# DNAFLOW GENERATES MOLECULAR ENSEMBLES BY MULTISCALE PHYSICS-INFORMED FLOW MATCHING

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

## ABSTRACT

Macromolecular function emerges from structures and their conformational ensembles, whose high-dimensional distributions remain difficult to learn from limited atomistic simulation data. Here, we introduce DnaFlow, a physics-informed flow matching framework and architecture that decomposes molecular fluctuations across spatial scales, focusing on the fast generation of structural ensembles of double-stranded DNA. Repeated forward passes generate atomistic ensembles that reproduce the sequence-dependent conformational distributions obtained from state-of-the-art atomistic molecular dynamics simulations. The central idea is to decompose the chain's local deformation field across spatial frequencies, assigning each scale a distinct generative pathway: slow bending modes are modeled with a latent prior derived from a polymer model, specifically the worm-like chain (WLC), high-frequency roughness by a calibrated noise field, and sequence-specific equilibrium structure with the flow matching objective. By learning only local, length-independent statistics while assigning global fluctuations to a physical prior, DnaFlow enables sequence-dependent extrapolation to chain lengths beyond those observed during training, extending the scope of existing generative models of DNA. More broadly, our approach provides a general framework for multiscale generative modeling that combines physical priors with learned, system-specific conformational distributions.

## 1 INTRODUCTION

The sequence of bases along a DNA double helix encodes not only the proteins that carry most of the functions sustaining life, but also the molecule's own mechanical properties (Rief et al., 1999; Travers, 2004). Base pair (bp) composition dictates DNA's rigidity, and thereby the energetics of genome compaction and DNA–protein interaction (Widom, 2001; Morozov et al., 2009; Afek et al., 2020). Flexibility in turn governs fundamental processes such as transcription efficiency and is associated with disease (Pearson & Sinden, 1998). Understanding sequence-dependent structure therefore means obtaining a distribution of conformations that approximates thermodynamic equilibrium, not a single structure.

The conformational landscape of DNA is typically assessed experimentally by means of structural biology or *in silico* by molecular dynamics (MD) and quantum-mechanical calculations (Lankaš et al., 2003; Dans et al., 2019), all limited in throughput. MD is the established route to a full distribution (Tuckerman & Martyna, 1999), affordable to the microsecond regime for short duplexes ($\sim$20 bp). The Ascona B-DNA Consortium's hexABC dataset (Battistini et al., 2026) pushes this to 380 trajectories of at least ten microseconds over all hexamers. Beyond the computational resources this demands, the sampling problem for DNA is combinatorial: complete sequence coverage grows as $4^k$ in the context width $k$, while long-range correlations along DNA (Li et al., 1994) demand chains far longer than the context being covered. Sequence-dependent studies have consequently been confined to short duplexes or to coarse-grained models (Petkevičiūtė et al., 2014; Becker & Everaers, 2007; Skoruppa & Schiessel, 2025; Šulc et al., 2012), and generating high-quality structural ensembles across length regimes would lift this bottleneck.

Whereas structure prediction for proteins (Jumper et al., 2021; Baek et al., 2021; Lin et al., 2023; Yim et al., 2023b) and RNA (Tarafder & Bhattacharya, 2026; Shen et al., 2024; Baek et al., 2024) is well developed, the literature on learning atomistic DNA structures, or their distribution, is sparse



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.

(Abramson et al., 2024; Li et al., 2024; Liebl, 2026), possibly due to a lack of data spanning larger ranges of DNA length and composition. We therefore build our training libraries by hierarchical chain growth (HCG), which stacks duplexes from randomly sampled frames of atomistic MD until the target oligomer sequence is reached, and has been shown to scale and to reproduce state of the art MD and experimental results (Manuscript in preparation).

**Our approach.** We adopt the Riemannian flow matching framework of Yim et al. (2024) to generate double-stranded DNA structures invariant under the special Euclidean group $\mathrm{SE}(3)$. The architecture is designed around one question: how local geometry translates into global fluctuations in a sequence-dependent manner. Exploiting the symmetries intrinsic to a duplex, we build the atomistic strands around a simultaneously generated centerline rod whose fluctuations match the reference ensemble. This generates accurate DNA ensembles of more than a hundred base pairs within seconds on a single GPU.

## 2 Background and Preliminaries

### 2.1 Flow Matching on Riemannian Manifolds

Conditional flow matching (CFM) on a Riemannian manifold $\mathcal{M}$ requires care in defining the interpolation between a noise sample $x_0 \sim p_0$ and a data sample $x_1 \sim q$, with $p_0$ a tractable prior, $q$ the data distribution, and $t \in [0,1]$ the *flow time*. Since $\mathcal{M}$ has no global vector space structure in general, the interpolation is taken along geodesics, whose tangent spaces $T_x\mathcal{M}$ are locally linear. Riemannian flow matching (RFM) learns the conditional vector field $u_t(x|x_1) \in T_x\mathcal{M}$ by regressing a neural network $u_t^\theta$ onto it, where $\theta$ collects every trainable weight of the model[1]. Sampling integrates the ODE $\mathrm{d}x_t/\mathrm{d}t = u_t^\theta(x_t)$ from $t=0$ to $t=1$ with a first-order Euler scheme whose steps are taken along geodesics, so the iterates stay on $\mathcal{M}$. This transports randomly distributed frames $T = (\boldsymbol{R}, \boldsymbol{x})$, with $\boldsymbol{R} \in \mathrm{SO}(3)$ and $\boldsymbol{x} \in \mathbb{R}^3$, onto the DNA double helix[2]. On $\mathrm{SE}(3)$ we take the product metric of Yim et al. (2023b), which lets rotations and translations be treated independently, and set $p_0 = \mathcal{U}(\mathrm{SO}(3)) \otimes \mathcal{N}(0, \boldsymbol{I}_3)$, so that the geodesic path $T_t = \exp_{T_0}(t \log_{T_0}(T_1))$ is a linear interpolation along the geodesics of each factor: the convex combination $\boldsymbol{x}_t = (1-t)\boldsymbol{x}_0 + t\boldsymbol{x}_1$ for translations, and $\boldsymbol{R}_t = \exp_{\boldsymbol{R}_0}(t \log_{\boldsymbol{R}_0}(\boldsymbol{R}_1))$ for rotations. Invariance under $\mathrm{SE}(3)$ additionally requires the induced probability path to be centered (Yim et al., 2024; Köhler et al., 2020), so every molecule's center of mass is restricted to the corresponding subspace of $\mathbb{R}^{N\times 3}$ in the dataset and during integration. We cover only what our approach needs: the metric, and the reparameterized loss are in Appendix A.1, with the full treatment in Chen & Lipman (2024); Yim et al. (2024; 2023b).

### 2.2 Worm-Like Chain Model

The worm-like chain (WLC) (Kratky & Porod, 1949) is widely accepted to describe the behavior of DNA in solution (Marko & Siggia, 2002; Bustamante et al., 2003). A polymer of contour length $L$ resists bending with the energy $H_{\mathrm{bend}} = \frac{\epsilon}{2}\int_0^L (\partial_s \boldsymbol{t}(s))^2 \,\mathrm{d}s = \frac{\epsilon}{2}\int_0^L \boldsymbol{\kappa}^2(s)\,\mathrm{d}s$, with $s$ the arc length along the chain, $\boldsymbol{t}(s) = \partial_s \boldsymbol{r}(s)$ the unit tangent vector along the parametric curve $\boldsymbol{r}(s)$, and $\boldsymbol{\kappa}(s) = \partial_s \boldsymbol{t}(s)$ the *curvature vector* (Razbin & Benetatos, 2023). For a three-dimensional WLC the tangent correlation function is $\langle \boldsymbol{t}(s)\cdot\boldsymbol{t}(0)\rangle = e^{-s/L_p}$, where the persistence length $L_p$ describing the polymer's stiffness (Baschnagel et al., 2016) is proportional to the bending rigidity $\epsilon = k_B T L_p$ (Pilyugina et al., 2017). The angle enclosed by the two terminal tangents, $\Theta = \angle\big(\boldsymbol{t}(0), \boldsymbol{t}(L)\big)$, is the chain's *bend angle*, for which equipartition gives $\langle \Theta^2 \rangle = 2L/L_p$ in the stiff limit, $L \ll L_o$; we use $\Theta$ throughout as the global measure of bending.

## 3 DnaFlow for Generation of Equilibrium Ensembles

Our architecture follows from three requirements: (1) atomistic structures that reproduce the reference diversity, (2) extrapolation to unseen lengths, and (3) sequence dependence. Atomistic detail comes from a torsion head on the backbone trunk; sequence dependence from 1D convolutions over

[1] Throughout, subscripts $\theta$ mark learned maps and hats the quantities the network predicts.

[2] Here, $T$ is a rigid-body transformation, unrelated to the tangent space $T_x\mathcal{M}$.

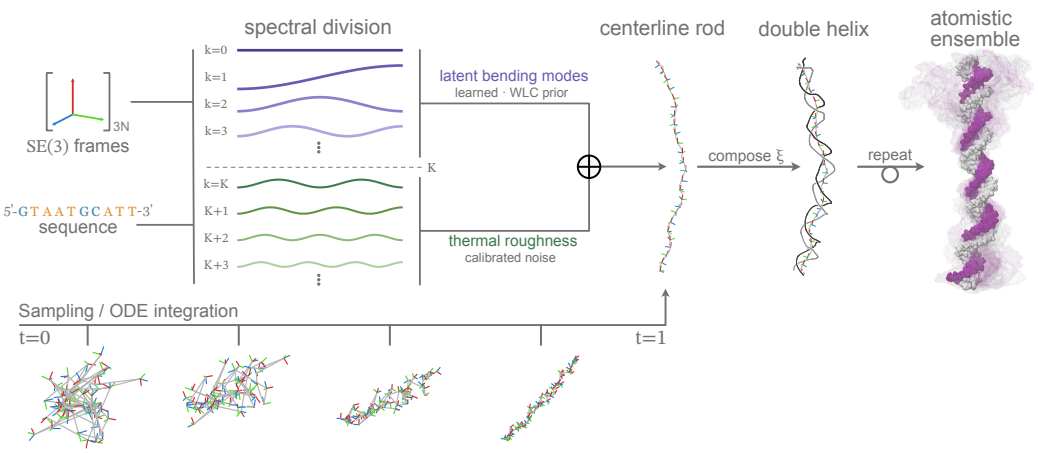



Figure 1: **Overview of DnaFlow's architecture.** Corrupted frames and the DNA sequence serve as inputs. DnaFlow learns the vector field which transports noisy frames to clean centerline rods and DNA strands. The architecture makes use of a spectral division to inject stochasticity at different scales of the latent space. Sampled rods provide the correct amount of fluctuations, onto which the DNA backbone is composed with base pair specific offsets $\boldsymbol{\xi}$. Repeating this procedure multiple times produces the ensemble.

local neighborhoods and base-pair embeddings. The challenge is to match ensemble statistics while preserving extrapolation, as sequence-dependent shape dispersion extends beyond the local context, with small changes in local geometry inducing large changes in global observables (Matsumoto & Olson, 2002; Gardasevic & Noy, 2024). This splits the problem in two: flow matching produces the sequence-specific equilibrium shape, while the spread of conformations around it comes from an explicit statistical model of the chain, never from supervision on diversity, which a training set of single conformations cannot provide (Figure 1, and Figure 4 for more details).

Base-pair frames along the centerline follow the convention of Olson et al. (2001), and the *step* frame is the symmetric geodesic midpoint of a base frame and the flipped frame of its partner on the antiparallel strand (Appendix A.3). Each step between consecutive base pairs then carries six numbers: a rotation vector $\boldsymbol{\omega}_i \in \mathbb{R}^3$ whose components are the rigid-base-pair rotations tilt, roll and twist, and a translation $\boldsymbol{\tau}_i \in \mathbb{R}^3$ whose components are shift, slide and rise, both expressed in that step's own frame; Figure 5 draws all six. We use *law* throughout for the conditional distribution of that six-vector given the sequence in a bounded neighborhood of step $i$, i.e., the per-step mean field together with the amplitude of the fluctuations about it. It is *local* because it reads a bounded window of sequence, and *length-independent* because the same law is applied at every step of every chain, whatever that chain's length.

### 3.1 Centerline Rod Model

That the global conformation of DNA can largely be reconstructed from local step parameters (El Hassan & Calladine, 1997; Lu & Olson, 2003) is the first inductive bias in our architecture. On top of the published FrameFlow architecture, we add sequence context with a strand-aware 1D convolution-derived embedding (receptive field 13 bp), edge-type embeddings that encode the spacing between the center base pair $i$ and its neighbors $j$, and rotary positional embeddings (Su et al., 2023), which encode relative distance and therefore repeating sequence motives. This paradigm is further reinforced by our local backbone propagation: a local frame-aware message passing along the strand's backbone, in which each frame's node embedding is updated from the displacements and relative rotations of its neighbours within a window $W$, expressed in its own local frame; the concatenation over neighbours deliberately imposes an ordering, and the aggregation is invariant under the group action of $\mathrm{SE}(3)$ (Appendix A.4, proof in Appendix A.5). The rod trunk produces a node derived rod embedding $\boldsymbol{g}$ consumed by the backbone trunk, and predicts the clean frame poses at the ODE endpoint $t = 1$ as rotations $\boldsymbol{R}_{t=1}$ together with per-step displacements $\boldsymbol{\tau}_{t=1}$ from a small

multi-layer perceptron (MLP), the tau head. The step translation is $\boldsymbol{\tau}_i = \boldsymbol{R}_{\mathrm{ref},i}^{\top}(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)$, with $\boldsymbol{R}_{\mathrm{ref},i}$ the step reference frame, the geodesic midpoint of $\boldsymbol{R}_i$ and $\boldsymbol{R}_{i+1}$. The centerline is recovered from the rotations by integrating these steps, $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{R}_{\mathrm{ref},i}\boldsymbol{\tau}_i$, so the accumulated rotations set the global shape while $\boldsymbol{\tau}$ stays local.

