Improving AlphaFLOW for Efficient Protein Ensembles Generation

Shaoning Li [*1] **Mingyu Li** [*2] **Yusong Wang** [3] **Xinheng He** [4] **Nanning Zheng** [3] **Jian Zhang** [2] **Pheng Ann Heng** [1]

Abstract

Investigating conformational landscapes of proteins is a crucial way to understand their biological functions and properties. AlphaFLOW stands out as a *sequence-conditioned* generative model that introduces flexibility into structure prediction models by fine-tuning AlphaFold under the flow-matching framework. Despite the advantages of efficient sampling afforded by flow-matching, AlphaFLOW still requires multiple runs of AlphaFold to finally generate one single conformation. Due to the heavy consumption of AlphaFold, its applicability is limited in sampling larger set of protein ensembles or the longer chains within a constrained timeframe. In this work, we propose a *feature-conditioned* generative model called AlphaFLOW-Lit to realize efficient protein ensembles generation. In contrast to the full fine-tuning on the entire structure, we focus solely on the light-weight structure module to reconstruct the conformation. AlphaFLOW-Lit performs on-par with AlphaFLOW and surpasses its distilled version without pretraining, all while achieving a significant sampling acceleration of around 47 times. The advancement in efficiency showcases the potential of AlphaFLOW-Lit in enabling faster and more scalable generation of protein ensembles.

1. Introduction

Exploring conformational landscapes is essential to capture the dynamic nature of protein structures, offering insights into their flexibility, biological function, and interactions. Traditionally, ensembles of conformational changes are collected through molecular dynamics (MD) simulations (Karplus & McCammon, 2002). While MD-base methods adhere to physical laws and will theoretically explore the entire landscape, they are time- and resource-intensive. To expedite this process, some methods focus on increasing the diversity of AlphaFold (Jumper et al., 2021), which is a powerful deep learning model for crystal structure prediction but falls short in accounting for conformational divergence. Specifically, these methods sample different multiple sequence alignments (MSAs) as input (Wayment-Steele et al., 2024) or enable the dropout function in AlphaFold (Wallner, 2023) during the inference process. Despite these inference interventions indeed bring some diversity to AlphaFold, they fall significantly short of generating enough conformational heterogeneity to thoroughly explore the protein landscape.

Recently, (Jing et al., 2024) harnessed the power of generative methods *flow matching*, and integrated this framework into AlphaFold, called AlphaFLOW. To be concrete, it treats AlphaFold as a powerful *sequence-conditioned* denoising model, which receives the noisy structures as templates and samples the protein ensembles from harmonic prior under a flow field. AlphaFLOW inherits the weights of AlphaFold, and was trained on general PDB then fine-tuned on different protein MD trajectories as a regression model, using loss functions similar to those in the original AlphaFold. Due to these enhancements, AlphaFLOW is much more flexible and diverse than the aforementioned inference intervention methods. It is the first method to ingeniously combine both the advantages of accurate structure prediction and the generative capability for conformation sampling.

However, limitations persist in sampling consumption. As shown in Fig. 1, since AlphaFLOW is trained by fully fine-tuning AlphaFold, generating the final structure $\hat{x}_0$ requires T denoising steps, which means running T times AlphaFold with additional embedders. Although flow matching method is relatively faster compared with other diffusion methods, as shown in Fig. 2(**A**), AlphaFLOW showcases cubic growth with the chain length, which leads to unacceptable time consumption and hinders its application for generating larger set of protein ensembles. While AlphaFLOW adopts diffusion distillation to reduce the generative process to a single forward pass, this approach compromises the level of the sampling performance.

[*]Equal contribution [1]Department of Computer Science and Engineering, The Chinese University of Hong Kong [2]Medicinal Chemistry and Bioinformatics Center, Shanghai Jiao Tong University School of Medicine [3]National Key Laboratory of Human-Machine Hybrid Augmented Intelligence and Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University [4]Shanghai Institute of Materia Medica, Chinese Academy of Sciences. Correspondence to: Pheng Ann Heng <pheng@cse.cuhk.edu.hk>, Jian Zhang <jian.zhang@sjtu.edu.cn>.

