# Cross-Gate MLP with Protein Complex Invariant Embedding Is a One-Shot Antibody Designer

**Cheng Tan[1,2*], Zhangyang Gao[1,2*], Lirong Wu[1,2], Jun Xia[1,2], Jiangbin Zheng[1,2], Xihong Yang[3], Yue Liu[3], Bozhen Hu[1,2], Stan Z. Li[2†]**

[1]Zhejiang University
[2]AI Lab, Research Center for Industries of the Future, Westlake University
[3]College of Computer, National University of Defense Technology
tancheng@westlake.edu.cn

## Abstract

Antibodies are crucial proteins produced by the immune system in response to foreign substances or antigens. The specificity of an antibody is determined by its complementarity-determining regions (CDRs), which are located in the variable domains of the antibody chains and form the antigen-binding site. Previous studies have utilized complex techniques to generate CDRs, but they suffer from inadequate geometric modeling. Moreover, the common iterative refinement strategies lead to an inefficient inference. In this paper, we propose a *simple yet effective* model that can co-design 1D sequences and 3D structures of CDRs in a one-shot manner. To achieve this, we decouple the antibody CDR design problem into two stages: (i) geometric modeling of protein complex structures and (ii) sequence-structure co-learning. We develop a novel macromolecular structure invariant embedding, typically for protein complexes, that captures both intra- and inter-component interactions among the backbone atoms, including C$\alpha$, N, C, and O atoms, to achieve comprehensive geometric modeling. Then, we introduce a simple cross-gate MLP for sequence-structure co-learning, allowing sequence and structure representations to implicitly refine each other. This enables our model to design desired sequences and structures in a one-shot manner. Extensive experiments are conducted to evaluate our results at both the sequence and structure levels, which demonstrate that our model achieves superior performance compared to the state-of-the-art antibody CDR design methods.

## Introduction

Antibodies are essential proteins that the immune system makes to fight foreign substances, or antigens (Raybould et al. 2019; Kong, Huang, and Liu 2023b; Shi et al. 2022). They have a Y-shape with two arms that can attach to specific antigens, such as bacteria. Once an antibody binds to an antigen, it marks the antigen for destruction by other cells in the immune system (Basu et al. 2019). The ability of antibodies to recognize and bind to antigens is crucial for the immune system's defense against infections and protection against future exposure to the same antigen (Maynard and Georgiou 2000; Akbar et al. 2022a). The complementarity-determining regions (CDRs) are parts of the variable domains of the antibody chains that vary and form the antigen-binding site that defines the specificity of the antibody (Kuroda et al. 2012). CDRs play a critical role in the recognition and binding of antigens by antibodies (Tiller and Tessier 2015).

Early works on antibody design focus on generating the sequences of CDRs without the corresponding structures (Saka et al. 2021; Alley et al. 2019; Shin et al. 2021), while a recent work (Jin et al. 2022) proposes a novel approach called RefineGNN, which enables the co-design of both the sequences and structures of antibody CDRs. DiffAb (Luo et al. 2022) proposes generating antibodies with high affinity to given antigen structures. MEAN (Kong, Huang, and Liu 2023a) further involves the light chain context information as a conditional input to generate CDRs. However, as more conditional contexts are introduced, designed models struggle to adequately capture the complex interactions between complementarity-determining regions (CDRs) and context information. This is due to an insufficient geometric modeling approach, which relies solely on considering the C$\alpha$ atoms or orientations in each residue.

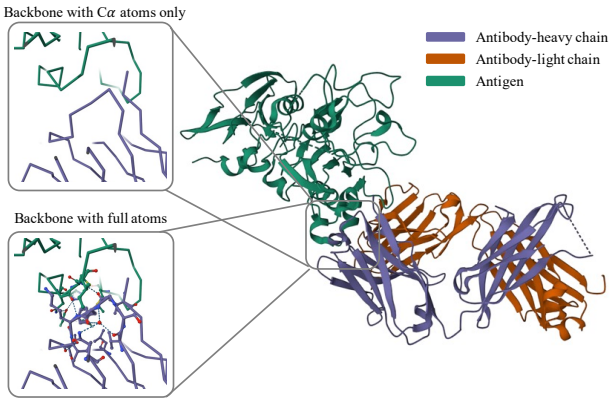



Figure 1: The backbone comprised solely of C$\alpha$ atoms provides a reduced amount of information compared to the backbone consisting of all atoms.

*These authors contributed equally.
†Corresponding author.

As demonstrated in Figure 1, the use of full backbone atoms provides a richer source of geometric information that is critical for the intricate interactions between components of antibody-antigen complexes. While some contemporary studies (Jin et al. 2022; Luo et al. 2022) have endeavored to integrate the orientations of amino acids, their capacity to furnish comprehensive information remains constrained without the utilization of full backbone atoms. Moreover, existing works rely on either iterative refinement (Jin et al. 2022; Fu and Sun 2022; Kong, Huang, and Liu 2023a) or diffusion sampling strategies (Luo et al. 2022) in the CDR decoding process, which leads to inefficient inference.

To address the limitations of existing methods, we propose a novel antibody CDR design model that co-learns the 1D sequences and 3D structures using a protein complex invariant embedding network and can decode the CDRs in a one-shot manner. Specifically, we decouple the antibody CDR design into a two-stage process: (i) **geometric modeling of protein structures** and (ii) **sequence-structure co-learning**. For the comprehensive geometric modeling, the protein complex invariant embedding constructs intra-component relationships inside the same component and inter-component interactions between different components of antigen-antibody complexes with full backbone atoms. Our approach explicitly models complete atomic-level geometric dependencies, which include not only C$\alpha$ atoms but also N, C, and O atoms in the protein backbone. This enables our model to capture the intricate dependencies between the CDRs and the contexts. Then, we introduce a simple cross-gate MLP to implicitly refine the sequence and structure representations by each other in a co-learning manner. The model thus does not require any explicit iterative refinement strategies that are computationally expensive. We evaluate our approach on three challenging tasks: sequence and structure modeling, antigen-binding CDR design, and binding affinity optimization, and demonstrate superior performance compared to state-of-the-art methods, indicating the effectiveness of our model.

## Related Work

**Protein Design** Several machine learning approaches in structure-based protein design use fragment-based and energy-based global features derived from protein structures (Hu et al. 2022; Chen and Arnold 2020; Wu et al. 2021; Kuhlman and Bradley 2019). A seminal work is (Ingraham et al. 2019)'s introduction of the formulation of fixed-backbone design as a structure-to-sequence translation problem. GVP (Jing et al. 2021) developed typical model architectures with translational and rotational equivariances. GCA (Tan et al. 2023) utilizes global attention to learn geometric representations from residue interactions. AlphaDesign (Gao et al. 2022) has established a benchmark based on AlphaFold DB (Varadi et al. 2022; Jumper et al. 2021). ESM-IF (Hsu et al. 2022) has augmented training data by incorporating predicted structures from AlphaFold2 (Jumper et al. 2021). ProteinMPNN (Dauparas et al. 2022) employs expressive features with message-passing networks similar to those used in (Ingraham et al. 2019)'s model. PiFold (Gao, Tan, and Li 2023b) introduces additional features and generates sequences in a one-shot manner. We focus on antibody design, a specific type of protein design, creating antibodies that bind to a target antigen with high affinity.

