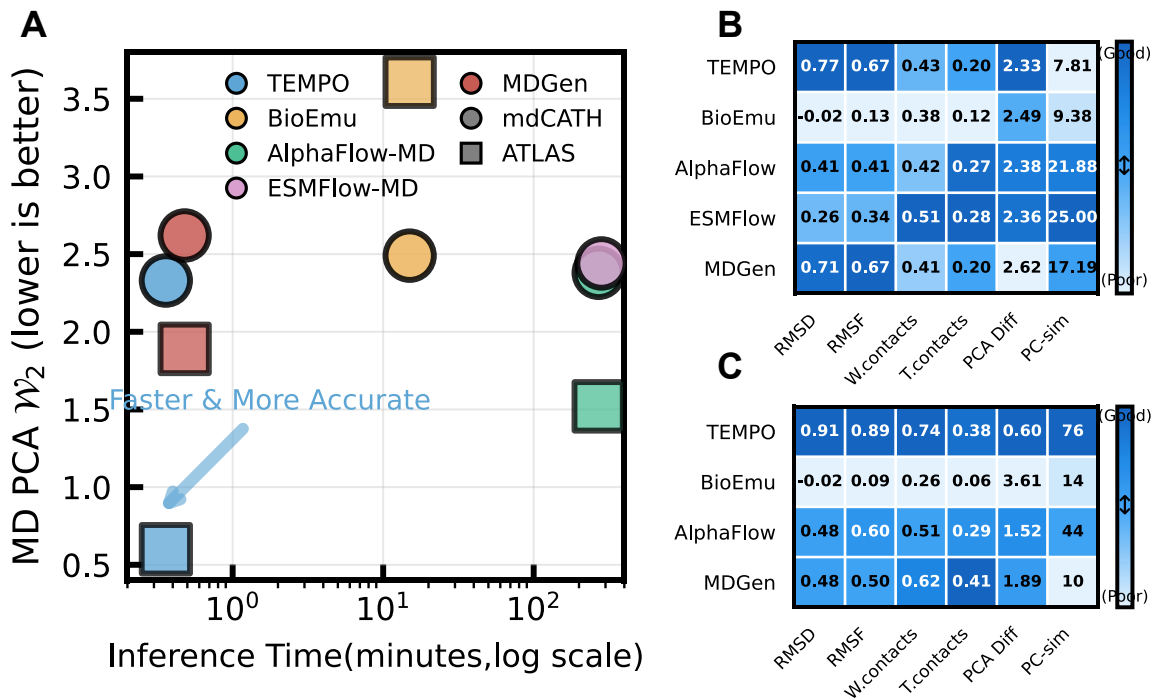


Figure 2: **TEMPO performance on MD-derived ensemble observables and generation time.** (A) Reported inference time versus PCA-space Wasserstein distance ($\mathcal{W}_2$) to the MD reference. Circle and square markers denote performance on the mdCATH and ATLAS datasets, respectively. (B) Performance summary on mdCATH. (C) Performance summary on ATLAS. Metric definitions and aggregation procedures are described in SI Appendix, Section S3.

ADK opening links conformational coverage to temporal ordering

ADK opening involves coordinated motion of the NMP and LID domains around a comparatively stable CORE [15–18]. To test frame ordering during this transition, we generated 10 independent 500-frame rollouts from closed 1AKE conditioned on open 4AKE.

Generated conformations formed a continuous corridor between the endpoints in RMSD space (Fig. 3A). Three recurrent regions represented progressively open NMP and LID conformations, while the CORE remained comparatively stable (Fig. 3B).

Using kinetic-map tICA [19], a model fitted jointly to the 10 rollouts identified tIC1 as the shared closed-to-open coordinate and tIC2 as a coordinate capturing variation among generated routes (Fig. 3C). Across the 10 rollouts, tIC1 showed a consistent overall progression with frame order. Closed-state-specific contacts decreased and open-state-specific contacts accumulated across States 1–3, while shared contacts preserved the intradomain scaffold (Fig. 3D). Mobility and deformation localized to the NMP/LID domains, hinges, and interfaces (SI Appendix, Figs. S2–S3), and six additional systems showed connected corridors in endpoint-RMSD space (SI Appendix, Fig. S4).

