Diffusion models with group symmetries for biomolecule generation

Wen-ran Li [*1] **Xavier F. Cadet** [*2] **David Medina-Ortiz** [3 4] **Mehdi D. Davari** [5] **Ramanathan Sowdhamini** [6 7 8] **Cédric Damour** [9] **Yu LI** [10] **Alain Miranville** [11] **Frederic Cadet** [1 12]

Abstract

Generative artificial intelligence is a promising approach and has achieved great success in various protein generation tasks. Notably, diffusion models stand out for their robust mathematical foundations and impressive generative capabilities, offering unique advantages in certain applications such as protein design. This review summarizes Special Euclidean group $SE(3)$-invariant and Euclidean group $E(3)$-invariant diffusion models tailored to the structural properties of proteins and small molecules, respectively. We examine why $SE(3)$-equivariant models are predominantly used in protein design, while $E(3)$-equivariant models are favored for molecule generation. Finally, we discuss future directions, including multiscale modeling, dynamic integration, and cross-domain applications, to advance biomolecular design and drug discovery.

[*]Equal contribution [1]University of Paris City & University of Reunion, INSERM, BIGR, DSIMB, F-75015 Paris, France [2]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA (Work done whilst at Department of Computing, Imperial College London) [3]Department of Computer Engineering, University of Magallanes, Punta Arenas, Chile [4]Centre for Biotechnology and Bioengineering, Universidad de Chile, Santiago, Chile [5]Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, Halle 06120, Germany [6]National Centre for Biological Science, TIFR, Bangalore, India [7]Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India [8]Institute of Bioinformatics and Applied Biotechnology, Bangalore, India [9]EnergyLab, EA 4079, Faculty of Sciences and Technology, University of Reunion, France [10]School of Information Science and Technology and Beijing Institute of Artificial Intelligence, China [11]Laboratoire de Mathématiques Appliquées du Havre (LMAH), University Le Havre Normandie, le Havre, France [12]PEACCEL, Artificial Intelligence Department, AI for Biologics, 75013, Paris, France. Correspondence to: Frederic Cadet <frederic.cadet.run@gmail.com>.

Proceedings of the Workshop on Generative AI for Biology at the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

1. Introduction

For decades, protein engineering and protein design tasks have been regarded as NP-hard optimization problems (Mukhopadhyay, 2014). Algorithm challenges continue to persist despite advances in computational methods (Pierce & Winfree, 2002). Researchers have been working to explore effective methods to bridge the sizeable gap. Due to their ability to learn complex patterns for large datasets, deep learning approaches have been applied to various tasks such as protein structure prediction, sequence design for specific functions, and *de novo* protein design (DNPD) (Watson et al., 2023).

Generative modeling is a subfield of machine learning (ML) that focuses on developing algorithms capable of generating new data samples that resemble the data distribution of a given training dataset. Successful applications of generative modeling have highlighted the potential of protein design by modeling the probability distribution of protein sequences. Diffusion models (Kloeden et al., 1992) have given amazing results for image, audio, and text synthesis, while being relatively simple to implement. These models use a parameterized Markov chain trained by variational inference, enabling the generation of samples that align with the data distribution within a finite time, providing a structured and efficient mechanism for generative tasks. Transitions of this chain are learned to reverse a diffusion process, which is a Markov chain that gradually adds noise to the data in the opposite direction of sampling until the signal is destroyed.

Diffusion models address key challenges faced by other generative approaches: They overcome the difficulty of accurately matching posterior distributions in Variational autoencoders (VAEs), mitigate the instability arising from the adversarial training objectives in Generative adversarial networks (GANs), and excel in protein generation tasks, particularly in producing structures with improved atom stability (Chen et al., 2024). Moreover, compared to the above two sets of models, diffusion models are based on the theory of Brownian processes; they are more suitable to represent fluctuations in protein sequence and structure changes (Cadet et al., 2019). This characteristic helps us deal with the complex problem of evolutionary dynamics in protein structures and sequences (Morcos, 2021).

Although there have been several surveys on diffusion models for biomolecule generation, they have not explored the difference between $SE(3)$ and $E(3)$, and also do not answer why these two types of models are usually used in protein design and molecule generation, respectively. The motivation for this work is to provide advanced and comprehensive insights into the development and comparison of diffusion models, explaining the advantages and disadvantages of the diffusion model with equivariant properties $SE(3)$ and $E(3)$, and the future directions and perspectives of the diffusion models to assist in protein design. We compare our survey with existing surveys in Table 1 in Appendix.

The main **contributions** of this review include:

- An accessible introduction to the fundamentals of diffusion models and group invariance.
- A fairly detailed overview of the applications of 13 models for protein structure design and 18 models for molecule generation.
- Discussion on the future development of diffusion models to assist in biomolecule design.

This work explores the generation of different biomolecules through diffusion models, emphasizing protein design.

2. Theoretical preparation

For a better understanding of the models, we provide some background on the protein structure and graph of the molecules, the symmetries of the $SE(3)$ group and the $E(3)$ group, and the theory of the diffusion model.

2.1. From structure to group

There are **3 "difference"** and **1 "why"** to help us better understand the relation of biomolecular structure and geometry graph.

Q: What is the **difference** between protein structure and molecule graph?

A: Protein structure has 4 levels. The simplest level of protein structure, primary structure, is simply the sequence of amino acids in a polypeptide chain. Amino acids, decoded from mRNA, have a common backbone with a heavy atom part $N-C_\alpha-C-O$. The next level of protein structure, secondary structure, refers to local folded structures that form within a polypeptide due to interactions between atoms of the backbone. The overall 3D structure of a polypeptide is called its tertiary structure, see Fig. 1 (a). The last-level quaternary structure consists of more than one amino acid chain. Protein structures are often stored as files with the suffix .pdb. In this work, we mainly discuss the tertiary structure.

A molecule graph is a labeled graph whose vertices correspond to the atoms of the compound and the edges correspond to chemical bonds. A molecule can be represented as $G = (A, B, X, Y)$, where A is the set of atoms, X represents the matrix of atom content, B is the set of bonds, and Y represents the matrix of bond content.

Q: What is the **difference** between $SE(3)$ group and $E(3)$ group?

A: Both $SE(3)$ group and $E(3)$ belong to the Lie group. A Lie group is a topological group that is also a differentiable manifold, and such that the composition and inverse operations $G \times G \to G$ and $G \to G$ are infinitely differentiable functions.

The Special Euclidean group $SE(3)$ is defined as:

$$SE(3) = \left\{ A \;\middle|\; A = \begin{bmatrix} R & \mathbf{r} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix}, R \in SO(3), \mathbf{r} \in \mathbb{R}^3 \right\}$$

where:

- R is a proper rotation matrix in $SO(3)$, satisfying $R^\top R = I$ and $\det(R) = 1$. For example, for a roration angle θ, a rotation matrix around the z-axis would be written:

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \in SO(3).$$

- $\mathbf{r}$ is a translation vector in $\mathbb{R}^3$.
- $\mathbf{0}_{1\times 3}$ denotes a 1×3 row vector of zeros.

SE(3) represents translation and rotation in Fig. 1 (b).

$E(3)$ is the notation for the Euclidean group that denotes the set of isometric transformations in Euclidean space, and the transformations in $E(3)$ include translation, rotation, and reflection, see Fig. 1 (b).

Q: What is the **difference** between invariance and equivariance?

Definition 2.1. (Duval et al., 2023) Denote the action of a group G on a space X by $\rho_g(x)$, for $\rho_g \in G$ and $x \in X$. If ρ_g acts on spaces X and Y, we say:

- A function $f : X \to Y$ is G-**invariant** if $f(\rho_g \cdot x) = f(x)$, i.e. the output remains unchanged under transformations of the input, see Figure 2 (a).
- A function $f : X \to Y$ is G-**equivariant** if $f(\rho_g \cdot x) = \rho'_g \cdot f(x)$, i. e. a transformation of the input must result in the output transforming correspondingly, see Figure 2 (b).

