
ChromWeave: Equivariant Generative Modeling of 3D Genome Organization

Shuming Liu¹
shuming3@andrew.cmu.edu

Hongwei Tu^{2,3}
hongweit@mit.edu

Lorenzo Bini^{1,4}
lorenzo.bini@unige.ch

Jian Ma¹
jianma@cs.cmu.edu

¹ Ray and Stephanie Lane Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

² Computer Science Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

³ Present Address: Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA

⁴ Department of Computer Science, University of Geneva, Geneva, Switzerland

Abstract

Understanding the three-dimensional (3D) ensemble of chromosome conformations is essential, as dynamic structural states underlie diverse processes such as gene regulation, cell-type identity, and disease. However, systematically modeling how 1D epigenomic features give rise to 3D chromosome ensembles remains an unsolved challenge. Inspired by advances in protein structure prediction and modern generative modeling, we introduce ChromWeave, a symmetry-aware generative framework that couples flow matching with an equivariant neural network to directly generate physically realistic 3D chromosome ensembles from epigenomic signals. By enforcing physical symmetries, ChromWeave achieves orders-of-magnitude speedup over traditional physics-based simulations while producing experimentally consistent structural distributions. Trained on single-cell MERFISH chromosome imaging data and raw 1D epigenomic features, ChromWeave learns a continuous-time equivariant velocity field that maps simple prior distributions to realistic conformational distributions, avoiding large-scale pretraining and remaining highly data-efficient. Our results show that ChromWeave accurately captures both polymeric properties of chromosomes and genomic organizational features such as domains and boundaries that closely match MERFISH measurements. Together, ChromWeave provides a computationally efficient and physically grounded approach for mapping 1D epigenomic features to 3D genome ensembles, with the potential to generalize across genomic regions and ultimately scale to whole-genome modeling.

1 Introduction

The three-dimensional (3D) organization of the genome is central to cellular functions. Within the nucleus, chromosome folding plays key roles in gene transcription, DNA replication, and other regulatory processes [1, 2, 3, 4, 5]. These 3D structures fluctuate across an ensemble of conformations, reflecting intrinsic dynamics and cell-to-cell variability [6, 7, 8, 9]. Capturing this structural heterogeneity is essential yet challenging because chromosomes are long, flexible polymers with complex interactions, and measurements are often indirect or incomplete.

Classical polymer models have revealed fundamental principles of chromosome folding [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22], but rely on coarse assumptions, struggle to capture sequence- or locus-specific effects, and are computationally demanding at genome scale. In contrast, recent advances in deep generative modeling have transformed structural biology. Breakthroughs in protein structure prediction, such as AlphaFold [23] and RosettaFold [24], demonstrate that complex spatial dependencies can be learned directly from sequence. BioEmu [25] and AlphaFlow [26] further extend the paradigm to structural ensemble generation, capturing conformational variability with high fidelity. We seek a similar sequence-to-structure paradigm for chromosomes, where 3D structure ensembles, rather than 2D contact maps, better represent nuclear architecture. Chromosomes are dynamic, and single conformations are less informative than the distributions of structures that define functional states [6, 27, 7, 28, 9]. Existing experimental methods remain limited: population Hi-C averages contacts [29, 30], single-cell Hi-C is sparse and noisy [31, 4], and chromatin tracing capture 3D conformations directly but with limited coverage and resolution [32, 33, 34, 9]. These limitations highlight the need for systematic, data-driven methods that can directly map 1D genomic features to 3D structural ensembles, with the ability to generalize across genomic regions. ChromoGen [35] moves in this direction by leveraging EPCOT [36] pretrained genomic embeddings and reconstructing 3D structures from 2D distance maps. However, distance maps remain a 2D surrogate: enforcing global metric consistency (e.g., the triangle inequality) is difficult. The all-pairs 2D formulation also scales poorly with increasing sequence length and cannot natively encode spatial perturbations or nuclear landmarks. These limitations underscore the need for end-to-end models that directly learn 3D ensembles while respecting symmetry constraints.

Here, we introduce ChromWeave, an equivariant flow-matching framework that learns chromosome ensembles directly in 3D from 1D raw epigenomic features, analogous to the sequence-to-structure paradigm in protein science. Unlike ChromoGen, which reconstructs conformations indirectly from 2D distance maps, ChromWeave performs end-to-end 3D generation without relying on pretrained embeddings, while incorporating polymeric constraints and physical symmetries. Equivariant flow matching ensures geometric consistency and implicitly provides data augmentation [37]. Direct 3D modeling further improves scalability through sparse graphs and enables the incorporation of spatial perturbations and nuclear landmarks into the generative process. As a result, ChromWeave avoids costly pretraining, remains computationally efficient, and generates realistic structural ensembles consistent with experimental measurements, including polymeric scaling, contact patterns, and domain boundaries, with the potential to generalize to unseen genomic regions.

2 Methods

2.1 Method overview

As illustrated in **Fig. 1A**, ChromWeave is a generative framework that maps 1D epigenomic features to 3D, physically realistic chromosome polymer ensembles, where each bead (monomer) represents a 250-kb (kilobase) genomic segment. In this work, we focus on chromosome 2 (Chr2) spanning approximately 240 Mb (megabases) in the hg38 genome assembly. At 250-kb resolution, this region is represented by 967 beads. ChromWeave directly models how genomic context shapes chromosome organization by encoding epigenomic features – including transcription factor (TF) and cohesin binding, transcription activity, histone modifications, and chromatin accessibility – into representative embeddings that condition the generative process. Starting from Gaussian-distributed initial coordinates, an equivariant flow-matching framework learns a velocity field that progressively transforms noisy conformations into realistic 3D chromosome structure ensembles (**Fig. 1B**). The generated ensembles are not only plausible but also approximate the true distribution of conformations measured by MERFISH [9]. This distributional perspective is essential because it connects 1D epigenomic signals with the intrinsically dynamic nature of 3D genome folding, establishing a direct link between epigenomic features and structural ensembles. The generated ensembles are then evaluated against experimental data using polymer scaling, contact patterns, and domain boundaries, providing biologically grounded assessment of 3D genome organization. While related generative techniques have been widely applied to protein modeling, we introduce them here for the first time in the context of genome architecture, establishing a framework for direct 3D generation of chromosome ensembles from epigenomic features. Below we summarize the core ideas of flow matching and the equivariant neural network, with further details on the model, training, and inference provided in Supplementary Note **B.2**.

2.2 Flow matching

Flow matching (FM) [38, 39, 40, 41, 42] trains a flow model that transforms samples from a simple prior distribution p_0 to a complex distribution p_1 approximating the data distribution. The evolution of distribution $p_t(\mathbf{r})$ under a velocity field $\mathbf{v}_t(\mathbf{r})$ follows the continuity equation [38, 42]

$$\frac{\partial p_t(\mathbf{r})}{\partial t} + \nabla_{\mathbf{r}} \cdot (p_t(\mathbf{r}) \mathbf{v}_t(\mathbf{r})) = 0. \quad (1)$$

FM learns a velocity model $\hat{\mathbf{v}}_t(\mathbf{r}_t; \theta)$ that approximates \mathbf{v}_t . Diffusion models [43, 44, 45] can be viewed as a relevant generative framework but with stochastic dynamics, whereas FM uses deterministic paths with faster and more stable convergence—especially suitable for chromosome modeling, since chromosomes are dynamic and only realistic structural ensembles are biologically meaningful.

In this work, we use an isotropic Gaussian distribution as prior p_0 , and a linear interpolation path

$$\mathbf{r}_t = (1 - t)\mathbf{r}_0 + t\mathbf{r}_1 \quad (\mathbf{r}_t \in \mathbb{R}^{L \times 3}, t \in [0, 1]), \quad (2)$$

where \mathbf{r}_t is the configuration of a chain with L beads at time $t \in [0, 1]$. To remove relative global translations and rotations between configurations \mathbf{r}_0 and \mathbf{r}_1 , thus improving transport efficiency, we align centered coordinates \mathbf{r}_0 to \mathbf{r}_1 using the Kabsch algorithm [46] before interpolation. Theorem 1 proves that this ensures the entire interpolation \mathbf{r}_t ($t \in [0, 1]$) remains aligned to \mathbf{r}_1 .

Training FM directly is intractable, since the true velocity field \mathbf{v}_t is unknown. However, it can be reformulated as conditional flow matching (CFM) [38, 42, 47], which provides an equivalent, tractable, and unbiased objective. CFM learns a velocity field $\hat{\mathbf{v}}_t(\mathbf{r}_t; \theta)$ by minimizing the CFM loss

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_{t \sim \mathcal{T}, \mathbf{r}_1 \sim q, \mathbf{r}_0 \sim p_0} \|\hat{\mathbf{v}}_t(\mathbf{r}_t; \theta) - \mathbf{v}_t(\mathbf{r}_t | \mathbf{r}_1)\|^2. \quad (3)$$

where the conditional velocity is

$$\mathbf{v}_t(\mathbf{r}_t | \mathbf{r}_1) = \left. \frac{d\mathbf{r}_t}{dt} \right|_{\mathbf{r}_0, \mathbf{r}_1} = \mathbf{r}_1 - \mathbf{r}_0 = \frac{\mathbf{r}_1 - \mathbf{r}_t}{1 - t}. \quad (4)$$

In practice, CFM loss is often reparameterized [48, 49, 26] to predict $\hat{\mathbf{r}}_1 \approx \mathbf{r}_1$, thus yielding

$$\hat{\mathbf{v}}_t(\mathbf{r}_t; \theta) = \frac{\hat{\mathbf{r}}_1(\mathbf{r}_t, t; \theta) - \mathbf{r}_t}{1 - t}. \quad (5)$$

At inference, the target sample is obtained by integrating an ordinary differential equation (ODE)

$$\hat{\mathbf{r}}_t^{\text{ODE}} = \mathbf{r}_0 + \int_0^t \hat{\mathbf{v}}_{t'}(\mathbf{r}_{t'}; \theta) dt'. \quad (6)$$

Here $\hat{\mathbf{r}}_1^{\text{ODE}}$ denotes the endpoint at $t = 1$ obtained by solving the ODE, in contrast to the one-shot prediction $\hat{\mathbf{r}}_1(\mathbf{r}_t, t; \theta)$. RMSD alignment is applied at each Euler step during integration to eliminate irrelevant global moves, thereby improving training stability [50, 42]. Further details on the training algorithm and inference procedure are provided in Supplementary Note B.2.