**Antibody Design** Early approaches to computational antibody design relied heavily on handcrafted and statistical energy function optimization, utilizing Monte Carlo simulation to iteratively update both sequences and structures (Pantazes and Maranas 2010; Lapidoth et al. 2015; Adolf-Bryfogle et al. 2018; Warszawski et al. 2019; Ruffolo, Gray, and Sulam 2021). However, these methods are often computationally expensive and may only reach a local energy minimum. As an alternative, deep generative models have become a more feasible option. In the early stages of deep generative antibody design, sequence-based methods (Alley et al. 2019; Saka et al. 2021; Shin et al. 2021; Akbar et al. 2022b) were introduced. RefineGNN (Jin et al. 2022) is the first deep generative model of CDR sequence-structure co-design. DiffAb (Luo et al. 2022) generates antibodies explicitly targeting specific antigen structures by utilizing diffusion models. CEM (Fu and Sun 2022) designs a constrained manifold to characterize the geometry constraints of the CDR loops. MEAN (Kong, Huang, and Liu 2023a) applies E(3)-equivariant message passing and attention mechanisms. Our model aims to enhance capturing geometrical correlations between antigens and antibodies by incorporating structural information through a protein complex invariant embedding. Additionally, we developed a model capable of generating both CDR sequences and structures in a one-shot manner.

**Generative Models for Molecules** Autoregressive models have gained popularity for generating graphs (Yang et al. 2023b,a; Wu et al. 2022b) in the context of biological molecules, as evidenced by studies such as GraphRNN (You et al. 2018), LGP-Net (Li et al. 2018), CG-VAE (Liu et al. 2018), and HierVAE (Jin, Barzilay, and Jaakkola 2020). Among these, G-SchNet (Gebauer, Gastegger, and Schütt 2019) generates edges sequentially, while Graphite (Grover, Zweig, and Ermon 2019) iteratively refines the adjacency matrix with given node labels. RefineGNN (Jin et al. 2022) combines autoregressive models with iterative refinement to generate complete graphs with both node and edge labels for antibody design. MEAN (Kong, Huang, and Liu 2023a) utilizes the multi-channel extension (Huang et al. 2022) of the E(n)-equivariant GNN (Satorras, Hoogeboom, and Welling 2021) to generate sequences and structures in a progressive full-shot decoding manner. Recent advances in diffusion models (Song and Ermon 2019; Ho, Jain, and Abbeel 2020; Cao et al. 2022) have motivated their development in molecular generation, and several works (Wu et al. 2022a; Trippe et al. 2022; Yang et al. 2023c; Gao, Tan, and Li 2023a) have successfully employed generative diffusion models in protein structure generation. In particular, DiffAb (Luo et al. 2022) has designed a diffusion probabilistic model specifically for antibody structure generation. While previous works have utilized either equivariant graph neural networks with iterative refinement or diffusion models with iterative sampling, we propose cross-gate MLPs that effectively capture geometric correlations and generate both CDR sequences and structures in a one-shot manner.

# Method

## Preliminaries

A protein complex is comprised of $N$ amino acids, which can be represented as characters in a sequence, denoted as $\mathcal{S} = \{s_i\}_{i=1}^{N}$. Each token $s_i$ in the sequence is referred to as a *residue*, with a value that can be any one of the 20 amino acids. The three-dimensional structure of the protein is represented by the backbone atom coordinates, denoted as $\mathcal{X} = \{\boldsymbol{x}_{i,\omega}\}_{i=1}^{N}$, where $\boldsymbol{x}_{i,\omega} \in \mathbb{R}^3$ and $\omega \in \{\mathrm{C}\alpha, \mathrm{N}, \mathrm{C}, \mathrm{O}\}$. The antibody-antigen complex, a common type of protein complex, can be represented by the pair $\mathcal{C} = (\mathcal{S}, \mathcal{X})$.

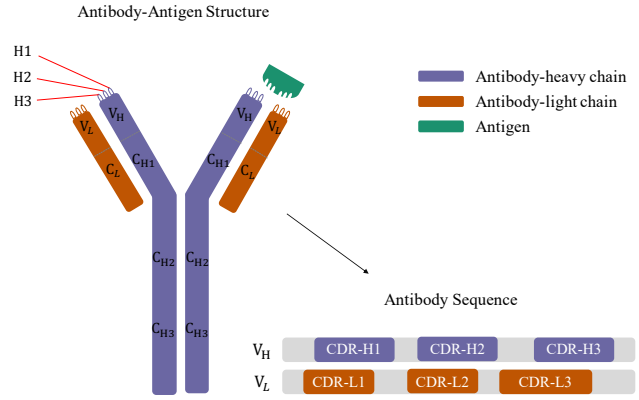



Figure 2: The schematic diagram of an antibody-antigen complex structure. Note that the antibody is a symmetric Y shape, each half of which contains a heavy and light chain. Here we focus on designing the CDR-H1, CDR-H2, and CDR-H3 loops in the heavy chain.

Specifically, an antibody is a protein with a symmetrical Y shape, as depicted in Figure 2. It consists of two identical H/L chains, each of which contains a *variable domain* (VH/VL) and several constant domains. The variable domain can be further divided into a *framework region* (Jin et al. 2022; Kuroda et al. 2012) and three *complementarity-determining regions* (CDRs), which play a crucial role in binding affinity to specific antigens. The heavy and light chains on each half of the antibody contain six CDR loops, namely CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, and CDR-L3, respectively. Prior research has formalized the antibody design problem as identifying CDRs that fit within a given framework region (Shin et al. 2021; Akbar et al. 2022b). More recent studies have indicated that CDRs in heavy chains are the most influential in determining antigen-binding affinity and are therefore the most difficult to design (Jin et al. 2022; Fischman and Ofran 2018). As a result, the context of the antigen and light chains are incorporated for better controlling the binding specificity of the generated antibodies.

