Antibody Design and Optimization with Multi-scale Equivariant Graph Diffusion Models for Accurate Complex Antigen Binding

Jiameng Chen[1] , **Xiantao Cai**[1*] , **Jia Wu**[2] and **Wenbin Hu**[3,1*]

[1]School of Computer Science, Wuhan University, Wuhan, China
[2]Department of Computing, Macquarie University, Sydney, Australia
[3]Wuhan University Shenzhen Research Institute, Shenzhen, China
{jiameng.chen, caixiantao, hwb}@whu.edu.cn, jia.wu@mq.edu.au

Abstract

Antibody design remains a critical challenge in therapeutic and diagnostic development, particularly for complex antigens with diverse binding interfaces. Current computational methods face two main limitations: (1) capturing geometric features while preserving symmetries, and (2) generalizing novel antigen interfaces. Despite recent advancements, these methods often fail to accurately capture molecular interactions and maintain structural integrity. To address these challenges, we propose **AbMEGD**, an end-to-end framework integrating **M**ulti-scale **E**quivariant **G**raph **D**iffusion for antibody sequence and structure co-design. Leveraging advanced geometric deep learning, AbMEGD combines atomic-level geometric features with residue-level embeddings, capturing local atomic details and global sequence-structure interactions. Its E(3)-equivariant diffusion method ensures geometric precision, computational efficiency, and robust generalizability for complex antigens. Furthermore, experiments using the SAbDab database demonstrate a 10.13% increase in amino acid recovery, 3.32% rise in improvement percentage, and a 0.062 Å reduction in root mean square deviation within the critical CDR-H3 region compared to DiffAb, a leading antibody design model. These results highlight AbMEGD's ability to balance structural integrity with improved functionality, establishing a new benchmark for sequence-structure co-design and affinity optimization. The code is available at: https://github.com/Patrick221215/AbMEGD.

1 Introduction

Antibodies are essential for targeting disease-related antigens in the therapeutic and diagnostic fields [Carter, 2006]. Their binding specificity and affinity are governed by complementarity-determining regions (CDRs), which exhibit high sequence and structural diversity [North *et al.*, 2011]. As a result, designing antibodies to effectively bind complex
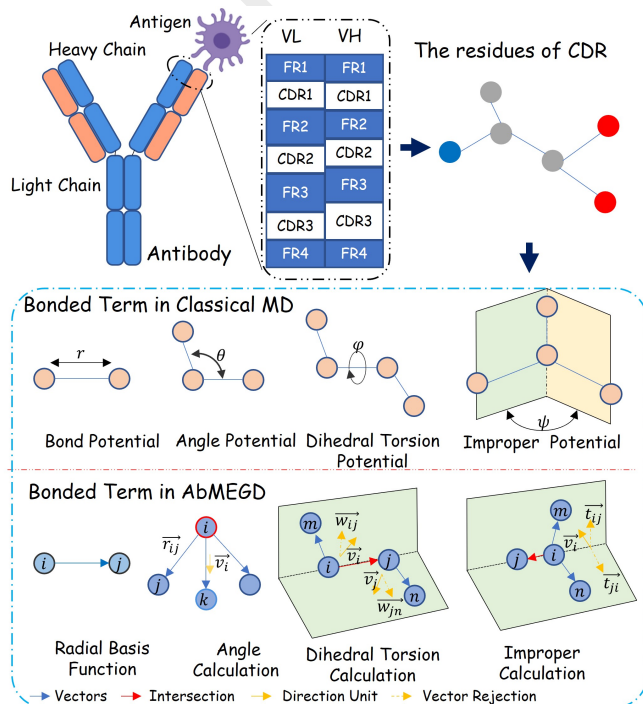
---

*Corresponding author

Figure 1: The antibody-antigen complex structure and CDRs involved in AbMEGD's geometric calculations. This figure translates bonded terms from classical molecular dynamics (MD)—such as bond length, bond angle, dihedral torsion, and improper angle—into efficient linear-time model operations.

antigens requires addressing two key challenges: (i) simplifying geometric feature extraction while preserving critical symmetries, and (ii) enhancing model generalization to novel antigen interfaces. This is because current methods often fail to capture precise molecular interactions and preserve structural integrity.

First,the complexity of extracting geometric features while accurately modeling symmetries essential for antibody functionality poses a critical challenge in therapeutic antibody design. Thus, accurate CDR modeling is crucial, as geometric attributes such as bond lengths, angles, and torsions regulate antigen-binding capabilities [Carter, 2006]. However, maintaining symmetries such as rotation and translation invariance

during feature extraction remains difficult. Traditional methods such as molecular dynamics (MD) and density functional theory provide precise geometric representations but are computationally restrictive in large-scale antibody modeling [Cai *et al.*, 2024; Luo *et al.*, 2022]. Consequently, advances in geometric deep learning (GDL) have introduced symmetry-based inductive biases, enabling accurate molecular modeling with smaller datasets [Shen *et al.*, 2023]. For example, equivariant graph neural networks (EGNNs), central to GDL, efficiently capture molecular properties by embedding symmetries (e.g., rotation and translation invariance) [Garcia Satorras *et al.*, 2021]. In comparison, innovations such as GearBind improve binding affinity predictions through multi-relational graph construction and contrastive pretraining [Cai *et al.*, 2024], while IgFold highlights the importance of structural precision in antibody design using advanced deep learning [Ruffolo *et al.*, 2023]. Additionally, models like AbX integrate evolutionary and physical constraints with geometric factors to optimize antibody sequences [Zhu *et al.*, 2024].

Second, generalizing antibody design to out-of-distribution antigen interfaces presents a significant challenge [Li *et al.*, 2024b; Li *et al.*, 2024a], complicating the development of effective antibodies. Antibody-antigen interactions are inherently unique and ad-hoc. This is because antibodies evolve rapidly to counter diverse pathogens without the long co-evolutionary histories typical of other protein-protein interactions [Adolf-Bryfogle *et al.*, 2018; Kong *et al.*, 2022]. As a result, predictive models often fail to adapt to novel antigens that differ substantially from the existing training datasets [Joubbi *et al.*, 2024]. As antibody design continues to evolve, integrating advanced computational techniques with a deeper understanding of structural dynamics will be crucial for overcoming these challenges and unlocking the full potential of therapeutic antibodies.

