

IGFLOW-LM: DE NOVO ANTIBODY DESIGN VIA JOINT FLOW MATCHING ON SE(3) AND PROTEIN LANGUAGE MODELS PROBABILITY FLOWS

Anonymous authors

Paper under double-blind review

ABSTRACT

In this work, we present IgFlow-LM, a multi-modal deep generative model for de novo antibody design based on a flow-matching framework that integrates protein language models (PLMs). By learning the joint distribution over SE(3)-equivariant structural flows and PLM-derived probabilistic flows, IgFlow-LM enables the coordinated generation of antibody 3D structures and latent embeddings in the PLM space. Experimental results demonstrate that, in unconditional design, IgFlow-LM generates antibody structures that closely resemble naturally occurring antibodies. IgFlow-LM generates antibodies closely resembling naturally observed ones, with backbone dihedral angles exhibiting strong agreement with reference antibody distributions and overall backbone conformations adhering more closely to physical constraints. Furthermore, we benchmark IgFlow-LM against baseline models on two commonly studied conditional CDR design tasks. IgFlow-LM demonstrates superior overall performance compared to baselines and generates CDR sequences with higher diversity.

1 INTRODUCTION

In the field of protein engineering, an important application is the design of immunoglobulins (antibodies) (Shanehsazzadeh et al., 2023; Ruffolo et al., 2024). Immunoglobulins play a central role in assisting the adaptive immune system to recognize and neutralize pathogens (Raybould et al., 2019; Kong et al., 2023a). They are composed of two heavy chains and two light chains, each of which consists of both constant and variable domains (Schroeder Jr & Cavacini, 2010). The variable regions contain highly diverse loops known as complementarity-determining regions (CDRs), which are critical for antigen recognition and binding (Narciso et al., 2011). Among these, the third CDR of the heavy chain (HCDR3) exhibits the greatest sequence variability and is the most actively involved in antigen binding, often serving as the primary determinant of antigen specificity (Tsuchiya & Mizuguchi, 2016). Consequently, antibody design efforts frequently focus on the conditional design of CDRs, with particular emphasis on accurately determining the HCDR3 region.

Generative models have recently gained significant prominence in protein engineering (Ingraham et al., 2019; Kucera et al., 2022; Alford et al., 2017). Given that antibody function is largely determined by its sequence (Whisstock & Lesk, 2003; Robinson, 2015), simultaneous optimization of both sequence and structure during generation is essential to enhance their consistency. Protein language models (PLMs) have been shown to capture a range of biologically relevant features, including structural (Rao et al., 2020; Lin et al., 2023) and functional properties (Nijkamp et al., 2023). By jointly modeling PLM latent space embeddings and antibody structures, generative models can leverage the evolutionary information encoded within PLMs to improve the foldability and functionality of generated sequences, while also enabling more effective exploration of the antibody structural landscape—particularly for tasks involving highly diverse regions such as complementarity-determining region (CDR) loops. This approach offers a dual advantage: It avoids the error propagation associated with traditional fixed backbone sequence design methods (Dreyer et al., 2023; Lu et al., 2024) and avoids the reduction in sequence diversity often caused by joint modeling with discrete modalities (Campbell et al., 2024; Nagaraj et al., 2024).

Motivated by this, we propose a multimodal deep generative model for antibody design, named IgFlow-LM. This model is built upon a conditional flow matching (CFM) framework that enables the joint generation of PLM latent embeddings and antibody structures. Specifically, we extend the flow-matching-based protein design model FrameFlow (Yim et al., 2023a) to simultaneously generate PLM latent embeddings and antibody structures, and train the model on the Structural Antibody Database (SAbDab) (Dunbar et al., 2014). For the choice of PLM, we adopt IgBert (Kenlay et al., 2024), which is specifically trained on the Observed Antibody Space (OAS) (Olsen et al., 2022), as the source of antibody sequence embeddings. IgBert is capable of taking both heavy and light chain sequences as input and producing their corresponding latent space representations. Our focus is on generating novel antibody variable domains, and we evaluate the performance of our model in the following tasks: (1) unconditional generation of paired heavy and light chains; and (2) framework-conditioned loop design of CDRs. In summary, the main contributions of this work are as follows:

- We propose IgFlow-LM, a multi-modal deep generative model for de novo antibody design that integrates PLM latent embeddings within a flow matching framework.
- We modify the architecture of the base FrameFlow model to enable joint generation of antibody structures and PLM latent embeddings, and further extend it to the task of framework-conditioned CDR loop design.
- Experimental results demonstrate that IgFlow-LM generates antibodies that better conform to physical constraints compared to baseline models, showcasing its effectiveness in antibody engineering applications.