The CDR is represented as a set of residues $\mathcal{R}$ given by:

$$\mathcal{R} := \{(s_i, \boldsymbol{x}_{i,\omega}) | i = \{p+1, ..., p+q\}\}, \quad (1)$$

The remainder context is represented as:

$$\mathcal{C} \backslash \mathcal{R} := \{(s_i, \boldsymbol{x}_{i,\omega}) | i = \{1, ..., N\} \backslash \{p+1, ..., p+q\}\}. \quad (2)$$

where $\mathcal{C}$ is the full antibody-antigen complex, and $\mathcal{C} \setminus \mathcal{R}$ is the context information that excludes the CDR residues. Our objective is to learn a mapping $\mathcal{F}_\Theta : \mathcal{C} \setminus \mathcal{R} \mapsto \mathcal{C}$, parameterized by $\Theta$, that generates the sequence and structure of a CDR consisting of $q$ amino acids with indices $p+1$ to $p+q$. The optimal parameters $\Theta^*$ of $\mathcal{F}_\Theta$ is obtained by:

$$\Theta^* = \arg\min_{\Theta} \mathcal{L}(\mathcal{F}_\Theta(\mathcal{C} \setminus \mathcal{R}), \mathcal{C}), \quad (3)$$

where $\mathcal{L}$ is a loss function that measures the difference between the generated and real antibody-antigen complex.

## Overview

The overall framework is depicted in Figure 3, named as **ADesigner**. The input to the model is the structure of the antibody-antigen complex, which is processed by a protein complex invariant embedding (PIE) module to obtain two sets of embeddings: one for intra-component interactions and the other for inter-component interactions. The PIE module is explained in detail in the following subsection.

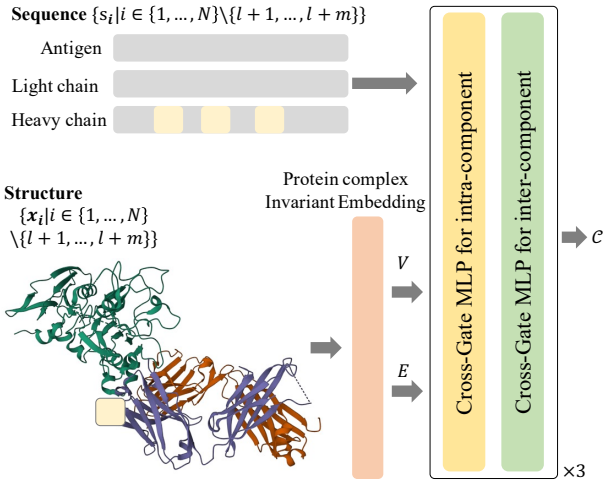



Figure 3: The overall framework of our model. The input is both the sequence and structure of the antibody-antigen complex. The CDRs are visually masked by light yellow blocks to highlight their generation by the model. The output consists of the complete sequence and structure of the antibody-antigen complex, including the generated CDRs.

The embeddings undergo processing through a sequence of cross-gate MLP modules, which take into account both the sequence and structure information. In each block of cross-gate MLP modules, there are two types of modules for processing intra- and inter-component interactions separately. The cross-gate MLP modules facilitate sequence-structure co-learning, allowing the representations to be refined and enriched in an implicit manner.

Ultimately, the learned embeddings are harnessed to generate the complete sequence and structure of the antibody-antigen complex. In contrast to prior methods that depend on explicit iterative decoding strategies, our framework directly outputs the generated result. This is made possible by the co-learning of sequence and structure information in the cross-

gate MLP modules. Overall, our framework provides a more efficient and effective approach to generating the antibody-antigen complex.

## Protein Complex Invariant Embedding

In order to obtain a comprehensive geometric model of the antibody-antigen complex, we introduce the concept of the *Protein complex Invariant Embedding* (PIE). Following previous works (Jin et al. 2022; Kong, Huang, and Liu 2023a), we represent the protein structure as a graph, where $S_i$ denotes the component of residue $i$. We define intra-component and inter-component edges as follows:

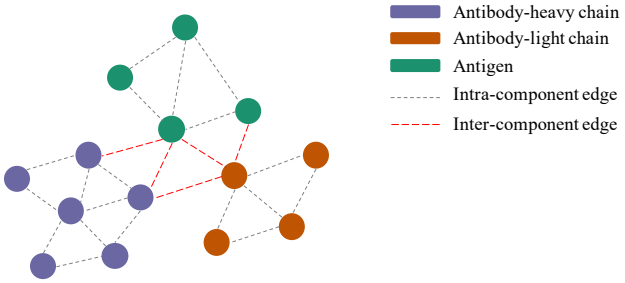



Figure 4: The schematic diagram of intra- and inter-component edges in the antibody-antigen complex.

**Definition 1.** An *intra-component edge* is defined as the edge between two residues in the same component if the distance between their C$\alpha$ atoms is less than a threshold $\delta_{in}$. For residue $i$, we denote the set of its intra-component edges as $\mathcal{E}_{in}(i) = \{j \mid \|\boldsymbol{x}_{i,\mathrm{C}\alpha} - \boldsymbol{x}_{j,\mathrm{C}\alpha}\|^2 < \delta_{in}, \forall S_i = S_j\}$.

**Definition 2.** An *inter-component edge* is defined as the edge between two residues in different components if the distance between their C$\alpha$ atoms is less than a threshold $\delta_{ex}$. For residue $i$, we denote the set of its inter-component edges as $\mathcal{E}_{ex}(i) = \{j \mid \|\boldsymbol{x}_{i,\mathrm{C}\alpha} - \boldsymbol{x}_{j,\mathrm{C}\alpha}\|^2 < \delta_{ex}, \forall S_i \neq S_j\}$.

The intra- and inter-component edges are defined for general protein complexes. In the case of our antibody-antigen complex, there are three components: the heavy chain, the light chain, and the antigen. The schematic diagram of intra- and inter-component edges is shown in Figure 4. Empirically, we set the thresholds $\delta_{in} = 8.0$Å and $\delta_{ex} = 12.0$Å.

Given that the intra- and inter-component edges have captured the component-level interactions, our focus now turns to the residue-level dependencies. We achieve this by transforming the protein complex structure coordinates into a graph. For each residue $i$, we define its node embedding as the distance encoding of its C$\alpha$ atom to the remaining backbone atoms:

$$V_i = \left\{ \mathrm{RBF}(\|\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{i,\gamma}\|) \;\middle|\; \omega, \gamma \in \{\mathrm{Ca}, \mathrm{N}, \mathrm{C}, \mathrm{O}\} \right\}, \quad (4)$$