To address these challenges, we propose AbMEGD, an end-to-end solution integrating Multi-scale Equivariant Graph Diffusion for co-designing antibody sequences and structures. Inspired by ViSNet [Wang *et al.*, 2024], which balances computational efficiency with effective geometric information utilization, AbMEGD incorporates geometry calculation and vector–scalar interactive equivariant message passing (ViS-MP). These strategies enable efficient modeling of atomic-level features such as bond angles, dihedral torsions, and improper angles, aligning with classical molecular dynamics force fields while ensuring geometric precision and scalability. AbMEGD combines atomic-level geometric features with residue-level embeddings in a multi-scale framework, capturing local atomic details and global sequence–structure interactions to maintain CDR structural fidelity. Leveraging ViS-MP, the model facilitates interactions between vector representations and scalar embeddings, fully exploiting geometric features while preserving E(3)-equivariance. This multi-scale architecture simplifies geometric feature extraction and enhances generalization to complex antigen interfaces. Unlike existing methods such as GEOAB [Lin *et al.*, 2024], which separate structure generation and affinity prediction, AbMEGD employs a unified framework that preserves essential geometric symmetries and addresses the computational challenges of high-throughput antibody design.

We validate AbMEGD using the SAbDab database [Dunbar *et al.*, 2014], which includes diverse antigen-antibody complexes. Compared to DiffAb, a leading model in antibody design, AbMEGD achieves a 10.13% increase in amino acid recovery (AAR), a 3.32% rise in improvement percentage (IMP), and a 0.062 Å reduction in root mean square deviation (RMSD) within the critical CDR-H3 region. These results highlight its capacity to generate accurate structures and high-affinity antibodies, particularly for complex antigens. AbMEGD also excels in CDR design and affinity optimization tasks. Furthermore, its robust generalizability supports high-affinity antibody design for novel and out-of-distribution antigens, establishing a new benchmark in balancing structural conservation with improved functionality.

In summary, our contributions are as follows:

- **Unified Generative Framework with Multi-Scale Equivariant Graph Diffusion:** We propose AbMEGD, an end-to-end framework that unifies sequence-structure co-design and affinity optimization. By leveraging a multi-scale architecture, AbMEGD captures global sequence-structure interactions and local atomic details, preserving geometric symmetry while ensuring accuracy and efficiency.

- **Superior Performance and Broad Applicability:** Evaluations on the SAbDab database confirm Ab-MEGD's superior performance in sequence-structure co-design and affinity optimization. The model achieves structural fidelity and enhances functionality, showcasing adaptability across antibody design tasks and setting a new benchmark in therapeutic antibody development.

2 Related Work

2.1 Computational Antibody Design

Traditional antibody design methods rely on statistical energy functions and Monte Carlo sampling for sequence-structure optimization [Adolf-Bryfogle *et al.*, 2018; Ruffolo *et al.*, 2021], but suffer from inefficiency and local optima due to the rugged energy landscape. They also struggle with the vast sequence-structure space, especially for complex antigens, limiting design diversity. Deep learning has enabled sequence-based generative models that treat proteins as language [Alley *et al.*, 2019; Shin *et al.*, 2021], offering efficient sequence generation but often neglecting structural cues critical for antigen binding. The first CDR co-design model [Jin *et al.*, 2021], targeting SARS-CoV-2, incorporated structure but depended on antigen-specific predictors, limiting generalizability. Moreover, it focused on backbone atoms while ignoring side-chain orientations essential for protein-protein interactions. In contrast, our method captures both backbone and side-chain features via multi-level representation, enabling a more complete modeling of the antigen-antibody interface for high-affinity binding.

2.2 Diffusion-based Generative Model

Diffusion probabilistic models (DDPMs) have gained traction for generating protein sequences and structures by denoising
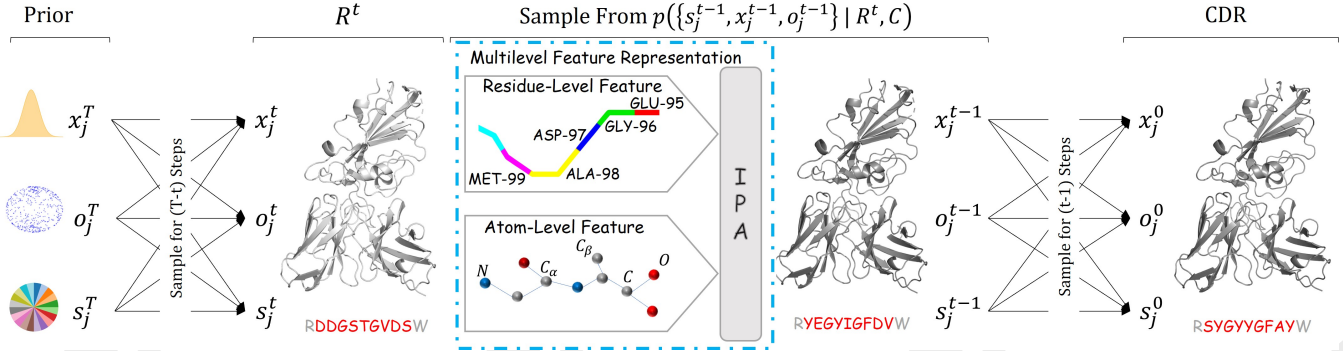
Figure 2: This figure depicts the generative diffusion process. At every iteration, the network uses the current state of the CDR as input to model the distribution for the sequences, positions, and orientations of the CDR in the subsequent step.

samples from a prior distribution [Sohl-Dickstein *et al.*, 2015; Song and Ermon, 2019; Ho *et al.*, 2020; Martinkus *et al.*, 2024]. Models like DiffAb [Luo *et al.*, 2022] and DIFF-FORCE [Kulytė *et al.*, 2024] have demonstrated superior efficiency in generating antibody sequences and structures. However, these models primarily focus on residue-level features and often fail to capture atomic-level interactions essential for high-specificity antibody-antigen binding. In comparison, our approach enhances existing DDPMs using VisNet-inspired graph-based modules [Wang *et al.*, 2024] to model atomic-level geometric features. This facilitates the effective representation of complex three-dimensional (3D) spatial relationships among the atoms in antibodies and antigens, optimizing antibody specificity and affinity through detailed atomic interaction modeling.

3 Methodology

This section is organized as follows: Section 3.1 introduces the notations used throughout this paper and formally defines the problem. Section 3.2 presents the multi-level feature generation and unified representation mechanism, integrating residue- and atom-level features. Finally, Section 3.3 details the diffusion processes tailored for modeling antibodies.

3.1 Problem Definitions and Notations

In this study, we represent each amino acid in a protein complex using three components: type, Cα atom coordinates, and orientation. Specifically, the amino acid type is denoted as $s_{r_i} \in \{ACDEFGHIKLMNPQRSTVWY\}$, where $i = 1, \ldots, N$ and N is the total number of amino acids in the complex. The spatial position of each residue is described by the Cα coordinates $\mathbf{x}_{r_i} \in R^3$, and the orientation is denoted as $\mathbf{O}_{r_i} \in SO(3)$, which captures the geometric directionality in 3D space.