2.3 Equivariant neural network

The internal geometry of chromosome structures should remain invariant under global translation and rotation. While RMSD alignment (Section 2.2) eliminates relative rigid-body motions between \mathbf{r}_0 and \mathbf{r}_1 , we further employ equivariant networks so that predictions transform correspondingly with globally moved inputs. Specifically, we train a denoising model $\hat{\mathbf{r}}_1(\mathbf{r}_t, t; \theta)$ with internal equivariant coordinate update layers as introduced in EGNN [51]:

$$\mathbf{r}^{l+1, i} \leftarrow \mathbf{r}^{l, i} + \sum_{j \in N(i)} w_{ij} \cdot \frac{\mathbf{r}^{l, i} - \mathbf{r}^{l, j}}{\|\mathbf{r}^{l, i} - \mathbf{r}^{l, j}\| + \Delta}. \quad (7)$$

The superscript l indexes message-passing layers, with $\mathbf{r}^{l=0} = \mathbf{r}_t$, and $\mathbf{r}^{l=l_{\max}} = \hat{\mathbf{r}}_1$. Superscripts i, j index chromosome beads, each representing a 250-kb genomic segment. $N(i)$ denotes the neighbors of node i . Although we use fully connected graphs in this study, the formulation naturally extends to sparse graphs. w_{ij} are model-predicted weights. $\Delta = 1$ for improved numerical stability [52, 51, 53]. Because the update depends only on relative positions, Equation 7 ensures that the output \mathbf{r}^{l+1} transforms consistently with the input \mathbf{r}^l under global motions, thereby preserving intrinsic 3D geometry. This equivariant design also acts as implicit data augmentation: all rigidly transformed variants of the same structure are inherently captured, thereby improving robustness and data efficiency. Further details are provided in Supplementary Note B.2.

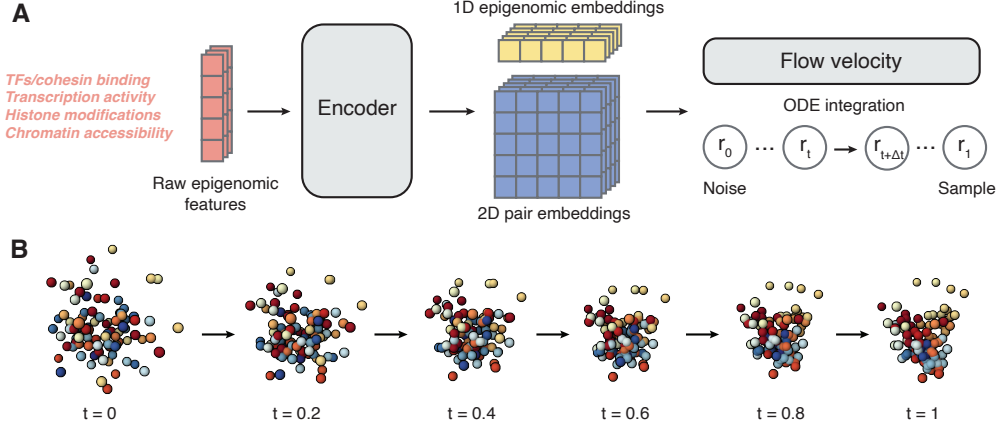


Figure 1: Overview of the ChromWeave framework. (A) Model architecture. Raw epigenomic features are encoded into 1D epigenomic and 2D pair embeddings, which are passed to a flow-based generative model. The learned flow velocity field defines an ODE that transforms an initial noisy conformation r_0 into a sample r_1 resembling realistic chromosome conformation. (B) Flow matching trajectory. Starting from Gaussian noise at $t = 0$, the flow progressively organizes beads (250-kb segments) into a structured conformation at $t = 1$. Beads are colored by genomic index, from red (low index) through yellow to blue (high index).

3 Results

We evaluated ChromWeave with an isotropic Gaussian noise prior. An overview of the ChromWeave framework, including the encoder, embedding modules, and flow-based generative model, is shown in **Fig. 1A**. An example flow-matching trajectory is shown in **Fig. 1B**. Because chromosomes are highly dynamic and the model is designed to capture distributions rather than single conformations, all evaluations are based on ensemble averages and statistical properties. We present results mainly for Chr2 beads 201-300, a region with clear and well-resolved structural features that highlights the model’s capabilities, and organize the analyses into two levels: (i) fundamental polymeric features of chromosome folding and (ii) genomic-specific structural characteristics. Analogous analyses for another region (Chr2 1-100) are provided in the Supplementary Figures.

Fundamental polymeric features. At the fine scale, ChromWeave-generated ensembles preserve bond connectivity: the bond-length distribution closely matches the true distribution (**Fig. S1A**, see also **Fig. S2A**). At the global scale, the radius of gyration R_g (see Supplementary Note B.3 for definition) agrees well with experimental values (**Fig. S1B**, see also **Fig. S2B**). These results indicate faithful capture of both local connectivity and large-scale compaction. Moreover, ChromWeave reproduces expected polymer scaling laws, including the power-law decay of contact probability (**Fig. 2A**, see also **Fig. S3A**) and the scaling of mean pairwise distance with sequence separation (**Fig. 2B**, see also **Fig. S3B**). These statistics capture the fundamental polymer physics of chromosomes, representing baseline behaviors any simple polymer would exhibit and providing a reference for experimental validation. These findings demonstrate that ChromWeave accurately models the polymeric nature of chromosomes.

Genomic-specific structural characteristics. Beyond generic polymer behavior, ChromWeave also captures locus-specific chromatin organization. Average contact maps generated by ChromWeave closely resemble the experimental ensemble (**Fig. 3A**), with a Pearson correlation of 0.9512 and a stratum-adjusted correlation coefficient (SCC) of 0.6697 (see Supplementary Note B.3 for definitions). Insulation score profiles (see Supplementary Note B.3), which quantify how strongly genomic loci are insulated from their neighbors and help identify chromatin domain boundaries, also closely match experimental data (**Fig. 3B**, $\text{corr} = 0.8302$). Comparable analyses for another region (Chr2 1-100) are shown in **Fig. S4**, where correlations are lower but the overall trends remain consistent. These results further demonstrate that ChromWeave not only preserves generic polymeric features but also reproduces biologically meaningful chromatin structure patterns such as domain boundaries, while achieving these high-fidelity reconstructions with practical efficiency.

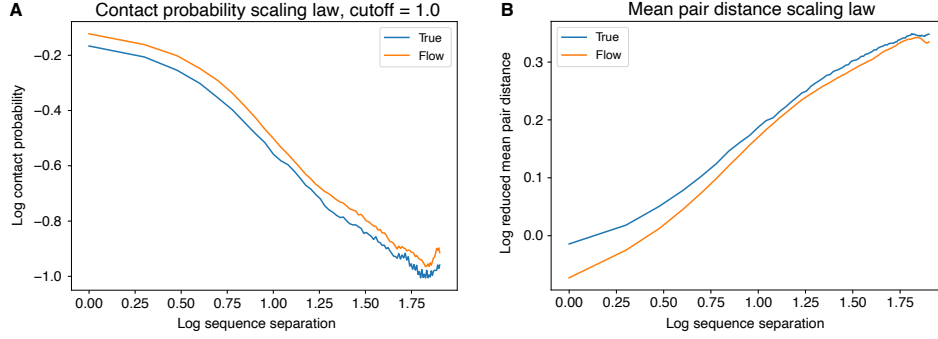


Figure 2: Scaling laws of ChromWeave-generated chromosome conformations (chr2, beads 201–300). (A) Contact probability versus sequence separation (cutoff = 1.0). (B) Reduced mean pairwise distance versus sequence separation. Both are shown on a base-10 log-log scale.

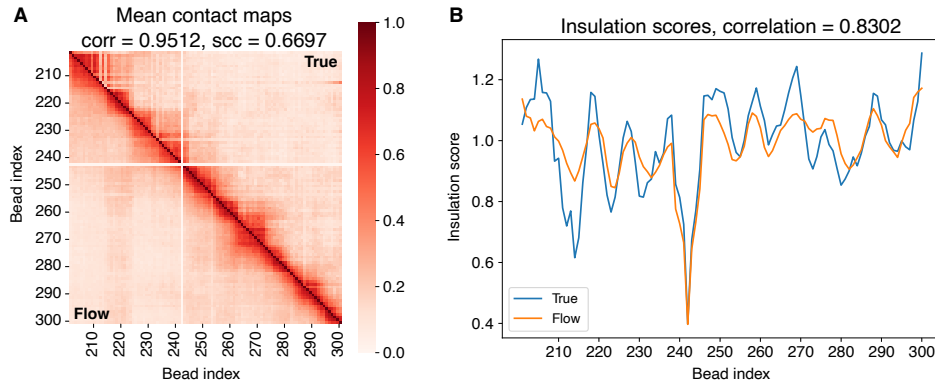


Figure 3: Comparison of contact maps and insulation scores. (A) Mean contact maps for chr2 (beads 201–300), comparing the true ensemble (upper triangle) with ChromWeave-generated conformations (lower triangle). Reported similarity metrics include Pearson correlation ($\text{corr} = 0.9512$) and stratum-adjusted correlation coefficient ($\text{scc} = 0.6697$). (B) Insulation score profiles for the same region. ChromWeave predictions (orange) closely match the true data (blue), with $\text{corr} = 0.8302$.

4 Conclusion

The ability to generate accurate and diverse 3D chromosome structure ensembles that reflect the correct underlying distribution is crucial for understanding genome function. In this work, we introduced ChromWeave, a data-driven generative framework that directly models chromosome conformations in 3D from raw epigenomic features using equivariant flow matching. Unlike approaches that rely on distance maps, expensive pretrained embeddings, or massive structure datasets, ChromWeave builds physically consistent ensembles end-to-end. The equivariant design and RMSD-based alignment make the model both data-efficient and practically usable, yielding high-fidelity structural distributions that closely match experimental MERFISH imaging data.

By combining generative modeling with equivariant architectures, ChromWeave offers an efficient alternative to polymer simulations and opens new opportunities for probing how epigenomic features shape 3D chromosome organization. Future directions include scaling the framework to entire genomes with memory-efficient sparse graphs, improving flow matching with advanced transport schemes, and performing ablations to identify which epigenomic features drive chromatin folding. Incorporating additional biological priors such as nuclear landmarks and perturbations could expand the framework’s scope, while integration with pretrained genomic models could enhance accuracy and generalization. In the long term, ChromWeave offers a path toward comprehensive, dynamic, and multi-scale modeling of genome organization, bringing us closer to decoding the relationship between 1D epigenomics and 3D structural dynamics.