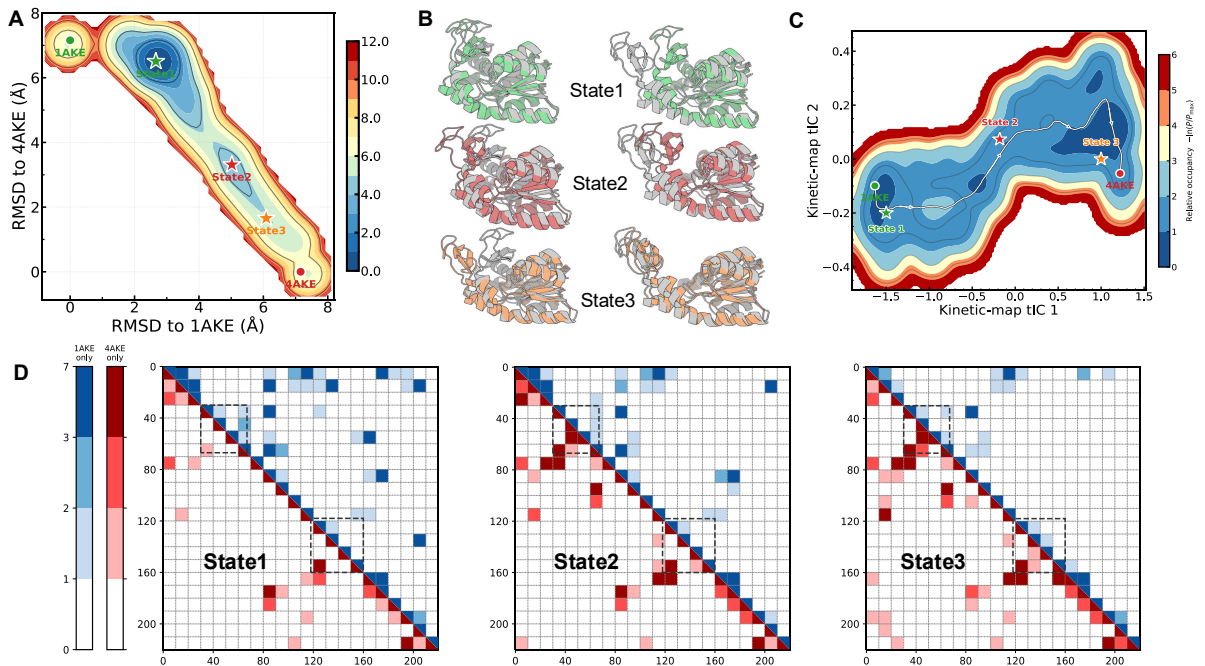


Figure 3: Endpoint-conditioned TEMPO generates an ordered adenylate kinase opening sequence. (A) Relative occupancy of 5,000 conformations from 10 closed-to-open rollouts projected by RMSD to closed 1AKE and open 4AKE. Color shows $-\ln(P/P_{\max})$, with zero denoting the highest generated density; stars mark three representative state clusters. (B) Representative frames from States 1–3 aligned to the closed and open structures. (C) Kinetic-map tICA at a five-frame lag, fitted jointly to all 10 rollouts. For clarity, contours show the two rollouts closest to the median route in tIC2; stars mark the same representative frames as in B, and arrowheads indicate increasing frame index. (D) Contact maps for States 1–3. Closed-specific contacts are shown in blue and open-specific contacts in red. Dashed boxes mark the NMP and LID domains; $C\alpha$ - $C\alpha$ cutoff, 10 Å.

Inference-time steering generates alternative local DFG backbone geometries

We next tested whether inference-time steering could generate alternative local DFG backbone geometries across a kinase family. The DFG motif at the start of the activation loop undergoes regulatory backbone rearrangements [20, 21], and DFG-out conformations are associated with the pocket recognized by type II inhibitors [22, 23]. Separate rollouts were directed toward predefined DFG-in- and DFG-out-like criteria for each of 100 kinases; no model parameters were updated.

For 68 kinases, the DFG-in-guided and DFG-out-guided rollouts each yielded at least one accepted frame (Fig. 4A). Accepted frame pairs occupied distinct regions of the DFG-Phe Ramachandran plot and showed local backbone displacement (Fig. 4A and B). The other 32 kinases did not yield both frame classes under the current rollout length and steering protocol. Two large-separation examples are shown in Fig. 4C. Because both

