EVORESOLVE: PRESERVING EVOLUTIONARY HETEROGENEITY IN PROTEIN CONFORMATION PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Evolution preserves a protein family's capacity for multiple functional conformations, leaving partially distinct structural constraints across its homologs. However, AlphaFold3 (AF3)-style folding models merge these signals by averaging MSA-row outer products, leaving the folding model with single evolutionary summary. Therefore, we introduce EVORESOLVE, which learns multiple evolutionary summaries before MSA aggregation. Motivated by the richer row-wise variation observed in multi-conformation proteins, EVORESOLVE organizes homolog evidence into compact, query-conditioned branches using sliced wasserstein transport. Multi-conformation training thus supervises the new aggregation process, encouraging branches to emphasize different evidence in support of alternative states. The resulting representations drive a shared folding backbone, producing multiple predictions from one searched MSA without target-specific alignment editing. Across various conformational benchmarks, EVORESOLVE achieves the lowest reported RMSD among the compared methods, indicating that it can retain state-relevant differences while preserving the evolutionary constraints across homologs.

1 INTRODUCTION

Evolution selects protein sequences for functional conformational repertoires (Juritz et al., 2013; Dishman et al., 2021). Across the whole evolutionary tree, members of the same protein family may differ in their preferred conformations or functions. Thankfully, multiple sequence alignment(MSA) collects these homologous sequences and records their patterns of conservation and correlated substitutions. That's why MSA input provides folding model with structural evidence beyond the query sequence alone (Morcos et al., 2011; Jumper et al., 2021; Abramson et al., 2024).

However, for protein families with multiple experimentally observed conformations, information behind MSA could be more difficult to interpret. Instead of a single dominant residue contact, alternative states exhibits diverse coupling representing each possible evolutionary directions. Therefore, a full-family MSA can superimpose multiple state-associated coupling patterns, some of which become detectable only at the subfamily level (Sfriso et al., 2016; Schafer & Porter, 2023). This mixture makes multi-conformation folding fundamentally different from recovering one consensus structure.

Modern folding models receive this heterogeneous population of homologs, but compress it into a single MSA-derived representation. In AF3, outer product mean (OPM) performs this compression by averaging row-wise outer products across the MSA (Abramson et al., 2024). As shown in Figure 2, multi-conformation protein families contain substantially greater and more multidirectional variation across homologs than single-conformation controls. We quantify this information mixture in Section 4.1.

One strategy for mitigating this collapse is to perturb or restructure the MSA input. This *input-space MSA editing* paradigm clusters, subsamples, or masks homologs to expose alternative evolutionary signals, but relies on discrete, target-dependent MSA operations (Del Alamo et al., 2022; Lazou et al., 2024; Wayment-Steele et al., 2024). A second strategy, *structure-space diversification*, perturbs latent features or generative trajectories after evolutionary aggregation, leaving the alloca-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

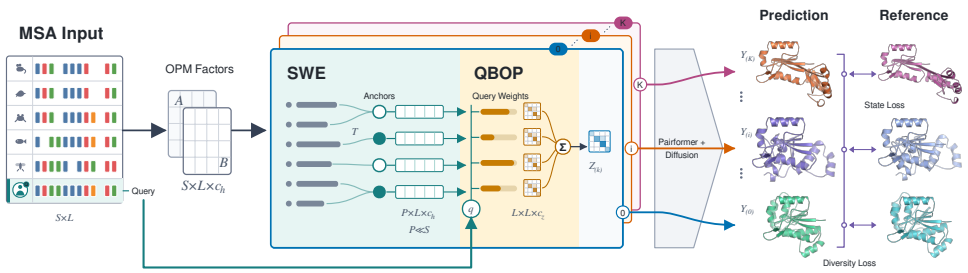

Figure 1: Overview of the EVORESOLVE's architecture. Branch-specific sliced Wasserstein embedding (SWE) compresses MSA-row features into pseudo-MSAs, and query-biased outer product (QBOP) aggregation constructs branch-specific pair representations. A shared Pairformer and diffusion module generate structural predictions, guided by state and diversity supervision.

tion of homolog evidence outside the learning problem (Lee et al., 2026; Richman et al., 2026). Instead, we introduce a *representation-space evolutionary routing* paradigm that learns multiple MSA-derived pair representations based on original MSA input within the folding model.

Under this paradigm, the central problem is not how to perturb one structure generator, but how to assign heterogeneous homolog evidence to distinct structural predictions. This requires two learned mappings: from MSA rows to evolutionary branches, and from those branches to experimentally observed conformational states. We use training data of multi-conformation structures to supervise both mappings end to end.

As illustrated in Figure 1, we instantiate representation-space routing with **EVORESOLVE**. Our latent-space outer product first uses branch-specific Sliced Wasserstein transport to map the MSA embedding from S rows to $P \ll S$ input-adaptive pseudo-rows, and only then constructs outer products. Within each branch, query-biased outer product (QBOP) weights these pseudo-rows by query relevance and produces an AF3-compatible residual pair representation. Permutation-invariant matching, state supervision, and ground-truth-adaptive diversity supervision align the learned evolutionary branches with distinct conformations without imposing arbitrary separation.

At inference, EVORESOLVE maps one searched MSA to a set of state-aligned structural predictions within a shared folding model. Its transport coupling exposes original-row-to-branch and row-to-anchor assignments, while QBOP provides query-conditioned weights over the latent pseudo-rows. Across multi-conformation evaluation benchmarks, EVORESOLVE leads baselines with either input-space editing or edited structure-space perturbing.

Our contributions are summarized as follows:

- We expose an evolutionary information bottleneck in AF3-style folding model by showing that multi-conformation proteins contain systematically richer and higher-rank variation across MSA-row outer-product representations.
- We introduce evolutionary routing and latent-space outer-product module that converts one searched MSA into compact, branch-specific evolutionary representations.
- We develop an end-to-end multi-conformation folding model in which experimental structures supervise homolog routing, evolutionary evidence weighting, and branch-to-state alignment.

2 PRELIMINARY

2.1 ALPHAFOLD3'S MSA MODULE AND THE OUTER PRODUCT MEAN

Let $X \in \mathcal{A}^{S \times L}$ denote a raw multiple sequence alignment (MSA) containing S aligned sequences of length L and vocabulary $\mathcal{A}$. AF3 (Abramson et al., 2024) first maps every categorical MSA entry to a continuous representation:

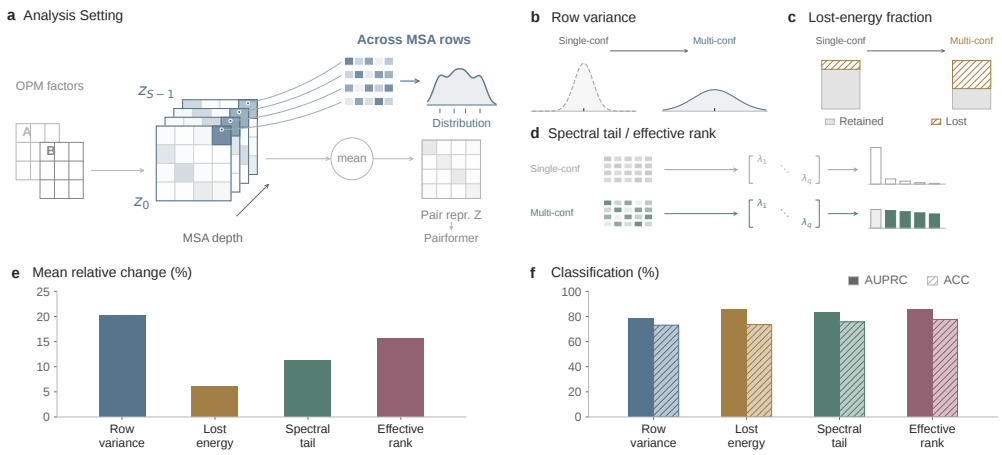

Figure 2: Characterizing row-wise variation lost through OPM aggregation. **(a)** Analysis of outer-product representations across MSA rows before mean aggregation. **(b–d)** Illustration of row variance, lost-energy fraction, spectral tail, and effective rank for single- and multi-conformation proteins. **(e)** Mean relative increases in these statistics for multi-conformation datasets compared with single-conformation controls. **(f)** Classification performance measured by AUPRC and accuracy using each statistic independently.

$$X \in \mathcal{A}^{S\times L} \xrightarrow{\text{MSAEmbed}} M = [m_{si}] \in \mathbb{R}^{S\times L\times c_m}, \qquad m_{si} \in \mathbb{R}^{c_m}. \tag{1}$$

The MSA representation is layer-normalized(LN) and projected by two independent bias-free linear layers to obtain the left and right OPM factors:

$$a_{si}, b_{si} = W_a\,\mathrm{LN}(m_{si}), W_b\,\mathrm{LN}(m_{si}) \in \mathbb{R}^{c_h} \qquad A, B = [a_{si}], [b_{si}] \in \mathbb{R}^{S\times L\times c_h} \tag{2}$$

For each MSA row s and residue pair (i, j), OPM constructs a row-wise outer product and then averages these outer products over the complete MSA-row dimension:

$$\mathcal{O}_{ij} = \frac{1}{S}\sum_{s=0}^{S-1}(a_{si} \otimes b_{sj}) \in \mathbb{R}^{c_h\times c_h} \qquad \mathcal{O} = [\mathcal{O}_{ij}] \in \mathbb{R}^{L\times L\times c_h\times c_h} \tag{3}$$

The mean in Eq. equation 3 is the central compression step of OPM. By averaging among MSA rows, this reduction makes full-MSA processing tractable, but it also replaces the distribution of row-specific coevolutionary patterns with a single mean. Consequently, distinct MSA subfamilies or conformational modes can be merged or mutually cancelled in later pairformer and diffusion modules.

$$Y = \mathrm{Diffusion}(..., \mathrm{Pairformer}(..., \mathrm{Linear}_{\mathrm{out}}(\mathcal{O}) \in \mathbb{R}^{L\times L\times c_z} \quad) \quad) \in \mathbb{R}^{N_{\mathrm{atom}}\times 3} \tag{4}$$

2.2 Query-Biased Outer Product

Query-biased outer product (Akiyama et al., 2025) replaces the uniform averaging in OPM with a weighted sum over MSA rows. Each weight λ_s is obtained from a query-conditioned attention score:

$$o_{ij} = \sum_{s=0}^{S-1}\lambda_s\,(a_{si} \otimes b_{sj}) \qquad \lambda_s = \mathrm{Softmax}\Big(\frac{1}{L\sqrt{d_q}}\sum_{i=1}^{L}\Big[(q_i^{(1)})^\top k_{si}^{(1)} - \theta(q_i^{(2)})^\top k_{si}^{(2)}\Big] \quad) \tag{5}$$

where θ is a learned scaling factor and the softmax is applied across the MSA-row dimension. Consequently, λ_s represents the relevance of row s to the query, with $\sum_s \lambda_s = 1$. The queries are

constructed from the query-row embeddings, while the keys are constructed from the embeddings of the MSA row being scored:

$$(q_i^{(1)}, q_i^{(2)}) = W_q \,\mathrm{LN}(m_{0i}) \qquad (k_{si}^{(1)}, k_{si}^{(2)}) = W_k \,\mathrm{LN}(m_{si}) \tag{6}$$

Thus, the attention pathway compares each homologous sequence with the query and uses the resulting score to weight its contribution to the outer product.