We consider the antigen structure and the antibody framework as given, focusing our design efforts on the CDRs of the antibody. Specifically, let the generated CDR consist of m amino acids, indexed from $c + 1$ to $c + m$, the set of residues in the CDR is denoted as $\mathcal{R} = \{(s_{r_j}, \mathbf{x}_{r_j}, \mathbf{O}_{r_j}) \mid j = c + 1, \ldots, c + m\}$. Our objective is to model the conditional distribution $P(\mathcal{R} \mid \mathcal{C})$, where $\mathcal{C} = \{(s_{r_i}, \mathbf{x}_{r_i}, \mathbf{O}_{r_i}) \mid$

$r_i \in \{1, \ldots, N\} \setminus \{c+1, \ldots, c+m\}\}$ represents the antigen-antibody framework excluding the CDR regions.

3.2 Multi-level Feature Generation and Unified Representation Mechanism

The overall workflow of AbMEGD is illustrated in Figure 2. This study presents an innovative multi-level feature generation and unified representation mechanism that seamlessly integrates residue- and atom-level features. This approach leverages the complementary nature of these features to simultaneously capture global structural information and fine-grained geometric details, enabling comprehensive characterization of antibody-antigen complexes.

Residue-Level Feature Generation. At the residue level, we employ a multi-layer perceptron (MLP) module to process features of individual residues and residue pairs. We define the following variables to represent the feature states of the CDRs: $\{s_{r_j}^t, \mathbf{x}_{r_j}^t, \mathbf{O}_{r_j}^t\}_{j=1}^{c+m} \cup \{s_{r_i}^t, \mathbf{x}_{r_i}^t, \mathbf{O}_{r_i}^t\}_{i \in \{1, \ldots, N\} \setminus \{c+1, \ldots, c+m\}}$, where $s_{r_j}^t$ denotes the amino acid type of residue j at time step t, $\mathbf{x}_{r_j}^t$ represents the Cα backbone coordinates, and $\mathbf{O}_{r_j}^t \in SO(3)$ indicates the orientation.

The single-residue features encapsulate each residue's chemical properties, geometric characteristics, and local structural environment. Thus, the features for residue r_i are defined as [Luo *et al.*, 2022]:

$$\mathbf{e}_{r_i} = \text{MLP}_{\text{res}}([\mathbf{t}_{r_i}, \mathbf{g}_{r_i}, \mathbf{d}_{r_i}, \mathbf{f}_{r_i}]), \qquad (1)$$

where $\mathbf{t}_{r_i}$ represents the amino acid type features, obtained by embedding the amino acid type s_{r_i} into a low-dimensional vector. $\mathbf{g}_{r_i}$ denotes the local geometric features, derived by constructing a local reference frame based on the backbone atomic coordinates $\mathbf{p}_{r_i}$. $\mathbf{d}_{r_i}$ encodes the backbone dihedral angles through angular encoding of ϕ_{r_i} and ψ_{r_i}, which describe the torsional conformation of the protein backbone. Finally, $\mathbf{f}_{r_i}$ represents the fragment type features, capturing local fragment-level structural information.

Pairwise residue features describe the interactions and geometric relationships between residues. The features for a residue pair (r_i, r_j) are represented by [Luo *et al.*, 2022]:

$$\mathbf{z}_{r_i, r_j} = \text{MLP}_{\text{pair}}([\mathbf{t}_{r_i, r_j}, \mathbf{r}_{r_i, r_j}, \mathbf{d}_{r_i, r_j}, \mathbf{o}_{r_i, r_j}]), \qquad (2)$$

where $\mathbf{t}_{r_i,r_j}$ represents the residue-pair type features, encoded by embedding the combination of the amino acid types s_{r_i} and s_{r_j}. $\mathbf{r}_{r_i,r_j}$ indicates the relative positional features of the residue pair, derived from the sequence index difference $\Delta_{r_i,r_j} = |r_i - r_j|$. $\mathbf{d}_{r_i,r_j}$ denotes the geometric distance features, calculated as the Euclidean distance between the backbone atoms of the two residues $\|\mathbf{p}_{r_i} - \mathbf{p}_{r_j}\|$ and further represented using Gaussian distribution encoding. Lastly, $\mathbf{o}_{r_i,r_j}$ refers to the orientation features of the residue pair, based on the relative orientation angles (ϕ_{r_i,r_j}, ψ_{r_i,r_j}). Specifically, ϕ_{r_i,r_j} and ψ_{r_i,r_j} capture spatial orientation and torsional relationships, ensuring rotationally invariant geometric properties.

Atom-Level Feature Generation. To enhance complex representation, we focus on heavy atoms (N, C_α, C, O, and side-chain C_β, excluding glycine which lacks C_β). The atom-level feature extraction process utilizes the positions of selected atoms $\mathbf{X} \in R^{A \times 3}$ and their atomic indices $\mathbf{Z} \in N^A$, where A is the number of heavy atoms.

Initially, an embedding block encodes the atomic indices and pairwise edge distances into a high-dimensional embedding space. Then, a series of ViS-MP blocks iteratively update the heavy atoms' scalar and vector representations based on their interactions. Subsequently, residual connections are incorporated between consecutive ViS-MP blocks [Wang *et al.*, 2024]. Thus, the final output includes node and edge features, representing the molecular structure at the atomic level.

Furthermore, the embedding process starts with the heavy atomic indices z_{a_i} of atom a_i, embedded along with the pairwise distances $e_{a_i a_j}$ using radial basis functions within a cutoff distance. The initial node and edge embeddings are defined as [Wang *et al.*, 2024]:

$$\mathbf{h}^0_{a_i}, \mathbf{f}^0_{a_i a_j} = \text{Embedding}(\mathbf{z}_{a_i}, \mathbf{z}_{a_j}, \mathbf{e}_{a_i a_j}), \quad a_j \in \mathcal{N}(a_i), \tag{3}$$

where $\mathcal{N}(a_i)$ denotes the set of first-order neighbors of atom a_i, and a_j represents one of its neighbors. The initial vector embedding $\mathbf{v}^0_{a_i}$ is initialized to $\mathbf{0}$ and projected into the embedding space as $\mathbf{v}^0_{a_i} \in R^{3 \times K}$, where K is the hidden dimension size.