Proceedings of the 41st International Conference on Machine Learning AI for Science workshop, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

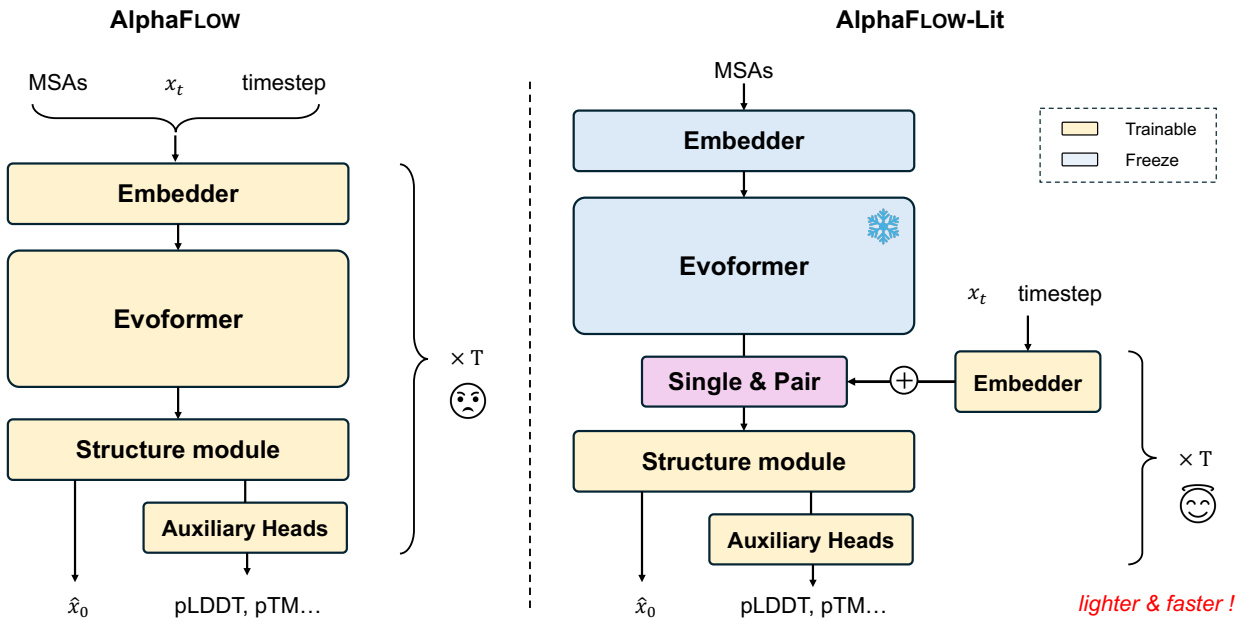

Figure 1. Model architecture of *sequence-conditioned* AlphaFLOW (*left*) and *feature-conditioned* AlphaFLOW-Lit (*right*). T: Denoising steps; x_t: Noisy structure; $\tilde{x}_0$: Predicted structure.

To address this issue, we propose a *feature-conditioned* generative model called AlphaFLOW-Lit, which can be treated as an efficient and *lighter* version of AlphaFLOW. As demonstrated in AlphaMissense (Cheng et al., 2023), features derived from MSAs encoder could be further utilized for variants classification. The lastest AlphaFold3 also employs these features to train a non-equivariant denoiser (Abramson et al., 2024) for the adaption of multi-modality such as nucleic acids, small molecules, ions, and modified residues. Also, as shown in (Jing et al., 2024), the MSAs have relatively minor impact on structural diversity compared with the flow matching framework. Inspired by these findings, as illustrated in Fig. 1, AlphaFLOW-Lit retains the original AlphaFold embedder and Evoformer blocks in a frozen state and is directly conditioned on computed single and pair features. Given that the remaining structure module and auxiliary heads are significantly lighter than the Evoformer block, compared to AlphaFLOW, AlphaFLOW-Lit can achieve a faster sampling process (around 47 times speedup) with the same number of denoising steps. When directly trained on ATLAS dataset (Vander Meersche et al., 2024) of protein MD trajectories, AlphaFLOW-Lit performs on-par with AlphaFLOW and surpasses its distilled version. We also provide the analysis of the generated MD ensembles from AlphaFLOW-Lit, including protein dynamics analysis, local arrangements within residue, and long-range correlations among residues to illustrate the diverse capabilities of AlphaFLOW-Lit.

2. Preliminary

In this section, we briefly introduce the flow matching framework and some details of AlphaFLOW.