### 3.2 Atomistic Strand Model

The DNA backbone is predicted conditioned on the generated centerline rod, the heuristic being that the rod carries most of the large scale ensemble variability. We additionally exploit the coarse geometric symmetry of the two strands wrapping around the centerline. Sequence context for each strand is incorporated in the $5' \to 3'$ direction, using the sequence embedding mechanism of the rod trunk, so that the model may recognize a recurring motif in either strand more easily. The edge type encoding in the backbone trunk is more elaborate, distinguishing strand membership as well as distance. The atomistic detail is derived from a torsion head, a small MLP reading the final backbone frame geometry, whose outputs are the torsion angles $\vartheta_{i,k}$ for each bond $k$ of nucleotide $i$, predicted as $(\cos\vartheta_{i,k}, \sin\vartheta_{i,k})$ on the unit circle so that the periodicity of an angle is respected (adapted from (Baek et al., 2024)). Additionally, the backbone's offset head predicts the local per-base offset $(\boldsymbol{\omega}, \boldsymbol{v})_i = \boldsymbol{\xi}_i \in \mathbb{R}^6$ from the neighboring base pair frame. Because our frames follow Olson et al. (2001), $\boldsymbol{\xi}$ is approximately half of the intra-base-pair parameter set: the rotation vector $\boldsymbol{\omega}$ carries buckle, propeller and opening, and the translation vector $\boldsymbol{v}$ the local base displacements shear, stretch and stagger. The composition of $\boldsymbol{\xi}_i$ with the rod frame of the same index is $T_i = T_i^{\mathrm{rod}} \circ \boldsymbol{\xi}_i = \big(\boldsymbol{R}_i^{\mathrm{rod}} \exp(\boldsymbol{\omega}_i),\ \boldsymbol{x}_i^{\mathrm{rod}} + \boldsymbol{R}_i^{\mathrm{rod}}\boldsymbol{v}_i\big)$, where the exponential map $\exp$ acts on rotation vectors and $\boldsymbol{v}$ is expressed in the rod's local reference frame. Each strand, A and B, composes onto the same centerline, with their own offsets $\boldsymbol{\xi}_i^A, \boldsymbol{\xi}_i^B$. Hence the base pair deformation is encoded by the two, while the coupling to the centerline prohibits the strands from diverging away from physical solutions and carries the large scale conformational changes to the atoms. At training time we compose the two strands onto the ground truth centerline, such that the offset target is exact and independent of the rod's error; during sampling, it is composed onto the predicted rod, so the atomistic structures inherit the ensemble properties.

### 3.3 Spectral Division

Dispersing FrameFlow's predictions by varying its initial noise alone produces valid double helices but does not disperse them correctly across length scales (Appendix Figure 6). This failure, with our requirement of sensitivity to sequence-dependent dispersion beyond the local ($\sim$ 5 bp) scale, motivates *spectral division*: decompose the local deformation field by spectral frequency and give each band exactly one owner. The chain's bending modes are carried by a latent $\boldsymbol{z}$ with a WLC prior, high-frequency roughness by a calibrated noise field, and the sequence-specific equilibrium shape by the rod trunk's flow-matching objective. The architecture operates on the local-frame inter-step rotation vector $\boldsymbol{\omega}_i = \log(\boldsymbol{R}_i^{\top}\boldsymbol{R}_{i+1})^{\vee} = \int_{s_i}^{s_{i+1}} \boldsymbol{\kappa}(s)\,\mathrm{d}s \approx \boldsymbol{\kappa}_i h \in \mathbb{R}^3$ with the logarithmic map and vee operator mapping from $\mathrm{SO}(3)$ to the Lie algebra $\mathfrak{se}(3)$, and the WLC curvature vector integrated over one base-pair step of mean rise $h \approx 3.3$ Å. Its three components in the local frame are the rigid-base-pair rotations tilt, roll and twist, indexed by $\alpha \in \{x, y, z\}$, the transverse axes $x, y$ carrying bending and $z$ carrying twist, measured in radians per step.

The split cannot be taken on $\boldsymbol{\omega}$ directly: viewed in the local frame, a smooth global bend is modulated by the equilibrium twist per step $\Omega_0 \approx 0.61$ rad/step, so that the duplex repeats every $p_h = 2\pi/\Omega_0 \approx 10.3$ bp and a global bend appears at a mode index growing linearly in $N$, which a fixed $K$ would stop covering. Demodulating by the accumulated helical phase $\varphi_i = \Omega_0\, i$ before projection removes this drift, $\tilde{\boldsymbol{\omega}}_i = \mathcal{D}\boldsymbol{\omega}_i := \boldsymbol{R}_z(\varphi_i)\,\boldsymbol{\omega}_i$: it turns the two transverse components back by the phase the frame has accumulated by step $i$ and leaves the twist channel untouched, and by being a per-step rotation it is orthogonal, with inverse $\mathcal{D}^{-1}$ undoing the same phase. The demodulated field is projected onto an orthonormal cosine basis, the type-II discrete cosine transform (DCT-II), written as a fixed matrix $\boldsymbol{C} \in \mathbb{R}^{N\times N}$ whose rows are indexed by the spatial frequency $k = 0, \ldots, N-1$ along the chain, $k = 0$ being the chain-wide mean and large $k$ step-to-step roughness. Projection is $\boldsymbol{\mu} = \boldsymbol{C}\tilde{\boldsymbol{\omega}}$ and re-expansion its exact transpose $\boldsymbol{C}^{\top}\boldsymbol{\mu}$ (Appendix A.7); before projection the twist channel alone is centered on the equilibrium twist per step, so that its zero-frequency coefficient captures the chain's over- or under-twist while the transverse ones keep the net bend (Appendix A.8). In *spectral division* we implement the split of that inverse at a cut $K = 8$, held fixed rather

than scaled with $N$ so that one latent serves every chain length

$$\tilde{\boldsymbol{\omega}}_i = \sum_{k=0}^{N-1} C_{ki}\,\boldsymbol{\mu}_k = \underbrace{\sum_{k<K} C_{ki}\,\boldsymbol{\mu}_k}_{\text{low-frequency } \textit{bending modes}} + \underbrace{\sum_{k=K}^{N-1} C_{ki}\,\boldsymbol{\mu}_k}_{\text{high-frequency } \textit{thermal modes}} .$$

The sum runs over every frequency, but only the low band is carried as a latent, and from here on the unindexed $\boldsymbol{\mu} = (\mu_{k\alpha})_{k<K} \in \mathbb{R}^{K\times 3}$ denotes that block alone. With $\boldsymbol{C}_{<K}$ the first $K$ rows of $\boldsymbol{C}$, the two bands are the images of the complementary orthogonal projectors $\mathcal{P}_{<K} = \mathcal{D}^{-1}\boldsymbol{C}_{<K}^{\top}\boldsymbol{C}_{<K}\mathcal{D}$ and $\boldsymbol{I} - \mathcal{P}_{<K}$, in body-frame coordinates. With $\mathcal{D}$ the demodulation operator of Section 3.3, they apply to $\boldsymbol{\omega}$ directly. Being orthogonal and complementary, the two bands cannot mix, so each has exactly one owner. Throughout, $\boldsymbol{\mu}$ is the parameter-free *measurement* made on clean conformations and $\hat{\boldsymbol{\mu}}$ the network's sequence-conditioned *prediction* of its ensemble average.

Orthonormality allows us to place a diagonal Gaussian prior $\mathcal{N}(\hat{\boldsymbol{\mu}}, \sigma_p^2)$ on the coefficients, whose spread $\sigma_p$ follows from WLC equipartition: each transverse axis carries the per-step variance $\sigma_{\omega\alpha}^2 = \langle \omega_{i\alpha}^2 \rangle = h/L_p$, a dimensionless constant independent of position and chain length, which the projection maps to a coefficient spread $\sigma_p[k,\alpha] = \sigma_{\omega\alpha}$ flat in $k$ and independent of $N$, so one calibrated prior serves every length. The macroscopic bend is then not calibrated but accumulated, recovering the WLC result $\langle \Theta^2 \rangle = 2L/L_p$: the $\sqrt{N}$ growth of the rod's bend follows from integrating a length-independent local law rather than from a length-dependent prior (Appendix A.9).

Sequence dependence of the predicted coefficients is incorporated by a 1D convolution stack (`SeqMeanFieldHead`) that emits a per-step residual field $\Delta\boldsymbol{\omega}(\mathrm{seq})$ on the equilibrium step and is fed through the same projection, $\hat{\boldsymbol{\mu}} = \boldsymbol{C}_{<K}\,\mathcal{D}(\boldsymbol{\omega}_{\mathrm{eq}} + \Delta\boldsymbol{\omega}(\mathrm{seq}))$, where $\boldsymbol{\omega}_{\mathrm{eq}}$ is the sequence-independent equilibrium step, calibrated once on the training ensemble, with $\boldsymbol{z} \in \mathbb{R}^{K\times 3}$ and $\boldsymbol{z} \sim \mathcal{N}(\hat{\boldsymbol{\mu}},\, \sigma_p^2)$ the spectral latent.

During training the projection is applied to *uncorrupted* rod frames, and $\boldsymbol{\mu}$ is perturbed to $\boldsymbol{z} = \boldsymbol{\mu} + \sigma_p\boldsymbol{\varepsilon}$, with $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ a standard normal draw, one per training sample, and read by both trunks through cross-attention in every invariant point attention (IPA) block, carrying bend information their corrupted input lacks. Simultaneously $\boldsymbol{\mu}$ is the regression target of the mean-field head, $\mathcal{L}_\mu = \big\langle \big((\mathrm{sg}[\boldsymbol{\mu}] - \hat{\boldsymbol{\mu}})/\sigma_p\big)^2 \big\rangle$, the stop-gradient $\mathrm{sg}[\cdot]$ fitting the head to the measurement. Its population optimum is the sequence-conditioned mean spectrum $\mathbb{E}[\boldsymbol{\mu}|\mathrm{seq}]$, with no precomputed per-sequence averages (Appendix A.10). At inference no clean conformation exists and the measurement is replaced by the regressed latent $\boldsymbol{z} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}, \sigma_p^2)$, which is why $\hat{\boldsymbol{\mu}}$ depends on sequence alone, the one quantity shared between training and sampling.

High-frequency roughness is supplied by a calibrated noise field: white noise, frozen along the ODE integration and independent per ensemble member, colored by the measured lag correlations of rod steps (Lavery et al., 2010) through a moving-average (MA) kernel. Two such fields $\boldsymbol{\eta}$ are drawn from separate kernels and enter the composition below. The rotational one is $\boldsymbol{\omega}_{\mathrm{th}} = (\boldsymbol{I} - \mathcal{P}_{<K})\,[\sigma_{\mathrm{th}} \odot \boldsymbol{\eta}]$, with $\odot$ the per-axis product and $\sigma_{\mathrm{th}} \in \mathbb{R}^3$ a true per-axis standard deviation of the per-step rotation in radians per step, trained against the reference rather than set by hand. The high-pass $\boldsymbol{I} - \mathcal{P}_{<K}$ places the field in the orthogonal complement of the bending subspace, leaving the band owned by the latent untouched. A separate translational field $\boldsymbol{\tau}_{\mathrm{th}}$ is needed because frame integration of the rotational field carries no translational information, which rise, slide and shift require; it is not high-passed, since the bending subspace is a property of the rotational field alone, and it is added to the rod trunk's $\boldsymbol{\tau}_{t=1}$ with a constant body-frame offset $\boldsymbol{\Delta}_\tau$, the mean step the end-to-end distance responds to. Both kernels, their orders, and the calibration of $\sigma_{\mathrm{th}}$ and $\boldsymbol{\Delta}_\tau$, are set out in Appendix A.11. Here, sequence enters through *width* rather than mean, a second convolutional head (`SeqLocalSigmaHead`) predicting the per-step $\log\sigma_{\mathrm{th}}$ profile as a normalized multiplicative shape along the chain.

At the terminal step $t = 1$ of the sampling ODE, both fields are composed additively in the Lie algebra of $\mathfrak{so}(3)$ into the step rotation emitted by the rod trunk,

$$\boldsymbol{\omega}_i^{\mathrm{step}} = \log\big(\boldsymbol{R}_i^{\top}\boldsymbol{R}_{i+1}\big)_{t=1}^{\vee} + \boldsymbol{\omega}_{z,i} + \boldsymbol{\omega}_{\mathrm{th},i}, \qquad \boldsymbol{\omega}_z = \mathcal{D}^{-1}\boldsymbol{C}_{<K}^{\top}\boldsymbol{z} - \mathcal{P}_{<K}\,\boldsymbol{\omega}_{t=1},$$

after which frames are re-integrated from the first, $\boldsymbol{R}_{i+1} = \boldsymbol{R}_i \exp(\boldsymbol{\omega}_i^{\mathrm{step}})$, and the centerline follows the step integration introduced in Section 3.1, $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{R}_{\mathrm{ref},i}(\boldsymbol{\tau}_i + \boldsymbol{\tau}_{\mathrm{th},i})$. Because $\boldsymbol{C}_{<K}^{\top}$

is the exact transpose of the projection $\boldsymbol{C}_{<K}$ and $\mathcal{D}$ is orthogonal, a prior draw re-expands at the correct physical amplitude, and subtracting $\mathcal{P}_{<K}\boldsymbol{\omega}_{t=1}$ leaves the band $k < K$ with a single owner, and since the field enters the step rotations, $k = 0$ is a coherent bend by construction. Composing in the Lie algebra rather than by right product is required, since every field is expressed in the local frame of the $i$-th step (Appendix A.7).

Our choices make sequence effects beyond the local scale accessible: the mean field $\Delta\boldsymbol{\omega}(\mathrm{seq})$ is a learned local pattern, but its projection onto the demodulated basis acts as a phase-sensitive detector. Since the approach mixes a bend amplitude of period $p$ into the sidebands $\Omega_0 \pm 2\pi/p$, the accumulation grows with $N$ when the modulation is phased with the helix and cancels otherwise (Appendix A.12). Hence sequence-directed curvature is amplified in the coordinate the latent owns, and phasing behavior enters as a low-$k$ shift of $\hat{\boldsymbol{\mu}}$. Every parameter of the decomposition, the per-step mean field, the prior $\sigma_p$, the thermal amplitudes, correlations and mean offset, is local and length-independent, so no global observable enters the objective: end-to-end distance, global bending and phasing sensitivity are *predictions* of the local law, not fitted quantities.

### 3.4 Training Setup

**Data.** The model is trained on a library of HCG-generated structures, energy-minimised under parmbsc1 (Ivani et al., 2016): 100 random sequences of 11–110 bp, 1000 conformations each, at 40 % GC content matching the human genome. Training uses the 11–80 bp subset with 100 sequences per length, validation split $90/10$ by sequence; all longer chains (81–110 bp) are never seen and serve as the held-out control for the extrapolation claim. Each sample is a *single* conformation encoded as a set of frames $\boldsymbol{T}$, so ensemble diversity is never supervised directly. A second, independent HexABC-derived set of 170 sequences with 544 decorrelated conformations each of central 16 bp segments from explicit-solvent molecular dynamics covers all 136 tetramers. Held out entirely, it parameterizes a rigid-base-pair *tetramer table*, the mean and covariance of the six step parameters per four-nucleotide context (Lankaš et al., 2003; Petkevičiūtė et al., 2014): the standard sequence-dependent elastic model of DNA and the reference for the A-tract phasing test.

