Exploit your Latents: Coarse-Grained Protein Backmapping with Latent Diffusion Models

Rongchao Zhang[1], Yu Huang[2*] Yiwei Lou[1], Yi Xin[3], Haixu Chen[4], Yongzhi Cao[1], Hanpin Wang[1]

[1]Key Laboratory of High Confidence Software Technologies (Peking University), Ministry of Education, School of Computer Science, Peking University, Beijing, China

[2]National Engineering Research Center for Software Engineering, Peking University, Beijing, China

[3]National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

[4]Institute of Geriatrics&National Clinical Research Center of Geriatrics Disease, Chinese PLA General Hospital, Beijing, China

rcz@stu.pku.edu.cn, hy@pku.edu.cn

Abstract

Coarse-grained (CG) molecular dynamics of proteins is a preferred approach to studying large molecules on extended time scales by condensing the entire atomic model into a limited number of pseudo-atoms and preserving the thermodynamic properties of the system. However, the significantly increased efficiency impedes the analysis of substantial physicochemical information, since high-resolution atomic details are sacrificed to accelerate simulation. In this paper, we propose LatCPB, a generative approach based on diffusion that enables high-resolution backmapping of CG proteins. Specifically, our model encodes an all-atom into discrete latent embeddings, aligned with learnable multimodal discrete priors for circumventing posterior collapse and maintaining the discrete properties of the protein sequence. During the generation, we further design a latent diffusion process within the continuous latent space due to the potential stochastics in the data. Moreover, LatCPB performs a contrastive learning strategy in latent space to separate feature representations of various molecules and conformations of the same molecule, thus enhancing the comprehension of molecular representational diversity. Experimental results demonstrate that LatCPB is able to backmap CG proteins effectively and achieve outstanding performance.

Introduction

Coarse-grained (CG) models are utilized to facilitate scalable molecular dynamics simulations by simplifying atomic properties and their interactions, thus reducing computational demands and accelerating simulations, particularly in processes like aggregation (Jones, Shmilovich, and Ferguson 2023) and cardiolipin-selective (Mohr et al. 2022). Simultaneously, the ongoing attention to reintegrating missing all-atom details back into the CG protein structures enables researchers to derive particular structural insights from CG simulations.

Proteins are not strictly static entities but rather ensembles of occasionally similar conformations. Transitions between these states may even occur over length scales from $1/10$ $\mathring{A}$ to nm and time scales from ns to s (Conti Nibali and Paciaroni 2023). Previous backmapping works primarily adopt geometric rules or random placement to generate initial structures and subsequent refinement through Monte Carlo relaxation or molecular dynamics simulations, which exhibit high repeatability. Consequently, the choice of scoring functions and relaxation methods in these approaches must be appropriate and partially empirical (Yang and Gómez-Bombarelli 2023), leading to varying optimization outcomes that might not fully reflect the actual behavior of biomolecules. AlphaFold2 (Jumper et al. 2021) has shown a sensation in structural biology by precisely addressing the protein structure prediction problem, facilitating the gradual integration of AI-generated content into interdisciplinary molecular modeling. Then, several studies (Wang et al. 2022; Yang and Gómez-Bombarelli 2023) have achieved successful backmapping performance by introducing condition-based autoencoders for modeling all-atom distributions.

Nevertheless, the atomic feature space is typically composed of diverse physical quantities like atomic types and coordinates (Xu et al. 2023). These features exhibit multimodal attributes of discrete, integer, and continuous variables, particularly in protein sequences that are codes formed by discrete values (Santos et al. 2023; Ingraham et al. 2019). Protein sequences exhibit a high degree of regularity in nature and shift rapidly among few stable conformations, which are typically maintained by specific geometrical conformations (Li et al. 1996; Škrbić et al. 2024; Roche and Royer 2018). Current generative approaches (Wang et al. 2022; Yang and Gómez-Bombarelli 2023) directly model all-atom conformations in the continuous domain feature space of atoms, failing to capture the high-dimensional intricacies of input features. Recently, diffusion models have achieved state-of-the-art performance by modeling in latent spaces, including applications in image generation (Ni et al. 2023; Singh, Gould, and Zheng 2023), voice synthesis (Lee, Chung, and Chung 2023; Takahashi, Singh, and Mitsufuji 2023), and molecular design (Xu et al. 2023; Huang et al. 2023). Up until the date of this work, the exploration of latent diffusion models for backmapping has remained scarce. Besides, due to the dynamic diversity of protein conformational trajectories, a multitude of feasible atomic configurations can be associated with a single CG structure (Shmilovich et al. 2022). There is suffering from identifying features of various molecules and conformations of the same molecule while dealing with new.

*Corresponding author.

In this paper, we formulate the CG protein backmapping problem as a series of latent diffusion steps and propose LatCPB, which progressively transfers a CG structure in proteins to its target all-atom conformation within a latent embedding space. Specifically, we first encode the input all-atom conformation and corresponding CG structure by an all-atom graph encoder and a CG structure encoder respectively to obtain input embeddings for alignment. The idea of autoencoder reconstruction is to represent high-dimensional data distributions through low-dimensional latent encodings, with continuous representations typically undergoing intrinsic discretization by the encoder or decoder (van den Oord, Vinyals, and Kavukcuoglu 2017; Razavi, van den Oord, and Vinyals 2019). For modeling molecular conformations, we employ an attention layer to transform continuous latent embeddings into a finite set of discrete embeddings to satisfy the decoding quality and generative capability constraints, where directly aligning discrete embeddings with a discrete prior latent distribution leads to over-regularization and posterior collapse (Truong, Salah, and Lauw 2021; Peng et al. 2021). Therefore, we construct the discrete prior as a learnable multimodal latent space, since it is more reasonable to assume that the observed data are generated from a multimodal rather than an unimodal subspace (Bai, Kong, and Gomes 2022). Furthermore, similar features that frequently co-occur should have proximate embeddings; while two features rarely appear together, their embeddings should be significantly separated (Bai, Kong, and Gomes 2022). We design a contrastive learning strategy to constrain and separate feature representations of various molecules and conformations of the same molecule. At last, due to the potential stochastics in the data generation process, we further perform the latent diffusion model in the continuous encoding layer instead of the discrete layer to ensure that the representation does not become overly specialized to a single conformation.

