# LEARNING PROTEIN-SPECIFIC ELASTIC INTERACTIONS FOR GLOBALLY COUPLED PROTEIN DYNAMICS

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Proteins function through conformational flexibility and collective motion that a single reference structure does not specify. Molecular dynamics (MD) simulation captures these dynamics but requires long trajectories; learning-based methods that predict covariances from structure avoid this cost, but they couple each residue pair through a single scalar and provide no mechanism from which globally consistent collective motion arises. Elastic network models (ENMs) offer such a mechanism, describing global dynamics through explicit inter-residue interactions, yet they prescribe these interactions with fixed cutoffs and hand-designed rules. We introduce **ENMulator**, which combines the two: a network maps sequence and structure to protein-specific elastic interactions, which parameterize an elastic network whose Hessian induces a globally coupled, residue-level displacement covariance. A complementary branch predicts residue-wise anisotropic covariances, and a blockwise geodesic congruence calibrates the residue marginals of the elastic covariance toward these predictions while exactly preserving the canonical correlations between residues. On the ATLAS test set, in both covariance and sampled-ensemble evaluation, it improves recovery of covariance geometry, collective modes, and correlated motion over prescribed ENMs and direct covariance predictors, reducing DCCM error by 17% relative to the next-best method, while remaining competitive on residue-wise fluctuations. These results support learned elastic interactions as a principled bridge from a single reference structure to globally coupled protein dynamics without simulation.

## 1 INTRODUCTION

Recent advances in protein structure prediction have made accurate static protein structures broadly accessible (Jumper et al., 2021; Tunyasuvunakool et al., 2021; Varadi et al., 2022; 2024). However, a single reference structure does not fully characterize protein function, which is often mediated by conformational flexibility, allosteric communication, and collective motion (Teague, 2003; Goodey & Benkovic, 2008; Bahar et al., 2010). Characterizing these functionally relevant dynamics requires modeling both residue-level fluctuations and correlated motions across the structure (Bahar et al., 1997; Atilgan et al., 2001; Nussinov et al., 2023). Molecular dynamics (MD) simulation provides the de facto physics-based reference for estimating these quantities from atomistic trajectories (Shaw et al., 2010; Hollingsworth & Dror, 2018; Durrant & McCammon, 2011). Yet obtaining reliable equilibrium fluctuation statistics typically requires long trajectories, making routine characterization at scale difficult (Lewis et al., 2025). This motivates efficient models that infer MD-derived dynamics from a single reference structure.

Learning-based approaches alleviate this computational bottleneck by inferring protein dynamics directly from sequence and structure. Some methods generate conformational ensembles or molecular trajectories (Jing et al., 2024a; Lewis et al., 2025; Wolf et al., 2025; Lu et al., 2024; Jing et al., 2024b), whereas others predict compact dynamics descriptors more efficiently (Kouba et al., 2025; Wayment-Steele et al., 2026; Bafna et al., 2026; Viliuga et al., 2026). In particular, DynaProt (Bafna et al., 2026) and BackFlip-2 (Viliuga et al., 2026) learn residue-wise anisotropic fluctuations and inter-residue coupling from MD-derived supervision, showing that dynamics predictions can adapt to individual proteins. They leave two gaps, however. First, their expressivity is limited: both couple each residue pair through a single scalar, so the three canonical correlations between two residues are forced to coincide, and the direction-dependent coupling observed in MD cannot be expressed. Sec-



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.

Table 1: **ENMulator models protein dynamics through learned, protein-specific elastic interactions.** Prescribed ENMs provide explicit mechanical interactions but use fixed rules, whereas direct predictors adapt to proteins without an explicit interaction model. ENMulator combines both.

| | **Direct predictor** (*e.g., DynaProt*) | **Prescribed ENM** (*e.g., ANM*) | **ENMulator (Ours)** |
|---|---|---|---|
| Protein-specific (learned) | ✓ | ✗ | ✓ |
| Mechanically structured | ✗ | ✓ | ✓ |

ond, they provide no mechanism from which a globally consistent covariance arises: cross-residue blocks are predicted pair by pair, so positive semidefiniteness and consistency across all residue pairs must be designed separately. This leaves a complementary challenge: to retain protein-specific adaptation while obtaining a globally consistent covariance by construction.

One explicit global mechanism is provided by classical elastic network models (ENMs), which have long offered an efficient, mechanically structured framework for analyzing global protein dynamics directly from native structures (Go et al., 1983; Brooks & Karplus, 1983; Tirion, 1996; Bahar et al., 2010; 2015). Models from the Gaussian network model (GNM) to the anisotropic network model (ANM) represent proteins as residues coupled by elastic interactions, which induce a single globally coupled covariance from which residue fluctuations and correlated motions arise (Haliloglu et al., 1997; Bahar et al., 1997; Atilgan et al., 2001). ENMs provide a compelling inductive bias for modeling the correlated global motions that connect protein structure to function: their low-frequency modes often align with experimentally observed conformational changes (Tama & Sanejouand, 2001; Yang et al., 2007). However, conventional ENMs prescribe the connectivity and strengths of these interactions using fixed cutoffs and hand-designed rules and cannot adapt them to protein-specific dynamics observed in MD (Koehl et al., 2021; Dubanevics & McLeish, 2022). Elastic parameters can instead be fitted to atomistic fluctuations (Moritsugu & Smith, 2007; Lyman et al., 2008), but this requires MD of the very protein whose dynamics the model is meant to replace.

To address this challenge, we propose **ENMulator**, which combines the explicit global coupling of ENMs with protein-specific learning: instead of prescribed rules, a network maps sequence and structure to a protein-specific stiffness for every elastic interaction, emulating the full $C_\alpha$ displacement covariance estimated from MD. For positive predicted stiffnesses, the induced base covariance is positive semidefinite, and its cross-residue blocks are not constrained to have equal canonical correlations. A complementary branch predicts residue-wise anisotropic covariances in structure-dependent local frames, and geodesic congruence calibrates the local marginals of the base covariance toward these predictions while preserving positive semidefiniteness and the canonical correlations between residues. The model is trained end-to-end with MD-derived covariance supervision. On the ATLAS test set, ENMulator attains the lowest Bures–Wasserstein distance to the MD covariance and the highest collective-mode overlap, and reduces DCCM MAE by $17\%$ relative to the next-best baseline ($0.138$ vs. $0.166$), with gains consistent across proteins.

Our contributions are summarized as follows:

- We introduce **ENMulator**, which bridges adaptive direct predictors and fixed-rule ENMs: a network maps sequence and structure to protein-specific elastic interactions, which induce a globally coupled $C_\alpha$ displacement covariance that is positive semidefinite by construction.
- Because elastic marginals are tied to the global Hessian, we calibrate them toward learned residue-wise covariances through a geodesic congruence, which improves local anisotropic fluctuations while leaving inter-residue canonical correlations unchanged.
- On the ATLAS test set, **ENMulator** improves recovery of covariance geometry, collective modes, and correlated motion in both covariance and sampled-ensemble evaluation. We establish a lower bound on coupling error for joint-covariance reconstructions with equal pairwise canonical correlations, and show that ENMulator falls below by recovering direction-dependent coupling.

## 2 Related Work

**Learning-based modeling of protein dynamics.** Recent learning-based approaches infer protein dynamics from sequence and structure either by generating conformational ensembles or trajectories (Jing et al., 2024a;b; Lewis et al., 2025; Wolf et al., 2025; Lu et al., 2024), or by predicting dynamics descriptors directly. Flexpert-3D predicts scalar measures of residue flexibility, such as RMSF (Kouba et al., 2025), whereas DynaProt models richer second-order dynamics through residue-wise anisotropic covariances and inter-residue coupling, learned with MD-derived supervision (Bafna et al., 2026). More recently, BackFlip-2 follows a closely related direct-prediction approach, predicting directional per-residue flexibility and pairwise dynamic correlations from an equilibrium structure with an SE(3)-equivariant model (Viliuga et al., 2026).

**Elastic network models and normal mode analysis.** Elastic network models (ENMs) provide a classical framework for approximating protein dynamics around a reference structure without explicitly sampling atomistic trajectories (Haliloglu et al., 1997; Bahar et al., 1997; Atilgan et al., 2001). In the anisotropic network model (ANM), residues are represented as nodes connected by harmonic interactions. The non-rigid eigenvectors of the resulting Hessian describe collective normal modes, and the inverses of its nonzero eigenvalues set the fluctuation variance along each mode (Atilgan et al., 2001; Cui & Bahar, 2005; Skjaerven et al., 2009). Classical ANM uses a uniform spring constant for contacts within a distance cutoff, whereas subsequent ENM variants prescribe distance-dependent or heterogeneous interaction strengths (Hinsen et al., 2000; Ming & Wall, 2005; Yang et al., 2009). Other approaches fit elastic parameters to atomistic fluctuations, typically from MD of the protein of interest (Moritsugu & Smith, 2007; Lyman et al., 2008). In contrast, ENMulator learns a mapping from sequence and structure to interaction strengths across proteins and predicts them for unseen proteins without simulation.

## 3 Elastic Network Models for Protein Dynamics

Let $\boldsymbol{X} \in \mathbb{R}^{N\times 3}$ denote the reference $C_\alpha$ structure, with $\boldsymbol{x}_i \in \mathbb{R}^3$ the $C_\alpha$ coordinate of residue $i$. For a conformation $\boldsymbol{X}' \in \mathbb{R}^{N\times 3}$ globally aligned to $\boldsymbol{X}$, let $\Delta\boldsymbol{x}_i = \boldsymbol{x}'_i - \boldsymbol{x}_i$ denote the displacement of residue $i$, and stack these displacements as

$$\Delta\boldsymbol{x} = \begin{bmatrix}\Delta\boldsymbol{x}_1^\top & \cdots & \Delta\boldsymbol{x}_N^\top\end{bmatrix}^\top \in \mathbb{R}^{3N}.$$

Our prediction target is the reference–centered $C_\alpha$ displacement covariance $\boldsymbol{\Sigma} = \mathbb{E}[\Delta\boldsymbol{x}\,\Delta\boldsymbol{x}^\top] \in \mathbb{R}^{3N\times 3N}$ over the aligned ensemble, taken about $\boldsymbol{X}$, which the harmonic approximation treats as a local equilibrium (Appendix C.1).

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote an undirected residue contact graph, whose nodes are residues and edges connect spatially neighboring $C_\alpha$ atoms, for example pairs satisfying $r_{ij} \le r_c$. For each edge $(i,j) \in \mathcal{E}$, define

$$r_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2, \qquad \boldsymbol{u}_{ij} = \frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{r_{ij}}, \qquad \boldsymbol{b}_{ij} = (\boldsymbol{e}_i - \boldsymbol{e}_j) \otimes \boldsymbol{u}_{ij} \in \mathbb{R}^{3N}, \tag{1}$$

where $\boldsymbol{e}_i \in \mathbb{R}^N$ is the $i$-th standard basis vector. Here, $\boldsymbol{b}_{ij}^\top \Delta\boldsymbol{x} = \boldsymbol{u}_{ij}^\top(\Delta\boldsymbol{x}_i - \Delta\boldsymbol{x}_j)$ is the relative displacement of residues $i$ and $j$ along their reference inter-residue direction. Given the edge geometry above, each edge $(i,j) \in \mathcal{E}$ is associated with a positive scalar stiffness $\gamma_{ij} > 0$. The term $\gamma_{ij}\boldsymbol{b}_{ij}\boldsymbol{b}_{ij}^\top$ represents the directional stiffness contributed by edge $(i,j)$, and summing these contributions over the contact graph gives the Hessian

$$\boldsymbol{H} = \sum_{(i,j)\in\mathcal{E}} \gamma_{ij}\boldsymbol{b}_{ij}\boldsymbol{b}_{ij}^\top. \tag{2}$$

Because $\gamma_{ij} > 0$, each edge contribution is positive semidefinite, and hence $\boldsymbol{H}$ is symmetric positive semidefinite. Under the harmonic approximation, the displacement covariance is $\boldsymbol{\Sigma}^{\mathrm{ENM}} = k_{\mathrm{B}}T\,\boldsymbol{H}^+$, where $(\cdot)^+$ denotes the Moore–Penrose pseudoinverse, which excludes the rigid-body null modes. We absorb $k_{\mathrm{B}}T$ into the stiffnesses and write $\boldsymbol{\Sigma}^{\mathrm{ENM}} = \boldsymbol{H}^+$.

The canonical ANM is the uniform-stiffness special case, $\gamma_{ij} = \gamma$ for all $(i,j) \in \mathcal{E}$. Other ENMs vary the stiffnesses with distance, residue properties, or structural context. Our method retains this Hessian construction but predicts positive, protein-specific stiffnesses $\gamma_{ij}$ from data.

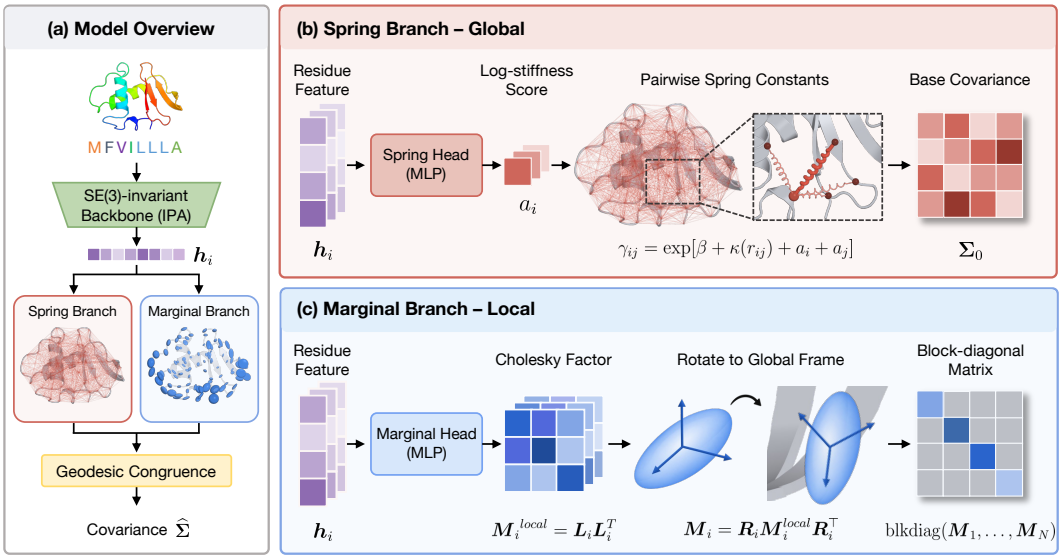


Figure 1: **Overview of ENMulator. (a)** The input structure and sequence are encoded into residue representations shared by a spring branch and a marginal branch. The outputs are combined through blockwise geodesic congruence to produce the final $C_\alpha$ displacement covariance $\widehat{\Sigma}$, whose diagonal blocks describe residue-wise anisotropic fluctuations and off-diagonal blocks describe cross-residue correlated motion. **(b)** The spring branch predicts pairwise edge stiffnesses on the residue contact graph and constructs the globally coupled base covariance $\Sigma_0$. **(c)** The marginal branch predicts residue-wise anisotropic covariances $\{M_i\}_{i=1}^N$ in local frames and rotates them into the global frame.

## 4 ENMulator: Learning Protein-Specific Elastic Interactions

Given an amino-acid sequence $\boldsymbol{s} = (s_1, \ldots, s_N)$ and a reference backbone structure with $C_\alpha$ coordinates $\boldsymbol{X}$, we predict the $C_\alpha$ displacement covariance $\boldsymbol{\Sigma} \in \mathbb{R}^{3N\times 3N}$ defined in Section 3. A spring branch induces a globally coupled base covariance $\boldsymbol{\Sigma}_0$, and a marginal branch predicts residue-wise anisotropic covariances $\{\boldsymbol{M}_i\}_{i=1}^N$. Blockwise geodesic congruence combines these into $\widehat{\boldsymbol{\Sigma}}$.

### 4.1 Learning Elastic Interactions and Residue-wise Marginals

A shared structure encoder $E_{\boldsymbol{\theta}}$, implemented with invariant point attention (IPA) (Jumper et al., 2021), maps sequence and structure to rotation-invariant residue representations $\boldsymbol{h}_i = [E_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{X})]_i$ for both branches.

**Learning protein-specific elastic interactions.** We connect residue pairs with $r_{ij} \le r_c$, together with all consecutive residues, to form an undirected contact graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with edge geometry defined by Eq. 1. The spring head combines bounded residue-wise log-stiffness scores with global and radial terms to predict positive, symmetric edge stiffnesses:

$$a_i = A \tanh\left(\frac{f_{\mathrm{spr}}(\boldsymbol{h}_i)}{A}\right), \qquad \gamma_{ij} = \exp[\beta + \kappa(r_{ij}) + a_i + a_j]\,. \tag{3}$$

Here, $A > 0$ bounds the residue scores, $\beta$ is a learned global log-stiffness offset, and $\kappa$ is a learned radial log-stiffness function shared across proteins, while the endpoint scores adapt interaction strengths to each protein. These stiffnesses define the Hessian in Eq. 2 on the edge geometry of the reference structure. Parameterizations of $f_{\mathrm{spr}}$ and $\kappa$ are provided in Appendix B.1.

Let $\boldsymbol{Q}_6 \in \mathbb{R}^{3N\times 6}$ contain an orthonormal basis for the six rigid-body modes of $\boldsymbol{X}$ (three translations and three infinitesimal rotations). Assuming $\boldsymbol{H}$ has no additional null modes, the base covariance is its pseudoinverse, which inverts $\boldsymbol{H}$ on the complement of the rigid-body subspace and assigns negligible variance to the rigid-body modes:

$$\boldsymbol{\Sigma}_0 = \boldsymbol{H}^+ = \left(\boldsymbol{H} + \boldsymbol{Q}_6\boldsymbol{Q}_6^\top\right)^{-1} - \boldsymbol{Q}_6\boldsymbol{Q}_6^\top. \tag{4}$$

**Predicting residue-wise anisotropic fluctuations.** To model local directional flexibility, the marginal branch predicts covariances in structure-dependent residue frames and rotates them into global coordinates. Let $\boldsymbol{R}_i \in \mathrm{SO}(3)$ denote a local frame constructed deterministically from the reference backbone geometry, as detailed in Appendix B.2. The branch maps $\boldsymbol{h}_i$ to a Cholesky factor $\boldsymbol{L}_i$ with strictly positive diagonal entries:

$$\boldsymbol{L}_i = f_{\mathrm{marg}}(\boldsymbol{h}_i), \qquad \boldsymbol{M}_i = \boldsymbol{R}_i \boldsymbol{L}_i \boldsymbol{L}_i^\top \boldsymbol{R}_i^\top. \tag{5}$$

Because $\boldsymbol{h}_i$ is rotation-invariant and $\boldsymbol{R}_i$ co-rotates with the input, a global rotation $\boldsymbol{Q}$ maps $\boldsymbol{M}_i$ to $\boldsymbol{Q}\boldsymbol{M}_i\boldsymbol{Q}^\top$, as required of a Cartesian covariance.