Moreover, the ViS-MP mechanism facilitates the interaction between scalar and vector features, enhancing the geometric representation of the molecular structure. Therefore, the mechanism's key operations expressed as follows [Wang *et al.*, 2024]:

$$m^l_{a_i} = \sum_{a_j \in \mathcal{N}(a_i)} \phi^s_m\left(h^l_{a_i}, h^l_{a_j}, f^l_{a_i,a_j}\right), \tag{4}$$

$$\mathbf{m}^l_{a_i} = \sum_{a_j \in \mathcal{N}(a_i)} \phi^v_m\left(m^l_{a_i,a_j}, \mathbf{r}_{a_i,a_j}, \mathbf{v}^l_{a_j}\right), \tag{5}$$

$$h^{l+1}_{a_i} = \phi^s_{\text{un}}\left(h^l_{a_i}, m^l_{a_i}, \langle \mathbf{v}^l_{a_i}, \mathbf{v}^l_{a_i} \rangle\right), \tag{6}$$

$$f^{l+1}_{a_i,a_j} = \phi^s_{\text{ue}}\left(f^l_{a_i,a_j}, \langle \text{Rej}_{\mathbf{r}_{a_i,a_j}}(\mathbf{v}^l_{a_i}), \text{Rej}_{\mathbf{r}_{a_j,a_i}}(\mathbf{v}^l_{a_j}) \rangle\right), \tag{7}$$

$$\mathbf{v}^{l+1}_{a_i} = \phi^v_{\text{un}}\left(\mathbf{v}^l_{a_i}, m^l_{a_i}, \mathbf{m}^l_{a_i}\right), \tag{8}$$

Here, $\mathbf{h}^l_{a_i}$ denotes the scalar embedding of atom a_i at layer l, $f^l_{a_i,a_j}$ stands for the edge feature between atoms a_i and a_j

at layer l, and $\mathbf{v}^l_{a_i}$ represents the vector embedding of atom a_i at layer l. Meanwhile, $m^l_{a_i}$ is the aggregated scalar message for atom a_i at layer l, computed by the scalar message function ϕ^s_m. Similarly, $\mathbf{m}^l_{a_i}$ is the aggregated vector message for atom a_i at layer l, computed by the vector message function ϕ^v_m. The update functions ϕ^s_{un} and ϕ^v_{un} respectively update the scalar and vector embeddings based on the aggregated messages and current state. Likewise, the edge features $f^{l+1}_{a_i,a_j}$ are updated using ϕ^s_{ue}, which incorporates the vector rejection of embeddings $\mathbf{v}^l_{a_i}$ and $\mathbf{v}^l_{a_j}$ along the relative position vector $\mathbf{r}_{a_i a_j}$, ensuring rotationally invariant geometric representation.

Notably, this approach updates node and edge features during message passing, providing a more comprehensive geometric representation compared to previous studies [Schütt *et al.*, 2021; Thölke and De Fabritiis, 2022].

Residue and Atom Feature Integration. The residue features $(\mathbf{e}_{r_i}, \mathbf{z}_{r_i,r_j})$ and atom features $(h_{a_i}, \mathbf{v}_{a_i})$ are integrated and input into an invariant point attention (IPA) [Jumper *et al.*, 2021] module to capture interactions between residue and atom features:

$$\mathbf{h}_{r_j} = \text{IPA}(\mathbf{e}_{r_i}, \mathbf{z}_{r_i,r_j}, h_{a_i}, \mathbf{v}_{a_i}) \tag{9}$$

Furthermore, the IPA module facilitates learning local-to-global relationships within molecular structures by leveraging the interplay between residue and atom features.

3.3 Diffusion Processes

Unlike traditional models focused on denoising molecular states, our approach incorporates multi-level molecular features, enhancing accuracy and structural fidelity by leveraging residue- and atom-level information. Let $\mathcal{R}^t = \{s^t_j, \mathbf{x}^t_j, \mathbf{O}^t_j\}^{c+m}_{j=c+1}$ denote the intermediate states of residue j at time step t. At $t = 0$, the system represents real data, including observed CDR sequences and structures, while at $t = T$, it corresponds to samples from the prior distribution. Thus, the forward diffusion process adds noise incrementally from $t = 0$ to $t = T$, while the generative process denoises in reverse. Hence, the diffusion processes for amino acid types s^t_j, coordinates $\mathbf{x}^t_j$, and orientations $\mathbf{O}^t_j$ are defined as follows:

Multinomial Diffusion for Amino Acid Types. The forward diffusion process for amino acid types follows a multinomial distribution, formulated as [Hoogeboom *et al.*, 2021]:

$$q(s^t_j|s^{t-1}_j) = \text{Multinomial}\Big((1 - \beta^{\text{type}}_t) \cdot \text{onehot}(s^{t-1}_j) + \beta^{\text{type}}_t \cdot \frac{1}{20} \cdot \mathbf{1}\Big), \tag{10}$$

where the onehot function maps an amino acid type into a 20-dimensional one-hot vector, and $\mathbf{1}$ represents an all-one vector. The parameter β^t_{type} denotes the probability of uniformly resampling an amino acid across 20 types. As $t \to T$, β^t_{type} approaches 1, driving the distribution closer to uniformity. Thus, the following probability density offers an efficient mechanism to perturb s^t_j at each timestep t during training [Hoogeboom *et al.*, 2021]:

$$q(s_j^t \mid s_j^0) = \text{Multinomial}\Big(\tilde{\alpha}_{\text{type}}^t \cdot \text{onehot}(s_j^0)$$
$$+ (1 - \tilde{\alpha}_{\text{type}}^t) \cdot \frac{1}{20} \cdot \mathbf{1}\Big), \qquad (11)$$

where $\tilde{\alpha}_{\text{type}}^t = \prod_{\tau=1}^t (1 - \beta_{\text{type}}^\tau)$.

The generative diffusion process is defined as:

$$p(s_j^{t-1} \mid \mathcal{R}^t, \mathcal{C}) = \text{Multinomial}\left(F(\mathcal{R}^t, \mathcal{C})[j]\right), \qquad (12)$$

where $F(\cdot)[j]$ represents a neural network model that takes as input both the structural context (comprising the antigen and antibody framework) and the CDR state from the previous step. This model predicts the probability distribution over the 20 amino acid types for the j-th residue in the CDR.

Unlike forward diffusion, the generative process relies on the structural context C and the CDR state from the previous step, which includes positions and orientations. Generative diffusion aims to approximate the posterior $q(s_j^{t-1} \mid s_j^t, s_j^0)$, derived from Eq.10 and Eq.11, for denoising purposes. Thus, the objective of training the generative diffusion process to predict amino acid types is to minimize the expected Kullback-Leibler (KL) divergence between Eq.12 and the posterior distribution. Let D_{KL}^j denote the KL divergence for the j-th term. Then, we define the loss function as:

$$L_{\text{type}}^t = E_{\mathcal{R}^t \sim \mathcal{P}}\left[\frac{1}{m}\sum_j D_{\text{KL}}^j\right], \qquad (13)$$

where $D_{\text{KL}}^j = D_{\text{KL}}\left(q(s_j^{t-1} \mid s_j^t, s_j^0)\|p(s_j^{t-1} \mid \mathcal{R}^t, C)\right)$.