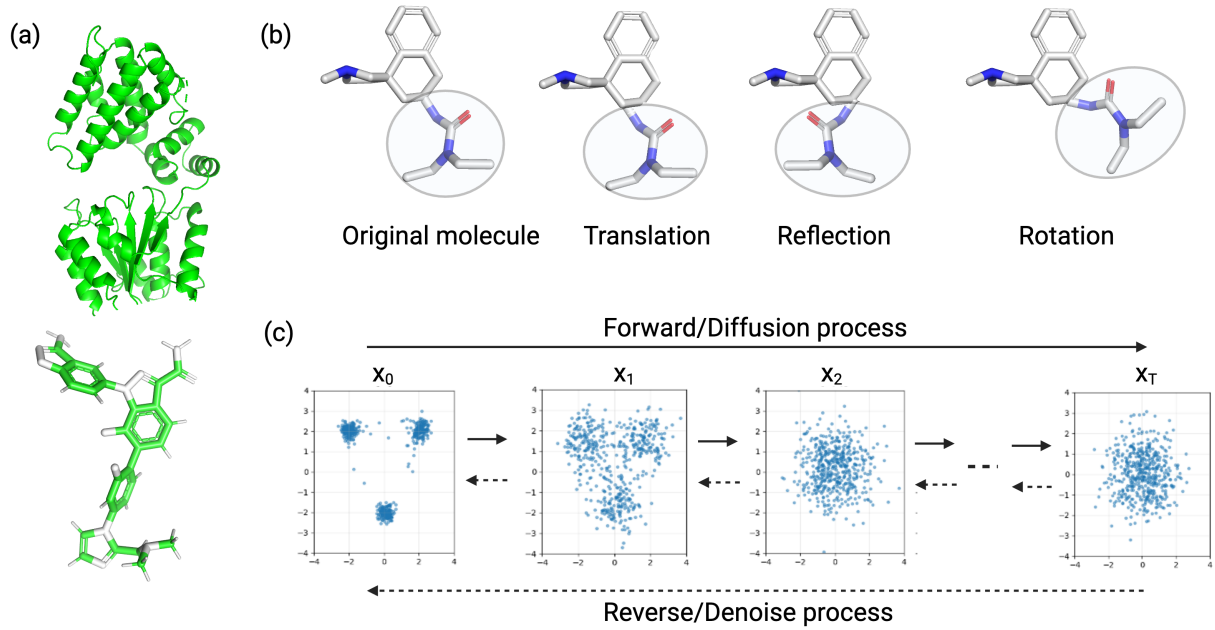

Figure 1. (a) Example of protein structure and molecule graph. (b) A molecule graph as an example to show its shape after translation, rotation and reflection. (c) Visualization of diffusion models operating on image generation. During the diffusion process, the image becomes blurred until it becomes a Gaussian distribution. The reverse process is a denoising process, and the image gradually becomes clear.

- A function $f : X \to Y$ which nether G-invariant nor G-equivariant is referred to as G-unconstrained. The transformation of the input results in an unknown change in the output.

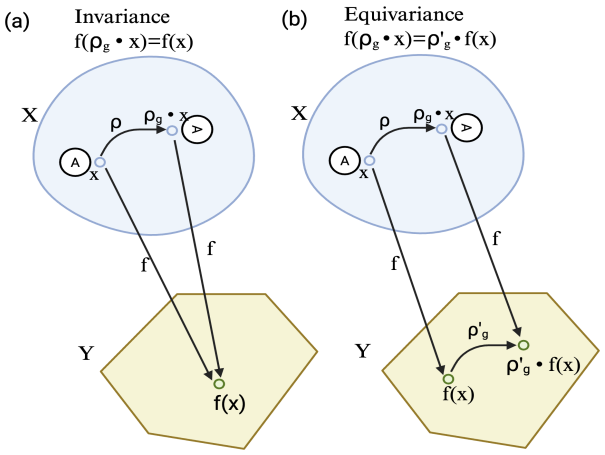

Figure 2. Illustration of invariant and equivariant. the meaning of f, ρ_g, X, Y and ρ'_g can be seen at Definition 2.1.

A function $f : \mathbb{R}^3 \to \mathbb{R}$ (or $\mathbb{R}^n$) is $SE(3)$-invariant if it is unchanged under any rigid transformation in $SE(3)$, which includes *rotations* and *translations*, i.e.,

$$\rho_g = Rx + \mathbf{r}$$

in Definition 2.1, where $R \in SO(3)$ for rotation, $\mathbf{r} \in \mathbb{R}^3$ for translation, $SE(3) = \{(R, \mathbf{r})\}$.

A function $f : \mathbb{R}^3 \to \mathbb{R}$ is $E(3)$-invariant if it is unchanged under any transformation in the Euclidean group $E(3)$, which includes *rotations*, *translations*, and *reflections*, i.e.,

$$\rho_g = Qx + \mathbf{r}$$

in Definition 2.1, where $Q \in O(3)$, $O(3)$ includes orthogonal matrices with $\det(Q) = \pm 1$, which is all the probability for rotation and reflection. $E(3) = \{(Q, \mathbf{r})\}$.

Q: Why $SE(3)$ be used predominantly for proteins while $E(3)$ is more common for molecules?

A: In addition to preserving rotational and translational transformations, $E(3)$ also includes reflection transformations, which align with the intrinsic reflection symmetry of certain small molecules. This makes $E(3)$ more suitable for describing molecular structures. In contrast, protein structures are governed by strict chirality constraints (e.g. L-form amino acids). Reflections would disrupt critical geometric properties, such as the planarity of peptide bonds (e.g., Ramachandran angle constraints) and are therefore physically invalid for biological macromolecules. Consequently, the $SE(3)$ group, which excludes reflections and preserves rigid-body motions (rotations and translations), is better suited for modeling protein structures.

2.2. Diffusion models

A diffusion model (Kingma & Gao, 2023; Nakkiran et al., 2024) is a deep generative model based on two stages: a forward diffusion stage and a reverse diffusion stage. In the forward diffusion stage, the input data are gradually perturbed over several steps by adding Gaussian noise. In the reverse phase, a model restores the original input data by learning to reverse the diffusion process step by step. Figure 1 (c) illustrates how a diffusion model works to generate scattered points that satisfy a specific distribution.

In discrete form, for a sufficiently large time $T > 0$, $t = 0, 1, ..., T$, with the random variable $x_0 \in \mathbb{R}^n$, where n is the dimension, the forward process iteratively adds isotropic Gaussian noise to the sample. The Gaussian transition kernel is set as:

$$q(x_t|x_{t-1}) = \mathcal{N}(\sqrt{1-\beta_t}x_{t-1}, \beta_t\mathbb{I}), \tag{1}$$

$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}), \tag{2}$$

where $\mathbb{I}$ is the identity matrix, β_t are chosen according to a fixed variance scheme. Noisy data x_t can be sampled directly from x_0:

$$x_t = \sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}\epsilon, \tag{3}$$

where $\epsilon \sim \mathcal{N}(0, I)$ and $\alpha_t = \prod_{s=1}^{t}(1-\beta_s)$.

While the reverse process, starting from noise $x_T \sim \mathcal{N}(0, \mathbb{I})$, aims to learn the process of denoising:

$$p_\theta(x_0) = p(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1}|x_t); \tag{4}$$

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma_\theta(x_t, t)), \tag{5}$$

Eq. (5) is defined as a Markov chain with learned Gaussian transitions stering at $p(x_T) = \mathcal{N}(x_T; 0, \mathbb{I})$, where

$$\mu_\theta(x_t, t) = \frac{1}{\sqrt{1-\beta_t}}(x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}}\sigma_\theta(x_t, t)),$$

the Denoising Diffusion Probabilistic Models (DDPM) aims to approximate ϵ using a parametric model structured as σ_θ. The objective function can be written as follows:

$$\theta^* = \arg\min_\theta \mathbb{E}_{x_0,t,\epsilon}[\|\epsilon - \sigma_\theta(\sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}\epsilon, t)\|^2].$$

In addition to DDPM, diffusion models have some other forms, such as score-based generative model (SGM) (Jo et al., 2022) and noise-conditional score networks (NCSN) (Guo et al., 2023), which we will not discuss further due to space limitations.

3. Diffusion model for protein generation

A protein is a sequence of amino acids (residues) linked into a chain that folds into a complex 3D structure under the influence of electrostatic forces. The protein backbone can be seen as N rigid bodies that contain four heavy atoms $N - C_\alpha - C - O$. This section discusses the generation of protein backbone.

A protein backbone is a continuous chain of atoms that runs throughout the length of a protein. Generating a backbone is a difficult task because a backbone should fulfill the following three criteria:

- **Physically realizable:** The sequence can be found to fold into the generated structure.
- **Functional:** We aim for conditional sampling under diverse functional constraints without retraining.
- **Generalizability:** The model has multiple application scenarios.

For the above criteria, we introduce several models that, in our assessment, meet the highest standards for protein backbone generation and discuss the effects of these models.