### 4.2 Blockwise Geodesic Congruence

In an elastic network, how much and in which directions each residue fluctuates is not a free choice: the marginal $\boldsymbol{P}_i = [\boldsymbol{\Sigma}_0]_{ii}$ is fixed by the same Hessian that sets the global coupling. So the spring branch cannot fit local anisotropic fluctuations without also changing collective motion. We therefore adjust each $\boldsymbol{P}_i$ toward the learned local covariance $\boldsymbol{M}_i$ using a residue-local linear transformation applied to the joint elastic fluctuations. Replacing diagonal blocks alone does not ensure positive semidefiniteness, so we transform the marginal and cross-covariance blocks together through a geodesic congruence. Assuming $\boldsymbol{P}_i \succ 0$, for a fixed $\alpha \in [0, 1]$, define:

$$\begin{aligned} \boldsymbol{D}_i &= \boldsymbol{P}_i^{1/2} \left( \boldsymbol{P}_i^{-1/2} \boldsymbol{M}_i \boldsymbol{P}_i^{-1/2} \right)^{\alpha/2} \boldsymbol{P}_i^{-1/2}, \\ \boldsymbol{D} &= \mathrm{blkdiag}(\boldsymbol{D}_1, \ldots, \boldsymbol{D}_N), \qquad \widehat{\boldsymbol{\Sigma}} = \boldsymbol{D}\boldsymbol{\Sigma}_0\boldsymbol{D}^\top. \end{aligned} \tag{6}$$

In coordinates whitened by $\boldsymbol{P}_i^{-1/2}$, $\boldsymbol{P}_i^{-1/2}\boldsymbol{D}_i\boldsymbol{P}_i^{1/2}$ is symmetric positive definite, so it contains no orthogonal factor, and $\boldsymbol{D}_i = \boldsymbol{I}$ at $\alpha = 0$. Each marginal follows the affine-invariant geodesic $\boldsymbol{P}_i \#_\alpha \boldsymbol{M}_i$ (Pennec et al., 2006):

$$\widehat{\boldsymbol{\Sigma}}_{ii} = \boldsymbol{P}_i \#_\alpha \boldsymbol{M}_i = \boldsymbol{P}_i^{1/2} \left( \boldsymbol{P}_i^{-1/2} \boldsymbol{M}_i \boldsymbol{P}_i^{-1/2} \right)^{\alpha} \boldsymbol{P}_i^{1/2}. \tag{7}$$

Thus, $\alpha = 0$ recovers $\boldsymbol{\Sigma}_0$, whereas $\alpha = 1$ sets each marginal to $\boldsymbol{M}_i$. Cross-covariances transform consistently as $\widehat{\boldsymbol{\Sigma}}_{ij} = \boldsymbol{D}_i[\boldsymbol{\Sigma}_0]_{ij}\boldsymbol{D}_j^\top$. Because every $\boldsymbol{D}_i$ is invertible, $\widehat{\boldsymbol{\Sigma}}$ stays positive semidefinite with the rank of $\boldsymbol{\Sigma}_0$ and keeps the inter-residue canonical correlations of $\boldsymbol{\Sigma}_0$ (Appendix B.3): the marginal branch sets how each residue fluctuates, whereas the learned elastic network alone sets their canonical coupling. We use $\alpha = 0.75$, selected on the validation split (Appendix G.2).

### 4.3 Training Objective and Ensemble Generation

**Training objective.** We train the model with a composite objective that supervises the full covariance, the residue-wise marginals, and the overall fluctuation scale:

$$\mathcal{L} = \mathcal{L}_{\mathrm{cov}} + \lambda_{\mathrm{s}}\mathcal{L}_{\mathrm{scale}} + \lambda_{\mathrm{d}}\mathcal{L}_{\mathrm{diag}} + \lambda_{\mathrm{m}}\mathcal{L}_{\mathrm{marg}}. \tag{8}$$

The full-covariance and scale terms are

$$\mathcal{L}_{\mathrm{cov}} = \rho\left( \frac{\|\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_F}{\|\boldsymbol{\Sigma}\|_F} \right), \qquad \mathcal{L}_{\mathrm{scale}} = \left( \log \frac{\mathrm{Tr}\,\widehat{\boldsymbol{\Sigma}}}{\mathrm{Tr}\,\boldsymbol{\Sigma}} \right)^2. \tag{9}$$

Here, $\rho(r) = r$ for $r \le 1$ and $1 + \log r$ otherwise, limiting the influence of poorly predicted proteins. The relative Frobenius error is bounded for under-scaled but unbounded for over-scaled predictions, whereas $\mathcal{L}_{\mathrm{scale}}$ penalizes both symmetrically. For residue-wise supervision, we use the squared log-Euclidean distance $d_{\log}(\boldsymbol{A}, \boldsymbol{B}) = \|\log \boldsymbol{A} - \log \boldsymbol{B}\|_F^2$ (Arsigny et al., 2007). $\mathcal{L}_{\mathrm{diag}}$ averages this distance between $\widehat{\boldsymbol{\Sigma}}_{ii}$ and $\boldsymbol{\Sigma}_{ii}$, whereas $\mathcal{L}_{\mathrm{marg}}$ applies it directly to $\boldsymbol{M}_i$. The two terms differ when $\alpha < 1$, because $\widehat{\boldsymbol{\Sigma}}_{ii} = \boldsymbol{P}_i \#_\alpha \boldsymbol{M}_i$ also depends on the elastic marginal $\boldsymbol{P}_i$. Training protocols are detailed in Appendix C.

**Covariance-based sampling.** The predicted covariance $\widehat{\boldsymbol{\Sigma}}$ defines a Gaussian distribution over $\mathrm{C}_\alpha$ displacements around $\boldsymbol{X}$. For the $k$-th sample, we draw a displacement vector:

$$\boldsymbol{d}^{(k)} = \widehat{\boldsymbol{\Sigma}}^{1/2}\mathbf{z}^{(k)}, \qquad \mathbf{z}^{(k)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_{3N}), \tag{10}$$

and form the sampled conformation $\widetilde{\boldsymbol{X}}^{(k)}$ with coordinates $\widetilde{\boldsymbol{x}}_i^{(k)} = \boldsymbol{x}_i + \boldsymbol{d}_i^{(k)}$, where $\boldsymbol{d}_i^{(k)} \in \mathbb{R}^3$ is the $i$-th block of $\boldsymbol{d}^{(k)}$. The matrix square root is computed by eigendecomposition, with small negative eigenvalues from finite-precision arithmetic clipped to zero.

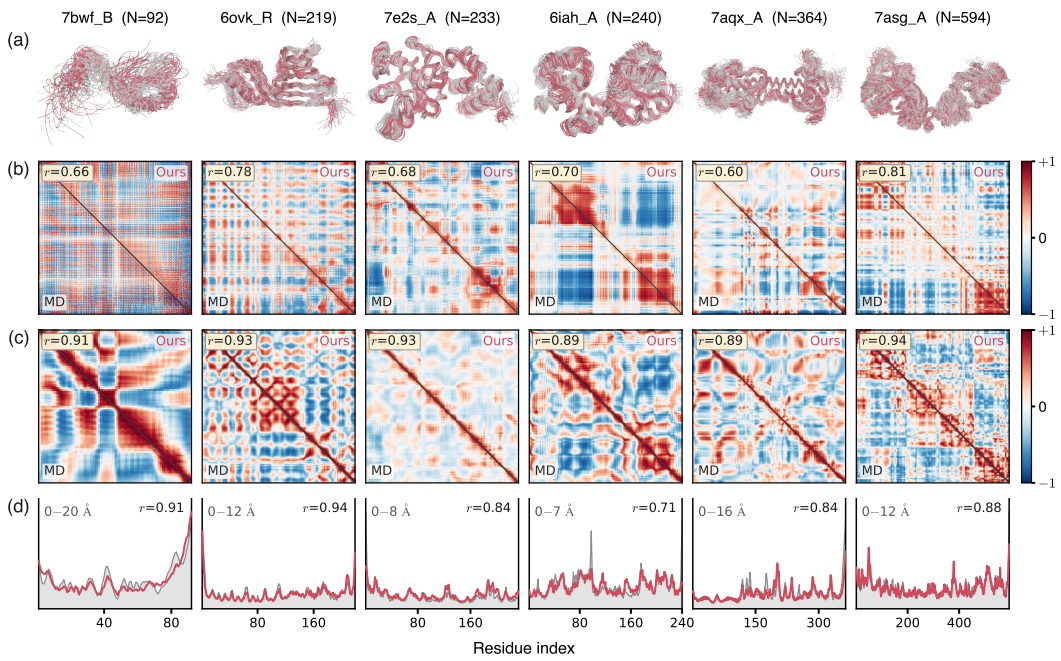


Figure 2: **ENMulator predictions of ensemble and covariance-derived dynamics compared with all-atom MD.** All four views are derived from the predicted covariance. **(a) Ensemble:** conformations sampled from the predicted covariance (red) are superposed with MD frames (grey) on the same reference backbone. **(b) Covariance** and **(c) DCCM:** each map compares MD (lower triangle) with the prediction (upper triangle); covariance maps are individually clipped for visualization. **(d) RMSF:** predicted and MD residue-wise fluctuation profiles are compared across residues. For covariance, DCCM, and RMSF, $r$ denotes the Pearson correlation with MD.

**Geometric relaxation.** Gaussian sampling reproduces the predicted second-order displacement statistics but does not enforce nonlinear geometric constraints. Raw conformations may therefore exhibit distorted consecutive $C_\alpha$ distances or steric clashes between sequence-distant residues. We apply a deterministic geometric relaxation to reduce these artifacts.

Let $d_i^0 = \|\boldsymbol{x}_{i+1} - \boldsymbol{x}_i\|_2$ denote the consecutive-residue distance in the reference structure. For a conformation $\boldsymbol{Y}$ with residue coordinates $\boldsymbol{y}_i$, we define the relaxation energy as:

$$\mathcal{E}_{\text{relax}}(\boldsymbol{Y}) = \sum_{i=1}^{N-1} \left(\|\boldsymbol{y}_{i+1} - \boldsymbol{y}_i\|_2 - d_i^0\right)^2 + \sum_{j-i\geq s_{\min}} [d_{\text{cl}} - \|\boldsymbol{y}_i - \boldsymbol{y}_j\|_2]_+^2 . \tag{11}$$

Here, $[z]_+ = \max(z, 0)$, $s_{\min}$ is the minimum sequence separation for clash evaluation, and $d_{\text{cl}}$ is the clash-distance threshold. The first term restores consecutive $C_\alpha$ distances toward their reference values, while the second penalizes clashes between sequence-distant residues. Starting from $\boldsymbol{Y} = \widetilde{\boldsymbol{X}}^{(k)}$, we minimize Eq. 11 for a fixed number of gradient steps to obtain $\boldsymbol{X}^{(k)}$. The relaxation leaves $\widehat{\boldsymbol{\Sigma}}$ unchanged and is applied identically to every covariance-based method (Appendix C.3).

## 5 Experiments

We evaluate ENMulator at the covariance level (Section 5.2), at the ensemble level (Section 5.3) and on the directional coupling that one-scalar-per-pair predictors cannot express (Section 5.4).

### 5.1 Experimental Setup and Baselines

We train and evaluate ENMulator on the ATLAS molecular-dynamics (MD) dataset (Vander Meersche et al., 2024), following the AlphaFlow (Jing et al., 2024a) primary split and the preprocessing

Table 2: **ENMulator recovers MD-derived covariance structure and correlated motion.** Across the ATLAS test split, our model achieves the strongest global covariance reconstruction and inter-residue DCCM agreement among the evaluated methods, while also recovering collective modes and residue-wise fluctuations. Values are medians across 82 test proteins; arrows indicate the preferred direction. For participation ratio, $(= x)$ denotes the median MD target, and values closer to this target are preferred. Cell color denotes best, second, third within each row.

| Metric | *Hand-designed ENM* GNM | ANM | ANM* | pfANM | Hinsen | *Learning-based methods* DynaProt | BackFlip-2 | **ENMulator** |
|---|---|---|---|---|---|---|---|---|
| *Global covariance accuracy* | | | | | | | | |
| Relative Frobenius (↓) | 0.978 | 0.984 | 0.962 | 0.988 | 0.990 | 0.954 | 0.910 | 0.874 |
| Bures–Wasserstein $\mathcal{W}_2$ norm. (↓) | 0.928 | 0.839 | 0.856 | 0.978 | 0.867 | 0.825 | 0.752 | 0.666 |
| Pearson correlation $r$ (↑) | 0.292 | 0.448 | 0.468 | 0.348 | 0.496 | 0.417 | 0.516 | 0.589 |
| *Collective modes and spectrum* | | | | | | | | |
| Top-3 subspace overlap (↑) | 0.266 | 0.551 | 0.593 | 0.474 | 0.603 | 0.441 | 0.527 | 0.639 |
| Top-10 subspace overlap (↑) | 0.448 | 0.564 | 0.602 | 0.512 | 0.636 | 0.526 | 0.625 | 0.663 |
| Participation ratio ($= 5.48$) | 50.34 | 46.03 | 33.11 | 213.70 | 19.52 | 40.69 | 17.17 | 8.98 |
| *Per-residue marginals (diagonal) and cross-residue coupling (off-diagonal)* | | | | | | | | |
| RMSF $r$ (↑) | 0.754 | 0.764 | 0.790 | 0.768 | 0.801 | 0.861 | 0.877 | 0.859 |
| RMSF MAE (↓) | 0.774 | 1.001 | 0.600 | 1.179 | 1.292 | 0.539 | 0.445 | 0.433 |
| DCCM $r$ (↑) | 0.566 | 0.779 | 0.791 | 0.668 | 0.824 | 0.579 | 0.797 | 0.856 |
| DCCM MAE (↓) | 0.233 | 0.205 | 0.197 | 0.246 | 0.166 | 0.235 | 0.167 | 0.138 |

protocol used by DynaProt (Bafna et al., 2026). The resulting split contains 1,242 training, 39 validation, and 82 test proteins. The MD target for each protein is the covariance of Section 3, computed from its aligned trajectories (Appendix C.1).

To test whether learned elastic interactions improve on established ENMs, we compare with GNM, ANM, ANM*, pfANM, and Hinsen, spanning uniform-contact, two-scale, and distance-dependent interaction schemes (Appendix D). We further compare with the learning-based covariance predictors DynaProt (Bafna et al., 2026) and BackFlip-2 (Viliuga et al., 2026). For ensemble evaluation, we additionally include the direct generation methods AlphaFlow (Jing et al., 2024a) and BBFlow (Wolf et al., 2025). We draw 250 $C_\alpha$ conformations per protein from each covariance-based method and reconstruct full-atom structures with PULCHRA (Rotkiewicz & Skolnick, 2008).

### 5.2 Recovering Covariance Structure and Collective Modes

Table 2 evaluates complementary properties of $\widehat{\boldsymbol{\Sigma}}$ against the MD covariance. Relative Frobenius error is an entrywise, scale-sensitive comparison, whereas the normalized Bures–Wasserstein distance compares the corresponding Gaussian distributions. Top-$k$ subspace overlap assesses the dominant collective-motion directions (Knyazev & Argentati, 2002), while participation ratio summarizes how broadly the predicted variance is distributed across modes. RMSF and DCCM summarize the diagonal blocks (residue fluctuations) and the normalized off-diagonal blocks (correlated motion) of the covariance, respectively. Metrics are defined in Appendix E.

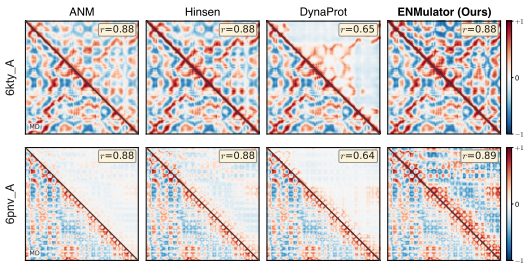


Figure 3: **DCCM predictions compared with all-atom MD.** For two ATLAS test proteins, `6kty_A` ($N = 196$) and `6pnv_A` ($N = 411$), each panel compares the predicted DCCM (upper triangle) with the MD DCCM (lower triangle).

ENMulator achieves the strongest overall recovery of the MD covariance, reducing relative Frobenius error (0.910 to 0.874) and Bures–Wasserstein distance (0.752 to 0.666) relative to the next-best method, while attaining the highest collective-mode agreement. Its participation ratio is also substantially closer to the MD target (8.98 versus 17.17; MD: 5.48). The clearest advantage is in inter-residue coupling: ENMulator reaches a DCCM correlation of 0.856 and MAE of 0.138, compared with 0.824 and 0.166 for the next-best method. At the residue level, it attains the lowest RMSF MAE, with RMSF correlation comparable to the learned baselines. These gains are broadly consis-

Table 3: **Covariance-based sampling translates predicted dynamics into realistic conformational ensembles.** Among covariance-based methods, ENMulator closely matches MD in ensemble spread and $C_\alpha$ RMSF, yields the strongest global and per-target DCCM agreement, and retains structural validity with a median chain break rate of zero. All covariance-based methods use the same sampling and geometric-relaxation protocol. Target-level values are medians across test proteins, whereas global correlations are computed after concatenating test-set results. Runtime is measured for sampling 250 conformations of `6kty_A`. Arrows indicate the preferred direction, and $(= x)$ denotes the MD reference value. The full metric set is reported in Appendix Table A16.

| Metric | *Hand-designed ENM* ANM | Hinsen | *Learning-based covariance methods* DynaProt | BackFlip-2 | **ENMulator** | *Direct generation* AlphaFlow | BBFlow |
|---|---|---|---|---|---|---|---|
| *Efficiency* | | | | | | | |
| Runtime (s) (↓) | 8.15 | 8.80 | 7.57 | 7.58 | 7.53 | 5,149.30 | 172.20 |
| Relative runtime (↓) | ×1.08 | ×1.17 | ×1.01 | ×1.01 | ×1.00 | ×684 | ×22.9 |
| # Parameters (↓) | — | — | 2.9M | 1.0M | 2.8M | 97M | 18.2M |
| *Ensemble spread and residue flexibility* | | | | | | | |
| Pairwise RMSD (= 2.91) | 0.79 | 2.51 | 2.01 | 2.29 | 2.38 | 2.18 | 3.09 |
| Pairwise RMSD $r$ (↑) | 0.64 | 0.19 | 0.71 | 0.81 | 0.83 | 0.94 | 0.82 |
| $C_\alpha$ RMSF (= 1.48) | 0.45 | 0.58 | 1.04 | 1.20 | 1.40 | 1.17 | 1.49 |
| Per-target RMSF $r$ (↑) | 0.77 | 0.81 | 0.87 | 0.88 | 0.87 | 0.92 | 0.90 |
| *Correlated motion and distributional accuracy* | | | | | | | |
| Global DCCM $r$ (↑) | 0.69 | 0.77 | 0.58 | 0.76 | 0.84 | 0.88 | 0.84 |
| Per-target DCCM $r$ (↑) | 0.74 | 0.82 | 0.68 | 0.80 | 0.86 | 0.89 | 0.87 |
| RMWD (↓) | 2.66 | 2.73 | 2.39 | 2.25 | 2.32 | 1.71 | 1.96 |
| Variance contrib. (↓) | 1.51 | 1.77 | 1.14 | 0.93 | 0.92 | 0.80 | 0.95 |
| MD PCA $\mathcal{W}_2$ (↓) | 1.93 | 1.81 | 1.81 | 1.67 | 1.63 | 1.29 | 1.30 |
| PC similarity $> 0.5$ (%) (↑) | 28.0 | 30.5 | 19.5 | 29.3 | 37.8 | 43.9 | 40.2 |
| *Ensemble observables and structural validity* | | | | | | | |
| Transient contacts $J$ (↑) | 0.30 | 0.26 | 0.30 | 0.31 | 0.35 | 0.47 | 0.28 |
| Conf. with $C_\alpha$–$C_\alpha$ break (%) (↓) | 0.00 | 0.40 | 0.80 | 3.20 | 0.00 | 1.60 | 0.00 |

tent across proteins against other baselines (Appendix G.1). Ablations support the intended division of roles: learned springs raise DCCM correlation from $0.779$ to $0.851$, whereas the marginal branch raises RMSF correlation from $0.764$ to $0.842$ with essentially unchanged DCCM (Appendix G.3). The advantage is not explained by global scale alone: even after oracle rescaling, the best classical baseline, Hinsen, has a higher Bures–Wasserstein distance than unscaled ENMulator ($0.702$ versus $0.666$; Table A14). Together, these results indicate that learned elastic interactions improve global covariance structure and correlated motion, while marginal calibration improves local flexibility. Figures 2 and 3 provide qualitative comparisons.

### 5.3 Recovering Correlated Motions in Sampled Ensembles

We next evaluate whether the covariance-level fidelity observed above translates into realistic sampled conformational ensembles. For each predicted covariance, we generate an ensemble by Gaussian sampling followed by geometric relaxation to improve local structural validity. Table 3 compares the resulting ensembles.

With efficiency comparable to the other covariance-based methods, ENMulator best recovers the MD ensemble scale, achieving the closest pairwise RMSD correlation and $C_\alpha$ RMSF to the MD targets and the highest global pairwise-RMSD correlation among covariance-based approaches. Although BackFlip-2 attains a marginally higher per-target RMSF correlation, ENMulator more faithfully reproduces the magnitude and protein-wise variation of residue flexibility. Direct-generation models achieve higher DCCM agreement on some metrics, but at greater sampling cost ($22.9\times$–$684\times$); meanwhile, ENMulator recovers transient contacts better than all covariance-based baselines. Together, these results show that accurate covariance prediction yields efficient, structurally valid ensembles with realistic flexibility and correlated motion.