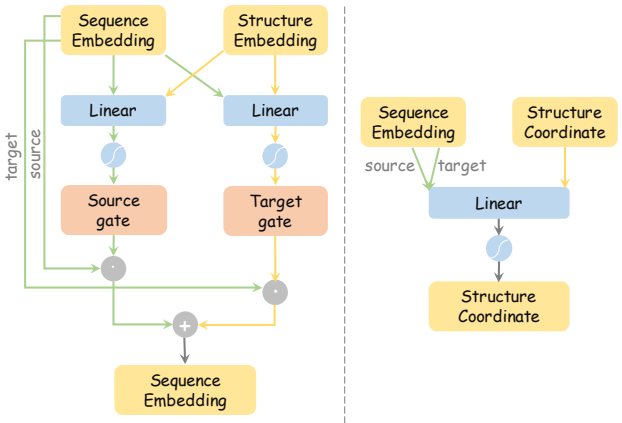
where $\mathrm{RBF}(\cdot)$ is a radial basis distance encoding function. Analogously, we define edge embedding as the distance encoding of pairwise backbone atoms in the neighboring residue. We also encode the directions to identify the relative positions between neighboring residues. Formally, the intra- and inter-component edge embeddings are as follows:

$$\begin{aligned} E_i^{in} = \Bigg\{ & \mathrm{RBF}(\|\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{j,\gamma}\|),\ \boldsymbol{Q}_i^T \frac{\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{j,\gamma}}{\|\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{j,\gamma}\|} \;\Bigg| \\ & \omega, \gamma \in \{\mathrm{Ca}, \mathrm{N}, \mathrm{C}, \mathrm{O}\}, j \in \mathcal{E}_{in} \Bigg\}, \\ E_i^{out} = \Bigg\{ & \mathrm{RBF}(\|\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{j,\gamma}\|),\ \boldsymbol{Q}_i^T \frac{\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{j,\gamma}}{\|\boldsymbol{x}_{i,\omega} - \boldsymbol{x}_{j,\gamma}\|} \;\Bigg| \\ & \omega, \gamma \in \{\mathrm{Ca}, \mathrm{N}, \mathrm{C}, \mathrm{O}\}, j \in \mathcal{E}_{out} \Bigg\}, \end{aligned} \quad (5)$$

where $\boldsymbol{Q}_i$ is a local coordinate system (Ingraham et al. 2019) of residue $i$.

## Cross-Gate MLP

To improve the efficiency of protein complex sequence-structure co-learning, we propose a novel Cross-Gate MLP (CGMLP) that updates sequence embeddings by incorporating both sequence and structure embeddings.




(a) sequence learning with Cross-Gate MLP

(b) structure learning with MLP

Figure 5: The schematic diagram of the sequence learning with Cross-Gate MLP and structure learning with MLP.

We define $\boldsymbol{h}_i^{(l)}$ as the source sequence embedding of residue $i$ in layer $l$, and its target sequence embedding $\boldsymbol{h}_j^{(l)}$ in layer $l$ is from its neighboring residue $j$. The coordinates of residue $i$ in layer $l$ are denoted as $\boldsymbol{Z}_i^{(l)} \in \mathbb{R}^{4\times 3}$, which contains the four types of backbone atoms $\{\boldsymbol{x}_{i,\omega} | \omega \in \{Ca, N, C, O\}\}$. The CGMLP is defined as follows:

$$m_{ij}^{(l)} = \mathrm{Concat}(\boldsymbol{h}_i^{(l)}, \boldsymbol{h}_j^{(l)}, V_i, V_j, E_i), \quad (6)$$

$$g_s^{(l)} = \sigma(\phi_s(m_{ij}^{(l)})), g_t^{(l)} = \sigma(\phi_t(m_{ij}^{(l)})), \quad (7)$$

$$m_{ij}^{(l+1)} = g_s^{(l)} \odot \boldsymbol{h}_j^{(l)} + g_t^{(l)} \odot \boldsymbol{h}_i^{(l)}, \quad (8)$$

$$\boldsymbol{h}_i^{(l+1)} = \phi_h(\boldsymbol{h}^{(l)}, \sum_{j\in\mathcal{E}} m_{ij}^{(l+1)}), \quad (9)$$

where $\phi_s(\cdot)$, $\phi_t(\cdot)$, and $\phi_h(\cdot)$ are MLPs, $\sigma(\cdot)$ is the Sigmoid activation function, $\odot$ is the element-wise multiplication, and $\mathrm{Concat}(\cdot)$ is the concatenation operation. Here, we

use $\mathcal{E}$ and $E_i$ without the subscript to denote both the set of *intra- and inter-component edges*, and *the edge embedding of residue* $i$, respectively, for convenience.

As illustrated in Figure 5(a), we utilize both the sequence and structure embeddings to obtain the latent message $m_{ij}^l$. The source and target gates are obtained by using $m_{ij}^l$ with two individual branches of an MLP and a sigmoid activation function. The refined latent message $m_{ij}^{l+1}$ is the sum of the source and target sequence embeddings weighted by the gates. Finally, the sequence embedding $\boldsymbol{h}_i^{(l+1)}$ of residue $i$ in layer $l+1$ is obtained by aggregating the refined latent message of its neighboring residues.

With the refined sequence embedding, we update the coordinates $\boldsymbol{Z}_i^{(l+1)}$ of residue $i$ in layer $l+1$ as follows:

$$m_{ij}^{\prime(l)} = \mathrm{Concat}(\boldsymbol{h}_i^{(l)}, \boldsymbol{h}_j^{(l)}, \frac{(\boldsymbol{Z}_i^{(l)} - \boldsymbol{Z}_j^{(l)})^T(\boldsymbol{Z}_i^{(l)} - \boldsymbol{Z}_j^{(l)})}{\|(\boldsymbol{Z}_i^{(l)} - \boldsymbol{Z}_j^{(l)})^T(\boldsymbol{Z}_i^{(l)} - \boldsymbol{Z}_j^{(l)})\|_F}), \tag{10}$$

$$\boldsymbol{Z}^{(l+1)} = Z^{(l)} + \frac{1}{|\mathcal{E}|}\sum_{j\in\mathcal{E}} \phi_z(m_{ij}^{\prime(l)})(\boldsymbol{Z}_i^{(l)} - \boldsymbol{Z}_j^{(l)}), \tag{11}$$

where, $\phi_z(\cdot)$ is a vanilla MLP. As illustrated in Figure 5(b), the coordinates of residue $i$ in layer $l+1$ are obtained by aggregating the refined latent messages from its neighboring residues. It's worth noting that we only use refined sequence embeddings to update the structure coordinates, rather than structure embeddings. The reason is that structure embeddings are directly influenced by the coordinates, and using them may lead to overfitting of structure learning.

## One-shot Decoding

Our method has been designed to perform sequence-structure co-learning, allowing us to directly output the CDR regions of the antibody-antigen complex without the need for any additional decoding process. Assuming there are $L$ layers, the predicted sequence $\hat{s}_i$ and structure $\hat{\boldsymbol{Z}}_i$ of the CDR regions are obtained as follows:

$$\begin{aligned} \hat{s}_i &= \mathrm{Argmax}(h_i^{(L)}), \\ \hat{\boldsymbol{Z}}_i &= \boldsymbol{Z}_i^{(L)}, \end{aligned} \tag{12}$$

where $\mathrm{Argmax}(\cdot)$ is the argmax operation.