Related Work

Backmapping of CG Proteins

Coarse-grained protein backmapping is the process of recovering from a coarse-grained representation to a high-resolution, atomic-level representation. As illustrated in Fig. 1, the position of an atom can be represented in internal coordinates (bond lengths, bond angles, and torsion angles) (Oenen, Dinu, and Liedl 2024; Li et al. 2023; Yang and Gómez-Bombarelli 2023). From the atoms in the protein backbone, the relative positions of other atoms within their vicinity can be determined through internal coordinates (Yang and Gómez-Bombarelli 2023; Grambow et al. 2023). Therefore, the generation of all-atom conformations involves the allocation of accurate 3D coordinates at the atomic level. Early efforts (Lombardi, Marti, and Capece 2016; Roel-Touris and Bonvin 2020) primarily relied on rule-based and scoring function-based sampling methods. These approaches typically constructed the initial structural framework using predefined geometric constraints, empirical rules, or physicochemical properties. The generated initial structure was then further refined through iterative optimization processes to ensure compliance with physicochemical constraints and achieve

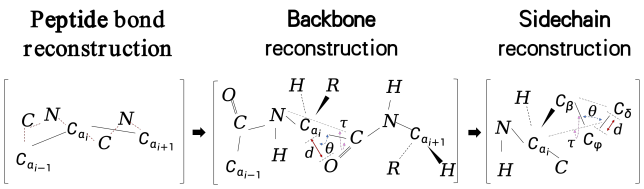


Figure 1: The position of an atom can be represented in internal coordinates (bond lengths, bond angles, and torsion angles) (Yang and Gómez-Bombarelli 2023; Li et al. 2023).

the lowest energy state, which is inefficient and tends to deviate from actual biophysical states due to excessive reliance on the choice of scoring functions. The introduction of AIGC offers a novel technical pathway for backmapping. Wang et al. (Wang et al. 2022) proposed a principled probabilistic formulation using AE, employing latent variables to approximate conditional distributions for modeling all-atom structural proteins. Although demonstrating promise in simple systems like alanine dipeptide and mini-proteins, this approach cannot be generalized beyond the chemical scope of their training. Yang et al. (Yang and Gómez-Bombarelli 2023) further utilized a conditional AE to model the internal coordinates of all atoms, achieving proficient reconstruction of bond topologies. However, these methods are based on AE architectures, which may fall short in performance and adaptability when dealing with complex proteins. We turn to latent diffusion models, which can handle high-dimensional data and complex structural dependencies more effectively.

Latent Diffusion Models

Different from traditional diffusion models that operate directly in data space (Sohl-Dickstein et al. 2015; Qin et al. 2023; Zhang et al. 2024), latent diffusion models (Kim and Kim 2024; Fabian, Tinaz, and Soltanolkotabi 2024) employ an encoder to map the raw data into latent space, perform the diffusion process in this space, and finally use a decoder to map the latent variables back to the original data. After achieving success in image generation (Ni et al. 2023; Singh, Gould, and Zheng 2023), latent diffusion models have been extended to molecular generation (Peng et al. 2023; Huang et al. 2023), exhibiting competitive or even better performance. Xu et al. (Xu et al. 2023) introduced the first latent diffusion model in the molecular geometry domain, consisting of autoencoders that map structures into continuous latent representations and diffusion models functioning within the latent space. Fu et al. (Fu et al. 2023) utilized an encoder to embed proteins into latent space, then employed a diffusion model to learn the distribution of latent protein representations, effectively generating novel protein backbone structures with high designability and efficiency. Wang et al. (Wang et al. 2024) trained diffusion models in latent space to generate molecules defined by gene expression profiles targeting biological activity, achieving outstanding performance on molecular generation benchmarks. Inspired by the successful application of latent diffusion models in molecular generation, we introduce it into the task of CG protein backmapping to explore a more precise generative approach.

Preliminary

Problem Formulation

Let $n \sim \mathcal{N}(0, I)$ be a Gaussian noise volume with shape $N_n \times A_n \times P_n$ where N_n, A_n, and P_n represent the number of beads in the CG structure, the number of all atoms in a bead, and the internal coordinates, respectively. For a given CG protein $\mathcal{M} = \{(x_m)|m \in \{1, \ldots, N\}\} \in \mathbb{R}^{N\times 3}$, with $\mathcal{Q} = \{x_j = (d_j, \theta_j, \tau_j)|j \in \{N+1, \ldots, N+A\}\} \in \mathbb{R}^{A\times 3}$ representing the real all-atom conformation of $\mathcal{M}$, the goal of the backmapping from CG protein to all-atom conformation is to learn a mapping that transforms a Gaussian noise container n into a synthetic all-atom conformation $\hat{\mathcal{Q}} = \{\hat{x}_j = (\hat{d}_j, \hat{\theta}_j, \hat{\tau}_j)|j \in \{N+1, \ldots, N+A\}\} \in \mathbb{R}^{A\times 3}$. Thus, the conditional distribution of $\hat{\mathcal{Q}}$ given $\mathcal{M}$ is equivalent to the conditional distribution of $\mathcal{Q}$ given $\mathcal{M}$, implying $p(\hat{\mathcal{Q}}|\mathcal{M}) = p(\mathcal{Q}|\mathcal{M})$. N and A represent the number of beads in the CG protein and the number of atoms in the corresponding all-atom conformation, excluding the beads, respectively. We coarse-grain the C_α atoms, as this approach has been widely applied in CG dynamical simulations (Zalewski, Kmiecik, and Koliński 2021; Badaczewska-Dawid, Kolinski, and Kmiecik 2020) and backmapping (Yang and Gómez-Bombarelli 2023).

Conditional Diffusion Model

Forward Process. The inspiration behind diffusion models (Ho, Jain, and Abbeel 2020; Chen et al. 2023) comes from nonequilibrium thermodynamics, formalized as two Markov chains: the forward diffusion process and the reverse process. In the forward diffusion process, the atom $x_j^{(0)}$ at time $t = 0$ is gradually infused with noise from a standard normal distribution. Following the variance schedule $\beta^1, \ldots, \beta^T (\beta^t \in (0, 1))$, and ultimately over T steps, the data distribution is transformed into a simple prior distribution $x_j^{(T)} \sim p(x_j^{(T)})$, formalized as follows:

$$q(x_j^{(1:T)}|x_j^{(0)}) = \prod_{t=1}^{T} q(x_j^{(t)}|x_j^{(t-1)}), \text{and} \tag{1}$$

$$q(x_j^{(t)}|x_j^{(t-1)}) = \mathcal{N}(x_j^{(t)}; \sqrt{1-\beta^t}x_j^{(t-1)}, \beta^t I), \tag{2}$$

where $\mathcal{N}$ is a Gaussian distribution, $x_j^{(t)}$ is obtained by adding noise to $x_j^{(t-1)}$ from time step $t-1$ to time step t. By marginalizing the joint distribution $q(x_j^{(1:T)}|x_j^{(0)})$, given $x_j^{(0)}$, $x_j^{(t)}$ can be directly obtained as follows:

$$q(x_j^{(t)}|x_j^{(0)}) = \mathcal{N}(x_j^{(t)}; \sqrt{\overline{\alpha}_t}x_j^{(0)}, (1-\overline{\alpha}_t)I), \tag{3}$$

where $\overline{\alpha}_t = \prod_{\hat{\tau}=1}^{t} 1 - \beta^{\hat{\tau}}$, thereby allowing us to sample from $q(x_j^{(t)}|x_j^{(0)})$ at any time step t : $x_j^{(t)} = \overline{\alpha}_t x_j^{(0)} + \sqrt{1-\overline{\alpha}_t}\epsilon$ and $\epsilon \sim \mathcal{N}(0, I)$. When $\sqrt{\overline{\alpha}_t} \to 0$ and $x_j^{(T)}$ approximates a standard Gaussian distribution, then $q(x_j^{(T)}) = \int q(x_j^{(T)}|x_j^{(0)})q(x_j^{(0)})dx_j^{(0)} \to \mathcal{N}(0, 1)$.