### 5.4 Evaluating Directional Residue-Pair Coupling

Section 1 argued that one-scalar-per-pair predictors cannot express direction-dependent coupling; here we quantify that limitation and test whether learned elastic interactions overcome it. DCCM measures whether two residues move in the same direction at the same time (Ichiye & Karplus,

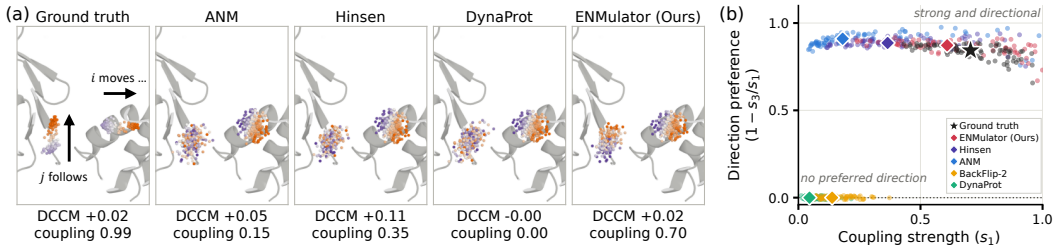


Figure 4: **Residue-pair coupling.** **(a)** Motion of one residue pair; each dot is an MD frame or a model sample, and the numbers report its DCCM and coupling strength $s_1$. As $i$ moves right (purple to orange), $j$ moves up in step. The two are tightly coupled although they do not move in parallel. **(b)** Coupling strength and direction preference of each test protein; large markers are medians.

Table 4: **ENMulator improves directional residue-pair coupling recovery.** Coupling strength $s_1$ measures the correlation along residues' best-matched directions, whereas direction preference $1 - s_3/s_1$ measures anisotropy between the strongest and weakest coupling direction. Coupling map is the Pearson correlation between predicted and MD $s_1$ across residue pairs, and coupling error $e_{\mathrm{cc}}$ is the mean absolute error of all three canonical correlations. Values are medians over 82 test proteins; $(= x)$ denotes MD. †Lower bound on $e_{\mathrm{cc}}$ for reconstructions constrained to equal canonical correlations within each residue pair (Appendix F).

| Method | Coupling model | Coupling strength $s_1$ $(=0.71)$ | Direction preference $1 - s_3/s_1$ $(=0.84)$ | Coupling map $r$ (↑) | Coupling error $e_{\mathrm{cc}}$ (↓) |
|---|---|---|---|---|---|
| DynaProt | one scalar per pair | 0.045 | 0.000 | 0.408 | 0.335 |
| BackFlip-2 | one scalar per pair | 0.138 | 0.000 | 0.494 | 0.280 |
| *Lower bound*† | one scalar per pair | – | 0.000 | – | $\geq 0.182$ |
| ANM | prescribed springs | 0.181 | 0.910 | 0.406 | 0.269 |
| Hinsen | prescribed springs | 0.365 | 0.886 | 0.580 | 0.185 |
| **ENMulator** | learned springs | 0.609 | 0.872 | 0.568 | 0.121 |

1991). However, residues can be tightly coupled even when DCCM $\approx 0$: if $j$ always moves up whenever $i$ moves right, the two are coupled without moving in parallel (Figure 4a). We therefore measure *coupling strength* $s_1$, the largest canonical correlation between two residue displacements, and *direction preference* $1 - s_3/s_1$, which is near 1 when coupling is concentrated along one pair of directions and 0 when it is similar in every direction (Appendix F). Both are invariant to per-residue fluctuation amplitudes and, by Section 4.2, isolate the coupling induced by the learned elastic interactions. In MD, coupling is typically strong and directional and remains appreciable with distance (Figure A3). DynaProt and BackFlip-2 represent each pair by one scalar, so $s_1 = s_2 = s_3$ by construction; this also bounds the coupling error of such model below by 0.182 (Table 4). ANM captures directional coupling but underestimates its strength. ENMulator more closely matches MD coupling strength with similar direction preference, and its coupling error (0.121) falls below this bound. Its coupling map correlation is on par with Hinsen (0.568 vs. 0.580) at a much lower coupling error (0.121 vs. 0.185). Its coupling error is also nearly constant from adjacent to long-range pairs, whereas uniform ANM stiffnesses over-couple nearby residues (Table A8).

## 6 Conclusion

We presented ENMulator, which learns protein-specific elastic interactions from sequence and structure in place of prescribing them. It improves covariance reconstruction, recovers correlated motion in sampled ensembles, and captures direction-dependent residue coupling, suggesting that learned elastic interactions are a promising route to protein dynamics from a single reference structure.

### AI USE STATEMENT

In this work, we used generative AI tools to provide feedback on research methodology and experimental design, implementing and debugging methods, cleaning and reformatting datasets, and assisting in interpreting experimental results. We have not used generative AI tools to assist in the writing and checking of proofs, formulate mathematical claims, generate synthetic datasets, develop theoretical models or conceptual frameworks, provide critical ingredients for proving mathematical claims, propose or refine hypotheses, assist with translation, or support qualitative or thematic data analysis; the remaining disclosure tasks are not applicable to this work. Additionally, we used generative AI tools for literature retrieval and discovery and to aid the drafting, revision, and polishing of manuscript text. AI was used to refine the authors' ideas, but not to determine methodological or experimental choices; all consequential decisions were made by the authors and supported by experimental validation. All LLM-generated code was verified by two authors. The authors also reviewed AI-assisted preprocessing code to ensure that it did not alter the underlying data, and independently checked all reported numerical results and their interpretations. All AI-assisted manuscript text, including technical details, was reviewed and revised by the authors. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

### ETHICS STATEMENT

This work develops a computational method for modeling protein conformational dynamics from protein sequence and structure. It uses publicly available molecular-dynamics data and does not involve human participants, personal data, animal experimentation, or clinical decision-making. We do not raise immediate ethical concerns in the present research setting. As with other protein-design and structure-modeling tools, downstream applications should be evaluated responsibly in their specific scientific and experimental contexts. We support responsible use through clear reporting of the evaluation protocol, limitations, and code availability.

### REPRODUCIBILITY STATEMENT

Appendix B specifies the elastic graph, radial kernel, local frames, prediction heads, and numerical details of the congruence. Appendix C gives the target construction, two-stage training schedule, loss weights, and optimizer settings, and Appendix C.3 the relaxation settings. Appendices D and E describe the baseline configurations, the ensemble-generation pipeline, and all evaluation metrics. All experiments use the public ATLAS dataset with the AlphaFlow split. The supplementary materials include the training and evaluation code of ENMulator.

## REFERENCES

Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2007.

Ali Rana Atilgan, Stewart R Durell, Robert L Jernigan, Melik C Demirel, Ozlem Keskin, and Ivet Bahar. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophysical Journal*, 80(1):505–515, 2001.

Mihir Bafna, Bowen Jing, and Bonnie Berger. Learning residue level protein dynamics with multiscale gaussians. In *ICLR*, 2026.

Ivet Bahar, Ali Rana Atilgan, and Burak Erman. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding and Design*, 2(3):173–181, 1997.

Ivet Bahar, Timothy R Lezon, Lee-Wei Yang, and Eran Eyal. Global dynamics of proteins: bridging between structure and function. *Annual Review of Biophysics*, 39:23–42, 2010.

Ivet Bahar, Mary Hongying Cheng, Ji Young Lee, Cihan Kaya, and She Zhang. Structure-encoded global motions and their role in mediating protein-substrate interactions. *Biophysical Journal*, 109(6):1101–1109, 2015.

B. Brooks and M. Karplus. Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proceedings of the National Academy of Sciences*, 80(21): 6571–6575, 1983.

Qiang Cui and Ivet Bahar (eds.). *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*. CRC Press, 2005.

Igors Dubanevics and Tom CB McLeish. Optimising elastic network models for protein dynamics and allostery: Spatial and modal cut-offs and backbone stiffness. *Journal of Molecular Biology*, 434(17):167696, 2022.

Jacob D Durrant and J Andrew McCammon. Molecular dynamics simulations and drug discovery. *BMC Biology*, 9(1):71, 2011.

N. Go, T. Noguti, and T. Nishikawa. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proceedings of the National Academy of Sciences*, 80(12):3696–3700, 1983.

Nina M Goodey and Stephen J Benkovic. Allosteric regulation and catalysis emerge via a common route. *Nature Chemical Biology*, 4(8):474–482, 2008.

Turkan Haliloglu, Ivet Bahar, and Burak Erman. Gaussian dynamics of folded proteins. *Physical Review Letters*, 79(16):3090–3093, 1997.

Konrad Hinsen, Andrei-Jose Petrescu, Serge Dellerue, Marie-Claire Bellissent-Funel, and Gerald R. Kneller. Harmonicity in slow protein dynamics. *Chemical Physics*, 261(1-2):25–37, 2000.

Scott A Hollingsworth and Ron O Dror. Molecular dynamics simulation for all. *Neuron*, 99(6): 1129–1143, 2018.

Toshiko Ichiye and Martin Karplus. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Structure, Function, and Bioinformatics*, 11(3):205–217, 1991.

Bowen Jing, Bonnie Berger, and Tommi Jaakkola. Alphafold meets flow matching for generating protein ensembles. In *ICML*, 2024a.

Bowen Jing, Hannes Stärk, Tommi Jaakkola, and Bonnie Berger. Generative modeling of molecular dynamics trajectories. In *NeurIPS*, 2024b.

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.

Andrew V Knyazev and Merico E Argentati. Principal angles between subspaces in an A-based scalar product: algorithms and perturbation estimates. *SIAM Journal on Scientific Computing*, 23 (6):2008–2040, 2002.

Patrice Koehl, Henri Orland, and Marc Delarue. Parameterizing elastic network models to capture the dynamics of proteins. *Journal of Computational Chemistry*, 42(23):1643–1661, 2021.

Petr Kouba, Joan Planas-Iglesias, Jiri Damborsky, Jiri Sedlar, Stanislav Mazurenko, and Josef Sivic. Learning to engineer protein flexibility. In *ICLR*, 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, Bozitao Zhong, Zuobai Zhang, and Jian Tang. Str2Str: A score-based framework for zero-shot protein conformation sampling. In *ICLR*, 2024.

Edward Lyman, Jim Pfaendtner, and Gregory A. Voth. Systematic multiscale parameterization of heterogeneous elastic network models of proteins. *Biophysical Journal*, 95(9):4183–4192, 2008.

Dengming Ming and Michael E. Wall. Allostery in a coarse-grained model of protein dynamics. *Physical Review Letters*, 95(19):198103, 2005.

Kei Moritsugu and Jeremy C. Smith. Coarse-grained biomolecular simulation with REACH: Realistic extension algorithm via covariance hessian. *Biophysical Journal*, 93(10):3460–3469, 2007.

Ruth Nussinov, Yonglan Liu, Wengang Zhang, and Hyunbum Jang. Protein conformational ensembles in function: roles and mechanisms. *RSC Chemical Biology*, 4(11):850–864, 2023.

Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.

Piotr Rotkiewicz and Jeffrey Skolnick. Fast procedure for reconstruction of full-atom protein models from reduced representations. *Journal of Computational Chemistry*, 29(9):1460–1465, 2008.

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.

Lars Skjaerven, Siv M. Hollup, and Nathalie Reuter. Normal mode analysis for proteins. *Journal of Molecular Structure: THEOCHEM*, 898(1–3):42–48, 2009.

Florence Tama and Y-H Sanejouand. Conformational change of proteins arising from normal mode calculations. *Protein Engineering*, 14(1):1–6, 2001.

Simon J Teague. Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery*, 2(7):527–541, 2003.

Monique M. Tirion. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Physical Review Letters*, 77(9):1905–1908, 1996.

Kathryn Tunyasuvunakool, Jonas Adler, Zachary Wu, Tim Green, Michal Zielinski, Augustin Žídek, Alex Bridgland, Andrew Cowie, Clemens Meyer, Agata Laydon, et al. Highly accurate protein structure prediction for the human proteome. *Nature*, 596(7873):590–596, 2021.

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.

Mihaly Varadi, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, Oana Stroe, Gemma Wood, Agata Laydon, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1):D439–D444, 2022.

Mihaly Varadi, Damian Bertoni, Paulyna Magana, Urmila Paramval, Ivanna Pidruchna, Malarvizhi Radhakrishnan, Maxim Tsenkov, Sreenath Nair, Milot Mirdita, Jingi Yeo, et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Research*, 52(D1):D368–D375, 2024.

Vsevolod Viliuga, Leif Seute, Matteo Tadiello, Nicolas Wolf, Frauke Gräter, and Arne Elofsson. Predicting directional flexibility in proteins. *arXiv preprint arXiv:2609.08474*, 2026.

Hannah K. Wayment-Steele, Gina El Nesr, Ramith Hettiarachchi, Adedolapo M. Ojoawo, Hasindu Kariyawasam, Sergey Ovchinnikov, and Dorothee Kern. Learning millisecond protein dynamics from what is missing in NMR spectra. *Nature*, 2026. Advance online publication.

Nicolas Wolf, Leif Seute, Vsevolod Viliuga, Simon Wagner, Jan Stühmer, and Frauke Gräter. Learning conformational ensembles of proteins based on backbone geometry. In *NeurIPS*, 2025.

Lei Yang, Guang Song, and Robert L Jernigan. How well can we understand large-scale protein motions using normal modes of elastic network models? *Biophysical Journal*, 93(3):920–929, 2007.

Lei Yang, Guang Song, and Robert L. Jernigan. Protein elastic network models and the ranges of cooperativity. *Proceedings of the National Academy of Sciences*, 106(30):12347–12352, 2009.

# A CODE

We provide the implementation of ENMulator for reproducibility. Please refer to `README.md` in supplementary material for instructions on running the code for training and evaluation. We will release the model weights to the public upon acceptance.

# B MODEL SPECIFICATION AND MATHEMATICAL DETAILS

## B.1 ELASTIC GRAPH AND RADIAL KERNEL

The elastic graph contains unordered residue pairs within a distance cutoff and always includes consecutive residues:

$$\mathcal{E} = \big\{(i,j) : i < j,\ r_{ij} \le r_c\big\} \cup \big\{(i,i+1) : 1 \le i < N\big\}. \tag{A1}$$

The latter term maintains chain connectivity even when spatial contacts are sparse. We use $r_c = 15$ Å. The Hessian sum in Eq. 2 is evaluated once per unordered pair.

The shared distance contribution $\kappa$ is a learned piecewise-constant kernel. Let $b(r) \in \{0, \dots, 15\}$ be the bin index obtained from the 15 interior edges of an equally spaced partition of $[3.5, 15]$ Å. Then

$$\kappa(r) = w_{b(r)}, \qquad b(r) = \mathrm{bucketize}\left(r; \left\{3.5 + \frac{11.5m}{16}\right\}_{m=1}^{15}\right). \tag{A2}$$

The 16 values $\{w_b\}$ are initialized to zero, as is the global log-stiffness offset $\beta$. Distances below the first interior edge and above the last interior edge use the first and final bins, respectively.

## B.2 RESIDUE-LOCAL FRAMES

For residue $i$, let $\boldsymbol{x}_i^{\mathrm{N}}$, $\boldsymbol{x}_i$, and $\boldsymbol{x}_i^{\mathrm{C}}$ denote the backbone N, $C_\alpha$, and C coordinates. We construct a right-handed local frame $\boldsymbol{R}_i \in \mathrm{SO}(3)$ from the two backbone bond vectors

$$\mathbf{v}_{i,1} = \boldsymbol{x}_i^{\mathrm{N}} - \boldsymbol{x}_i, \qquad \mathbf{v}_{i,2} = \boldsymbol{x}_i^{\mathrm{C}} - \boldsymbol{x}_i. \tag{A3}$$

Writing $\mathbf{e}_{i,1} = \mathbf{v}_{i,1}/\|\mathbf{v}_{i,1}\|_2$, we remove the component of $\mathbf{v}_{i,2}$ parallel to $\mathbf{e}_{i,1}$ and normalize the remainder:

$$\mathbf{e}_{i,2} = \frac{\mathbf{v}_{i,2} - (\mathbf{e}_{i,1}^\top \mathbf{v}_{i,2})\mathbf{e}_{i,1}}{\|\mathbf{v}_{i,2} - (\mathbf{e}_{i,1}^\top \mathbf{v}_{i,2})\mathbf{e}_{i,1}\|_2}, \qquad \boldsymbol{R}_i = \left[\mathbf{e}_{i,1} \quad \mathbf{e}_{i,2} \quad \mathbf{e}_{i,1} \times \mathbf{e}_{i,2}\right]. \tag{A4}$$

Under a global rotation $\boldsymbol{Q} \in \mathrm{SO}(3)$, this construction gives $\boldsymbol{R}_i \mapsto \boldsymbol{Q}\boldsymbol{R}_i$. Consequently, a covariance predicted in the local frame transforms as a Cartesian covariance after multiplication by $\boldsymbol{R}_i$, as used in Eq. 5.

## B.3 PROPERTIES OF BLOCKWISE GEODESIC CONGRUENCE

Equation 6 applies an invertible, residue-block-diagonal congruence to the elastic covariance. It therefore preserves positive semidefiniteness and rank:

$$\boldsymbol{z}^\top \widehat{\boldsymbol{\Sigma}} \boldsymbol{z} = (\boldsymbol{D}^\top \boldsymbol{z})^\top \boldsymbol{\Sigma}_0 (\boldsymbol{D}^\top \boldsymbol{z}) \ge 0, \qquad \mathrm{rank}(\widehat{\boldsymbol{\Sigma}}) = \mathrm{rank}(\boldsymbol{\Sigma}_0). \tag{A5}$$

Equation 7 further gives $\widehat{\boldsymbol{\Sigma}}_{ii} = \boldsymbol{P}_i \#_\alpha \boldsymbol{M}_i$: the final marginal equals the elastic marginal at $\alpha = 0$ and the learned marginal prediction at $\alpha = 1$. The same congruence preserves a precise measure of cross-residue dependence. For residues $i \ne j$, let

$$\boldsymbol{C}_{ij} = \boldsymbol{P}_i^{-1/2} [\boldsymbol{\Sigma}_0]_{ij} \boldsymbol{P}_j^{-1/2}$$

be the whitened elastic cross-covariance. The corresponding quantity after congruence satisfies

$$\widehat{\boldsymbol{C}}_{ij} := \widehat{\boldsymbol{\Sigma}}_{ii}^{-1/2} \widehat{\boldsymbol{\Sigma}}_{ij} \widehat{\boldsymbol{\Sigma}}_{jj}^{-1/2} = \boldsymbol{U}_i \boldsymbol{C}_{ij} \boldsymbol{U}_j^\top, \tag{A6}$$

Table A1: **Effect of congruence within the trained ENMulator model** ($\alpha = 0.75$). The base covariance $\boldsymbol{\Sigma}_0$ and the final covariance $\widehat{\boldsymbol{\Sigma}}$ are taken from the same model. Values are medians over 82 test proteins. Paired changes are evaluated per protein before aggregation and therefore need not equal the difference between displayed medians.

| Metric | Base $\boldsymbol{\Sigma}_0$ | Final $\widehat{\boldsymbol{\Sigma}}$ | Paired change [95% CI] |
|---|---|---|---|
| Marginal log-Euclidean ↓ | 2.649 | 1.983 | $-0.452$ [$-0.686$, $-0.282$] |
| RMSF MAE (Å) ↓ | 0.551 | 0.433 | $-0.049$ [$-0.087$, $-0.026$] |
| DCCM MAE (off-diagonal) ↓ | 0.141 | 0.138 | $-0.001$ [$-0.003$, $-6.3\times10^{-4}$] |
| Relative Frobenius error ↓ | 0.949 | 0.874 | $-0.060$ [$-0.076$, $-0.038$] |
| Normalized BW distance ↓ | 0.755 | 0.666 | $-0.073$ [$-0.085$, $-0.050$] |

where $\boldsymbol{U}_i = \widehat{\boldsymbol{\Sigma}}_{ii}^{-1/2}\boldsymbol{D}_i\boldsymbol{P}_i^{1/2}$ is orthogonal, and likewise for $\boldsymbol{U}_j$. Hence the singular values of $\boldsymbol{C}_{ij}$, i.e., the three canonical correlations between the two residue displacements, are exactly invariant under congruence. Thus, the elastic covariance determines the pairwise canonical-correlation spectra, while congruence calibrates local anisotropic fluctuations without changing those spectra.