Flow matching The flow matching framework begins with the continuous normalizing flow (CNF) ψ_t, defined as the solution of an ordinary differential equation (ODE) governed by a time-dependent vector field $u_t : \frac{d}{dt}\psi_t = u_t(\psi_t)$. Let x be a data point on a specific manifold, the CNF has an initial condition $\psi_0(x) = x$. Given two distributions p_0 and p_1, we can define a probability path p_t as their interpolation, which can be viewed as paths generated by u_t. To effectively learn the CNF, we make u_t tractable and by adopting the conditional probability path $p_t(x|x_1)$, which samples x_0 from prior distribution p_0 and interpolating it linearly with the data point x_1:

$$x_t = (1-t)\cdot x_0 + t \cdot x_1 \tag{1}$$

with the corresponding vector field:

$$u_t(x_t|x_1) = (x_1 - x_t)/(1-t) \tag{2}$$

This method is referred as conditional flow matching (CFM). We employ a neural network v_t^θ to learn the vector field. The objective of CFM can be written as:

$$\mathcal{L}_{\text{CFM}} = \mathbb{E}_{t,p(x_0),p(x_1)} \left\| v_t^\theta(x_t) - u_t(x_t|x_1) \right\|^2 \tag{3}$$

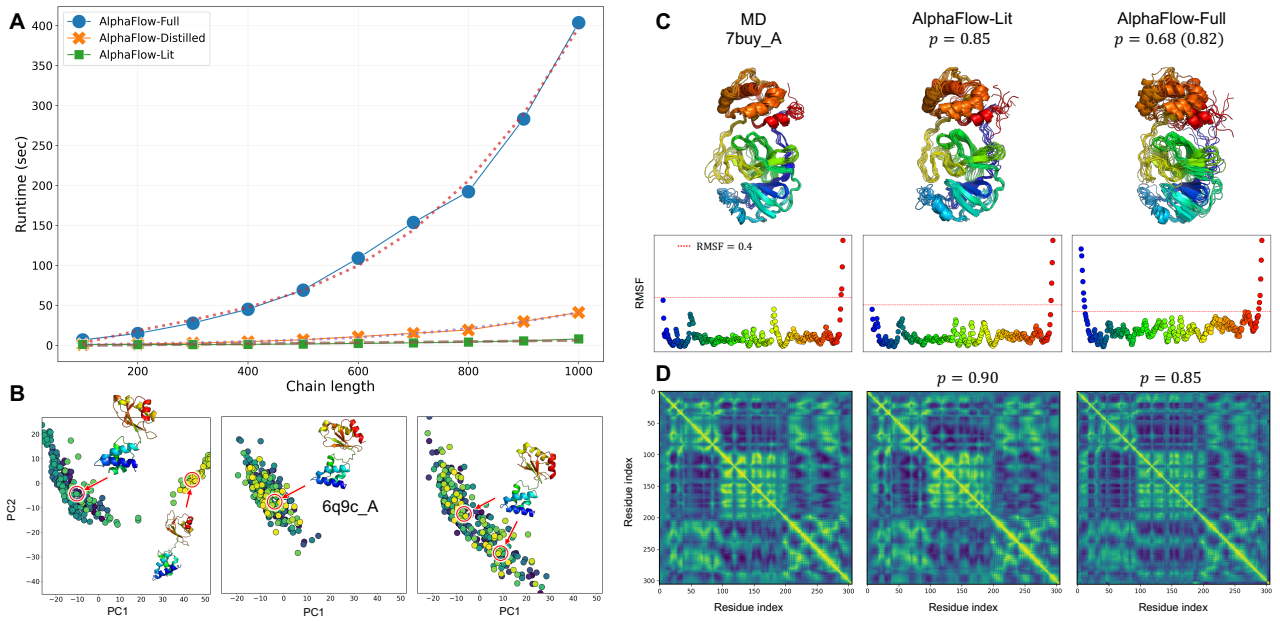

Figure 2. **Visualization of MD evaluation from MD, AlphaFLOW-Lit and AlphaFLOW.** (**A**) Runtime comparison corresponding to the sequence length and their fitted curves. (**B**) Principal components analysis (PCA) for `6q9c_A` ensembles. The representative structures are pointed out. (**C, D**) Ensembles of PDB ID `7buy_A` with Cα RMSF by residue index shown in insets, and their Dynamic cross-correlation matrix (DCCM).