The loss function is defined as a linear combination of the sequence loss and the structure loss. For the sequence loss, we use the cross-entropy loss between the predicted sequence and the ground truth sequence:

$$\mathcal{L}_{seq} = \frac{1}{q}\sum_{i=p+1}^{p+q} \ell_{ce}(s_i, \mathrm{Softmax}(h_i^{(L)})), \tag{13}$$

where $\ell_{ce}$ denotes the cross-entropy loss, and $\mathrm{Softmax}(\cdot)$ is the softmax activation function. For the structure loss, we use the differentiable L1 loss (Lai et al. 2018):

$$\mathcal{L}_{struct} = \frac{1}{q}\sum_{i=p+1}^{p+q} \sqrt{(\boldsymbol{Z}_i - \hat{\boldsymbol{Z}}_i)^2 + \epsilon^2}, \tag{14}$$

where $\epsilon$ is a small constant empirically set to $10^{-8}$. This loss function is more robust to outliers compared to the commonly used L2 loss, as it suppresses large errors using the $\epsilon^2$ term. Consequently, outliers do not have much influence on the total loss, making the network more stable. The overall loss is $\mathcal{L} = \mathcal{L}_{seq} + \lambda\mathcal{L}_{struct}$, where $\lambda = 0.8$ is a weight hyperparameter that balances the sequence and structure loss.

# Experiments

We evaluate our model on three challenging antibody design tasks using the common experimental setups from previous works (Jin et al. 2022; Kong, Huang, and Liu 2023a; Fu and Sun 2022). These tasks include: (i) generative task on the Structural Antibody Database (Dunbar et al. 2014), (ii) antigen-binding CDR-H3 design using an existing antibody design benchmark of 60 antibody-antigen complexes from (Adolf-Bryfogle et al. 2018), and (iii) antigen-antibody binding affinity optimization that redesigns CDR-H3 of antibodies on the Structural Kinetic and Energetic database of Mutant Protein Interactions (Jankauskaitė et al. 2019).

**Baselines** We compare our model to recent state-of-the-art approaches, including (i) **LSTM**-based approach by (Saka et al. 2021; Akbar et al. 2022b) that generates the amino acid sequence in an autoregressive manner without structure modeling, leveraging a long short-term memory (LSTM) network; (ii) **C-LSTM** implemented by (Kong, Huang, and Liu 2023a) that considers the entire context of the antibody-antigen complex, built upon LSTM; (iii) **RefineGNN** proposed by (Jin et al. 2022) that takes the 3D geometry for antibody CDR design. This approach unravels the amino acid sequence in an autoregressive manner and iteratively refines its predicted global structure. (iv) **C-RefineGNN** implemented by (Kong, Huang, and Liu 2023a) that extends RefineGNN by accommodating the entire antibody-antigen complex. (v) **MEAN** proposed by (Kong, Huang, and Liu 2023a), which is related to but distinct from our method. It takes less geometric information into account and requires an iterative refinement strategy. We used the default setup of each method, training the models for 20 epochs with Adam optimizer and a learning rate of $10^{-3}$. We used the checkpoint with the lowest validation loss for testing.

**Metrics** We evaluate the results from two perspectives, i.e., sequence modeling, and structure modeling. For sequence modeling, we employ Amino Acid Recovery (AAR) that measures the overlapping rate between the predicted sequences and ground truths. For structure modeling, we employ Root Mean Squared Deviation (RMSD) between the predicted structures and ground truths. We report the TM-score (Zhang and Skolnick 2004; Xu and Zhang 2010) that measures the global structural similarity in the second task.

## Sequence and Structure Modeling

We evaluate our model with baseline approaches on the Structural Antibody Database (SAbDab) (Dunbar et al. 2014), which contains 3,127 complexes consisting of heavy chains, light chains, and antigens. Following (Jin et al. 2022; Kong, Huang, and Liu 2023a), the dataset is split into training, validation, and testing sets according to the clustering

| Method | CDR-H1 | | CDR-H2 | | CDR-H3 | |
|---|---|---|---|---|---|---|
| | AAR (↑) | RMSD (↓) | AAR (↑) | RMSD(↓) | AAR (↑) | RMSD(↓) |
| LSTM | 49.98±5.20% | - | 28.50±1.55% | - | 15.69±0.91% | - |
| C-LSTM | 40.93±5.41% | - | 29.24±1.08% | - | 15.48±1.17% | - |
| RefineGNN | 39.40±5.56% | 3.22±0.29 | 37.06±3.09% | 3.64±0.40 | 21.13±1.59% | 6.00±0.55 |
| C-RefineGNN | 33.19±2.99% | 3.25±0.40 | 33.53±3.23% | 3.69±0.56 | 18.88±1.37% | 6.22±0.59 |
| DiffAB | 61.34±1.98% | 1.02±0.66 | 37.66±1.89% | 1.20±0.09 | 25.79±1.52% | 3.02±0.11 |
| MEAN | 58.29±7.26% | 0.98±0.16 | 47.15±3.09% | 0.95±0.05 | 36.38±3.08% | 2.21±0.16 |
| ADesigner | **64.34**±3.37% | **0.82**±0.12 | **55.52**±3.36% | **0.79**±0.06 | **37.37**±2.33% | **1.97**±0.19 |
| Improvement | +6.05% | +16.33% | +8.37% | +16.84% | +0.98% | +10.86% |

Table 1: The mean (standard deviation) of 10-fold cross-validation results for 1D sequence and 3D structure modeling on the SAbDab dataset.

of CDRs to maintain the generalization test. The total numbers of clusters for CDR-H1, CDR-H2, and CDR-H3 are 765, 1093, and 1659, respectively. The clusters are split into training, validation, and testing sets with a ratio of 8:1:1. We report the results of 10-fold cross-validation in Table 1.

It can be observed that our proposed method surpasses all other methods in terms of AAR and RMSD scores for all three CDR regions. Furthermore, our proposed method outperforms MEAN by a significant margin, with an average improvement of over 5.13% in AAR and over 14.68% in RMSD. These results demonstrate the efficacy of our proposed approach in modeling both the sequence and structure of CDRs, making it a promising method for the sequence and structure modeling of antibody-antigen complexes.

## Antigen-Binding CDR-H3 Design

We validated our approach for designing CDR-H3 loops with desired antigen-binding capabilities using the well-established RAbD benchmark dataset (Adolf-Bryfogle et al. 2018). For a comprehensive comparison, the widely-adopted conventional method, RosettaAD (Adolf-Bryfogle et al. 2018), is also incorporated as a benchmark. Rigorous training was undertaken using the extensive SAbDab database of antibody-antigen complexes, carefully excluding any entries bearing significant structural homology to complexes present in the RAbD test set. A detailed analysis of the results is provided in Table 2.

| Method | CDR-H3 | | |
|---|---|---|---|
| | AAR (↑) | TM-score (↑) | RMSD (↓) |
| RosettaAD | 22.50% | 0.9435 | 5.52 |
| LSTM | 22.36% | - | - |
| C-LSTM | 22.18% | - | - |
| RefineGNN | 29.79% | 0.8308 | 7.55 |
| C-RefineGNN | 28.90% | 0.8317 | 7.21 |
| MEAN | 36.77% | 0.9812 | 1.81 |
| ADesigner | **40.94%** | **0.9850** | **1.55** |

Table 2: The performance of CDR-H3 design on the RAbD benchmark using amino acid recovery (AAR), TM-score, and RMSD metrics.