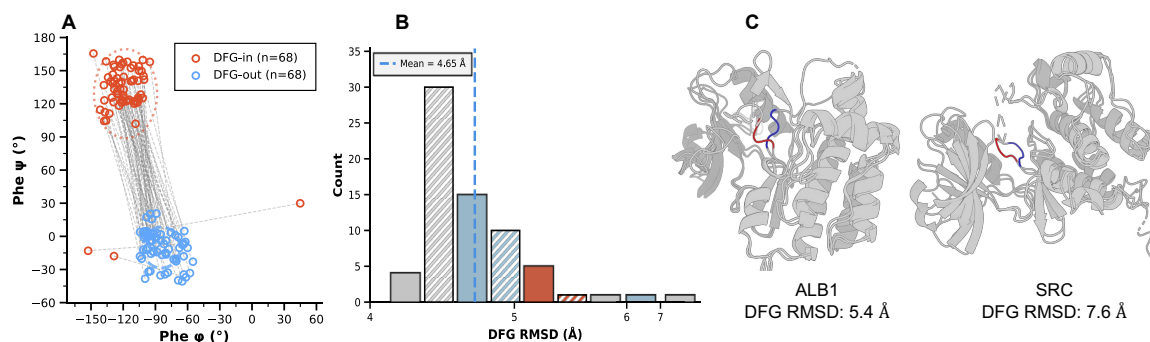


Figure 4: **Inference-time steering generates alternative local DFG backbone geometries across a kinase panel.** (A) Ramachandran distribution of DFG-Phe for the 68 kinases in which the DFG-in-guided rollout yielded an accepted in-like frame and the DFG-out-guided rollout yielded an accepted out-like frame. Red circles denote DFG-in-like frames, blue circles denote DFG-out-like frames, and gray lines connect paired frames from the same kinase. (B) Maximal DFG-region backbone RMSD between accepted in-like and out-like frames for each kinase. (C) Local DFG structures from two kinases selected for large separation between accepted frame classes. The DFG motif is shown in blue (in-like) and red (out-like), and the remainder of the kinase is gray.

the steering and acceptance criteria target local DFG geometry, this analysis evaluates controllability rather than spontaneous switching.

Kinase rollouts show directional variation in contact similarity and B-factor agreement

We next used unguided rollouts to compare opposite-state contact similarity and residue-resolved flexibility across kinase panels. For 104 kinases with paired active and inactive structures, we generated 400-frame rollouts from each endpoint and calculated Q-scores relative to the opposite-state reference [24] (Fig. 5A). Because active and inactive structures share a large contact core, Q-scores were treated as continuous similarity measures rather than state assignments.

Directional differences in mean Q-score were at most 0.05 for 63% of kinases, 0.05–0.10 for 22%, and greater than 0.10 for 15% (SI Appendix, Fig. S5). These values describe asymmetry in opposite-state contact similarity rather than transition barriers.

Across 252 kinases, generated and crystallographic flexibility profiles had a mean Spearman correlation of 0.358 and a mean z-score RMSE of 1.16; 66.3% had $\rho \geq 0.3$, and 22% had $\rho \geq 0.5$ (Fig. 5B and C). Agreement varied across the panel, with PDB ID 5TQ8 shown as a high-correlation example (Fig. 5D and E).