AlphaFLOW To integrate the AlphaFold, which is a regression model that directly outputs x_1, into the flow matching framework, AlphaFLOW reparameterizes the neural vector field as:

$$v_t^\theta(x_t) = (\text{AlphaFold}(x_t) - x_t)/(1-t) \tag{4}$$

It allows the objective to be rewritten as learning the expectation of x_1. Consequently, AlphaFLOW can employ the similar regression loss function (e.g. FAPE) to optimize the neural network. AlphaFLOW introduces two key innovations: (a) It employs the 3D coordinates of its β-carbons (α-carbon for glycine) to describe the noisy structure and the prior distribution is defined over the β-carbons coordinates as a harmonic prior (Jing et al., 2023); (b) AlphaFLOW treats x_t as features (similar to templates), and the denoising process does not directly apply to the spatial domain as in prevailing SE(3) generative models (Yim et al., 2023; Bose et al., 2023; Li et al., 2024). Instead, it starts from the identity rigids, which is the same as AlphaFold. These contributions make AlphaFLOW as a new paradigm for utilizing AlphaFold within different frameworks or applications.

3. Method

AlphaFLOW-Lit follows the same ideas of AlphaFLOW but introduces some modifications to the input pipeline. AlphaFLOW-Lit is a *feature-conditioned* generative model, that is to say, it is conditioned on the single and pair features after the Evoformer blocks to generate diverse conformations. As illustrated in Fig. 1, the AlphaFold embedders (including the original input embedder, recycling embedder, extra MSA embedder, and extra MSA stack) and Evoformer are kept frozen. The input embedding module for noisy structures x_t and timesteps is similar to that of AlphaFLOW but with little modifications. The single and pair output of input embedding module are derived from the torsion angles (if designated) and contact map of the noisy structure, respectively. These outputs are followed by a Linear layer initialized with zeros before summation with the features after Evoformer blocks. This is similar to the zero convolution in ControlNet (Zhang et al., 2023), designed to minimally disrupt the pretrained weights at the outset. The detailed algorithm for input embedding module is described in Appendix A Algorithm 2. It is worth noting that the torsion angles in AlphaFold are represented in 8 rigids groups with $\sin - \cos$ formats, indicating rotation towards the coordinates of the former group. As a result, these angles are invariant to rigid transformations, eliminating the need to rotate the predicted structure after RMSD alignment. The training procedure for AlphaFLOW-Lit is the same as for AlphaFLOW, and the inference procedure is provided in Algorithm 1. We keep the Algorithm notations same as (Jing et al., 2024). The acceleration primarily results from the pre-computation of single and pair features. In contrast to

AlphaFLOW, AlphaFLOW-Lit does not necessitate running Evoformer blocks at each denoising step but only conducts this process once at the inception. Under this circumstances, the denoising network in AlphaFLOW-Lit is a lightweight structure module that is conditioned on single and pair features rather than MSAs. It distinguishes AlphaFLOW-Lit as feature-conditioned and contributes to its efficiency and speed in generating protein emsembles.

Algorithm 1 INFERENCE

Input: Sequence and MSA (A, M)
Output: Sampled all-atom structure $\hat{S}$
Sample $\mathbf{x}_0 \sim \mathrm{HarmonicPrior}(\mathrm{length}(A))$
$\mathbf{f}_0^{\mathrm{tor}} = \emptyset$
$\mathbf{s}_i^{\mathrm{evo}}, \mathbf{z}_{ij}^{\mathrm{evo}} \leftarrow \mathrm{Evoformer}(\mathrm{AlphaFoldEmbedder}(A, M))$
for $n \leftarrow 0$ to $N-1$ **do**
  Let $t \leftarrow n/N$ and $s \leftarrow t + 1/N$
  $\mathbf{s}_i, \mathbf{z}_{ij} \leftarrow \mathrm{InputEmbedder}(\mathbf{x}_t, \mathbf{f}_t^{\mathrm{tor}}, t)$
  $\mathbf{s}_i, \mathbf{z}_{ij} \mathrel{+}= \mathbf{s}_i^{\mathrm{evo}}, \mathbf{z}_{ij}^{\mathrm{evo}}$
  Predict $\hat{S} \leftarrow \mathrm{StructureModule}(\mathbf{s}_i, \mathbf{z}_{ij}, A)$
  if $n = N-1$ **then**
    return $\hat{S}$
  end if
  Extract $\hat{\mathbf{x}}_1 \leftarrow \mathrm{BetaCarbons}(\hat{S})$
  Align $\mathbf{x}_t \leftarrow \mathrm{RMSDAlign}(\mathbf{x}_t, \hat{\mathbf{x}}_1)$
  Interpolate $\mathbf{x}_s \leftarrow \frac{s-t}{1-t} \cdot \hat{\mathbf{x}}_1 + \frac{1-s}{1-t} \cdot \mathbf{x}_t$
  if embed_angles **then**
    Extract $\mathbf{f}_t^{\mathrm{tor}} \leftarrow \mathrm{TorsionAngle}(\hat{S})$
  else
    $\mathbf{f}_t^{\mathrm{tor}} \leftarrow \emptyset$
  end if
end for