The results clearly demonstrate the superior performance of our method compared to all other techniques, as evidenced by the improved accuracy across AAR, TM-score, and RMSD metrics. This highlights the efficacy of our approach for designing CDR-H3 loops that closely recapitulate native antigen-binding topologies.

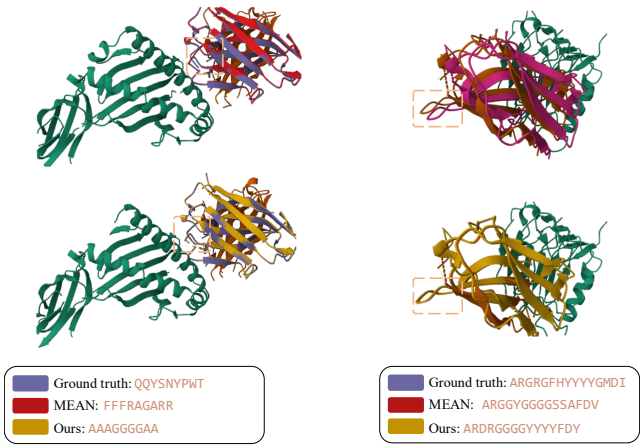



Figure 6: The designed examples of CDR-H3.

To provide further insights, we visually examine the designed CDR-H3 structures for two representative antibody test cases (PDB IDs: 1w72 and 3h3b) in Figure 6. For the first example, the loops designed by MEAN and our method yield RMSDs of 0.27Å and 0.19Å respectively, compared to the native structure. While MEAN performs reasonably well, it still contains inaccuracies in some details. In contrast, our approach predicts a structure that closely matches the native one. In the second more challenging example, the CDR-H3 loops designed by MEAN and our approach have RMSDs of 2.46Å and 1.19Å respectively. MEAN fails on this difficult sample, whereas our method predicts a structure that approximates the correct native conformation.

| Method | CDR-H1 | | CDR-H2 | | CDR-H3 | |
|---|---|---|---|---|---|---|
| | AAR (↑) | RMSD (↓) | AAR (↑) | RMSD(↓) | AAR (↑) | RMSD(↓) |
| ADesigner | 64.34±3.37% | 0.82±0.12 | 55.52±3.36% | 0.79±0.06 | 37.37±2.33% | 1.97±0.19 |
| w/o PIE | 60.27±6.69% | 0.95±0.16 | 49.14±2.96% | 0.98±0.23 | 35.78±2.43% | 2.17±0.21 |
| w/o CGMLP | 62.59±5.09% | 1.07±0.17 | 52.50±3.78% | 0.97±0.18 | 36.14±2.56% | 2.00±0.21 |

Table 3: Ablation of our proposed method on the SAbDab dataset.

## Affinity Optimization

We thoroughly evaluated the efficacy of our methodology for optimizing the binding affinity of antibody-antigen complexes through the simultaneous sequence and conformational tuning of the crucial CDR-H3 loop region. To predict the binding energy ($\Delta\Delta$G) after optimization, we utilized the pre-trained deep geometric network (Shan et al. 2022) and followed the same protocol as in a previous study (Kong, Huang, and Liu 2023a). We incorporated Iterative Target Augmentation (Yang et al. 2020) (ITA) into the optimization process. The results are presented in Table 4.

| Method | $\Delta\Delta$G (↓) |
|---|---|
| Random | +1.52 |
| LSTM | -1.48 |
| C-LSTM | -1.83 |
| RefineGNN | -3.98 |
| C-RefineGNN | -3.79 |
| MEAN | -5.33 |
| ADesigner | **-10.78** |

Table 4: Average affinity change after optimization. The lower is better.

Our proposed method achieved superior results to the previous state-of-the-art method MEAN, with a notably lower $\Delta\Delta$G of below $-10\text{kcal/mol}$, indicating a substantial binding affinity between the optimized antibody and the antigen. We visualize two optimized examples (PDB IDs: 1kip, $\Delta\Delta G = -10.60$; 4jpk, $\Delta\Delta G = -12.09$) in Figure 7.

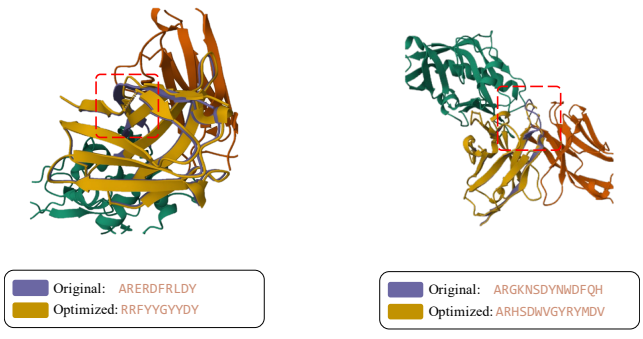



Figure 7: The optimized examples of CDR-H3.

## Ablation Study

We conducted ablation studies as summarized in Table 3. Specifically, we examined the impact of removing the protein complex invariant embedding (w/o PIE) and replacing the cross-gate MLP with equivariant graph neural networks (Satorras, Hoogeboom, and Welling 2021; Kong, Huang, and Liu 2023a) (w/o CGMLP). Our results show that PIE provides rich information that plays a critical role in our model. Furthermore, the CGMLP consistently improved the performance across AAR and RMSD metrics. These findings demonstrate the effectiveness of our approach.

## Training/Inference Efficiency

We conducted a comparison of the training and inference efficiency. Training efficiency was the training time of one full epoch on the SAbDab training set, while inference efficiency was the inference time on the SAbDab testing set, including postprocessing steps that may result in more time overhead. Table 5 shows that our method outperforms RefineGNN and MEAN in both training and inference efficiency. The training efficiency of our method is 16s, which is 27.27% faster than MEAN. Meanwhile, our method's inference efficiency is 24s, which is 14.29% faster than RefineGNN and 14.29% faster than MEAN. These results demonstrate that our one-shot design is highly efficient and effective in optimizing antibody sequence and structure design.

| Method | Training efficiency (↓) | Inference efficiency (↓) |
|---|---|---|
| RefineGNN | 218s | 47s |
| MEAN | 22s | 28s |
| ADesigner | 14s | 24s |
| Improvement | +36.36% | +14.29% |

Table 5: The training and inference efficiency comparison.