2 RELATED WORKS

Diffusion models progressively transform prior samples into meaningful outputs through a step-by-step denoising process (Ho et al., 2020; Vincent, 2011; Sohl-Dickstein et al., 2015), and have been widely applied in image (Ho et al., 2020; Zhang et al., 2024), text (Li et al., 2022; Hoogeboom et al., 2021), and molecular generation (Watson et al., 2023; Lianza et al., 2024; Hoogeboom et al., 2022; Jing et al., 2020). Recently, continuous normalizing flows based on ordinary differential equations (ODEs) (Pooladian et al., 2023; Chen et al., 2018) have emerged as a promising alternative to diffusion models. In particular, CFM (Lipman et al., 2022) directly learns the ODE-based probabilistic path from a prior distribution to a target distribution, avoiding the computationally expensive simulation procedures required by diffusion models (Lipman et al., 2022; Song et al., 2020a; Chen & Lipman, 2023; Ben-Hamu et al., 2022; De Bortoli et al., 2022).

Generative models have demonstrated remarkable performance in protein-related applications (Watson et al., 2023; Yim et al., 2023a; Jain et al., 2022; Khan et al., 2023; Chen et al., 2024; Trippe et al., 2022; Luo et al., 2025). Current antibody design approaches can be broadly categorized into three classes: sequence design (Dreyer et al., 2023; Hie et al., 2024), structure design (Cutting et al., 2025; Bennett et al., 2023), and sequence-structure co-design (Nagaraj et al., 2024; Jin et al., 2021; Kong et al., 2023b). Sequence-structure co-design has become increasingly prevalent in recent antibody modeling efforts. However, existing methods for generating antibody sequence-structure pairs typically operate directly on discrete sequences, an approach that may be suboptimal for preserving structural diversity. Such discretization can lead to premature convergence toward commonly observed sequence-structure motifs, thereby compromising the exploration of diverse solutions (Nagaraj et al., 2024).

3 METHODOLOGY

In this section, we first review the concept of the CFM framework. Subsequently, we describe the extension of FrameFlow to antibody design, enabling the joint generation of two modalities: antibody structures and PLMs latent embeddings. Figure 1 provides an overview of IgFlow-LM, illustrating the architecture and workflow of the proposed multi-modal generative framework. The code for IgFlow-LM is available at <https://anonymous.4open.science/r/IgFlow-LM-8848/>.

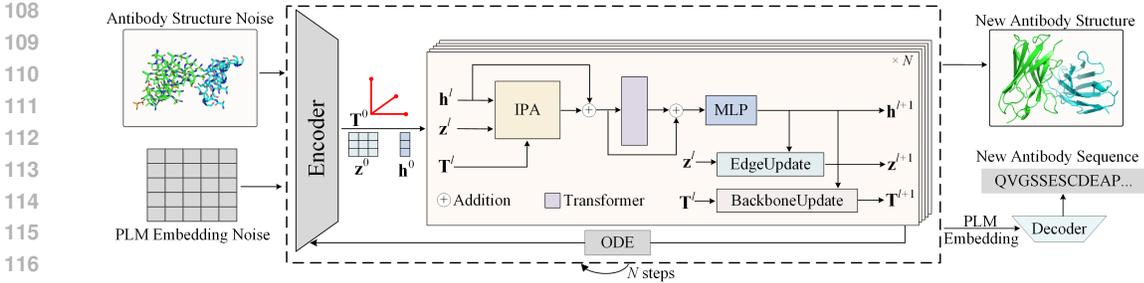


Figure 1: Overview of the IgFlow-LM architecture. IgFlow-LM takes as input initial noise-perturbed antibody structures and noisy PLM latent embeddings. An encoder first processes these inputs together with the current time step to produce initial frames \mathbf{T}^0 , node embeddings \mathbf{h}^0 , and edge embeddings \mathbf{z}^0 . Invariant Point Attention (IPA) updates the node features by integrating information from the explicit residue frames \mathbf{T}^l , as well as the PLM embedding and timestep information encoded in the input embeddings. The updated features are added to the input node features and subsequently fed into a Transformer block to capture inter-residue dependencies. Subsequently, the output is summed with the embeddings from the IPA module and fed into a multi-layer perceptron (MLP) to obtain the updated node embeddings \mathbf{h}^{l+1} . Additionally, the updated node embeddings are used to refine the edge embeddings \mathbf{z}^l and the frame representations \mathbf{T}^l . Through multiple steps of ODE integration, the model generates a novel antibody sequence-structure pairs.