**Objective.** All nine loss terms and their weights, the schedule and the remaining hyperparameters are given in Appendix A.13. The $\mathrm{SE}(3)$ conditional flow-matching terms are inherited unchanged from FrameFlow (Yim et al., 2024), and the additions are the rod flow-matching loss, a strand-to-centerline consistency term, the spectral mean loss $\mathcal{L}_\mu$, the per-step width regression, and the thermal calibration.

**Optimization.** The model has 2.45 M parameters (1.07 M rod trunk, 1.10 M backbone trunk, the remainder the latent cross-attention and the two sequence heads), and is trained with AdamW for 50 epochs on a single GPU. Each batch holds conformers of one sequence, so all members share a chain length and no padding is required.

## 4 Experiments

We now show that a correct local law yields the right global statistics, extrapolates past the lengths seen in training, and is sequence-specific.

### 4.1 Ensemble Quality

We assess generated ensembles over the $M$ members of one ensemble at a fixed sequence. The three *local* observables are the step-wise spreads of tilt, roll and twist averaged along the chain, the four *global* ones the across-member mean and standard deviation of the bend angle $\Theta$ between the terminal tangents and of the end-to-end distance $d_{\mathrm{e2e}}$. Carrying different units, each of the seven $o_k$ is scored as a dimensionless ratio to its reference and condensed into the mean absolute log-ratio $D_7 = \frac{1}{7}\sum_{k=1}^{7}|\log(\hat{o}_k/o_k)|$, with $\hat{o}_k$ the generated and $o_k$ the reference value, symmetric in over- and under-dispersion and zero at exact agreement (Appendix A.14).

Figure 2A shows those components for one held-out 50 bp sequence, $M = 1000$ generated members against its 1000 reference conformations. Beyond the mean and spread statistics, the full distributions of bend angle and end-to-end distance follow the reference closely, and per-base-pair positional fluctuations agree within their confidence intervals (Figure 2B–D). Projected onto two principal

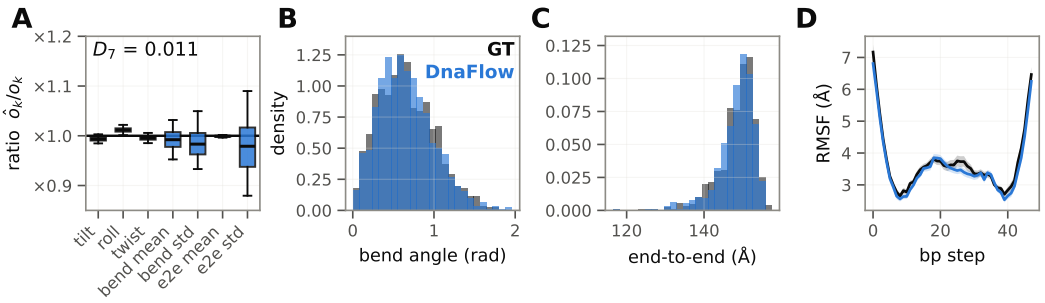



Figure 2: **Free-ensemble statistics against the reference ensemble.** One held-out 50 bp sequence, 1000 generated samples (blue) against its hierarchical chain-growth reference (ground truth GT, black). **(A)** Generated-to-reference ratio $\hat{o}_k/o_k$ for seven observables entering $D_7$, against the line of exact agreement. **(B,C)** Distributions of the global bend angle and the end-to-end distance. **(D)** Per-position root-mean-square fluctuation after rigid body alignment. In A the boxes span the interquartile range of 400 bootstrap resamples of both ensembles and the whiskers, like the bands in D, the 95 % interval.

components of the held-out reference, the generated samples disperse over the reference distribution and reach its extreme conformations, so the samples are in distribution (Appendix Figure 6). However, correct centerline statistics do not imply chemically sound atomic structures. Hence, all-atom structures are scored separately, on steric overlap, pairing geometry and the sugar substates, through the identical rebuild (Appendix A.15).

### 4.2 Extrapolation to Unseen Lengths

Extrapolation is assessed at the held-out lengths of the random sequence dataset (Table 1). Inside the training range, $D_7$ stays between 0.036 and 0.058, with every ratio within 10 % of the reference except the end-to-end spread. Past it the score holds at 90 and 100 bp and then degrades at 110 bp, where roll and twist over-disperse together. The global observables hold even there. The end-to-end spread scatters between 0.90 and 0.98 with no trend in length, the term the sequences resolve worst (Figure 2A). Re-scoring the same draws restricted to the reference-covered region leaves the score at 110 bp essentially unchanged, 0.071 to 0.070: the rotational over-dispersion does sit in a tail of out-of-distribution members, roll falling from 1.13 to 1.08 and twist from 1.09 to 1.06, but the global terms move the other way, so the aggregate does not improve (Appendix A.17). This tail is what grows with length: the fraction of members inside the covered region falls from 98 % in the training range to 95, 92 and 68 % at 90, 100 and 110 bp, so the extrapolation is visible in the coverage well before it is visible in a moment.

| $N$ | seqs | tilt | roll | twist | $\text{bend}_{\text{mean}}$ | $\text{bend}_{\text{std}}$ | $\text{e2e}_{\text{mean}}$ | $\text{e2e}_{\text{std}}$ | $D_7 \pm \text{se}$ |
|---|---|---|---|---|---|---|---|---|---|
| 20 | 5 | 1.10 | 0.97 | 1.02 | 1.01 | 1.03 | 1.00 | 0.92 | $0.049 \pm 0.009$ |
| 40 | 5 | 1.03 | 0.97 | 1.00 | 0.96 | 0.96 | 1.00 | 0.92 | $0.058 \pm 0.011$ |
| 60 | 5 | 0.97 | 1.02 | 1.00 | 1.00 | 1.01 | 0.99 | 0.98 | $0.036 \pm 0.003$ |
| 80 | 5 | 0.94 | 1.03 | 0.99 | 1.01 | 0.99 | 1.00 | 0.90 | $0.051 \pm 0.006$ |
| 90† | 5 | 0.94 | 1.05 | 1.03 | 0.98 | 1.00 | 0.99 | 0.95 | $0.046 \pm 0.004$ |
| 100† | 5 | 0.95 | 1.08 | 1.02 | 0.99 | 1.00 | 0.99 | 0.97 | $0.057 \pm 0.010$ |
| 110† | 5 | 0.99 | 1.13 | 1.09 | 0.93 | 0.97 | 0.97 | 0.91 | $0.071 \pm 0.006$ |

Table 1: **Generated ensembles versus the reference across duplex length.** Each length is scored on 1000 members, split evenly over five held-out sequences. The seven columns are the terms $D_7$ averages, each a ratio of the generated statistic to the hierarchical chain-growth reference, so 1.00 is perfect agreement, and $D_7$ is the ensemble discrepancy, lower being better, with its standard error over the five. Columns and score are combined over sequences differently, so $D_7$ is not recoverable from the row (Appendix A.14). †Held-out length, outside the 11–80 bp training range.

### 4.3 SEQUENCE DEPENDENCE

Specificity is assessed on the generated ensembles, not on the model's sequence-dependent heads. We condition on $S = 40$ held-out 90 bp sequences, a length outside the training range so that every candidate is unseen by construction, draw $M = 64$ members each through the production path, and re-measure the six step parameters with the estimators used on the reference. The object every panel compares is the ensemble-mean *step field* of one sequence $s$, $\hat{\boldsymbol{f}}_i^{(s)} = \mathrm{mean}_m\big[\boldsymbol{p}_i^{(m)}\big]$ with $\boldsymbol{p}_i^{(m)} = (\boldsymbol{\omega}_i, \boldsymbol{\tau}_i)^{(m)} \in \mathbb{R}^6$ the step parameters of member $m$ at step $i$, taken over the chain's steps with the two outermost at each end dropped. $\boldsymbol{f}^{(s)}$ is the same quantity measured on that sequence's reference ensemble, and $\bar{\boldsymbol{f}} = S^{-1}\sum_s \boldsymbol{f}^{(s)}$ is the *sequence-blind* field that every sequence shares. Two read-outs follow, because the first saturates.

*Identification* (Figure 3A) matches each generated field against the reference field of all $S$ candidates under a squared distance $D(s, s')$ in which every component is divided by the reference's own sampling error, so components carrying no signal cannot dominate through their units (Appendix A.18). Every ensemble is closest to the sequence it was conditioned on, $\arg\min_{s'} D(s, s') = s$ in all 40 cases against a chance rate of 0.025. *Field agreement* (Figure 3B) is the paired coefficient of determination taken against the sequence-blind field, $R^2 = 1 - \sum_s \|\hat{\boldsymbol{f}}^{(s)} - \boldsymbol{f}^{(s)}\|^2 \big/ \sum_s \|\boldsymbol{f}^{(s)} - \bar{\boldsymbol{f}}\|^2$, so that $R^2 = 0$ is no better than predicting every sequence by the average field and 1 is exact. The sequence-dependent latent achieves a correlation of $+0.873$ when evaluated over the full spatial field, and $+0.670$ when both the predictions and targets are restricted to the $K = 8$ spatial Fourier modes owned by the latent. This performance exceeds the rigid-base-pair tetramer baseline, which reaches $+0.823$ on the reference data itself (with a split-half ceiling of $+0.988$) and $+0.783$ when evaluated on samples generated through the identical DnaFlow sampling pipeline (Appendix A.18). The effect is not confined to the mean: the same identification run on the per-step $\log\sigma$ profile picks its own sequence in all 40 cases.

Base order carries the specificity, not composition. Permuting the conditioning sequence at fixed composition prevents identification, 3 of 40, and drives $R^2$ to $-0.831$, below the zero the sequence-blind field alone would give. Resolved per step, Figure 3C draws the roll component of that field for one representative sequence as a deviation from the field all $S$ share, and scores every sequence according to the per-step profile $r^{(s)}$, i.e., the correlation of the generated and reference roll profiles along the chain, each centered on its own chain mean. Thus, $r$ reads the shape of the profile and not its offset. The generated field follows the reference within its sampling error where the permutation does not, at $r = +0.872 \pm 0.004$ against $+0.010 \pm 0.020$ over the 40 sequences. The same comparison for the other five components is shown in Appendix Figure 8.

Agreement of the local fields does not fix what the steps add up to: the bend a field accumulates over many turns is determined by the helical phase which carries each local bend into the global frame. Whether the model's per-step field adds over many turns rather than canceling is what the spectral division is built to deliver. We assessed it on phased A-tract constructs, eight $A_5$ tracts at center-to-center period $P$ over 8–13 bp against purine-matched controls with the run interrupted, giving chains of 61–96 bp whose long-period end lies past the training range, and read out the *excess bend* $\Theta_{\mathrm{ex}}(P)$ (Figure 3D). Against a tetramer table fitted to independent explicit-solvent trajectories the model never saw, which has a peak excess bend of $47.6^\circ$ on its own helical repeat $p_h = 10.43$ bp, the bend for the full sampler delivers $31.2 \pm 1.5^\circ$, peaks at $10.12 \pm 0.18$ bp against its own $p_h = 10.20$ bp, and tracks the independent law per step at $r = +0.77$. Falling short of the full amplitude is expected, since the model can only carry a wedge its data contains: the same table fitted to the HCG training library has a peak bend of $28.8^\circ$, 60 % of the independent reference, and the model reaches 108 % of that. The per-step field correlation highlights the difference between the models and controls profile, $+0.77$ against $-0.24$ (Appendix A.19). Nonetheless, it is evident that the local law accumulates in phase over roughly seven helical turns of a construct unlike anything in the training library, though neither the phasing response nor the sequence class enters the objective.

## 5 RELATED WORK

Normalizing flows have been employed to model the statistical fluctuation of centerline base pair steps, derived from MD simulations of sequences containing all possible dinucleotide steps (Liebl,

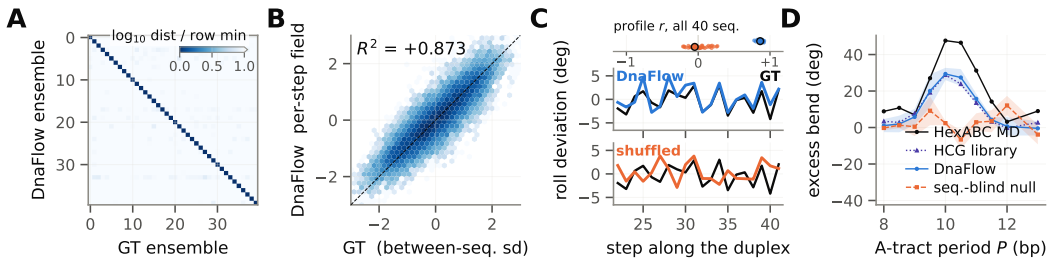



Figure 3: **Sequence dependence of the generated ensemble.** All panels use one checkpoint, sequences absent from training, and the measurements of Section 4.3; panel details in Appendix A.18. **(A)** Cross-identification: rows are generated ensembles, columns their reference ensembles. A dark diagonal indicates every ensemble is closest to its own sequence. **(B)** The same fields resolved per step, in units of the between-sequence standard deviation (dashed line, identity). **(C)** Ensemble-mean roll of one representative held-out sequence over 20 steps, as a deviation from the field all 40 share: the reference (black) against the generated (blue, above) and permutation (orange, below) arms, with a strip carrying both arms' profile $r$ for all 40. **(D)** Phased A-tract excess bend versus period $P$: the independent HexABC reference (black), the same tetramer law fitted to the HCG training library (purple), the model (blue), and the model with its base-identity embedding re-drawn (orange), bands spanning three sampling seeds.

2026), where the invertible transformation from noise to the distribution of the six helical parameters allows energies to be evaluated through Boltzmann inversion. This model relies on a sparse dataset, three sequences with 100 nanoseconds each, which only allows for fitting next-neighbor effects, hence global ensemble properties are not discussed and no atomistic structures obtained. Deep DNAshape (Li et al., 2024) incorporates larger context with flanking regions into predictions of DNA shape and has been demonstrated to predict transcription factor DNA interactions, but does not scale beyond local context and does not produce atomistic structures. AlphaFold3 predicts atomistic structures of DNA and DNA-bound complexes, but has not been trained to model larger sequences and their behavior in solution accurately (Abramson et al., 2024) (Figure 6). Modern approaches of learning the distribution of molecular structures, such as BioEmu (Lewis et al., 2025) and Boltzmann generators (Schebek et al., 2026) have so far only been applied to proteins and systems relevant for material sciences.