4. Experiments

We directly train AlphaFLOW-Lit on ALTAS MD trajectories (Vander Meersche et al., 2024) without pretraining on the PDB. Similar to AlphaFLOW, we use 1265/39/82 ensembles for the training, validation, and test splits, respectively. All multiple sequence alignments (MSAs) are derived from OpenProteinSet (Ahdritz et al., 2024). For sequences not present in OpenProteinSet, we use MMseqs2 (Steinegger & Söding, 2017) to search the UniRef30 and ColabDB databases (Mirdita et al., 2022). The initial weight of AlphaFLOW-Lit is from the AlphaFold's publicly available weights. ALTAS provides three parallel trajectories with 10,001 frames for each protein. We subsample the trajectories with a stride of 100 frames to create the training set (300 frames in total). During training, we uniformly sample one frame at each step. Since the Evoformer blocks are frozen, we set the weight of masked MSA loss to 0.

We generate 250 samples for 82 targets in the test set. AlphaFLOW-Full refers to AlphaFLOW with 10 consecutive denoising steps. AlphaFLOW-Distilled denotes its distilled with a single forward denoising step. AlphaFLOW-Lit employs the full denoising steps. The protein ensembles of AlphaFLOW-Full and AlphaFLOW-Distilled are downloaded from its public repository[1]. We assess the sampling runtime based on the protein length for these methods and approximate their consumption curve. To evaluate the effectiveness of each method, we first investigate the protein dynamics, considering both the general dynamics indicated by the pairwise root-mean-square deviation (RMSD) and the essential dynamics uncovered through principal components analysis (PCA) (Amadei et al., 1993). Furthermore, we assess the detailed capability of each method at the residue resolution of protein dynamics by systematic comparisons of local arrangements within residues and motional correlations among residues. The results are presented in Table 1 and Fig. 2.

Runtime comparison In Figure 2, we depict the relation between runtime of sampling and sequence length ranging from 100AA to 1,000AA in increments of 100. AlphaFLOW-Lit demonstrate superior scalability, maintaining consistently low runtime across increasing protein lengths. AlphaFLOW exhibit cubic growth in runtime, indicating its inefficiency for longer chains. This inefficiency could be attributed to the cubic complexity of attention in the Evoformer block. While AlphaFLOW-Distilled performs better than AlphaFLOW-Full due to its single forward inference, it still shows moderate increases in runtime as protein length grows. In summary, AlphaFLOW-Lit surpasses AlphaFLOW by 6 to 51 times (47 times in average) and AlphaFLOW-Distilled by 2 to 4 times (3.8 times in average), making it the most efficient configuration and highlighting its potential for generating a larger set of protein ensembles.

Protein dynamics analysis For each conformational ensemble, the general dynamics are quantified as the average C_α-RMSD between any pair of conformations. Using this measurement, the AlphaFLOW-Lit ensembles demonstrate the strongest Pearson correlation with the ground truth ensembles produced by classic MD, while maintaining a comparable level of diversity in the conformational ensembles relative to AlphaFLOW-Full. In contrast, AlphaFLOW-Distilled loosely match the general dynamics with the ground truth and does not achieve the same level of diversity. Also, we assess the essential dynamics of proteins by projecting the ensembles onto the first two principal components (PCs) derived from PCA. Two common featurization methods for proteins are utilized: aligned C_α absolute coordinates and pairwise C_α internal distances. The differences in the distributions are quantified using the mean Jensen-Shannon divergence (JSD) for each PC

[1] https://github.com/bjing2016/alphaflow

Table 1. **Evaluation on MD ensembles.** We compare the predicted ensemble from AlphaFLOW-Lit and AlphaFLOW with the ground truth MD ensemble according to various metrics. For Pairwise RMSD and Per-target RMSF, the ground truth values (from the MD ensembles) are in parenthesis. Following (Jing et al., 2024), the median across the 82 test ensembles is reported. p: Pearson correlation; JSD: Jensen-Shannon divergence.