2.3 Sliced Wasserstein Embedding

Sliced Wasserstein embedding (SWE) maps a variable-size set to a fixed-size, permutation-invariant representation while retaining more distributional information than arithmetic averaging (Naderi-Alizadeh & Singh, 2026). Let

$$M = \{m_s\}_{s=0}^{S-1}, \qquad N = \{n_t\}_{t=0}^{\mathcal{T}-1}, \qquad W = \{w_r\}_{r=0}^{R-1}, \qquad m_s, n_t, w_r \in \mathbb{R}^d, \tag{7}$$

where M is the set of input MSA-row embeddings, N is a fixed-size anchor set with generally $\mathcal{T} \ll S$, and W is a set of unit slicing directions. SWE produces

$$E = \mathrm{SWE}(M; N, W) = [e_{tr}] \in \mathbb{R}^{\mathcal{T}\times R}, \qquad \mathbb{R}^{S\times d} \xrightarrow{\mathrm{SWE}} \mathbb{R}^{\mathcal{T}\times R}. \tag{8}$$

Thus, the variable MSA-row dimension S is replaced by the fixed support dimension $\mathcal{T}$. For each slicing direction w_r, let π_r and σ_r denote the permutations that sort the projected MSA embeddings and anchors, respectively:

$$w_r^\top m_{\pi_r(0)} \le \cdots \le w_r^\top m_{\pi_r(S-1)}, \qquad w_r^\top n_{\sigma_r(0)} \le \cdots \le w_r^\top n_{\sigma_r(\mathcal{T}-1)}. \tag{9}$$

Let $\mathrm{QInterp}_{S\to\mathcal{T}}$ denote rank-preserving quantile interpolation from an ordered sequence of S projected values to $\mathcal{T}$ equally spaced quantile values. The complete mapping from the sorted projections to the SWE output is

$$e_{\sigma_r(t),r} = \left[\mathrm{QInterp}_{S\to\mathcal{T}}\left(\left\{w_r^\top m_{\pi_r(s)}\right\}_{s=0}^{S-1}\right)\right]_t - w_r^\top n_{\sigma_r(t)}, \qquad \begin{smallmatrix} t=0,\ldots,\mathcal{T}-1, \\ r=0,\ldots,R-1. \end{smallmatrix} \tag{10}$$

The interpolation operator first summarizes the S ordered source projections at T quantile positions. The t-th resampled source quantile is then matched with the t-th ordered anchor, and their signed displacement along w_r is stored at the original anchor index $\sigma_r(t)$.

3 EvoResolve

3.1 Architecture Overview

Our central motivation is to recover K distinct coevolutionary signals from MSA outer-product features and use the resulting latent modes to predict alternative protein conformations. To reduce memory consumption while preserving heterogeneous evolutionary patterns, we introduce K branch-specific SWE modules before the outer-product operation. Each SWE branch compresses the original MSA representation into a fixed-depth pseudo-MSA. The outer products of the pseudo-MSA rows are subsequently aggregated by a query-biased outer-product (QBOP) module. QBOP compares each pseudo-MSA row with the original query representation and learns its contribution to the branch-specific pair representation. Each branch predicts a new pair embedding $Z^{(k)}$, which is used as input before the shared Pairformer. Starting from the two OPM factors $A, B \in \mathbb{R}^{S\times L\times c_h}$ defined in Section 2.1, the proposed multi-conformation pathways are summarized as:

$$(A, B) \xrightarrow{\mathrm{SWE}} \left\{\left(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}\right)\right\}_{k=1}^{K} \xrightarrow{\mathrm{QBOP}} \left\{Z^{(k)}\right\}_{k=1}^{K} \xrightarrow{\mathrm{Pairformer + Diffusion}} \left\{Y^{(k)}\right\}_{k=1}^{K}. \tag{11}$$

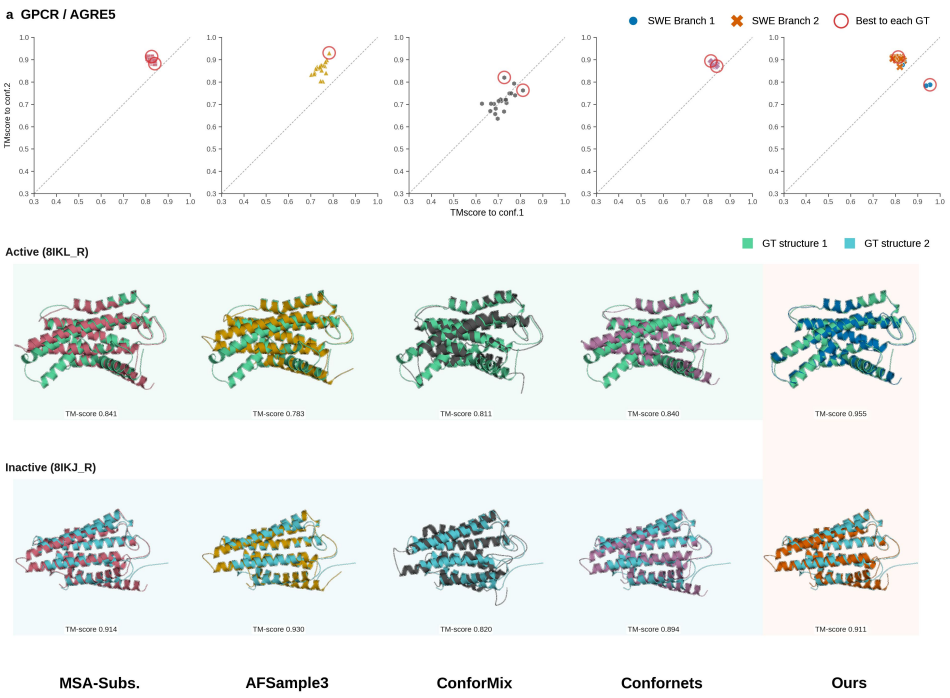

Figure 3: Conformation prediction for protein AGRE5 from the GPCR benchmark. EVORESOLVE achieves the most accurate prediction of the active conformation among the compared methods, with its two structural branches capturing the active and inactive states, respectively (reference structures: PDB entries 8IKL_R and 8IKJ_R).

3.2 PRE-OUTER-PRODUCT: SWE AGGREGATION

Starting from the OPM factors $A, B \in \mathbb{R}^{S\times L\times c_h}$, we construct one outer-product-aware descriptor for each MSA row:

$$d_s = f_{\text{desc}}\left(\frac{1}{L}\sum_{i=1}^{L} a_{si}, \frac{1}{L}\sum_{i=1}^{L} b_{si}\right) \in \mathbb{R}^{D_h}, \qquad f_{\text{desc}} : \mathbb{R}^{c_h} \times \mathbb{R}^{c_h} \to \mathbb{R}^{D_h}, \tag{12}$$

where D_h is the hidden dimension of the SWE representation. The descriptor and anchor construction are detailed in Appendix A.2. For branch k, the three arguments of SWE are

$$M_{\text{row}} = \{d_s\}_{s=0}^{S-1} \in \mathbb{R}^{S\times D_h}, \qquad N^{(k)} = \{n_{kp}\}_{p=1}^{P} \in \mathbb{R}^{P\times D_h}, \qquad W = I_{D_h}. \tag{13}$$

The model contains K independent sets of learnable base anchors, one for each conformation branch, the branch-specific compression is

$$E^{(k)} = \text{SWE}\left(M_{\text{row}}; N^{(k)}, W\right), \qquad \left(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}\right) = \text{Decode}\left(E^{(k)}, A, B\right). \tag{14}$$

The decoder uses the source-to-anchor mapping T learned during SWE computation to pool the original OPM factors, as detailed in Appendix A.1. The decoded factors satisfy $\widetilde{A}^{(k)}, \widetilde{B}^{(k)} \in \mathbb{R}^{P\times L\times c_h}$. Consequently,

$$\mathbb{R}^{S\times L\times c_h} \xrightarrow{\text{SWE}+\text{Decode}} \mathbb{R}^{K\times P\times L\times c_h}, \qquad P \ll S. \tag{15}$$

3.3 DURING-OUTER-PRODUCT: QUERY-CONDITIONED SEQUENCE WEIGHTING

Following Section 2.2, we write the original query-biased outer product as a compact operator:

$$O = \text{QBOP}(A, B, q), \qquad q = [m_{0i}]_{i=1}^{L} \in \mathbb{R}^{L\times c_m}. \tag{16}$$

Here, QBOP uses the query-conditioned scores to weight MSA-row outer products, we adapt this operator to the SWE-compressed pseudo-MSA factors. For each conformational branch, the pseudo factors $(\widetilde{A}^{(k)}, \widetilde{B}^{(k)})$ are weighted by a QBOP-style attention module whose query descriptor is constructed from the OPM factors:

$$Z^{(k)} = \mathrm{Linear}_{\mathrm{out}}\left(\mathrm{QBOP}\left(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}, q\right)\right), \qquad q = [a_{0i}; b_{0i}]_{i=1}^{L} \in \mathbb{R}^{L\times 2c_h}. \tag{17}$$

Thus, instead of weighting the original S MSA rows, the adapted QBOP module weights the P SWE pseudo-rows and aggregates their outer-product features. Consequently,

$$\mathbb{R}^{K\times P\times L\times c_h} \xrightarrow{\mathrm{QBOP}+\mathrm{Linear}_{\mathrm{out}}} \mathbb{R}^{K\times L\times L\times c_z}, \qquad P \ll S. \tag{18}$$

3.4 Post-Outer-Product: Structure Guidance

Our training objective uses experimentally observed multi-conformation structures to teach the MSA module how to extract different coevolutionary signals from the same MSA input. Given one MSA, the model produces K branch-specific pair representations $\{Z^{(k)}\}_{k=1}^{K}$ and predicts K corresponding structures $\{Y^{(k)}\}_{k=1}^{K}$. The supervision encourages each branch to recover a valid conformational state, while preventing all branches from encoding the same pair-space mode.

State loss. Because branch identities are exchangeable, we first match predicted structures to ground-truth conformations with a detached assignment, and then supervise the assigned state using the frozen distogram head $\mathcal{D}$:

$$\pi^* = \arg\min_{\pi\in\mathfrak{S}_K} \frac{1}{K}\sum_{k=1}^{K} \mathcal{L}_{\mathrm{TM}}\left(Y^{(k)}, Y_{gt}^{\pi(k)}\right), \qquad \mathcal{L}_{\mathrm{state}}^{(k)} = \frac{\sum_{i,j}\xi_{ij}\,\mathrm{CE}\left(\mathcal{D}(Z_{ij}^{(k)}), y_{ij}^{\pi^*(k)}\right)}{\sum_{i,j}\xi_{ij}+\epsilon}. \tag{19}$$

Here $\xi_{ij} \in [0,1]$ is the variable-region pair weight and $y_{ij}^{\pi^*(k)}$ is the distance-bin label from the assigned ground-truth conformation.