## 6 Conclusion

We presented DnaFlow, a lightweight flow-matching model that builds both strands around a simultaneously generated centerline rod whose fluctuations are organized by spectral division. Every global observable we report is a prediction, not a fit: training sees one conformation at a time and no global statistic enters the objective. Held-out sequences within the training lengths are reproduced to a few percent in every observable. Specificity enters through base order rather than composition, each ensemble identifying its own sequence and tracking its per-step mean field. Being local and length-independent, the law also accumulates coherently on sequences unlike any in the training library: phased A-tract constructs recover the excess bend its training library contains, indicating that the learned geometry transfers across data modalities. While the force field, and ion concentrations are identical between our GT, and the HexABC data, some of the discrepancy may be due to the difference in ion parameters, and the water model employed. The limits are equally clear. What drift there is past the training range sits in the per-step rotational widths while every global observable holds, and screening out-of-distribution members recovers some of it, so what grows with length is a tail rather than a shifted ensemble. Our references are themselves force-field and sampling limited, so we claim not the true Boltzmann ensemble but that the architecture faithfully learns the ensemble it is shown. Neither limit is intrinsic to the paradigm, and the recipe is not specific to DNA: decompose the local deformation field by scale and give each band exactly one owner. We expect it to apply wherever atomistic structures follow a reasonable coarse-grained model, and for DNA it already turns a combinatorial sampling problem into a single training run, yielding sequence-specific, length-transferable ensembles at atomistic resolution in seconds.

### AI USE STATEMENT

In this work, we used generative AI tools for optimizing written code and generating new modules. The theoretical concepts of this manuscript were formulated and conceptualized by us and thereafter challenged by a LLM. The draft of this manuscript has been written by the authors and only afterwards been polished using AI to improve the readability and discuss which concepts may be unclear to an unfamiliar reader. Everything generated by an AI was double checked, and we take responsibility for the final content of this work.

### REPRODUCIBILITY STATEMENT

We provide the models code, and evaluation scripts used to reproduce the contents of this manuscript. The code to generate the required dataset will be published elsewhere. Therefore, these results can easily be reproduced by researchers.

## REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, Sebastian W Bodenstein, David A Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė, Eirini Arvaniti, Charles Beattie, Ottavia Bertolli, Alex Bridgland, Alexey Cherepanov, Miles Congreve, Alexander I Cowen-Rivers, Andrew Cowie, Michael Figurnov, Fabian B Fuchs, Hannah Gladman, Rishub Jain, Yousuf A Khan, Caroline M R Low, Kuba Perlin, Anna Potapenko, Pascal Savy, Sukhdeep Singh, Adrian Stecula, Ashok Thillaisundaram, Catherine Tong, Sergei Yakneen, Ellen D Zhong, Michal Zielinski, Augustin Žídek, Victor Bapst, Pushmeet Kohli, Max Jaderberg, Demis Hassabis, and John M Jumper. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016):493–500, June 2024.

Ariel Afek, Honglue Shi, Atul Kaushik Rangadurai, Harshit Sahay, Alon Senitzki, Suela Xhani, Mimi Fang, Raul Salinas, Zachery Mielko, Miles A. Pufall, Gregory M. K. Poon, Tali E. Haran, Maria A. Schumacher, Hashim M. Al-Hashimi, and Raluca Gordân. DNA mismatches reveal conformational penalties in protein-DNA recognition. *Nature*, 587:291 – 296, 2020.

Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, Claudia Millán, Hahnbeom Park, Carson Adams, Caleb R Glassman, Andy DeGiovanni, Jose H Pereira, Andria V Rodrigues, Alberdina A van Dijk, Ana C Ebrecht, Diederik J Opperman, Theo Sagmeister, Christoph Buhlheller, Tea Pavkov-Keller, Manoj K Rathinaswamy, Udit Dalwadi, Calvin K Yip, John E Burke, K Christopher Garcia, Nick V Grishin, Paul D Adams, Randy J Read, and David Baker. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, August 2021.

Minkyung Baek, Ryan McHugh, Ivan Anishchenko, Hanlun Jiang, David Baker, and Frank DiMaio. Accurate prediction of protein-nucleic acid complexes using RoseTTAFoldNA. *Nat. Methods*, 21 (1):117–121, January 2024.

Jörg Baschnagel, Hendrik Meyer, Joachim Wittmer, Igor Kulić, Hervé Mohrbach, Falko Ziebert, Gi-Moon Nam, Nam-Kyung Lee, and Albert Johner. Semiflexible chains at surfaces: Worm-like chains and beyond. *Polymers (Basel)*, 8(8):E286, August 2016.

Federica Battistini, Miłosz Wieczór, Adam Hospital, Marco Pasi, Juan Pablo Arcón, Israel Serrano-Chacón, Alba Sala, Subhamoy Deb, Agustín García-Doñate, Matthew Burman, Elliot W Chan, Liwei Chang, Gabriela da Rosa, Jorge R Espinosa, Gia Linh Hoang, Kazi A Hossain, Michał Jurkowski, Romain Poupon, Rahul Sharma, Ran Sun, Thomas C Bishop, Paolo Carloni, Thomas E Cheatham, 3rd, Rosana Collepardo-Guevara, Jacek Czub, Pablo D Dans, Sarah A Harris, Charles Laughton, Rodrigo Galindo-Murillo, John H Maddocks, Agnes Noy, Alberto Pérez, Daiva Petkevičiūtė-Gerlach, and Modesto Orozco. hexABC seeking the physical code of DNA. *Nat. Commun.*, June 2026.

Nils B. Becker and Ralf Everaers. From rigid base pairs to semiflexible polymers: Coarse-graining DNA. *Phys. Rev. E*, 76:021923, Aug 2007. doi: 10.1103/PhysRevE.76.021923.

Carlos Bustamante, Zev Bryant, and Steven B Smith. Ten years of tension: single-molecule DNA mechanics. *Nature*, 421(6921):423–427, January 2003.

Ricky T. Q. Chen and Yaron Lipman. Flow matching on general geometries. In *The Twelfth International Conference on Learning Representations*, 2024.

Vincent B. Chen, W. Bryan Arendall, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson, and David C. Richardson. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D*, 66(1):12–21, 2010. doi: 10.1107/S0907444909042073.

Pablo D Dans, Alexandra Balaceanu, Marco Pasi, Alessandro S Patelli, Daiva Petkevičiūtė, Jürgen Walther, Adam Hospital, Genís Bayarri, Richard Lavery, John H Maddocks, and Modesto Orozco. The static and dynamic structural heterogeneities of B-DNA: extending Calladine-Dickerson rules. *Nucleic Acids Res.*, 47(21):11090–11102, December 2019.

M A El Hassan and C R Calladine. Conformational characteristics of DNA: empirical classifications and a hypothesis for the conformational behaviour of dinucleotide steps. *Philos. Trans. A Math. Phys. Eng. Sci.*, 355(1722):43–100, January 1997.

Tania Gardasevic and Agnes Noy. The impact of sequence periodicity on DNA mechanics: investigating the origin of a-tract's curvature. *Nanoscale*, 16(39):18410–18420, October 2024.

Ivan Ivani, Pablo D. Dans, Agnes Noy, Alberto Pérez, Ignacio Faustino, Adam Hospital, Jürgen Walther, Pau Andrio, Ramon Goñi, Alexandra Balaceanu, Guillem Portella, Federica Battistini, Josep Lluis Gelpí, Carlos González, Michele Vendruscolo, Charles A. Laughton, Sarah A. Harris, David A. Case, and Modesto Orozco. Parmbsc1: a refined force field for DNA simulations. *Nature Methods*, 13(1):55–58, 2016. doi: 10.1038/nmeth.3690.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A A Kohl, Andrew J Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, August 2021.

Jonas Köhler, Leon Klein, and Frank Noe. Equivariant flows: Exact likelihood generative learning for symmetric densities. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 5361–5370. PMLR, 13–18 Jul 2020.

O. Kratky and G. Porod. Röntgenuntersuchung gelöster fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas*, 68(12):1106–1122, 1949. doi: https://doi.org/10.1002/recl.19490681203.

Filip Lankaš, Jiří Šponer, Jörg Langowski, and Thomas E. Cheatham. DNA basepair step deformability inferred from molecular dynamics simulations. *Biophysical Journal*, 85(5):2872–2883, 2003. ISSN 0006-3495. doi: https://doi.org/10.1016/S0006-3495(03)74710-9.

Richard Lavery, Krystyna Zakrzewska, David Beveridge, Thomas C. Bishop, David A. Case, III Cheatham, Thomas, Surjit Dixit, B. Jayaram, Filip Lankas, Charles Laughton, John H. Maddocks, Alexis Michon, Roman Osman, Modesto Orozco, Alberto Perez, Tanya Singh, Nada Spackova, and Jiri Sponer. A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Research*, 38(1): 299–313, 01 2010. ISSN 0305-1048. doi: 10.1093/nar/gkp834.

Sarah Lewis, Tim Hempel, José Jiménez-Luna, Michael Gastegger, Yu Xie, Andrew Y. K. Foong, Victor García Satorras, Osama Abdin, Bastiaan S. Veeling, Iryna Zaporozhets, Yaoyi Chen, Soojung Yang, Adam E. Foster, Arne Schneuing, Jigyasa Nigam, Federico Barbero, Vincent Stimper, Andrew Campbell, Jason Yim, Marten Lienen, Yu Shi, Shuxin Zheng, Hannes Schulz, Usman Munir, Roberto Sordillo, Ryota Tomioka, Cecilia Clementi, and Frank Noé. Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science*, 389(6761):eadv9817, 2025. doi: 10.1126/science.adv9817.

Jinsen Li, Tsu-Pei Chiu, and Remo Rohs. Predicting DNA structure using a deep learning method. *Nat. Commun.*, 15(1):1243, February 2024.

Wentian Li, Thomas G. Marr, and Kunihiko Kaneko. Understanding long-range correlations in DNA sequences. *Physica D: Nonlinear Phenomena*, 75(1):392–416, 1994. ISSN 0167-2789.

Korbinian Liebl. Deciphering DNA's sequence-dependent structure and deformability with normalizing flows. *J. Chem. Theory Comput.*, 22(1):456–468, January 2026.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan Dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, March 2023.

Xiang-Jun Lu and Wilma K Olson. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, 31(17):5108–5121, September 2003.

John F. Marko and Eric D. Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, 05 2002. ISSN 0024-9297. doi: 10.1021/ma00130a008.

Atsushi Matsumoto and Wilma K. Olson. Sequence-dependent motions of DNA: A normal mode analysis at the base-pair level. *Biophysical Journal*, 83(1):22–41, 2002. ISSN 0006-3495. doi: https://doi.org/10.1016/S0006-3495(02)75147-3.

Alexandre V. Morozov, Karissa Fortney, Daria A. Gaykalova, Vasily M. Studitsky, Jonathan Widom, and Eric D. Siggia. Using DNA mechanics to predict in vitro nucleosome positions and formation energies. *Nucleic Acids Research*, 37:4707 – 4722, 2009.

W K Olson, M Bansal, S K Burley, R E Dickerson, M Gerstein, S C Harvey, U Heinemann, X J Lu, S Neidle, Z Shakked, H Sklenar, M Suzuki, C S Tung, E Westhof, C Wolberger, and H M Berman. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, 313(1):229–237, October 2001.

Christopher E Pearson and Richard R Sinden. Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Current Opinion in Structural Biology*, 8(3):321–330, 1998. ISSN 0959-440X. doi: https://doi.org/10.1016/S0959-440X(98)80065-1.

D. Petkevičiūtė, M. Pasi, O. Gonzalez, and J.H. Maddocks. cgDNA: a software package for the prediction of sequence-dependent coarse-grain free energies of b-form DNA. *Nucleic Acids Research*, 42(20):e153–e153, 11 2014. ISSN 0305-1048. doi: 10.1093/nar/gku825.

Ekaterina Pilyugina, Brad Krajina, Andrew J Spakowitz, and Jay D Schieber. Buckling a semiflexible polymer chain under compression. *Polymers (Basel)*, 9(3):E99, March 2017.

Mohammadhosein Razbin and Panayotis Benetatos. Elasticity of a grafted rod-like filament with fluctuating bending stiffness. *Polymers (Basel)*, 15(10):2307, May 2023.

Randy J. Read, Paul D. Adams, W. Bryan Arendall, Axel T. Brunger, Paul Emsley, Robbie P. Joosten, Gerard J. Kleywegt, Eugene B. Krissinel, Thomas Lütteke, Zbyszek Otwinowski, Anastassis Perrakis, Jane S. Richardson, William H. Sheffler, Janet L. Smith, Ian J. Tickle, Gert Vriend, and Peter H. Zwart. A new generation of crystallographic validation tools for the protein data bank. *Structure*, 19(10):1395–1412, 2011. ISSN 0969-2126. doi: https://doi.org/10.1016/j.str.2011.08.006.

Matthias Rief, Hauke Clausen-Schaumann, and Hermann Eduard Gaub. Sequence-dependent mechanics of single DNA molecules. *Nature Structural Biology*, 6:346–349, 1999.

Maximilian Schebek, Frank Noé, and Jutta Rogal. Scalable boltzmann generators for equilibrium sampling of large-scale materials. *Nat. Commun.*, 17(1), June 2026.

Tao Shen, Zhihang Hu, Siqi Sun, Di Liu, Felix Wong, Jiuming Wang, Jiayang Chen, Yixuan Wang, Liang Hong, Jin Xiao, Liangzhen Zheng, Tejas Krishnamoorthi, Irwin King, Sheng Wang, Peng Yin, James J Collins, and Yu Li. Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nat. Methods*, 21(12):2287–2298, December 2024.

Enrico Skoruppa and Helmut Schiessel. Systematic coarse-graining of sequence-dependent structure and elasticity of double-stranded DNA. *Phys. Rev. Res.*, 7:013044, Jan 2025. doi: 10.1103/PhysRevResearch.7.013044.

Jianlin Su, Yu Lu, Shengfeng Pan, Ahmed Murtadha, Bo Wen, and Yunfeng Liu. Roformer: Enhanced transformer with rotary position embedding, 2023.

Sumit Tarafder and Debswapna Bhattacharya. RNAbpFlow: base pair-augmented SE(3) flow matching for conditional RNA 3D structure generation. *Nat. Methods*, 23(7):1349–1358, July 2026.

A. A. Travers. The structural basis of DNA flexibility. *Philosophical Transactions of the Royal Society of London, Series A: Physical and Engineering Sciences*, 362(1820):1423–1438, 07 2004.

Mark E. Tuckerman and Glenn J. Martyna. Understanding modern molecular dynamics: techniques and applications. *The Journal of Physical Chemistry B*, 104(2):159–178, 12 1999. ISSN 1520-6106. doi: 10.1021/jp992433y.

Jonathan Widom. Role of DNA sequence in nucleosome stability and dynamics. *Quarterly Reviews of Biophysics*, 34:269 – 324, 2001.

J. Michael Word, Simon C. Lovell, Thomas H. LaBean, Hope C. Taylor, Michael E. Zalis, Brent K. Presley, Jane S. Richardson, and David C. Richardson. Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *Journal of Molecular Biology*, 285(4):1711–1733, 1999. doi: 10.1006/jmbi.1998.2400.