3.1. Physically realizable model: Diffusion on $SE(3)$ group

$SE(3)$ is the notation for the special Euclidean 3D group that includes translational and rotational isometric transformations and keeps the volume constant. Proteins are not static objects; They naturally exist in an equilibrium of conformations. The mathematical framework is particularly relevant for modeling molecular systems, where maintaining spatial invariance is crucial for accurate predictions.

Building on this principle, **RFDiffusion** (Watson et al., 2023) repurposes RoseTTAFold to perform reverse diffusion. RFDiffusion uses RoseTTAFold's $SE(3)$-equivariant architecture to preserve isometric transformations during structure generation, ensuring geometric consistency; see Fig. 3. By fine-tuning RoseTTAFold All-Atom (RFAA) (Krishna et al., 2024), a neural network for predicting biomolecular structures, to diffusion denoising tasks, **RFDiffusionAA** generates folded protein structures surrounding the small molecule from random residue distributions. **ProteinGenerator** (Lisanza et al., 2023) is a RoseTTAFold-based sequence space diffusion model that simultaneously generates protein sequences and structures. ProteinGenerator exhibits a lower success rate than RFDiffusion in producing long, structurally accurate sequences, likely due to inherent challenges in sequence-space diffusion compared to structure-space approaches.

A group that is a differentiable manifold is called a Lie

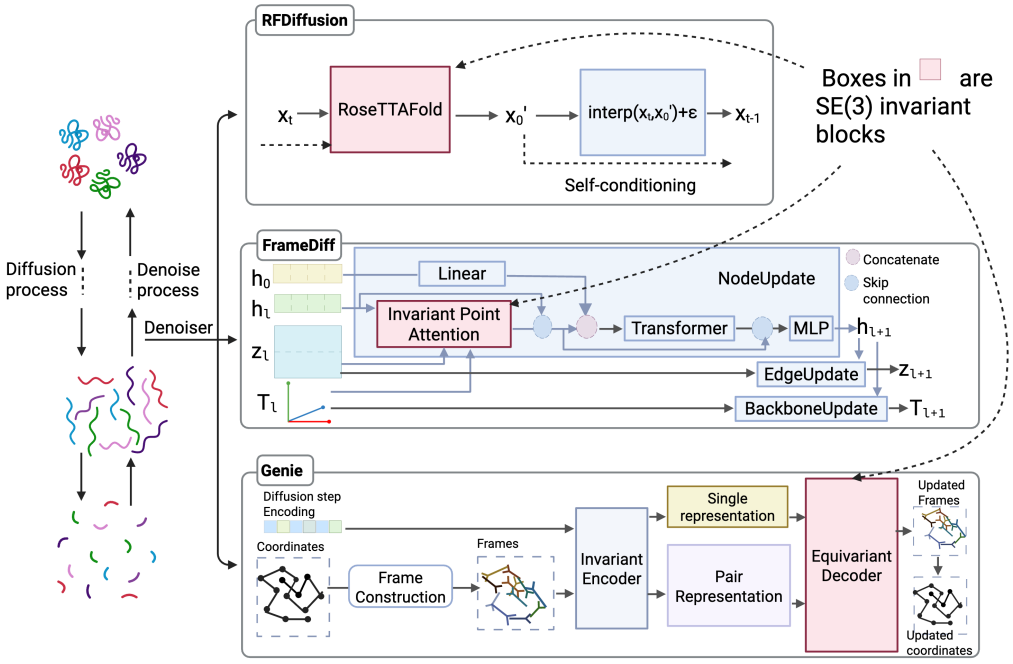

Figure 3. $SE(3)$ equivariant diffusion models for protein structure generation. RFDiffusion, FrameDiff and Genie utilize RoseTTAFold, IPA and $SE(3)$-equivariant denoiser as a single step of the denoise process in the diffusion model, respectively. Boxes in pink color are $SE(3)$ equivariant blocks. $SE(3)$ equivariant keeps the frames of each amino acid physically stable.

group. **FrameDiff** (Yim et al., 2023) is a diffusion model in the Lie group $SE(3)_0^N$ for the generation of protein backbones. The superscript N indicates that the product space consists of N independent $SE(3)$ elements; the subscript 0 indicates the constraint on the translation component: the translation components of all $SE(3)$ transformations must satisfy the centralization condition, that is, the average value of the translation components is fixed at the origin, $\sum_{i=1}^{N} x_i = 0$. This model applies Invariant Point Attention (IPA) (Jumper et al., 2021) to keep the updates of residues in coordinate space that are $SE(3)$-invariant, see Fig. 3.

FrameDiff is also used for inpainting protein structures and motif scaffolding, named **FrameDiPT** (Zhang et al., 2023a) and **TDS** (Wu et al., 2024), respectively. **VFN-Diff** (Mao et al., 2023) replaces the IPA in FrameDiff with Vector Field Networks (VFN), which is also a $SE(3)$ equivariant model. Following the settings and benchmarks of FrameDiff, VFN-Diff significantly outperforms FrameDiff in terms of designability (67.04% vs. 53.58%) and diversity (66.54% vs. 51.98%).

Genie (Lin & AlQuraishi, 2023) integrates the SE(3)-equivariant reasoning framework of IPA with DDPMs, developing an SE(3)-equivariant denoiser $\epsilon_\theta(F(x_t), t)$ for protein generation, see Fig. 3. **Genie2** (Lin et al., 2024b) extended Genie to motif scaffolding, and introduced a novel multi-motifs framework that designs co-occurring motifs without needing to specify inter-motif positions and orientations in advance.

The $N - C_\alpha - C$ frame of each amino acid means that for, N: amino group; C_α: side group, also called R group; C: carboxyl group, their coordinates jointly constitute the geometric configuration of a residue in a protein backbone. Figure 3 illustrates how RFdiffusion, FrameDiff, and Genie incorporate $SE(3)$-equivariant neural networks into their denoiser architectures. This kind of architecture will keep the $N - C_\alpha - C$ frame of each amino acid invariant to global rotations and translations.

As special subsets of $SE(3)$ equivariant models, some protein generation models such as **ProtDiff-SMCDiff** (Trippe et al., 2023) satisfy $E(3)$ equivariance. This kind of model additionally keeps consistency for permutation and translation.

3.2. Model with strong functionality

Protein design projects often involve complex and composite requirements that vary over time. **Chroma** (Ingraham et al., 2023) explores a programmable generative process with custom energy functions, which aims to make the generated protein have desired properties and functions, such

as symmetry, substructure, shape and semantics.

3.3. Model with generalizability

AlphaFold2 (AF2) (Jumper et al., 2021) comprises two main strategies: Evoformer provides mechanisms for the exchange of information within the Multiple Sequence Alignment (MSA) and pair representations that enable direct reasoning about spatial and evolutionary relationships; IPA is used to update a set of residue neural activations without changing the 3D positions. AlphaFold2 was much more accurate than the competing methods in CASP14.

AlphaFold3 replaces the IPA of AlphaFold2 with Diffusion model, which reconstructs coordinates from the residue level to the atomic level. AlphaFold3 (AF3) (Abramson et al., 2024) exhibits strong generalizability and versatility, expanding beyond protein generation to handle diverse molecular tasks, including the prediction of ligand and RNA structures.

AlphaFold3 takes a larger step in the generation of biomolecules. It has many more application scenarios: ligand docking, protein-nucleic acid complexes, covalent modifications, and protein complexes. With AF3, it is possible to handle a more diverse biomolecular space. In CASP16 (Elofsson, 2025), AF3 performs comparably to top predictors for proteins and complexes, with average GDT-TS and DockQ scores indicating high model quality.

4. Small molecule generation

Similarly to proteins, small molecules can be represented in both 1D linear formats and 3D graph-based formats. While 1D representations (e.g., SMILES and InChI) are primarily used for rapid retrieval and standardized data exchange, 3D graph representations are widely adopted in AI-driven molecular modeling. Consequently, all molecular generation models discussed in this work employ graph-based representations.

The topic of generating molecules using diffusion models is equivalent to the following question: *How to generate attributed graphs using diffusion models?* To answer this question, there are two main challenges:

- **Complex dependency:** Dependency between nodes and edges.
- **Non-unique representations:** Order of the nodes is not fixed.