Diversity loss. The diversity loss sets the required branch separation from the structural difference between the ground-truth conformations. Let $t_{\mathrm{gt}} \in [0,1]$ be the TM similarity between two ground-truth states. The constants τ_{full} and τ_{ignore} are hyperparameters: below τ_{full} the two states are treated as fully different, while above τ_{ignore} no diversity margin is required. With maximum target distance η, we define

$$d_{\mathrm{target}} = \eta\ \mathrm{clip}\left(\frac{\tau_{\mathrm{ignore}} - t_{\mathrm{gt}}}{\tau_{\mathrm{ignore}} - \tau_{\mathrm{full}}}, 0, 1\right). \tag{20}$$

Let CLN denote channel-wise LayerNorm over the c_z pair channels, used only to measure representation distance independently of feature scale. For branch pair (k,l),

$$d_{\mathrm{pair}}^{(k,l)} = \sqrt{\frac{1}{L^2 c_z}\sum_{i,j}\left\|\mathrm{CLN}(Z_{ij}^{(k)}) - \mathrm{CLN}(Z_{ij}^{(l)})\right\|_2^2 + \epsilon}, \qquad \mathcal{L}_{\mathrm{div}}^{(k,l)} = \left[d_{\mathrm{target}} - d_{\mathrm{pair}}^{(k,l)}\right]_+^2. \tag{21}$$

Thus, structurally distinct conformations impose stronger separation between learned MSA-derived pair representations, whereas similar conformations impose little or no separation.

Final objective. The native structure loss $\mathcal{L}_{\mathrm{native}}^{(k)}$ is the original AF3 folding objective evaluated against the assigned ground-truth state $Y_{gt}^{\pi^*(k)}$. For general K, the diversity term is averaged over all unordered branch pairs:

$$\mathcal{L}_{\mathrm{final}} = \lambda_{\mathrm{native}}\frac{1}{K}\sum_{k=1}^{K}\mathcal{L}_{\mathrm{native}}^{(k)} + \lambda_{\mathrm{state}}\frac{1}{K}\sum_{k=1}^{K}\mathcal{L}_{\mathrm{state}}^{(k)} + \lambda_{\mathrm{div}}\frac{2}{K(K-1)}\sum_{1\le k<l\le K}\mathcal{L}_{\mathrm{div}}^{(k,l)}. \tag{22}$$

4 Experiments

Motivated by the information-mixture view developed above, we first characterize the heterogeneous row-wise coevolutionary patterns that OPM mean aggregation compresses into a single pair representation. We then evaluate EvoResolve on multi-conformation prediction and intrinsically disordered proteins (IDPs) before ablation study.

Table 1: Comparison of EVORESOLVE and baseline methods across six protein conformational benchmarks using global C_α RMSD (Å). Lower values indicate better performance. The best result in each column is shown in **bold**.

Method	Membrane	CrypticPocket	DomainMotion	Foldswitching	GPCR	Kinases
Edited MSA Input						
AF-cluster	4.54	2.87	3.79	3.47	3.06	3.44
MSA-subsampling	3.35	1.72	2.80	2.30	1.88	2.11
AFSample3	4.08	2.37	3.24	3.28	2.74	2.57
Raw MSA Input						
Protenix-v1	3.31	1.77	3.07	2.62	1.89	2.11
BioEmu	3.84	2.34	3.22	2.86	2.77	2.42
Conformix	3.79	2.47	3.37	3.06	3.07	2.45
ConforNets	3.38	1.77	3.05	2.67	1.89	2.10
EVORESOLVE	**3.20**	**1.59**	**2.75**	**2.19**	**1.80**	**1.99**

4.1 OPM AGGREGATION ANALYSIS

Setup. The multi-conformation cohort spans membrane transport, GPCR activation, kinase switching, domain motion, cryptic-pocket opening, and fold switching, and is compared with pooled CASP15 (Kryshtafovych et al., 2023) and CASP16 (Yuan et al., 2026) single-conformation controls. To make fair comparison, we generate all MSAs using the same search protocol (Team et al., 2026). From the projected OPM features immediately before row averaging, we compute four complementary statistics: row variance, lost-energy fraction, spectral tail, and effective rank. Their definitions, derivations, and the complete evaluation protocol are provided in Appendix D.3.

Results. As shown in Figure 2 and Table 7, every multi-conformation category exhibits higher values than the pooled controls for all four statistics. The increase in row variance indicates stronger absolute dispersion of homolog-specific pair features around their mean. The larger spectral tail further shows that this variation is not concentrated along one dominant direction but extends across additional spectral directions. Consistently, the higher effective rank corresponds to a flatter eigenvalue spectrum in which more directions contribute substantially. Thus, multi-conformation proteins exhibit not only greater row-wise variation but also a more multidirectional organization of that variation. Each statistic also discriminates multi-conformation proteins from the controls when used independently.

4.2 UNCONDITIONAL MULTI-CONFORMATION PREDICTION

Task and evaluation protocol. We evaluate the ability of each method to recover both experimentally observed conformations of a protein rather than repeatedly predicting only one state. Training data are constructed from CoDNaS (Monzon et al., 2013; 2016), with conformer selection, quality control, and train–test sequence decontamination detailed in Appendix C.1. We implement EVORESOLVE by fine-tuning a Protenix-v1 backbone on these training conformation structures (Team et al., 2026). At inference time, complete retrieved MSA is supplied to the model without target-specific clustering, sub-sampling, or masking. We compute the global Cα RMSD, measured in Å after structural superposition, between each prediction and each of reference structures.

Evaluation datasets. We evaluate on complementary benchmarks spanning diverse conformational changes: **Membrane** covers inward- and outward-facing transporter states (Del Alamo et al., 2022); **CrypticPocket** contains ligand-induced pocket opening (Cimermancic et al., 2016; Meller et al., 2023); **DomainMotion** focuses on large-scale hinge and domain rearrangements (Lee et al., 2025; Lewis et al., 2025); **Foldswitching** contains proteins with two experimentally observed folds (Porter & Looger, 2018; Schafer & Porter, 2023); **GPCR** and **Kinases** cover functionally important activation-state transitions (Kooistra et al., 2021; Kanev et al., 2021). Detailed information provided in Table 4.

Table 2: Evaluation of predicted intrinsically disordered protein ensembles against Protein Ensemble Database references. Mean MAEs (mean absolute error) are evaluated for each evaluation metric. The best result in each column is shown in **bold**, and the second-best result is underlined.

Method	PwD.	Rg.	Cp.	R_{ee}	Shape	Flory-ν	SASA	ens-dRMS
BioEmu	**7.748**	4.794	0.657	23.121	3.612	0.184	613.042	11.912
ConforMix	7.859	4.657	0.596	23.519	3.667	0.179	**499.292**	11.800
ConforNets	8.165	4.878	0.238	23.870	3.582	0.166	647.169	12.079
AF-Cluster	8.735	4.970	0.292	26.540	3.702	0.191	595.640	12.945
MSA-Subs.	8.290	4.527	**0.190**	24.338	3.860	0.167	531.687	12.465
AFSample3	8.416	4.671	0.216	24.900	3.721	0.168	517.937	12.783
EVORESOLVE	7.799	**4.230**	0.211	**22.939**	**3.482**	**0.159**	517.773	**11.757**

Baselines. Among methods that do not apply MSA editing, **ConforNets** diversifies predictions through channel-wise affine modulation of pre-Pairformer pair representations (Lee et al., 2026). **BioEmu** generates conformational ensembles using a diffusion model trained with equilibrium structural data (Lewis et al., 2025), whereas **ConforMix** steers inference through diffusion-time guidance, filtering, and enhanced sampling (Richman et al., 2026). Among methods that explicitly edit the MSA, **AF-Cluster** partitions homologous sequences into similarity-based clusters and folds the resulting sub-alignments independently (Wayment-Steele et al., 2024); **MSA-Subs.** randomly subsamples MSA rows to vary the available coevolutionary evidence (Del Alamo et al., 2022); and **AFSample3** randomly masks MSA columns searching for alternative structures (Kalakoti & Wallner, 2026). Further methodological and implementation details are summarized in Table 5.

Main results. As shown in Table 1, EVORESOLVE achieves the lowest RMSD on all evaluated benchmarks. Notably, although EVORESOLVE directly uses the complete, unedited MSA, it consistently outperforms methods that explicitly cluster, subsample, or mask the MSA to expose alternative conformational signals. This comparison demonstrates that our aggregation module learns to identify and selectively route heterogeneous coevolutionary evidence already present in the full alignment. Consequently, EVORESOLVE provides a learned alternative to target-specific MSA manipulation and makes more effective use of the available evolutionary information without requiring an external selection over edited alignments. Case study shown in Figure 3.

4.3 PREDICTION OF INTRINSICALLY DISORDERED PROTEINS

Dataset and evaluation. We further evaluate EVORESOLVE on intrinsically disordered proteins (IDPs), which generally do not adopt a single well-defined tertiary structure (Oldfield & Dunker, 2014; Trivedi & Nagarajaram, 2022). Following Lu et al. (2025), we compare generated predictions with representative structures from the Protein Ensemble Database (Lazar et al., 2021) using mean absolute error(MAE) in ranges of evaluation metrics: local pair geometry (PwD. and Cp.), global size and shape (R_g, R_{ee}, and Shape), polymer scaling (Flory-ν), solvent exposure (SASA), and ensemble-level distance consistency (ens-dRMS). Detailed definitions are provided in Appendix D.4 (Lazar et al., 2020; Shadman et al., 2024; Shrake & Rupley, 1973). We compare EVORESOLVE against the same baselines used above.

Results. As shown in Table 2, EVORESOLVE achieves the best result on most metrics and ranks second on all others. These results indicate that EVORESOLVE captures complementary local and global characteristics of IDP conformations rather than improving only a single property.

4.4 ABLATION STUDY

We conduct an ablation study to disentangle the contributions of the Sliced-Wasserstein Embedding (SWE) module and the Query-Biased Outer Product (QBOP) module. All variants use the same Protenix-v1 backbone, training objective and other experimental setttings. We compare the following variants: **Protenix-v1**, the frozen backbone without a multiconformation router; **SWE-only**, SWE compresses the input MSA into two branch-specific pseudo-MSAs, which are passed directly to the outer product mean operation using uniform aggregation; **QBOP-only**, QBOP learns

Table 3: Ablation study of the SWE and QBOP modules on the Membrane, CrypticPocket, and DomainMotion benchmarks. The best result in each column is shown in **bold**.

Method	Membrane		CrypticPocket		DomainMotion	
	TM-score ↑	RMSD ↓	TM-score ↑	RMSD ↓	TM-score ↑	RMSD ↓
Protenix-v1	0.790	3.29	0.890	1.75	0.742	3.04
SWE-only	0.832	3.23	0.901	1.63	0.761	2.78
QBOP-only	0.829	3.26	0.901	1.62	0.764	2.75
EVORESOLVE	**0.840**	**3.20**	**0.902**	**1.60**	**0.772**	**2.74**

branch-specific query-dependent weights over the original MSA rows; and EVORESOLVE, SWE first produces branch-specific pseudo-MSAs, after which QBOP adaptively weights their evolutionary evidence before pair-representation construction.

From Table 3, removing either SWE or QBOP consistently degrades performance relative to the complete model. This result shows that SWE provides useful evidence separation while QBOP preserves query-dependent weighted aggregation. Both module contributes to structured representation selection and evolutionary routing.

5 RELATED WORKS

MD(Molecular Dynamic)-data Based Conformation Prediction MD-based approaches use explicit simulations or MD-derived conformational data to improve the physical grounding and coverage of protein conformational prediction. AlphaPPImd trains a Transformer on MD trajectories of protein–protein complexes to generate unsampled conformations beyond the simulated time window (Wang et al., 2024a). AI^{2}BMD learns a machine-learning force field from quantum-level conformational data and uses it to accelerate ab initio-like MD simulations of protein folding and unfolding (Wang et al., 2024b). BioEmu combines large-scale MD trajectories, experimental structures, and stability measurements to emulate protein equilibrium ensembles (Lewis et al., 2025). However, these methods remain dependent on the coverage, quality, and physical assumptions of their simulation data or learned force fields, while extensive MD-based data generation can still be computationally demanding.