Jason Yim, Andrew Campbell, Andrew Y. K. Foong, Michael Gastegger, José Jiménez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S. Veeling, Regina Barzilay, Tommi Jaakkola, and Frank Noé. Fast protein backbone generation with se(3) flow matching, 2023a.

Jason Yim, Brian L. Trippe, Valentin De Bortoli, Emile Mathieu, Arnaud Doucet, Regina Barzilay, and Tommi Jaakkola. Se(3) diffusion model with application to protein backbone generation, 2023b.

Jason Yim, Andrew Campbell, Emile Mathieu, Andrew Y. K. Foong, Michael Gastegger, Jose Jimenez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S. Veeling, Frank Noe, Regina Barzilay, and Tommi Jaakkola. Improved motif-scaffolding with SE(3) flow matching. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856.

Petr Šulc, Flavio Romano, Thomas E. Ouldridge, Lorenzo Rovigatti, Jonathan P. K. Doye, and Ard A. Louis. Sequence-dependent thermodynamics of a coarse-grained DNA model. *The Journal of Chemical Physics*, 137(13):135101, 10 2012. ISSN 0021-9606. doi: 10.1063/1.4754132.

# A Appendix

## A.1 SE(3) Metric and Reparameterized Flow-Matching Loss

The metric on SE(3) referred to in Section 2.1 is

$$\langle (u, v), (u', v') \rangle_{\mathrm{SE}(3)} = \langle u, u' \rangle_{\mathrm{SO}(3)} + \langle v, v' \rangle_{\mathbb{R}^3},$$

built from the canonical metrics on $\mathrm{SO}(3)$ and $\mathbb{R}^3$ (Yim et al., 2023b), which is what lets rotations and translations be treated independently, and $\exp_T$ and $\log_T$ are available analytically on both factors. After reparameterization the RFM loss depends only on the model's predictions and the sample they were conditioned on, and is a $1/(1-t)^2$-reweighted squared error on translations and rotations in the tangent space at $\boldsymbol{R}_t$, over the $N$ frames corrupted at time $t$. In full it is

$$\mathcal{L}_{\mathrm{SE}(3)} = \mathbb{E}_t, p_0, q \left[ \frac{1}{(1-t)^2} \sum_{n=1}^{N} \left\{ \left\| \hat{\boldsymbol{x}}_1^n(\boldsymbol{T}_t, t) - \boldsymbol{x}_1^n \right\|_{\mathbb{R}^3}^2 + \left\| \log_{\boldsymbol{R}_t^n} \hat{\boldsymbol{R}}_1^n(\boldsymbol{T}_t, t) - \log_{\boldsymbol{R}_t^n} \boldsymbol{R}_1^n \right\|_{\mathrm{SO}(3)}^2 \right\} \right]$$

where $\boldsymbol{T}_t := [T^1, \ldots, T^N] \in \mathrm{SE}(3)^N$ is the collection of $N$ frames corrupted at time $t$, $\boldsymbol{T}_0 \sim p_0$, $\boldsymbol{T}_1 \sim q$, and $t \sim \mathcal{U}([0, 1))$ (Yim et al., 2023a).

## A.2 Model Architecture

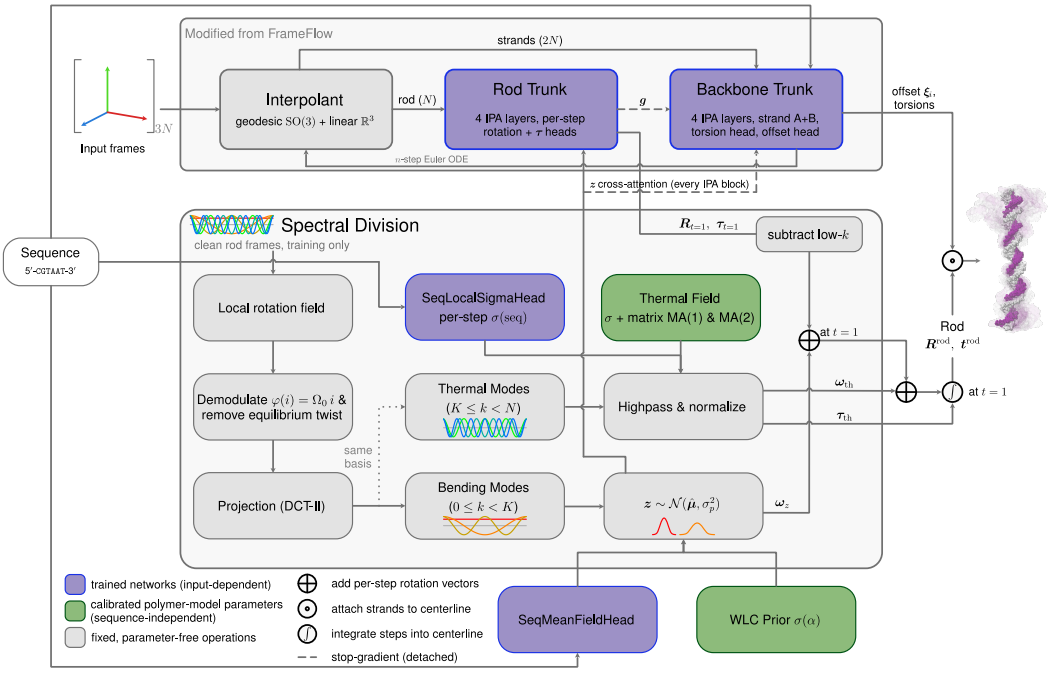



Figure 4: **Schema of DnaFlow's architecture.** Input frames are corrupted by geodesic $\mathrm{SO}(3)$ and linear $\mathbb{R}^3$ interpolation and denoised by FrameFlow-derived modules: a rod trunk over the $N$ centerline steps and a backbone trunk over the $2N$ nucleobase frames, which emits the per-base offsets $\boldsymbol{\xi}_i$ and the torsions. Both trunks cross-attend to the spectral latent $\boldsymbol{z}$, and the backbone additionally reads the detached rod embedding $\boldsymbol{g}$. Sections 3.1–3.3 describe the components, the latent's low-$k$ and thermal high-$k$ bands, and the terminal $t = 1$ composition that refines the centerline before the strands are built on it.

## A.3 Base-pair and Step Frame Conventions

Base pair frames along the centerline of a DNA double helix are defined according to the convention published by Olson et al. (2001). Individual bases are rigid-body fitted to the standardized coordinates using least squares. The $x$ axis is consequently orthogonal to the bisector of the $\mathrm{C1}' \cdots \mathrm{C1}'$

axis, the $y$-axis parallel to it, and $z$ is defined as their cross product $z = x \times y$ pointing in the $5' \to 3'$ direction along the helical axis. The base pair *step* frame of the helix centerline is then obtained by flipping the local frame of the base on the antiparallel strand, and then calculating the geodesic midpoint between the two frames. This resulting frame is a symmetric base pair step frame, which is therefore independent of the arbitrary selection of reference base. The six parameters of a *step* are in turn referred to the geodesic midpoint of the two consecutive base-pair frames it joins, the CEHS/3DNA construction (El Hassan & Calladine, 1997; Lu & Olson, 2003), which is what makes tilt, roll and twist symmetric in the two base pairs rather than biased toward either of them. Figure 5 draws all six in this convention.

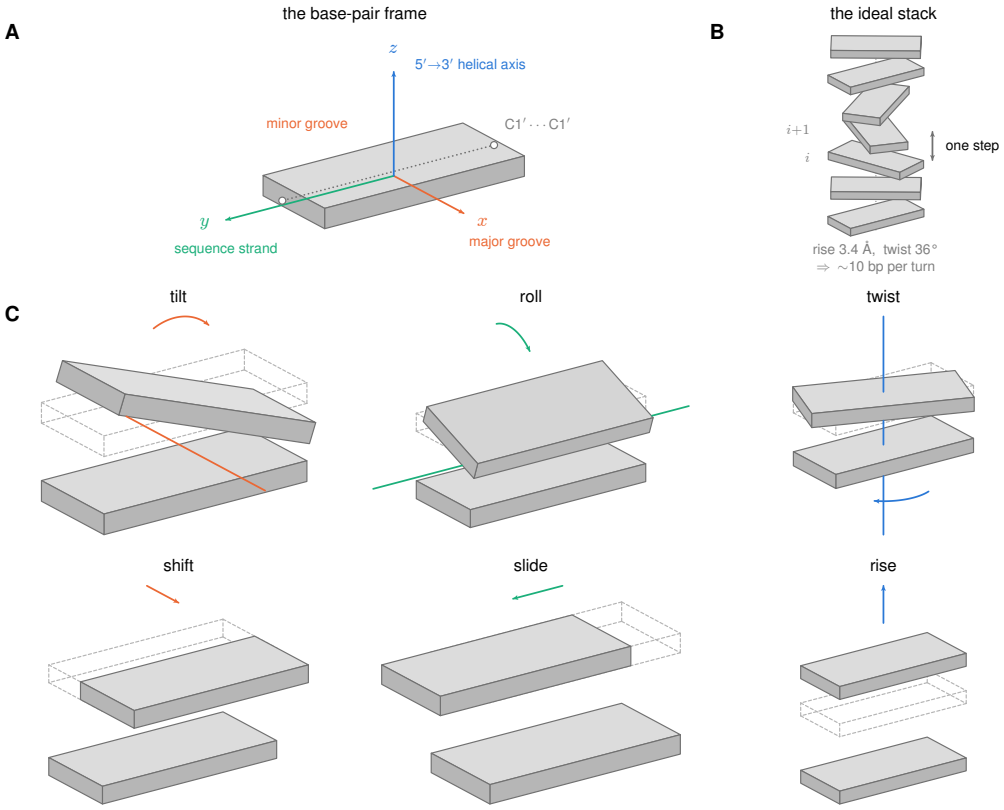



Figure 5: **The six rigid-base-pair step parameters.** A schematic, in the convention used throughout this work; one colour per body axis. **(A)** One base pair, drawn as a slab, and the body frame the six are measured in: $x$ is the pseudo-dyad, pointing toward the major groove, $y$ the long axis, parallel to the $\mathrm{C1}' \cdots \mathrm{C1}'$ vector and pointing toward the sequence strand, and $z = x \times y$ the $5' \to 3'$ helical direction. **(B)** The ideal B-DNA stack that the six parameters are deviations from, and the two base pairs $i, i+1$ that one step $(\boldsymbol{\omega}_i, \boldsymbol{\tau}_i)$ joins. **(C)** Each parameter on its own: the three rotations $\boldsymbol{\omega}_i$, tilt, roll and twist about $x$, $y$ and $z$, on the upper row, and the three translations $\boldsymbol{\tau}_i$, shift, slide and rise along $x$, $y$ and $z$, below.

### A.4 Local Frame-aware Propagation

In the local backbone propagation of Section 3.1, each frame $i$ considers its neighborhood $j$ up to a window size $W$. Frame $i$'s node embedding $\boldsymbol{h}_i \in \mathbb{R}^{c_s}$, a vector of $c_s$ learned scalar channels, is updated by expressing the displacement of $j$ in $i$'s local frame:

$$\boldsymbol{h}_i' = \boldsymbol{h}_i + f_\theta\Big(\boldsymbol{h}_i, \bigoplus_{j=i\pm1,\ldots,i\pm W} \big[\boldsymbol{R}_i^\top(\boldsymbol{x}_j - \boldsymbol{x}_i), \log(\boldsymbol{R}_i^\top \boldsymbol{R}_j)^\vee\big]\Big),$$

where $f_\theta$ is an MLP whose weights are part of $\theta$, and relative rotations are parameterized as rotation vectors in the Lie algebra $\mathfrak{so}(3)$, as denoted by the logarithmic map and vee operator $\log(\cdot)^\vee$ map-

ping a rotation matrix $\boldsymbol{R} \in \mathrm{SO}(3)$ to a rotation vector $\boldsymbol{\omega} \in \mathfrak{so}(3) \simeq \mathbb{R}^3$, the axis of rotation scaled by its angle in radians. The concatenation $\oplus$ deliberately imposes an ordering and is therefore not permutation invariant, unlike the regular message passing paradigm.

### A.5 Proof of $\mathrm{SE}(3)$ Invariance

The $\mathrm{SE}(3)$-invariance of the proposed `LocalFrameProp` from Section A.4 is a result of constructing its geometric features in the local reference frame of residue $i$. Formally, define $\boldsymbol{d}_{ij} := \boldsymbol{R}_i^\top(\boldsymbol{x}_j - \boldsymbol{x}_i)$ the displacement of frame $T_j$ in $i$'s local coordinate system. Consider a rigid transformation $(\boldsymbol{Q}, \boldsymbol{r}) = \rho(g), g \in \mathrm{SE}(3)$ acting on frames according to $\boldsymbol{x}_i' = \boldsymbol{Q}\boldsymbol{x}_i + \boldsymbol{r}$ and $\boldsymbol{R}_i' = \boldsymbol{Q}\boldsymbol{R}_i$ the transformed rotation, it follows:

$$\begin{aligned}
\boldsymbol{d}_{ij}' &= (\boldsymbol{R}_i')^\top(\boldsymbol{x}_j' - \boldsymbol{x}_i') \\
&= (\boldsymbol{Q}\boldsymbol{R}_i)^\top[(\boldsymbol{Q}\boldsymbol{x}_j + \boldsymbol{r}) - (\boldsymbol{Q}\boldsymbol{x}_i + \boldsymbol{r})] \\
&= \boldsymbol{R}_i^\top \boldsymbol{Q}^\top \boldsymbol{Q}(\boldsymbol{x}_j - \boldsymbol{x}_i) \\
&= \boldsymbol{R}_i^\top(\boldsymbol{x}_j - \boldsymbol{x}_i).
\end{aligned}$$

Therefore, $\boldsymbol{R}_i^\top(\boldsymbol{x}_j - \boldsymbol{x}_i)$ is $\mathrm{SE}(3)$-invariant. This holds true for $\log(\boldsymbol{R}_i^\top \boldsymbol{R}_j)^\vee$ as well, since:

$$\begin{aligned}
\log((\boldsymbol{R}_i')^\top \boldsymbol{R}_j')^\vee &= \log((\boldsymbol{Q}\boldsymbol{R}_i)^\top \boldsymbol{Q}\boldsymbol{R}_j)^\vee \\
&= \log(\boldsymbol{R}_i^\top \boldsymbol{Q}_i^\top \boldsymbol{Q}\boldsymbol{R}_j)^\vee \\
&= \log(\boldsymbol{R}_i^\top \boldsymbol{R}_j)^\vee.
\end{aligned}$$

Hence, our `LocalFrameProp` maps frame representations to invariant geometric features and subsequently to invariant scalar embedding updates. It's inclusion preserves the $\mathrm{SE}(3)$-invariance of the node embeddings and therefore does not break the symmetry of the surrounding FrameFlow architecture.