C_α **Coordinate Diffusion.** To accurately capture the hierarchical spatial configurations between residues and atoms, our framework models the forward diffusion of normalized C_α coordinates $\mathbf{x}_j^t$ as a Gaussian distribution:

$$q(\mathbf{x}_j^t \mid \mathbf{x}_j^{t-1}) = \mathcal{N}\left(\mathbf{x}_j^t \mid \sqrt{1 - \beta_{\text{pos}}^t} \cdot \mathbf{x}_j^{t-1}, \beta_{\text{pos}}^t \mathbf{I}\right), \quad (14)$$

$$q\left(\mathbf{x}_j^t \mid \mathbf{x}_j^0\right) = \mathcal{N}\left(\mathbf{x}_j^t \mid \sqrt{\bar{\alpha}_{\text{pos}}^0} \cdot \mathbf{x}_j^0, \left(1 - \bar{\alpha}_{\text{pos}}^0\right)\mathbf{I}\right), \quad (15)$$

where the diffusion rate β_{pos}^t determines the noise level at each timestep t, increasing from 0 to 1 as the process progresses. The cumulative scaling factor, $\bar{\alpha}_{\text{pos}}^t = \prod_{\tau=1}^t (1 - \beta_{\text{pos}}^\tau)$. By applying the reparameterization trick [Ho *et al.*, 2020], the generative diffusion process is formalized as:

$$p(\mathbf{x}_j^{t-1} \mid \mathcal{R}^t, C) = \mathcal{N}\left(\mathbf{x}_j^{t-1} \mid \mu_p(\mathcal{R}^t, C), \beta_{\text{pos}}^t \mathbf{I}\right), \quad (16)$$

where the mean $\mu_p(\mathcal{R}^t, C)$ is computed as:

$$\mu_p(\mathcal{R}^t, C) = \frac{1}{\sqrt{\alpha_{\text{pos}}^t}}\left(\mathbf{x}_j^t - \frac{\beta_{\text{pos}}^t}{\sqrt{1 - \bar{\alpha}_{\text{pos}}^t}} G(\mathcal{R}^t, C)[j]\right). \qquad (17)$$

where $G(\cdot)[j]$ is a neural network model that predicts the standard Gaussian noise $\epsilon_j \sim \mathcal{N}(0, \mathbf{I})$, which is added to the scaled coordinate of amino acid j, $\sqrt{\bar{\alpha}_{\text{pos}}^0}\mathbf{x}_j^0$, following the reparameterization process described in Eq.15: $\mathbf{x}_j^t = \sqrt{\bar{\alpha}_{\text{pos}}^0}\mathbf{x}_j^0 + \sqrt{1 - \bar{\alpha}_{\text{pos}}^0}\epsilon_j$.

The objective function for training the generative process is the expected mean squared error (MSE) between G and ϵ_j. This is derived by aligning the distribution p with the posterior $q(x_j^{t-1} \mid x_j^t, x_j^0)$ [Ho *et al.*, 2020]:

$$L_{\text{pos}}^t = E\left[\frac{1}{m}\sum_j \left\|\epsilon_j - G(\mathcal{R}^t, C)[j]\right\|^2\right]. \qquad (18)$$

SO(3)-Based Orientation Diffusion. We propose a diffusion framework based on the $SO(3)$, extending the method [Leach *et al.*, 2022] by incorporating multi-level features. The diffusion process on $SO(3)$ is modeled as an isotropic Gaussian distribution, with dynamics governed by the diffusion parameter β_{ori}^t. This process transitions molecular orientations from well-defined real states at $t = 0$ to near-uniform distributions at $t = T$, simulating the gradual addition of noise [Leach *et al.*, 2022]:

$$T_j^t = \mathbf{O}_j^t \cdot \text{ScaleRot}\left(\sqrt{\bar{\alpha}_{\text{ori}}^t}, \mathbf{O}_j^0, 1 - \bar{\alpha}_{\text{ori}}^t\right), \qquad (19)$$

$$q(\mathbf{O}_j^t \mid \mathbf{O}_j^0) = \mathcal{IG}_{SO(3)}\left(T_j^t\right), \qquad (20)$$

$\mathcal{IG}_{SO(3)}$ represents the isotropic Gaussian distribution on $SO(3)$, parameterized by a mean rotation and a scalar variance [Leach *et al.*, 2022; Matthies *et al.*, 1988; Nikolayev and Savyolov, 1997]. Then, the ScaleRot operation modifies the rotation matrix by scaling its rotation angle while keeping the rotation axis fixed [Gallier and Xu, 2003]. $\bar{\alpha}_{\text{ori}}^t = \prod_{\tau=1}^t (1 - \beta_{\text{ori}}^\tau)$, where β_{ori}^τ represents the variance that increases with the diffusion step t. The conditional distribution used for the orientation generation process is defined as:

$$p(\mathbf{O}_j^{t-1} \mid \mathcal{R}^t, \mathcal{C}) = \mathcal{IG}_{SO(3)}\left(\mathbf{O}_j^{t-1} \cdot H(\mathcal{R}^t, \mathcal{C})[j], \beta_{\text{ori}}^t\right), \qquad (21)$$

Here, $H(\cdot)[j]$ represents a neural network designed to denoise orientations, producing the denoised orientation matrix for amino acid j. Training the conditional distribution involves aligning the predicted orientation from $H(\cdot)$ with the real orientation. Consequently, the training objective is formulated to minimize the expected discrepancy, measured using the inner product between the real and predicted orientation matrices:

$$L_{\text{ori}}^t = E\left[\frac{1}{m}\sum_j \left\|\left(O_j^0\right)^\top \hat{O}_j^{t-1} - I\right\|_F^2\right], \qquad (22)$$

where $\hat{\mathbf{O}}_j^{t-1} = H(\cdot)[j]$ represents the predicted orientation matrix for amino acid j.