Generative Models for Conformation Prediction Generative approaches make the sampling architecture itself the main source of conformational diversity, typically by learning a sequence-conditioned distribution over protein structures. EigenFold uses diffusion over protein eigenmodes to generate structure ensembles from a protein sequence (Jing et al., 2023). ConfDiff introduces force-guided SE(3) diffusion to improve the physical fidelity of generated conformations (Wang et al., 2024c). Mac-Diff combines locality-aware modal alignment with diffusion and uses PDB pretraining followed by MD fine-tuning to generate diverse conformational ensembles (Wang et al., 2026). Despite their efficient sampling, these methods face an input-information bottleneck, as their conformational coverage remains constrained by the structural information encoded in conditioning inputs such as MSAs and templates.

6 CONCLUSION

We introduced EVORESOLVE, a branch-specific architecture that transforms one searched MSA into multiple compact evolutionary representations before outer-product aggregation. Within EVORESOLVE, sliced Wasserstein embedding preserves distributional variation across homologs, while query-biased weighting and structure-guided supervision associate different branches with alternative conformations. The resulting model performs strongly across diverse multi-conformation benchmarks and improves characterization of intrinsically disordered proteins. Overall, by constructing multiple evolutionary representations before homolog evidence is aggregated, EVORESOLVE provides promising direction to extend folding model from single-structure toward multiple conformation prediction.

AI USE STATEMENT

In this work, we used OpenAI Codex to assist with language polishing, literature search, and code debugging. All AI-assisted text, references, and code were manually reviewed by the authors for accuracy, relevance, and correctness. References suggested through AI-assisted searches were verified against their original sources, and AI-assisted code was inspected and tested before being used in our experiments.

Generative AI tools were not used to generate the research ideas, network architecture, training configurations, experimental design, or selection of evaluation metrics presented in this work. They were also not used to draw scientific conclusions from the experimental results. Tasks outside those described above were either performed directly by the authors or were not applicable to this work.

The authors have reviewed all AI-assisted work and take full responsibility for the final content of this paper, including all text, claims, references, code, and other artifacts produced with the assistance of generative AI.

ETHICS STATEMENT

This work develops methods for predicting and analyzing multiple protein conformations, with the aim of supporting and accelerating scientific discovery. It does not involve human participants, human-derived private information, clinical interventions, or animal experiments. All data used in this study were obtained from established scientific data sources and were processed under their applicable terms of use. The datasets were carefully curated and filtered to minimize overlap and information leakage between training and evaluation data.

REPRODUCIBILITY STATEMENT

To support reproducibility, Section 3 describes the architecture and training objectives of EVORESOLVE , with additional mathematical details in Appendix A. Appendix C documents the training data construction, conformer selection, train–test filtering, and evaluation datasets. Appendix D provides experimental settings, hyperparameters, baseline information, and evaluation details. Since implementation are subject to institutional confidentiality, future release is contingent on approval.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.

Yo Akiyama, Zhidian Zhang, Milot Mirdita, Martin Steinegger, and Sergey Ovchinnikov. Scaling down protein language modeling with msa pairformer. *bioRxiv*, pp. 2025–08, 2025.

Sarah Alamdari, Nitya Thakkar, Rianne Van Den Berg, Neil Tenenholtz, Robert Strome, Alan M Moses, Alex X Lu, Nicolò Fusi, Ava P Amini, and Kevin K Yang. Protein generation with evolutionary diffusion: sequence is all you need. *BioRxiv*, pp. 2023–09, 2023.

Peter Cimermancic, Patrick Weinkam, T Justin Rettenmaier, Leon Bichmann, Daniel A Keedy, Rahel A Woldeyes, Dina Schneidman-Duhovny, Omar N Demerdash, Julie C Mitchell, James A Wells, et al. Cryptosite: expanding the druggable proteome by characterization and prediction of cryptic binding sites. *Journal of molecular biology*, 428(4):709–719, 2016.

Diego Del Alamo, Davide Sala, Hassane S Mchaourab, and Jens Meiler. Sampling the conformational landscapes of transporters and receptors with alphafold2. *BioRxiv*, pp. 2021–11, 2021.

Diego Del Alamo, Davide Sala, Hassane S Mchaourab, and Jens Meiler. Sampling alternative conformational states of transporters and receptors with alphafold2. *elife*, 11:e75751, 2022.

Acacia F Dishman, Robert C Tyler, Jamie C Fox, Andrew B Kleist, Kenneth E Prehoda, M Madan Babu, Francis C Peterson, and Brian F Volkman. Evolution of fold switching in a metamorphic protein. *Science*, 371(6524):86–90, 2021.

Bowen Jing, Ezra Erives, Peter Pao-Huang, Gabriele Corso, Bonnie Berger, and Tommi Jaakkola. Eigenfold: Generative protein structure prediction with diffusion models. *arXiv preprint arXiv:2304.02198*, 2023.

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

Ezequiel Juritz, Nicolas Palopoli, Maria Silvina Fornasari, Sebastian Fernandez-Alberti, and Gustavo Parisi. Protein conformational diversity modulates sequence divergence. *Molecular biology and evolution*, 30(1):79–87, 2013.

Yogesh Kalakoti and Björn Wallner. Afsample2 predicts multiple conformations and ensembles with alphafold2. *Communications biology*, 8(1):373, 2025.

Yogesh Kalakoti and Björn Wallner. Afsample3: Generating and selecting multiple conformational states with alphafold3. *bioRxiv*, pp. 2026–01, 2026.

Georgi K Kanev, Chris de Graaf, Bart A Westerman, Iwan JP de Esch, and Albert J Kooistra. Klifs: an overhaul after the first 5 years of supporting kinase research. *Nucleic acids research*, 49(D1): D562–D569, 2021.

Albert J Kooistra, Stefan Mordalski, Gáspár Pándy-Szekeres, Mauricio Esguerra, Alibek Mamyrbekov, Christian Munk, György M Keserű, and David E Gloriam. Gpcrdb in 2021: integrating gpcr sequence, structure and function. *Nucleic Acids Research*, 49(D1):D335–D343, 2021.

Gil Koren, Sagi Meir, Lennard Holschuh, Haydyn DT Mertens, Tamara Ehm, Nadav Yahalom, Adina Golombek, Tal Schwartz, Dmitri I Svergun, Omar A Saleh, et al. Intramolecular structural heterogeneity altered by long-range contacts in an intrinsically disordered protein. *Proceedings of the National Academy of Sciences*, 120(30):e2220180120, 2023.

Andriy Kryshtafovych, Torsten Schwede, Maya Topf, Krzysztof Fidelis, and John Moult. Critical assessment of methods of protein structure prediction (casp)—round xv. *Proteins: Structure, Function, and Bioinformatics*, 91(12):1539–1549, 2023.

Tamas Lazar, Mainak Guharoy, Wim Vranken, Sarah Rauscher, Shoshana J Wodak, and Peter Tompa. Distance-based metrics for comparing conformational ensembles of intrinsically disordered proteins. *Biophysical Journal*, 118(12):2952–2965, 2020.

Tamas Lazar, Elizabeth Martínez-Pérez, Federica Quaglia, András Hatos, Lucía B Chemes, Javier A Iserte, Nicolás A Méndez, Nicolás A Garrone, Tadeo E Saldaño, Julia Marchetti, et al. Ped in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic acids research*, 49(D1):D404–D411, 2021.

Maria Lazou, Omeir Khan, Thu Nguyen, Dzmitry Padhorny, Dima Kozakov, Diane Joseph-McCarthy, and Sandor Vajda. Predicting multiple conformations of ligand binding sites in proteins suggests that alphafold2 may remember too much. *Proceedings of the National Academy of Sciences*, 121(48):e2412719121, 2024.

Minji Lee, Colin Kalicki, Minkyu Jeon, Aymen Qabel, Alisia Fadini, and Mohammed AlQuraishi. Confornets: Latents-based conformational control in openfold3. *ArXiv*, pp. arXiv–2604, 2026.

Myeongsang Lee, Joseph W Schafer, Jeshuwin Prabakaran, Devlina Chakravarty, Madeleine F Clore, and Lauren L Porter. Large-scale predictions of alternative protein conformations by alphafold2-based sequence association. *Nature communications*, 16(1):5622, 2025.

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.

Jiarui Lu, Xiaoyin Chen, Stephen Lu, Chence Shi, Hongyu Guo, Yoshua Bengio, and Jian Tang. Structure language models for protein conformation generation. In *International Conference on Learning Representations*, volume 2025, pp. 98645–98678, 2025.

Chang Ma, Haiteng Zhao, Lin Zheng, Jiayi Xin, Qintong Li, Lijun Wu, Zhi-Hong Deng, Yang Young Lu, Qi Liu, Sheng Wang, et al. Retrieved sequence augmentation for protein representation learning. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 1738–1767, 2024.

Artur Meller, Soumendranath Bhakat, Shahlo Solieva, and Gregory R Bowman. Accelerating cryptic pocket discovery using alphafold. *Journal of chemical theory and computation*, 19(14):4355, 2023.

Alexander Miguel Monzon, Ezequiel Juritz, María Silvina Fornasari, and Gustavo Parisi. Codnas: a database of conformational diversity in the native state of proteins. *Bioinformatics*, 29(19): 2512–2514, 2013.

Alexander Miguel Monzon, Cristian Oscar Rohr, María Silvina Fornasari, and Gustavo Parisi. Codnas 2.0: a comprehensive database of protein conformational diversity in the native state. *Database*, 2016:baw038, 2016.

Faruck Morcos, Andrea Pagnani, Bryan Lunt, Arianna Bertolino, Debora S Marks, Chris Sander, Riccardo Zecchina, José N Onuchic, Terence Hwa, and Martin Weigt. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proceedings of the National Academy of Sciences*, 108(49):E1293–E1301, 2011.

Navid NaderiAlizadeh and Rohit Singh. Evopool: Evolution-guided pooling of protein language model embeddings. *bioRxiv*, pp. 2026–02, 2026.

Pascal Notin, Mafalda Dias, Jonathan Frazer, Javier Marchena-Hurtado, Aidan N Gomez, Debora Marks, and Yarin Gal. Tranception: protein fitness prediction with autoregressive transformers and inference-time retrieval. In *International Conference on Machine Learning*, pp. 16990–17017. PMLR, 2022.

Christopher J Oldfield and A Keith Dunker. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annual review of biochemistry*, 83(1):553–584, 2014.

Lauren L Porter and Loren L Looger. Extant fold-switching proteins are widespread. *Proceedings of the National Academy of Sciences*, 115(23):5968–5973, 2018.

Roshan M Rao, Jason Liu, Robert Verkuil, Joshua Meier, John Canny, Pieter Abbeel, Tom Sercu, and Alexander Rives. Msa transformer. In *International conference on machine learning*, pp. 8844–8856. PMLR, 2021.

Daniel D Richman, Jessica Karaguesian, Carl-Mikael Suomivuori, and Ron Dror. Unlocking hidden biomolecular conformational landscapes in diffusion models at inference time. *Advances in neural information processing systems*, 38:167662–167682, 2026.