## Conclusions and Limitations

In this paper, we develop a simple yet effective antibody designer for antibody sequence and structure design based on the entire context of the antibody-antigen complex. By leveraging comprehensive geometric modeling with a novel macromolecular invariant embedding tailored for protein complexes, and enabling sequence-structure co-learning through a simple cross-gate MLP, our approach achieves competitive results on various antibody-related tasks. A limitation is that our method is currently limited to in silico design; we leave wet-lab validation to future work.

## Acknowledgements

This work was supported by the National Key R&D Program of China (2022ZD0115100), the National Natural Science Foundation of China (U21A20427), the Competitive Research Fund (WU2022A009) from the Westlake Center for Synthetic Biology and Integrated Bioengineering.

## References

Adolf-Bryfogle, J.; Kalyuzhniy, O.; Kubitz, M.; Weitzner, B. D.; Hu, X.; Adachi, Y.; Schief, W. R.; and Dunbrack Jr, R. L. 2018. RosettaAntibodyDesign (RAbD): A general framework for computational antibody design. *PLoS computational biology*, 14(4): e1006112.

Akbar, R.; Bashour, H.; Rawat, P.; Robert, P. A.; Smorodina, E.; Cotet, T.-S.; Flem-Karlsen, K.; Frank, R.; Mehta, B. B.; Vu, M. H.; et al. 2022a. Progress and challenges for the machine learning-based design of fit-for-purpose monoclonal antibodies. In *MAbs*, volume 14, 2008790. Taylor & Francis.

Akbar, R.; Robert, P. A.; Weber, C. R.; Widrich, M.; Frank, R.; Pavlović, M.; Scheffer, L.; Chernigovskaya, M.; Snapkov, I.; Slabodkin, A.; et al. 2022b. In silico proof of principle of machine learning-based antibody design at unconstrained scale. In *MAbs*, volume 14, 2031482. Taylor & Francis.

Alley, E. C.; Khimulya, G.; Biswas, S.; AlQuraishi, M.; and Church, G. M. 2019. Unified rational protein engineering with sequence-based deep representation learning. *Nature methods*, 16(12): 1315–1322.

Basu, K.; Green, E. M.; Cheng, Y.; and Craik, C. S. 2019. Why recombinant antibodies—benefits and applications. *Current opinion in biotechnology*, 60: 153–158.

Cao, H.; Tan, C.; Gao, Z.; Chen, G.; Heng, P.-A.; and Li, S. Z. 2022. A survey on generative diffusion model. *arXiv preprint arXiv:2209.02646*.

Chen, K.; and Arnold, F. H. 2020. Engineering new catalytic activities in enzymes. *Nature Catalysis*, 3(3): 203–213.

Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I.; Courbet, A.; de Haas, R. J.; Bethel, N.; et al. 2022. Robust deep learning–based protein sequence design using ProteinMPNN. *Science*, 378(6615): 49–56.

Dunbar, J.; Krawczyk, K.; Leem, J.; Baker, T.; Fuchs, A.; Georges, G.; Shi, J.; and Deane, C. M. 2014. SAbDab: the structural antibody database. *Nucleic acids research*, 42(D1): D1140–D1146.

Fischman, S.; and Ofran, Y. 2018. Computational design of antibodies. *Current opinion in structural biology*, 51: 156–162.

Fu, T.; and Sun, J. 2022. Antibody complementarity determining regions (cdrs) design using constrained energy model. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 389–399.

Gao, Z.; Tan, C.; Li, S.; et al. 2022. AlphaDesign: A graph protein design method and benchmark on AlphaFoldDB. *arXiv preprint arXiv:2202.01079*.

Gao, Z.; Tan, C.; and Li, S. Z. 2023a. DiffSDS: A language diffusion model for protein backbone inpainting under geometric conditions and constraints. *arXiv preprint arXiv:2301.09642*.

Gao, Z.; Tan, C.; and Li, S. Z. 2023b. PiFold: Toward effective and efficient protein inverse folding. In *ICLR*.

Gebauer, N.; Gastegger, M.; and Schütt, K. 2019. Symmetry-adapted generation of 3d point sets for the targeted discovery of molecules. *NeurIPS*, 32.

Grover, A.; Zweig, A.; and Ermon, S. 2019. Graphite: Iterative generative modeling of graphs. In *ICML*, 2434–2444. PMLR.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *NeurIPS*, 33: 6840–6851.

Hsu, C.; Verkuil, R.; Liu, J.; Lin, Z.; Hie, B.; Sercu, T.; Lerer, A.; and Rives, A. 2022. Learning inverse folding from millions of predicted structures. In *ICML*, 8946–8970. PMLR.

Hu, B.; Xia, J.; Zheng, J.; Tan, C.; Huang, Y.; Xu, Y.; and Li, S. Z. 2022. Protein Language Models and Structure Prediction: Connection and Progression. arXiv:2211.16742.

Huang, W.; Han, J.; Rong, Y.; Xu, T.; Sun, F.; and Huang, J. 2022. Equivariant Graph Mechanics Networks with Constraints. In *ICLR*.

Ingraham, J.; Garg, V.; Barzilay, R.; and Jaakkola, T. 2019. Generative models for graph-based protein design. *NeurIPS*, 32.

Jankauskaitė, J.; Jiménez-García, B.; Dapkūnas, J.; Fernández-Recio, J.; and Moal, I. H. 2019. SKEMPI 2.0: an updated benchmark of changes in protein–protein binding energy, kinetics and thermodynamics upon mutation. *Bioinformatics*, 35(3): 462–469.

Jin, W.; Barzilay, R.; and Jaakkola, T. 2020. Hierarchical generation of molecular graphs using structural motifs. In *ICML*, 4839–4848. PMLR.

Jin, W.; Wohlwend, J.; Barzilay, R.; and Jaakkola, T. S. 2022. Iterative Refinement Graph Neural Network for Antibody Sequence-Structure Co-design. In *ICLR*.

Jing, B.; Eismann, S.; Suriana, P.; Townshend, R. J. L.; and Dror, R. 2021. Learning from Protein Structure with Geometric Vector Perceptrons. In *ICLR*.

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

Kong, X.; Huang, W.; and Liu, Y. 2023a. Conditional Antibody Design as 3D Equivariant Graph Translation. In *ICLR*.

Kong, X.; Huang, W.; and Liu, Y. 2023b. End-to-End Full-Atom Antibody Design. *arXiv preprint arXiv:2302.00203*.

Kuhlman, B.; and Bradley, P. 2019. Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology*, 20(11): 681–697.

Kuroda, D.; Shirai, H.; Jacobson, M. P.; and Nakamura, H. 2012. Computer-aided antibody design. *Protein engineering, design & selection*, 25(10): 507–522.