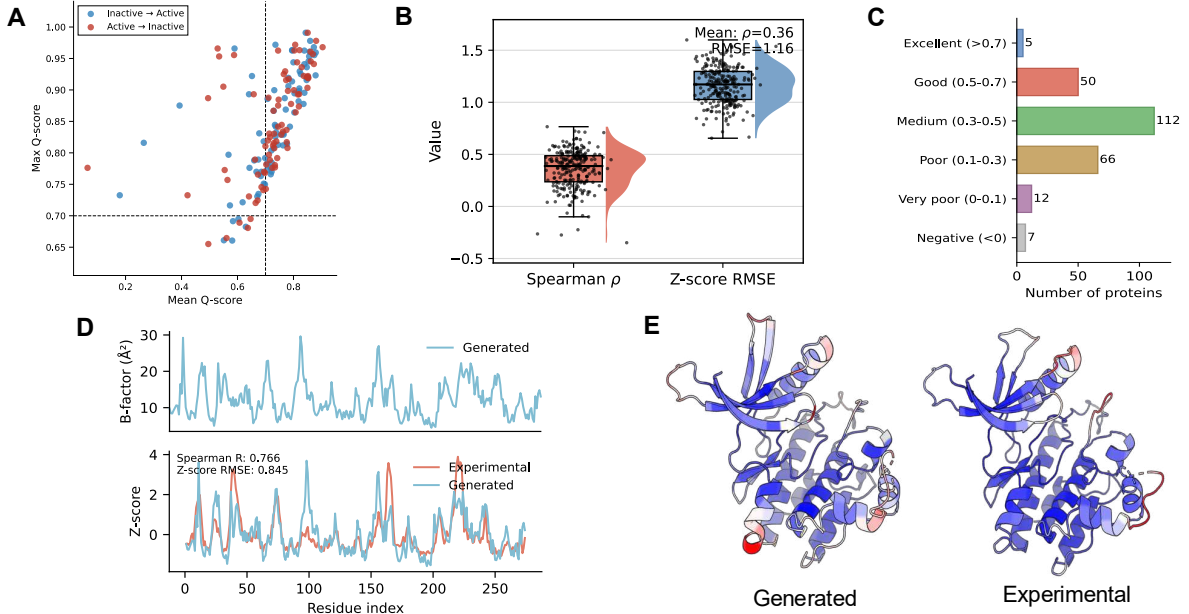


Figure 5: Kinase rollouts show directional variation in contact similarity and B-factor agreement. (A) Opposite-state contact similarity for 104 kinases. Each point represents one 400-frame rollout initiated from an inactive (blue) or active (red) conformation and is plotted by its mean and maximum Q-score relative to the opposite reference. Dashed reference lines, if retained in the artwork, are visual guides only and are not used for state assignment. (B) Agreement between residue-level fluctuations from generated rollouts and crystallographic B-factors across 252 kinases, summarized by Spearman correlation (ρ) and z-score RMSE. Larger ρ denotes greater rank agreement, and smaller RMSE denotes closer normalized profiles. (C) Fractions of kinases within the displayed Spearman-correlation ranges; these bins are descriptive and do not define validated performance categories. (D) High-correlation example, PDB ID 5TQ8. Top, per-residue B-factors derived from the generated rollout; bottom, z-score-normalized experimental and generated profiles (Spearman ρ = 0.766, z-score RMSE = 0.845). (E) Residue-wise B-factor patterns mapped onto the generated (left) and experimental (right) structures.

Exploratory transfer to GPCR backbone variability

To test transfer beyond soluble proteins, we applied the mdCATH-trained model without GPCR-specific updates to six receptors spanning Classes A–C [25]. TEMPO generated one 500-frame backbone rollout per receptor. In the two examples shown, generated conformations overlapped prominent regions of the receptor-only MD projection; variation was concentrated in the N terminus, extracellular loops, and intracellular helix 8, while the transmembrane core remained comparatively stable (Fig. 6). Results for the other four receptors are shown in SI Appendix, Fig. S6.

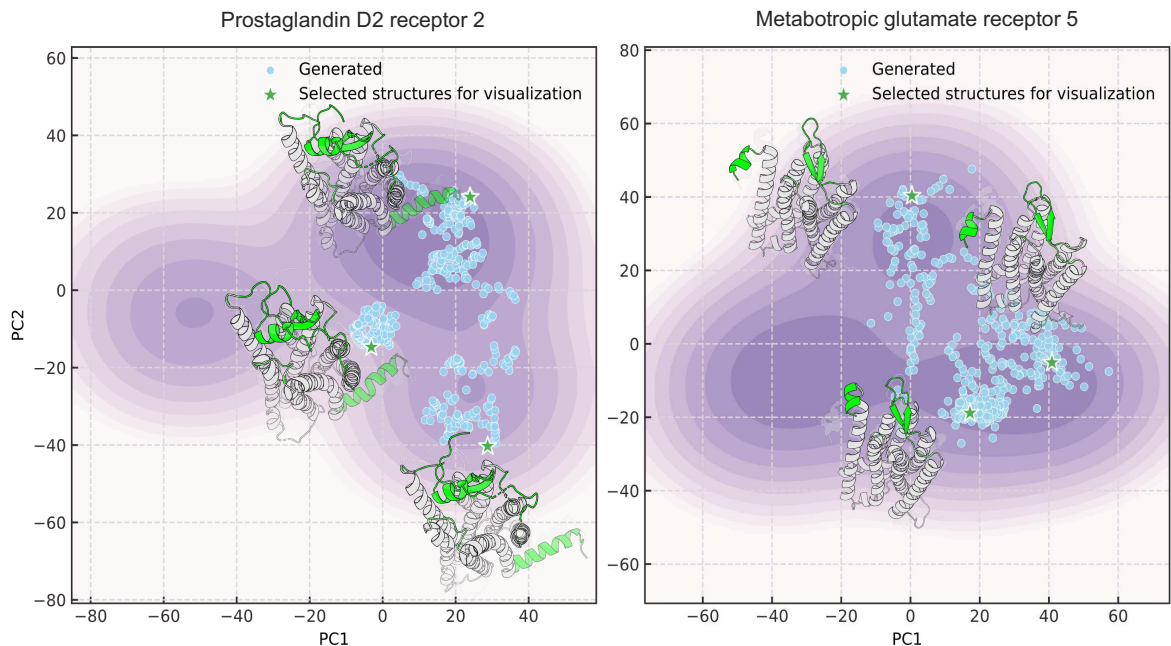


Figure 6: **Exploratory transfer to GPCR backbone conformational variability.** Two receptors from the six-protein test set are shown: prostaglandin D₂ receptor 2 (left) and metabotropic glutamate receptor 5 (right). Purple contours denote the projected density of the receptor-only MD reference in PC1–PC2 space, blue points denote TEMPO-generated conformations, and green stars mark three generated conformations selected for structural visualization. Regions that vary most among the displayed conformations are highlighted in green.