Joseph W Schafer and Lauren L Porter. Evolutionary selection of proteins with two folds. *Nature communications*, 14(1):5478, 2023.

Pedro Sfriso, Miquel Duran-Frigola, Roberto Mosca, Agustí Emperador, Patrick Aloy, and Modesto Orozco. Residues coevolution guides the systematic identification of alternative functional conformations in proteins. *Structure*, 24(1):116–126, 2016.

Hossain Shadman, Jesse D Ziebarth, Caleb E Gallops, Ray Luo, Zhengxin Li, Hai-Feng Chen, and Yongmei Wang. Map conformational landscapes of intrinsically disordered proteins with polymer physics quantities. *Biophysical Journal*, 123(10):1253–1263, 2024.

Andrew Shrake and John A Rupley. Environment and exposure to solvent of protein atoms. lysozyme and insulin. *Journal of molecular biology*, 79(2):351–371, 1973.

Protenix Team, Yuxuan Zhang, Chengyue Gong, Hanyu Zhang, Wenzhi Ma, Zhenyu Liu, Xinshi Chen, Jiaqi Guan, Lan Wang, Yanping Yang, et al. Protenix-v1: Toward high-accuracy open-source biomolecular structure prediction. *bioRxiv*, pp. 2026–02, 2026.

Rakesh Trivedi and Hampapathalu Adimurthy Nagarajaram. Intrinsically disordered proteins: an overview. *International journal of molecular sciences*, 23(22):14050, 2022.

Timothy Truong Jr and Tristan Bepler. Poet: A generative model of protein families as sequences-of-sequences. *Advances in Neural Information Processing Systems*, 36:77379–77415, 2023.

UniCorNE. UNICORNE: Unified conformational ensemble database. `https://unicornedb.org/`, a.

UniCorNE. UNICORNE documentation. `https://unicornedb.org/docs`, b. Benchmark leaderboard and UNICORNE_BENCH description.

Baoli Wang, Chenglin Wang, Jingyang Chen, Danlin Liu, Changzhi Sun, Jie Zhang, Kai Zhang, and Honglin Li. Conditional diffusion with locality-aware modal alignment for generating diverse protein conformational ensembles. *Nature Machine Intelligence*, 8(3):415–434, 2026.

Jianmin Wang, Xun Wang, Yanyi Chu, Chunyan Li, Xue Li, Xiangyu Meng, Yitian Fang, Kyoung Tai No, Jiashun Mao, and Xiangxiang Zeng. Exploring the conformational ensembles of protein–protein complex with transformer-based generative model. *Journal of chemical theory and computation*, 20(11):4469–4480, 2024a.

Tong Wang, Xinheng He, Mingyu Li, Yatao Li, Ran Bi, Yusong Wang, Chaoran Cheng, Xiangzhen Shen, Jiawei Meng, He Zhang, et al. Ab initio characterization of protein molecular dynamics with ai2bmd. *Nature*, 635(8040):1019–1027, 2024b.

Yan Wang, Lihao Wang, Yuning Shen, Yiqun Wang, Huizhuo Yuan, Yue Wu, and Quanquan Gu. Protein conformation generation via force-guided se (3) diffusion models. *arXiv preprint arXiv:2403.14088*, 2024c.

Hannah K Wayment-Steele, Adedolapo Ojoawo, Renee Otten, Julia M Apitz, Warintra Pitsawong, Marc Hömberger, Sergey Ovchinnikov, Lucy Colwell, and Dorothee Kern. Predicting multiple conformations via sequence clustering and alphafold2. *Nature*, 625(7996):832–839, 2024.

Rongqing Yuan, Jing Zhang, Andriy Kryshtafovych, R Dustin Schaeffer, Jian Zhou, Qian Cong, and Nick V Grishin. Casp16 protein monomer structure prediction assessment. *Proteins: Structure, Function, and Bioinformatics*, 94(1):86–105, 2026.

Yang Zhang and Jeffrey Skolnick. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4):702–710, 2004.

A Mathematical Explanations

A.1 SWE-Induced Balanced Coupling and Pseudo-MSA Decoding

This section makes the decoder in Eq. equation 14 explicit. Because the slicing matrix is $W = I_{D_h}$, the output of each SWE branch can be written as

$$E^{(k)} = \left[e_p^{(k)}\right]_{p=1}^{P} \in \mathbb{R}^{P\times D_h}, \qquad \mathcal{E} = \left\{E^{(k)}\right\}_{k=1}^{K}. \tag{23}$$

Each row $e_p^{(k)} \in \mathbb{R}^{D_h}$ is the axis-aligned quantile-displacement vector associated with anchor slot p in branch k. Hence it lies in the same coordinate system as the MSA-row descriptor d_s and can be used as a decoder key. The complete batched decoding map is

$$(\mathcal{E}, A, B) \xrightarrow{\text{Decode}} \left(T^\star[\mathcal{E}], \widetilde{A}, \widetilde{B}\right), \qquad \begin{array}{l} T^\star[\mathcal{E}] \in \mathbb{R}^{S\times K\times P}, \\ \widetilde{A}, \widetilde{B} \in \mathbb{R}^{K\times P\times L\times c_h}. \end{array} \tag{24}$$

The branchwise notation in Eq. equation 14 denotes the corresponding slice of this batched operator:

$$\text{Decode}\left(E^{(k)}, A, B\right) \equiv \left[\text{Decode}\left(\mathcal{E}, A, B\right)\right]_k. \tag{25}$$

We first normalize the source descriptors and the SWE decoder keys:

$$\widehat{d}_s = \frac{d_s}{\|d_s\|_2 + \epsilon_E}, \qquad \widehat{e}_p^{(k)} = \frac{e_p^{(k)}}{\|e_p^{(k)}\|_2 + \epsilon_E}, \tag{26}$$

where $\epsilon_E > 0$ is a numerical-stability constant. The SWE-induced source-to-slot cost is then

$$C_{s,kp}[\mathcal{E}] = \frac{1}{D_h}\left\|\widehat{d}_s - \widehat{e}_p^{(k)}\right\|_2^2. \tag{27}$$

Unlike an independently constructed descriptor-to-anchor cost, $C[\mathcal{E}]$ depends directly on the numerical SWE output. We obtain its balanced entropic transport realization as

$$T^\star[\mathcal{E}] = \text{Sinkhorn}_{\mathbf{a},\mathbf{b}}\left(-\frac{C[\mathcal{E}]}{\tau_T}\right) \in \mathbb{R}^{S\times K\times P}, \qquad a_s = 1, \qquad b_{kp} = \frac{S}{KP}. \tag{28}$$

Here $\text{Sinkhorn}_{\mathbf{a},\mathbf{b}}$ denotes alternating row and column normalization with prescribed marginals $\mathbf{a}$ and $\mathbf{b}$. With a finite number of normalization steps, the resulting coupling is row-stochastic and approximately balanced across the KP SWE slots:

$$\sum_{k=1}^{K}\sum_{p=1}^{P} T^\star_{s,kp}[\mathcal{E}] = 1, \qquad \sum_{s=0}^{S-1} T^\star_{s,kp}[\mathcal{E}] \approx \frac{S}{KP}. \tag{29}$$

Thus, $T^\star[\mathcal{E}]$ is not a second representation computed independently of SWE. It is a deterministic transport realization of $\mathcal{E}$, and gradients through the decoder propagate from the pseudo-MSA factors to the SWE embeddings.

For each branch, the induced coupling first defines a branch-level context. Let

$$\omega_{sk}[\mathcal{E}] = \sum_{p=1}^{P} T^\star_{s,kp}[\mathcal{E}]. \tag{30}$$

The corresponding OPM factor summaries are

$$\bar{a}_{ki} = \frac{\sum_{s=0}^{S-1}\omega_{sk}[\mathcal{E}]a_{si}}{\sum_{s=0}^{S-1}\omega_{sk}[\mathcal{E}] + \epsilon}, \qquad \bar{b}_{ki} = \frac{\sum_{s=0}^{S-1}\omega_{sk}[\mathcal{E}]b_{si}}{\sum_{s=0}^{S-1}\omega_{sk}[\mathcal{E}] + \epsilon}. \tag{31}$$

Given a learnable branch embedding u_k, the branch-specific factor residuals are

$$\delta a_{ki} = \alpha\tanh f_A\left([\bar{a}_{ki}, \bar{b}_{ki}, u_k]\right), \qquad \delta b_{ki} = \alpha\tanh f_B\left([\bar{b}_{ki}, \bar{a}_{ki}, u_k]\right). \tag{32}$$

Finally, the pseudo-MSA factors are decoded as the barycentric projections of the original OPM factors under the SWE-induced coupling:

$$\widetilde{a}_{kpi} = \frac{\sum_{s=0}^{S-1} T^{\star}_{s,kp}[\mathcal{E}] a_{si}}{\sum_{s=0}^{S-1} T^{\star}_{s,kp}[\mathcal{E}] + \epsilon} + \delta a_{ki}, \qquad \widetilde{b}_{kpi} = \frac{\sum_{s=0}^{S-1} T^{\star}_{s,kp}[\mathcal{E}] b_{si}}{\sum_{s=0}^{S-1} T^{\star}_{s,kp}[\mathcal{E}] + \epsilon} + \delta b_{ki}. \tag{33}$$

Eq. 27– 33 therefore provide the explicit realization of $\left(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}\right) = \mathrm{Decode}\left(E^{(k)}, A, B\right)$ used in Eq. 14. In particular, the learnable anchors $N^{(k)}$ affect the decoded factors through $E^{(k)} = \mathrm{SWE}\left(M_{\mathrm{row}}; N^{(k)}, W\right)$; the decoder does not construct a separate coupling directly from $N^{(k)}$. When factor refinement is disabled, the residual terms in Eq. 33 are omitted.

A.2 MSA-Row Descriptor and Anchor Construction

This section derives the following mapping:

$$\begin{matrix} A \in \mathbb{R}^{S \times L \times c_h}, \\ B \in \mathbb{R}^{S \times L \times c_h} \end{matrix} \quad \longrightarrow \quad \begin{matrix} M_{\mathrm{row}} \in \mathbb{R}^{S \times D_h}, \\ \left\{N^{(k)}\right\}_{k=1}^{K} \in \mathbb{R}^{K \times P \times D_h}. \end{matrix} \tag{34}$$

For each MSA row s, we first average the two OPM factors along the residue dimension:

$$\bar{a}_s = \frac{1}{L} \sum_{i=1}^{L} a_{si}, \qquad \bar{b}_s = \frac{1}{L} \sum_{i=1}^{L} b_{si}, \qquad \bar{a}_s, \bar{b}_s \in \mathbb{R}^{c_h}. \tag{35}$$

The descriptor network used in Eq. equation 12 is

$$\begin{aligned} \ell_s &= W_A \bar{a}_s, & r_s &= W_B \bar{b}_s, \\ h_s &= \mathrm{LayerNorm}\left(\ell_s \odot r_s + \frac{\ell_s + r_s}{2}\right), & d_s &= h_s + \mathrm{MLP}(h_s), \end{aligned} \tag{36}$$

where $W_A, W_B \in \mathbb{R}^{D_h \times c_h}$ and $\ell_s, r_s, h_s, d_s \in \mathbb{R}^{D_h}$. The MLP consists of two linear layers separated by a SiLU activation. The multiplicative term makes d_s sensitive to interactions between the two OPM factors.