For the first challenge, diffusion models need to define the atomic positions $x_i \in \mathbb{R}^3$ and the atomic types $a_i = \{C, N, O, ...\}$ and specify independent forward processes for each data type,

$$p_t(x_t|x_0) = \mathcal{N}(x_t|\alpha_t x_t, \sigma_t \mathbb{I}), \tag{6}$$
$$p_t(a_t|a_0) = \mathcal{N}(a_t|\alpha_t a_t, \sigma_t \mathbb{I}), \tag{7}$$

If $G_t = (x_t, a_t)$, then $p_t(G_t|G_0) = \mathcal{N}(x_t|\alpha_t G_t, \sigma_t \mathbf{I})$, and the continuous forward process is represented as

$$dG_t = f_t(G_t)dt + g_t(G_t)d\omega_t,$$

the reverse-time diffusion process is represented as:

$$\begin{cases} dx_t = [f_{1,t}(x_t) - g_{1,t}^2 \nabla_{x_t} \log p_t(G_t)]dt + g_{1,t} d\bar{\omega}_1, \\ da_t = [f_{2,t}(x_t) - g_{2,t}^2 \nabla_{a_t} \log p_t(G_t)]dt + g_{2,t} d\bar{\omega}_2. \end{cases} \tag{8}$$

We use $s_\theta^x(G_t)$, $s_\theta^a(G_t)$ to approximate $\nabla_{x_t} \log p_t(G_t)$, $\nabla_{a_t} \log p_t(G_t)$ respectively, and train the neural network to jointly approximate the score functions of the constituent processes:

$$\mathcal{L} = \mathbb{E}_{x_t, a_t}[\|s_\theta^a(G_t) - \nabla_{a_t} \log p_t(G_t))\| + \|s_\theta^x(G_t) - \nabla_{x_t} \log p_t(G_t)\|].$$

For the second challenge, the nodes representing the atoms may contain information about the atom type as well as its 3D spatial coordinates. It is desirable to process the latter part of the features in a manner that would transform in the same way as the molecule is transformed in space, in other words, be equivariant to the Euclidean group $E(3)$ of rigid motions. In the following two subsections, we discuss how diffusion models capture the system of positional equivariance, such as $SE(3)$ equivariance and $E(3)$ equivariance.

4.1. Diffusion model on $SE(3)$ group for molecule

In isolation, a molecule and its mirror image share the same internal features and properties, regardless of its chirality. Since ML datasets often showcase molecules in isolation, E(3)-equivariance is desirable. However, molecular functionality is most often conferred by intermolecular interactions with surrounding components, meaning that a molecule's properties may differ from those of its mirror image. In such cases, we no longer require equivariance to reflections, making SE(3)-equivariance desirable (Duval et al., 2023).

Although in section 2, we have discussed that $SE(3)$ equivariant models are more appropriate to be used in protein design, there are still several $SE(3)$ equivariant models used for molecule generation without considering their intrinsic reflection symmetry. Here, we discuss their applications.

GeoDiff (Xu et al., 2022) integrates the diffusion model with graph field networks (GFN), an equivariant convolutional layer, to generate stable conformations, the difference

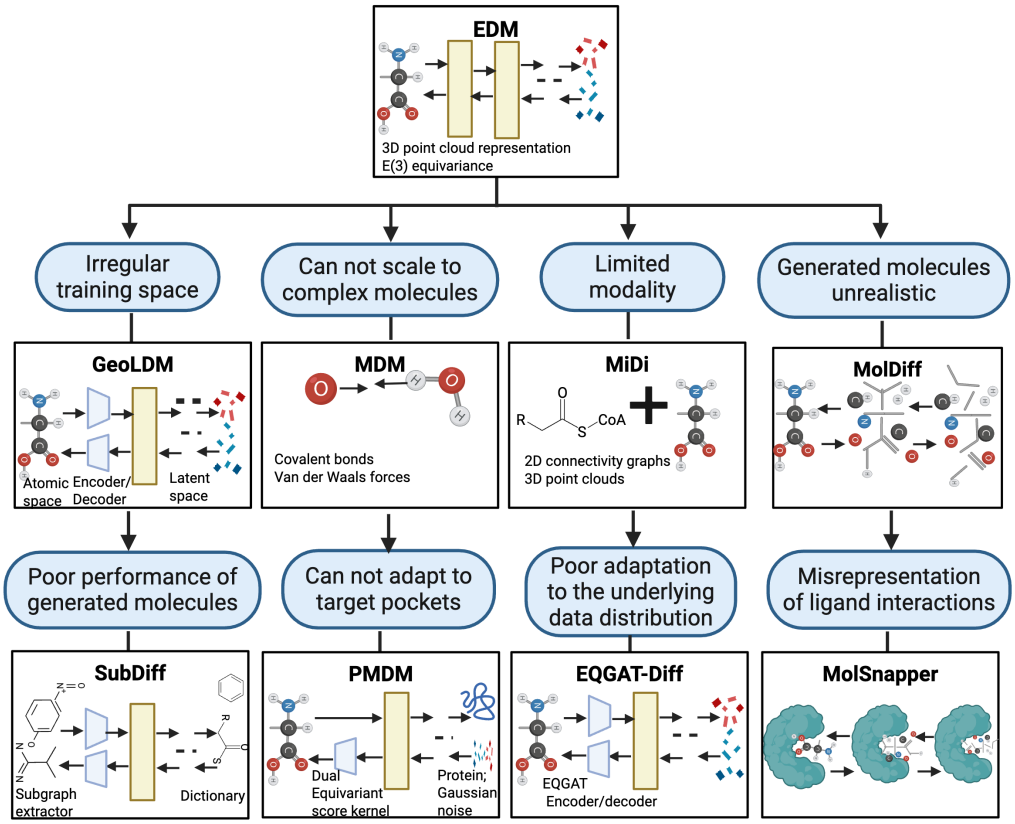

Figure 4. Overview of EDM and its extensions for molecular generation. The top box represents the foundational EDM model. The blue box indicates that the limitations present in the previous model are addressed by the subsequent model.

being that the GNN is $SE(3)$-invariant. According to (Garcia Satorras et al., 2021), it is impossible to have a non-zero distribution that is invariant to translation, since it cannot integrate to one. For dealing with this problem, GeoDiff leverages distributions on the linear subspace to diffusion models. In the subspace, $\sum_i x_i = 0$, i.e. the center of gravity is always zero, the normal distribution $\mathcal{N}_x$ over this subspace and its likelihood can be expressed as:

$$\mathcal{N}_x(x|\mu, \sigma^2\mathbb{I}) = (\sqrt{2\pi\sigma})^{-(M-1)\cdot n} \exp(-\frac{1}{2\sigma^2}\|x-\mu\|^2)$$

where μ lies in the same subspace as x.

However, GeoDiff treats atoms as individual particles, overlooking the substructure of molecules, which contains information on properties. Considering the above problem, **SubGDiff** (Zhang et al., 2024) introduces a discrete binary distribution to the diffusion process, where a mask vector sampling from the distribution can be used to select a subset of the atoms (i.e. subgraph) to determine which substructure the noise should be added to at the current time step. SubGDiff also uses GFN as the denoising network for conformation generation. By using the same data, with 500 steps, SubGDiff achieves much better performance than GeoDiff with 5000 steps on 5 out of 8 metrics, which implies that it can accelerate the sampling efficiency.

After applying GeoDiff's zero center of gravity trick, the Geometry-Complete Diffusion Model (GCDM) (Morehead & Cheng, 2024) parametrizes the transition function using an $SE(3)$-equivariant neural network to assign the same likelihood to a generated molecule regardless of arbitrary rotations or translations in 3D simulation. Following the noise process in GCDM, **DiffSBDD** (Schneuing et al., 2023) formulates a structure-based drug design (SBDD). The nodes have both geometric atomic coordinates x as well as nuclear type features h. DiffSBDD uses a simple implementation of EGNN to update features h and coordinates x.

Both **TargetDiff** (Guan et al., 2023) and **DiffBP** (Lin et al., 2024a) propose a target-aware molecular diffusion process with a $SE(3)$-equivariant GNN denoiser. The training and sampling procedures in TargetDiff are aligned in non-autoregressive diffusion models and $SE(3)$ equivariant fashion. DiffBP generates molecules with high protein affinity, appropriate sizes, and favorable drug-like profiles.

4.2. Models based on EGNNs

$E(3)$ group is the group in $\mathbb{R}^3$ with rotations, reflections, and translations. The graph neural networks equivariant to rotations, translations, reflections and permutations are called $E(n)$- Equivariant Graph Neural Networks (EGNNs). Because the molecule is mainly on 3D space, in this work we mainly discuss the case $n = 0$, i.e., $E(3)$- Equivariant Graph Neural Networks.