References

- [1] Job Dekker and Leonid Mirny. The 3d genome as moderator of chromosomal communication. *Cell*, 164(6):1110–1121, 2016.
- [2] M Jordan Rowley and Victor G Corces. Organizational principles of 3d genome architecture. *Nature Reviews Genetics*, 19(12):789–800, 2018.
- [3] Shuming Liu, Advait Athreya, Zhuohan Lao, and Bin Zhang. From nucleosomes to compartments: physicochemical interactions underlying chromatin organization. *Annual review of biophysics*, 53, 2024.
- [4] Yang Zhang, Lorenzo Boninsegna, Muyu Yang, Tom Misteli, Frank Alber, and Jian Ma. Computational methods for analysing multiscale 3d genome organization. *Nature Reviews Genetics*, 25(2):123–141, 2024.
- [5] Job Dekker and Leonid A Mirny. The chromosome folding problem and how cells solve it. *Cell*, 187(23):6424–6450, 2024.
- [6] Takashi Nagano, Yaniv Lubling, Tim J Stevens, Stefan Schoenfelder, Eitan Yaffe, Wendy Dean, Ernest D Laue, Amos Tanay, and Peter Fraser. Single-cell hi-c reveals cell-to-cell variability in chromosome structure. *Nature*, 502(7469):59–64, 2013.
- [7] Bogdan Bintu, Leslie J Mateo, Jun-Han Su, Nicholas A Sinnott-Armstrong, Mirae Parker, Seon Kinrot, Kei Yamaya, Alistair N Boettiger, and Xiaowei Zhuang. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*, 362(6413):eaau1783, 2018.
- [8] Longzhi Tan, Dong Xing, Chi-Han Chang, Heng Li, and X Sunney Xie. Three-dimensional genome structures of single diploid human cells. *Science*, 361(6405):924–928, 2018.
- [9] Jun-Han Su, Pu Zheng, Seon S Kinrot, Bogdan Bintu, and Xiaowei Zhuang. Genome-scale imaging of the 3d organization and transcriptional activity of chromatin. *Cell*, 182(6):1641–1659, 2020.
- [10] Pierre-Gilles De Gennes. *Scaling concepts in polymer physics*. Cornell university press, 1979.
- [11] Alexei R Khokhlov, Alexander Yu Grosberg, and Vijay S Pande. *Statistical physics of macromolecules*, volume 1. Springer, 1994.
- [12] Michael Rubinstein and Ralph H Colby. *Polymer physics*. Oxford university press, 2003.
- [13] Adrian L Sanborn, Suhas SP Rao, Su-Chen Huang, Neva C Durand, Miriam H Huntley, Andrew I Jewett, Ivan D Bochkov, Dharmaraj Chinnappan, Ashok Cutkosky, Jian Li, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences*, 112(47):E6456–E6465, 2015.
- [14] Bin Zhang and Peter G Wolynes. Topology, structures, and energy landscapes of human chromosomes. *Proceedings of the National Academy of Sciences*, 112(19):6062–6067, 2015.
- [15] Geoffrey Fudenberg, Maxim Imakaev, Carolyn Lu, Anton Goloborodko, Nezar Abdennur, and Leonid A Mirny. Formation of chromosomal domains by loop extrusion. *Cell reports*, 15(9):2038–2049, 2016.
- [16] Michele Di Pierro, Bin Zhang, Erez Lieberman Aiden, Peter G Wolynes, and José N Onuchic. Transferable model for chromosome architecture. *Proceedings of the National Academy of Sciences*, 113(43):12168–12173, 2016.
- [17] Johan H Gibcus, Kumiko Samejima, Anton Goloborodko, Itaru Samejima, Natalia Naumova, Johannes Nuebler, Masato T Kanemaki, Linfeng Xie, James R Paulson, William C Earnshaw, et al. A pathway for mitotic chromosome formation. *Science*, 359(6376):eaao6135, 2018.

- [18] Simona Bianco, Darío G Lupiáñez, Andrea M Chiariello, Carlo Annunziatella, Katerina Kraft, Robert Schöpflin, Lars Wittler, Guillaume Andrey, Martin Vingron, Ana Pombo, et al. Polymer physics predicts the effects of structural variants on chromatin architecture. *Nature genetics*, 50(5):662–667, 2018.
- [19] Johannes Nuebler, Geoffrey Fudenberg, Maxim Imakaev, Nezar Abdennur, and Leonid Mirny. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Biophysical Journal*, 114(3):30a, 2018.
- [20] Yifeng Qi and Bin Zhang. Predicting three-dimensional genome organization with chromatin states. *PLoS computational biology*, 15(6):e1007024, 2019.
- [21] Eric R Schultz, Jay L Kaplan, Yiheng Wu, Soren Kyhl, Rebecca Willett, and Juan J de Pablo. Current advances in genome modeling across length scales. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 15(3):e70024, 2025.
- [22] D Thirumalai, Guang Shi, Sucheol Shin, and Changbong Hyeon. Organization and dynamics of chromosomes. *Annual Review of Physical Chemistry*, 76, 2025.
- [23] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Židek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.
- [24] Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, 2021.
- [25] Sarah Lewis, Tim Hempel, José Jiménez-Luna, Michael Gastegger, Yu Xie, Andrew YK Foong, Victor García Satorras, Osama Abidin, Bastiaan S Veeling, Iryna Zaporozhets, et al. Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science*, page eadv9817, 2025.
- [26] Bowen Jing, Bonnie Berger, and Tommi Jaakkola. Alphafold meets flow matching for generating protein ensembles. *arXiv preprint arXiv:2402.04845*, 2024.
- [27] Tim J Stevens, David Lando, Srinjan Basu, Liam P Atkinson, Yang Cao, Steven F Lee, Martin Leeb, Kai J Wohlfahrt, Wayne Boucher, Aoife O’Shaughnessy-Kirwan, et al. 3d structures of individual mammalian genomes studied by single-cell hi-c. *Nature*, 544(7648):59–64, 2017.
- [28] Elizabeth H Finn, Gianluca Pegoraro, Hugo B Brandão, Anne-Laure Valton, Marlies E Oomen, Job Dekker, Leonid Mirny, and Tom Misteli. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell*, 176(6):1502–1515, 2019.
- [29] Erez Lieberman-Aiden, Nynke L Van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R Lajoie, Peter J Sabo, Michael O Dorschner, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950):289–293, 2009.
- [30] Suhas SP Rao, Miriam H Huntley, Neva C Durand, Elena K Stamenova, Ivan D Bochkov, James T Robinson, Adrian L Sanborn, Ido Machol, Arina D Omer, Eric S Lander, et al. A 3d map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7):1665–1680, 2014.
- [31] Tianming Zhou, Ruochi Zhang, and Jian Ma. The 3d genome structure of single cells. *Annual Review of Biomedical Data Science*, 4(1):21–41, 2021.
- [32] Kok Hao Chen, Alistair N Boettiger, Jeffrey R Moffitt, Siyuan Wang, and Xiaowei Zhuang. Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348(6233):aaa6090, 2015.
- [33] Jeffrey R Moffitt, Junjie Hao, Guiping Wang, Kok Hao Chen, Hazen P Babcock, and Xiaowei Zhuang. High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proceedings of the National Academy of Sciences*, 113(39):11046–11051, 2016.

- [34] Chenglong Xia, Jean Fan, George Emanuel, Junjie Hao, and Xiaowei Zhuang. Spatial transcriptome profiling by merfish reveals subcellular rna compartmentalization and cell cycle-dependent gene expression. *Proceedings of the National Academy of Sciences*, 116(39):19490–19499, 2019.
- [35] Greg Schuette, Zhuohan Lao, and Bin Zhang. Chromogen: Diffusion model predicts single-cell chromatin conformations. *Science Advances*, 11(5):eadr8265, 2025.
- [36] Zhenhao Zhang, Fan Feng, Yiyang Qiu, and Jie Liu. A generalizable framework to comprehensively predict epigenome, chromatin organization, and transcriptome. *Nucleic Acids Research*, 51(12):5931–5947, 2023.
- [37] Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219*, 2018.
- [38] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.
- [39] Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*, 2022.
- [40] Michael S Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *arXiv preprint arXiv:2303.08797*, 2023.
- [41] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.
- [42] Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *arXiv preprint arXiv:2302.00482*, 2023.
- [43] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pages 2256–2265. pmlr, 2015.
- [44] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.
- [45] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.
- [46] Wolfgang Kabsch. A solution for the best rotation to relate two sets of vectors. *Foundations of Crystallography*, 32(5):922–923, 1976.
- [47] Yaron Lipman, Marton Havasi, Peter Holderrieth, Neta Shaul, Matt Le, Brian Karrer, Ricky TQ Chen, David Lopez-Paz, Heli Ben-Hamu, and Itai Gat. Flow matching guide and code. *arXiv preprint arXiv:2412.06264*, 2024.
- [48] 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. *arXiv preprint arXiv:2302.02277*, 2023.
- [49] Jason Yim, Andrew Campbell, Andrew YK Foong, Michael Gastegger, José Jiménez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S Veeling, Regina Barzilay, Tommi Jaakkola, et al. Fast protein backbone generation with se (3) flow matching. *arXiv preprint arXiv:2310.05297*, 2023.
- [50] Leon Klein, Andreas Krämer, and Frank Noé. Equivariant flow matching. *Advances in Neural Information Processing Systems*, 36:59886–59910, 2023.
- [51] Victor Garcia Satorras, Emiel Hoogetboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pages 9323–9332. PMLR, 2021.