Reverse Process. Assuming there exists a reverse process that can incrementally denoise the noise variable $x_j^{(T:1)}$ under the guidance of condition $\mathcal{C}$ and approximate the target data $x_j^{(0)}$, starting from a standard Gaussian distribution, and following a Markov chain from $t = T$ back to $t = 0$, as follows:

$$p_\theta(x_j^{(0:T-1)}|\mathcal{Q}^T, \mathcal{C}) = \prod_{t=1}^{T} p_\theta(x_j^{(t-1)}|\mathcal{Q}^t, \mathcal{C}), \text{and} \tag{4}$$

$$p_\theta(x_j^{(t-1)}|\mathcal{Q}^t, \mathcal{C}) = \mathcal{N}(x_j^{(t-1)}; \mu_\theta(\mathcal{Q}^t, \mathcal{C}, t), \sigma_\theta(\mathcal{Q}^t, \mathcal{C}, t)), \tag{5}$$

where θ represents a learnable parameter, and $\mathcal{Q}^t = [x_{N+1}^{(t)}, \ldots, x_{N+A}^{(t)}] \in \mathbb{R}^{A\times 3}$ is the all-atom conformation sampled at time step t.

Latent Diffusion

Given an atom $x_j^{(0)}$ and an autoencoder that includes an encoder $\mathcal{E}$ and a decoder $\mathcal{D}$, the corresponding latent representation $z_j^{(0)} = \mathcal{E}(x_j^{(0)})$ can be encoded. By replacing the atomic data x_j in Eq. (1) and (4) with the latent representation z_j, the latent diffusion and denoising processes can be computed as:

$$q(z_j^{(1:T)}|z_j^{(0)}) = \prod_{t=1}^{T} q(z_j^{(t)}|z_j^{(t-1)}), \text{and} \tag{6}$$

$$p_\theta(z_j^{(0:T-1)}|\mathcal{Q}^T, \mathcal{C}) = \prod_{t=1}^{T} p_\theta(z_j^{(t-1)}|\mathcal{Q}^t, \mathcal{C}). \tag{7}$$

During the inference stage, the final output atoms are reconstructed from the denoised latent $\hat{x}_j^{(0)} = \mathcal{D}(\hat{z}_j^{(0)})$, with $\hat{z}_j^{(0)}$ being sampled and denoised using Eq. (5):

$$\hat{z}_j^{(t-1)} = \frac{1}{\sqrt{1-\beta^t}}(\hat{z}_j^{(t)} - \frac{\beta^t}{\sqrt{1-\overline{\alpha}_t}}\epsilon_\theta(\mathcal{Q}^t, \mathcal{C})) + \rho_t\epsilon_t, \tag{8}$$

where $\epsilon_j \sim \mathcal{N}(0, I)$.

Methodology

Discrete All-Atom Representations

For a given all-atom protein structure, we begin learning the mapping between the all atoms and their discrete representations to maintain the discrete properties of the protein sequence (Ingraham et al. 2019; Santos et al. 2023). This is achieved by learning an autoencoder for a discrete latent space. Specifically, to represent a C_α atom $x_i \in \mathbb{R}^{S\times 3}$ and its corresponding all atom $\mathcal{Q}_i \in \mathbb{R}^{N_i\times S\times 3}$, we train a all-atom structure encoder $\mathcal{E}$ that yields a low-dimensional parameter $\hat{z}_i$, pertaining to the corresponding distribution $\mathcal{E}(\mathcal{Q}_i, x_i)$. Where S represents the number of conformations in a protein molecule. We employ a transformer (Gorishniy et al. 2021; Song et al. 2019) followed by a Sigmoid nonlinearity σ to regularize the parameter $\hat{zt}_i \in \mathbb{R}^{S\times P} = \psi(\mathcal{E}(\mathcal{Q}_i, x_i))$, where P represents the dimensions of latent encoding. To obtain a discrete representation of the atoms, we perform polynomial sampling, as follows:

$$\hat{y}_i = \delta(\text{Multinomial}(\hat{zt}_i, 1), K), \tag{9}$$

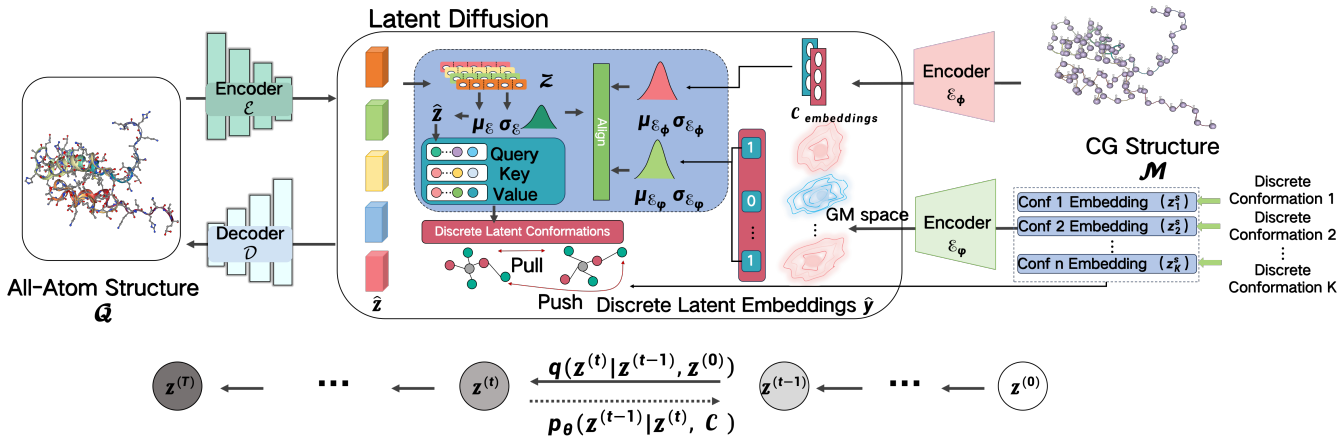


Figure 2: Overview of the LatCPB for backmapping synthesis of CG protein conformations. The diffusion process at each step is performed in the latent space of the autoencoder.

where $\delta(\cdot)$ denotes a function that transforms embeddings into an K-dimensional one-hot vector.