$E(3)$ Equivariant diffusion model (**EDM**) (Hoogeboom et al., 2022) learns a diffusion model that is equivariant to translation and rotation. It jointly operates on continuous (atom coordinates) and categorical features (atom types) in the denoising phase. **DiffLinker** (Igashov et al., 2024) leverages EDM and develops diffusion models for the design of molecular linkers.

Context-guided diffusion (**CGD**) (Klarner et al., 2024) can consistently generate novel, near-out-of-distribution (near-OOD) molecules with desirable properties. CGD also applies to EDM for material design following the setup of **GaUDI** (guided diffusion model for inverse molecular design) (Weiss et al., 2023), which can discover molecules better than existing ones. Selective iterative latent variable refinement (**SILVR**) (Runcie & Mey, 2023) combines Iterative Latent Variable Refinement (ILVR) and EDM to perform fragment merging and linker generation.

By building point-structured latent codes with invariant scalars and equivariant tensors, **GeoLDM** (Xu et al., 2023) can effectively learn latent representations while preserving roto-translational equivariance. It also circumvents the limitations of EDM on irregular training surfaces. Subgraph latent diffusion model (**SubDiff**) (Yang et al., 2024) performs subgraph-level encoding in the diffusion process and is used for 3D molecular generation tasks. For unconditional generation tasks, SubDiff is generally better than EDM and GeoLDM.

EDM represents molecular geometries as point clouds, which makes it difficult to capture the abundance of local constraint relations between adjacent atoms with no explicit indications for chemical bonds. Molecular Diffusion Model (**MDM**) (Huang et al., 2022) tackles this drawback by treating pairs of atoms with atomic spacing below the specified threshold covalently bonded. It also points out the lack of consideration for interatomic relations in GCDM, and addresses the scalability issue by introducing the Dist-transition Block. Pocket based Molecular Diffusion Model (**PMDM**) (Huang et al., 2024) introduces equivariant kernels to MDM to simulate the local chemical boned graph and the global distant graph.

MiDi (Vignac et al., 2023) utilizes the adaptive noise schedule and relaxedEGNN (rEGNN) to generate 3D molecules. MiDi outperformed EDM in 2D metrics while obtaining similar 3D metrics for the generated conformers. **EQGAT-diff** (Le et al., 2024) takes Equivariant Graph Attention Networks (EQGAT) as the component of the diffusion model to carry out the *de novo* 3D molecule design. EQGAT-diff employs rotation equivariant vector features that can be interpreted as learnable vector bundles, for which the denoising networks of EDM and MiDi are lacking.

Taking advantage of the strong relationship between the types and lengths of the bonds to guide the generation of atom positions, **MolDiff** (Peng et al., 2023) produces high-quality 3D molecular graphs and effectively addresses the problem of atom bond inconsistency with the E(3)-equivariant diffusion model. Because MolDiff models and diffuses the bonds of molecules, it exceeds SILVR and EDM in the generation of molecules with better validity. (Ziv et al., 2024b) extends MolDiff to structure-based drug design and creates a model called **MolSnapper**, which can sample molecules for given pockets. Compared with MolDiff, MolSnapper generates molecules better tailored to fit the given binding site, achieving a high structural and chemical similarity to the original molecules.

A full overview of the developments based on EDM can be seen in Figure 3. The examples above show that the combination of EGNN and diffusion model has been widely used in the generation of proteins and small molecules. EGNN is also used alone for the identification of protein binding sites (Sestak et al., 2024). But EGNN is not always optimal if EGNN and Geometric Vector Perceptron (GVP) are both integrated with **Keypoint Diffusion** (Dunn & Koes, 2023), a diffusion model for *de novo* ligand design: the GVP keypoint model can approach all-atom levels of performance, while the EGNN keypoint model may exhibit poor performance on structure representations where a node cannot be adequately described by a single point-mass i.e., residue or fragment point clouds.

5. Discussion

Here, we highlight several landmark models:

- The IPA in AlphaFold2 satisfies the property of $SE(3)$ equivariant, but was replaced by the diffusion transformer in **AlphaFold3**. Therefore, Alphafold3 does not satisfy the properties of an equivariant.
- The reverse diffusion in **RFDiffusion** is composed of RoseTTAFold. This model inherits the good properties of RoseTTAFold, making the generated model physically realizable.
- **FrameDiff** is the first model to introduce $SE(3)$ manifolds into protein structure generation problems. The properties of the $SE(3)$ group provide a mathematical basis for the expression of structural information.

- As a better type of $SE(3)$ equivariant, $E(3)$ equivariant is widely used in the generation of small molecules. The most successful example so far is **EDM**.

Due to the large size and complexity of protein structures, most current protein models can only satisfy SE(3) equivariance but do not have as good properties as E(3) equivariance. How to establish a diffusion model in the E(3) group to complete protein generation is a topic we can study in the future.

While progress in the field has demonstrated that diffusion models can accelerate early-stage drug discovery, challenges remain in adapting such workflows to real-world discovery campaigns:

- Addressing synthesizability is an ongoing challenge because many proposed ideas may not have known synthetic routes, and a chemist can only triage a function of proposed ideas.
- Despite various widely adopted evaluation metrics, measuring and comparing the performance of diffusion models remains a major challenge given the lack of ground-truth and universal metrics.
- Complex dynamics. Cohesive models tend to be static and ignore the fact that proteins and ligands are amphipathic, which is a factor that should be considered when analyzing protein functions.
- Protein structure prediction models typically predict static structures as seen in PDB, not the dynamical behavior of biomolecular systems in solution.

What are potential directions the community could consider exploring further?

- RFDiffusion and ProteinGenerator, which adapt the diffusion model with the traditional model, RoseTTAFold, have done a variety of tasks, such as peptide binder generation, motif-scaffolding, and sequence-structure codesign. We can explore more applications of these two models.
- Alongside $SE(3)$-invariant models for protein structure and $E(3)$-invariant models for small molecule, there are some $O(3)$-invariant models for crystal structure prediction (Jiao et al., 2023) and $SO(3)$-invariant models for antibody generation (Zhu et al., 2024; Ucar et al., 2024) worth paying attention.
- Traditional models are more analytical and closely match the physical properties of proteins. We can use them for more fruitful tasks, such as protein-nucleic acid and protein-ligand interactions.
- Can $E(n)$-Equivariant Topological Neural Networks (ETNN) deformations of EGNN (Battiloro et al., 2024) and NequIP (Batzner et al., 2022) be applied to the generation of molecules? Can EGNN be used to study peptide structures?

6. Conclusion

This review comprehensively summarizes the application of the diffusion model for bioengineering. It captures the progression of AI model architectures, highlighting the emergence of EGNN and diffusion models as game changers in recent works. Diffusion Models are particularly promising generative frameworks.

Acknowledgements

W. Li is supported by a PhD grant from the Region Reunion and European Union (FEDER-FSE 2021/2027) file number 2023062, tiers 345879. X. F. Cadet is supported by the UKRI CDT in AI for Healthcare http://ai4health.io, (Grant No. P/S023283/1), UK. PEACCEL was supported through a research program partially co-funded by the European Union (UE) and Region Reunion (FEDER). DM-O acknowledges ANID for the project "Subvención a instalación en la academia convocatoria año 2022", Folio 85220004. DM-O gratefully acknowledges support from the Centre for Biotechnology and Bioengineering - CeBiB (PIA project FB0001 and AFB240001, ANID, Chile). DM-O acknowledges ANID for FONDECYT INICIACIÓN project 11250295. MDD acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - within the Priority Program Molecular Machine Learning SPP2363 (Project Number 497207454). RS acknowledges funding and support provided by JC Bose Fellowship (SB/S2/JC-071/2015) from Science and Engineering Research Board, India and Bioinformatics Centre Grant funded by Department of Biotechnology, India (BT/PR40187/BTIS/137/9/2021).

Impact Statement

This review of SE(3)/E(3) diffusion models for proteins and small molecules can accelerate drug discovery and enzyme engineering. Benefits include faster responses to emerging diseases and reduced wet-lab waste. Risks: dual-use potential for creating harmful bioagents. Dataset bias may sideline under-studied pathogens, worsening global inequities. Large-scale pre-training carries significant carbon costs. Mitigations: staged model release, automated biosecurity screening, and broader training data. Adopt green AI practices to reduce energy use. Engage ethicists and regulators to guide safe deployment. Responsible governance will maximize health gains while minimizing misuse.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A. I., Cowie, A., Figurnov, M., Fuchs, F. B., Gladman, H., Jain, R., Khan, Y. A., Low, C. M. R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E. D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., and Jumper, J. M. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, June 2024. ISSN 0028-0836, 1476-4687.