To reduce the influence of redundant MSA rows, let $\widehat{d}_s = d_s / \|d_s\|_2$ and define

$$q_s = \frac{\left[1 + \sum_{u=0}^{S-1} \exp\left(\beta\left(\widehat{d}_s^{\top} \widehat{d}_u - 1\right)\right)\right]^{-1}}{\sum_{v=0}^{S-1} \left[1 + \sum_{u=0}^{S-1} \exp\left(\beta\left(\widehat{d}_v^{\top} \widehat{d}_u - 1\right)\right)\right]^{-1}}, \qquad c_M = \sum_{s=0}^{S-1} q_s d_s \in \mathbb{R}^{D_h}. \tag{37}$$

The model maintains learnable base anchors

$$N_0 = \left\{N_0^{(k)}\right\}_{k=1}^{K} \in \mathbb{R}^{K \times P \times D_h}. \tag{38}$$

An anchor-prediction network maps the MSA context to an input-dependent offset:

$$\Delta N = \mathrm{reshape}_{K \times P \times D_h}\left(g_{\mathrm{anchor}}(c_M)\right) \in \mathbb{R}^{K \times P \times D_h}. \tag{39}$$

The final branch-specific anchor sets are

$$\left\{N^{(k)}\right\}_{k=1}^{K} = \mathrm{Sort}_p\left(N_0 + \Delta N\right) \in \mathbb{R}^{K \times P \times D_h}. \tag{40}$$

Thus, the learnable base anchors provide independent parameters for the K branches, while their offsets adapt to the input MSA.

B Algorithms

The training and inference procedures of EvoResolve are summarized in Algo. 1 and 2, respectively.

Algorithm 1 Training EVORESOLVE for unconditional multi-conformation prediction

Require: Raw MSA X containing S sequences of target length L
Require: Non-MSA Protenix features Φ
Require: Ground-truth training conformations $\{Y_{\mathrm{gt}}^{(r)}\}_{r=1}^{K}$
Require: Branch anchors $\{N^{(k)}\}_{k=1}^{K}$, anchor dimension D_h, and diffusion noises $\{\varepsilon^{(k)}\}_{k=1}^{K}$
Require: Variable-region pair weights W_{var} and loss weights $\lambda_{\mathrm{native}}, \lambda_{\mathrm{state}}, \lambda_{\mathrm{div}}$

$M \leftarrow \mathrm{MSAEmbed}(X)$
$A, B \leftarrow \mathrm{OPMFactors}(M)$
$q \leftarrow [a_{0i}; b_{0i}]_{i=1}^{L}$
$M_{\mathrm{row}} \leftarrow \left[f_{\mathrm{desc}} \left(\frac{1}{L} \sum_{i=1}^{L} a_{si}, \frac{1}{L} \sum_{i=1}^{L} b_{si} \right) \right]_{s=0}^{S-1}$
for $k = 1, \ldots, K$ **do**
    $(E^{(k)}, T^{(k)}) \leftarrow \mathrm{SWE}\left(M_{\mathrm{row}}; N^{(k)}, I_{D_h}\right)$
    $(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}) \leftarrow \mathrm{Decode}\left(E^{(k)}, T^{(k)}, A, B\right)$
    $Z^{(k)} \leftarrow \mathrm{Linear}_{\mathrm{out}}\left(\mathrm{QBOP}\left(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}, q\right)\right)$
    $Y^{(k)} \leftarrow \mathrm{Diffusion}\left(\mathrm{Pairformer}(Z^{(k)}, \Phi), ...; \varepsilon^{(k)}\right)$
end for
$\pi^{*} \leftarrow \arg\min_{\pi \in \mathfrak{S}_K} \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}_{\mathrm{TM}}\left(\mathrm{sg}[Y^{(k)}], Y_{\mathrm{gt}}^{\pi(k)}\right)$ ▷ Detached branch matching
for $k = 1, \ldots, K$ **do**
    $\mathcal{L}_{\mathrm{native}}^{(k)} \leftarrow \mathrm{AF3Loss}\left(Y^{(k)}, Y_{\mathrm{gt}}^{\pi^{*}(k)}\right)$
    $\mathcal{L}_{\mathrm{state}}^{(k)} \leftarrow \mathrm{StateLoss}\left(Z^{(k)}, Y_{\mathrm{gt}}^{\pi^{*}(k)}, W_{\mathrm{var}}\right)$ ▷ Eq. equation 19
end for
for all $1 \le k < l \le K$ **do**
    $\mathcal{L}_{\mathrm{div}}^{(k,l)} \leftarrow \mathrm{DiversityLoss}\left(Z^{(k)}, Z^{(l)}, Y_{\mathrm{gt}}^{\pi^{*}(k)}, Y_{\mathrm{gt}}^{\pi^{*}(l)}\right)$ ▷ Eqs. equation 20–equation 21
end for
$\mathcal{L}_{\mathrm{final}} \leftarrow \frac{\lambda_{\mathrm{native}}}{K} \sum_{k=1}^{K} \mathcal{L}_{\mathrm{native}}^{(k)} + \frac{\lambda_{\mathrm{state}}}{K} \sum_{k=1}^{K} \mathcal{L}_{\mathrm{state}}^{(k)} + \frac{2\lambda_{\mathrm{div}}}{K(K-1)} \sum_{1 \le k < l \le K} \mathcal{L}_{\mathrm{div}}^{(k,l)}$
$\mathcal{L}_{\mathrm{final}}.\mathrm{backward}()$

EVORESOLVE performs unconditional multi-conformation prediction from a raw MSA X and the corresponding non-MSA Protenix features Φ. All branches receive the same X and Φ, while learned branch anchors $\{N^{(k)}\}_{k=1}^{K}$ induce branch-specific pair representations $\{Z^{(k)}\}_{k=1}^{K}$. Each pair representation is subsequently processed by the shared Pairformer and diffusion modules to produce a predicted conformation $Y^{(k)}$.

No conformation label or target structure is supplied as a model input. During training, the ground-truth conformations are used only for detached permutation matching and for computing the native, state, and diversity losses. The variable-region pair weights W_{var} restrict the state-sensitive supervision to conformationally variable residue pairs. During inference, neither branch matching nor structural supervision is required; the trained model directly produces the unordered prediction set $\{Y^{(k)}\}_{k=1}^{K}$.

C DATASETS

C.1 TRAINING DATASETS

Data source. Candidate conformations were collected from CoDNaS, which groups experimentally determined structures of the same protein obtained under different experimental conditions (Monzon et al., 2013; 2016).

Algorithm 2 Unconditional multi-conformation inference with EVORESOLVE

Require: Raw MSA X containing S sequences of target length L
Require: Non-MSA Protenix features Φ
Require: Trained branch anchors $\{N^{(k)}\}_{k=1}^{K}$ and parameters inside f_{desc}, Decode and QBOP
Require: Anchor dimension D_h, diffusion noises $\{\varepsilon^{(k)}\}_{k=1}^{K}$
Ensure: Unordered set of predicted conformations $\{Y^{(k)}\}_{k=1}^{K}$

$M \leftarrow \mathrm{MSAEmbed}(X)$
$A, B \leftarrow \mathrm{OPMFactors}(M)$
$q \leftarrow [a_{0i}; b_{0i}]_{i=1}^{L}$
$M_{\mathrm{row}} \leftarrow \left[f_{\mathrm{desc}} \left(\frac{1}{L} \sum_{i=1}^{L} a_{si}, \frac{1}{L} \sum_{i=1}^{L} b_{si} \right) \right]_{s=0}^{S-1}$
for $k = 1, \ldots, K$ **do**
    $(E^{(k)}, T^{(k)}) \leftarrow \mathrm{SWE}\left(M_{\mathrm{row}}; N^{(k)}, I_{D_h}\right)$
    $(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}) \leftarrow \mathrm{Decode}\left(E^{(k)}, T^{(k)}, A, B\right)$
    $Z^{(k)} \leftarrow \mathrm{Linear}_{\mathrm{out}}\left(\mathrm{QBOP}\left(\widetilde{A}^{(k)}, \widetilde{B}^{(k)}, q\right)\right)$
    $Y^{(k)} \leftarrow \mathrm{Diffusion}\left(\mathrm{Pairformer}(Z^{(k)}, \Phi), \Phi; \varepsilon^{(k)}\right)$
end for
return $\{Y^{(k)}\}_{k=1}^{K}$

Selection of structurally diverse conformers. For each target t, let $\mathcal{S}_t = \{s_1, \ldots, s_{n_t}\}$ denote its set of available experimental structures, and let K_t, where $2 \le K_t \le n_t$, denote the desired number of conformers. We compute the TM-score between each pair of structures (Zhang & Skolnick, 2004). The initial base set is formed by the pair with the lowest TM-score:

$$(i^\star, j^\star) = \arg \min_{1 \le i < j \le n_t} \mathrm{TM}(s_i, s_j), \qquad \mathcal{B}_t^{(2)} = \{s_{i^\star}, s_{j^\star}\}. \tag{41}$$

We retain targets for which the selected pair satisfies

$$0.5 \le \mathrm{TM}(s_{i^\star}, s_{j^\star}) \le 0.9. \tag{42}$$

Starting from this pair, we iteratively expand the base set. At each iteration, every remaining structure is compared with all structures already included in the base set. The structure with the lowest mean TM-score to the current base set is selected:

$$s_t^{(m+1)} = \arg \min_{s \in \mathcal{S}_t \setminus \mathcal{B}_t^{(m)}} \frac{1}{m} \sum_{b \in \mathcal{B}_t^{(m)}} \mathrm{TM}(s, b), \qquad m = 2, \ldots, K_t - 1. \tag{43}$$

The selected structure is then added to the base set:

$$\mathcal{B}_t^{(m+1)} = \mathcal{B}_t^{(m)} \cup \left\{ s_t^{(m+1)} \right\}. \tag{44}$$

The resulting set $\mathcal{B}_t^{(K_t)}$ contains the K_t structurally diverse conformers selected for target t.

Train–test sequence and structure filtering. A training target and all its conformers are excluded if its query has at least 80% global sequence identity to any test query or if any of its available conformers has a TM-score of at least 0.85 against any test reference structure.

C.2 EVALUATION DATASET

Summary of evaluation datasets in Table 4.

D EXPERIMENTAL DETAILS

D.1 HYPERPARAMETER SELECTION

The principal settings are summarized below.

Table 4: Summary of evaluation datasets for unconditional conformation prediction.

Dataset	# Targets	Mean length (aa)	Conformational Change	Reference
Membrane	15	479.33	Inward-open ↔ outward-open transporters	(Lee et al., 2026; Kalakoti & Wallner, 2025)
CrypticPocket	34	301.12	Apo ↔ holo cryptic-pocket opening	(Lee et al., 2026; Lewis et al., 2025; Cimermancic et al., 2016; Meller et al., 2023)
DomainMotion	22	349.77	Large-scale hinge or domain motion	(Lee et al., 2026; Lewis et al., 2025)
Foldswitching	15	280.00	Two distinct experimentally observed folds	(Lee et al., 2026; Porter & Looger, 2018)
GPCR	51	307.55	Inactive ↔ active; TM5–TM7 rearrangement	(Lee et al., 2026; Kooistra et al., 2021)
Kinases	20	333.50	DFG-in ↔ DFG-out; activation-loop motion	(Lee et al., 2026; Kanev et al., 2021)

Table 5: Detailed information and implementation links for baseline methods. The methods are grouped according to whether the original MSA is retained or edited.