- [52] Victor Garcia Satorras, Emiel Hoogetboom, Fabian Fuchs, Ingmar Posner, and Max Welling. E (n) equivariant normalizing flows. *Advances in Neural Information Processing Systems*, 34:4181–4192, 2021.
- [53] Emiel Hoogetboom, Victor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pages 8867–8887. PMLR, 2022.
- [54] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [55] Zhuoran Qiao, Weili Nie, Arash Vahdat, Thomas F Miller III, and Animashree Anandkumar. State-specific protein–ligand complex structure prediction with a multiscale deep generative model. *Nature Machine Intelligence*, 6(2):195–208, 2024.
- [56] Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.
- [57] Ke Liu, Weian Mao, Shuaike Shen, Xiaoran Jiao, Zheng Sun, Hao Chen, and Chunhua Shen. Floating anchor diffusion model for multi-motif scaffolding. *arXiv preprint arXiv:2406.03141*, 2024.
- [58] Jeremy Wohlwend, Gabriele Corso, Saro Passaro, Noah Getz, Mateo Reveiz, Ken Leidal, Wojtek Swiderski, Liam Atkinson, Tally Portnoi, Itamar Chinn, et al. Boltz-1 democratizing biomolecular interaction modeling. *BioRxiv*, pages 2024–11, 2025.
- [59] Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.
- [60] Yaowei Jin, Qi Huang, Ziyang Song, Mingyue Zheng, Dan Teng, and Qian Shi. P2dflow: A protein ensemble generative model with se (3) flow matching. *Journal of Chemical Theory and Computation*, 21(6):3288–3296, 2025.
- [61] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999. PMLR, 2016.
- [62] Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco S Cohen. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. *Advances in Neural information processing systems*, 31, 2018.
- [63] Taco S Cohen, Mario Geiger, and Maurice Weiler. A general theory of equivariant cnns on homogeneous spaces. *Advances in neural information processing systems*, 32, 2019.
- [64] Minkai Xu, Alexander S Powers, Ron O Dror, Stefano Ermon, and Jure Leskovec. Geometric latent diffusion models for 3d molecule generation. In *International Conference on Machine Learning*, pages 38592–38610. PMLR, 2023.
- [65] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):2453, 2022.
- [66] Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. Se (3)-transformers: 3d roto-translation equivariant attention networks. *Advances in neural information processing systems*, 33:1970–1981, 2020.
- [67] Brandon Anderson, Truong Son Hy, and Risi Kondor. Cormorant: Covariant molecular neural networks. *Advances in neural information processing systems*, 32, 2019.
- [68] Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International conference on machine learning*, pages 9377–9388. PMLR, 2021.

- [69] Johannes Gasteiger, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems*, 34:6790–6802, 2021.
- [70] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.
- [71] Yi-Lun Liao and Tess Smidt. Equiformer: Equivariant graph attention transformer for 3d atomistic graphs. *arXiv preprint arXiv:2206.11990*, 2022.
- [72] Philipp Thölke and Gianni De Fabritiis. Torchmd-net: equivariant transformers for neural network based molecular potentials. *arXiv preprint arXiv:2202.02541*, 2022.
- [73] Weian Mao, Muzhi Zhu, Zheng Sun, Shuaike Shen, Lin Yuanbo Wu, Hao Chen, and Chunhua Shen. De novo protein design using geometric vector field networks. *arXiv preprint arXiv:2310.11802*, 2023.
- [74] Hongwei Tu, Yanqiang Han, Zhilong Wang, An Chen, Kehao Tao, Simin Ye, Shiwei Wang, Zhiyun Wei, and Jinjin Li. Rotnet: a rotationally invariant graph neural network for quantum mechanical calculations. *Small Methods*, 8(1):2300534, 2024.
- [75] Job Dekker, Marc A Marti-Renom, and Leonid A Mirny. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews Genetics*, 14(6):390–403, 2013.
- [76] Boyan Bonev and Giacomo Cavalli. Organization and function of the 3d genome. *Nature Reviews Genetics*, 17(11):661–678, 2016.
- [77] Tom Misteli. The self-organizing genome: principles of genome architecture and function. *Cell*, 183(1):28–45, 2020.
- [78] Mariano Barbieri, Mita Chotalia, James Fraser, Liron-Mark Lavitas, Josée Dostie, Ana Pombo, and Mario Nicodemi. Complexity of chromatin folding is captured by the strings and binders switch model. *Proceedings of the National Academy of Sciences*, 109(40):16173–16178, 2012.
- [79] Kartik Kamat, Zhuohan Lao, Yifeng Qi, Yuchuan Wang, Jian Ma, and Bin Zhang. Compartmentalization with nuclear landmarks yields random, yet precise, genome organization. *Biophysical Journal*, 122(7):1376–1389, 2023.
- [80] Lao Zhuohan, Kartik D Kamat, Jiang Zhongling, and Bin Zhang. Opennucleome for high-resolution nuclear structural and dynamical modeling. *eLife*, 13, 2024.
- [81] Davide Baù, Amartya Sanyal, Bryan R Lajoie, Emidio Capriotti, Meg Byron, Jeanne B Lawrence, Job Dekker, and Marc A Marti-Renom. The three-dimensional folding of the α -globin gene domain reveals formation of chromatin globules. *Nature structural & molecular biology*, 18(1):107–114, 2011.
- [82] Ruochi Zhang, Tianming Zhou, and Jian Ma. Multiscale and integrative single-cell hi-c analysis with higashi. *Nature biotechnology*, 40(2):254–261, 2022.
- [83] Lorenzo Boninsegna, Asli Yildirim, Guido Polles, Yuxiang Zhan, Sofia A Quinodoz, Elizabeth H Finn, Mitchell Guttman, Xianghong Jasmine Zhou, and Frank Alber. Integrative genome modeling platform reveals essentiality of rare contact events in 3d genome organizations. *Nature methods*, 19(8):938–949, 2022.
- [84] ENCODE Project Consortium et al. An integrated encyclopedia of dna elements in the human genome. *Nature*, 489(7414):57, 2012.
- [85] Jesse R Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–380, 2012.

- [86] Matthias Merkenschlager and Elphège P Nora. Ctf and cohesin in genome folding and transcriptional gene regulation. *Annual review of genomics and human genetics*, 17(1):17–43, 2016.
- [87] Stefan Schoenfelder, Tom Sexton, Lyubomira Chakalova, Nathan F Cope, Alice Horton, Simon Andrews, Sreenivasulu Kurukuti, Jennifer A Mitchell, David Umlauf, Daniela S Dimitrova, et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nature genetics*, 42(1):53–61, 2010.
- [88] Anshul Kundaje, Wouter Meuleman, Jason Ernst, Misha Bilenky, Angela Yen, Pouya Kheradpour, Zhizhuo Zhang, Alireza Heravi-Moussavi, Yaping Liu, Viren Amin, et al. Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539):317, 2015.
- [89] Jason D Buenrostro, Paul G Giresi, Lisa C Zaba, Howard Y Chang, and William J Greenleaf. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, dna-binding proteins and nucleosome position. *Nature methods*, 10(12):1213–1218, 2013.
- [90] Jason D Buenrostro, Beijing Wu, Ulrike M Litzenburger, Dave Ruff, Michael L Gonzales, Michael P Snyder, Howard Y Chang, and William J Greenleaf. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, 523(7561):486–490, 2015.
- [91] Alan P Boyle, Sean Davis, Hennady P Shulha, Paul Meltzer, Elliott H Margulies, Zhiping Weng, Terrence S Furey, and Gregory E Crawford. High-resolution mapping and characterization of open chromatin across the genome. *Cell*, 132(2):311–322, 2008.
- [92] Robert E Thurman, Eric Rynes, Richard Humbert, Jeff Vierstra, Matthew T Maurano, Eric Haugen, Nathan C Sheffield, Andrew B Stergachis, Hao Wang, Benjamin Vernot, et al. The accessible chromatin landscape of the human genome. *Nature*, 489(7414):75–82, 2012.
- [93] Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.
- [94] Divya Nori and Wengong Jin. Rnaflow: Rna structure & sequence design via inverse folding-based flow matching. *arXiv preprint arXiv:2405.18768*, 2024.
- [95] Jason Yim, Andrew Campbell, Emile Mathieu, Andrew YK Foong, Michael Gastegger, José Jiménez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S Veeling, Frank Noé, et al. Improved motif-scaffolding with se (3) flow matching. *ArXiv*, pages arXiv–2401, 2024.
- [96] Emily Crane, Qian Bian, Rachel Patton McCord, Bryan R Lajoie, Bayly S Wheeler, Edward J Ralston, Satoru Uzawa, Job Dekker, and Barbara J Meyer. Condensin-driven remodelling of x chromosome topology during dosage compensation. *Nature*, 523(7559):240–244, 2015.
- [97] Open2C, Nezar Abdennur, Sameer Abraham, Geoffrey Fudenberg, Ilya M Flyamer, Aleksandra A Galitsyna, Anton Goloborodko, Maxim Imakaev, Betul A Oksuz, Sergey V Venev, et al. Cooltools: enabling high-resolution hi-c analysis in python. *PLOS Computational Biology*, 20(5):e1012067, 2024.
- [98] Tao Yang, Feipeng Zhang, Galip Gürkan Yardımcı, Fan Song, Ross C Hardison, William Stafford Noble, Feng Yue, and Qunhua Li. Hicrep: assessing the reproducibility of hic data using a stratum-adjusted correlation coefficient. *Genome research*, 27(11):1939–1949, 2017.

A Related Work

Our work is situated at the intersection of deep generative models for 3D structures and equivariant neural networks, with a specific application to 3D genome organization.

Generative models for 3D molecular structures. The field of 3D molecular structure modeling has advanced rapidly since the breakthroughs of AlphaFold2 [23] and RosettaFold [24], which established attention-based architectures [54] as powerful tools for protein folding. Building on these foundations, diffusion- and flow-based generative methods [44, 45, 38, 41, 39] emerged as a second stage, combining stable training with comprehensive mode coverage to enable high-fidelity, diverse generation of complex molecular geometries. These methods provide systematic approaches for modeling molecular conformations and binding poses [55, 56, 57, 58], and have been successfully applied to de novo protein design, such as RFDiffusion [59]. While such approaches can generate multiple plausible conformations, they are often anchored to experimentally determined structures and therefore do not capture the full range of intermediate states beyond experimental measurements. More recently, a third stage of models has extended this framework by incorporating molecular dynamics trajectories [26, 60, 25], enabling the modeling of structural ensembles and distributions that include transient and intermediate states. Together, these advances mark a transition from predicting isolated conformations to modeling full distributions of biologically relevant structures.

Equivariant neural networks. Symmetry is an intrinsic property of physical systems: when a system is translated or rotated, its physical observables transform in a well-defined manner. Equivariant neural networks encode these transformation rules directly into their architectures, ensuring physically consistent outputs while reducing the need for explicit data augmentation, improving data efficiency. Early works on group-equivariant and steerable convolutional neural networks (CNNs) [61, 62, 63] established the foundational framework of equivariant models. Tensor field networks (TFNs) were an early breakthrough, introducing tensor representations to achieve rotational equivariance on point clouds. Equivariant graph neural networks (EGNNs) [51] further simplified the design, enabling broad application in molecular and geometric learning tasks [52, 53, 64]. In molecular dynamics, NequIP [65] marked a major leap by replacing handcrafted functions of invariant descriptors with an end-to-end equivariant force field of unprecedented accuracy. Many other architectures have extended these ideas and focused on scalability and richer interactions [66, 67, 68, 69, 70, 71, 72, 73, 74]. Taken together, these advances trace a clear trajectory from theory to state-of-the-art performance, with future efforts focused on simpler, more efficient representations that scale to large systems and extend to multimodal domains.