To obtain a prior for the discrete latent representation, we design a multimodal latent space. Through an encoder $\mathcal{E}_\varphi$, we directly map each discrete conformation $\hat{y}_{i,k}$ of conformation k to an individual latent Gaussian distribution $\mathcal{N}(\mu_k, \mathrm{diag}(\sigma_k^2))$. By sharing the same encoder, the randomly initialized embeddings z_k^s are learnable during the training process, where each discrete conformation activates a positive Gaussian $(\hat{y}_{i,k} = 1)$ and forms a Gaussian mixture subspace (Bai, Kong, and Gomes 2022). The definition of the probability distribution within the subspace is as follows:

$$p_\varphi(\hat{z}_{i,s}|\hat{y}_i) = \frac{1}{\sum_k \hat{y}_{i,k}} \sum_{k=1}^{K} \mathbb{1}\{\hat{y}_{i,k}=1\} \mathcal{N}(\hat{z}_{i,s}|\mu_k, \mathrm{diag}(\sigma_k^2)), \quad (10)$$

where $\mathbb{1}(\cdot)$ is the indicator function. The multimodal latent space can be aligned using the evidence lower bound, as follows:

$$\mathcal{L}_M^t \approx \log q_\theta(\hat{z}_{i,s}|Q_i, x_i) - \log p_\varphi(\hat{z}_{i,s}|\hat{y}_i), \quad (11)$$

where the all-atom structure encoder $\mathcal{E}$ is parameterized by θ. The reconstruction of the all-atom conformation is obtained using a decoder $\mathcal{D}$, denoted as $\hat{Q}_i$. The training objective is as follows:

$$\mathcal{L}_R^t = \sum_{i=1}^{N} \sum_{m}^{\|\mathcal{C}\|} \omega_m \mathcal{C}[m](\hat{Q}_i, Q_i), \quad (12)$$

where $\mathcal{C}$ represents a collection of loss functions (Yang and Gómez-Bombarelli 2023), including the reconstruction error of atomic Cartesian coordinates $\|\hat{X}_i - X_i\|_2^2$, bond length error $(\hat{d}_i, d_i)^2$, bond angle error $\sqrt{2(1-(\cos(\hat{\theta}_i - \theta_i)))} + \epsilon$, and torsion angle error $\sqrt{2(1-(\cos(\hat{\tau}_i - \tau_i)))} + \epsilon$. ϵ is a smoothing term, set to 10^{-7}. By decoding in the continuous embedding layer, it ensures that the representation does not become overly specialized to a single conformation. Additionally, to ensure the chemical validity of generated structures, we introduce a spatial collision loss (Yang and Gómez-Bombarelli 2023), calculated as follows:

$$\mathcal{L}_E^t = \sum_{j=N}^{N+A} \sum_{\hat{Y}_j \in G(\hat{X}_j)} \max(2.0 - \|\hat{X}_j - \hat{Y}_j\|_2^2, 0.0), \quad (13)$$

where $G(\hat{X}_j)$ is the set of atoms within a cutoff distance of $5\mathring{A}$ from atom $\hat{X}_j$.

Latent Contrastive Learning

Driven by the success of contrastive learning in tasks (Cherti et al. 2023; Dong et al. 2023; Lou et al. 2024) such as image-language pre-training, we employ a contrastive learning strategy (Bai, Kong, and Gomes 2022) to constrain and differentiate the feature representations of different molecules and different conformations of the same molecule. Let $G = \{1, \ldots, K\}$, and define $P(\hat{y}_i) \equiv P\{k \in G : \hat{y}_{i,k} = 1\}$ with respect to the mapping pair $(x_i, \hat{y}_i)$. For a protein mixture $\mathcal{M}$, contrastive learning can be formulated as:

$$\mathcal{L}_C^t = \frac{1}{|N|} \sum_{(x_i,\hat{y}_i)\in\mathcal{M}} \frac{1}{|P(\hat{y}_i)|} \times \sum_{p\in P(\hat{y}_i)} -\log \frac{\mathrm{sim}(\hat{z}_i, z_p^s)}{\sum_{t\in G} \mathrm{sim}(\hat{z}_i, z_t^s)}, \quad (14)$$

where $\mathrm{sim}(z_1, z_2) = \exp(z_1 \cdot z_2/\tau)$ is a function that measures the similarity between two embeddings, z_p^s and z_t^s represent the embeddings of discrete representations. τ is the temperature parameter that controls the scale of the product.

Then the overall loss $\mathcal{L}^t$ for training the mixture autoencoder can be given by:

$$\mathcal{L}^t = \lambda_M \mathcal{L}_M^t + \lambda_R \mathcal{L}_R^t + \lambda_E \mathcal{L}_E^t + \lambda_C \mathcal{L}_C^t + \lambda_K \mathcal{L}_K^t, \quad (15)$$

where λ_* are hyper-parameters for the weights of each item. $\mathcal{L}_K^t$ is the Kullback Leibler term derived from the difference between coarse-grained prior encoder $\mathcal{E}_\phi$ and encoder $\mathcal{E}$.

Metric	Method	Structure			
		PED00055 (55 frames)	**PED00090 (27 frames)**	**PED00151 (140 frames)**	**PED00218 (20 frames)**
RMSD ($\mathring{A}$ **;** $\downarrow$)	CGVAE	2.093	2.134	2.390	1.967
	GenZProt	1.871	**0.029**	1.711	1.727
	LatCPB	**1.695**	1.758	**1.539**	**1.563**
GED ($\downarrow$)	CGVAE	0.212	0.204	0.376	0.140
	GenZProt	0.054	0.070	0.019	0.030
	LatCPB	**0.039**	**0.055**	**0.015**	**0.020**
KL ($\downarrow$)	CGVAE	0.369	0.424	0.316	0.331
	GenZProt	0.568	0.758	0.535	0.541
	LatCPB	**0.153**	**0.124**	**0.108**	**0.087**

Table 1: Quantitative comparison among different methods on the PED00055, PED00090, PED00151 and PED00218 structures. We report the mean values of all frames in each set. The best results in three metrics are highlighted in bold.

Conditional Latent Diffusion Model

To ensure complex CG-all-atom alignment and the intrinsic sequence independence among atoms during the diffusion process, we upgrade the diffusion process with additional semantic priors. Equipped with the autoencoders $\mathcal{E}$ and $\mathcal{D}$, we are able to represent structures $\mathcal{Q}$ and $\mathcal{M}$ through low-dimensional latent variables $\hat{z}_i$, while preserving their topological and chemical characteristics. Then, a transformer-based latent diffusion model is trained to synthesize a latent sequence. Compared to the initial atomic features of high-dimensional complex data, the encoded latent low-dimensional space significantly aids the likelihood-based generative model (Xu et al. 2023).