Method	Setting	Main Mechanism	Links	Reference
Raw MSA Input				
ConforNets	Raw MSA	Channel-wise affine modulation of pre-Pairformer pair representations		(Lee et al., 2026)
BioEmu	Raw MSA	Diffusion-based generation of protein equilibrium conformational ensembles		(Lewis et al., 2025)
ConforMix	Raw MSA	Inference-time twisted diffusion with guidance, filtering, and enhanced sampling		(Richman et al., 2026)
Edited MSA Input				
AF-Cluster	clustering	Clustering MSA sequences by similarity and folding individual clusters		(Wayment-Steele et al., 2024)
MSA-Subs.	shadowing	Random subsampling of MSA rows over different alignment depths		(Del Alamo et al., 2021)
AFSample3	masking	Random masking of non-query MSA columns while preserving the query row		(Kalakoti & Wallner, 2026)

- **Conformation branches and MSA depth.** Since each evaluation target contains two experimentally observed conformations, we set the number of conformation branches to $K = 2$. The input MSA depth was fixed at $S = 5{,}120$. Targets whose raw MSA depth was below this value were excluded. Each SWE branch compressed the input into $P = 128$ pseudo-MSA rows.
- **Architecture dimensions.** The original outer-product factors and pair representation used $c_h = 32$ and $c_z = 128$, respectively. We set the SWE descriptor and anchor dimension to $D_h = 512$. The row-descriptor MLP had dimensions $512 \to 2{,}048 \to 512$.
- **Architecture operations.** The SWE transport temperature was set to 0.2, and the transport matrix was computed using eight Sinkhorn iterations. For QBOP, the attention temperature and differential-attention coefficient were both set to 1.0.
- **Loss weights.** The nominal post-warm-up objective weights were

$$\lambda_{\text{native}} = 0.6, \qquad \lambda_{\text{state}} = 0.2, \qquad \lambda_{\text{div}} = 0.2. \tag{45}$$

- **Diversity thresholds.** We used

$$\tau_{\text{full}} = 0.6, \qquad \tau_{\text{ignore}} = 0.9, \qquad \eta = 0.2. \tag{46}$$

  Thus, ground-truth conformations with $t_{\text{gt}} \leq \tau_{\text{full}}$ imposed the full target separation η, whereas pairs with $t_{\text{gt}} \geq \tau_{\text{ignore}}$ imposed no diversity margin. The target separation was linearly interpolated between these two thresholds.

D.2 Computational Resources

All model training was performed on NVIDIA H20 GPUs using bfloat16 precision. Our method was implemented by fine-tuning Protenix (Team et al., 2026), an open-source reproduction of AlphaFold 3. Input MSAs were searched and prepared using the Protenix-potoco pipeline.

D.3 OPM Aggregation Analysis: Details and Derivations

This section provides the evaluation protocol and mathematical derivations for the OPM aggregation analysis in Section 4.1. The analysis characterizes the magnitude and dimensional structure of row-wise information that is not retained by conventional OPM mean aggregation.

Evaluation data. The multi-conformation set contains 157 proteins with multiple experimentally resolved structures: 15 membrane proteins, 51 GPCRs, 20 kinases, 22 proteins with domain motions, 34 proteins with cryptic pockets, and 15 fold-switching proteins. The control group pools the

CASP15 and CASP16 targets, each represented by one experimental reference structure. The same MSA searching and OPM feature-extraction procedure is applied to both groups.

All statistics are first computed at the protein level and then averaged within each dataset. For a statistic X, the relative change reported for dataset g is

$$\Delta_g(X) = \frac{\overline{X}_g - \overline{X}_{\mathrm{ctrl}}}{\overline{X}_{\mathrm{ctrl}}} \times 100\%, \tag{47}$$

where $\overline{X}_{\mathrm{ctrl}}$ is the pooled mean of the CASP15 and CASP16 controls.

For binary discrimination, each protein-level statistic is used independently as a scalar score, with larger values indicating a multi-conformation protein. Accuracy and AUPRC are reported as the mean over 20 repetitions of stratified five-fold cross-validation.

Row-wise OPM representation. For MSA row s and residue pair (i, j), we define the projected row-wise OPM feature as

$$r_{sij} = \mathrm{Linear}_{\mathrm{out}}\left(a_{si} \otimes b_{sj}\right) \in \mathbb{R}^{c_z}. \tag{48}$$

Its mean over the S MSA rows is

$$\bar{r}_{ij} = \frac{1}{S}\sum_{s=0}^{S-1} r_{sij} = \mathrm{Linear}_{\mathrm{out}}\left(\mathcal{O}_{ij}\right), \tag{49}$$

where $\mathcal{O}_{ij}$ is the conventional OPM output defined in Eq. equation 3. The second equality follows from the linearity of the output projection. Thus, $\bar{r}_{ij}$ is retained by the downstream pair representation, whereas differences among the individual r_{sij} are not directly retained.

For each residue pair, we collect the row-wise features into

$$R_{ij} = \begin{bmatrix} r_{0ij}^{\top} \\ \vdots \\ r_{(S-1)ij}^{\top} \end{bmatrix} \in \mathbb{R}^{S \times c_z}. \tag{50}$$

This matrix admits the decomposition

$$R_{ij} = \mathbf{1}_S \bar{r}_{ij}^{\top} + \widetilde{R}_{ij}, \qquad \widetilde{R}_{ij} = \begin{bmatrix} (r_{0ij} - \bar{r}_{ij})^{\top} \\ \vdots \\ (r_{(S-1)ij} - \bar{r}_{ij})^{\top} \end{bmatrix}. \tag{51}$$

Because

$$\mathbf{1}_S^{\top} \widetilde{R}_{ij} = 0, \tag{52}$$

averaging along the row dimension removes the complete centered component $\widetilde{R}_{ij}$. Table 6 summarizes the four statistics used to quantify its magnitude and structure. The spectral notation in the table is defined formally below.

Row variance. For residue pair (i, j), row variance is the centered energy per MSA row and per feature channel:

$$V_{ij} = \frac{1}{Sc_z}\sum_{s=0}^{S-1} \|r_{sij} - \bar{r}_{ij}\|_2^2 = \frac{\left\|\widetilde{R}_{ij}\right\|_{\mathrm{F}}^2}{Sc_z}. \tag{53}$$

The protein-level statistic averages this quantity over the evaluated residue pairs $\mathcal{P}$:

$$V_{\mathrm{row}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[V_{ij}\right]. \tag{54}$$

A larger value indicates greater absolute disagreement among MSA rows before mean aggregation.

Table 6: Definitions and interpretations of four statistics characterizing row-wise variation removed by OPM mean aggregation.

Statistic	Definition	Interpretation
Row variance	$V_{\text{row}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[\frac{1}{Sc_z}\sum_{s=0}^{S-1}\|r_{sij}-\bar{r}_{ij}\|_2^2\right]$	Absolute magnitude of row-wise variation before averaging
Lost-energy fraction	$E_{\text{lost}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[\frac{\sum_{s=0}^{S-1}\|r_{sij}-\bar{r}_{ij}\|_2^2}{\sum_{s=0}^{S-1}\|r_{sij}\|_2^2}\right]$	Fraction of total row energy not retained by the mean
Spectral tail	$T_{\text{spec}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[1-\frac{\lambda_{ij,1}}{\text{tr}(K_{ij})}\right]$	Fraction of centered variation outside its leading direction
Effective rank	$R_{\text{eff}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[\exp\left(-\sum_{k=1}^{q_{ij}}\pi_{ij,k}\log\pi_{ij,k}\right)\right]$	Effective number of contributing row-wise directions

Table 7: OPM variation statistics and their ability to distinguish multi-conformation proteins from single-conformation controls. Dataset rows report percentage changes relative to the pooled mean of the CASP15 and CASP16 controls. The final two rows report mean AUPRC and classification accuracy using each statistic independently.

Dataset	Row Variance	Lost Energy	Spectral Tail	Effective Rank
Membrane	+18.40%	+7.62%	+10.39%	+13.64%
GPCR	+26.21%	+7.88%	+16.77%	+21.67%
Kinases	+24.12%	+5.54%	+13.06%	+18.58%
Domain motion	+17.15%	+6.09%	+9.46%	+14.05%
Cryptic pocket	+19.73%	+5.12%	+10.15%	+14.57%
Fold switching	+15.23%	+4.58%	+8.10%	+11.23%
AUPRC (%)	78.77	85.29	83.55	**85.93**
ACC (%)	73.08	73.57	75.70	**77.64**

Lost-energy fraction. The mean and centered components in Eq. equation 51 are orthogonal. Consequently,

$$\|R_{ij}\|_{\text{F}}^2 = S\,\|\bar{r}_{ij}\|_2^2 + \left\|\widetilde{R}_{ij}\right\|_{\text{F}}^2. \tag{55}$$

The fraction of row-wise energy absent from the mean is therefore

$$\begin{aligned} E_{ij}^{\text{lost}} &= \frac{\left\|\widetilde{R}_{ij}\right\|_{\text{F}}^2}{\|R_{ij}\|_{\text{F}}^2} \\ &= \frac{\sum_{s=0}^{S-1}\|r_{sij}-\bar{r}_{ij}\|_2^2}{\sum_{s=0}^{S-1}\|r_{sij}\|_2^2} \\ &= 1 - \frac{S\,\|\bar{r}_{ij}\|_2^2}{\sum_{s=0}^{S-1}\|r_{sij}\|_2^2}. \end{aligned} \tag{56}$$

The corresponding protein-level statistic is

$$E_{\text{lost}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[E_{ij}^{\text{lost}}\right]. \tag{57}$$

Unlike row variance, lost energy is normalized by the total feature energy and therefore measures the relative amount of variation removed by averaging.

Spectrum of the centered row features. To characterize the dimensional structure of the removed variation, we form the centered row Gram matrix

$$K_{ij} = \widetilde{R}_{ij}\widetilde{R}_{ij}^{\top} \in \mathbb{R}^{S\times S}. \tag{58}$$

Its trace equals the total centered energy:

$$\mathrm{tr}(K_{ij}) = \left\|\widetilde{R}_{ij}\right\|_{\mathrm{F}}^2. \tag{59}$$

Let its positive eigenvalues be

$$\lambda_{ij,1} \geq \cdots \geq \lambda_{ij,q_{ij}} > 0, \qquad q_{ij} = \mathrm{rank}(K_{ij}) \leq \min(S-1, c_z). \tag{60}$$

Their normalized weights are

$$\pi_{ij,k} = \frac{\lambda_{ij,k}}{\sum_{\ell=1}^{q_{ij}} \lambda_{ij,\ell}} = \frac{\lambda_{ij,k}}{\mathrm{tr}(K_{ij})}. \tag{61}$$

Spectral tail. The leading eigenvalue captures the strongest direction of row-wise variation. The remaining fraction of the centered energy is

$$T_{ij}^{\mathrm{spec}} = 1 - \pi_{ij,1} = 1 - \frac{\lambda_{ij,1}}{\mathrm{tr}(K_{ij})}. \tag{62}$$

Averaging over residue pairs gives

$$T_{\mathrm{spec}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[T_{ij}^{\mathrm{spec}}\right]. \tag{63}$$

A small spectral tail indicates that most row-wise variation follows one dominant direction. A larger value indicates that more of the variation lies in additional directions.