Computational modeling of 3D genome organization. Chromosome organization is intrinsically dynamic and complex, involving characteristic modes such as loop extrusion, domain boundaries, and compartmental segregation that shape transcription, regulation, and replication [75, 76, 2, 77, 3, 5]. The advent of Hi-C provided the first genome-wide maps of chromatin contacts [29], establishing the foundation for subsequent computational modeling. Early efforts combined polymer physics with experimental constraints or molecular dynamics to reproduce key principles [78, 13, 14, 15, 16, 17, 18, 19, 20, 79, 80], but these simulations often underrepresent the heterogeneous and complex influences of sequence, epigenome, proteins, and RNAs. Therefore, data-driven approaches have emerged to better incorporate diverse and biologically relevant factors. These methods frame 3D genome organization as a statistical modeling task, learning probabilistic ensembles of structures or population-level contact patterns directly from genomic features and high-throughput assays [81, 82, 4]. Direct modeling of 3D structures provides more extensive knowledge of chromosome organization than 2D contact maps. While data-driven energy-based simulations [14, 16, 20] and integrative genome modeling platforms infer structural ensembles from experimental data [83], the computational cost of iterative sampling and optimization becomes prohibitive as system size increases. Looking ahead, unifying diverse data modalities with physics-informed priors in dynamic probabilistic models and integrating them with modern generative frameworks, for example, applications such as ChromoGen [35], offers a promising path toward more comprehensive and mechanistic descriptions of chromosome organization at large scales.

B Supplementary information

In this section, we provide additional details for the datasets, methods, and results.

B.1 Data preparation

Chromosome 3D structures We use high-resolution MERFISH data of chromosome 2 (Chr2) from [9]. Each MERFISH probe covers 50kb. However, due to gaps between probes, we represent each bead as a 250 kb segment. The bead position is determined by the centroid of its central 50 kb sub-segment. If the central 50kb region is not probed, we label the bead as “unknown”, indicating that while the bead physically exists, its 3D position is experimentally unresolved.

We used Chr2 region 0.15-241.9 Mb (hg38 assembly), which yields 967 beads in all (each bead representing 250 kb). To normalize bond lengths, all 3D positions are scaled by the average bond length computed from the available bonds in the training set. For Chr2, the training set average bond length is 630.88 nm. Unless otherwise specified, reported distances are unitless and expressed relative to this average bond length. The Chr2 MERFISH dataset is divided into training, validation, and test sets with fractions of 80%, 10%, 10%, respectively.

Meanwhile, we observed that a small fraction of beads in the experimental structures are connected by abnormally long bonds. These long bonds may not reflect true polymer conformations, but instead arise from experimental limitations such as probe dropout, low fluorescence signal, or displacement errors during repeated imaging (as discussed in [9]). Such artifacts can introduce artificially stretched neighboring bead distances (i.e., bond lengths).

Raw genomic features Since the MERFISH Chr2 data was measured in IMR90 cells, we downloaded processed ENCODE files for IMR90 using GRCh38 as the reference genome [84]. We treated the values from these processed files as raw genomic features for downstream modeling. Specifically, we included the following features: CTCF, RAD21, SMC3, POLR2A, H3K27Ac, H3K4me3, H3K36me3, H3K9me3, H3K4me1, H3K27me3, DNase-seq, and ATAC-seq. CTCF and cohesin subunits (RAD21, SMC3) play crucial roles in establishing topologically associating domains (TADs) and loop extrusion [85, 86]. Active transcriptional machinery (POLR2A) reflects sites of high transcriptional activity, which often colocalize with accessible chromatin and contribute to the spatial clustering of transcriptionally active regions (“transcription factories”) [87, 85]. Histone modifications further distinguish functional chromatin states: active marks (e.g., H3K27ac, H3K4me3, H3K36me3) highlight regulatory elements and euchromatic compartments, whereas repressive marks (e.g., H3K9me3, H3K27me3) are enriched in heterochromatic regions [88]. Because chromatin of similar type tends to spatially associate [29, 30], these euchromatin and heterochromatin markers are highly informative for 3D genome organization. Accessibility assays (DNase-seq, ATAC-seq) further capture open chromatin regions that often coincide with regulatory hubs [89, 90, 91, 92]. Together, these features are closely associated with genome organization and are widely available, suggesting that our model may generalize to other chromosomes if the same features are provided. Future ablation studies can further identify key factors that contribute most to chromosome organization at the target scale.

Batch loader When training, we sample subchains of lengths uniformly from range [96, 128]. Since samples of varying lengths are concatenated in one batch, a valid mask of size (B, N_{\max}) is constructed, where B is batch size and N_{\max} is the longest chain length in the batch. This valid mask distinguishes valid beads from padded ghost beads. Note unknown beads are still valid and labeled as True or 1 in the valid mask.

B.2 Method details

Overview We first describe the model components (encoder and decoder) with their pseudocode (Algorithms 1, 2, 3, 4, 5). We then specify the flow-matching objective, RMSD alignment strategy, and loss. Finally, we present the training and inference procedures (Algorithms 6, 7). Masks for padded beads are applied in batched settings but omitted in pseudocode for brevity.

B.2.1 Model components

Encoder: ESM folding block The encoder transforms raw genomic features and relative index into epigenomic (\mathbf{h}) and pair (\mathbf{z}) embeddings. With pair-biased attention (Algorithm 1) and triangular modules (introduced in AlphaFold [23], with slight modifications), the ESM folding block (Algorithm 2) refines inputs to achieve representative 1D epigenomic and 2D pair embeddings.

Algorithm 1 Pair-biased attention

Require: Input epigenomic embedding \mathbf{h} , pair embedding \mathbf{z} , bool variables `agg-pair`, `use-gate`

$$\mathbf{y} \leftarrow \text{LayerNorm}(\mathbf{h}) \quad \triangleright \mathbf{y}, \mathbf{h} \in \mathbb{R}^{L \times d_{\text{seq}}}$$

$$\mathbf{q}^i, \mathbf{k}^i, \mathbf{v}^i \leftarrow \text{Reshape}(\text{LinearNoBias}(\mathbf{y})) \quad \triangleright \mathbf{q}^i, \mathbf{k}^i, \mathbf{v}^i \in \mathbb{R}^{L \times d_{\text{head}}}, i \in \{1, \dots, n_{\text{heads}}\}$$

$$\mathbf{b}_{\text{pair}}^i \leftarrow \text{LinearNoBias}(\mathbf{z}) \quad \triangleright \mathbf{b}_{\text{pair}}^i \in \mathbb{R}^{L \times L}$$

$$\mathbf{A}^i \leftarrow \mathbf{q}^i \cdot (\mathbf{k}^i)^\top / \sqrt{d_{\text{head}}} + \mathbf{b}_{\text{pair}}^i \quad \triangleright \mathbf{A}^i \in \mathbb{R}^{L \times L}$$

$$\mathbf{W}^i \leftarrow \text{Softmax}(\mathbf{A}^i) \quad \triangleright \mathbf{W}^i \in \mathbb{R}^{L \times L}$$

$$\mathbf{v}_{\text{agg}}^i \leftarrow \mathbf{W}^i \mathbf{v}^i \quad \triangleright \mathbf{v}_{\text{agg}}^i \in \mathbb{R}^{L \times d_{\text{head}}}$$

if `agg-pair` **then**

$$d_{\text{context}} = n_{\text{heads}} \cdot (d_{\text{head}} + d_{\text{pair}})$$

$$(\mathbf{z}_{\text{agg}}^i)_{jd} = \sum_k \mathbf{W}_{jk}^i \cdot \mathbf{z}_{jkd} \quad \triangleright \mathbf{z} \in \mathbb{R}^{L \times L \times d_{\text{pair}}}, \mathbf{z}_{\text{agg}}^i \in \mathbb{R}^{L \times d_{\text{pair}}}$$

$$\mathbf{c} \leftarrow \text{Concat}[\mathbf{v}_{\text{agg}}^1, \dots, \mathbf{v}_{\text{agg}}^{n_{\text{heads}}}, \mathbf{z}_{\text{agg}}^1, \dots, \mathbf{z}_{\text{agg}}^{n_{\text{heads}}}] \quad \triangleright \mathbf{c} \in \mathbb{R}^{L \times d_{\text{context}}}$$

else

$$d_{\text{context}} = n_{\text{heads}} \cdot d_{\text{head}}$$

$$\mathbf{c} \leftarrow \text{Concat}[\mathbf{v}_{\text{agg}}^1, \dots, \mathbf{v}_{\text{agg}}^{n_{\text{heads}}}] \quad \triangleright \mathbf{c} \in \mathbb{R}^{L \times d_{\text{context}}}$$

end if

if `use-gate` **then**

$$\mathbf{g} \leftarrow \text{Sigmoid}(\text{Linear}(\mathbf{y})) \quad \triangleright \mathbf{g} \in \mathbb{R}^{L \times d_{\text{context}}}$$

$$\mathbf{c} \leftarrow \mathbf{g} \odot \mathbf{c}$$

end if

$$\mathbf{h}_{\text{out}} \leftarrow \text{Linear}(\mathbf{c}) \quad \triangleright \mathbf{h}_{\text{out}} \in \mathbb{R}^{L \times d_{\text{seq}}}$$

return \mathbf{h}_{out}

Algorithm 2 ESM folding block, following ESMFold [93] with some modifications

Require: Epigenomic embedding \mathbf{h} , pair embedding \mathbf{z} , `PairBiasedAttention` with `agg-pair` = `False`, `use-gate` = `True`

$$\mathbf{h} \leftarrow \mathbf{h} + \text{Dropout}_{0.1}(\text{PairBiasedAttention}(\mathbf{h}, \mathbf{z}))$$

$$\mathbf{h} \leftarrow \mathbf{h} + \text{Dropout}_{0.1}(\text{Linear}(\text{ReLU}(\text{Linear}(\text{LayerNorm}(\mathbf{h}))))))$$

$$\mathbf{q}, \mathbf{k} \leftarrow \text{Linear}(\text{LayerNorm}(\mathbf{h}))$$

$$\mathbf{z}_{ij} \leftarrow \mathbf{z}_{ij} + \text{Linear}(\text{Concat}[\mathbf{q}_i \odot \mathbf{k}_j, \mathbf{q}_i - \mathbf{k}_j]) \quad \triangleright i, j \in \{1, \dots, L\}$$