Battiloro, C., Karaismailoğlu, E., Tec, M., Dasoulas, G., Audirac, M., and Dominici, F. E(n) equivariant topological neural networks, July 2024.

Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., and Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, May 2022. ISSN 2041-1723.

Bento, A. P., Hersey, A., Félix, E., Landrum, G., Gaulton, A., Atkinson, F., Bellis, L. J., De Veij, M., and Leach, A. R. An open source chemical structure curation pipeline using rdkit. *Journal of Cheminformatics*, 12:1–16, 2020.

Cadet, X. F., Dehak, R., Chin, S. P., and Bessafi, M. Nonlinear dynamics analysis of protein sequences. application to cyp450. *Entropy*, 21(9):852, 2019.

Chen, M., Mei, S., Fan, J., and Wang, M. An overview of diffusion models: Applications, guided generation, statistical rates and optimization, April 2024.

Dunn, I. and Koes, D. R. Accelerating inference in molecular diffusion models with latent representations of protein structure. *ArXiv*, 2023.

Duval, A., Mathis, S. V., Joshi, C. K., Schmidt, V., Miret, S., Malliaros, F. D., Cohen, T., Liò, P., Bengio, Y., and Bronstein, M. A hitchhiker's guide to geometric gnns for 3d atomic systems. *arXiv preprint arXiv:2312.07511*, 2023.

Elofsson, A. Alphafold3 at casp16. *bioRxiv*, pp. 2025–04, 2025.

Gao, Z., Tan, C., Zhang, Y., Chen, X., Wu, L., and Li, S. Z. Proteininvbench: Benchmarking protein inverse folding on diverse tasks, models, and metrics. *Advances in Neural Information Processing Systems*, 36, 2024.

Garcia Satorras, V., Hoogeboom, E., Fuchs, F., Posner, I., and Welling, M. E (n) equivariant normalizing flows. *Advances in Neural Information Processing Systems*, 34: 4181–4192, 2021.

Guan, J., Qian, W. W., Peng, X., Su, Y., Peng, J., and Ma, J. 3d equivariant diffusion for target-aware molecule generation and affinity prediction, March 2023.

Guo, Z., Liu, J., Wang, Y., Chen, M., Wang, D., Xu, D., and Cheng, J. Diffusion models in bioinformatics and computational biology. *Nature Reviews Bioengineering*, 2(2):136–154, October 2023. ISSN 2731-6092.

Haas, J., Barbato, A., Behringer, D., Studer, G., Roth, S., Bertoni, M., Mostaguir, K., Gumienny, R., and Schwede, T. Continuous automated model evaluation (cameo) complementing the critical assessment of structure prediction in casp12. *Proteins: Structure, Function, and Bioinformatics*, 86(S1):387–398, 2018. doi: https://doi.org/10.1002/prot.25431. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.25431.

Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887. PMLR, 2022.

Huang, H., Sun, L., Du, B., and Lv, W. Conditional diffusion based on discrete graph structures for molecular graph generation, May 2023a.

Huang, H., Sun, L., Du, B., and Lv, W. Learning joint 2d & 3d diffusion models for complete molecule generation, June 2023b.

Huang, L., Zhang, H., Xu, T., and Wong, K.-C. Mdm: Molecular diffusion model for 3d molecule generation, September 2022.

Huang, L., Xu, T., Yu, Y., Zhao, P., Chen, X., Han, J., Xie, Z., Li, H., Zhong, W., Wong, K.-C., and Zhang, H. A dual diffusion model enables 3d molecule generation and lead optimization based on target pockets. *Nature Communications*, 15(1):2657, March 2024. ISSN 2041-1723.

Igashov, I., Stärk, H., Vignac, C., Schneuing, A., Satorras, V. G., Frossard, P., Welling, M., Bronstein, M., and Correia, B. Equivariant 3d-conditional diffusion model for molecular linker design. *Nature Machine Intelligence*, 6 (4):417–427, April 2024. ISSN 2522-5839.

Ingraham, J. B., Baranov, M., Costello, Z., Barber, K. W., Wang, W., Ismail, A., Frappier, V., Lord, D. M., Ng-Thow-Hing, C., Van Vlack, E. R., Tie, S., Xue, V., Cowles, S. C., Leung, A., Rodrigues, J. V., Morales-Perez, C. L., Ayoub, A. M., Green, R., Puentes, K., Oplinger,

F., Panwar, N. V., Obermeyer, F., Root, A. R., Beam, A. L., Poelwijk, F. J., and Grigoryan, G. Illuminating protein space with a programmable generative model. *Nature*, 623(7989):1070–1078, November 2023. ISSN 0028-0836, 1476-4687.

Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y., and Liu, Y. Crystal structure prediction by joint equivariant diffusion. *Advances in Neural Information Processing Systems*, 36:17464–17497, 2023.

Jo, J., Lee, S., and Hwang, S. J. Score-based generative modeling of graphs via the system of stochastic differential equations, June 2022.

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. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, August 2021. ISSN 0028-0836, 1476-4687.

Kingma, D. P. and Gao, R. Understanding diffusion objectives as the elbo with simple data augmentation, September 2023.

Klarner, L., Rudner, T. G. J., Morris, G. M., Deane, C. M., and Teh, Y. W. Context-guided diffusion for out-of-distribution molecular and protein design, July 2024.

Kloeden, P. E., Platen, E., Kloeden, P. E., and Platen, E. *Stochastic differential equations*. Springer, 1992.

Kovtun, D., Akdel, M., Goncearenco, A., Zhou, G., Holt, G., Baugher, D., Lin, D., Adeshina, Y., Castiglione, T., Wang, X., Marquet, C., McPartlon, M., Geffner, T., Rossi, E., Corso, G., Stärk, H., Carpenter, Z., Kucukbenli, E., Bronstein, M., and Naef, L. Pinder: The protein interaction dataset and evaluation resource, July 2024.

Krishna, R., Wang, J., Ahern, W., Sturmfels, P., Venkatesh, P., Kalvet, I., Lee, G. R., Morey-Burrows, F. S., Anishchenko, I., Humphreys, I. R., et al. Generalized biomolecular modeling and design with rosettafold all-atom. *Science*, 384(6693):eadl2528, 2024.

Le, T., Cremer, J., Noé, F., Clevert, D.-A., and Schütt, K. Navigating the design space of equivariant diffusion-based generative models for de novo 3d molecule generation. *arXiv preprint arXiv:2309.17296*, 2023.

Le, T., Cremer, J., Noe, F., and Clevert, D.-A. Navigating the design space of equivariant diffusion-based generative models for de novo 3d molecule generation. 2024.

Lin, H., Huang, Y., Zhang, O., Ma, S., Liu, M., Li, X., Wu, L., Wang, J., Hou, T., and Li, S. Z. Diffbp: Generative diffusion of 3d molecules for target protein binding, July 2024a.

Lin, Y. and AlQuraishi, M. Generating novel, designable, and diverse protein structures by equivariantly diffusing oriented residue clouds, June 2023.

Lin, Y., Lee, M., Zhang, Z., and AlQuraishi, M. Out of many, one: Designing and scaffolding proteins at the scale of the structural universe with genie 2, May 2024b.

Lisanza, S. L., Gershon, J. M., Tipps, S., Arnoldt, L., Hendel, S., Sims, J. N., Li, X., and Baker, D. Joint generation of protein sequence and structure with rosettafold sequence space diffusion. *bioRxiv*, pp. 2023–05, 2023.

Mao, W., Zhu, M., Sun, Z., Shen, S., Wu, L. Y., Chen, H., and Shen, C. De novo protein design using geometric vector field networks, October 2023.

Montalvão, R. W., Pitt, W. R., Pinheiro, V. B., and Blundell, T. L. Melodia: A python library for protein structure analysis. *Bioinformatics*, 40(7):btae468, July 2024. ISSN 1367-4811.

Morcos, F. Information theory in molecular evolution: From models to structures and dynamics, 2021.

Morehead, A. and Cheng, J. Geometry-complete diffusion for 3d molecule generation and optimization. *Communications Chemistry*, 7(1):150, July 2024. ISSN 2399-3669.

Mukhopadhyay, M. A brief survey on bio inspired optimization algorithms for molecular docking, July 2014.

Nakkiran, P., Bradley, A., Zhou, H., and Advani, M. Step-by-step diffusion: An elementary tutorial, June 2024.