		AlphaFLOW-Full	AlphaFLOW-Lit	AlphaFLOW-Distilled
Protein dynamics	Pairwise RMSD (= 2.90)	2.89	2.43	1.94
	Pairwise RMSD $p \uparrow$	0.49	0.58	0.49
	PCA C_α coordinates $JSD \downarrow$	0.43	0.46	0.51
	PCA C_α pairwise distance $JSD \downarrow$	0.48	0.52	0.56
Local arrangements within residues	Per-target RMSF (= 1.94)	1.88	1.65	1.34
	Per-target RMSF $p \uparrow$	0.75	0.77	0.71
	Stable contacts $JSD \downarrow$	0.84	0.83	0.79
	Dihedral distributions $JSD \downarrow$	0.47	0.51	0.57
Long-range correlations among residues	DCCM $p \uparrow$	0.78	0.78	0.74

between the predicted and true ensembles. Likewise, AlphaFLOW-Lit exhibits essential dynamics distributions that are comparable to those of AlphaFLOW-Full, surpassing the performance of AlphaFLOW-Distilled. We visualize an example `6q9c_A` in Fig. 2(**B**) and extract the representative structures. Both AlphaFLOW-Full and AlphaFLOW-Lit highly align with one of the principal component distributions of molecular dynamics (MD) However, they do not capture the other distribution, which, although relatively minor, is still significant.

Local arrangements analysis Allostery, which has been coined the second secret of life after genetic codes, is a fundamental mechanism underlying most protein dynamics (Fenton, 2008). To further evaluate the practical effects of each method in identifying residues that undergo conformational changes in the local environment—changes that are likely critical for allostery—we performed a multifaceted analysis at the residue resolution, including thermally averaged flexibility, residual contact probabilities, and key dihedral angles distribution. In terms of thermally averaged flexibility, we calculate the root-mean-square fluctuation (RMSF) at the residue level, represented by C_α. AlphaFLOW-Lit achieves a strong Pearson correlation of 0.77 between the predicted and actual ensembles within a target, whereas AlphaFLOW-Distilled only exhibits a moderate Pearson correlation of 0.71. We visualize the RMSF of `7buy_A` in Fig. 2(**C**). We observe that the decrease in Pearson correlation of AlphaFLOW is due to the high diversity of the structure's end point. If we exclude the first 5 residues and recompute the RMSF Pearson correlation (values in parentheses), both AlphaFLOW and AlphaFLOW-Lit will yield identical results. In addition, contact probability analysis is utilized to elucidate the conformational rearrangements concerning the relative positions and orientations of structural motifs. Meanwhile, their internal conformational displacements are more accurately represented through variations in key dihedral angles. For contact probability analysis, a stable contact is defined as C_α pairs that maintain contact (with a threshold of 7 Å) in over 85% of the conformational ensembles. The Jaccard similarity (JS) between the contact residue pairs is calculated between the predicted and the ground truth sets. Regarding key dihedral angles, in addition to the backbone phi and psi angles, the chi1 values are also included, as they may indicate significant side chain reorientations relevant to the formation of key polar or non-polar interactions. The results show that contact and dihedral distributions exhibit moderate consistency between the actual and predicted ensembles generated by AlphaFLOW-Full and AlphaFLOW-Lit, unlike the larger inconsistencies observed with AlphaFLOW-Distilled.

Long-range correlations analysis Finally, we analyze the motional correlations similarity among residues by calculating the dynamic cross-correlation map (DCCM) using the conformational ensembles generated by different methods (Hünenberger et al., 1995). Such correlations could reveal pivot residues that mediate long-range allosteric coupling (McClendon et al., 2009). AlphaFLOW-Lit shows higher similarity in these matrices compared to AlphaFLOW-Distilled, underscoring its superior ability to capture couplings even among long-range residues. We visualize the DCCM of `7buy_A` in Fig. 2(**D**).

5. Conclusion

We propose AlphaFLOW-Lit, an improved version of AlphaFLOW for efficient protein ensembles generation. Compared with AlphaFLOW, AlphaFLOW-Lit is a feature-conditioned generative model that eliminates the heavy reliance on MSAs encoding blocks and utilizes computed features to produce a diverse range of conformations. By directly training on ATLAS, AlphaFLOW-Lit performs on-par with AlphaFLOW while outperforming its distilled version, all while achieving a substantial acceleration in sampling speed of around 47 times. In addition, we conduct a thor-

ough analysis of protein dynamics, local arrangements, and long-range coupling within the generated ensembles. The advantages of AlphaFLOW-Lit make it capable of generating a larger set of protein ensembles, enabling us to more effectively explore the protein landscape using deep learning techniques.