Lai, W.-S.; Huang, J.-B.; Ahuja, N.; and Yang, M.-H. 2018. Fast and accurate image super-resolution with deep laplacian pyramid networks. *TPAMI*, 41(11): 2599–2613.

Lapidoth, G. D.; Baran, D.; Pszolla, G. M.; Norn, C.; Alon, A.; Tyka, M. D.; and Fleishman, S. J. 2015. Abdesign: A n algorithm for combinatorial backbone design guided by natural conformations and sequences. *Proteins: Structure, Function, and Bioinformatics*, 83(8): 1385–1406.

Li, Y.; Vinyals, O.; Dyer, C.; Pascanu, R.; and Battaglia, P. 2018. Learning deep generative models of graphs. *arXiv preprint arXiv:1803.03324*.

Liu, Q.; Allamanis, M.; Brockschmidt, M.; and Gaunt, A. 2018. Constrained graph variational autoencoders for molecule design. *NeurIPS*, 31.

Luo, S.; Su, Y.; Peng, X.; Wang, S.; Peng, J.; and Ma, J. 2022. Antigen-Specific Antibody Design and Optimization with Diffusion-Based Generative Models for Protein Structures. In *NeurIPS*.

Maynard, J.; and Georgiou, G. 2000. Antibody engineering. *Annual review of biomedical engineering*, 2(1): 339–376.

Pantazes, R.; and Maranas, C. D. 2010. OptCDR: a general computational method for the design of antibody complementarity determining regions for targeted epitope binding. *Protein Engineering, Design & Selection*, 23(11): 849–858.

Raybould, M. I.; Marks, C.; Krawczyk, K.; Taddese, B.; Nowak, J.; Lewis, A. P.; Bujotzek, A.; Shi, J.; and Deane, C. M. 2019. Five computational developability guidelines for therapeutic antibody profiling. *Proceedings of the National Academy of Sciences*, 116(10): 4025–4030.

Ruffolo, J. A.; Gray, J. J.; and Sulam, J. 2021. Deciphering antibody affinity maturation with language models and weakly supervised learning. *arXiv preprint arXiv:2112.07782*.

Saka, K.; Kakuzaki, T.; Metsugi, S.; Kashiwagi, D.; Yoshida, K.; Wada, M.; Tsunoda, H.; and Teramoto, R. 2021. Antibody design using LSTM based deep generative model from phage display library for affinity maturation. *Scientific reports*, 11(1): 1–13.

Satorras, V. G.; Hoogeboom, E.; and Welling, M. 2021. E (n) equivariant graph neural networks. In *ICML*, 9323–9332. PMLR.

Shan, S.; Luo, S.; Yang, Z.; Hong, J.; Su, Y.; Ding, F.; Fu, L.; Li, C.; Chen, P.; Ma, J.; et al. 2022. Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization. *Proceedings of the National Academy of Sciences*, 119(11): e2122954119.

Shi, C.; Wang, C.; Lu, J.; Zhong, B.; and Tang, J. 2022. Protein sequence and structure co-design with equivariant translation. *arXiv preprint arXiv:2210.08761*.

Shin, J.-E.; Riesselman, A. J.; Kollasch, A. W.; McMahon, C.; Simon, E.; Sander, C.; Manglik, A.; Kruse, A. C.; and Marks, D. S. 2021. Protein design and variant prediction using autoregressive generative models. *Nature communications*, 12(1): 2403.

Song, Y.; and Ermon, S. 2019. Generative modeling by estimating gradients of the data distribution. *NeurIPS*, 32.

Tan, C.; Gao, Z.; Xia, J.; Hu, B.; and Li, S. Z. 2023. Generative de novo protein design with global context. In *ICASSP*.

Tiller, K. E.; and Tessier, P. M. 2015. Advances in antibody design. *Annual review of biomedical engineering*, 17: 191–216.

Trippe, B. L.; Yim, J.; Tischer, D.; Broderick, T.; Baker, D.; Barzilay, R.; and Jaakkola, T. 2022. Diffusion probabilistic modeling of protein backbones in 3D for the motif-scaffolding problem. *arXiv preprint arXiv:2206.04119*.

Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*, 50(D1): D439–D444.

Warszawski, S.; Borenstein Katz, A.; Lipsh, R.; Khmelnitsky, L.; Ben Nissan, G.; Javitt, G.; Dym, O.; Unger, T.; Knop, O.; Albeck, S.; et al. 2019. Optimizing antibody affinity and stability by the automated design of the variable light-heavy chain interfaces. *PLoS computational biology*, 15(8): e1007207.

Wu, K. E.; Yang, K. K.; Berg, R. v. d.; Zou, J. Y.; Lu, A. X.; and Amini, A. P. 2022a. Protein structure generation via folding diffusion. *arXiv preprint arXiv:2209.15611*.

Wu, L.; Lin, H.; Tan, C.; Gao, Z.; and Li, S. Z. 2021. Self-supervised learning on graphs: Contrastive, generative, or predictive. *TKDE*.

Wu, L.; Xia, J.; Gao, Z.; Lin, H.; Tan, C.; and Li, S. Z. 2022b. Graphmixup: Improving class-imbalanced node classification by reinforcement mixup and self-supervised context prediction. In *ECML-PKDD*, 519–535. Springer.

Xu, J.; and Zhang, Y. 2010. How significant is a protein structure similarity with TM-score= 0.5? *Bioinformatics*, 26(7): 889–895.

Yang, K.; Jin, W.; Swanson, K.; Barzilay, R.; and Jaakkola, T. 2020. Improving molecular design by stochastic iterative target augmentation. In *ICML*, 10716–10726. PMLR.

Yang, X.; Jin, J.; Wang, S.; Liang, K.; Liu, Y.; Wen, Y.; Liu, S.; Zhou, S.; Liu, X.; and Zhu, E. 2023a. DealMVC: Dual Contrastive Calibration for Multi-view Clustering. In *ACM Multimedia*.

Yang, X.; Liu, Y.; Zhou, S.; Wang, S.; Tu, W.; Zheng, Q.; Liu, X.; Fang, L.; and Zhu, E. 2023b. Cluster-guided Contrastive Graph Clustering Network. In *AAAI*, volume 37, 10834–10842.

Yang, X.; Tan, C.; Liu, Y.; Liang, K.; Wang, S.; Zhou, S.; Xia, J.; Li, S. Z.; Liu, X.; and Zhu, E. 2023c. CONVERT: Contrastive Graph Clustering with Reliable Augmentation. In *ACM Multimedia*.

You, J.; Ying, R.; Ren, X.; Hamilton, W.; and Leskovec, J. 2018. Graphrnn: Generating realistic graphs with deep auto-regressive models. In *ICML*, 5708–5717. PMLR.

Zhang, Y.; and Skolnick, J. 2004. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4): 702–710.