Given a CG protein mixture $\mathcal{M}$ and its corresponding all-atom $\mathcal{Q}$, we initially employ the CG structure encoder $\mathcal{E}_\phi$ to compute the latent encoding sequence from the first C_α atom to the N-th C_α atom, $Z^C = \{z_1, \ldots, z_N\} \in \mathbb{R}^{N\times S\times K_1}$, and the encoding sequence within the all-atom structure as $\hat{Z}^A = \{\hat{z}_1, \ldots, \hat{z}_N\} \in \mathbb{R}^{N\times S\times K_2}$, through the training of $\mathcal{E}$. K_1 and K_2 are the dimensions of the latent vectors. Z^C is used as a condition for the diffusion model. The atom $z_j^{(0)} \in \hat{Z}^A$ is mapped to a standard Gaussian noise volume $n \sim \mathcal{N}(0, I)$ by gradually introducing Gaussian noise through the forward process of the diffusion model. Specifically, the denoising model $\epsilon_\theta(z_j^{(t)}, \mathcal{M})$ is trained to predict the noise ϵ_j added to $z_j^{(t)}$, based on a transformer network, with the loss defined as follows:

$$\mathcal{L}_{DM} = \mathbb{E}_{z_j^{(0)},\epsilon_j,t}[\|\epsilon_j - \epsilon_\theta(z_j^{(t)}, \mathcal{M})\|^2], \qquad (16)$$

where the time step t is sampled from $\{1, \ldots, T\}$.

Equivariant Graph Neural Networks

Molecular Graph Construction. Due to the inherent graph-like structure of protein molecules, which allows their complex and dynamic 3D conformations to be effectively encoded in graph models. Following prior work (Yang and Gómez-Bombarelli 2023), we establish a protein molecular graph with residues and atoms as nodes, and their identities as initial node features. In constructing the graph model, edges are established based on the distances between atoms, where atom pairs within 9 $\mathring{A}$ and C_α atom pairs within 21 $\mathring{A}$ undergo global message passing. To enhance message passing between local structures, we also establish additional connections between atoms within the same residue and between pairs of residues. Such information transfer on three levels is effective in capturing the spatial geometry of proteins. Then, the encoder $\mathcal{E}$ performs message passing over the entire molecular graph, while the CG encoder $\mathcal{E}_\phi$ captures only the interaction information between C_α atom pairs.

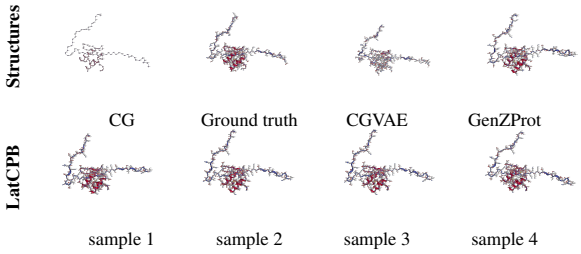


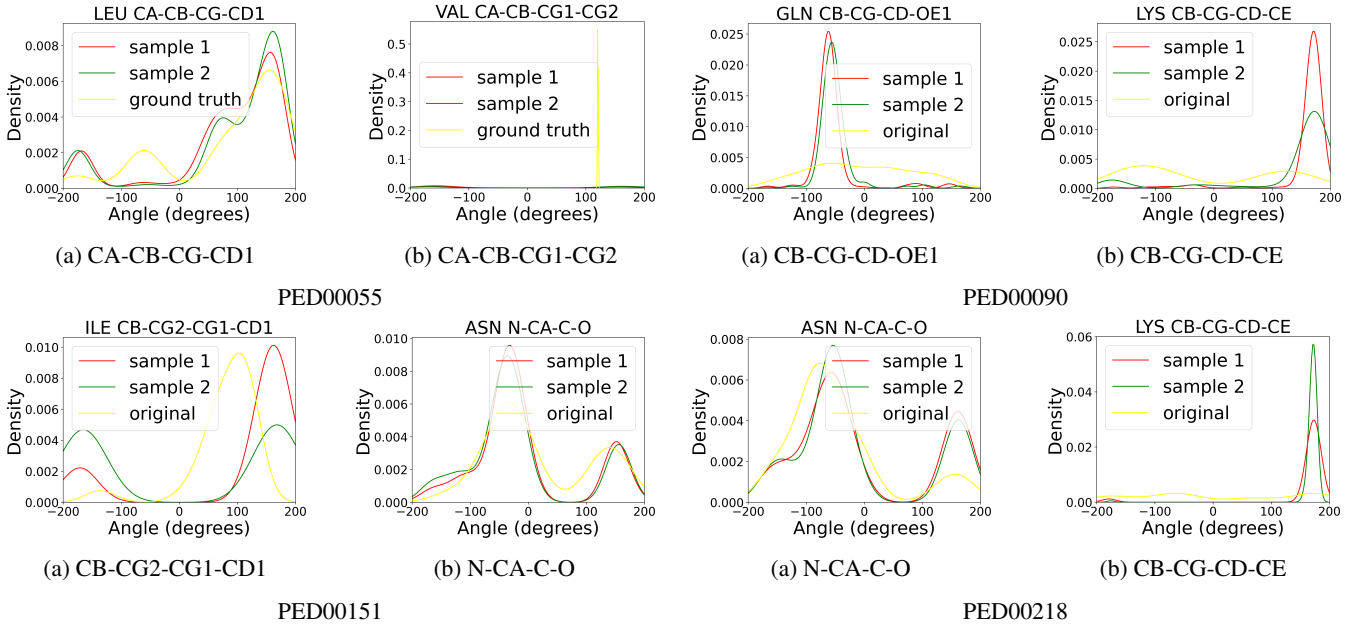
Figure 3: The qualities of the generated 3D molecules.

Messaging with Neural Networks. The encoder network aims to encode the all-atom structure and contextual interaction information: $\mathcal{M} \cup \mathcal{Q}$ into the latent space and reconstruct the structure $\hat{\mathcal{Q}}$. Then, the denoising network performs denoising the noisy all-atom latent features at time step t, guided by the CG structure representation. Specifically, since the spatial coordinates of each atom exhibit symmetry and equivariance, following (Yang and Gómez-Bombarelli 2023), we employ two SE(3) neural networks (Geiger et al. 2022; Corso et al. 2023) to implement encoders $\mathcal{E}$ and $\mathcal{E}_\phi$. For the decoder $\mathcal{D}$, to accurately model the joint distribution of internal coordinates, it is preferable to allow flexibility within physical ranges. Since bond lengths follow an unimodal Gaussian distribution with a smaller variance (Yang and Gómez-Bombarelli 2023), we utilize a lookup table based on residue types. Network $\mathcal{D}_\phi$ employs message passing and pooling operations on node feature vectors, followed by multilayer perceptron layers.

Structures	DL	CL	RMSD	GED	KL
PED00055 (55 frames)	-	-	1.743	0.042	0.212
	✓	-	1.736	0.062	0.229
	✓	✓	**1.695**	**0.039**	**0.153**
PED00090 (27 frames)	-	-	1.810	0.108	0.942
	✓	-	1.806	0.144	0.443
	✓	✓	**1.758**	**0.055**	**0.124**

Structures	DL	CL	RMSD	GED	KL
PED00151 (140 frames)	-	-	1.557	**0.010**	0.212
	✓	-	1.549	0.022	0.167
	✓	✓	**1.539**	0.015	**0.108**
PED00218 (20 frames)	-	-	1.620	0.033	0.094
	✓	-	1.595	0.044	0.168
	✓	✓	**1.563**	**0.020**	**0.087**

Table 2: Ablation study for each component of LatCPB.