This invariance is specific to canonical correlations. Cartesian DCCM, covariance eigenvectors, and the rigid-body null space may change under the residue-dependent maps in Eq. 6. Congruence is therefore a local marginal calibration of the elastic covariance, not a transformation that preserves every global covariance observable. Table A1 isolates the effect of congruence within the trained ENMulator model by comparing its base covariance $\boldsymbol{\Sigma}_0$ with its final covariance $\widehat{\boldsymbol{\Sigma}}$. Congruence improves marginal accuracy and the global covariance distances while leaving DCCM nearly unchanged, consistent with the division of roles in Section 4.2. It does not improve every covariance property: the median trace ratio changes from $1.04$ to $0.76$.

### B.4 Rigid-Body Modes

The rigid-body basis $\boldsymbol{Q}_6$ in Eq. 4 contains three translations and three infinitesimal rotations about the centroid of the reference $C_\alpha$ coordinates. We construct these six vectors explicitly and orthonormalize them with QR decomposition. When the elastic Hessian has no additional null modes, Eq. 4 is the Moore–Penrose pseudoinverse of $\boldsymbol{H}$; the base covariance therefore satisfies $\boldsymbol{\Sigma}_0\boldsymbol{Q}_6 = \boldsymbol{0}$. Consecutive-residue edges maintain chain connectivity, but do not by themselves rule out additional infinitesimal mechanisms.

Congruence preserves the rank of $\boldsymbol{\Sigma}_0$, but need not preserve its original null space:

$$\ker \widehat{\boldsymbol{\Sigma}} = \boldsymbol{D}^{-\top} \ker \boldsymbol{\Sigma}_0 = \boldsymbol{D}^{-\top} \operatorname{range}(\boldsymbol{Q}_6). \tag{A7}$$

Thus, residue-dependent anisotropic maps can assign variance to the original rigid-body directions even though the final covariance remains singular with the same rank. We quantify this effect by the fraction of total variance in the original rigid subspace,

$$r_{\mathrm{rigid}} = \frac{\operatorname{tr}(\boldsymbol{Q}_6^\top \widehat{\boldsymbol{\Sigma}} \boldsymbol{Q}_6)}{\operatorname{tr}(\widehat{\boldsymbol{\Sigma}})}. \tag{A8}$$

At $\alpha = 0.75$, the median $r_{\mathrm{rigid}}$ is $8.7\times10^{-3}$ across the 82 test proteins (95th percentile: $2.4\times10^{-2}$). This leakage is introduced by the residue-local calibration; for $\boldsymbol{\Sigma}_0$ and for the aligned MD covariances, $r_{\mathrm{rigid}}$ is at floating-point precision. We retain the unprojected covariance throughout the reported evaluation. Projecting $\widehat{\boldsymbol{\Sigma}}$ onto the complement of $\operatorname{range}(\boldsymbol{Q}_6)$ would remove this leakage, but would also alter the prescribed marginals and the canonical correlations preserved by congruence.

All covariance solves and matrix powers are evaluated in float64. The eigenvalue floor used for spectral operations is inactive for every block in the reported test evaluation.

### B.5 Shared Encoder and Prediction Heads

The shared encoder $E_{\boldsymbol{\theta}}$ takes a learned embedding of each residue type (21 types) and the backbone frames as input and produces rotation-invariant, 128-dimensional residue representations $\boldsymbol{h}_i$. It

consists of eight invariant-point attention layers, each with eight attention heads of scalar width 32, four query points, eight value points, and transition multiplier 4. Dropout with rate 0.1 is used in the shared encoder and spring head.

**Spring head.** The spring head maps each $\boldsymbol{h}_i$ to one scalar through two hidden layers of width 128 with SiLU activations, followed by a scalar output layer. Its endpoint-additive form $a_i + a_j$ guarantees symmetry of $\gamma_{ij}$ while retaining a residue-wise head. The scalar output layer is initialized to zero; the residue term is bounded as $a_i = 6\tanh(\widetilde{a}_i/6)$, i.e., $A = 6$ in Eq. 3, while the global offset $\beta$ and the distance-kernel parameters are initialized at zero and remain unbounded.

**Marginal head.** The marginal head maps $\boldsymbol{h}_i$ to six scalars and arranges them into a lower-triangular Cholesky factor

$$\boldsymbol{L}_i = \begin{pmatrix} \mathrm{softplus}(u_{i,1}) + \epsilon & 0 & 0 \\ u_{i,4} & \mathrm{softplus}(u_{i,2}) + \epsilon & 0 \\ u_{i,5} & u_{i,6} & \mathrm{softplus}(u_{i,3}) + \epsilon \end{pmatrix}. \tag{A9}$$

The strictly positive diagonal makes $\boldsymbol{L}_i\boldsymbol{L}_i^\top$ positive definite. We use $\epsilon = 10^{-4}$. The marginal head is a 128-wide SiLU hidden layer followed by six outputs. Because $\boldsymbol{h}_i$ is invariant, $\boldsymbol{L}_i\boldsymbol{L}_i^\top$ is a covariance in the local frame, and Eq. 5 maps it to global coordinates.

# C Training and Data Protocol

## C.1 Reference-Centered Target Covariances

For each trajectory, we first align every frame to its corresponding reference $C_\alpha$ structure $\boldsymbol{X}_{\mathrm{ref}}$ (Section 3). We then represent frame $t$ by the stacked displacement vector $\Delta\boldsymbol{x}^{(t)} \in \mathbb{R}^{3N}$ relative to that reference. The training target is the reference-centered displacement second moment

$$\boldsymbol{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \Delta\boldsymbol{x}^{(t)}\Delta\boldsymbol{x}^{(t)\top}. \tag{A10}$$

Writing $\bar{\boldsymbol{d}} = T^{-1}\sum_t \Delta\boldsymbol{x}^{(t)}$, this target decomposes as

$$\boldsymbol{\Sigma} = \frac{1}{T}\sum_t (\Delta\boldsymbol{x}^{(t)} - \bar{\boldsymbol{d}})(\Delta\boldsymbol{x}^{(t)} - \bar{\boldsymbol{d}})^\top + \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^\top.$$

It therefore includes displacement of the trajectory mean from the reference. We use *displacement covariance* for this reference-centered second moment under the harmonic approximation, rather than assuming that it equals the mean-centered MD covariance. The matrix-level RMSF and DCCM descriptors below follow this same convention. Using a common reference fixes the global coordinate gauge across the trajectory and matches the prediction setting, in which ENMulator receives the reference structure as input. The alignment removes global translations and, to first order, rotations; consequently, the full target covariance is rank deficient, although its residue-wise $3 \times 3$ blocks are typically positive definite. We therefore evaluate the full-covariance loss in the ambient matrix space, while applying log-domain losses only to the residue-wise blocks.

## C.2 Optimization and Model Selection

We train ENMulator in two stages. The first stage trains the shared encoder and marginal branch alone, which provides a stable initialization of residue-wise anisotropic fluctuations. The second stage activates the spring branch and blockwise congruence and jointly optimizes the local marginals and the globally coupled covariance. Table A2 lists the loss weights in the two stages.

Stage 1 and Stage 2 run for 20 and 100 epochs, respectively. We use AdamW with $(\beta_1, \beta_2) = (0.9, 0.999)$, five epochs of linear warm-up, and cosine decay to $10^{-6}$. The radial kernel and global stiffness parameter use learning rate 0.05 without weight decay. All other non-marginal parameters use learning rate $10^{-3}$; the marginal head uses $10^{-3}$ in Stage 1 and $10^{-4}$ in Stage 2. We use gradient clipping at 1.0 and construct batches subject to $\sum N^3 \le 4\times10^8$, with at most 16 proteins per batch. We report the Stage 2 model with the lowest validation loss; it has 2.816M parameters and trains in about 1.5 hours on a single NVIDIA A100 GPU. The interpolation strength $\alpha = 0.75$ is selected on the validation split.

Table A2: **Two-stage training objective for ENMulator.**

| Loss term | Stage 1 | Stage 2 |
|---|---|---|
| Full-covariance loss, $\mathcal{L}_{\mathrm{cov}}$ | 0 | 1 |
| Calibrated-marginal loss, $\lambda_{\mathrm{d}}\mathcal{L}_{\mathrm{diag}}$ | 0 | 0.5 |
| Marginal-head loss, $\lambda_{\mathrm{m}}\mathcal{L}_{\mathrm{marg}}$ | 0.5 | 0.5 |
| Scale loss, $\lambda_{\mathrm{s}}\mathcal{L}_{\mathrm{scale}}$ | 0 | 0.1 |

Table A3: **Geometric-relaxation settings used for the reported ensembles.**

| Setting | Value |
|---|---|
| Gradient steps | 60 |
| Learning rate | 0.05 |
| Clash threshold $d_{\mathrm{cl}}$ | 3.6 Å |
| Minimum sequence separation $s_{\min}$ | 2 |
| Candidate refresh interval | 10 |
| Candidate margin | 1.0 Å |

### C.3 Hyperparameters of Geometric Relaxation

Starting from each raw sample $\widetilde{\boldsymbol{X}}^{(k)}$, we minimize the energy in Eq. 11 with respect to its $C_\alpha$ coordinates. The target distance for each consecutive pair is measured from the input reference structure. We periodically update the candidate clash pairs; pairs outside the clash threshold contribute zero through the hinge penalty. The relaxation settings are summarized in Table A3 and are applied uniformly to all covariance-based methods.

### C.4 Effect of Geometric Relaxation

We first separate the effects of the chain and clash terms using matched draws. As shown in Table A4, the chain term substantially reduces broken consecutive $C_\alpha$ pairs, while adding the clash term reduces the median clash rate to zero. The ensemble-level motion measures change only slightly between the two relaxed settings.

We vary the number of relaxation steps while keeping the initial draws fixed (Table A5). The median fraction of conformations with a clash falls from $32.4\%$ to zero by step 10, whereas chain continuity improves more gradually. The median fraction with at least one $C_\alpha$ break decreases from $100.0\%$ before relaxation to $26.6\%$ at step 20 and reaches zero at step 60. We therefore use 60 steps, although some individual targets retain breaks.

The same relaxation improves agreement with MD in the correlation-based ensemble measures. Per-target RMSF correlation rises from $0.8573$ for raw samples to $0.8730$ at steps 30–40 and remains at $0.8726$ after 60 steps. Global DCCM correlation increases from $0.8424$ to $0.8448$. Thus, the chosen setting improves local geometry while retaining the correlation gains relative to raw samples.

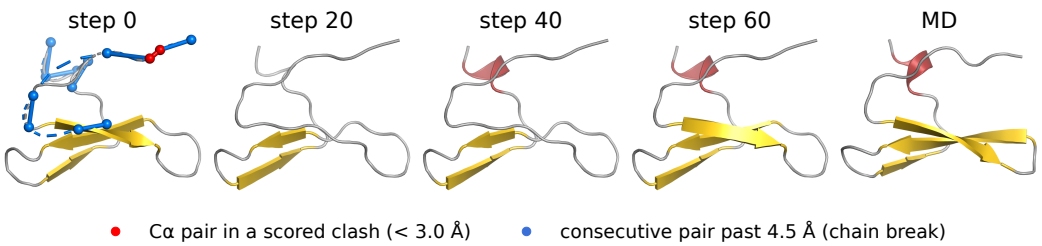


Figure A1: **Example of geometric relaxation across steps.** A sampled conformation of test protein 7lp1_A is shown before relaxation and after 20, 40, and 60 steps, alongside an MD conformation. Blue markers indicate consecutive $C_\alpha$ pairs farther than 4.5 Å; red markers indicate scored $C_\alpha$ clashes below 3.0 Å.

Table A4: **Effect of the relaxation terms under matched draws.** Each row starts from the same sampled conformations. Clash is the median fraction of conformations with a scored clash, whereas broken $C_\alpha$ pairs is the median fraction of consecutive residue pairs classified as breaks. Values are aggregated over 82 test proteins.

| Relaxation | Clash (%) ↓ | Broken $C_\alpha$ pairs (%) ↓ | Pairwise RMSD (Å) | Global DCCM $r$ ↑ | Per-target RMSF $r$ ↑ |
|---|---|---|---|---|---|
| None | 32.4 | 16.39 | 2.51 | 0.842 | 0.857 |
| Chain, 60 steps | 4.0 | 0.00 | 2.38 | 0.845 | 0.873 |
| Chain + clash, 60 steps | 0.0 | 0.00 | 2.38 | 0.845 | 0.873 |

Table A5: **Effect of relaxation steps under matched draws.** Step 0 denotes raw samples. Clash and break conformation rates, broken-pair rates, and per-target RMSF correlations are medians over 82 test proteins. Bold indicates the best reported value.

| Steps | 0 | 10 | 20 | 30 | 40 | 50 | 60 |
|---|---|---|---|---|---|---|---|
| Conformations with clash (%) ↓ | 32.4 | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** |
| Conformations with $C_\alpha$ break (%) ↓ | 100.0 | 93.0 | 26.6 | 4.4 | 0.8 | 0.4 | **0.0** |
| Broken consecutive $C_\alpha$ pairs (%) ↓ | 16.387 | 1.647 | 0.330 | 0.047 | 0.006 | 0.001 | **0.000** |
| Per-target RMSF $r$ ↑ | 0.8573 | 0.8689 | 0.8728 | **0.8730** | **0.8730** | 0.8728 | 0.8726 |
| Global DCCM $r$ ↑ | 0.8424 | 0.8443 | 0.8446 | 0.8447 | 0.8447 | **0.8448** | **0.8448** |

Figure A1 illustrates the coordinate-level changes for one test protein. The raw sample contains visible chain breaks; successive relaxation steps restore a more continuous backbone while retaining a conformation distinct from the MD reference.

# D Baselines, Evaluation, and Runtime Protocol

## D.1 Classical ENM Baselines

We evaluate GNM, ANM, ANM*, pfANM, and Hinsen using a common covariance assembly and pseudoinverse implementation. Each baseline is specified by an interaction graph over $C_\alpha$ atoms and a spring rule $\gamma_{ij}$ for each interacting pair. For cutoff-based models, sequence-adjacent residues are always connected irrespective of their spatial distance. Table A6 summarizes the configurations used in the reported experiments.

**Configuration and scale calibration.** For GNM and ANM, we use interaction cutoffs of 7.3 Å and 13 Å, respectively. pfANM and Hinsen use all-pairs interaction graphs according to their prescribed interaction rules. ANM* uses a 15 Å graph with a two-scale interaction rule; its relative nonlocal stiffness is selected on the training split and then fixed to $\eta^* = 0.1$ for evaluation.

## D.2 Learning-based Baselines

For the learning-based covariance baselines, we use the released pretrained models without additional training, fine-tuning, or post-hoc covariance-scale calibration.

**DynaProt.** For DynaProt (Bafna et al., 2026), we use the official implementation and pretrained checkpoint released by the authors. DynaProt predicts a $3 \times 3$ marginal covariance for each residue together with an $N \times N$ residue–residue coupling matrix. We use the joint covariance constructed by the official implementation directly for matrix-level evaluation. For ensemble-level evaluation, this covariance is passed to the common sampling and post-sampling processing pipeline.

**BackFlip-2.** For BackFlip-2 (Viliuga et al., 2026), we follow the publicly released implementation and use the `backflip-2.1` checkpoint trained on ATLAS. We construct the joint covariance from the predicted per-residue covariance blocks and pairwise DCCM, converting covariance units from $nm^2$ to $Å^2$ before evaluation. The resulting covariance is evaluated directly at the matrix level and is used as input to the same ensemble-generation pipeline described below.

Table A6: **Classical ENM baseline configurations.** Interaction rules, graph definitions, and fitted configurations used in the reported experiments.

| Method | Interaction rule | Graph | Configuration |
|---|---|---|---|
| GNM | $\gamma_{ij} = c$ | 7.3 Å | $c^* = 0.201$ |
| ANM | $\gamma_{ij} = c$ | 13 Å | $c^* = 0.291$ |
| | $\gamma_{ij} = c$ for $\lvert i-j\rvert \le 2$; | | |
| ANM* | $\gamma_{ij} = c\eta$ otherwise | 15 Å | $\eta^* = 0.1,\ c^* = 0.101$ |
| pfANM | $\gamma_{ij} = c\, r_{ij}^{-2}$ | All pairs | $c^* = 7.4 \times 10^{-3}$ |
| | $\gamma_{ij} = c(A\tilde{r} - B)$ for $\tilde{r} < 0.4$; | | $A = 8.6 \times 10^5,\ B = 2.39 \times 10^5,$ |
| Hinsen | $\gamma_{ij} = c\, k_0 \tilde{r}^{-6}$ otherwise, $\quad \tilde{r} = r_{ij}/10$ (nm) | All pairs | $k_0 = 128,\ c^* = 5.1 \times 10^{-4}$ |

### D.3 Ensemble Generation and Reconstruction

For every covariance-based method included in the ensemble comparison, we draw $K = 250$ $C_\alpha$ conformations per protein from the predicted covariance. The same sampling and post-sampling geometric-relaxation procedure is used for the classical ENM baselines, DynaProt, BackFlip-2, and ENMulator.

Gaussian covariance sampling reproduces the predicted second-order displacement statistics but does not enforce local geometric constraints. We therefore apply the same deterministic relaxation to every sampled conformation. The relaxation penalizes deviations of consecutive $C_\alpha$ distances from their values in the reference structure and steric clashes between sequence-distant residues. It is applied only after sampling and does not modify the predicted covariance itself.

We use 60 relaxation steps with learning rate $0.05$ and clash threshold $d_{\mathrm{cl}} = 3.6$ Å. The remaining relaxation settings are summarized in Table A3. After relaxation, each $C_\alpha$ conformation is reconstructed into a full-atom structure using PULCHRA v3.06 (Rotkiewicz & Skolnick, 2008). All reported full-atom ensemble metrics are computed from these reconstructed structures.

### D.4 Runtime and Computational Cost

**Measurement protocol.** We measured all methods on a commercial workstation with an NVIDIA GeForce RTX 4090 GPU (24 GB), an Intel Core i9-14900KF CPU, and 62 GB of host memory, using PyTorch 2.5.1, CUDA 12.1, and eight CPU threads. PULCHRA v3.06 runs single-threaded. For each of the 82 ATLAS test proteins, we performed one untimed warm-up followed by five timed runs and report the per-protein median. Summary values across proteins are medians of these per-protein measurements. Timings are wall-clock times and exclude model initialization, checkpoint loading, and initial input parsing; all graph construction, covariance assembly, sampling, relaxation, backbone reconstruction, and output operations listed in Table A7 are included in their corresponding measurements.

For GPU measurements, we disabled gradient tracking and explicitly synchronized the device immediately before and after each timed region so that asynchronous CUDA execution was included in the reported wall-clock time. Backbone reconstruction and output were timed separately on each method's saved relaxed conformations, with three timed runs per protein (method order rotated between runs) after a single warm-up. End-to-end totals are the per-protein sum of the two medians.

**Covariance-prediction latency.** Figure A2 reports the time required to map a reference structure to the complete covariance matrix. ENMulator requires 7.6 ms per protein at the median and 62 ms for the largest test protein ($N = 724$). The fitted constant term is approximately 3.5 ms and accounts for most of the runtime below $N \approx 150$. At larger $N$, the float64 factorization and inverse contribute an increasing fraction of the runtime. Consistent with this transition, the effective local exponent $d \log t / d \log N$ of the fitted curve increases from 0.42 at $N = 100$ to 2.14 at $N = 700$, approaching the cubic regime of the dense linear-algebra operations.