Effective rank. Effective rank is the exponential of the entropy of the normalized eigenvalue spectrum:

$$R_{ij}^{\mathrm{eff}} = \exp\left(-\sum_{k=1}^{q_{ij}} \pi_{ij,k} \log \pi_{ij,k}\right). \tag{64}$$

The protein-level effective rank is

$$R_{\mathrm{eff}} = \mathbb{E}_{(i,j)\in\mathcal{P}}\left[R_{ij}^{\mathrm{eff}}\right]. \tag{65}$$

It approaches one when a single direction dominates and increases when the centered energy is distributed more evenly across several directions.

Detailed result interpretation. Every multi-conformation category exhibits an increase in all four statistics relative to the pooled controls. The 21.48% increase in pooled row variance indicates that multi-conformation proteins contain stronger absolute disagreement among MSA rows. The simultaneous increases in lost energy, spectral tail, and effective rank show that a larger fraction of the representation is removed and that this variation is distributed across more independent directions.

The single-statistic classification results provide a complementary protein-level test. The dimensional organization of the centered row features is particularly informative for distinguishing the two groups. These quantities remain representation-level diagnostics. They do not assign individual MSA rows to particular experimental conformations, nor do they establish a one-to-one correspondence between spectral directions and structural states. They instead show that reducing all row-wise features to a single mean removes a signal associated with conformational multiplicity. This observation motivates the architecture in Section 3.1, which retains multiple compact row-wise summaries and routes them to distinct structural branches before conformation prediction.

D.4 Evaluation Metrics for Intrinsically Disordered Proteins

Common protocol and mean absolute error (MAE). For each target of length L, let $\mathcal{X} = \{X^{(m)}\}_{m=1}^{M}$ denote the generated ensemble and $\mathcal{Y} = \{Y^{(n)}\}_{n=1}^{N}$ the PED reference ensemble, where $M = 20$ and N may vary across targets. The Cα coordinates are $X^{(m)} = \{\mathbf{r}_i^{(m)}\}_{i=1}^{L}$ and $Y^{(n)} = \{\mathbf{y}_i^{(n)}\}_{i=1}^{L}$, with corresponding residue indices across conformations. All conformations within each ensemble receive equal weight. We evaluate residue pairs

$$\mathcal{P} = \{(i,j) : 1 \leq i < j \leq L,\ j - i \geq 3\}. \tag{66}$$

For any scalar structural descriptor f, the target-level error is the absolute difference between its ensemble means:

$$e_f = \left|\frac{1}{M}\sum_{m=1}^{M} f\left(X^{(m)}\right) - \frac{1}{N}\sum_{n=1}^{N} f\left(Y^{(n)}\right)\right|. \tag{67}$$

Different descriptors are as follows:

Pairwise distance (PwD.). For each eligible residue pair, the generated and reference distances are

$$d_{ij}^{\mathrm{gen},(m)} = \left\| \mathbf{r}_i^{(m)} - \mathbf{r}_j^{(m)} \right\|_2, \qquad d_{ij}^{\mathrm{ref},(n)} = \left\| \mathbf{y}_i^{(n)} - \mathbf{y}_j^{(n)} \right\|_2. \tag{68}$$

PwD. compares the ensemble-mean distance vectors:

$$e_{\mathrm{PwD}} = \frac{1}{|\mathcal{P}|} \sum_{(i,j)\in\mathcal{P}} \left| \frac{1}{M} \sum_{m=1}^{M} d_{ij}^{\mathrm{gen},(m)} - \frac{1}{N} \sum_{n=1}^{N} d_{ij}^{\mathrm{ref},(n)} \right|. \tag{69}$$

Radius of gyration (R_g). For any conformation $Z = \{\mathbf{z}_i\}_{i=1}^{L}$ in either ensemble, we calculate the radius of gyration from equal-mass Cα coordinates:

$$\bar{\mathbf{z}} = \frac{1}{L} \sum_{i=1}^{L} \mathbf{z}_i, \qquad R_g(Z) = \sqrt{\frac{1}{L} \sum_{i=1}^{L} \|\mathbf{z}_i - \bar{\mathbf{z}}\|_2^2}. \tag{70}$$

The Rg. column reports e_{R_g} from Eq. 67, comparing the mean radius of gyration of the generated and reference ensembles (Lu et al., 2025).

Contact map (Cp.). A residue pair is considered in contact when its Cα distance is below $8.0\,\text{Å}$. Contact probabilities are estimated separately within the two ensembles:

$$P_{ij}^{\mathrm{gen}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\left\{ d_{ij}^{\mathrm{gen},(m)} < 8.0\,\text{Å} \right\}, \qquad P_{ij}^{\mathrm{ref}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\left\{ d_{ij}^{\mathrm{ref},(n)} < 8.0\,\text{Å} \right\}. \tag{71}$$

Using a pseudocount of 0.01 before taking logarithms, the contact-map error is

$$e_{\mathrm{Cp}} = \frac{1}{|\mathcal{P}|} \sum_{(i,j)\in\mathcal{P}} \left| \log\big(P_{ij}^{\mathrm{gen}} + 0.01\big) - \log\big(P_{ij}^{\mathrm{ref}} + 0.01\big) \right|. \tag{72}$$

End-to-end distance (R_{ee}) and shape. For each conformation Z, we calculate the end-to-end distance and instantaneous shape ratio:

$$R_{ee}(Z) = \|\mathbf{z}_L - \mathbf{z}_1\|_2, \qquad R_s(Z) = \frac{R_{ee}(Z)^2}{R_g(Z)^2}. \tag{73}$$

Here, R_{ee} measures chain extension, while the dimensionless R_s characterizes end-to-end extension relative to overall size. Following (Shadman et al., 2024), we average the per-conformation ratios within each ensemble rather than taking a ratio of ensemble-averaged quantities. The R_{ee} and Shape columns report $e_{R_{ee}}$ and e_{R_s}, respectively, using Eq. 67.

Flory exponent (ν). For each conformation Z and sequence separation $s \in \{3, \ldots, L-1\}$, we calculate

$$R_Z(s) = \sqrt{\frac{1}{L-s} \sum_{i=1}^{L-s} \|\mathbf{z}_{i+s} - \mathbf{z}_i\|_2^2}. \tag{74}$$

We estimate the conformation-specific exponent $\nu(Z)$ as the least-squares slope in

$$\log R_Z(s) = a_Z + \nu(Z) \log s. \tag{75}$$

This descriptor characterizes internal-distance scaling (Koren et al., 2023). We fit the exponent separately for each conformation in both ensembles and report the absolute difference between their mean fitted exponents, e_ν, using Eq. 67.

Solvent-accessible surface area (SASA). We compute SASA separately for every generated and reference conformation using the Shrake–Rupley algorithm (Shrake & Rupley, 1973), with a $1.4\,\text{Å}$ solvent probe and 960 sphere points per atom. For each target, we use a consistent set of corresponding backbone heavy atoms drawn from $\{\mathrm{N}, \mathrm{CA}, \mathrm{C}, \mathrm{O}, \mathrm{OXT}\}$ and present across all evaluated conformations and methods. Their atomic areas are summed to obtain the total backbone SASA of each conformation in Å^2. The SASA column reports the absolute difference between the ensemble-mean values according to Eq. 67.

Ensemble distance RMS (ens-dRMS). Following Lazar et al. (2020), we summarize each ensemble using the median distance of every eligible residue pair:

$$\widetilde{d}_{ij}^{\mathrm{gen}} = \underset{m=1,\dots,M}{\mathrm{median}}\, d_{ij}^{\mathrm{gen},(m)}, \qquad \widetilde{d}_{ij}^{\mathrm{ref}} = \underset{n=1,\dots,N}{\mathrm{median}}\, d_{ij}^{\mathrm{ref},(n)}. \tag{76}$$

The ensemble distance RMS is

$$e_{\text{ens-dRMS}} = \sqrt{\frac{1}{|\mathcal{P}|} \sum_{(i,j)\in\mathcal{P}} \left(\widetilde{d}_{ij}^{\mathrm{gen}} - \widetilde{d}_{ij}^{\mathrm{ref}}\right)^2}. \tag{77}$$

This metric is invariant to rigid-body rotation and translation and compares the median internal-distance geometry of the two ensembles. Together, these metrics assess ensemble summary statistics; they do not fully characterize differences between conformational distributions.

E More Related Works

MSA understanding and representative learning. MSA-based models use homologous sequences for contact inference, representation learning, and generation. MSA Transformer and MSA Pairformer encode cross-sequence dependencies for contact and pair-representation learning (Rao et al., 2021; Akiyama et al., 2025), while AlphaFold2's Evoformer couple MSA, pair, and structural representations within end-to-end folding networks (Jumper et al., 2021). Related approaches incorporate homologs through retrieval or family-level modeling for representation learning, variant-effect prediction, and sequence generation (Ma et al., 2024; Notin et al., 2022; Truong Jr & Bepler, 2023; Alamdari et al., 2023). EvoPool further uses sliced-Wasserstein pooling to summarize homolog-derived embeddings (NaderiAlizadeh & Singh, 2026). Guided by this broader line of work, EVORESOLVE adapts query-biased aggregation from MSA Pairformer and SWE-based compression from EvoPool to the OPM pathway, integrating them with multi-branch structural supervision and a shared folding model for alternative-state prediction.

F Can EvoResolve Support Variable-K Prediction?

The model presented in the main paper uses a fixed number of conformation branches: a model trained with K branches also predicts exactly K conformations at inference time. We have implemented an extension that supports mixed-K training and autoregressive inference. Instead of maintaining a separate SWE–QBOP route for every output conformation, the improved model reuses one shared route and generates conformations sequentially. A lightweight history adaptor summarizes the conformations and MSA-derived signals selected in earlier steps, and uses this history to redirect the next rollout toward a new conformational mode. During training, proteins are presented at every feasible cardinality up to $K_{\max} = 6$, allowing the same model to learn from targets with different numbers of experimental structures. At inference time, the shared model is unrolled autoregressively until the requested number of conformations is produced. We evaluate UniCorNE targets after grouping them by the number of available ground-truth conformations in Table 8.

Evaluation dataset. UniCorNE organizes experimentally determined protein structures into conformational ensembles associated with UniProt entries, providing a resource for studying structural variation within individual proteins (UniCorNE, a). Its associated benchmark, UniCorNE-bench, provides multiple experimental reference conformations per target for evaluating protein ensemble prediction methods (UniCorNE, b). To evaluate prediction beyond the two-conformation setting, we select targets from UniCorNE-bench with at least three reference conformations ($K \geq 3$), where K denotes the number of available ground-truth conformations for each target. We report results separately for $K = 3$, 4, 5, 6, and $K \geq 7$.

Table 8: TM-score on UniCorNE, grouped by the number of experimentally observed reference conformations (K). Higher values indicate better performance. The best result in each column is shown in **bold**.

Method	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K \geq 7$
Edited MSA Input					
AF-Cluster	0.601	0.588	0.564	0.624	0.532
MSA-Subsampling	0.678	0.664	0.646	0.698	**0.603**
AFSample3	0.611	0.602	0.585	0.633	0.533
Raw MSA Input					
BioEmu	0.635	0.626	0.604	0.643	0.586
ConforMix	0.602	0.594	0.566	0.617	0.574
ConforNets	0.665	0.655	0.658	0.700	0.593
EVORESOLVE(Mixed K)	**0.684**	**0.670**	**0.659**	**0.704**	0.598