(a) CA-CB-CG-CD1 (b) CA-CB-CG1-CG2 (a) CB-CG-CD-OE1 (b) CB-CG-CD-CE

PED00055 PED00090

(a) CB-CG2-CG1-CD1 (b) N-CA-C-O (a) N-CA-C-O (b) CB-CG-CD-CE

PED00151 PED00218

Figure 4: KDE plots of torsion angles from the structures generated.

The denoising network $\epsilon_\theta(z_j^{(t)}, \mathcal{M})$ is implemented based on the transformer architecture (Vaswani et al. 2017).

Experiments

Experimental Setup

Dataset. We evaluate the validity of our method using the PED protein dataset (Lazar et al. 2021), a structural collection of intrinsically disordered proteins (IDPs). PED is currently the only database focused on representing the diversity of IDP collections, focusing on biologically interesting protein regions with conformational collections. Since IDPs are notoriously sensitive to conditioning, these alternative collections may provide very valuable insights into the conditional disorder of these proteins.

Following previous work (Yang and Gómez-Bombarelli 2023), we exclude 17 metal ion-binding complexes, 4 nucleotide-binding complexes, 5 cofactor-binding complexes and 1 D-amino acid protein. In addition, 13 proteins modelled or measured under unnatural conditions are excluded, as well as 8 protein post-translational modifications other than phosphorylation and oxidation.

Implementation Details. As our approach utilizes three C_α as anchors to rebuild backbone nitrogen and carbon, it is incapable of reconstructing the atomic positions of terminal residues (Yang and Gómez-Bombarelli 2023). Consequently, all terminal residues are masked. To prevent certain entries from being overly represented in the model, a subset is extracted from entries exceeding 500 frames. In the experiments, we utilize approximately 10,000 frames as the training set and select about 240 frames as test data, which come from four different structures: PED00055, PED00090, PED00151, and PED00218. We report the average performance of the model to ensure robustness of the results. The implementation environment is PyTorch 2.1 version and the Adam optimizer is applied to train the model with the learning rate of 10^{-3} and decayed to zero with a scheduler.

To evaluate the performance, we compare our model with the prior arts that focus on model improvement, including CGVAE (Wang et al. 2022) and GenZProt (Yang and Gómez-Bombarelli 2023). For fair comparisons, we reproduce all methods under the same implemental environment. In our experiments, we evaluate the generated all-atom structures using three different metrics: Root Mean Square Distance (RMSD), Graph Edit Distance (GED) and Kullback–Leibler

(KL). RMSD calculates the root-mean-square deviation between the reference structure and the structure generated by the model to evaluate the accuracy of structural alignment. GED measures the graph edit distance between the generated molecular graph and the reference molecular graph. KL is used to evaluate the statistical consistency between the generated structure and the experimentally obtained reference protein structure.

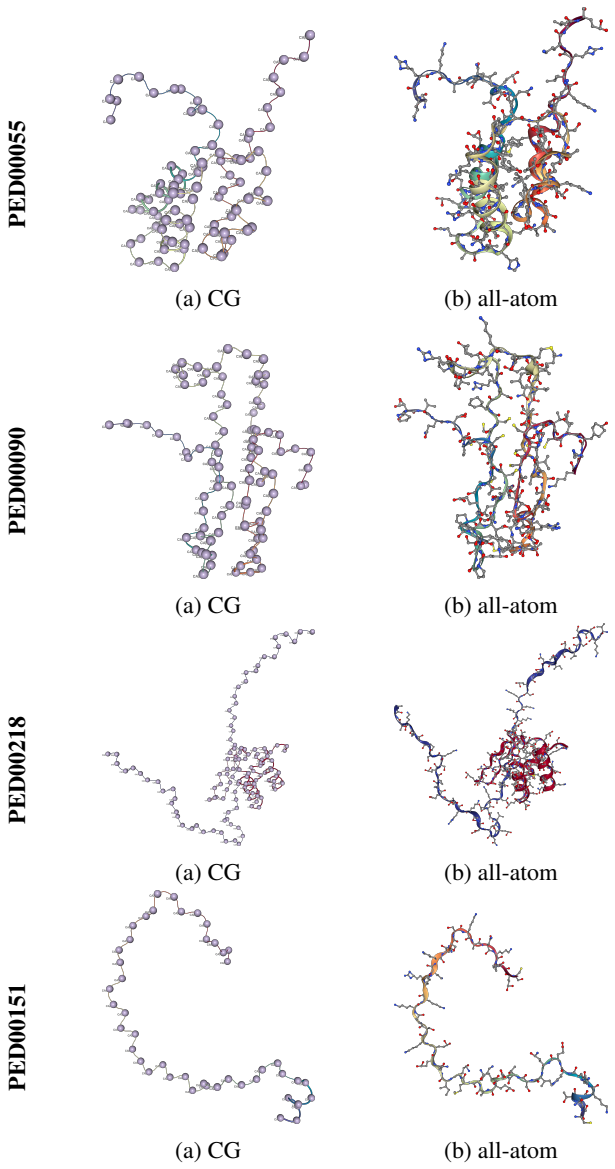


Figure 5: Assessment of quality in the backmapping from CG to all-atom structures for four proteins: PED00055, PED00090, PED00218, and PED00151.

Results and Analysis

As shown in Table 1, highlight that LatCPB excels in the RMSD metric on PED00055, PED00151, and PED00218, underscoring its superior capability in spatial structure reconstruction. While GenZProt achieves a notably lower RMSD of 0.029 $\mathring{A}$ on PED00090, LatCPB consistently performs excellently on the remaining other protein structures. LatCPB exhibits the lowest GED across all tested proteins, particularly on PED00151. For KL divergence, LatCPB also shows the lowest values in all test cases, reflecting high consistency between its generated protein structures and the reference structures obtained experimentally.

Fig. 3 shows that the nodes within the PED00218 generated by CGVAE are dispersed, indicating its limited ability to maintain complex internal interactions and spatial proximity. Although GenZProt restores some details, it falls short of LatCPB in maintaining overall structural consistency and accurate alignment. Fig. 4 illustrates the torsion angle distributions between ground truth and synthetic proteins. For VAEs, an inherent issue is that the learning objective is to minimize the reverse KL divergence, which often leads to the model overlooking less frequent but biologically significant patterns. In contrast, LatCPB, by introducing randomness and an improved CL objective, is better able to capture the distribution of the data, including those rarer or more complex patterns.

Ablation Studies

We conduct ablation studies on the LatCPB to evaluate the impacts of the contrastive learning (CL) and discrete latent space (DL) components on its performance. As shown in Table 2, with DL enabled alone, we notice an improvement in performance, albeit limited. When DL and CL are used together, performance across all metrics reaches its peak. DL provides a structural enhancement base that bolsters the model's capability to explore the data space, while CL improves the model's discrimination ability, further enhancing its performance. These components are particularly effective in enhancing performance and adaptability when processing proteins with complex spatial structures.