### A.6 Sampling Coverage

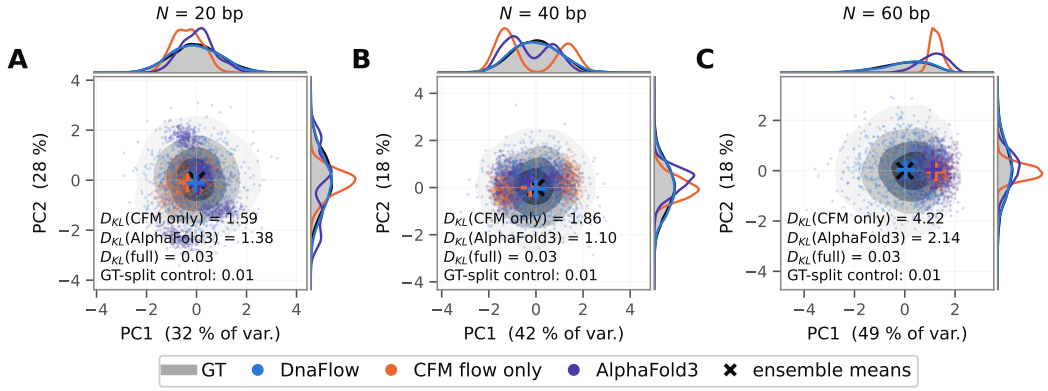



Figure 6: **Ground truth coverage of generated ensembles, at three chain lengths.** **(A)** $N = 20$ bp, **(B)** 40 bp, **(C)** 60 bp. Centerline ensembles projected onto the first two principal components of the hierarchical chain-growth reference at that length, pooled over five held-out sequences, 200 members per sequence and arm against that sequence's 1000 reference conformations. Each sequence keeps its own basis and is divided by the reference standard deviation along each component, so the pooled reference is isotropic with unit width (grey fill: highest-density regions enclosing 25, 50, 75 and 95 % of the mass), and everything the other clouds do relative to it is the signal. Blue, the full sampler; orange, conditional flow matching (CFM) using the same flow with both inference-time corrections removed, i.e., the spectral latent held at its sequence-conditioned mean and the thermal field switched off, leaving the conditional flow-matching field with only its initial noise. Both are drawn from the same seed, such that the two clouds differ by the ablation alone. Purple, AlphaFold3 run as an ensemble on the same held-out sequences (50 seeds $\times$ 4 diffusion samples per duplex), measured into centerline frames by the same estimator and projected into the same reference basis, so a published atomistic predictor enters the comparison in its own right. Crosses mark ensemble means. Axes are in units of the reference standard deviation and are drawn at equal aspect. The curves on the top and right edges are the PC1 and PC2 marginals of the same clouds, where a width mismatch reads off directly. Each panel gives $D_{\mathrm{KL}}(\text{reference} \,\|\, \text{arm})$ in nats, which charges an arm for deviations from the distribution of reference conformations. For reference, the ground truth (GT) split control is computed between the reference and an equally sized subsample of itself.

### A.7 Cosine Basis and Composition Seam

Viewed in the local frame a smooth global bend is modulated by the equilibrium twist per step $\Omega_0$, the rate at which the local frame turns about the helical axis, so a global bend appears near the mode index $k^* = \Omega_0 N/\pi$, which grows linearly in $N$. This is what the demodulation $\mathcal{D}$ of Section 3.3 removes: $\boldsymbol{R}_z(\varphi)$ is the rotation by $\varphi$ about the base-pair normal, so $\mathcal{D}$ turns the two transverse components back by the accumulated helical phase and leaves the twist channel untouched, $\tilde{\omega}_{z,i} = \omega_{z,i}$.

The demodulated field is then projected with the type-II discrete cosine transform. We write it as a fixed matrix $\boldsymbol{C} \in \mathbb{R}^{N\times N}$ rather than as a function, so that projection and re-expansion are transposes of one another,

$$C_{ki} = \gamma_k \sqrt{\frac{2}{N}} \cos\left[\frac{\pi k\,(i+\frac{1}{2})}{N}\right], \qquad \gamma_0 = \tfrac{1}{\sqrt{2}},\ \gamma_{k>0} = 1, \qquad \boldsymbol{C}^\top \boldsymbol{C} = \boldsymbol{I}_N,$$

acting on each axis independently. Its rows are indexed by the spatial frequency $k = 0, \ldots, N-1$ along the chain, $k = 0$ being the chain-wide mean and large $k$ step-to-step roughness, and its columns by the step index $i$. Projection is $\boldsymbol{\mu} = \boldsymbol{C}\tilde{\boldsymbol{\omega}}$, i.e., $\mu_{k\alpha} = \sum_i C_{ki}\,\tilde{\omega}_{i\alpha}$, and orthonormality makes it invertible with $\boldsymbol{C}^{-1} = \boldsymbol{C}^\top$, so the re-expansion $\tilde{\boldsymbol{\omega}} = \boldsymbol{C}^\top\boldsymbol{\mu}$ is exact and reuses the same entries with only the summed index changed, $\tilde{\omega}_{i\alpha} = \sum_k C_{ki}\,\mu_{k\alpha}$, since $(\boldsymbol{C}^\top)_{ik} = C_{ki}$. Read

along $i$ at fixed $k$, the row $C_{k\cdot}$ is the $k$-th cosine profile along the chain and $\boldsymbol{\mu}_k \in \mathbb{R}^3$ the per-axis amplitude it carries, so re-expansion is a superposition of profiles.

At the terminal composition of Section 3.3 the two fields are added in $\mathfrak{so}(3)$ rather than composed as a right product. Composing in the Lie algebra is required: every field on that seam is referenced to frame $i$ by the definition of $\boldsymbol{\omega}$.

### A.8 Twist Referencing of Spectral Coordinates

The demodulation $\tilde{\boldsymbol{\omega}}_i = \boldsymbol{R}_z(\varphi_i)\boldsymbol{\omega}_i$ fixes the helical axis, so the twist channel passes through unchanged, $\tilde{\omega}_{z,i} = \omega_{z,i}$. Before projection we subtract the equilibrium mean twist per step $\Omega_0^{\mathrm{eq}} = \langle \omega_z \rangle$, measured once over the training ensemble,

$$\tilde{\omega}_{z,i} = \omega_{z,i} - \Omega_0^{\mathrm{eq}},$$

while leaving the two transverse channels uncentered. The asymmetry is deliberate. On the transverse axes the zero-frequency component *is* the net bend of the chain, a conformational degree of freedom that must survive into the latent. On the twist axis the signal is dominated by the deterministic $\approx 0.61$ rad/step, so centering on a constant leaves a residual that is still a physical coordinate: $\mu_{0z}/\sqrt{N} = \langle \omega_z \rangle_{\mathrm{chain}} - \Omega_0^{\mathrm{eq}}$ is the chain's over- or under-twist relative to equilibrium.

### A.9 WLC Equipartition and Prior Amplitude

The discretization of the rod's bending energy on steps of mean rise $h$ is

$$\frac{H}{k_{\mathrm{B}}T} = \frac{L_p}{2}\int_0^L \boldsymbol{\kappa}_\perp^2(s)\,\mathrm{d}s \;=\; \frac{L_p}{2h}\sum_{i=0}^{N-1} \boldsymbol{\omega}_{\perp,i}^2,$$

with $\boldsymbol{\kappa}_\perp$ the two transverse (bending) components of the curvature vector and $\boldsymbol{\omega}_{\perp,i}$ their per-step integrals. Equipartition assigns each transverse axis the per-step variance $\sigma_{\omega\alpha}^2 = \langle \omega_{i\alpha}^2 \rangle = h/L_p$, a dimensionless constant independent of the position $i$ and of the total chain length. Because the projection is orthonormal, the coefficient variance of every mode equals this per-step variance, $\sigma_p[k,\alpha] = \sigma_{\omega\alpha}$: flat in $k$ and independent of $N$, so a single calibrated prior serves all chain lengths. In the continuous setting, with basis functions orthonormal on $\mathrm{L}^2[0,L]$, the same argument gives $\langle \kappa_k^2 \rangle = 1/L_p$, related to the discrete coefficient by $\mu_{k\alpha} = \sqrt{h}\,\kappa_{k\alpha}$. The macroscopic bend then follows without separate calibration: since $\sum_i \tilde{\omega}_{i\alpha} = \sqrt{N}\,\mu_{0\alpha}$, the accumulated bend has variance $N\sigma_{\omega\alpha}^2 = L/L_p$ per axis, recovering the WLC result $\langle \Theta^2 \rangle = 2L/L_p$ for uncorrelated steps. The $\sqrt{N}$ growth of the rod's bend is therefore obtained by integrating a length-independent local law, not by means of a length-dependent prior.

### A.10 Population Optimum of Spectral Mean Loss

The mean-field loss is

$$\mathcal{L}_\mu = \left\langle \left( \frac{\mathrm{sg}[\boldsymbol{\mu}] - \hat{\boldsymbol{\mu}}}{\sigma_p} \right)^2 \right\rangle,$$

with $\mathrm{sg}[\cdot]$ the stop gradient, $\boldsymbol{\mu}$ the coefficients measured from the clean frames of a single conformation and $\hat{\boldsymbol{\mu}}$ the head's sequence-conditioned prediction. Division by $\sigma_p$ removes the scale, so all three axes and all chain lengths are trained under a single weight. Its target is the projection of a *single* conformation, whose deviation from the sequence's conditional mean is by construction the fluctuation the prior $\sigma_p$ describes. Squared-error regression therefore has its population optimum at the conditional mean $\mathbb{E}[\boldsymbol{\mu}|\mathrm{seq}]$, which is the sequence-dependent mean spectrum. It is recovered without ever forming per-sequence averages over the training set, and hence without requiring more than one conformation per sequence in a batch.

### A.11 Correlated Thermal Kernels

DNA has known correlations between neighboring base pair steps (Lavery et al., 2010), which we model explicitly with a moving-average kernel. Let $\delta \boldsymbol{y}_i$ be the fluctuation of a per-step quantity

about its mean (rotations $\boldsymbol{\omega}_i$ or translations $\boldsymbol{\tau}_i$) and

$$\boldsymbol{\Gamma}_m = \left\langle \delta \boldsymbol{y}_i \, \delta \boldsymbol{y}_{i+m}^\top \right\rangle$$

its lag-$m$ covariance measured on the training ensemble, so that $\boldsymbol{\Gamma}_0$ carries the per-axis magnitudes $\sigma_\alpha = \sqrt{\Gamma_{0,\alpha\alpha}}$ on its diagonal and $\boldsymbol{\Gamma}_{m>0}$ the correlation between steps $m$ apart. The correlated field is generated from white noise $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_3)$ as

$$\boldsymbol{\eta}_i = \sum_{l=0}^{Q} \boldsymbol{M}_l \, \boldsymbol{\varepsilon}_{i-l}, \qquad \boldsymbol{\Gamma}_m = \sum_{l=0}^{Q-m} \boldsymbol{M}_l \, \boldsymbol{M}_{l+m}^\top, \qquad 0 \le m \le Q,$$

with kernel matrices $\boldsymbol{M}_l \in \mathbb{R}^{3\times 3}$ and order $Q$, an MA($Q$) process: each step's noise is a fixed linear combination of the last $Q+1$ independent draws, so it is correlated out to lag $Q$ and white beyond it.

For rotations, $Q = 1$ with diagonal $\boldsymbol{M}_l$ suffices. DNA's step-to-step bending anti-correlation is coherent in the lab frame, and the body frame between two consecutive steps has twisted by $\Omega_0$, hence in body-frame coordinates the correlation appears as an antisymmetric roll–tilt cross term, which a diagonal kernel cannot express. What that term controls is the bend accumulated over many steps, a low-frequency quantity owned by the latent's calibrated prior, and the thermal field is high-passed to $k \ge K$ so it never reaches that band. For step translations the cross terms are important, since the contour length responds to the component along the mean step and therefore mixes all three axes. This kernel is a full-matrix MA(2), giving $6+9+9 = 24$ free coefficients instead of the nine of a per-axis kernel. We parameterize the kernel matrices rather than the covariances they produce, which guarantees a valid correlated field throughout optimization and lets the coefficients be calibrated by gradient descent without constraining the covariances. $\boldsymbol{M}_0$ is taken lower triangular with positive diagonal to fix the non-uniqueness of the covariance factorization. Where the scale is carried differs by channel. The diagonal rotational kernel is normalized to unit marginal variance and its per-axis scale held separately in $\sigma_{\mathrm{th}}$, which is what makes $\sigma_{\mathrm{th}}$ a standard deviation in radians per step rather than a free gain. The full-matrix translational kernel cannot be split this way, since factoring out a per-axis scale would rescale the cross terms with it, so there the amplitude stays inside $\boldsymbol{M}_l$ and $\boldsymbol{\Gamma}_0$ carries it on its diagonal. One correction is needed for $\sigma_{\mathrm{th}}$ to remain a standard deviation of what actually lands: the high-pass is an orthogonal projection removing $K$ of $N$ modes, so it retains only a fraction $1 - K/N$ of a white field's variance per high-passed axis, an attenuation that depends on chain length. We therefore rescale the high-passed field by $\sqrt{(1-K/N_{\mathrm{ref}})/(1-K/N)}$, referenced to the length $N_{\mathrm{ref}} = 60$ at which $\sigma_{\mathrm{th}}$ was calibrated, which removes the length dependence while leaving the calibrated value untouched at $N_{\mathrm{ref}}$.

The construction has $\mathbb{E}[\boldsymbol{\eta}_i] = \mathbf{0}$ for every $\{\boldsymbol{M}_l\}$, so for translations the first moment is a separate parameter: a single body-frame vector $\boldsymbol{\Delta}_\tau$ added to $\boldsymbol{\eta}_i$ at the same point in the composition, matched to the reference per-step mean. The two are identified independently by construction, since the covariance target centers $\boldsymbol{\tau}$ per position across ensemble members and is therefore blind to a constant, while no kernel can move a mean. Both are needed because the two moments enter the accumulated length independently. Writing $\boldsymbol{\tau}_i = \boldsymbol{m}_i + \delta\boldsymbol{\tau}_i$ with mean step $\boldsymbol{m}_i = \mathbb{E}[\boldsymbol{\tau}_i]$ and covariance $\boldsymbol{\Sigma}_i = \langle \delta\boldsymbol{\tau}_i \delta\boldsymbol{\tau}_i^\top \rangle$, the expected length of an individual step is, to second order,

$$\mathbb{E}[\|\boldsymbol{\tau}_i\|] = \|\boldsymbol{m}_i\| + \frac{\operatorname{tr}\boldsymbol{\Sigma}_i - \hat{\boldsymbol{m}}_i^\top \boldsymbol{\Sigma}_i \hat{\boldsymbol{m}}_i}{2\|\boldsymbol{m}_i\|} + \mathcal{O}\big(\|\delta\boldsymbol{\tau}\|^3/\|\boldsymbol{m}\|^2\big), \qquad \hat{\boldsymbol{m}}_i = \boldsymbol{m}_i/\|\boldsymbol{m}_i\|.$$

Therefore the contour length, being a sum of step norms, depends on the mean step and the transverse fluctuation variance.