ENMulator and GPU ANM solve the same Hessian system. Consequently, the combined Cholesky-factorization and inverse time is approximately 4.0 ms for each method at the median. ENMulator additionally spends 3.3 ms in the IPA backbone, prediction heads, and marginal congruence, making

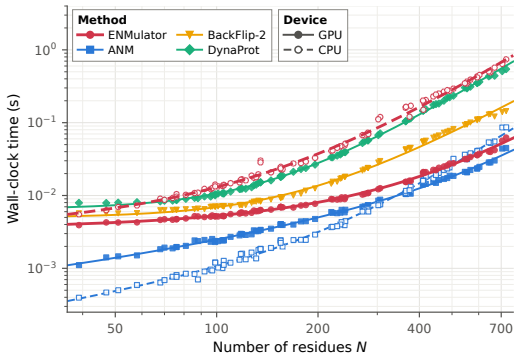


Figure A2: **Covariance-prediction time versus protein length.** Each marker represents one of the 82 ATLAS test proteins and shows the median wall-clock time over five timed runs after one untimed warm-up.

Table A7: **Stepwise runtime of ENMulator ensemble generation.** Wall-clock time (*ms*) per protein for 250 conformations. The "All" and length-bin columns report medians; "Share" is computed from times summed over all 82 proteins. Independently computed row medians need not sum exactly to the displayed subtotals or total.

| Step | All $n = 82$ | $N \leq 150$ $n = 33$ | $150 < N \leq 300$ $n = 20$ | $N > 300$ $n = 29$ | Share (%) |
|---|---|---|---|---|---|
| *Structure → covariance (GPU)* | | | | | |
| Contact graph and rigid-body basis | 0.65 | 0.47 | 0.71 | 0.93 | $< 0.01$ |
| IPA backbone (8 blocks) | 2.5 | 2.4 | 2.5 | 6.5 | 0.04 |
| Spring and marginal heads | 0.23 | 0.22 | 0.24 | 0.25 | $< 0.01$ |
| Hessian assembly | 0.09 | 0.08 | 0.09 | 0.11 | $< 0.01$ |
| Cholesky factorization | 2.0 | 0.97 | 2.3 | 5.0 | 0.02 |
| Inverse → $H^{+}$ | 2.1 | 0.98 | 2.7 | 11 | 0.05 |
| Marginal congruence | 0.66 | 0.43 | 0.76 | 2.2 | 0.01 |
| *Subtotal (one forward call)* | *7.6* | *5.2* | *8.8* | *26* | *0.1* |
| *Sampling (CPU)* | | | | | |
| Transfer of $\hat{\Sigma}$ to host | 0.24 | 0.10 | 0.31 | 1.2 | $< 0.01$ |
| 250 Gaussian draws | 14 | 3.8 | 19 | 124 | 0.5 |
| Relaxation (60 steps) | 98 | 56 | 112 | 430 | 2.1 |
| *Subtotal* | *111* | *61* | *133* | *554* | *2.6* |
| *Backbone reconstruction and output (CPU)* | | | | | |
| Write 250 $C_\alpha$ input files | 164 | 136 | 166 | 211 | 1.5 |
| PULCHRA (250 calls) | 6,708 | 4,563 | 7,011 | 17,190 | 85.1 |
| Parse PULCHRA output | 504 | 394 | 516 | 733 | 4.7 |
| Write multi-model PDB | 498 | 263 | 567 | 1,216 | 5.9 |
| *Subtotal* | *7,784* | *5,370* | *8,316* | *19,814* | *97.3* |
| **Total ensemble generation** | **7,892** | **5,430** | **8,463** | **20,562** | **100.0** |

GPU ANM $1.64\times$ faster for covariance prediction. CPU ANM is faster for small proteins, whereas the nonnegative polynomial fits intersect at $N \approx 506$. Under the same hardware and timing protocol, ENMulator is faster than BackFlip-2 (12.1 ms) and DynaProt (25.2 ms).

**End-to-end ensemble generation.** Table A7 decomposes the wall-clock time required to generate and write 250 reconstructed backbone conformations. The median end-to-end runtime is 7.9 s per protein. When time is aggregated over all 82 test proteins, covariance prediction accounts for 0.12% of the total, sampling and relaxation for 2.6%, and backbone reconstruction and output for 97.3%. The 250 PULCHRA calls alone account for 85.1% of the aggregate runtime.

### D.5 Uncertainty estimation

We report uncertainty as the bootstrap standard error over test proteins. Each bootstrap replicate resamples the 82 test targets with replacement and recomputes every reported statistic exactly as in the table: the median over targets for per-target metrics, the cross-target error, fraction or correlation where the metric is defined that way, and pooling over the residues of the drawn targets for global metrics. For pooled metrics, a protein drawn more than once contributes all of its residues each time, so residues within a protein are not treated as independent samples. The same $B = 10{,}000$ resampled target sets are shared by all methods and all tables. Generated ensembles and the MD reference are held fixed, so the reported error reflects only the choice of test proteins. Standard errors following this protocol are given in Table A14, Table A15, and Table A16; parameter counts and runtimes are not resampled.

## E Covariance and Ensemble Evaluation Metrics

### E.1 Evaluation Objects and Aggregation

We evaluate predicted dynamics at two complementary levels. *Covariance evaluation* compares the predicted $C_\alpha$ displacement covariance $\widehat{\mathbf{\Sigma}} \in \mathbb{R}^{3N\times 3N}$ directly with the MD reference $\mathbf{\Sigma}^\star$, and thus assesses the learned second-order fluctuation model before any sampling. *Ensemble evaluation* compares coordinates after drawing $K = 250$ conformations and, for covariance-based methods, applying geometric relaxation. This distinction separates errors in the predicted covariance itself from changes introduced by finite sampling and post-processing. All coordinates are aligned to the common reference $C_\alpha$ structure. Unless stated otherwise, target-level quantities are summarized by their median over the 82 test proteins; global RMSF and DCCM correlations are instead computed after concatenating entries across targets, and therefore additionally reflect calibration across proteins of different sizes.

### E.2 Covariance-Level Metrics

**Global covariance accuracy.** We compare the complete predicted and MD covariance matrices using relative Frobenius error, full-matrix Pearson correlation, and normalized Bures–Wasserstein distance:

$$e_{\mathrm{rel}} = \frac{\|\widehat{\mathbf{\Sigma}} - \mathbf{\Sigma}^\star\|_F}{\|\mathbf{\Sigma}^\star\|_F}, \qquad r_{\mathrm{cov}} = \mathrm{Pearson}\Big(\{\widehat{\Sigma}_{ab}\}_{a,b=1}^{3N}, \{\Sigma^\star_{ab}\}_{a,b=1}^{3N}\Big), \tag{A11}$$

$$\frac{d_{\mathrm{BW}}(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}^\star)}{\sqrt{\mathrm{tr}(\mathbf{\Sigma}^\star)}}, \qquad d^2_{\mathrm{BW}}(\widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}^\star) = \mathrm{tr}\,\widehat{\mathbf{\Sigma}} + \mathrm{tr}\,\mathbf{\Sigma}^\star - 2\,\mathrm{tr}\left[\left((\mathbf{\Sigma}^\star)^{1/2}\widehat{\mathbf{\Sigma}}(\mathbf{\Sigma}^\star)^{1/2}\right)^{1/2}\right]. \tag{A12}$$

Relative Frobenius error is an entrywise, scale-sensitive comparison, while Bures–Wasserstein distance measures the geometry of the corresponding zero-mean Gaussian distributions. For the scale diagnostics in Table A14, we also use the optimal global rescaling $c_{\mathrm{F}} = \langle \widehat{\mathbf{\Sigma}}, \mathbf{\Sigma}^\star\rangle_F / \|\widehat{\mathbf{\Sigma}}\|_F^2$ and the trace ratio $r_{\mathrm{tr}} = \mathrm{tr}(\widehat{\mathbf{\Sigma}})/\,\mathrm{tr}(\mathbf{\Sigma}^\star)$.

**Collective modes and spectrum.** Top-$k$ subspace overlap compares the dominant collective-motion directions of the prediction and MD (Knyazev & Argentati, 2002). Given orthonormal bases $\boldsymbol{U}_k$ and $\boldsymbol{V}_k$ of their leading eigenspaces, respectively, the overlap is the mean singular value of $\boldsymbol{U}_k^\top \boldsymbol{V}_k$. We also report the participation ratio $\mathrm{PR}(\mathbf{\Sigma}) = (\sum_j \max(\lambda_j, 0))^2 / \sum_j \max(\lambda_j, 0)^2$, which summarizes the effective number of modes carrying variance; values closer to the MD target are preferred. Table A14 further resolves scale-adjusted eigenvalue error and relative covariance error by sequence separation.

**Residue flexibility and correlated motion.** For a covariance $\mathbf{\Sigma}$, residue RMSF and DCCM are defined by

$$f_i = \sqrt{\mathrm{tr}\,\mathbf{\Sigma}_{ii}}, \qquad C_{ij} = \frac{\mathrm{tr}\,\mathbf{\Sigma}_{ij}}{\sqrt{\mathrm{tr}\,\mathbf{\Sigma}_{ii}\,\mathrm{tr}\,\mathbf{\Sigma}_{jj}}}. \tag{A13}$$

For the reference-centered MD target, $f_i$ is the root mean squared displacement from the reference and includes any mean shift. Likewise, $C_{ij}$ is a normalized displacement second moment, not

necessarily a mean-centered correlation. These descriptors test local displacement magnitude and normalized cross-residue coupling under the stated reference convention. We report Pearson correlation and MAE for both descriptors. Covariance-level DCCM scores use only off-diagonal entries because the unit diagonal contains no information about inter-residue dependence.

### E.3 Ensemble-Level Metrics

**Ensemble spread and residue flexibility.** For each covariance-based method, we sample and geometrically relax 250 conformations using the common protocol in Appendix D.3; AlphaFlow (Jing et al., 2024a) and BBFlow (Wolf et al., 2025) generate coordinates directly and do not receive this post-processing. Ensemble spread is measured by mean pairwise RMSD,

$$D_{\text{pair}}(\mathcal{E}) = \frac{1}{K^2}\sum_{k=1}^{K}\sum_{l=1}^{K}\text{RMSD}(\boldsymbol{X}^{(k)}, \boldsymbol{X}^{(l)}), \tag{A14}$$

and by its Pearson correlation with the corresponding MD quantity across targets. Pairwise RMSD measures ensemble diversity. We compute RMSF from the sampled coordinate covariance and report absolute value, MAE, and target-level and global correlation where applicable. Target-level RMSF correlation measures whether a method recovers the within-protein flexibility profile, whereas global correlation additionally tests whether fluctuation amplitudes are calibrated across proteins.

**Correlated motion and distributional accuracy.** Sampled DCCM is computed from the coordinate covariance of each generated ensemble. Per-target DCCM correlation is evaluated within each protein and summarized by its median, while global DCCM correlation is computed after concatenating targets. Following the AlphaFlow evaluation protocol (Jing et al., 2024a), these ensemble DCCM correlations include the unit diagonal; this convention is intentionally distinct from the off-diagonal-only covariance analysis above. We additionally report root mean Wasserstein distance (RMWD), which compares generated and MD coordinate distributions residue by residue. If $\widehat{\boldsymbol{\mu}}_i, \widehat{\boldsymbol{S}}_i$ and $\boldsymbol{\mu}_i^\star, \boldsymbol{S}_i^\star$ denote the sample means and covariances of the aligned generated and MD ensembles, respectively, then

$$d_{\text{trans}}^2 = \frac{1}{N}\sum_i \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i^\star\|_2^2, \qquad d_{\text{var}}^2 = \frac{1}{N}\sum_i d_{\text{BW}}^2(\widehat{\boldsymbol{S}}_i, \boldsymbol{S}_i^\star), \qquad \text{RMWD} = \sqrt{d_{\text{trans}}^2 + d_{\text{var}}^2}. \tag{A15}$$

The variance contribution evaluates local anisotropic fluctuations, but because RMWD discards off-diagonal covariance blocks, it is not evidence for recovery of correlated motion.

**Low-dimensional structure and structural validity.** We report MD-PCA and joint-PCA $\mathcal{W}_2$ distances to compare the generated and MD ensembles in a two-dimensional principal-component representation. We additionally report the fraction of targets whose leading-PC similarity, defined as the absolute cosine between the leading generated and MD loading vectors, exceeds $0.5$. Weak and transient contact Jaccard scores, as well as the target-aware clash criterion, follow the AlphaFlow evaluation protocol. Contacts use an $8$ Å $C_\alpha$ cutoff; weak contacts are present in the reference structure but occur in fewer than $90\%$ of MD frames, whereas transient contacts are absent from the reference structure but occur in more than $10\%$ of MD frames. A chain break is an adjacent $C_\alpha$ separation larger than the maximum observed for that target in MD. For each target, we compute the fraction of generated conformations with at least one clash or chain break and summarize these fractions across targets. Runtime, relative runtime, and parameter count in Table 3 measure implementation efficiency; runtime is measured for sampling 250 conformations of `6kty_A`.

## F Directional Coupling Between Residue Pairs

Section 5.4 evaluates whether a model recovers directional dependence between residue displacements, beyond the same-direction motion measured by DCCM. For a residue pair $i \neq j$, we normalize its cross-covariance by the two marginal covariances,

$$\boldsymbol{K}_{ij} = \boldsymbol{\Sigma}_{ii}^{-1/2}\boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1/2}, \tag{A16}$$

and denote its singular values by $s_1(i,j) \geq s_2(i,j) \geq s_3(i,j)$. With mean-centered covariances, $s_1$ is the largest Pearson correlation attainable between projections of the two residue displacements. With the reference-centered second moments used here, it is instead the largest normalized cross second moment of such projections; the two interpretations agree when the mean displacement is zero. We retain the canonical-correlation terminology for the associated Gaussian covariance model. The score $1 - s_3/s_1$, termed direction preference, measures anisotropy between the strongest and weakest coupling directions. Because it does not use $s_2$, it does not distinguish one strong direction from two.

This quantity complements DCCM. DCCM averages Cartesian co-movement through the trace of the cross-covariance block, and is consequently most responsive when two residues move in similar directions. In contrast, canonical correlations allow the coupled directions to differ between residues. For example, a residue that moves upward whenever another moves rightward can have strong directional dependence despite near-zero DCCM. More generally, $|\mathrm{DCCM}_{ij}| \leq s_1(i,j)$, so a small DCCM does not imply weak residue-pair coupling.

For each protein, we compute four quantities over all residue pairs $i < j$: the median coupling strength $s_1$; the direction-preference score $1 - \mathrm{median}(s_3/s_1)$; the Pearson correlation between predicted and MD values of $s_1$ across pairs; and the coupling error

$$e_{\mathrm{cc}} = \frac{1}{3|\mathcal{P}|} \sum_{(i,j)\in\mathcal{P}} \sum_{k=1}^{3} |s_k(i,j) - s_k^\star(i,j)|\,, \tag{A17}$$

where $\mathcal{P}$ is the set of residue pairs and $s_k^\star$ denotes the MD value. Table 4 reports the median of each target-level quantity across the 82 test proteins.

The lower bound applies to reconstructions whose whitened cross-covariance blocks have three equal singular values. A scalar-coupling construction with this property is

$$\boldsymbol{\Sigma}_{ij} = \rho_{ij} \boldsymbol{L}_i \boldsymbol{L}_j^\top, \qquad \boldsymbol{\Sigma}_{ii} = \boldsymbol{L}_i \boldsymbol{L}_i^\top, \qquad \rho_{ii} = 1,$$

where the $\boldsymbol{L}_i$ are invertible. Then $\boldsymbol{\Sigma}_{ii}^{-1/2} \boldsymbol{L}_i$ is orthogonal, so $s_1 = s_2 = s_3 = |\rho_{ij}|$. This is a restriction of the reconstruction, not a consequence of using one scalar parameter per pair in every possible model. For a given MD pair, minimizing over a shared singular value gives

$$\min_{q\geq 0} \frac{1}{3} \sum_{k=1}^{3} |q - s_k^\star| = \frac{s_1^\star - s_3^\star}{3},$$

with a minimum at $q = s_2^\star$. Averaging this bound over pairs within each target and then taking the median over targets yields $0.182$. This pairwise relaxation does not require that the pairwise optima form a jointly valid covariance, and thus is a lower bound rather than an attainable joint optimum. Table 4 compares the reported scalar reconstructions with this restricted family; the bound does not apply to arbitrary scalar-parameterized interaction models.

For ENMulator, the blockwise congruence changes residue marginals but preserves the canonical correlations of the elastic covariance (Appendix B.3). The analysis in Table 4 therefore isolates the directional coupling induced by the learned elastic interactions. This does not imply that DCCM is invariant under congruence: DCCM and canonical correlation summarize different aspects of the same cross-covariance block.

Table A8 resolves coupling error by sequence separation. ENMulator's error ranges from $0.093$ to $0.125$ across the reported bands, whereas uniform ANM springs have their largest errors for nearby pairs. These absolute errors alone do not distinguish over-coupling from under-coupling.

# G Ablation Studies and Additional Analyses

## G.1 Paired Per-Target Comparison

Table 2 reports medians over the 82 test proteins. To test whether the differences are consistent across proteins, Table A9 compares ENMulator with BackFlip-2, DynaProt, and ANM target by target. For each metric we report the fraction of targets on which ENMulator is better and the median

Table A8: **Coupling error by sequence separation.** Values are medians of $e_{cc}$ over 82 test proteins.

| Method | 1 | 2–4 | 5–12 | 13–32 | 33–96 | $\geq 97$ |
|---|---|---|---|---|---|---|
| DynaProt | 0.168 | 0.269 | 0.321 | 0.324 | 0.331 | 0.327 |
| BackFlip-2 | 0.122 | 0.204 | 0.279 | 0.278 | 0.282 | 0.276 |
| ANM | 0.637 | 0.487 | 0.317 | 0.264 | 0.255 | 0.240 |
| **ENMulator** | 0.093 | 0.122 | 0.125 | 0.121 | 0.122 | 0.119 |

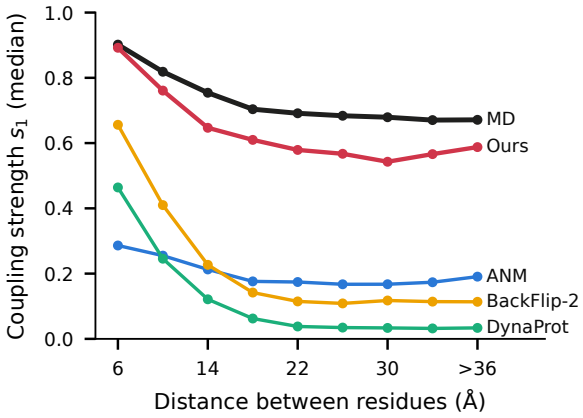


Figure A3: **Directional coupling strength across spatial distance.** Median $s_1$ as a function of binned $C_\alpha$ distance for residue pairs with $|i - j| \geq 3$, aggregated over the 82 test proteins.

paired improvement with a 95% bootstrap confidence interval (10,000 resamples of targets). Positive improvements favor ENMulator for every metric. Against BackFlip-2, ENMulator improves Bures–Wasserstein distance, top-10 subspace overlap, and DCCM metrics on 72–78% of targets; the corresponding win rate for top-3 overlap is 63%. The median paired improvements in these metrics have confidence intervals excluding zero. The relative Frobenius error and the RMSF metrics, in contrast, do not differ consistently between the two methods. Against DynaProt and ANM, ENMulator improves nearly every metric on most targets. Per-target values for Hinsen were not retained, so it is omitted from this analysis.

### G.2 Training-Time $\alpha$ Trade-off

Each row in Table A10 is a separately trained model with the displayed $\alpha$ fixed throughout training and inference. All five models share the same architecture and training protocol and differ only in $\alpha$. Covariance metrics are computed on the same 82 test targets, and ensemble metrics use the same 250 random draws per target and the same 60-step relaxation. Our default $\alpha = 0.75$, selected on the validation split, gives the best trace ratio and global ensemble correlations, whereas $\alpha = 0.50$ gives the lowest relative Frobenius error and the highest per-target RMSF correlation. Covariance metrics and per-target RMSF correlations are medians over the 82 test targets; global correlations pool entries across targets, and pairwise-RMSD correlation is computed across targets.