Method	AAR (%) ↑			IMP (%) ↑			RMSD (Å) ↓		
	H1	H2	H3	H1	H2	H3	H1	H2	H3
RAbD	22.85	25.50	22.14	43.88	**53.50**	23.25	2.261	1.641	**2.900**
DiffAb	**67.63**	44.20	28.70	42.84	30.26	20.42	<u>1.248</u>	**1.108**	3.481
DIFFFORCE	60.78	<u>53.51</u>	<u>29.52</u>	<u>49.45</u>	<u>36.81</u>	**30.22**	1.561	1.401	3.612
AbMEGD	<u>66.96</u>	**55.40**	**38.93**	**55.11**	24.05	<u>23.74</u>	**1.169**	<u>1.130</u>	<u>3.419</u>

Table 1: Evaluation of the generated antibody CDRs (sequence-structure co-design) by different models. The best result for each metric is highlighted in **bold**, while the second-best result is indicated with <u>underlining</u> for clearer comparison.

The Overall Training Objective. By summing Eq.13, 18, and 22, and taking the expectation over t, the final training objective function is formulated as:

$$L = E_{t \sim \text{Uniform}(1,...,T)} \left[L_{\text{type}}^t + L_{\text{pos}}^t + L_{\text{ori}}^t \right]. \quad (23)$$

Training begins by selecting a random time step t and generating noisy states $\{s_j^t, \mathbf{x}_j^t, \mathbf{O}_j^t\}_{j=c+1}^{c+m} \sim p$ through the diffusion process, which incorporates noise into the data as described in Eq.11, 15, and 20. The model parameters are updated by computing the loss based on the perturbed data and applying backpropagation.

4 Experiments

We propose AbMEGD, an end-to-end framework that integrates multi-scale equivariant graph diffusion for co-designing antibody sequences and structures. It addresses two key tasks: sequence-structure co-design (Section 4.1) and antibody optimization (Section 4.2).

4.1 Sequence-Structure Co-design

We evaluated our model using data from the SAbDab [Dunbar *et al.*, 2014], collected up to September 2024. Antibody-antigen complexes with resolutions worse than 4Å and those binding non-protein antigens were excluded. Antibodies were clustered by 50% CDR-H3 sequence identity, as in [Luo *et al.*, 2022], and five clusters (19 antibody-antigen complexes) formed the test set, including antigens from SARS-CoV-2, MERS, and influenza. The remaining clusters formed the training set.

For evaluation, we removed the CDRs from the test set, generating sequences and structures for the removed regions. We fixed the CDR length to the original value (though other lengths could be explored). Each model generated 100 samples per CDR, with all generated and native structures refined using AMBER99 force field [Lindorff-Larsen *et al.*, 2010] in OpenMM [Eastman *et al.*, 2017] and Rosetta [Alford *et al.*, 2017].

Baseline Models. To evaluate performance, as described in Section 2.1, we selected three baseline models for comparison. DiffAb [Luo *et al.*, 2022] represents a diffusion-based generative model specifically designed for antibody structure and sequence modeling. DIFFFORCE [Kulytė *et al.*, 2024] leverages a diffusion framework with force-guided

sampling to enhance the quality of generated molecular structures. Additionally, we included the traditional energy-based method RABD [Adolf-Bryfogle *et al.*, 2018], which utilizes energy minimization to design antibodies and optimize binding affinities.

Evaluation Metrics. The models are evaluated using four key metrics: (1) AAR measures how accurately the generated sequences match the native amino acid sequences; (2) RMSD quantifies the structural deviation between the generated and native CDRs by comparing the Cα atom coordinates;
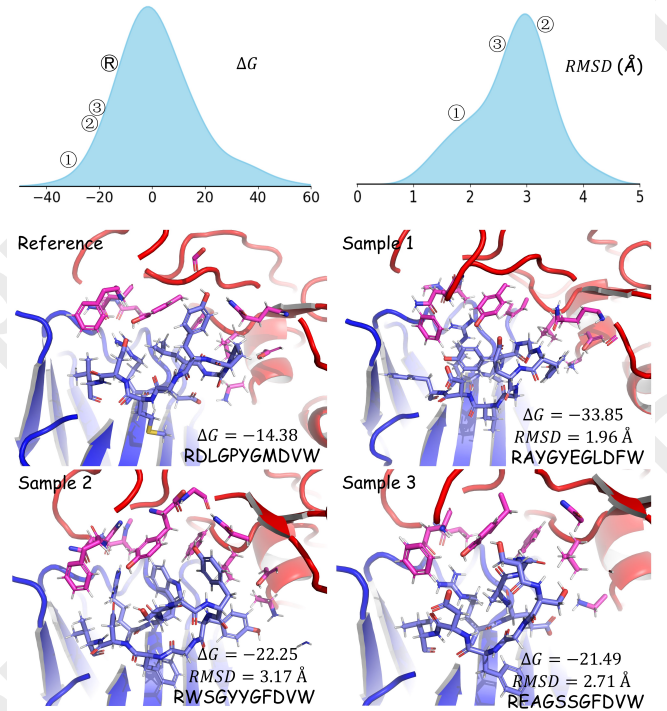

Figure 3: Examples of CDR-H3 designed using a sequence-structure co-design approach, showcasing the distributions of interaction energy (ΔG) and RMSD. The antigen-antibody complex is derived from PDB entry 7chf, with the antigen being the SARS-CoV-2 RBD. Sample 1 exhibits better complementarity with the antigen, while Sample 3 displays less effective binding, potentially explaining the variation in their interaction energy (ΔG).

and (3) IMP indicates the proportion of designed antibodies that achieve improved binding energy (ΔG) compared to the baseline, calculated using the InterfaceAnalyzer module in Rosetta.

Table 1 compares four antibody design models across three metrics. AbMEGD excels overall, leading in AAR for H2 and H3, with competitive results in H1. It achieves the lowest RMSD in H1, indicating high structural accuracy, and performs well in H2 and H3 RMSD. While DIFFFORCE slightly outperforms AbMEGD in IMP for H2 and H3, AbMEGD leads in H1, balancing binding energy improvement with sequence recovery. DIFFFORCE demonstrates its strength in affinity optimization, particularly in H3, while DiffAb performs well in AAR and achieves the lowest RMSD in H2 but lags in IMP. RAbD shows moderate IMP in H2 but low AAR, indicating poor sequence recovery, with favorable RMSD only in H3 due to reliance on the Rosetta energy function. In contrast, AbMEGD achieves robust energy improvements without Rosetta supervision. Building on DiffAb, which is widely recognized as a leading model in antibody design, AbMEGD significantly improves performance in the critical CDR-H3 region, achieving a 10.13% increase in AAR (from 28.70% to 38.93%), a 3.32% boost in IMP (from 20.42% to 23.74%), and a 0.062 Å reduction in RMSD (from 3.481 Å to 3.419 Å). These results highlight AbMEGD's capability to refine sequence-structure co-design and affinity optimization, surpassing current models. Figure 3 illustrates three designed examples of CDR-H3 interacting with the SARS-CoV-2 RBD. Among these, Sample 1 exhibits the strongest binding affinity, as reflected by its lowest binding energy, and demonstrates superior structural complementarity with the antigen. In contrast, Sample 3 shows a higher binding energy compared to the reference, and its CDR shape appears less compatible with the antigen, indicating weaker interactions.
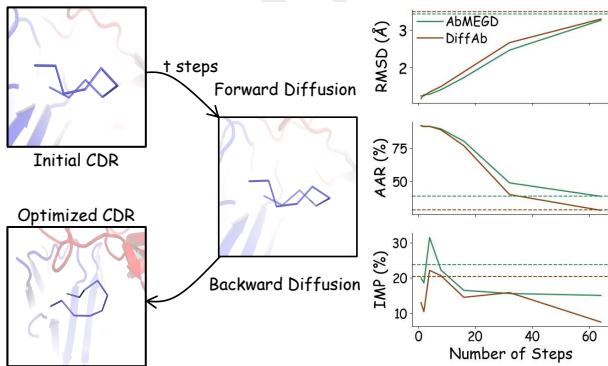