### A.12 Phase-sensitive Accumulation of Periodic Local Law

This argument is complex-valued, so throughout this appendix $\mathrm{i} = \sqrt{-1}$ is the imaginary unit and the base-pair step is indexed by $n$ rather than by the $i$ used elsewhere.

Write the transverse local-frame curvature as the complex quantity $w_n = \omega_{x,n} + \mathrm{i}\omega_{y,n}$, so that demodulation is $\tilde{w}_n = w_n e^{\mathrm{i}\Omega_0 n}$ and the lowest coefficient is the accumulated deformation, $\sum_n \tilde{w}_n = \sqrt{N}\,\mu_0$. Let the sequence impose a wedge of fixed direction $\chi$ in the local frame whose amplitude is modulated at period $p$,

$$w_n = \big(a_0 + a_1 \cos(2\pi n/p + \phi)\big)\, e^{\mathrm{i}\chi},$$

with mean amplitude $a_0$, modulation depth $a_1$ and phase offset $\phi$. Substituting and expanding the cosine gives three geometric sums,

$$\sum_{n=0}^{N-1} \tilde{w}_n = e^{\mathrm{i}\chi}\left[a_0 \sum_n e^{\mathrm{i}\Omega_0 n} + \frac{a_1}{2} e^{\mathrm{i}\phi} \sum_n e^{\mathrm{i}(\Omega_0 + 2\pi/p)\, n} + \frac{a_1}{2} e^{-\mathrm{i}\phi} \sum_n e^{\mathrm{i}(\Omega_0 - 2\pi/p)\, n}\right],$$

therefore, the modulation is mixed by the demodulation phase into the sidebands $\Omega_0 \pm 2\pi/p$. Each sum obeys $\left|\sum_{n=0}^{N-1} e^{\mathrm{i}\psi n}\right| = \left|\sin(N\psi/2)/\sin(\psi/2)\right| \le 1/|\sin(\psi/2)|$ for a phase increment $\psi$, so it is bounded independently of $N$ unless its exponent vanishes modulo $2\pi$, in which case it equals $N$. We call that condition *resonance*: the phase increment is a whole number of turns, $\psi \equiv 0 \pmod{2\pi}$, so $e^{\mathrm{i}\psi n} = 1$ for every $n$ — every step adds with the same phase and the sum grows as $N$. The unmodulated term never resonates, because $\Omega_0 \neq 0$. A sideband resonates exactly when $2\pi/p = \Omega_0$, i.e., when the modulation period matches the helical repeat $p = p_h = 2\pi/\Omega_0 \approx 10.3\,\mathrm{bp}$. On resonance the accumulated deformation therefore grows linearly in $N$ and $|\mu_0|$ grows as $\sqrt{N}$. Off resonance the accumulation cancels in this idealized example. A local, length-independent sequence law thus produces a macroscopic bend only when it is phased with the helix, which is why A-tract phasing enters as a low-$k$ shift of $\hat{\boldsymbol{\mu}}$, and why the peak of the phasing response is set by the law's *own* twist rate.

### A.13 Training Objective and Optimization

The model is trained with AdamW at learning rate $10^{-4}$ and weight decay $10^{-4}$, a three-epoch linear warm-up followed by a cosine schedule, for 50 epochs of 12 600 steps at batch size 250 on a single GPU, at $\approx 3$ h per epoch.

Table 2 lists every term of the objective with its weight. Strand- and rod-derived frames are corrupted at a shared flow time $t \sim \mathcal{U}([0.01, 1))$.

| term | $w$ | note |
|---|---|---|
| translation CFM | 2.0 | inherited from FrameFlow, $1/(1-t)^2$ reweighting, clipped at $t = 0.9$ |
| rotation CFM | 0.5 | as above |
| pseudo-atom + distogram | 0.25 | auxiliary, applied only for $t > 0.5$, where the structure is resolved |
| torsion | 1.0 | follows Baek et al. (2024) |
| rod CFM | 1.0 | the same $\mathrm{SE}(3)$ objective on the centerline chain |
| strand–centerline | 0.25 | ties each predicted strand to the predicted rod at the ground-truth $\boldsymbol{\xi}$ |
| spectral mean $\mathcal{L}_\mu$ | 1.0 | $\sigma_p$-normalized regression onto $\boldsymbol{\mu}$ from clean frames |
| per-step width | 1.0 | $\log\sigma$ regression onto the measured fluctuation profile, reads as $1 - R^2$ |
| thermal calibration | 1.0 | linearly ramped over the first 12 epochs, symmetric squared log-ratio |

Table 2: **Loss terms and weights.** The thermal calibration matches the per-step directional fluctuation, and the translational first and second moments, of a no-gradient rollout to cached local statistics of the training ensemble, penalizing over- and under-dispersion equally. The rollout is integrated at the step count used for sampling: the decoder's own across-member covariance depends on the integration depth, so a shorter probe would calibrate the field against an ensemble the sampler does not produce. Every rod term is a length-independent local statistic, so no global observable enters the objective.

### A.14 Ensemble Discrepancy $D_7$

Comparing two ensembles means comparing several observables that carry different units and differ in magnitude, so each is scored as a dimensionless ratio to its reference and averaged in log space,

$$D_7 \;=\; \frac{1}{7}\sum_{k=1}^{7} \left|\log \frac{\hat{o}_k}{o_k}\right|,$$

where $o_k$ is the reference value of observable $k$ and $\hat{o}_k$ its generated counterpart. The seven terms are the three per-step rotational standard deviations (tilt, roll, twist) together with the mean and standard deviation of both the global bend angle and the end-to-end distance. The logarithm makes the score symmetric in over- and under-dispersion, and $D_7 = 0$ is exact agreement.

Writing $\mathrm{mean}_m$ and $\mathrm{sd}_m$ for the mean and the standard deviation taken across the $M$ members $m = 1, \ldots, M$ of one ensemble at a fixed step, the three *local* observables are $o_\alpha = \frac{1}{N-1}\sum_{i=0}^{N-2} \mathrm{sd}_m[\omega_{i\alpha}^{(m)}]$ for $\alpha \in \{\mathrm{tilt}, \mathrm{roll}, \mathrm{twist}\}$, and the four *global* ones are $\mathrm{mean}_m[X]$ and $\mathrm{sd}_m[X]$ of the two per-member scalars $X \in \{\Theta, d_{\mathrm{e2e}}\}$: the bend angle $\Theta$ between the terminal tangents, and the end-to-end distance $d_{\mathrm{e2e}}^{(m)} = \|\boldsymbol{x}_{N-1}^{(m)} - \boldsymbol{x}_0^{(m)}\|$.

Tables 1 and 4 score five sequences at each length, and the two reductions are taken in the opposite order. An observable column is the geometric mean of the per-sequence ratios, a ratio being multiplicative: the arithmetic mean of 1.25 and 0.8 is 1.025 where the two errors are equal and opposite. $D_7$ is instead evaluated per sequence and averaged arithmetically, being already a mean of absolute logarithms. Deviations of opposite sign therefore cancel in a column and not in the score, and recomputing $D_7$ from a printed row understates it, by nearly threefold at 60 bp and not at all at 110 bp, where the departures have become a shared drift rather than sign-random.

### A.15 Atomistic Structure Quality

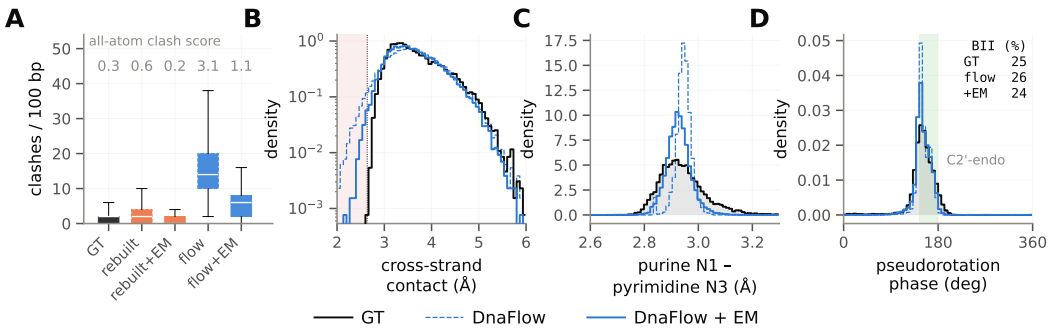



Figure 7: **Atomistic structure quality on one held-out 50 bp sequence**, 200 generated structures against 200 reference conformations. Panel A carries five arms as boxes, and panels B–D draw three of those arms as curves, the reference ensemble's own coordinates (black) and the production sampler through this work's all-atom rebuild, before the energy minimization (dashed) and after it (solid). The two rebuild arms, the same reference frames and torsions passed through that identical rebuild, appear in A alone: they carry zero model error, so they separate the price of the representation from the error of the model. **(A)** Steric clashes per 100 bp, per structure, at MolProbity's threshold (Chen et al., 2010) of a van der Waals overlap $> 0.4$ Å on heavy atoms, on the radii Word et al. (1999) defines it against. Whiskers span the central 95 %, and the number over each box is that arm's all-atom clash score per 1000 atoms, as the panel states. **(B)** Closest cross-strand base contact, against the closest the reference ensemble itself ever visits (2.64 Å, dotted, shaded below). **(C)** Purine N1 to pyrimidine N3 hydrogen bond. **(D)** Sugar pseudorotation phase, C2′-endo range shaded, with the BII backbone population inset.

All aforementioned metrics score the models base-pair frames, and are not sensitive to the atomistic detail, which is composed onto them. A duplex can carry clean centerline statistics, while its sugars clash. Since the building procedure places rigid ideal groups, intra-group bonds are exact by construction and carry no information about the model's performance. We evaluated five non-trivial quantities, each with the reference ensemble: steric overlap, the closest cross-strand base contact, the phosphodiester connecting bond, Watson–Crick pairing, and the sugar and backbone substates. Figure 7 draws four of them, the connecting bond being a closure residual rather than a prediction and treated in Appendix A.16.

The sugars show a median pseudo rotation phase of 149° against the reference's 152° and 99.9 % C2′-endo against 97.7 %. The BII backbone population is 25 % against 25 %, and $\mathrm{C1}' - \mathrm{C1}'$ and both groove widths match within their spreads. Steric overlap shows 15.8 clashes per 100 bp against the reference's 1.6, all of them cross-strand and concentrated on the minor-groove edge, the guanine amino group and the C2 positions facing it, a clash score (Chen et al., 2010) of 3.1 against 0.3, and 2.4 % of cross-strand base contacts below the closest the reference ever visits. They are also mildly concentrated at the ends, 10 % of them falling in the terminal two base pairs of each strand, 8 %

of the chain. Hydrogens are re-placed from each structure's own heavy atoms by one procedure before the all-atom score, so it compares skeletons rather than three ways of obtaining hydrogens. Appendix A.16 reads the clash counts against the thresholds in common use and treats the linkage, which is a closure residual rather than a prediction. We included a baseline for the error caused by the rebuilding kinematics from frames, which already induces 2.7 clashes per 100 bp at zero model error. These small clashes are to be expected from a deep learning approach that does not penalize clashes explicitly. Because our ground truth is primarily constrained by the repulsive arm of the Lennard-Jones potential to reduce clashes, we included a short energy minimization to yield a fair comparison: parmbsc1 (Ivani et al., 2016) with GB implicit solvent, minimized to an RMS force of 100 kJ mol$^{-1}$ nm$^{-1}$ with no iteration cap, under a harmonic restraint of 1000 kJ mol$^{-1}$ nm$^{-2}$ on every heavy atom.

What that costs is measured on the reference, which is why the reference's own rebuilt structures are pushed through the identical settings. Read back out of the refined atoms, the minimization moves the reference's ensemble by $D_7 = 0.019$ and the generated one by 0.030: it does stiffen what it touches, by a comparable amount in both arms and by less than the model's own $D_7$ in Table 1. The clash score of 1.1 per 1000 atoms sits well inside the $<5$ reported for experimental structures (Read et al., 2011), hence the derived atomistic structures are chemically sound after an inexpensive refinement, while the ensemble they were read from is moved by less than the model's own error.

### A.16 Reading Clash Counts Against Thresholds in Common Use

There is no single accepted steric threshold, and the spread between the published ones is far larger than the difference between any two of the ensembles above, so the criterion has to be stated before a count means anything. Table 3 scores the *same* coordinates three ways.

| criterion | unit | HCG | rebuilt | rebuilt+EM | generated | generated+EM |
|---|---|---|---|---|---|---|
| overlap $> 0.4$ Å, heavy | per 100 bp | 1.61 | 2.71 | 0.85 | 15.77 | 6.06 |
| clash score, all-atom | per 1000 atoms | 0.32 | 0.55 | 0.17 | 3.06 | 1.09 |
| $d < 2.0$ Å, heavy | per 100 bp | 0.00 | 0.00 | 0.00 | 0.08 | 0.01 |

Table 3: **The same structures under three steric criteria.** One held-out 50 bp sequence, 200 structures per arm. The first two are MolProbity's, on the radii of Word et al. (1999), and the third is the flat heavy-atom cutoff.

The flat 2 Å cutoff barely sees this defect. On the Word radii it is a van der Waals overlap of 1.1–1.6 Å depending on the atom pair, three to four times deeper than MolProbity's 0.4 Å. Under it all three reference arms score exactly zero and the generated ensemble 0.08 per 100 bp, less than a hundredth of what the strict criterion counts on the same coordinates: the criterion is satisfied by any structure whose atoms are not nearly superimposed. After refinement the generated arm falls to 0.01, so a paper quoting a flat cutoff would report essentially no clashes for any arm here and would compress a tenfold gap into a difference from zero.

The reference is the scale, not zero. HCG's own coordinates give 1.61 clashes per 100 bp, so the target is the reference's rate rather than zero. On the all-atom clash score, the quantity wwPDB validation reports, the refined samples reach 1.1 per 1000 atoms, in the range occupied by well-resolved experimental structures, against 0.3 for the reference's own coordinates and 0.2 for its rebuilt structures put through the identical refinement.