### G.3 Complementary Roles of Elastic Interactions and Marginal Calibration

We examine whether the two branches make complementary contributions that match their intended roles. Table A11 shows that learned protein-specific elastic interactions are the primary driver of correlated-motion accuracy: they raise DCCM correlation from 0.779 to 0.851 without the marginal branch. The marginal branch mainly improves residue-wise flexibility, raising RMSF correlation from 0.764 to 0.842 on a uniform elastic network while leaving DCCM nearly unchanged. The spring branch establishes the elastic covariance's canonical-correlation spectra, and congruence calibrates local anisotropic marginals while preserving those spectra; the two together give the lowest global covariance distances. Appendix G.4 further decomposes the spring parameterization.

Table A9: **Paired per-target comparison of covariance metrics.** For each baseline, "Win" is the percentage of the 82 test targets on which ENMulator is better, and $\Delta$ is the median paired improvement of ENMulator over the baseline (positive favors ENMulator) with a 95% bootstrap confidence interval.

| Metric | *vs. BackFlip-2* Win (%) | *vs. BackFlip-2* $\Delta$ [95% CI] | *vs. DynaProt* Win (%) | *vs. DynaProt* $\Delta$ [95% CI] | *vs. ANM* Win (%) | *vs. ANM* $\Delta$ [95% CI] |
|---|---|---|---|---|---|---|
| Relative Frobenius (↓) | 57 | 0.022 [−0.009, 0.040] | 78 | 0.052 [0.039, 0.080] | 85 | 0.092 [0.070, 0.127] |
| Bures–Wasserstein $\mathcal{W}_2$ norm. (↓) | 78 | 0.055 [0.039, 0.073] | 88 | 0.122 [0.090, 0.160] | 93 | 0.160 [0.150, 0.187] |
| Top-3 subspace overlap (↑) | 63 | 0.063 [0.018, 0.140] | 82 | 0.111 [0.070, 0.204] | 74 | 0.079 [0.035, 0.111] |
| Top-10 subspace overlap (↑) | 72 | 0.056 [0.038, 0.074] | 87 | 0.113 [0.086, 0.140] | 93 | 0.077 [0.056, 0.110] |
| RMSF $r$ (↑) | 54 | 0.002 [−0.008, 0.013] | 59 | 0.016 [−0.002, 0.033] | 87 | 0.087 [0.056, 0.111] |
| RMSF MAE (↓) | 54 | 0.008 [−0.030, 0.052] | 76 | 0.104 [0.056, 0.138] | 94 | 0.560 [0.453, 0.669] |
| DCCM $r$ (↑) | 78 | 0.051 [0.036, 0.071] | 98 | 0.227 [0.196, 0.267] | 87 | 0.060 [0.042, 0.075] |
| DCCM MAE (↓) | 76 | 0.022 [0.013, 0.031] | 96 | 0.074 [0.060, 0.097] | 94 | 0.053 [0.044, 0.068] |

Table A10: **Training-time $\alpha$ trade-off.** Each row is a separately trained model with $\alpha$ fixed throughout training. Bold marks the best value in each column.

| $\alpha$ | *Covariance* Rel. Fro. (↓) | *Covariance* Trace ratio ($\to 1$) | *Ensemble* (all ↑) Global DCCM $r$ | Global RMSF $r$ | Per-target RMSF $r$ | Pairwise RMSD $r$ |
|---|---|---|---|---|---|---|
| 0.25 | 0.880 | 0.660 | 0.834 | 0.687 | 0.864 | 0.633 |
| 0.50 | **0.868** | 0.660 | 0.842 | 0.745 | **0.876** | 0.699 |
| **0.75** | 0.874 | **0.761** | **0.845** | **0.825** | 0.873 | **0.832** |
| 0.90 | 0.890 | 0.750 | 0.835 | 0.791 | 0.856 | 0.795 |
| 1.00 | 0.877 | 0.596 | 0.836 | 0.778 | 0.865 | 0.762 |

### G.4 Decomposing the Learned Spring Parameterization

Table A12 decomposes the learned elastic interaction into the shared distance kernel $\kappa(r)$ and the residue-wise term $a_i$. The distance kernel provides the largest individual improvement over uniform stiffnesses, whereas the residue term is less effective in isolation but improves performance when combined with the kernel. Thus, $\kappa(r)$ captures how interaction strength varies with distance across proteins, while $a_i$ adapts this pattern to protein- and residue-specific environments. Their combination gives the most accurate covariance, correlated motion, and residue-flexibility predictions. The learned distance kernel alone already reaches the DCCM correlation of Hinsen ($0.824$), a hand-designed distance-dependent ENM, and the full model exceeds it.

### G.5 Training objective ablation

We vary the loss terms of Eq. 8, compression $\rho$, and stage-1 warm-up as listed in Table A13. Other settings are fixed: $\alpha = 0.75$, full congruence, a 120-epoch budget, and the same ATLAS split. All checkpoints are re-scored together.

Training with $\mathcal{L}_{\mathrm{cov}}$ alone (a) underestimates fluctuation scale (trace ratio $0.454$). Adding $\mathcal{L}_{\mathrm{scale}}$ (a→b) improves the trace ratio to $0.809$, while worsening BW distance ($0.709 \to 0.813$). Marginal-head supervision (b→c) improves both BW distance and RMSF correlation. Adding $\mathcal{L}_{\mathrm{diag}}$ (c→d) further improves BW distance ($0.711 \to 0.663$), but worsens trace calibration and leaves RMSF correlation similar.

Warm-up (d vs. ENMulator) improves RMSF correlation and top-10 overlap, but slightly worsens BW distance. These single-run comparisons show trade-offs rather than uniform superiority of the adopted objective.

### G.6 Effect of Per-Target Global Rescaling

Table A14 examines whether the covariance errors of existing methods can be explained by a single global amplitude mismatch. For each target $p$, we grant every method the oracle rescaling $c_p^\star = \arg\min_c \|c\widehat{\mathbf{\Sigma}}_p - \mathbf{\Sigma}_p^\star\|_F$. This diagnostic minimizes Frobenius error using the unavailable MD target and cannot be selected at deployment; the fitted scale is not generally optimal for BW distance or RMSF MAE. Even under this oracle correction, the best classical ENM, Hinsen, reaches

Table A11: **Ablation of learned springs and marginal congruence.** We compare uniform ANM stiffnesses ($\gamma_{ij} \equiv 1$) with learned protein-specific stiffnesses, with and without congruence ($\alpha = 0.75$ or 0). All metrics compare the predicted covariance with the MD covariance. Congruence acts on the residue marginals: it raises RMSF correlation by 0.078 over ANM while leaving DCCM correlation at the ANM level. The learned stiffnesses recover the correlated motion, and both branches together give the lowest global covariance distances. Medians over 82 test targets.

| Variant | Spring branch | Marginal branch | *Distance to MD covariance* | | *Correlation with MD* | |
|---|---|---|---|---|---|---|
| | | | Rel. Fro. (↓) | BW $\mathcal{W}_2$ (↓) | DCCM $r$ (↑) | RMSF $r$ (↑) |
| ANM (neither) | ✗ | ✗ | 0.984 | 0.839 | 0.779 | 0.764 |
| Marginal only | ✗ | ✓ | 0.932 | 0.789 | 0.778 | 0.842 |
| Spring only | ✓ | ✗ | 0.914 | 0.690 | 0.851 | 0.819 |
| **ENMulator (Ours)** | ✓ | ✓ | 0.874 | 0.666 | 0.856 | 0.859 |

Table A12: **Ablation of the learned spring parameterization.** All ablation rows use the same structure encoder and congruence setting while varying the learnable terms in $\gamma_{ij} = \exp[\beta + \kappa(r_{ij}) + a_i + a_j]$, where $\kappa(r_{ij})$ is distance-dependent and $a_i$ is residue-specific. The last two rows are classical ENMs, shown for reference.

| Spring parameterization | $\kappa(r)$ | $a_i$ | Rel. Fro. (↓) | DCCM $r$ (↑) | RMSF $r$ (↑) |
|---|---|---|---|---|---|
| uniform, $\gamma \equiv 1$ | ✗ | ✗ | 0.932 | 0.778 | 0.842 |
| + distance kernel $\kappa(r)$ | ✓ | ✗ | 0.886 | 0.824 | 0.846 |
| + residue term $a_i$ | ✗ | ✓ | 0.904 | 0.795 | 0.841 |
| **ENMulator (Ours)** | ✓ | ✓ | 0.874 | 0.856 | 0.859 |
| *Reference: classical ENMs* | | | | | |
| ANM (classical) | ✗ | ✗ | 0.984 | 0.779 | 0.764 |
| Hinsen (classical) | ✗ | ✗ | 0.990 | 0.824 | 0.801 |

normalized BW distance 0.702, still above ENMulator's unscaled value of 0.666. Moreover, global rescaling cannot correct errors in the relative strength of different modes or residue-wise fluctuations; for some methods it improves full-matrix error while worsening RMSF MAE. ENMulator's median Frobenius-optimal scale factor ($c_p^\star = 0.947$) is close to one, whereas its median trace ratio (0.761) indicates underestimation of total reference-centered displacement variance. These diagnostics measure different aspects of scale agreement.

### G.7 Extended Covariance Comparison and Ensemble Comparison

We provide Table A15 as an extension of Table 2 with bootstrap standard errors. Table A16 expands Table 3 to include all evaluated baselines and supplementary ensemble metrics.

### G.8 Limitations

ENMulator inherits the local harmonic approximation underlying elastic network models. Its raw sampling distribution is Gaussian around a single reference structure; subsequent geometric relaxation can change this distribution and its covariance. The model may not capture multimodal conformational landscapes or strongly anharmonic transitions, for which methods that explicitly model non-Gaussian structural distributions may be better suited. Our experiments are restricted to $C_\alpha$-level dynamics and the ATLAS benchmark; performance for all-atom representations, alternative force fields, and more diverse dynamical regimes remains to be established. Our matrix-level target includes mean displacement from the reference, so its descriptors need not equal mean-centered MD fluctuation statistics. Comparisons with pretrained predictors are therefore specific to this evaluation target. We evaluate static displacement and ensemble statistics, not temporal kinetics.

Table A13: **Training objective ablation.** Values are medians over 82 ATLAS test proteins, evaluated on predicted covariances. Auxiliary loss weights are either adopted or zero; $\mathcal{L}_{\mathrm{cov}}$ is unchanged across configurations. Arrows indicate the preferred direction.

| | **ENMulator** | (a) | (b) | (c) | (d) |
|---|---|---|---|---|---|
| $\lambda_{\mathrm{s}}$ ($\mathcal{L}_{\mathrm{scale}}$) | 0.1 | 0 | 0.1 | 0.1 | 0.1 |
| $\lambda_{\mathrm{d}}$ ($\mathcal{L}_{\mathrm{diag}}$) | 0.5 | 0 | 0 | 0 | 0.5 |
| $\lambda_{\mathrm{m}}$ ($\mathcal{L}_{\mathrm{marg}}$) | 0.5 | 0 | 0 | 0.5 | 0.5 |
| Stage-1 warm-up | yes | no | no | no | no |
| Relative Frobenius (↓) | 0.874 | 0.897 | 0.900 | 0.887 | 0.884 |
| Scale-free rel. Frobenius (↓) | 0.819 | 0.858 | 0.860 | 0.829 | 0.825 |
| Bures–Wasserstein distance (↓) | 0.666 | 0.709 | 0.813 | 0.711 | 0.663 |
| Top-10 subspace overlap (↑) | 0.663 | 0.646 | 0.623 | 0.661 | 0.657 |
| RMSF $r$ (↑) | 0.859 | 0.797 | 0.754 | 0.846 | 0.844 |
| DCCM $r$ (↑) | 0.856 | 0.829 | 0.801 | 0.849 | 0.856 |

Table A14: **Effect of per-target scalar rescaling.** For each target $p$, $c_p^\star = \arg\min_c \|c\widehat{\boldsymbol{\Sigma}}_p - \boldsymbol{\Sigma}_p^\star\|_F$ is fitted against the MD reference. Metrics marked $(*)$ therefore receive one oracle scale parameter that is unavailable at deployment. The middle block compares raw and rescaled values for the same metrics; the final block contains diagnostics defined only after applying $c_p^\star$. Values are medians over the 82 test targets.

| Metric | *Hand-designed ENM* GNM | ANM | ANM* | pfANM | Hinsen | *Learning-based methods* DynaProt | BackFlip-2 | **ENMulator** |
|---|---|---|---|---|---|---|---|---|
| *The fitted factor* | | | | | | | | |
| Optimal global rescaling $c^*$ (= 1) | 0.883±0.127 | 9.086±1.411 | 1.304±0.216 | 2.640±0.550 | 0.820±0.686 | 1.814±0.166 | 1.230±0.069 | 0.947±0.072 |
| *What rescaling changes (same metric, raw above, $c^*$ below)* | | | | | | | | |
| Relative Frobenius, raw (↓) | 0.978±0.004 | 0.984±0.003 | 0.962±0.009 | 0.988±0.005 | 0.990±0.005 | 0.954±0.006 | 0.910±0.018 | 0.874±0.015 |
| Relative Frobenius* (↓) | 0.964±0.002 | 0.903±0.014 | 0.888±0.013 | 0.945±0.006 | 0.876±0.011 | 0.925±0.009 | 0.873±0.016 | 0.819±0.016 |
| Bures–Wasserstein, norm., raw (↓) | 0.928±0.016 | 0.839±0.009 | 0.856±0.007 | 0.978±0.016 | 0.867±0.019 | 0.825±0.013 | 0.752±0.012 | 0.666±0.019 |
| Bures–Wasserstein, norm.* (↓) | 0.838±0.003 | 0.802±0.008 | 0.767±0.009 | 0.881±0.008 | 0.702±0.010 | 0.788±0.010 | 0.716±0.008 | 0.624±0.016 |
| RMSF MAE, raw (↓) | 0.774±0.036 | 1.001±0.068 | 0.600±0.032 | 1.179±0.135 | 1.292±0.169 | 0.539±0.049 | 0.445±0.024 | 0.433±0.038 |
| RMSF MAE* (↓) | 0.661±0.060 | 0.794±0.057 | 0.781±0.059 | 1.392±0.096 | 0.572±0.051 | 0.526±0.044 | 0.410±0.036 | 0.380±0.039 |
| *Defined only at $c^*$* | | | | | | | | |
| Top-10 eigenvalue rel. error* (↓) | 0.694±0.013 | 0.582±0.024 | 0.551±0.018 | 0.735±0.022 | 0.508±0.023 | 0.625±0.023 | 0.522±0.018 | 0.440±0.015 |
| Rel. error, $\lvert i-j\rvert = 1$ (↓) | 0.889±0.009 | 0.893±0.014 | 0.857±0.020 | 0.919±0.005 | 0.827±0.014 | 0.760±0.021 | 0.692±0.023 | 0.757±0.021 |
| Rel. error, $\lvert i-j\rvert \in [2,4]$ (↓) | 0.886±0.013 | 0.907±0.021 | 0.879±0.018 | 0.969±0.007 | 0.857±0.016 | 0.797±0.013 | 0.765±0.023 | 0.777±0.034 |
| Rel. error, $\lvert i-j\rvert \in [5,12]$ (↓) | 0.969±0.007 | 0.921±0.018 | 0.921±0.018 | 0.964±0.007 | 0.899±0.020 | 0.937±0.008 | 0.895±0.011 | 0.858±0.025 |
| Rel. error, $\lvert i-j\rvert \in [13,32]$ (↓) | 1.009±0.004 | 0.865±0.023 | 0.867±0.021 | 0.946±0.011 | 0.855±0.016 | 0.986±0.005 | 0.951±0.009 | 0.839±0.016 |
| Rel. error, $\lvert i-j\rvert \in [33,96]$ (↓) | 1.001±0.002 | 0.889±0.019 | 0.890±0.018 | 0.941±0.012 | 0.876±0.014 | 0.987±0.003 | 0.947±0.006 | 0.836±0.019 |
| Rel. error, $\lvert i-j\rvert \geq 97$ (↓) | 1.001±0.002 | 0.915±0.024 | 0.909±0.024 | 0.953±0.010 | 0.891±0.017 | 0.994±0.002 | 0.969±0.007 | 0.854±0.028 |
| Rel. error, all off-diagonal (↓) | 0.972±0.002 | 0.899±0.015 | 0.889±0.016 | 0.951±0.007 | 0.880±0.012 | 0.941±0.006 | 0.891±0.011 | 0.822±0.017 |

### G.9 Additional Qualitative Results

We additionally provide the full comparison on all 82 ATLAS test targets for the predicted covariance (Figure A4), the DCCM (Figure A5), and the per-residue RMSF (Figure A6).

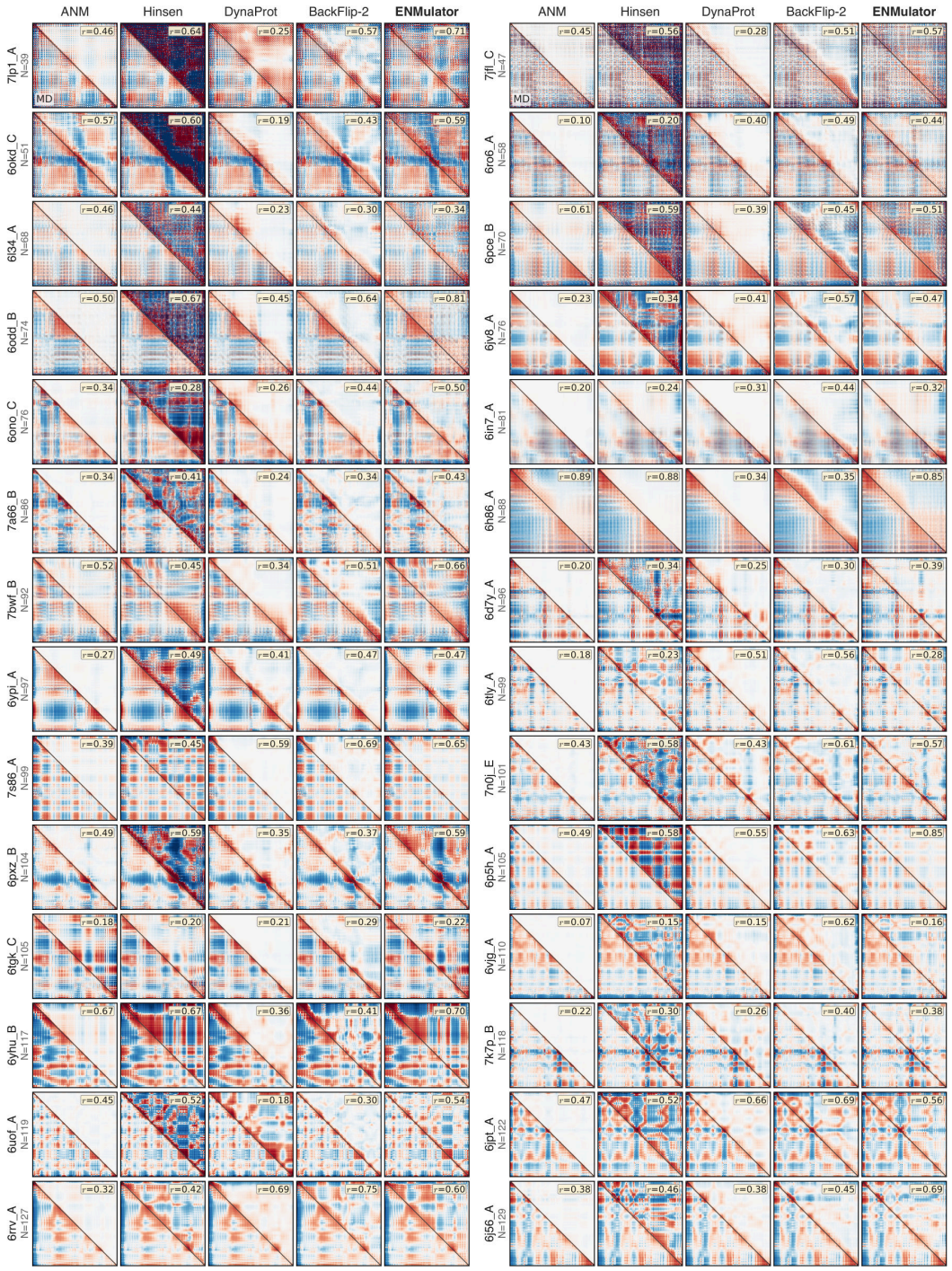


Figure A4: **Predicted covariance on all 82 ATLAS test targets.** Each block of five panels is one target, ordered by increasing $N$. Every panel is split, with MD below the diagonal and the column's method above it. Colors show the covariance on a symmetric-log scale, normalized separately for each target by the 99.5th percentile of $|\Sigma^{\mathrm{MD}}|$ (blue $-1$, red $+1$); no method is rescaled. $r$ is the Pearson correlation with MD over the off-diagonal entries.