Figure 4: (a) The antibody fenoptimization algorithm applies a forward diffusion process to perturb the initial CDR over t steps, followed by a backward diffusion process to refine it into the optimized CDR. (b) The IMP, RMSD, and SeqID metrics of the CDRs were evaluated across varying optimization steps. The dashed lines indicate the performance of CDRs generated through de novo design. At $t = 4$, the optimized CDRs achieve an IMP score comparable to that of the de novo designs while retaining a structural similarity to the original CDR.

4.2 Antibody Optimization

Steps	RMSD (Å) ↓		AAR (%) ↑		IMP (%) ↑	
	AbMEGD	DiffAb	AbMEGD	DiffAb	AbMEGD	DiffAb
1	1.166	1.239	91.63	91.60	20.00	13.04
2	1.274	1.244	90.97	91.23	18.64	10.53
4	1.284	1.342	91.18	91.19	31.48	22.18
8	1.404	1.490	89.27	88.58	22.24	20.66
16	1.728	1.880	80.09	76.70	16.53	14.56
32	2.458	2.657	48.97	40.24	15.64	15.89
64	3.253	3.286	38.91	28.44	15.11	7.58
T	3.419	3.481	38.93	28.70	23.74	20.42

Table 2: Evaluation of optimized CDR-H3s with different optimization steps. This table presents metrics including RMSD (Å), AAR (%), and IMP (%) to assess the performance of the optimization approach. Unlike redesigning methods, the optimization process improves binding energy while retaining structural similarity of the optimized CDR to the original one. Results for RMSD, AAR, and IMP are reported for AbMEGD and DiffAb. The accompanying line plot in Figure 4b further illustrates these results.

Our model is designed to optimize existing antibodies, providing an important application in the pharmaceutical domain. The optimization process begins by perturbing the CDR sequence and structure for t steps using the forward diffusion process. Afterward, starting from the $(T - t)$-th step (with t steps remaining) of the generative diffusion process, the sequences are denoised to produce a set of optimized antibodies. This procedure is illustrated in Figure 4a.

We focus specifically on optimizing the CDR-H3 antibody region in the test set with varying t values. For each antibody and t, the CDR is independently perturbed 100 times, yielding 100 optimized CDRs distinct from the original CDR. We evaluate the performance of our model and DiffAb by reporting the percentage of antibodies with IMP, RMSD, and AAR of the optimized CDRs relative to the original antibody. Additionally, we compare the optimized antibodies with *de novo* antibodies ($t = T = 100$) introduced in Section 4.1.

As depicted in Table 2 and Figure 4b , our optimization method produces antibodies with superior binding energy (IMP) as measured by the Rosetta energy function. At $t = 4$, the optimized CDRs reach an IMP score comparable to that of de novo antibodies, while maintaining significantly higher structural similarity to the original CDRs. The RMSD and AAR results further demonstrate that our model achieves a better trade-off between binding energy improvement and structural conservation compared to DiffAb, making it more suitable for practical applications.

5 Conclusion

In this study, we propose the AbMEGD, an advanced diffusion-based generative model addressing key challenges in antibody design. By integrating residue- and atomic-level features into a unified framework, AbMEGD achieves high performance in sequence-structure co-design and antibody optimization. The experimental results confirm its ability to balance structural conservation and enhance functionality during antibody refinement.

Ethical Statement

There are no ethical issues.

Acknowledgments

References

[Adolf-Bryfogle *et al.*, 2018] Jared Adolf-Bryfogle, Oleks Kalyuzhniy, Michael Kubitz, Brian D Weitzner, Xiaozhen Hu, Yumiko Adachi, William R Schief, and Roland L Dunbrack Jr. Rosettaantibodydesign (rabd): A general framework for computational antibody design. *PLoS computational biology*, 14(4):e1006112, 2018.

[Alford *et al.*, 2017] Rebecca F Alford, Andrew Leaver-Fay, Jeliazko R Jeliazkov, Matthew J O'Meara, Frank P DiMaio, Hahnbeom Park, Maxim V Shapovalov, P Douglas Renfrew, Vikram K Mulligan, Kalli Kappel, et al. The rosetta all-atom energy function for macromolecular modeling and design. *Journal of chemical theory and computation*, 13(6):3031–3048, 2017.

[Alley *et al.*, 2019] Ethan C Alley, Grigory Khimulya, Surojit Biswas, Mohammed AlQuraishi, and George M Church. Unified rational protein engineering with sequence-based deep representation learning. *Nature methods*, 16(12):1315–1322, 2019.

[Cai *et al.*, 2024] Huiyu Cai, Zuobai Zhang, Mingkai Wang, Bozitao Zhong, Quanxiao Li, Yuxuan Zhong, Yanling Wu, Tianlei Ying, and Jian Tang. Pretrainable geometric graph neural network for antibody affinity maturation. *Nature communications*, 15(1):7785, 2024.

[Carter, 2006] Paul J Carter. Potent antibody therapeutics by design. *Nature reviews immunology*, 6(5):343–357, 2006.

[Dunbar *et al.*, 2014] James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, Jiye Shi, and Charlotte M Deane. Sabdab: the structural antibody database. *Nucleic acids research*, 42(D1):D1140–D1146, 2014.

[Eastman *et al.*, 2017] Peter Eastman, Jason Swails, John D Chodera, Robert T McGibbon, Yutong Zhao, Kyle A Beauchamp, Lee-Ping Wang, Andrew C Simmonett, Matthew P Harrigan, Chaya D Stern, et al. Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*, 13(7):e1005659, 2017.