$$\mathbf{z} \leftarrow \mathbf{z} + \text{DropoutRowwise}_{0.2}(\text{TriangleMultiplicationOutgoing}(\mathbf{z}))$$

$$\mathbf{z} \leftarrow \mathbf{z} + \text{DropoutRowwise}_{0.2}(\text{TriangleMultiplicationIncoming}(\mathbf{z})) \quad \triangleright \text{Here ESMFold}$$

applies `DropoutColumnwise`, but we use `DropoutRowwise` as AlphaFold [23] does

$$\mathbf{z} \leftarrow \mathbf{z} + \text{DropoutRowwise}_{0.2}(\text{TriangleAttentionStartingNode}(\mathbf{z}))$$

$$\mathbf{z} \leftarrow \mathbf{z} + \text{DropoutColumnwise}_{0.2}(\text{TriangleAttentionEndingNode}(\mathbf{z}))$$

$$\mathbf{z} \leftarrow \mathbf{z} + \text{Dropout}_{0.1}(\text{Linear}(\text{ReLU}(\text{Linear}(\text{LayerNorm}(\mathbf{z}))))))$$

return \mathbf{h}, \mathbf{z}

Decoder: Attention-EGNN Given epigenomic and pair embeddings (i.e. \mathbf{h} and \mathbf{z}) produced by the encoder, as well as time t and input positions \mathbf{r} , the embeddings (Algorithm 3) and positions are updated iteratively with attention-based EGNN network (Algorithm 4, 5). Inspired by the invariant point attention (IPA) algorithm in AlphaFold [23], the network architecture of bioemu [25] and FrameDiff [48], this framework applies the spatial pair distance information as attention bias, updates epigenomic and pair embeddings in a graph network style, and finally use equivariant update [51] for positions to preserve 3D symmetry.

Algorithm 3 Fourier flow time encoding, inspired by [54, 25]

Require: $t \in [0, 1]$, model size d_{model}

$$\omega_i = 10000^{-2i/d_{\text{model}}}, i \in \{1, \dots, n\}, n = d_{\text{model}}/2 \quad \triangleright d_{\text{model}} \text{ is even}$$

$$\omega = (\omega_1, \dots, \omega_n)$$

$$\phi = 1000 \cdot t \cdot \omega$$

return $\text{Concat}(\sin(\phi), \cos(\phi))$

Algorithm 4 Attention-EGNN block

Require: Epigenomic embedding \mathbf{h} , pair embedding \mathbf{z} , positions \mathbf{r} , PairBiasedAttention with `agg-pair = True, use-gate = False`

Update epigenomic and pair embeddings:

$$\mathbf{z}_{ij} \leftarrow \mathbf{z}_{ij} + \text{DistanceEncoding}(|\mathbf{r}_i - \mathbf{r}_j|) \quad \triangleright i, j \in \{1, \dots, L\}$$

$$\mathbf{h} \leftarrow \mathbf{h} + \text{Dropout}_{0.1}(\text{PairBiasedAttention}(\mathbf{h}, \mathbf{z}))$$

$$\mathbf{h} \leftarrow \mathbf{h} + \text{Dropout}_{0.1}(\text{Linear}(\text{Dropout}_{0.1}(\text{GELU}(\text{Linear}(\text{LayerNorm}(\mathbf{h}))))))$$

$$\mathbf{h}' \leftarrow \text{Linear}(\mathbf{h}) \quad \triangleright \mathbf{h}' \in \mathbb{R}^{L \times (d_{\text{seq}}/2)}, d_{\text{seq}} \text{ is even}$$

$$\mathbf{m}_{ij} \leftarrow \text{Concat}[\mathbf{h}'_i, \mathbf{h}'_j, \mathbf{z}_{ij}] \quad \triangleright \mathbf{m}_{ij} \in \mathbb{R}^{d_{\text{seq}} + d_{\text{pair}}}$$

$$\mathbf{m} \leftarrow \mathbf{m} + \text{SiLU}(\text{Linear}(\text{SiLU}(\text{Linear}(\text{LayerNorm}(\mathbf{m}))))))$$

$$\mathbf{z} \leftarrow \text{Linear}(\mathbf{m})$$

Equivariant update of positions:

$$\mathbf{r}'_{ij} \leftarrow (\mathbf{r}_i - \mathbf{r}_j) / (|\mathbf{r}_i - \mathbf{r}_j| + \Delta) \quad \triangleright \Delta = 1$$

$$\mathbf{p}_{ij} \leftarrow \text{Concat}[\mathbf{h}_i, \mathbf{h}_j, \mathbf{z}_{ij}]$$

$$\mathbf{w}_{ij} \leftarrow \tanh(\text{Linear}(\text{SiLU}(\text{Linear}(\text{SiLU}(\text{Linear}(\text{LayerNorm}(\mathbf{p}_{ij}))))))) \quad \triangleright \mathbf{w}_{ij} \in [-1, 1]$$

$$\mathbf{r}_i \leftarrow \mathbf{r}_i + \sum_j \mathbf{w}_{ij} \mathbf{r}'_{ij}$$

return $\mathbf{h}, \mathbf{r}, \mathbf{z}$

Algorithm 5 Attention-EGNN

Require: Epigenomic embedding \mathbf{h} , positions \mathbf{r} , pair embedding \mathbf{z} , time t

$$\mathbf{h} \leftarrow \text{Linear}(\text{LayerNorm}(\mathbf{h})) + \text{TimeEncoding}(t) \quad \triangleright \text{TimeEncoding is Algorithm 3}$$

$$\mathbf{z} \leftarrow \text{LinearNoBias}(\text{LayerNorm}(\mathbf{z}))$$

for i in $1, \dots, n_{\text{blocks}}^{\text{decoder}}$ **do**

$$\mathbf{h}, \mathbf{r}, \mathbf{z} \leftarrow \text{AttentionEGNNBlock}_i(\mathbf{h}, \mathbf{r}, \mathbf{z})$$

end for

return $\mathbf{h}, \mathbf{r}, \mathbf{z}$

B.2.2 RMSD alignment and loss

Flow matching with RMSD alignment We apply flow matching (FM) [38, 39, 40, 41] to train a time-dependent velocity field model $\hat{\mathbf{v}}_t(\mathbf{r}, \theta)$, where θ denotes training parameters, which defines a

mapping between two distributions. A general overview of FM is provided in Section 2.2. In our setting, we must handle 3D symmetries without being affected by arbitrary global translations or rotations. To address this, we perform RMSD alignment with the Kabsch algorithm [46] during both training and inference. Existing studies have proved the effectiveness of applying certain alignment step to preserve correct symmetry and facilitate efficient transformation [50, 26, 58, 94, 60]. Following AlphaFlow [26]: During training we align \mathbf{r}_0 to \mathbf{r}_1 before interpolation. During inference, we align $\hat{\mathbf{r}}_t^{\text{ODE}}$ to $\hat{\mathbf{r}}_1$ before Euler step integration. Importantly, we establish the following result:

Theorem 1. *If \mathbf{r}_0 is RMSD-aligned to \mathbf{r}_1 , then the linear interpolation $\mathbf{r}_t = (1-t)\mathbf{r}_0 + t\mathbf{r}_1$ ($t \in [0, 1]$) remains RMSD-aligned to \mathbf{r}_1 .*

This theorem ensures consistency between training and inference: during training, \mathbf{r}_0 is aligned to \mathbf{r}_1 , so the interpolation \mathbf{r}_t preserves alignment, thereby allowing for the alignment of $\mathbf{r}_t^{\text{ODE}}$ to $\hat{\mathbf{r}}_1$ during inference.

Proof. Suppose \mathbf{r}_t ($t \in [0, 1]$) $\in \mathbb{R}^{N \times 3}$ is the matrix of interpolated coordinates. The Kabsch algorithm [46] first translates both point sets \mathbf{r} and \mathbf{r}_1 so their centroids coincide at the origin, then finds the optimal rotation \mathbf{R}^* to minimize $\|\mathbf{r}\mathbf{R} - \mathbf{r}_1\|_F^2$ (here \mathbf{r} and \mathbf{r}_1 are already centered at the origin).

$$\|\mathbf{r}\mathbf{R} - \mathbf{r}_1\|_F^2 = \text{tr}(\mathbf{r}^\top \mathbf{r}) + \text{tr}(\mathbf{r}_1^\top \mathbf{r}_1) - 2\text{tr}(\mathbf{R}^\top \mathbf{r}^\top \mathbf{r}_1). \quad (8)$$

Therefore, minimizing $\|\mathbf{r}\mathbf{R} - \mathbf{r}_1\|_F^2$ over all rotations \mathbf{R} is equivalent to maximizing $\text{tr}(\mathbf{R}^\top \mathbf{r}^\top \mathbf{r}_1)$.

Since \mathbf{r}_0 is already aligned to \mathbf{r}_1 , thus \mathbf{r}_0 and \mathbf{r}_1 are already centered at the origin. The linear interpolation $\mathbf{r}_t = (1-t)\mathbf{r}_0 + t\mathbf{r}_1$ is also centered at the origin. To show that \mathbf{r}_t is already aligned to \mathbf{r}_1 , we prove that identity rotation $\mathbf{R} = \mathbf{I}$ maximizes $\text{tr}(\mathbf{R}^\top \mathbf{r}_t^\top \mathbf{r}_1)$. We compute

$$\text{tr}(\mathbf{R}^\top \mathbf{r}_t^\top \mathbf{r}_1) = (1-t)\text{tr}(\mathbf{R}^\top \mathbf{r}_0^\top \mathbf{r}_1) + t\text{tr}(\mathbf{R}^\top \mathbf{r}_1^\top \mathbf{r}_1). \quad (9)$$

Since \mathbf{r}_0 is already aligned to \mathbf{r}_1 , $\text{tr}(\mathbf{R}^\top \mathbf{r}_0^\top \mathbf{r}_1) \leq \text{tr}(\mathbf{r}_0^\top \mathbf{r}_1)$ with equality at identity rotation $\mathbf{R} = \mathbf{I}$.