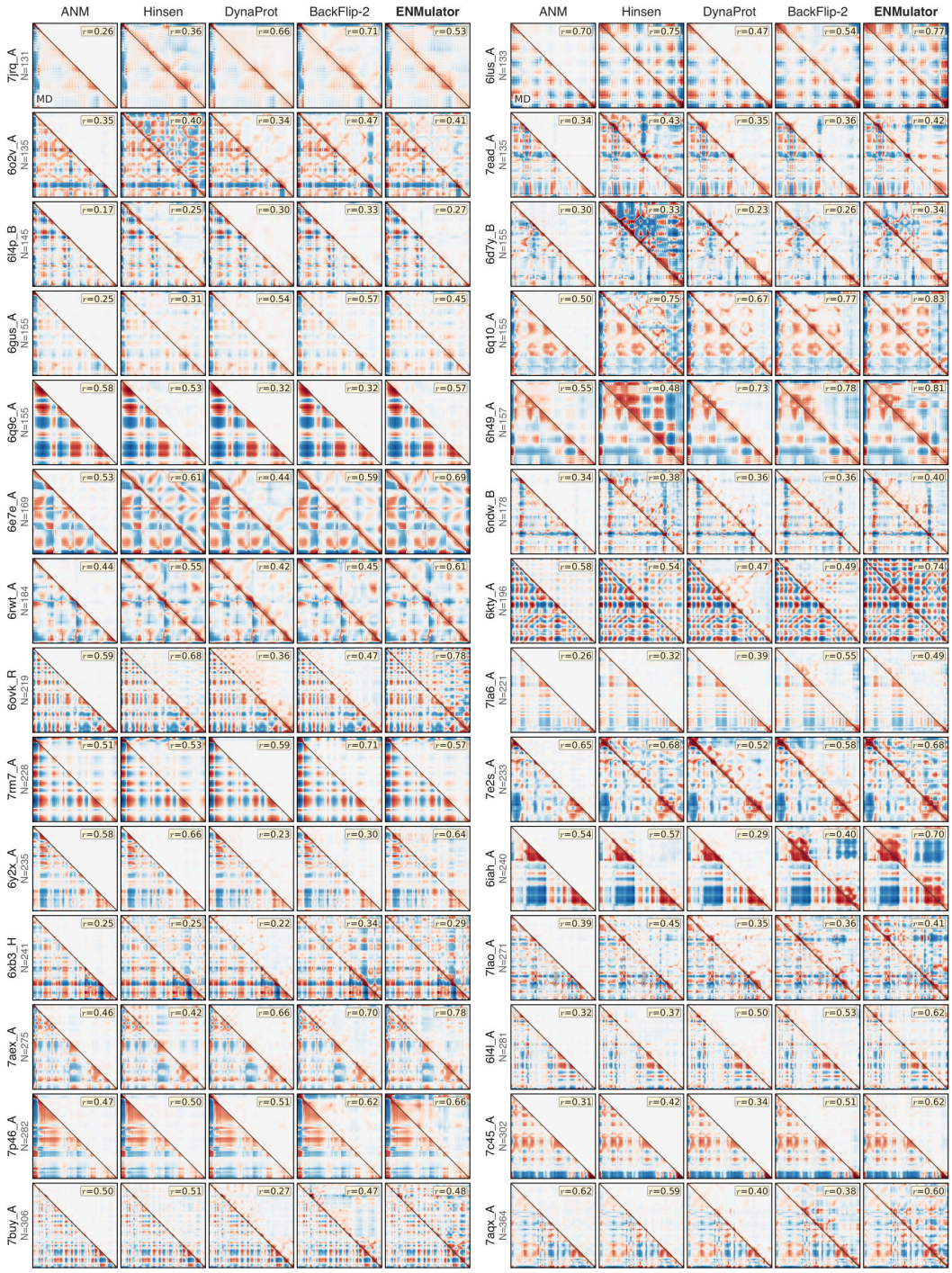


Figure A4: Predicted covariance on all 82 ATLAS test targets (continued).

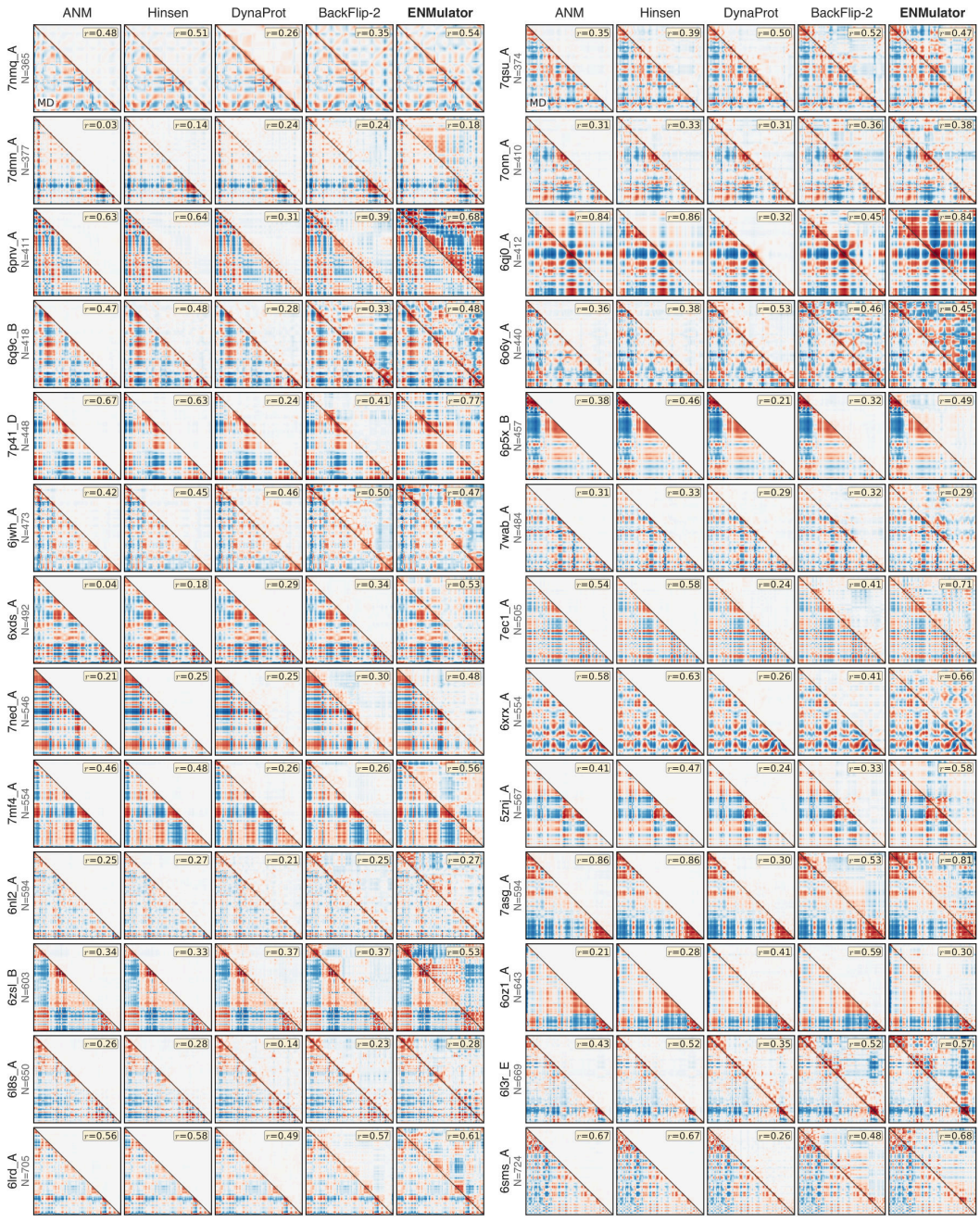


Figure A4: Predicted covariance on all 82 ATLAS test targets (continued).

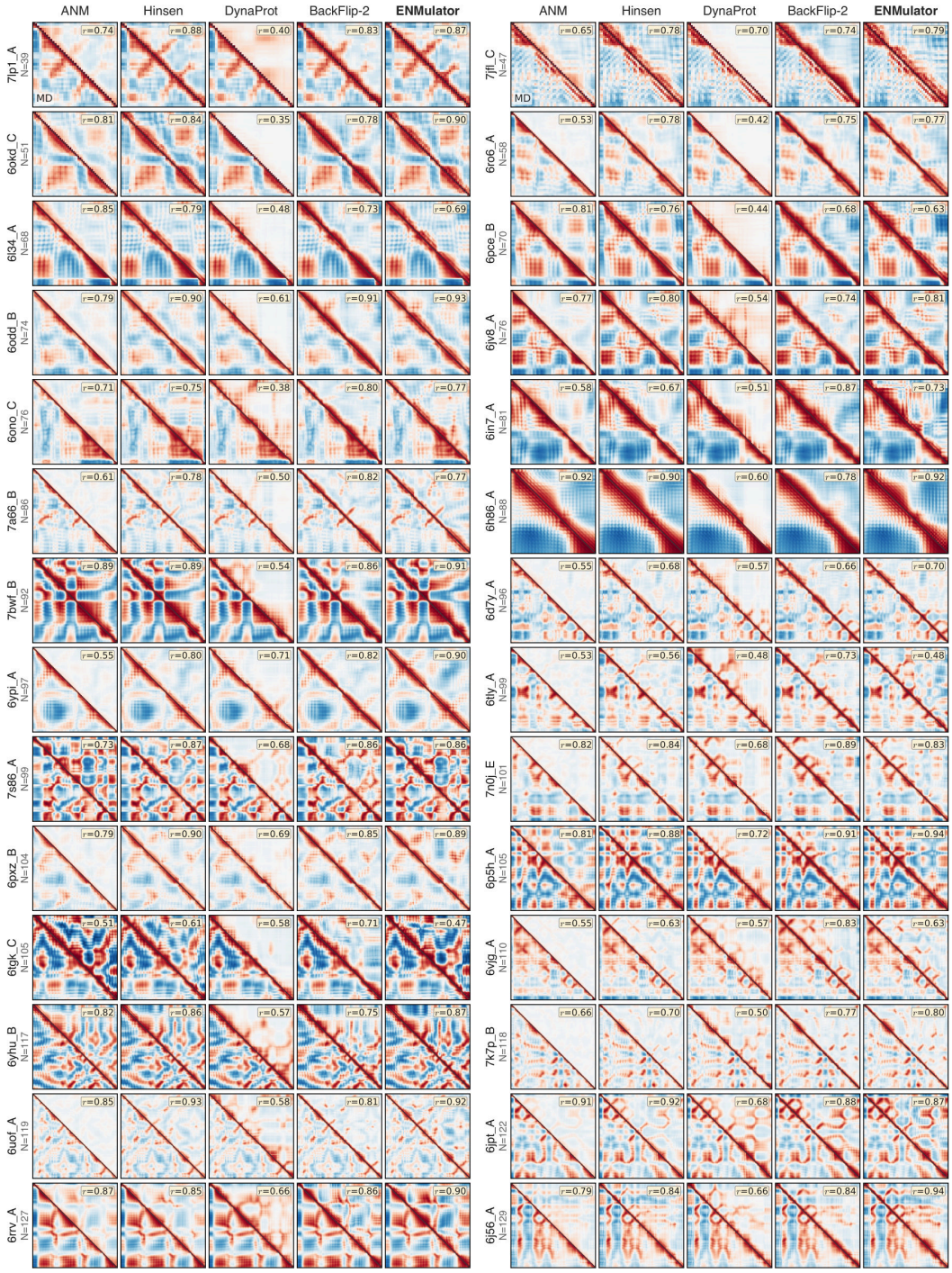


Figure A5: **Dynamic cross-correlation on all 82 ATLAS test targets.** Layout as in Figure A4; MD is below the diagonal and the prediction is above it and colors from $-1$ (blue) to $+1$ (red). $r$ is the Pearson correlation with MD over the off-diagonal entries.

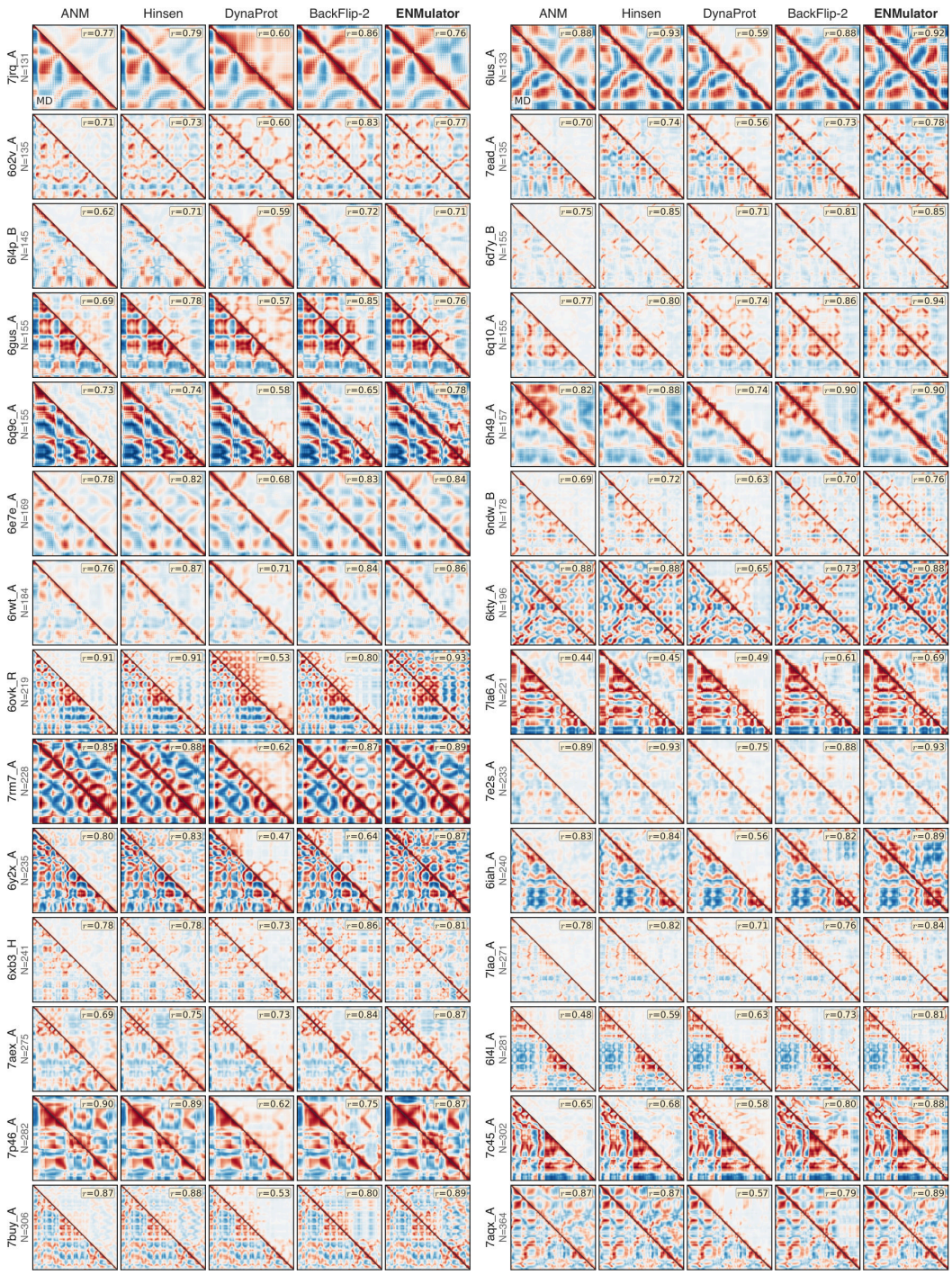


Figure A5: Dynamic cross-correlation on all 82 ATLAS test targets (continued).

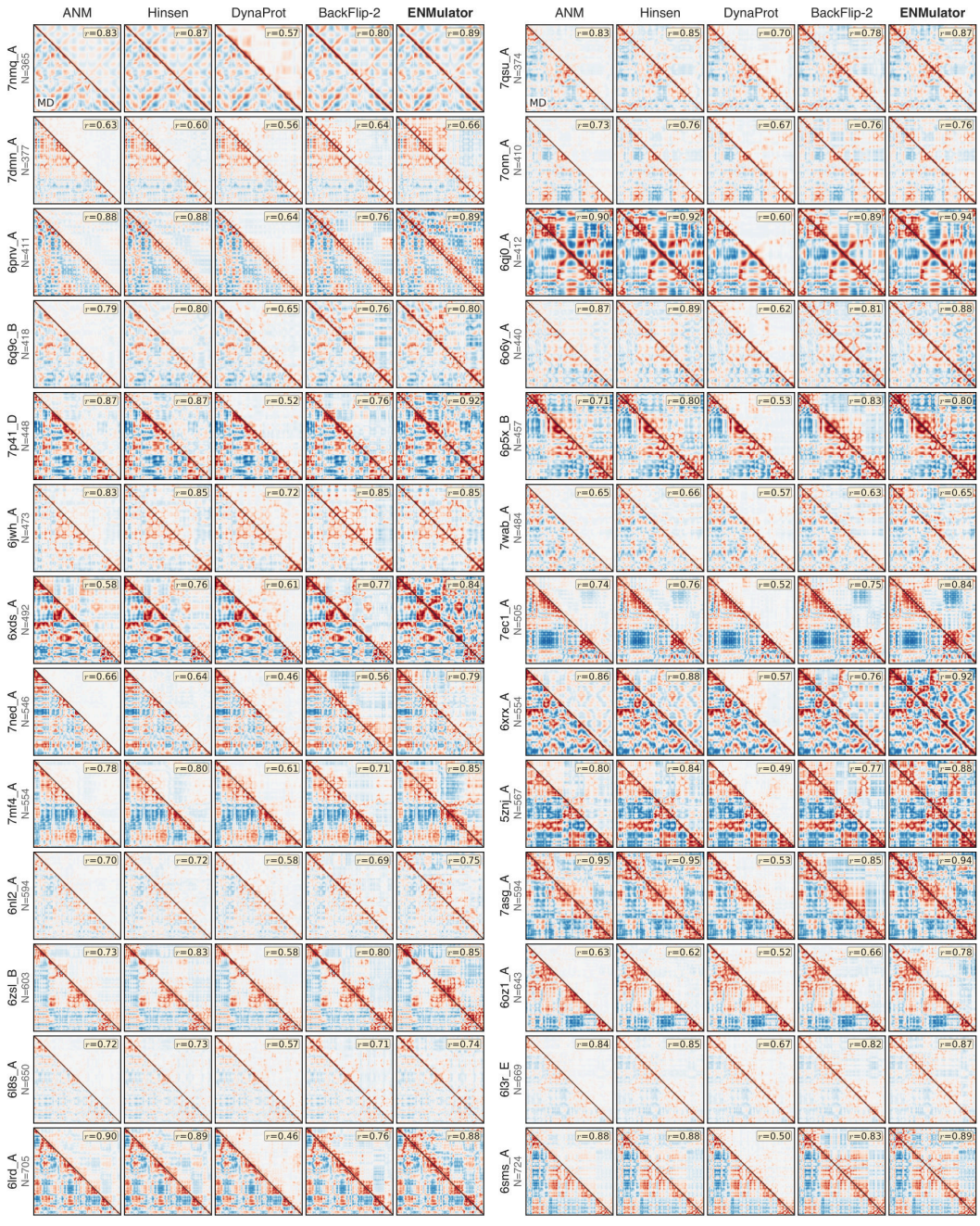


Figure A5: Dynamic cross-correlation on all 82 ATLAS test targets (continued).