[Gallier and Xu, 2003] Jean Gallier and Dianna Xu. Computing exponentials of skew-symmetric matrices and logarithms of orthogonal matrices. *International Journal of Robotics and Automation*, 18(1):10–20, 2003.

[Garcia Satorras *et al.*, 2021] Victor Garcia Satorras, Emiel Hoogeboom, Fabian Fuchs, Ingmar Posner, and Max Welling. E (n) equivariant normalizing flows. *Advances in Neural Information Processing Systems*, 34:4181–4192, 2021.

[Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[Hoogeboom *et al.*, 2021] Emiel Hoogeboom, Didrik Nielsen, Priyank Jaini, Patrick Forré, and Max Welling. Argmax flows and multinomial diffusion: Learning categorical distributions. *Advances in Neural Information Processing Systems*, 34:12454–12465, 2021.

[Jin *et al.*, 2021] Wengong Jin, Jeremy Wohlwend, Regina Barzilay, and Tommi Jaakkola. Iterative refinement graph neural network for antibody sequence-structure co-design. *arXiv preprint arXiv:2110.04624*, 2021.

[Joubbi *et al.*, 2024] Sara Joubbi, Alessio Micheli, Paolo Milazzo, Giuseppe Maccari, Giorgio Ciano, Dario Cardamone, and Duccio Medini. Antibody design using deep learning: from sequence and structure design to affinity maturation. *Briefings in Bioinformatics*, 25(4):bbae307, 2024.

[Jumper *et al.*, 2021] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.

[Kong *et al.*, 2022] Xiangzhe Kong, Wenbing Huang, and Yang Liu. Conditional antibody design as 3d equivariant graph translation. *arXiv preprint arXiv:2208.06073*, 2022.

[Kulytė *et al.*, 2024] Paulina Kulytė, Francisco Vargas, Simon Valentin Mathis, Yu Guang Wang, José Miguel Hernández-Lobato, and Pietro Liò. Improving antibody design with force-guided sampling in diffusion models. *arXiv preprint arXiv:2406.05832*, 2024.

[Leach *et al.*, 2022] Adam Leach, Sebastian M Schmon, Matteo T Degiacomi, and Chris G Willcocks. Denoising diffusion probabilistic models on so (3) for rotational alignment. 2022.

[Li *et al.*, 2024a] Kun Li, Xiuwen Gong, Jia Wu, and Wenbin Hu. Contrastive learning drug response models from natural language supervision. In Kate Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 2126–2134. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Main Track.

[Li *et al.*, 2024b] Kun Li, Weiwei Liu, Yong Luo, Xiantao Cai, Jia Wu, and Wenbin Hu. Zero-shot learning for preclinical drug screening. In Kate Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 2117–2125. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Main Track.

[Lin *et al.*, 2024] Haitao Lin, Lirong Wu, Yufei Huang, Yun-fan Liu, Odin Zhang, Yuanqing Zhou, Rui Sun, and Stan Z. Li. GeoAB: Towards realistic antibody design and reliable affinity maturation. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 30346–30361. PMLR, 21–27 Jul 2024.

[Lindorff-Larsen *et al.*, 2010] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L Klepeis, Ron O Dror, and David E Shaw. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950–1958, 2010.

[Luo *et al.*, 2022] Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific antibody design and optimization with diffusion-based generative models for protein structures. *Advances in Neural Information Processing Systems*, 35:9754–9767, 2022.

[Martinkus *et al.*, 2024] Karolis Martinkus, Jan Ludwiczak, Wei-Ching Liang, Julien Lafrance-Vanasse, Isidro Hotzel, Arvind Rajpal, Yan Wu, Kyunghyun Cho, Richard Bonneau, Vladimir Gligorijevic, et al. Abdiffuser: full-atom generation of in-vitro functioning antibodies. *Advances in Neural Information Processing Systems*, 36, 2024.

[Matthies *et al.*, 1988] Siegfried Matthies, J Muller, and GW Vinel. On the normal distribution in the orientation space. *Texture, Stress, and Microstructure*, 10(1):77–96, 1988.

[Nikolayev and Savyolov, 1997] Dmitry I Nikolayev and Tatjana I Savyolov. Normal distribution on the rotation group so (3). *Texture, Stress, and Microstructure*, 29(3-4):201–233, 1997.

[North *et al.*, 2011] Benjamin North, Andreas Lehmann, and Roland L Dunbrack Jr. A new clustering of antibody cdr loop conformations. *Journal of molecular biology*, 406(2):228–256, 2011.

[Ruffolo *et al.*, 2021] Jeffrey A Ruffolo, Jeffrey J Gray, and Jeremias Sulam. Deciphering antibody affinity maturation with language models and weakly supervised learning. *arXiv preprint arXiv:2112.07782*, 2021.

[Ruffolo *et al.*, 2023] Jeffrey A Ruffolo, Lee-Shin Chu, Sai Pooja Mahajan, and Jeffrey J Gray. Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Nature communications*, 14(1):2389, 2023.

[Schütt *et al.*, 2021] Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pages 9377–9388. PMLR, 2021.

[Shen *et al.*, 2023] Cong Shen, Jiawei Luo, and Kelin Xia. Molecular geometric deep learning. *Cell reports methods*, 3(11), 2023.

[Shin *et al.*, 2021] Jung-Eun Shin, Adam J Riesselman, Aaron W Kollasch, Conor McMahon, Elana Simon, Chris Sander, Aashish Manglik, Andrew C Kruse, and Debora S Marks. Protein design and variant prediction using autoregressive generative models. *Nature communications*, 12(1):2403, 2021.

[Sohl-Dickstein *et al.*, 2015] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pages 2256–2265. PMLR, 2015.

[Song and Ermon, 2019] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems*, 32, 2019.

[Thölke and De Fabritiis, 2022] Philipp Thölke and Gianni De Fabritiis. Torchmd-net: equivariant transformers for neural network based molecular potentials. *arXiv preprint arXiv:2202.02541*, 2022.

[Wang *et al.*, 2024] Yusong Wang, Tong Wang, Shaoning Li, Xinheng He, Mingyu Li, Zun Wang, Nanning Zheng, Bin Shao, and Tie-Yan Liu. Enhancing geometric representations for molecules with equivariant vector-scalar interactive message passing. *Nature Communications*, 15(1):313, 2024.

[Zhu *et al.*, 2024] Tian Zhu, Milong Ren, and Haicang Zhang. Antibody design using a score-based diffusion model guided by evolutionary, physical and geometric constraints. In *Forty-first International Conference on Machine Learning*, 2024.