To show that $\text{tr}(\mathbf{R}^\top \mathbf{r}_1^\top \mathbf{r}_1) \leq \text{tr}(\mathbf{r}_1^\top \mathbf{r}_1)$, as $\mathbf{r}_1^\top \mathbf{r}_1$ is symmetric positive semidefinite, its eigendecomposition is $\mathbf{r}_1^\top \mathbf{r}_1 = \mathbf{Q}^\top \mathbf{D} \mathbf{Q}$, where \mathbf{D} is diagonal with non-negative diagonal elements, and \mathbf{Q} is orthogonal. Define $\mathbf{S} = \mathbf{Q} \mathbf{R} \mathbf{Q}^\top$, then \mathbf{S} is orthogonal, and

$$\text{tr}(\mathbf{R}^\top \mathbf{r}_1^\top \mathbf{r}_1) = \text{tr}(\mathbf{R}^\top \mathbf{Q}^\top \mathbf{D} \mathbf{Q}). \quad (10)$$

Use the property that $\text{tr}(\mathbf{A} \mathbf{B}) = \text{tr}(\mathbf{B} \mathbf{A})$, we get

$$\text{tr}(\mathbf{R}^\top \mathbf{Q}^\top \mathbf{D} \mathbf{Q}) = \text{tr}(\mathbf{Q} \mathbf{R}^\top \mathbf{Q}^\top \mathbf{D}) = \text{tr}(\mathbf{S}^\top \mathbf{D}) = \sum_{i=1}^3 \mathbf{S}_{ii} \mathbf{D}_{ii}. \quad (11)$$

Since \mathbf{S} is orthogonal, its diagonal elements satisfy $|\mathbf{S}_{ii}| \leq 1$. Thus,

$$\text{tr}(\mathbf{R}^\top \mathbf{r}_1^\top \mathbf{r}_1) = \sum_{i=1}^3 \mathbf{S}_{ii} \mathbf{D}_{ii} \leq \sum_{i=1}^3 \mathbf{D}_{ii} = \text{tr}(\mathbf{D}) = \text{tr}(\mathbf{Q} \mathbf{Q}^\top \mathbf{D}) = \text{tr}(\mathbf{Q}^\top \mathbf{D} \mathbf{Q}) = \text{tr}(\mathbf{r}_1^\top \mathbf{r}_1). \quad (12)$$

The upper bound is achieved when $\mathbf{S} = \mathbf{I} \Leftrightarrow \mathbf{R} = \mathbf{I}$. Therefore, $\text{tr}(\mathbf{R}^\top \mathbf{r}_t^\top \mathbf{r}_1)$ achieves maximum when $\mathbf{R} = \mathbf{I}$, so \mathbf{r}_t is aligned to \mathbf{r}_1 . \square

Loss function The canonical loss function for flow matching is the CFM loss (Equation 3). As noted in Section 2.2, the CFM loss can be reparameterized as position-wise loss \mathcal{L}_{pos} , training the model to predict a clean structure $\hat{\mathbf{r}}_1$ close to \mathbf{r}_1 given \mathbf{r}_t and t . To better capture biologically meaningful contacts, we introduce an additional auxiliary pairwise loss $\mathcal{L}_{\text{pair}}$. The overall loss function $\mathcal{L}(\hat{\mathbf{r}}_1, \mathbf{r}_1)$ is a weighted combination of the two loss terms:

$$\mathcal{L} = \beta \mathcal{L}_{\text{pos}}(\hat{\mathbf{r}}_1, \mathbf{r}_1) + (1 - \beta) \mathcal{L}_{\text{pair}}(\hat{\mathbf{r}}_1, \mathbf{r}_1), \quad (13)$$

where β is a hyperparameter that balances their relative contributions. The position-wise loss \mathcal{L}_{pos} is the primary loss, while the auxiliary pairwise loss underscores the important pair contacts, which are closely relevant to genomic functions. Similar auxiliary losses have also been applied in protein generative models to promote physical backbone configurations and essential inter-residue contacts [48, 95].

To describe the position-wise loss in detail, we first classify beads into 3 categories:

1. **Unknown beads:** beads whose positions are not experimentally measured. They are excluded from the loss, although the model still predicts their coordinates.
2. **Not well-bonded beads:** known beads connected with abnormally long bonds (> 5.0), likely reflecting experimental artifacts.
3. **Well-bonded beads:** all other beads.

The position-wise term is defined as

$$w_i = \begin{cases} 0, & \text{unknown bead,} \\ 0.2, & \text{not well-bonded bead,} \\ 1, & \text{otherwise.} \end{cases} \quad (14)$$

$$\mathcal{L}_{\text{pos}} = \frac{\sum_i w_i \|\hat{\mathbf{r}}_1 - \mathbf{r}_1\|^2}{\sum_i \mathbf{1}[w_i > 0]}. \quad (15)$$

Here $\mathbf{1}[\cdot]$ denotes the indicator function, which equals 1 if the condition is true and 0 otherwise. Thus, $\sum_i \mathbf{1}[w_i > 0]$ counts how many beads contribute to the position-wise loss. Since \mathbf{r}_1 is aligned to $\hat{\mathbf{r}}_1$ to eliminate arbitrary translation or rotation, the position-wise loss is well-defined.

Bead pairs are categorized similarly:

1. **Unknown pairs:** any pair involving at least one unknown bead.
2. **Bad pairs:** any known pair involving at least one not well-bonded bead.
3. **Good pairs:** all other known pairs.

The pairwise loss is defined as

$$\mathcal{L}_{\text{pair}} = \frac{\sum_{i < j} w_{ij} (\hat{d}_{ij} - d_{ij})^2}{\sum_{i < j} \mathbf{1}[w_{ij} > 0]}, \quad (16)$$

where \hat{d}_{ij} means the predicted pair distance in $\hat{\mathbf{r}}_1$, and d_{ij} is the corresponding distance in ground-truth \mathbf{r}_1 . The weight factor w_{ij} is given by

$$\begin{aligned} w_{ij} = & \mathbf{1}[|i - j| \leq n_{\text{local}} \text{ or } d_{ij} \leq d_{\text{cutoff}}] \\ & \cdot \mathbf{1}[\text{Pair}(i, j) \text{ is known}] \\ & \cdot (0.2 + 0.8 \cdot \mathbf{1}[\text{Pair}(i, j) \text{ is good}]) \end{aligned} \quad (17)$$

Thus, only known pairs (i, j) that are sequentially close ($|i - j| \leq n_{\text{local}}$) or spatially close ($d_{ij} \leq d_{\text{cutoff}}$) contribute to the loss. Bad pairs are down-weighted by a factor of 0.2. The denominator $\sum_{i < j} \mathbf{1}[w_{ij} > 0]$ counts the number of contributing pairs. The best performance was achieved when the noise distribution p_0 was Gaussian and $\beta = 0.98$. This corresponds to the results reported in the paper.

B.2.3 Training and inference

Algorithms 6 and 7 summarize the detailed training and inference procedures. The hyperparameters of the model reported in the main text are: $d_{\text{seq}}^{\text{in}} = 12, d_{\text{seq}} = d_{\text{pair}} = 80, d_{\text{hidden}} = 160, n_{\text{heads}} = 8, n_{\text{blocks}}^{\text{encoder}} = 8, n_{\text{local}} = 10, d_{\text{cutoff}} = 1.0$. Inspired by [26], the time sampling schedule \mathcal{T} is defined as follows: with probability 0.2, $t = 0$; otherwise, t is sampled uniformly from $[0, 1]$.

B.3 Analysis

Here we introduce the metrics and methods used for analysis.

Radius of gyration By viewing each bead with equal weight, the radius of gyration (R_g) is defined as

$$R_g = \sqrt{\frac{\sum_{i=1}^N \|\mathbf{r}_i - \bar{\mathbf{r}}\|^2}{N}}, \quad \bar{\mathbf{r}} = \frac{\sum_{i=1}^N \mathbf{r}_i}{N}. \quad (18)$$

Here N is the total number of beads. R_g captures the global size of the conformation.

Algorithm 6 Training algorithm

Require: Raw epigenomic features of full sequence \mathbf{x} , full chain structures $\{\mathbf{R}\}$

for all $\{\mathbf{R}\}$ **do**

Preprocessing:

Uniformly sample subchain length $L \sim \mathcal{U}\{96, 97, \dots, 128\}$

Uniformly sample continuous subchain indices S with given length L

Get subchain positions: $\mathbf{r}_1 \leftarrow \mathbf{R}[S]$

$\triangleright \mathbf{r}_1 \in \mathbb{R}^{L \times 3}$

Get subchain raw genomic features: $\tilde{\mathbf{x}} \leftarrow \mathbf{x}[S]$

$\triangleright \tilde{\mathbf{x}} \in \mathbb{R}^{L \times d_{\text{seq}}^{\text{in}}}$

Encoder:

$\mathbf{h}_i \leftarrow \text{SiLU}(\text{Linear}(\text{LayerNorm}(\tilde{\mathbf{x}}_i)))$

$\triangleright \mathbf{h}_i \in \mathbb{R}^{d_{\text{seq}}}$

$\mathbf{z}_{ij} \leftarrow \text{AbsoluteRelativeIndexEmbedding}(|i - j|)$ \triangleright Bucketized embedding, $\mathbf{z}_{ij} \in \mathbb{R}^{d_{\text{pair}}}$,

for i in $1, \dots, n_{\text{blocks}}^{\text{encoder}}$ **do**

$\mathbf{h}, \mathbf{z} \leftarrow \text{ESMFoldingBlock}_i(\mathbf{h}, \mathbf{z})$

$\triangleright \mathbf{h} \in \mathbb{R}^{L \times d_{\text{seq}}}, \mathbf{z} \in \mathbb{R}^{L \times L \times d_{\text{pair}}}$

end for

Decoder / flow:

Sample $t \sim \mathcal{T}$

$\triangleright \mathcal{T}$ is a uniform distribution in $[0, 1]$

Sample $\mathbf{r}_0 \sim p_0$

$\triangleright \mathbf{r}_0 \in \mathbb{R}^{L \times 3}, p_0$ is a 3D isotropic Gaussian

Align: $\mathbf{r}_0 \leftarrow \text{RMSDAlign}(\mathbf{r}_0, \mathbf{r}_1)$

Interpolate $\mathbf{r}_t \leftarrow t\mathbf{r}_0 + (1 - t)\mathbf{r}_1$

$\triangleright \mathbf{r}_t$ is also aligned to \mathbf{r}_1

$\mathbf{h}, \hat{\mathbf{r}}_1, \mathbf{z} \leftarrow \text{AttentionEGNN}(\mathbf{h}, \mathbf{r}_t, \mathbf{z}, t)$

Optimize loss $\mathcal{L} = \text{Loss}(\hat{\mathbf{r}}_1, \mathbf{r}_1)$

end for

Insulation score The insulation score was first proposed in [96] to identify topologically associating domain (TAD) boundaries. In this work, we follow the implementation provided in cooltools [97], which computes the mean contact frequency within a local square (“diamond”) window of size w along the Hi-C. For an average contact map $C \in \mathbb{R}^{L \times L}$ and window size w , the local mean contact at position i is defined as

$$I_i = \frac{\sum_{j=l_i}^i \sum_{k=i}^{h_i} C_{jk}}{\sum_{j=l_i}^i \sum_{k=i}^{h_i} 1}, \quad (19)$$

$$l_i = \max(1, i - w + 1), \quad h_i = \min(L, i + w - 1), \quad (20)$$

and normalized by the median across all bins to get the insulation score S ,

$$S_i = \frac{I_i}{\text{median}(I)}. \quad (21)$$

Each bin in our data represents 250 kb, and we set $w = 4$, corresponding to ~ 2 Mb neighborhood around each bin. Valleys in insulation score profile indicate reduced cross-domain contacts and are interpreted as candidate TAD boundaries.