Discussion

TEMPO is a hierarchical autoregressive model for generating ordered protein-backbone sequences. Across benchmark proteins, it reproduced several MD-derived ensemble observables, and the single-scale ablation showed that slow anchors substantially reduce drift during long rollouts.

The application studies illustrate complementary uses of the generated sequences. Endpoint conditioning organized ADK opening along a shared time-lagged coordinate with progressive contact exchange, whereas inference-time steering generated alternative local DFG geometries across a kinase panel. Unguided kinase and exploratory GPCR analyses further showed that the same mdCATH-trained model can support family-scale comparisons and transfer to proteins not used for training. Conditioned and steered rollouts assess structural ordering and controllability; unconditioned rollouts characterize the learned conformational prior.

TEMPO remains backbone-centric and omits side chains, ligands, membranes, ions, and solvent. Its 1- and 20-ns frame labels reflect the sampling strides used in training rather than calibrated kinetic time, and conditioned or steered rollouts are not unbiased samples. Extending the model to include environmental context and validating generated

states with state-discriminating observables, energetic calculations, or experiments will be important. Within these limits, TEMPO provides a rapid complement to physics-based simulation for comparing ordered conformational changes across proteins.

Materials and Methods

Protein backbone representation

Consistent with standard representations in protein modeling [26], each residue is represented by an SE(3) backbone frame and backbone torsion angles; full definitions are provided in SI Appendix, Section S1. The SE(3) frame defines a local coordinate system through a rotation, represented by a unit quaternion $\mathbf{q} \in \mathbb{R}^4$, and a translation $\mathbf{t} \in \mathbb{R}^3$, both constructed from the backbone N, C α , and C atoms. Backbone torsions (ϕ, ψ, ω) are encoded as sine-cosine pairs to avoid discontinuities at $\pm\pi$. The per-residue state at time t is

$$\mathbf{X}_t^l = [q_w, q_x, q_y, q_z, t_x, t_y, t_z, \sin \phi, \cos \phi, \sin \psi, \cos \psi, \sin \omega, \cos \omega], \quad (1)$$

where l indexes residues.

SDE-inspired autoregressive dynamics

We parameterize conformational evolution with an SDE-inspired discrete update [27–29]. Given a two-frame context $\text{ctx}_t = (\mathbf{X}_{t-\Delta t}, \mathbf{X}_t)$ [30, 31], the model predicts

$$\mathbf{X}_{t+\Delta t} = \mathbf{X}_t + f_\theta(\text{ctx}_t)\Delta t + \sigma\sqrt{\Delta t}\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \quad (2)$$

where f_θ is a neural network. Because each update is conditioned on multiple snapshots, the formulation is not first-order Markovian in $\mathbf{X}_t$ alone. Full training and inference procedures are provided in SI Appendix, Algorithms S1 and S2.

Multiscale hierarchical framework

TEMPO separates generation into slow and fast temporal resolutions, each represented by an independently trained autoregressive network. The slow-scale model first generates anchor conformations at the $\Delta t_s = 20$ ns sampling stride. Each 20-frame fast-scale segment is initialized from the corresponding slow-scale anchor, after which the fast-scale model generates the remaining frames autoregressively at the $\Delta t_f = 1$ ns stride. Frame indices therefore follow the sampling strides used to construct the training trajectories.

Spatiotemporal neural architecture

The slow- and fast-scale models share the same architecture but are trained separately. Spatial encoding uses Invariant Point Attention (IPA) [26] to extract geometry-aware features from SE(3) frames and sequence embeddings, followed by multi-head self-attention layers for inter-residue interactions. Temporal encoding employs gated recurrent unit (GRU) layers to model frame-to-frame dependencies. The final output layer predicts per-residue conformational updates, which are applied autoregressively to obtain $\mathbf{X}_{t+\Delta t}$.