### A.17 Screening Out-of-distribution Ensemble Members

Table 4 re-scores the *same* draws as Table 1, restricted to those members that fall inside the region covered by the reference ensemble. The comparison isolates how much of the discrepancy at a given length is contributed by a tail of out-of-distribution outliers rather than by a systematic shift of the whole ensemble. In the training range the two scores agree to within 0.001, so those ensembles carry no such tail, and at 90 and 100 bp screening still has no noticeable effect, 0.046 to 0.045 and 0.057 to 0.056. At 110 bp outlier removal leaves $D_7$ at 0.070 against 0.071. The rotational spreads do respond, roll falling from 1.13 to 1.08, twist from 1.09 to 1.06 and tilt from 0.99 to 0.98, but the

global terms move the other way over the same restriction, the mean bend from $0.93$ to $0.91$ and the end-to-end spread from $0.91$ to $0.88$, so the aggregate score does not improve: at the longest chain the discrepancy is no longer carried by a removable tail alone. The screen is a per-axis $99.5$th percentile of the reference roughness and a member has to clear all three axes, so a model matching the reference exactly would still keep only $0.995^3 = 98.5\,\%$ of its draws. Against that calibration the kept fraction is $98\,\%$ through the training range and $95$, $92$ and $68\,\%$ at $90$, $100$ and $110$ bp: the extrapolation shows up as a growing tail long before it shows up in a moment.

| $N$ | members | tilt | roll | twist | $\text{bend}_{\text{mean}}$ | $\text{bend}_{\text{std}}$ | $\text{e2e}_{\text{mean}}$ | $\text{e2e}_{\text{std}}$ | $D_7 \pm \text{se}$ |
|---|---|---|---|---|---|---|---|---|---|
| 20 | all (1000) | 1.10 | 0.97 | 1.02 | 1.01 | 1.03 | 1.00 | 0.92 | $0.049 \pm 0.009$ |
| | screened | 1.09 | 0.97 | 1.02 | 1.01 | 1.03 | 1.00 | 0.92 | $0.049 \pm 0.009$ |
| 40 | all (1000) | 1.03 | 0.97 | 1.00 | 0.96 | 0.96 | 1.00 | 0.92 | $0.058 \pm 0.011$ |
| | screened | 1.02 | 0.97 | 1.00 | 0.96 | 0.95 | 1.00 | 0.91 | $0.058 \pm 0.011$ |
| 60 | all (1000) | 0.97 | 1.02 | 1.00 | 1.00 | 1.01 | 0.99 | 0.98 | $0.036 \pm 0.003$ |
| | screened | 0.96 | 1.01 | 1.00 | 1.00 | 1.01 | 0.99 | 0.98 | $0.037 \pm 0.003$ |
| 80 | all (1000) | 0.94 | 1.03 | 0.99 | 1.01 | 0.99 | 1.00 | 0.90 | $0.051 \pm 0.006$ |
| | screened | 0.94 | 1.03 | 0.99 | 1.00 | 0.99 | 1.00 | 0.90 | $0.052 \pm 0.006$ |
| 90† | all (1000) | 0.94 | 1.05 | 1.03 | 0.98 | 1.00 | 0.99 | 0.95 | $0.046 \pm 0.004$ |
| | screened | 0.94 | 1.05 | 1.03 | 0.99 | 0.99 | 0.99 | 0.94 | $0.045 \pm 0.003$ |
| 100† | all (1000) | 0.95 | 1.08 | 1.02 | 0.99 | 1.00 | 0.99 | 0.97 | $0.057 \pm 0.010$ |
| | screened | 0.95 | 1.06 | 1.01 | 0.99 | 1.00 | 0.99 | 0.98 | $0.056 \pm 0.011$ |
| 110† | all (1000) | 0.99 | 1.13 | 1.09 | 0.93 | 0.97 | 0.97 | 0.91 | $0.071 \pm 0.006$ |
| | screened | 0.98 | 1.08 | 1.06 | 0.91 | 0.96 | 0.98 | 0.88 | $0.070 \pm 0.003$ |

Table 4: **Full draw versus the screened re-score, across duplex length.** $1000$ members per length, split evenly over five held-out sequences. The *all* and *screened* rows score the same draw, the latter restricted to members inside the reference-covered region. Columns and $D_7$ are combined over the five sequences as in Table 1, so here too the score is not recoverable from the row. †Held-out length, outside the $11$–$80$ bp training range.

### A.18 Sequence-dependent Baselines

The identification distance of Figure 3A is

$$D(s, s') = \sum_{i,\alpha} \left( \frac{\hat{f}_{i\alpha}^{(s)} - f_{i\alpha}^{(s')}}{\sigma_{i\alpha}} \right)^2,$$

where $\sigma_{i\alpha}$ is the reference's own half-split sampling standard deviation of that component, and the per-step profile score of Figure 3C is

$$r^{(s)} = \text{corr}_i\big(\hat{f}_{i\alpha}^{(s)},\, f_{i\alpha}^{(s)}\big), \qquad \alpha = \text{roll},$$

each profile centred on its own chain mean, so $r^{(s)}$ uses the shape of the profile as input and not its offset. The paired field $R^2$ is defined in Section 4.3; its sum of squares runs over all steps, degrees of freedom and sequences at once. In Figure 3A the cells are the $\log_{10}$ distance over the row minimum; in Figure 3C the reference band is twice its own sampling spread and the drawn sequence is indicated by a ring in the strip. The scale is set by what a rigid-base-pair table of each context order achieves on the same measurement, computed on the reference data itself, and Table 5 gives that ladder beside the arms.

The tetramer table pushed through the identical sampling pipeline is the conservative comparison, and the model comes out ahead of it, $+0.873$ against $+0.783$, on a measurement that favours the table twice over. Its parameters are per-context mean step parameters, obtained by direct averaging over the same reference library and covering all $256$ tetramers, whereas the model must reproduce them indirectly, by sampling, from an objective that never forms a per-sequence average. And the

| | arm | top-1 | field $R^2$ | low-$k$ $R^2$ |
|---|---|---|---|---|
| generated | model | **1.000** | **+0.873** | **+0.670** |
| | tetramer table, same pipeline | **1.000** | +0.783 | +0.510 |
| | permuted sequence | 0.075 | −0.831 | −1.092 |
| table order, on the reference | composition only | – | −0.062 | – |
| | dinucleotide | – | +0.569 | – |
| | tetramer | – | +0.823 | – |
| | reference split-half ceiling | – | +0.988 | – |

Table 5: **Where the model sits among rigid-base-pair tables.** $S = 40$ held-out 90 bp sequences, 64 members each, chance top-1 $= 0.025$. The upper block is measured on generated ensembles, the lower block on the reference data directly, so the two are on the same axis but not the same pipeline.

score is a per-step local field, exactly what a table of fixed context width represents, so it charges nothing for accumulation — whereas the model's receptive field is 13 bp and its accumulation unbounded, which is what Section 4.3's last two paragraphs test.

Resolved per degree of freedom the model's field agreement is uneven, strongest on the in-plane translations and twist (+0.909 slide, +0.886 twist, +0.867 shift) and weakest on rise (+0.637) and roll (+0.672), with tilt between them at +0.754. Roll is the axis the spectral latent owns, so this is the same asymmetry that the low-$k$ column reports. Figure 8 resolves the same comparison per step, drawing Figure 3C for each of the six in turn.

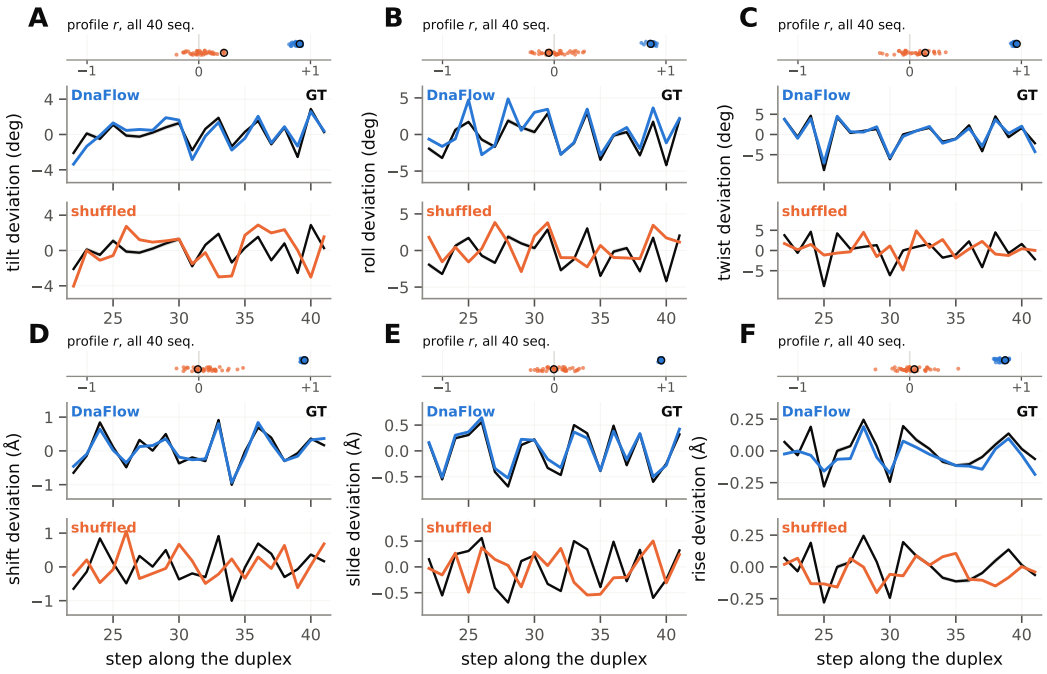



Figure 8: **Per-step field agreement in all six step parameters.** Figure 3C draws roll, because roll is the soft bending mode and the one the sequence-directed wedge lives in; this plate draws the same comparison for each of the six, rotations above and translations below, in the order of the frame convention. Each cell sets the reference (black, band twice its half-split sampling spread) against one arm per row, generated above and permuted sequence below, as the ensemble-mean deviation from the field all $S = 40$ sequences share, and carries a strip of both arms' profile correlation $r$ over all 40 sequences with the drawn one ringed. The held-out sequence and the window of steps are chosen on roll exactly as Figure 3C chooses them and then handed to all six cells.

### A.19 PHASED A-TRACT CONSTRUCTS

**The constructs.** Each molecule holds $8$ copies of $A_5$ whose centres are spaced by a period $P$ taken over $8, 8.5, \ldots, 12, 13$ bp, so the chain is $7P + 5$ base pairs long and the set spans 61–96 bp, the long-period end of it past the 11–80 bp training range. Half-integer periods alternate the two neighbouring spacings, which is how the underlying experiments build them. Each construct is paired with a purine-matched control in which the run is interrupted, so composition is held and only the *run* is destroyed, and the *excess bend* $\Theta_{\mathrm{ex}}(P)$ is the turn of the ensemble-mean helical axis between its terminal tangents, taken over an $11$ bp window so the helix's own precession is not read as curvature, of the construct minus that of its control. Every arm is measured through the same estimator on its own ensemble mean, so an arm that generates no coherent wedge scores zero rather than its own noise.

Appendix A.12 shows that the accumulation is a Dirichlet kernel in the mismatch between the modulation period and the arm's own helical repeat. The kernel resonates for *any* coherent wedge of the right period (i.e., local bends add up to a large in-plane bend), whatever direction that wedge points in the local frame. Thus, a model with a wrong-but-consistent sequence law still produces a peak, and it produces it at its own twist rate rather than at the reference's. This is exactly what the ablated arm does, and it peaks $1.32$ bp off its own repeat rather than on it. Table 6 shows the four setups. The reference is the rigid-base-pair tetramer law of the HexABC set, fitted to explicit-solvent trajectories the model never saw, restricted to its rotation field so that it is comparable to the model's. The training-library law is the identical fit performed on the HCG library the checkpoint was trained on, and it is what the model's amplitude should be read against: the model cannot carry a wedge its data does not contain, so the reference-minus-library gap is a property of the data and only the library-minus-model gap is the model's. It is a proxy for that ceiling rather than a bound, since the model sees the library's structures and not a tetramer fit of them, and supra-tetramer A-tract structure in the library would sit above the curve; hierarchical chain growth stacks short fragments and attenuates exactly that, which is what makes the proxy a fair one. The full sampler is the production path. The null re-draws the trunk's base-identity embedding at the learned row norm, which is the only ablation available on the generative path.

| arm | own $p_h$ (bp) | peak (bp) | excess (deg) | $r$(field) |
|---|---|---|---|---|
| HexABC MD, independent | 10.43 | $10.43 \pm 0.00$ | $47.6 \pm 0.0$ | – |
| training-library law | 10.21 | $10.03 \pm 0.00$ | $28.8 \pm 0.0$ | $+0.93$ |
| model, full sampler | 10.20 | $10.12 \pm 0.18$ | $31.2 \pm 1.5$ | $+0.77$ |
| sequence-blind null | 9.58 | $10.90 \pm 0.19$ | $15.0 \pm 3.9$ | $-0.24$ |

Table 6: **Phased A-tract response, by arm.** Peak and excess are over three independent sampling seeds. *own* $p_h$ is the arm's own helical repeat, measured from its own ensemble, and the peak is meaningful only relative to it. $r$(field) is the per-step correlation of the arm's roll field against the independent tetramer table over eight random 60-mers, which is the quantity that separates a learned law from a coherent accident.

The full sampler recovers $65\,\%$ of the reference amplitude, $108\,\%$ of the training-library law's, and peaks $0.08$ bp from its own repeat. Here the null misses its own repeat by $1.32$ bp, so the period does separate the two arms in this run, but it is not what the claim rests on: the kernel resonates for any coherent wedge, and a null that happened to accumulate one would peak as cleanly as the model. The test is which wedge each arm accumulates, and that is the field. The model's per-step roll correlates with the independent table at $r = +0.77$ and recovers $0.61$ of its sign-carrying A-tract roll contrast. The null does not correlate with it at all, $r = -0.24$.

The missing amplitude with respect to HexABC can not be attributed to the model. The library law reaches $28.8^\circ$ against the reference's $47.6^\circ$ with the per-step shape it is scored on intact, $r = +0.93$ against the reference, so two fifths of the reference amplitude is missing from the data before the model is asked for anything; the model then sits $2.3^\circ$ *above* that ceiling, $1.6$ times its own sampling standard deviation, so the entire gap to the independent reference is a property of the data. What the model does not match is the *shape* of the library's wedge in detail: its sequence-dependent roll field is smaller than the table's in raw spread, $0.030$ against $0.040$ rad, carrying $0.57$ of the table's amplitude on the component that correlates with it. Separately the two laws do not share a helical

repeat. On the A-tract probe the training library twists at $0.615$ rad/step and the reference table at $0.602$, so their repeats are $p_h = 10.21$ and $10.43$ bp. The model reads back $0.616$ rad/step from its own ensemble, $p_h = 10.20$: it reproduces the repeat of the library it was trained on to $0.1\,\%$, so the offset from the reference is a property of the training data and not a calibration the model missed. Over eight tracts a mismatch of that size costs coherence at the reference's period even when the local law is right, which is why the two arms are compared each on its own repeat.