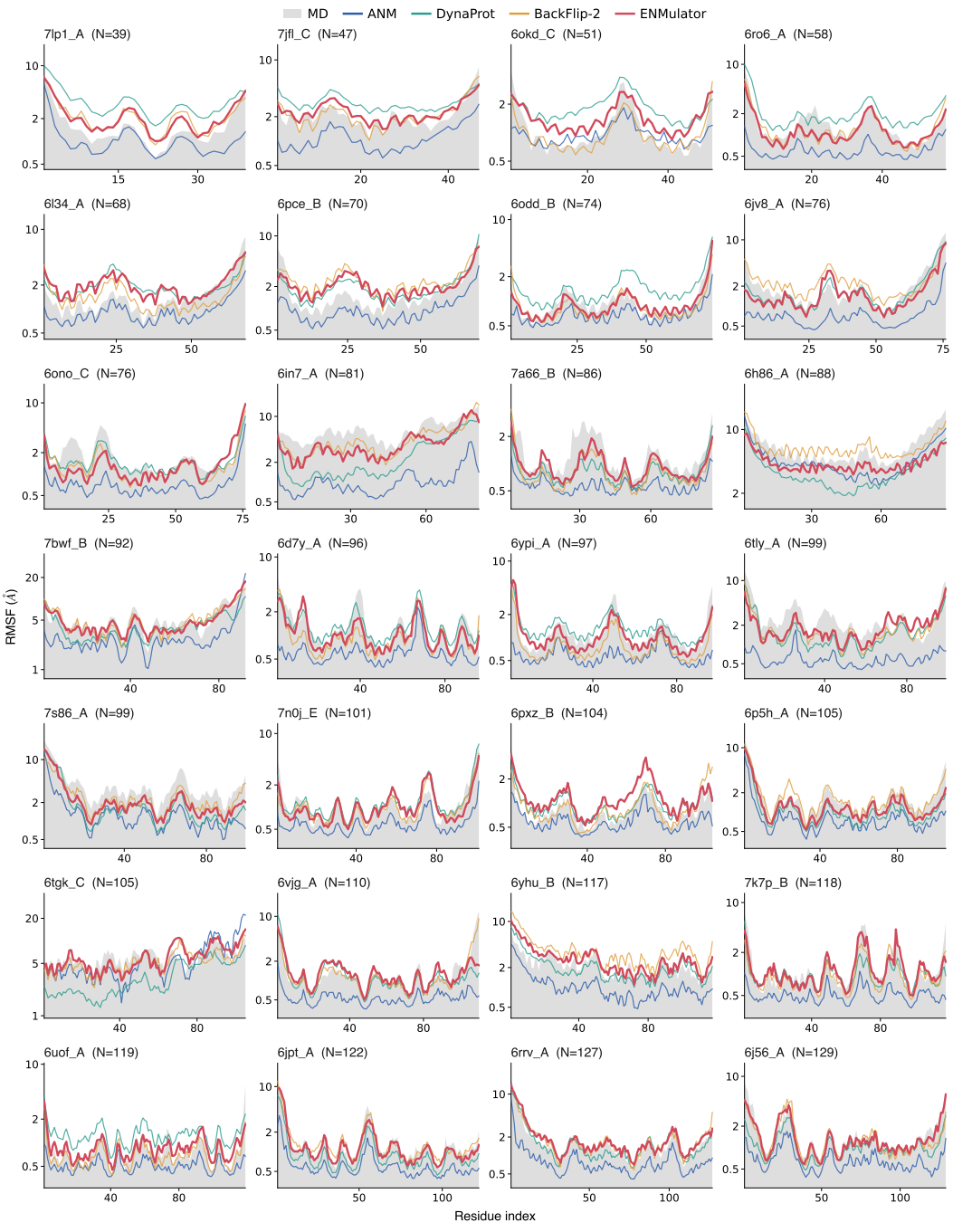


Figure A6: **Per-residue RMSF on all 82 ATLAS test targets.** Panels are ordered by increasing $N$, as in Figure A4. MD is shown as the gray area; for readability, only ANM and the learned predictors are overlaid. RMSF is computed from each predicted covariance without rescaling and shown on a logarithmic axis.

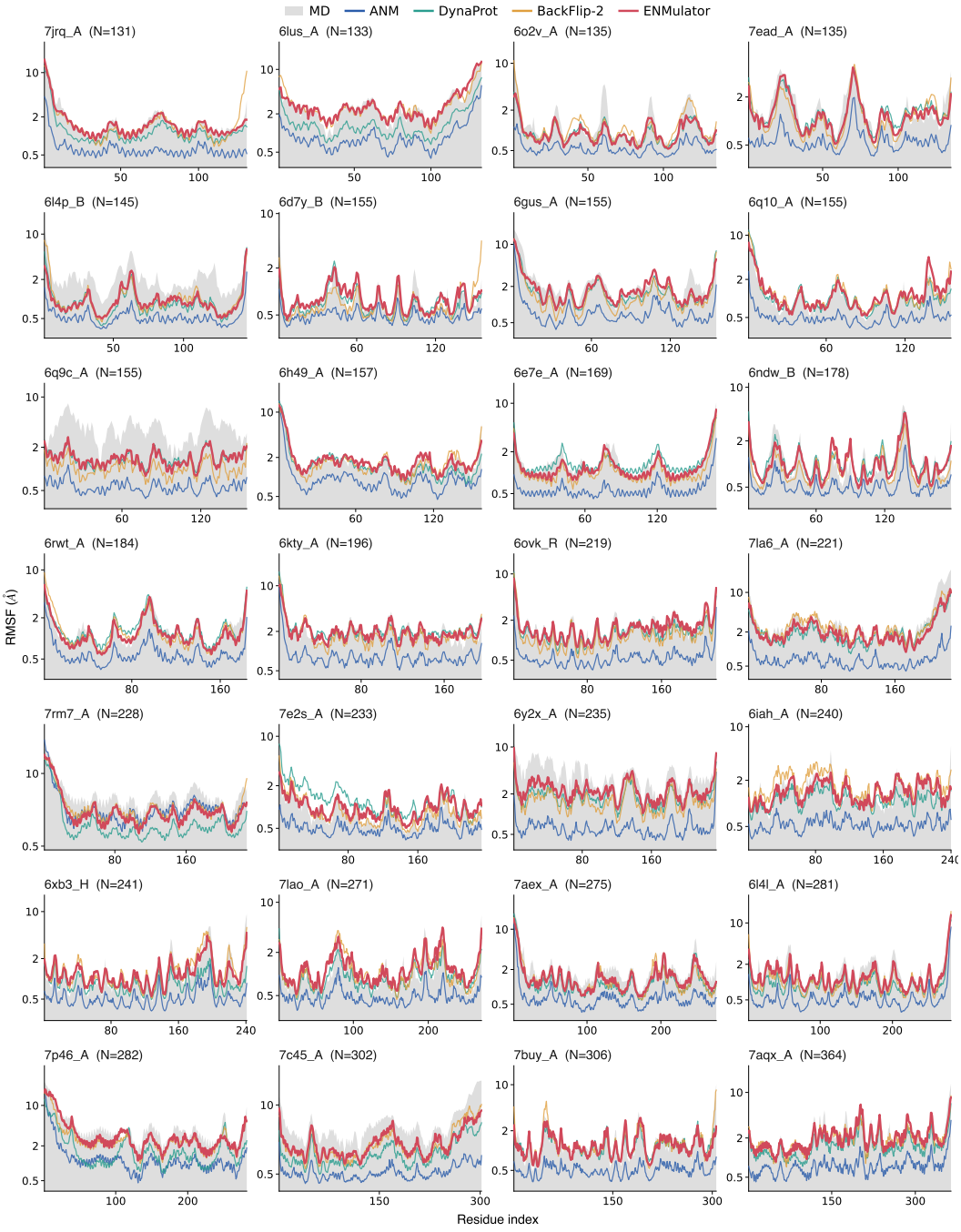


Figure A6: Per-residue RMSF on all 82 ATLAS test targets (continued).

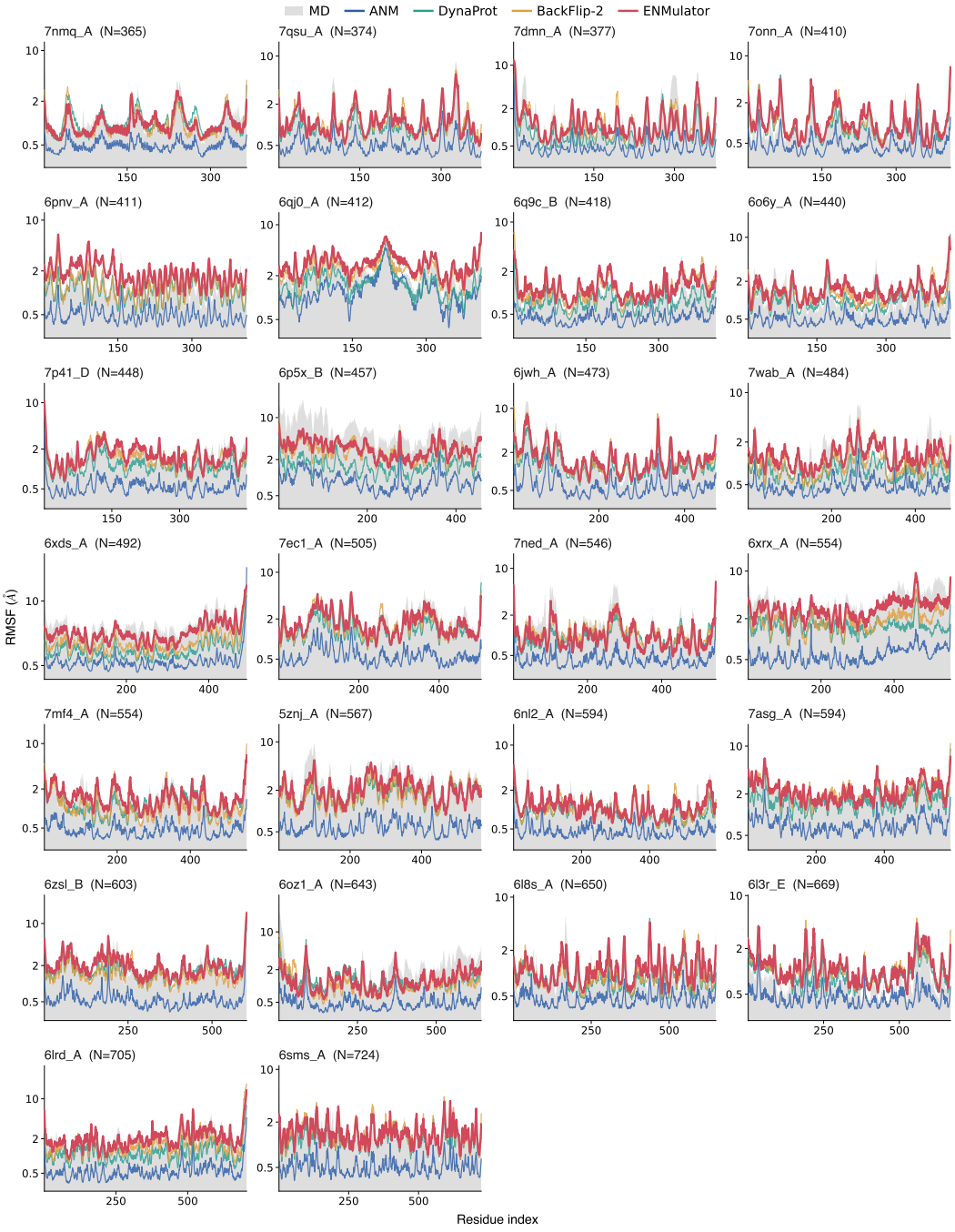


Figure A6: Per-residue RMSF on all 82 ATLAS test targets (continued).

Table A15: Extension of Table 2 with bootstrap standard errors (±) over the 82 test targets.

| Metric | *Hand-designed ENM* GNM | ANM | ANM* | pfANM | Hinsen | *Learning-based methods* DynaProt | BackFlip-2 | **ENMulator** |
|---|---|---|---|---|---|---|---|---|
| *Global covariance accuracy* | | | | | | | | |
| Relative Frobenius (↓) | 0.978±0.004 | 0.984±0.003 | 0.962±0.009 | 0.988±0.005 | 0.990±0.005 | 0.954±0.006 | 0.910±0.018 | 0.874±0.015 |
| Bures–Wasserstein $\mathcal{W}_2$ norm. (↓) | 0.928±0.016 | 0.839±0.009 | 0.856±0.007 | 0.978±0.016 | 0.867±0.019 | 0.825±0.013 | 0.752±0.012 | 0.666±0.019 |
| Pearson correlation $r$ (↑) | 0.292±0.011 | 0.448±0.026 | 0.468±0.022 | 0.348±0.017 | 0.496±0.017 | 0.417±0.023 | 0.516±0.031 | 0.589±0.024 |
| *Collective modes and spectrum* | | | | | | | | |
| Top-3 subspace overlap (↑) | 0.266±0.016 | 0.551±0.021 | 0.593±0.030 | 0.474±0.025 | 0.603±0.028 | 0.441±0.024 | 0.527±0.017 | 0.639±0.023 |
| Top-10 subspace overlap (↑) | 0.448±0.012 | 0.564±0.010 | 0.602±0.012 | 0.512±0.007 | 0.636±0.007 | 0.526±0.026 | 0.625±0.017 | 0.663±0.008 |
| Participation ratio (= 5.48) | 50.34±5.73 | 46.03±10.05 | 33.11±8.37 | 213.70±39.41 | 19.52±3.13 | 40.69±4.19 | 17.17±2.17 | 8.98±1.47 |
| *Per-residue marginals (diagonal) and off-diagonal coupling* | | | | | | | | |
| RMSF $r$ (↑) | 0.754±0.018 | 0.764±0.030 | 0.790±0.030 | 0.768±0.034 | 0.801±0.021 | 0.861±0.028 | 0.877±0.017 | 0.859±0.017 |
| RMSF MAE (↓) | 0.774±0.036 | 1.001±0.068 | 0.600±0.032 | 1.179±0.135 | 1.292±0.169 | 0.539±0.049 | 0.445±0.024 | 0.433±0.038 |
| DCCM $r$ (↑) | 0.566±0.016 | 0.779±0.016 | 0.791±0.015 | 0.668±0.023 | 0.824±0.019 | 0.579±0.011 | 0.797±0.013 | 0.856±0.012 |
| DCCM MAE (↓) | 0.233±0.011 | 0.205±0.006 | 0.197±0.007 | 0.246±0.015 | 0.166±0.006 | 0.235±0.018 | 0.167±0.006 | 0.138±0.008 |

Table A16: **Full version of Table 3.** All metrics and all baselines, including the additional ANM variant ANM*.

| Metric | *Hand-designed ENM* ANM | ANM* | Hinsen | *Learning-based covariance methods* DynaProt | BackFlip-2 | **ENMulator** | *Direct generation* AlphaFlow | BBFlow |
|---|---|---|---|---|---|---|---|---|
| *Efficiency* | | | | | | | | |
| # Parameters (↓) | — | — | — | 2.9M | 1.0M | 2.8M | 97M | 18.2M |
| Runtime (s) (↓) | 8.15 | 12.47 | 8.80 | 7.57 | 7.58 | 7.53 | 5,149.30 | 172.20 |
| *Ensemble spread and residue flexibility* | | | | | | | | |
| Pairwise RMSD (= 2.91 ± 0.23) | 0.79±0.05 | 2.12±0.13 | 2.51±0.38 | 2.01±0.08 | 2.29±0.13 | 2.38±0.17 | 2.18±0.17 | 3.09±0.23 |
| Pairwise RMSD $r$ (↑) | 0.64±0.09 | 0.65±0.09 | 0.19±0.12 | 0.71±0.08 | 0.81±0.05 | 0.83±0.04 | 0.94±0.02 | 0.82±0.06 |
| Pairwise RMSD MAE (↓) | 2.34±0.18 | 1.36±0.21 | 2.22±0.22 | 1.48±0.18 | 1.00±0.13 | 0.91±0.13 | 0.92±0.10 | 0.78±0.10 |
| $C_\alpha$ RMSF (= 1.48 ± 0.11) | 0.45±0.01 | 1.26±0.04 | 0.58±0.08 | 1.04±0.04 | 1.20±0.06 | 1.40±0.08 | 1.17±0.08 | 1.49±0.11 |
| Global RMSF $r$ (↑) | 0.56±0.06 | 0.58±0.05 | 0.38±0.08 | 0.73±0.04 | 0.81±0.03 | 0.82±0.03 | 0.91±0.01 | 0.84±0.03 |
| Global RMSF MAE (↓) | 1.50±0.15 | 0.90±0.12 | 1.47±0.15 | 0.96±0.14 | 0.75±0.11 | 0.65±0.09 | 0.62±0.07 | 0.62±0.08 |
| Per-target RMSF $r$ (↑) | 0.77±0.03 | 0.81±0.03 | 0.81±0.02 | 0.87±0.03 | 0.88±0.02 | 0.87±0.01 | 0.92±0.01 | 0.90±0.01 |
| Per-target RMSF MAE (↓) | 1.08±0.07 | 0.52±0.03 | 1.16±0.14 | 0.57±0.05 | 0.47±0.03 | 0.41±0.03 | 0.41±0.05 | 0.41±0.06 |
| *Correlated motion and distributional accuracy* | | | | | | | | |
| Global DCCM $r$ (↑) | 0.69±0.03 | 0.71±0.03 | 0.77±0.02 | 0.58±0.02 | 0.76±0.01 | 0.84±0.01 | 0.88±0.01 | 0.84±0.01 |
| Global DCCM MAE (↓) | 0.19±0.01 | 0.19±0.01 | 0.17±0.01 | 0.22±0.01 | 0.17±0.01 | 0.14±0.01 | 0.12±0.01 | 0.15±0.00 |
| Per-target DCCM $r$ (↑) | 0.74±0.02 | 0.77±0.02 | 0.82±0.01 | 0.68±0.02 | 0.80±0.01 | 0.86±0.01 | 0.89±0.01 | 0.87±0.01 |
| Per-target DCCM MAE (↓) | 0.19±0.01 | 0.18±0.01 | 0.17±0.01 | 0.20±0.01 | 0.16±0.01 | 0.14±0.01 | 0.12±0.00 | 0.15±0.00 |
| RMWD (↓) | 2.66±0.19 | 2.43±0.19 | 2.73±0.17 | 2.39±0.14 | 2.25±0.14 | 2.32±0.15 | 1.71±0.13 | 1.96±0.16 |
| Translation contrib. (↓) | 2.07±0.13 | 2.01±0.13 | 2.06±0.14 | 2.05±0.13 | 2.00±0.13 | 2.05±0.13 | 1.49±0.10 | 1.63±0.12 |
| Variance contrib. (↓) | 1.51±0.08 | 1.15±0.07 | 1.77±0.13 | 1.14±0.08 | 0.93±0.06 | 0.92±0.08 | 0.80±0.07 | 0.95±0.09 |
| MD PCA $\mathcal{W}_2$ (↓) | 1.93±0.17 | 1.72±0.18 | 1.81±0.13 | 1.81±0.15 | 1.67±0.14 | 1.63±0.17 | 1.29±0.11 | 1.30±0.09 |
| Joint PCA $\mathcal{W}_2$ (↓) | 2.48±0.16 | 2.32±0.17 | 2.43±0.14 | 2.37±0.13 | 2.18±0.15 | 2.19±0.18 | 1.55±0.12 | 1.73±0.14 |
| PC similarity > 0.5 (%) (↑) | 28.0±4.9 | 28.0±4.9 | 30.5±5.0 | 19.5±4.3 | 29.3±5.0 | 37.8±5.3 | 43.9±5.5 | 40.2±5.4 |
| *Ensemble observables and structural validity* | | | | | | | | |
| Weak contacts $J$ (↑) | 0.39±0.01 | 0.54±0.01 | 0.42±0.02 | 0.55±0.01 | 0.56±0.01 | 0.56±0.01 | 0.62±0.01 | 0.57±0.01 |
| Transient contacts $J$ (↑) | 0.30±0.01 | 0.26±0.01 | 0.26±0.01 | 0.30±0.02 | 0.31±0.01 | 0.35±0.01 | 0.47±0.01 | 0.28±0.01 |
| Conf. with $C_\alpha$–$C_\alpha$ break (%) (↓) | 0.00±0.00 | 0.00±0.10 | 0.40±0.73 | 0.80±0.56 | 3.20±2.08 | 0.00±0.09 | 1.60±0.45 | 0.00±0.00 |