Norton, T. and Bhattacharya, D. Sifting through the noise: A survey of diffusion probabilistic models and their applications to biomolecules. *arXiv preprint arXiv:2406.01622*, 2024.

Peng, X., Guan, J., Liu, Q., and Ma, J. Moldiff: Addressing the atom-bond inconsistency problem in 3d molecule diffusion generation, May 2023.

Pierce, N. A. and Winfree, E. Protein design is np-hard. *Protein Engineering, Design and Selection*, 15(10):779–782, October 2002. ISSN 1741-0134, 1741-0126.

Runcie, N. T. and Mey, A. S. Silvr: Guided diffusion for molecule generation. *Journal of Chemical Information and Modeling*, 63(19):5996–6005, October 2023. ISSN 1549-9596, 1549-960X.

Schneuing, A., Du, Y., Harris, C., Jamasb, A., Igashov, I., Du, W., Blundell, T., Lió, P., Gomes, C., Welling, M., Bronstein, M., and Correia, B. Structure-based drug design with equivariant diffusion models, June 2023.

Sestak, F., Schneckenreiter, L., Brandstetter, J., Hochreiter, S., Mayr, A., and Klambauer, G. Vn-egnn: E (3)-equivariant graph neural networks with virtual nodes enhance protein binding site identification. *arXiv preprint arXiv:2404.07194*, 2024.

Trippe, B. L., Yim, J., Tischer, D., Baker, D., Broderick, T., Barzilay, R., and Jaakkola, T. Diffusion probabilistic modeling of protein backbones in 3d for the motif-scaffolding problem, March 2023.

Ucar, T., Malherbe, C., and Gonzalez, F. Exploring log-likelihood scores for ranking antibody sequence designs. In *NeurIPS 2024 Workshop on AI for New Drug Modalities*, 2024.

Vignac, C., Osman, N., Toni, L., and Frossard, P. Midi: Mixed graph and 3d denoising diffusion for molecule generation, June 2023.

Wang, C., Zhong, B., Zhang, Z., Chaudhary, N., Misra, S., and Tang, J. Pdb-struct: A comprehensive benchmark for structure-based protein design.

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., Lauko, A., De Bortoli, V., Mathieu, E., Ovchinnikov, S., Barzilay, R., Jaakkola, T. S., DiMaio, F., Baek, M., and Baker, D. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, August 2023. ISSN 0028-0836, 1476-4687.

Weiss, T., Mayo Yanes, E., Chakraborty, S., Cosmo, L., Bronstein, A. M., and Gershoni-Poranne, R. Guided diffusion for inverse molecular design. *Nature Computational Science*, 3(10):873–882, 2023.

Wu, L., Trippe, B., Naesseth, C., Blei, D., and Cunningham, J. P. Practical and asymptotically exact conditional sampling in diffusion models. *Advances in Neural Information Processing Systems*, 36, 2024.

Xu, M., Yu, L., Song, Y., Shi, C., Ermon, S., and Tang, J. Geodiff: A geometric diffusion model for molecular conformation generation, March 2022.

Xu, M., Powers, A., Dror, R., Ermon, S., and Leskovec, J. Geometric latent diffusion models for 3d molecule generation, May 2023.

Yang, K., Zhou, Z., Li, L., Wang, P., Wang, X., and Wang, Y. Subdiff: Subgraph latent diffusion model. *OpenReview*, 2024.

Yim, J., Trippe, B. L., De Bortoli, V., Mathieu, E., Doucet, A., Barzilay, R., and Jaakkola, T. Se(3) diffusion model with application to protein backbone generation, May 2023.

Zhang, C., Leach, A., Makkink, T., Arbesú, M., Kadri, I., Luo, D., Mizrahi, L., Krichen, S., Lang, M., Tovchigrechko, A., Lopez Carranza, N., Şahin, U., Beguir, K., Rooney, M., and Fu, Y. Framedipt: Se(3) diffusion model for protein structure inpainting, November 2023a.

Zhang, J., Liu, Z., Wang, Y., and Li, Y. Subgdiff: A subgraph diffusion model to improve molecular representation learning, May 2024.

Zhang, M., Qamar, M., Kang, T., Jung, Y., Zhang, C., Bae, S.-H., and Zhang, C. A survey on graph diffusion models: Generative ai in science for molecule, protein and material. 2023b.

Zheng, Z., Zhang, B., Zhong, B., Liu, K., Li, Z., Zhu, J., Yu, J., Wei, T., and Chen, H.-F. Scaffold-lab: Critical evaluation and ranking of protein backbone generation methods in a unified framework, February 2024.

Zhu, T., Ren, M., and Zhang, H. Antibody design using a score-based diffusion model guided by evolutionary, physical and geometric constraints. In *Forty-first International Conference on Machine Learning*, 2024.

Ziv, Y., Marsden, B., and Deane, C. Molsnapper: Conditioning diffusion for structure based drug design. *bioRxiv*, pp. 2024–03, 2024a.

Ziv, Y., Marsden, B., and Deane, C. M. Molsnapper: Conditioning diffusion for structure based drug design, March 2024b.

A. Outline of Appendix

This appendix includes:

- **Table 1**: Comparison with other existing review papers.
- **Table 2:** Information list of the models mentioned in main text.
- **Section B:** Description of the benchmark for protein and small molecule, respectively.
- List of abbreviations.

Here we list the existing reviews on diffusion models for protein design, along with their frameworks and applications; see Table 1.

Table 1. Comparison of this review with existing surveys on Diffusion model for biomolecule generation: Frameworks and applications are enumerated

Surveys	**Frameworks**					**Applications**			
	Categorization	Benchmarks	Challenges	Future Works	mathematics behind	Protein generation	Peptide design	Molecule generation	Protein-ligand interaction
Ours	✓	✓	✓	✓	✓	✓	✓	✓	✓
(Norton & Bhattacharya, 2024)	✓	✓	✗	✓	✓	✓	✗	✗	✓
(Guo et al., 2023)	✓	✗	✗	✓	✗	✓	✗	✓	✓
(Zhang et al., 2023b)	✓	✓	✓	✗	✓	✓	✗	✓	✓

Models mentioned in this review have been implemented as open-source tools. We list their task, input, output, dataset for training, data size, and code link in Table 2. There are 12 models for protein design, and 18 models for small molecule generation, and 9 models for protein-ligand interaction, i.e., 30 models in total. This table may help users with their research problems and help developers further improve them.

Table 2: List of 30 models mentioned in the manuscript with their task, input, output, dataset, data size, code link and reference.

Task	**Paper**	**Input**	**output**	**Dataset**	**Data Size**	**Code**	**Ref**
Protein	RFDiffusion	structures	structures	PDB	-	code	(Watson et al., 2023)
	RFAA	sequence	structures	PDB	121,800	code	(Krishna et al., 2024)
	FrameDiff	structures	structures	PDB	20,312 back-bones	code	(Yim et al., 2023)
	FrameDiPT	structures	structures; Full-atom	RCEB; PDB	9K clusters	code	(Zhang et al., 2023a)
	TDS	structures	structures	-	-	code	(Wu et al., 2024)

	SMCDiff	motif	scaffolds	PDB	4,269	code	(Trippe et al., 2023)
	VFN-Diff	structures	structures	PDB	-	code	(Mao et al., 2023)
	Genie	structures	structures	SCOPe	195,214	code	(Lin & AlQuraishi, 2023)
	Genie2	structures	structures	PDB; AFDB	588,570 structures	code	(Lin et al., 2024b)
	Chroma	sequence	structures	PDB, UniProt, PFAM	28,819 structures	code	(Ingraham et al., 2023)
	AlphaFold3	sequence; SMILES	structures	PDB 2021	41,000,000 structures	code	(Abramson et al., 2024)
	PG	sequences	structures, sequences	-	-	code	(Lisanza et al., 2023)
Molecule	EDM	structures	structures	QM9; GEOM-Drugs	100K	-	(Hoogeboom et al., 2022)
	MDM	geometries	geometries	QM9; GEOM	290K	-	(Huang et al., 2022)
	GCDM	3D graph	3D graph	QM9; GEOM-Drugs	100K	code	(Morehead & Cheng, 2024)
	DiffSBDD	pockets	ligands	CrossDocked; Binding MOAD	-	code	(Schneuing et al., 2023)
	GeoLDM	geometries	structures	QM9; GEOM-Drugs	-	code	(Xu et al., 2023)
	MiDi	graph structures	graph	QM9; GEOM-Drugs	-	code	(Vignac et al., 2023)
	DiffLinker	structures	Molecule structures	ZINC, CASF, GEOM	185,678 examples	code	(Igashov et al., 2024)
	PMDM	Molecule, protein pocket	molecule structures	CrossDocked	22.5 million docked protein-ligand pairs	code	(Huang et al., 2024)