Limitation and future work As illustrated in the Experiment section, AlphaFLOW-Lit exhibits less diversity compared to AlphaFLOW-Full, likely due to the absence of pretraining on the PDB or insufficient training on MD trajectories. This will be addressed in future work. Additionally, in the PCA analysis of example `6q9c_A`, both models fail to capture the additional conformation present in the ground truth MD distribution. Enhancing their capability to capture such nuances will be a focus of our future research.

Acknowledgement

This paper is supported by National Key Research and Development Program of China (2023YFF1205103), National Natural Science Foundation of China (81925034), National Natural Science Foundation of China (62088102) and a grant from the Hong Kong Innovation and Technology Fund (Project No. ITS/241/21). We thank the open-source codebases from OpenFold Team and Bowen Jing (AlphaFLOW). We thank Dr. Le Zhuo for valuable discussions.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, pp. 1–3, 2024.

Ahdritz, G., Bouatta, N., Floristean, C., Kadyan, S., Xia, Q., Gerecke, W., O'Donnell, T. J., Berenberg, D., Fisk, I., Zanichelli, N., et al. Openfold: Retraining alphafold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods*, pp. 1–11, 2024.

Amadei, A., Linssen, A. B., and Berendsen, H. J. Essential dynamics of proteins. *Proteins: Structure, Function, and Bioinformatics*, 17(4):412–425, 1993.

Bose, J., Akhound-Sadegh, T., FATRAS, K., Huguet, G., Rector-Brooks, J., Liu, C.-H., Nica, A. C., Korablyov, M., Bronstein, M. M., and Tong, A. Se (3)-stochastic flow matching for protein backbone generation. In *The Twelfth International Conference on Learning Representations*, 2023.

Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., Pritzel, A., Wong, L. H., Zielinski, M., Sargeant, T., et al. Accurate proteome-wide missense variant effect prediction with alphamissense. *Science*, 381(6664):eadg7492, 2023.

Fenton, A. W. Allostery: an illustrated definition for the 'second secret of life'. *Trends in biochemical sciences*, 33 (9):420–425, 2008.

Hünenberger, P., Mark, A., and Van Gunsteren, W. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *Journal of molecular biology*, 252(4):492–503, 1995.

Jing, B., Erives, E., Pao-Huang, P., Corso, G., Berger, B., and Jaakkola, T. S. Eigenfold: Generative protein structure prediction with diffusion models. In *ICLR 2023-Machine Learning for Drug Discovery workshop*, 2023.

Jing, B., Berger, B., and Jaakkola, T. Alphafold meets flow matching for generating protein ensembles. *arXiv preprint arXiv:2402.04845*, 2024.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

Karplus, M. and McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature structural biology*, 9(9):646–652, 2002.

Li, S., Wang, Y., Li, M., Shao, B., Zheng, N., Jian, Z., and Tang, J. F^3low: Frame-to-frame coarse-grained molecular dynamics with se(3) guided flow matching. In *ICLR 2024 Workshop on Generative and Experimental Perspectives for Biomolecular Design*, 2024.

McClendon, C. L., Friedland, G., Mobley, D. L., Amirkhani, H., and Jacobson, M. P. Quantifying correlations between allosteric sites in thermodynamic ensembles. *Journal of chemical theory and computation*, 5(9):2486–2502, 2009.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. Colabfold: making protein folding accessible to all. *Nature methods*, 19(6):679–682, 2022.

Steinegger, M. and Söding, J. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology*, 35(11):1026–1028, 2017.

Vander Meersche, Y., Cretin, G., Gheeraert, A., Gelly, J.-C., and Galochkina, T. Atlas: protein flexibility description from atomistic molecular dynamics simulations. *Nucleic Acids Research*, 52(D1):D384–D392, 2024.

Wallner, B. Afsample: improving multimer prediction with alphafold using massive sampling. *Bioinformatics*, 39(9): btad573, 2023.

Wayment-Steele, H. K., Ojoawo, A., Otten, R., Apitz, J. M., Pitsawong, W., Hömberger, M., Ovchinnikov, S., Colwell, L., and Kern, D. Predicting multiple conformations via sequence clustering and alphafold2. *Nature*, 625(7996): 832–839, 2024.