Training objective

The training loss combines mean-squared error reconstruction with a steric-clash penalty:

$$\mathcal{L}_{\text{total}} = \|\mathbf{X}_{t+\Delta t} - \hat{\mathbf{X}}_{t+\Delta t}\|^2 + \lambda \sum_{(i,j) \in \mathcal{P}_{\text{nonadj}}} \text{ReLU}(1.2\text{\AA} - \|\mathbf{r}_i - \mathbf{r}_j\|), \quad (3)$$

where $\mathbf{r}_i$ denotes backbone atom positions and $\mathcal{P}_{\text{nonadj}}$ is the set of non-adjacent atom pairs; the second term penalizes pairs closer than 1.2 Å. The model is trained autoregressively with teacher forcing using an epoch-dependent, beta-distributed noise schedule over [0.01, 1.0]. The schedule initially favors noise scales near the lower end of this range and gradually approaches uniform sampling across the full range as training proceeds. At inference, the slow-scale model retains comparable noise levels, whereas the fast-scale model uses $\sigma_{\text{noise}} = 5.0$ to promote diverse local sampling. Here, σ_{noise} is an internal model hyperparameter rather than a physical temperature, displacement, or kinetic parameter. Model parameters and additional implementation details are provided in SI Appendix, Table S1.

Datasets and evaluation design

We trained on 320 K trajectories from 4,000 randomly sampled mdCATH proteins using three independent seeds.

The ADK, kinase, and GPCR proteins used in downstream analyses were not included in the mdCATH training set. The kinase panels were derived from a published collection of human kinase conformational ensembles [32]. ADK used the endpoint-conditioned model described below. The kinase flexibility and cross-state analyses used the mdCATH-trained base model, with an additional inference-time restraint only for DFG steering. The exploratory GPCR analysis also used the base model without GPCR-specific parameter updates. This design tests transfer to proteins not used for training but does not, by itself, establish homology-level out-of-distribution generalization.

Inference times in Fig. 2A correspond to the evaluation setting used for each comparator; the approximately 750-fold ratio reported in the Results is relative to AlphaFlow.

Endpoint-conditioned generation

For targeted transitions, generation is conditioned on a user-specified target conformation $\mathbf{X}^*$. The endpoint-conditioned model uses the same prediction loss as the base model while providing $\mathbf{X}^*$ as an additional input to the slow- and fast-scale networks at each step. A learnable encoder E_ψ maps the target structure to a per-residue embedding $\mathbf{e}^* = E_\psi(\mathbf{X}^*) \in \mathbb{R}^{L \times d}$, which is added to the latent state before spatial and temporal encoding:

$$\mathbf{h}_{\text{latent}}^{\text{cond}} = \mathbf{h}_{\text{latent}} + \mathbf{e}^*. \quad (4)$$

At inference, $\mathbf{X}^*$ is held fixed throughout the rollout, biasing the generated sequence toward the specified target conformation.

For ADK, we generated 10 independent 500-frame rollouts from closed 1AKE while conditioning on open 4AKE. Generated-state occupancy was summarized in the two-dimensional space defined by RMSD to the endpoints using the score $-\ln(P/P_{\text{max}})$. This score summarizes occupancy within the endpoint-conditioned ensemble and is not an equilibrium free-energy estimate. Clustering in the endpoint-RMSD projection identified three representative states, and one frame from each cluster was used for structural visualization, tICA projection, and contact-map analysis. For tICA, each seed was treated as a separate trajectory, frames were aligned on CORE C α atoms, and a common kinetic-map basis was fitted at a lag of five frames. Detailed tICA, flexibility, and strain analyses are provided in SI Appendix, Section S5.

Inference-time steering of local DFG geometry

To direct kinase rollouts toward alternative local DFG backbone geometries, we applied a steering procedure after each autoregressive generation step without updating model parameters. The procedure refines the DFG motif and neighboring residues with a multi-term objective that biases local torsions and C α displacements toward the specified DFG-in- or DFG-out-like criterion while constraining peptide geometry, continuity to flanking residues, steric overlap, and frame-to-frame changes. The remaining residues are held fixed. Accepted frames were classified using prespecified Ramachandran-basin signatures of DFG-Phe and additional local geometric criteria. Full objective definitions and hyperparameters are provided in SI Appendix, Section S2.

Kinase cross-state contact analysis

For 104 kinases with paired active and inactive structures, we generated rollouts from each endpoint and calculated a Q-score relative to the opposite reference structure. We report the mean and maximum opposite-state Q-score within each 400-frame rollout. To summarize directional asymmetry, we classified the absolute differences between mean Q-

scores from active- and inactive-initiated rollouts as ≤ 0.05 , $0.05\text{--}0.10$, or > 0.10 . Because the score includes contacts shared by both endpoint structures and the maximum is an extreme-value summary, these measures are descriptive rather than direct evidence of state transitions or dwell times. Q-score definitions are provided in SI Appendix, Section S3.