EQGAT-Diff	structures	molecule structures	QM9; GEOM-Drugs; CrossDocked; PubChem3D	-	code	(Le et al., 2023)
DiffBP	binding site	molecule structures	CrossDocked	10,000 protein-ligand paired samples	-	(Lin et al., 2024a)
Keypoint Diffusion	molecule structures	ligands	BindingMOAD	40,000	code	(Dunn & Koes, 2023)
Geodiff	molecular graphs	molecular conformations	QM9; GEOM-Drugs	200,000 conformations	code	(Xu et al., 2022)
TargetDiff	binding site	binding molecules	CrossDocked2020	100,000 complexes	code	(Guan et al., 2023)
MolDiff	molecular structures	molecular structures	QM9; GEOM-Drugs	231,523 molecules	code	(Peng et al., 2023)
MolSnapper	Protein-ligand complex	molecules	CrossDocked; Binding MOAD	-	code	(Ziv et al., 2024a)
CGD	molecule graph	molecule graph	Zink	250 000 small molecules	code	(Klarner et al., 2024)
GDSS	graph structures	structures	QM9 and ZINC250k	10,000 molecules	code	(Jo et al., 2022)
CDGS	graph	graph	ZINC250k; QM9	383,340 molecules	code	(Huang et al., 2023a)
JODO	graph	graph	QM9; GEOM-Drugs; ZINC250k; MOSES	2,621,542 molecules	code	(Huang et al., 2023b)
SubGDiff	molecular graph	graph	PCQM4Mv2	3.4 million molecules	code	(Zhang et al., 2024)
GaUDI	molecular graph	graph	cc-PBH; PAS	509,000 molucules	code	(Weiss et al., 2023)
SILVR	multiple superimposed fragments	graph	COVID Moonshot dataset	-	code	(Runcie & Mey, 2023)

SubDiff	subgraph	generative graph	GEOM-Drug; QM9	-	-	(Yang et al., 2024)

B. Benchmarks

B.1. Benchmarks for protein

To evaluate the performance of the models for protein backbone generation, it is crucial to establish and utilize robust benchmarks. These benchmarks not only facilitate the assessment of different generation methods, but also provide a standardized framework for comparing their strengths and limitations across various criteria. In the following, we outline several widely used benchmarks for evaluating protein backbone generation methods.

- PDB-struct (Wang et al.) suggests that encoder-decoder methods generally outperform structure-prediction-based methods in terms of refoldability, recovery, and stability metrics.
- Scaffold-Lab (Zheng et al., 2024) focuses on the evaluation of unconditional generation across metrics such as designability, novelty, diversity, efficiency, and structural properties.
- Melodia (Montalvão et al., 2024) is a Python library with a complete set of components devised for protein structural analysis and visualization using the differential geometry of three-dimensional curves and knot theory. Residue-wise confidence predicted local distance different test (pLDDT) and pairwise confidence predicted alignment error (PAE).
- PINDER (Kovtun et al., 2024) offers substantial advancement in the field of deep learning-based protein-protein docking and complex modeling by addressing key limitations of existing training and benchmark datasets.
- ProteinInvBench (Gao et al., 2024) is a benchmark for protein design, which comprises extended protein design tasks, integrated models, and diverse evaluation metrics (see Fig. 5).

B.2. Benchmarks for molecule generation

The goal of unconstrained molecular generation is to generate molecules that are:

- Valid and unique. Validity is the percentage of valid molecules measured by RDKit (Bento et al., 2020); Uniqueness is the percentage of unique molecules among valid molecules.
- Based on a chemical distribution corresponding to the training set.
- Novel and diverse. Novelty is the percentage of valid molecules not found in the training set. Diversity is the opposite of recovery and is meaningless if we measure it alone. If we examine sequence diversity and structural sc-TM together, we could gain a more comprehensive understanding of the designable protein space. To expand sequence diversity, we need to allow perturbations in the conformation of the protein backbone.

Continuous Automated Model Evaluation (CAMEO) (Haas et al., 2018) ligand-docking evaluation, publishes weekly benchmarking results based on models collected during a 4-day prediction window and evaluates their performance. The Frachet ChemNetDistance (FCD) measures the similarity between molecules in the training set and in the test set using the embedding learned by a neural network.

List of Abbreviations

$SE(3)$ special euclidean 3D group

AF2 AlphaFold2

AF3 AlphaFold3

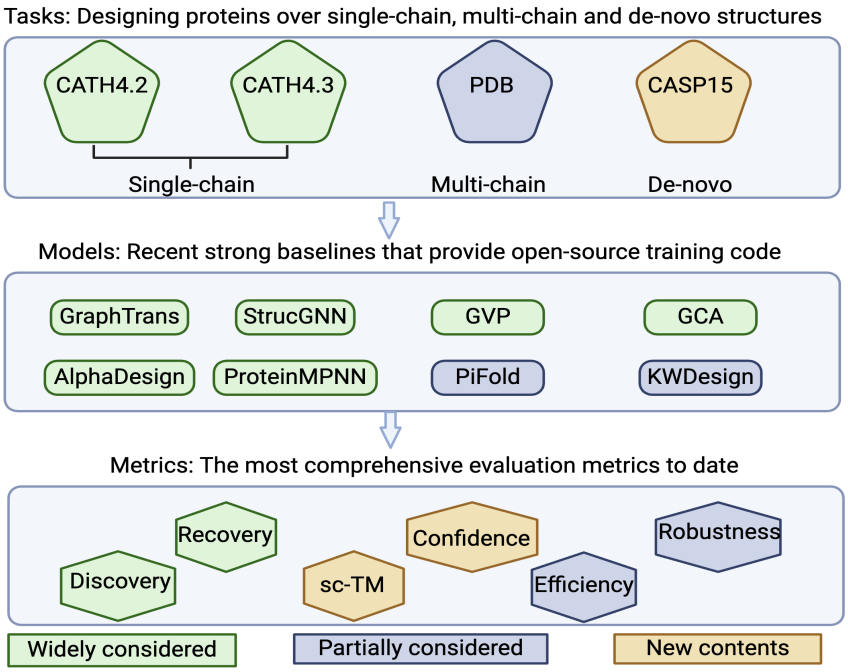

Figure 5. The framework of ProteinInvBench (Gao et al., 2024): tasks ⇒ models ⇒ metrics. Green, blue, yellow: widely considered, partially considered, newly introduced contents.

AMPs Antimicrobial peptides

CDGS Conditional Diffusion model based on discrete Graph Structures

CGD context-guided diffusion

DDPM Denoising Diffusion Probabilistic Models

DFM Discrete Flow Model

DisCo-Diff Discrete-Continuous Latent Variable Diffusion Models

DNPD De Novo Protein Design

EDM $E(3)$ equivariant diffusion model

EGNN $E(3)$ Equivariant Graph Neural Networks

EQGAT Equivariant Graph Attention Networks

ETNN E(n)-Equivariant Topological Neural Networks

FrameDipT FrameDiff inPainTing

GANs Generative Adversarial Networks

GCDM Geometry-Complete Diffusion Model

GDSS Graph Diffusion via the System of Stochastic differential equations

GeoLDM Geometric Latent Diffusion Models

GNN Graph Neural Network

GVP Geometric Vector Perception

IPA Invariant Point Attention

JODO joint 2D and 3D diffusion models

MiDi Mixed Graph+3D Denoising Diffusion

MSE Mean Square Error

NequIP Neural Equivariant Interatomic Potentials

ODE Ordinary Differential Equation

OOD out-of-distribution

PAE Predicted Alignment Error

pLDDT Predicted Local Distance Different Test

PLM ProteinLanguage Model

rEGNN relaxedEGNN

RFAA RoseTTAFold All-Atom

RFDiffusion RoseTTAFold Diffusion

RMSD Root Mean Square Deviation

SBDD Structure-based Drug Design

SDE Stochastic Differential Equation

SGM Score-based Generative Models

SI Supplementary Information

SILVR Selective Iterative Latent Variable Refinement

SubDiff subgraph latent diffusion model

TDS Twisted Diffusion Sampler

VAEs Variational Autoencoders

VFN Vector Field Networks