Yim, J., Campbell, A., Foong, A. Y., Gastegger, M., Jiménez-Luna, J., Lewis, S., Satorras, V. G., Veeling, B. S., Barzilay, R., Jaakkola, T., et al. Fast protein backbone generation with se (3) flow matching. *arXiv preprint arXiv:2310.05297*, 2023.

Zhang, L., Rao, A., and Agrawala, M. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3836–3847, 2023.

A. Method Details.

We highlight the difference with AlphaFLOW in yellow. $\text{init} = \text{'final'}$ indicates the weight and bias (if has) of Linear is initialized with 0. Other notations are kept same as (Jumper et al., 2021) and (Jing et al., 2024).

Algorithm 2 INPUTEMBEDDING

Input: Beta carbon coordinates $\mathbf{x} \in \mathbb{R}^{N\times 3}$, Torsion angles $\mathbf{f}^{\text{tor}}$, time $t \in [0, 1]$
Output: Input pair embedding $\mathbf{z} \in \mathbb{R}^{N\times N\times 64}$
$\mathbf{z}_{ij} \leftarrow \|\mathbf{x}_i - \mathbf{x}_j\|$
$\mathbf{z}_{ij} \leftarrow \text{Bin}(\mathbf{z}_{ij}, \text{min} = 3.25\ \text{Å}, \text{max} = 50.75\ \text{Å}, N_{\text{bins}} = 39)$
$\mathbf{z}_{ij} \leftarrow \text{Linear}(\text{concat}(\text{OneHot}(\mathbf{z}_{ij}), \mathbf{f}_{ij}^{\text{mask_2d}}))$
for $l \leftarrow 1$ to $N_{\text{blocks}} = 4$ **do**
  $\{\mathbf{z}\}_{ij} \mathrel{+}= \text{TriangleAttentionStartingNode}(\mathbf{z}_{ij}, c = 64, N_{\text{head}} = 4)$
  $\{\mathbf{z}\}_{ij} \mathrel{+}= \text{TriangleAttentionEndingNode}(\mathbf{z}_{ij}, c = 64, N_{\text{head}} = 4))$
  $\{\mathbf{z}\}_{ij} \mathrel{+}= \text{TriangleMultiplicationOutgoing}(\mathbf{z}_{ij}, c = 64)$
  $\{\mathbf{z}\}_{ij} \mathrel{+}= \text{TriangleMultiplicationIncoming}(\mathbf{z}_{ij}, c = 64)$
  $\{\mathbf{z}\}_{ij} \mathrel{+}= \text{PairTransition}(\mathbf{z}_{ij}, n = 2)$
end for
$\mathbf{z}_{ij} \leftarrow \text{Linear}(\mathbf{z}_{ij}, \text{init} = \text{'final'})$
$\mathbf{z}_{ij} \mathrel{+}= \text{Linear}(\text{GaussianFourierEmbedding}(t, d = 128), \text{init} = \text{'final'})$
if embed_angles and $\mathbf{f}^{\text{tor}} \neq \emptyset$ **then**
  $\mathbf{s}_i \leftarrow \text{Linear}(\text{concat}(\text{OneHot}(\mathbf{f}_i^{\text{tor}}, \mathbf{f}^{\text{mask_tor}})))$
  $\mathbf{s}_i \leftarrow \text{Linear}(\mathbf{s}_i, \text{init} = \text{'final'})$
else
  $\mathbf{s}_i \leftarrow \mathbf{0}$
end if

B. Runtime comparison

Table 2. Sampling runtime across AlphaFLOW, AlphaFLOW-Lit and AlphaFLOW-Distilled in Fig. 2(**A**). Proteins are selected from ATLAS. All methods are conducted on a single A100 GPU. Runtime values are reported in second.

PDB ID	Seq. length	AlphaFLOW-Full	AlphaFLOW-Distilled	AlphaFLOW-Lit
5h6x_A	100	6.63	0.86	0.76
2q9r_A	200	14.75	1.48	0.81
2v4b_B	300	27.74	2.76	0.85
1ru4_A	400	44.98	4.46	1.10
2d5b_A	500	68.97	6.82	1.50
6zsl_B	603	108.96	10.75	2.20
6lrd_A	705	153.57	14.93	3.00
4ys0_A	824	192.23	19.06	3.94
3nci_A	903	283.16	29.83	5.44
1gte_D	1025	403.68	40.97	7.89