Conclusions and Limitations

In this work, we propose a CG protein backmapping model based on latent diffusion. Our model is capable of extensively performing CG backmapping tasks and achieves competitive performance. A major limitation of this work is its reliance on the corresponding CG structures. More comprehensive datasets should be included for evaluation. Thus, we leave it for future work to collect more validated proteins and design an effective model to backmap more diverse all-atomic structures. Another limitation is that it remains unclear whether the all-atom structures of the currently produced mixtures have reasonable binding energies. More efforts are needed to design all-atomic structures with biological activity.

Acknowledgments

This paper was supported by National Key Research and Development Program of China (2021YFF1201100), National Natural Science Foundation of China under Grants (62172016 and 61932001), Beijing Nova Program and Scientific and Technological Innovation Project of China Academy of Chinese Medical Sciences (CI2023C062YLL).

References

Badaczewska-Dawid, A. E.; Kolinski, A.; and Kmiecik, S. 2020. Computational reconstruction of atomistic protein structures from coarse-grained models. *Comput. Struct. Biotechnol. J.*, 18: 162–176.

Bai, J.; Kong, S.; and Gomes, C. P. 2022. Gaussian Mixture Variational Autoencoder with Contrastive Learning for Multi-Label Classification. In Chaudhuri, K.; Jegelka, S.; Song, L.; Szepesvári, C.; Niu, G.; and Sabato, S., eds., *International Conference on Machine Learning*, 1383–1398.

Chen, X.; He, J.; Han, X.; and Liu, L. 2023. Efficient and Degree-Guided Graph Generation via Discrete Diffusion Modeling. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International Conference on Machine Learning*, 4585–4610.

Cherti, M.; Beaumont, R.; Wightman, R.; Wortsman, M.; Ilharco, G.; Gordon, C.; Schuhmann, C.; Schmidt, L.; and Jitsev, J. 2023. Reproducible Scaling Laws for Contrastive Language-Image Learning. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2818–2829.

Conti Nibali, V.; and Paciaroni, A. 2023. Virtual issue on fast dynamics and function of biomolecules. *J. Phys. Chem. Lett.*, 14(20): 4786–4788.

Corso, G.; Stärk, H.; Jing, B.; Barzilay, R.; and Jaakkola, T. S. 2023. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. In *The Eleventh International Conference on Learning Representations*.

Dong, X.; Bao, J.; Zheng, Y.; Zhang, T.; Chen, D.; Yang, H.; Zeng, M.; Zhang, W.; Yuan, L.; Chen, D.; Wen, F.; and Yu, N. 2023. MaskCLIP: Masked Self-Distillation Advances Contrastive Language-Image Pretraining. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10995–11005.

Fabian, Z.; Tinaz, B.; and Soltanolkotabi, M. 2024. Adapt and Diffuse: Sample-adaptive Reconstruction via Latent Diffusion Models. In *Forty-first International Conference on Machine Learning*.

Fu, C.; Yan, K.; Wang, L.; Au, W. Y.; McThrow, M.; Komikado, T.; Maruhashi, K.; Uchino, K.; Qian, X.; and Ji, S. 2023. A Latent Diffusion Model for Protein Structure Generation. In Villar, S.; and Chamberlain, B., eds., *Learning on Graphs Conference*.

Geiger, M.; Smidt, T.; M., A.; Miller, B. K.; Boomsma, W.; Dice, B.; Lapchevskyi, K.; Weiler, M.; Tyszkiewicz, M.; Batzner, S.; Madisetti, D.; Uhrin, M.; Frellsen, J.; Jung, N.; Sanborn, S.; Wen, M.; Rackers, J.; Rød, M.; and Bailey, M. 2022. e3nn/e3nn: 2022-04-13.

Gorishniy, Y.; Rubachev, I.; Khrulkov, V.; and Babenko, A. 2021. Revisiting Deep Learning Models for Tabular Data. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y. N.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems*, 18932–18943.

Grambow, C. A.; Weir, H.; Diamant, N. L.; Tseng, A. M.; Biancalani, T.; Scalia, G.; and Chuang, K. V. 2023. RINGER: Rapid Conformer Generation for Macrocycles with Sequence-Conditioned Internal Coordinate Diffusion. In *The Eleventh International Conference on Learning Representations*.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising Diffusion Probabilistic Models. In Larochelle, H.; Ranzato, M.; Hadsell, R.; Balcan, M.; and Lin, H., eds., *Advances in Neural Information Processing Systems*.

Huang, L.; Zhang, H.; Xu, T.; and Wong, K. 2023. MDM: Molecular Diffusion Model for 3D Molecule Generation. In Williams, B.; Chen, Y.; and Neville, J., eds., *Thirty-Seventh AAAI Conference on Artificial Intelligence*, 5105–5112.

Ingraham, J.; Garg, V. K.; Barzilay, R.; and Jaakkola, T. S. 2019. Generative Models for Graph-Based Protein Design. In Wallach, H. M.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E. B.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, 15794–15805.

Jones, M. S.; Shmilovich, K.; and Ferguson, A. L. 2023. DiAMoNDBack: Diffusion-denoising Autoregressive Model for Non-Deterministic Backmapping of Cα protein traces. *J. Chem. Theory Comput.*, 19(21): 7908–7923.

Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; and Hassabis, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873): 583–589.

Kim, J.; and Kim, T. 2024. Arbitrary-Scale Image Generation and Upsampling Using Latent Diffusion Model and Implicit Neural Decoder. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2024, Seattle, WA, USA, June 16-22, 2024*, 9202–9211. IEEE.

Lazar, T.; Martínez-Pérez, E.; Quaglia, F.; Hatos, A.; Chemes, L. B.; Iserte, J. A.; Méndez, N. A.; Garrone, N. A.; Saldaño, T. E.; Marchetti, J.; Rueda, A. J. V.; Bernadó, P.; Blackledge, M.; Cordeiro, T. N.; Fagerberg, E.; Forman-Kay, J. D.; Fornasari, M. S.; Gibson, T. J.; Gomes, G.-N. W.; Gradinaru, C. C.; Head-Gordon, T.; Jensen, M. R.; Lemke, E. A.; Longhi, S.; Marino-Buslje, C.; Minervini, G.; Mittag, T.; Monzon, A. M.; Pappu, R. V.; Parisi, G.; Ricard-Blum, S.; Ruff, K. M.; Salladini, E.; Skepö, M.; Svergun, D.; Vallet, S. D.; Varadi, M.; Tompa, P.; Tosatto, S. C. E.; and Piovesan, D. 2021. PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic Acids Res.*, 49(D1): D404–D411.

Lee, J.; Chung, J. S.; and Chung, S. 2023. Imaginary Voice: Face-Styled Diffusion Model for Text-to-Speech. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1–5.

Li, H.; Helling, R.; Tang, C.; and Wingreen, N. 1996. Emergence of preferred structures in a simple model of protein folding. *Science*, 273(5275): 666–669.