Pearson correlation coefficient For two arrays of equal lengths x and y , the Pearson correlation coefficient is defined as

$$\text{corr}(x, y) = \frac{1}{n} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma(x)\sigma(y)}, \quad (22)$$

where \bar{x} is the mean of x , and $\sigma(x)$ is the standard deviation of x . To ensure the variance is well-defined, $n \geq 2$ is required.

Algorithm 7 Inference algorithm for structure generation

Require: Raw epigenomic features of target sequence $\tilde{\mathbf{x}}$

Preprocessing and Encoder:

$\mathbf{h}_i \leftarrow \text{SiLU}(\text{Linear}(\text{LayerNorm}(\tilde{\mathbf{x}}_i)))$ $\triangleright \mathbf{h}_i \in \mathbb{R}^{d_{\text{seq}}}$

$\mathbf{z}_{ij} \leftarrow \text{AbsoluteRelativeIndexEmbedding}(|i - j|)$ $\triangleright \mathbf{z}_{ij} \in \mathbb{R}^{d_{\text{seq}}}$

for i in $1, \dots, n_{\text{blocks}}^{\text{encoder}}$ **do**

$\mathbf{h}, \mathbf{z} \leftarrow \text{ESMFoldingBlock}_i(\mathbf{h}, \mathbf{z})$

end for

\triangleright Refined embeddings \mathbf{h}, \mathbf{z} are now ready for the decoder

Decoder / Integration Loop:

Sample an initial noisy structure $\mathbf{r}_0 \sim p_0$

$\triangleright \mathbf{r}_0 \in \mathbb{R}^{N \times 3}$

Set number of integration steps n , time step $\Delta t = 1/n$

Initialize: $\hat{\mathbf{r}}_0^{\text{ODE}} \leftarrow \mathbf{r}_0$

for t in $0, \Delta t, \dots, (n-1)\Delta t$ **do**

$\mathbf{h}, \hat{\mathbf{r}}_1, \mathbf{z} \leftarrow \text{AttentionEGNN}(\mathbf{h}, \hat{\mathbf{r}}_t^{\text{ODE}}, \mathbf{z}, t)$

$\hat{\mathbf{r}}_t^{\text{ODE}} \leftarrow \text{RMSDAlign}(\hat{\mathbf{r}}_t^{\text{ODE}}, \hat{\mathbf{r}}_1)$

\triangleright RMSD align \mathbf{r}_t to $\hat{\mathbf{r}}_1$

$\hat{\mathbf{v}}_t \leftarrow (\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_t^{\text{ODE}})/(1 - t)$

$\hat{\mathbf{r}}_{t+\Delta t}^{\text{ODE}} \leftarrow \hat{\mathbf{r}}_t^{\text{ODE}} + \hat{\mathbf{v}}_t \cdot \Delta t$

end for

return $\hat{\mathbf{r}}_1^{\text{ODE}}$

Stratum-adjusted correlation coefficient (SCC) When comparing two contact maps $C, C' \in \mathbb{R}^{L \times L}$, a direct correlation over all entries is biased by genomic distance, since loci that are sequentially close tend to exhibit higher contact frequencies. To address this, the stratum-adjusted correlation coefficient (SCC) [98] is defined as a weighted correlation across strata:

$$w_d = \begin{cases} \text{length}(\{C_{i,i+d}\}), & \text{if } \text{length}(\{C_{i,i+d}\}) \geq 2 \text{ and } \rho_d \in [-1, 1] \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

$$\rho_d = \text{corr}(\{C_{i,i+d}\}, \{C'_{i,i+d}\}) \quad (w_d > 0), \quad (24)$$

$$\text{scc}(C, C') = \frac{\sum_{d=1}^{L-1} w_d \rho_d}{\sum_{d=1}^{L-1} w_d}. \quad (25)$$

Here, length denotes the number of elements in the stratum, and unknown contacts are omitted. $w_d = 0$ if the correlation within the d -th stratum is not well-defined. SCC reduces the bias from short-range contacts and provides a more reliable measure of global structural similarity.

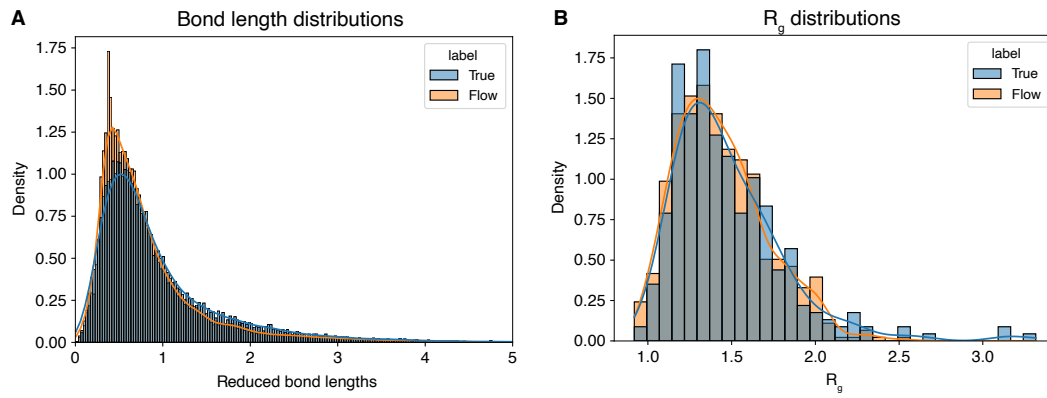


Figure S1: Bond length and radius of gyration distributions of Chr2 201-300, which is the same region of the results shown in the main text.

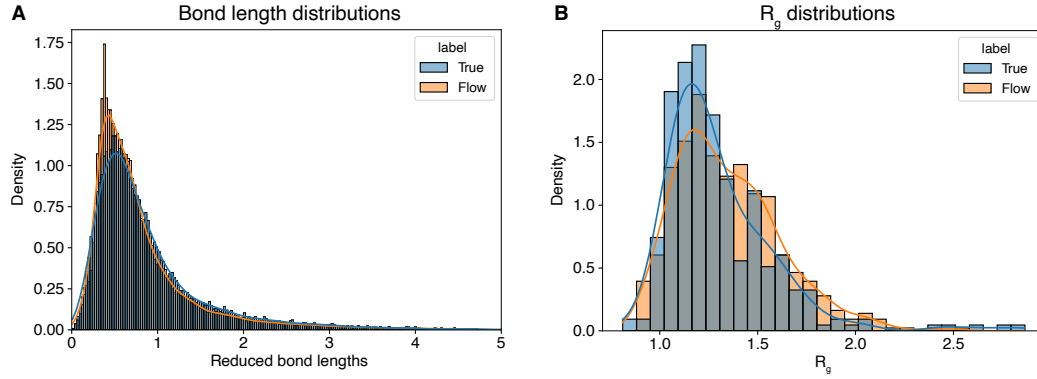


Figure S2: Bond length and radius of gyration distributions of Chr2 1-100, obtained using the same model as in the main text. This figure parallels **Fig. S1** but applies the analysis to a different genomic region.

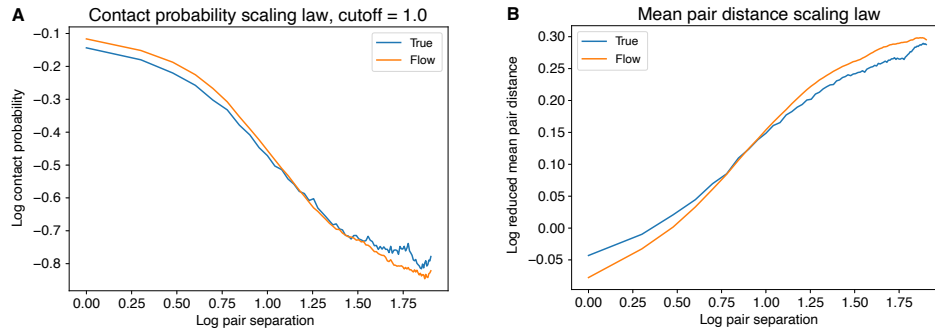


Figure S3: Scaling laws of ChromWeave-generated chromosome conformations of Chr2 1-100, obtained using the same model as in the main text. This figure parallels **Fig. 2** but applies the analysis to a different genomic region.

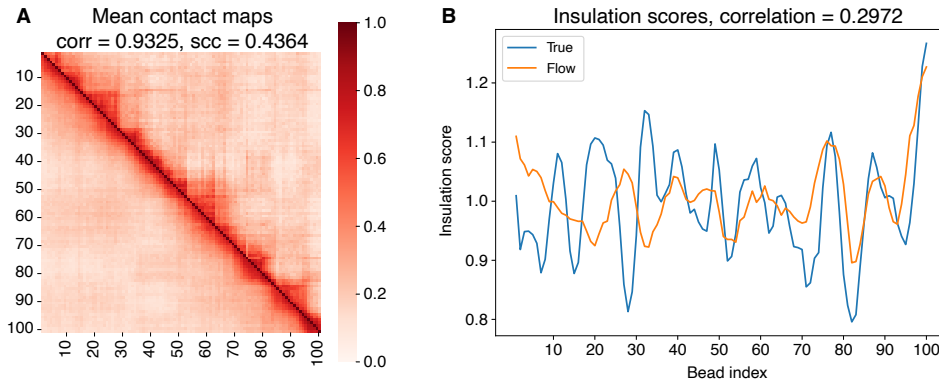


Figure S4: Comparison of contact maps and insulation scores for Chr2 1-100, obtained using the same model as in the main text. This figure parallels **Fig. 3** but applies the analysis to a different genomic region.