Crystallographic flexibility comparison

For comparison with crystallographic B-factors, per-residue C α RMSF values were computed from TEMPO-generated ensembles after sequence alignment and least-squares C α superposition to a reference frame [33]. RMSF values were converted to predicted B-factors using $B_i = \frac{8\pi^2}{3}\text{RMSF}_i^2$, and experimental C α B-factors were extracted from the corresponding PDB structures. Agreement was assessed using Spearman correlation between residue-wise profiles and RMSE between z-score-normalized profiles. Because crystallographic B-factors also reflect crystal packing, static disorder, and refinement [34], this comparison evaluates relative spatial patterns rather than absolute solution-state fluctuations. Full definitions are provided in SI Appendix, Section S3.

Exploratory GPCR transfer analysis

We selected six GPCR complexes from GPCRmd, with two receptors each from Classes A, B, and C. Nonprotein components were removed to form receptor-only reference trajectories, from which 500 frames at the 1-ns sampling stride were retained. Starting from the first processed frame, the mdCATH-trained base model generated one 500-frame backbone rollout per receptor without GPCR-specific retraining. Principal components were fitted to each receptor-only MD reference, and generated frames were projected into the same PC1–PC2 coordinates for qualitative comparison. This analysis does not evaluate membrane-, ligand-, solvent-, or ion-dependent dynamics.

Data Availability

TEMPO training and inference code and the analysis scripts are available in the TEMPO GitHub repository <https://github.com/LOGO-CUHKSZ/TEMPO>. A time-stamped code snapshot and the processed data underlying the figures, including generated-ensemble summaries and evaluation outputs, will be deposited on Zenodo upon publication. During peer review, these materials are available to editors and reviewers upon request.

Acknowledgments

This work was supported by Changping Laboratory and in part by the Guangdong Provincial Key Laboratory of Mathematical Foundations for Artificial Intelligence (2023B1212010001).

References

- [1] Diego U Ferreiro, Elizabeth A Komives, and Peter G Wolynes. Frustration in biomolecules. *Quarterly reviews of biophysics*, 47(4):285–363, 2014.
- [2] Katherine Henzler-Wildman and Dorothee Kern. Dynamic personalities of proteins. *Nature*, 450(7172):964–972, 2007.
- [3] David D Boehr, Ruth Nussinov, and Peter E Wright. The role of dynamic conformational ensembles in biomolecular recognition. *Nature chemical biology*, 5(11):789–796, 2009.
- [4] Frank Noé and Stefan Fischer. Transition networks for modeling the kinetics of conformational change in macromolecules. *Current opinion in structural biology*, 18(2):154–162, 2008.
- [5] David E Shaw, Paul Maragakis, Kresten Lindorff-Larsen, Stefano Piana, Ron O Dror, Michael P Eastwood, Joseph A Bank, John M Jumper, John K Salmon, Yibing Shan, et al. Atomic-level characterization of the structural dynamics of proteins. *Science*, 330(6002):341–346, 2010.
- [6] Bowen Jing, Bonnie Berger, and Tommi Jaakkola. Alphafold meets flow matching for generating protein ensembles. In *Proceedings of the 41st International Conference on Machine Learning*, pages 22277–22303, 2024.
- [7] Sarah Lewis, Tim Hempel, José Jiménez-Luna, Michael Gastegger, Yu Xie, Andrew YK Foong, Victor García Satorras, Osama Abdin, Bastiaan S Veeling, Iryna Zaporozhets, et al. Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science*, 389(6761):eadv9817, 2025.
- [8] Bowen Jing, Hannes Stärk, Tommi Jaakkola, and Bonnie Berger. Generative modeling of molecular dynamics trajectories. *Advances in Neural Information Processing Systems*, 37:40534–40564, 2024.
- [9] Hans Frauenfelder, Stephen G Sligar, and Peter G Wolynes. The energy landscapes and motions of proteins. *Science*, 254(5038):1598–1603, 1991.