Li, J.; Zhang, O.; Lee, S.; Namini, A.; Liu, Z. H.; Teixeira, J. M. C.; Forman-Kay, J. D.; and Head-Gordon, T. 2023. Learning correlations between internal coordinates to improve 3D Cartesian coordinates for proteins. *J. Chem. Theory Comput.*, 19(14): 4689–4700.

Lombardi, L. E.; Marti, M. A.; and Capece, L. 2016. CG2AA: backmapping protein coarse-grained structures. *Bioinform.*, 32(8): 1235–1237.

Lou, Y.; Zhang, J.; Xu, D.; Cao, Y.; Wang, H.; and Huang, Y. 2024. No-Reference MRI Quality Assessment via Contrastive Representation: Spatial and Frequency Domain Perspectives. In *2024 IEEE International Conference on Multimedia and Expo (ICME)*, 1–6. IEEE.

Mohr, B.; Shmilovich, K.; Kleinwächter, I. S.; Schneider, D.; Ferguson, A. L.; and Bereau, T. 2022. Data-driven discovery of cardiolipin-selective small molecules by computational active learning. *Chem. Sci.*, 13(16): 4498–4511.

Ni, H.; Shi, C.; Li, K.; Huang, S. X.; and Min, M. R. 2023. Conditional Image-to-Video Generation with Latent Flow Diffusion Models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 18444–18455.

Oenen, K.; Dinu, D. F.; and Liedl, K. R. 2024. Determining internal coordinate sets for optimal representation of molecular vibration. *J. Chem. Phys.*, 160(1).

Peng, J.; Liu, D.; Xu, S.; and Li, H. 2021. Generating Diverse Structure for Image Inpainting With Hierarchical VQ-VAE. In *IEEE Conference on Computer Vision and Pattern Recognition*, 10775–10784.

Peng, X.; Guan, J.; Liu, Q.; and Ma, J. 2023. MolDiff: Addressing the Atom-Bond Inconsistency Problem in 3D Molecule Diffusion Generation. In *International Conference on Machine Learning*, 27611–27629.

Qin, Y.; Zheng, H.; Yao, J.; Zhou, M.; and Zhang, Y. 2023. Class-Balancing Diffusion Models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 18434–18443.

Razavi, A.; van den Oord, A.; and Vinyals, O. 2019. Generating Diverse High-Fidelity Images with VQ-VAE-2. In Wallach, H. M.; Larochelle, H.; Beygelzimer, A.; d'Alché Buc, F.; Fox, E. B.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, 14837–14847.

Roche, J.; and Royer, C. A. 2018. Lessons from pressure denaturation of proteins. *J. R. Soc. Interface*, 15(147): 20180244.

Roel-Touris, J.; and Bonvin, A. M. J. J. 2020. Coarse-grained (hybrid) integrative modeling of biomolecular interactions. *Comput. Struct. Biotechnol. J.*, 18: 1182–1190.

Santos, J. E.; Fox, Z. R.; Lubbers, N.; and Lin, Y. T. 2023. Blackout Diffusion: Generative Diffusion Models in Discrete-State Spaces. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International Conference on Machine Learning*, 9034–9059.

Shmilovich, K.; Stieffenhofer, M.; Charron, N. E.; and Hoffmann, M. 2022. Temporally coherent backmapping of molecular trajectories from coarse-grained to atomistic resolution. *J. Phys. Chem. A*, 126(48): 9124–9139.

Singh, J.; Gould, S.; and Zheng, L. 2023. High-Fidelity Guided Image Synthesis with Latent Diffusion Models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5997–6006.

Škrbić, T.; Giacometti, A.; Hoang, T. X.; Maritan, A.; and Banavar, J. R. 2024. III. Geometrical framework for thinking about globular proteins: Turns in proteins. *Proteins*.

Sohl-Dickstein, J.; Weiss, E. A.; Maheswaranathan, N.; and Ganguli, S. 2015. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. In *International Conference on Machine Learning*, 2256–2265.

Song, W.; Shi, C.; Xiao, Z.; Duan, Z.; Xu, Y.; Zhang, M.; and Tang, J. 2019. AutoInt: Automatic Feature Interaction Learning via Self-Attentive Neural Networks. In Zhu, W.; Tao, D.; Cheng, X.; Cui, P.; Rundensteiner, E. A.; Carmel, D.; He, Q.; and Yu, J. X., eds., *International Conference on Information and Knowledge Management*, 1161–1170.

Takahashi, N.; Singh, M. K.; and Mitsufuji, Y. 2023. Hierarchical Diffusion Models for Singing Voice Neural Vocoder. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1–5.

Truong, Q.; Salah, A.; and Lauw, H. W. 2021. Bilateral Variational Autoencoder for Collaborative Filtering. In Lewin-Eytan, L.; Carmel, D.; Yom-Tov, E.; Agichtein, E.; and Gabrilovich, E., eds., *The Fourteenth ACM International Conference on Web Search and Data Mining*, 292–300.

van den Oord, A.; Vinyals, O.; and Kavukcuoglu, K. 2017. Neural Discrete Representation Learning. In Guyon, I.; von Luxburg, U.; Bengio, S.; Wallach, H. M.; Fergus, R.; Vishwanathan, S. V. N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, 6306–6315.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is All you Need. In Guyon, I.; von Luxburg, U.; Bengio, S.; Wallach, H. M.; Fergus, R.; Vishwanathan, S. V. N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, 5998–6008.

Wang, C.; Ong, H. H.; Chiba, S.; and Rajapakse, J. C. 2024. GLDM: hit molecule generation with constrained graph latent diffusion model. *Briefings Bioinform.*, 25(3).

Wang, W.; Xu, M.; Cai, C.; Miller, B. K.; Smidt, T. E.; Wang, Y.; Tang, J.; and Gómez-Bombarelli, R. 2022. Generative Coarse-Graining of Molecular Conformations. In Chaudhuri, K.; Jegelka, S.; Song, L.; Szepesvári, C.; Niu, G.; and Sabato, S., eds., *International Conference on Machine Learning*, 23213–23236.

Xu, M.; Powers, A. S.; Dror, R. O.; Ermon, S.; and Leskovec, J. 2023. Geometric Latent Diffusion Models for 3D Molecule Generation. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International Conference on Machine Learning*, 38592–38610.

Yang, S.; and Gómez-Bombarelli, R. 2023. Chemically Transferable Generative Backmapping of Coarse-Grained Proteins. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International Conference on Machine Learning*, 39277–39298.

Zalewski, M.; Kmiecik, S.; and Koliński, M. 2021. Molecular Dynamics scoring of protein-peptide models derived from coarse-grained docking. *Molecules*, 26(11): 3293.

Zhang, R.; Huang, Y.; Lou, Y.; Ding, W.; Cao, Y.; and Wang, H. 2024. Synergistic Attention-Guided Cascaded Graph Diffusion Model for Complementarity Determining Region Synthesis. *IEEE Transactions on Neural Networks and Learning Systems*, 1–12.