- [10] Antonio Mirarchi, Toni Giorgino, and Gianni De Fabritiis. mdcath: A large-scale md dataset for data-driven computational biophysics. *Scientific Data*, 11(1):1299, 2024.
- [11] Yann Vander Meersche, Gabriel Cretin, Aria Gheeraert, Jean-Christophe Gelly, and Tatiana Galochkina. Atlas: protein flexibility description from atomistic molecular dynamics simulations. *Nucleic acids research*, 52(D1):D384–D392, 2024.
- [12] Charles C David and Donald J Jacobs. Principal component analysis: a method for determining the essential dynamics of proteins. In *Protein dynamics: Methods and protocols*, pages 193–226. Springer, 2013.
- [13] Berk Hess. Similarities between principal components of protein dynamics and random diffusion. *Physical Review E*, 62(6):8438, 2000.
- [14] Urmi Doshi, Michael J Holliday, Elan Z Eisenmesser, and Donald Hamelberg. Dynamical network of residue–residue contacts reveals coupled allosteric effects in recognition, catalysis, and mutation. *Proceedings of the National Academy of Sciences*, 113(17):4735–4740, 2016.
- [15] Oliver Beckstein, Elizabeth J Denning, Juan R Perilla, and Thomas B Woolf. Zipping and unzipping of adenylate kinase: atomistic insights into the ensemble of open↔closed transitions. *Journal of molecular biology*, 394(1):160–176, 2009.
- [16] Osamu Miyashita, José Nelson Onuchic, and Peter G Wolynes. Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proceedings of the National Academy of Sciences*, 100(22):12570–12575, 2003.
- [17] Osamu Miyashita, Peter G Wolynes, and Jose N Onuchic. Simple energy landscape model for the kinetics of functional transitions in proteins. *The journal of physical chemistry B*, 109(5):1959–1969, 2005.
- [18] Paul C Whitford, Osamu Miyashita, Yaakov Levy, and José N Onuchic. Conformational transitions of adenylate kinase: switching by cracking. *Journal of molecular biology*, 366(5):1661–1671, 2007.
- [19] Frank Noé and Cecilia Clementi. Kinetic distance and kinetic maps from molecular dynamics simulation. *Journal of chemical theory and computation*, 11(10):5002–5011, 2015.
- [20] Peter Man-Un Ung, Rayees Rahman, and Avner Schlessinger. Redefining the protein kinase conformational space with machine learning. *Cell chemical biology*, 25(7):916–924, 2018.

- [21] Yibing Shan, Markus A Seeliger, Michael P Eastwood, Filipp Frank, Huafeng Xu, Morten Ø Jensen, Ron O Dror, John Kuriyan, and David E Shaw. A conserved protonation-dependent switch controls drug binding in the abl kinase. *Proceedings of the National Academy of Sciences*, 106(1):139–144, 2009.
- [22] Vivek Modi and Roland L Dunbrack Jr. Defining a new nomenclature for the structures of active and inactive kinases. *Proceedings of the National Academy of Sciences*, 116(14):6818–6827, 2019.
- [23] Barun Okram, Advait Nagle, Francisco J Adrián, Christian Lee, Pingda Ren, Xia Wang, Taebo Sim, Yongping Xie, Xing Wang, Gang Xia, et al. A general strategy for creating “inactive-conformation” abl inhibitors. *Chemistry & biology*, 13(7):779–786, 2006.
- [24] 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.
- [25] David Aranda-García, Tomasz Maciej Stepniewski, Mariona Torrens-Fontanals, Adrian García-Recio, Marta Lopez-Balastegui, Brian Medel-Lacruz, Adrián Morales-Pastor, Alejandro Peralta-García, Miguel Dieguez-Eceolaza, David Sotillo-Nuñez, et al. Large scale investigation of gpcr molecular dynamics data uncovers allosteric sites and lateral gateways. *Nature communications*, 16(1):2020, 2025.
- [26] 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.
- [27] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.
- [28] Sichun Yang, José N Onuchic, and Herbert Levine. Effective stochastic dynamics on a protein folding energy landscape. *The Journal of chemical physics*, 125(5), 2006.
- [29] J Copperman and MG Guenza. Predicting protein dynamics from structural ensembles. *The Journal of chemical physics*, 143(24), 2015.
- [30] Robert Zwanzig. Memory effects in irreversible thermodynamics. *Physical Review*, 124(4):983, 1961.
- [31] Yulia Rubanova, Ricky TQ Chen, and David K Duvenaud. Latent ordinary differential equations for irregularly-sampled time series. *Advances in neural information processing systems*, 32, 2019.

- [32] Dhruv Kumar Chaurasiya and Athi N Naganathan. Diverse conformational ensembles define the shared folding-allosteric landscapes of protein kinases. *Biophysical Journal*, 125(2):557–571, 2026.
- [33] Andrew R Leach. *Molecular modelling: principles and applications*. Pearson education, 2001.
- [34] KN Trueblood, H-B Bürgi, H Burzlaff, JD Dunitz, CM Gramaccioli, HH Schulz, U Shmueli, and SC Abrahams. Atomic displacement parameter nomenclature. report of a subcommittee on atomic displacement parameter nomenclature. *Foundations of Crystallography*, 52(5):770–781, 1996.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TEMPONCSIAppendix.pdf](#)