3.1 SE(3)-FLOW MATCHING FOR ANTIBODY STRUCTURE GENERATION

In this study, we employ SE(3)-flow matching to model the generative process of continuous backbone frameworks. The backbone atomic coordinates of the i -th residue in an antibody scaffold are parameterized by a rigid transformation $\mathbf{T}^i \in \text{SE}(3)$ and three fixed heavy atoms $N^*, C^*, C_\alpha^* \in \mathbb{R}^3$ (Jumper et al., 2021; Köhler et al., 2020):

$$[N^i, C^i, C_\alpha^i] = \mathbf{T}^i \cdot [N^*, C^*, C_\alpha^*] \quad (1)$$

where \mathbf{T}^i is an element of the special Euclidean group (the set of valid translations and rotations in Euclidean space). That is, the backbone of an antibody of length L can be parameterized as $\mathbf{T} = [T^{(1)}, \dots, T^{(L)}]$, with $\mathbf{T} \in \text{SE}(3)^L$. Each frame can be decomposed as $T^i = (r^i, x^i)$, where $x^i \in \mathbb{R}^3$ represents the translation vector corresponding to the C_α^* atom, and $r^i \in \text{SO}(3)$ is a 3×3 rotation matrix, which can be derived from relative atomic positions via the Gram-Schmidt process. Following the framework of FrameFlow (Yim et al., 2023a), we initialize the distribution $p_0(\mathbf{T}_0) = p_{\text{noise}}(\mathbf{T}_0)$ and learn a time-dependent vector field $v_\theta(\mathbf{T}_t, t)$ to evolve these frames. This vector field describes a structure flow ordinary differential equation with initial condition $\mathbf{T}_0 \sim p_0$:

$$\frac{d\mathbf{T}_t}{dt} = v_\theta(\mathbf{T}_t, t) \quad (2)$$

The target distribution is defined as the distribution of antibody backbone structures $p_1(\mathbf{T}_1)$. We approximate the SE(3)-equivariant flow by fitting a conditional vector field $u_t(\mathbf{T}_t | \mathbf{T}_1, \mathbf{T}_0)$ using a neural network. This conditional vector field represents an interpolation between the initial (noisy) frames and the final target frames through a series of probability distributions. In other words, the objective loss aims to minimize the discrepancy between the predicted final target and the true target, given the noisy initial condition:

$$\mathcal{L}_{\text{SE}(3)} = \mathbb{E}_{t, p_1(\mathbf{T}_1), p_0(\mathbf{T}_0)} [\|u_t(\mathbf{T}_t | \mathbf{T}_1, \mathbf{T}_0) - v_\theta(\mathbf{T}_t, t)\|^2] \quad (3)$$

After training, we can perform flexible sampling of antibody structures by reverse integrating the ODE with a specified number of time steps and a variance schedule. We provide a more complete description in the Appendix.

3.2 PLM-BASED PROBABILITY FLOWS MATCHING FOR PLMS EMBEDDINGS

Since the latent embeddings provided by PLMs are continuous variables, they are inherently more suitable for continuous FM-based methods (Hu et al., 2024). To leverage the semantically meaning-

ful latent space offered by PLMs, we can directly define the mapping between the initial distribution $p_0 = p_{\text{noise}}(\mathbf{x}_0)$ and the target distribution $p_1 = p(\mathbf{x})$ as a time-dependent vector field:

$$\frac{d\mathbf{x}_t}{dt} = v_\theta(\mathbf{x}_t, t) \quad (4)$$

This time-dependent vector field describes an ODE with an initial condition $\mathbf{x}_0 \sim p_0$. The target distribution corresponds to the distribution of PLM latent embeddings $p_1(\mathbf{x}_1)$. Similar to the SE(3)-equivariant flow, we approximate this PLM embedding flow using a conditional vector field $u_t(\mathbf{x}_t|\mathbf{x}_1, \mathbf{x}_0)$, which is modeled by a neural network. This conditional vector field represents an interpolation between the initial (noisy) embedding and the final target embedding through a series of probability distributions. A more detailed description of this process is provided in the Appendix. As with the SE(3)-equivariant flow, the objective loss aims to minimize the discrepancy between the predicted final target and the true target, given the noisy initial condition:

$$\mathcal{L}_{\text{PLM}} = \mathbb{E}_{t, p_1(\mathbf{x}_1), p_0(\mathbf{x}_0)} = [\|u_t(\mathbf{x}_t|\mathbf{x}_1, \mathbf{x}_0) - v_\theta(\mathbf{x}_t, t)\|^2] \quad (5)$$

After training, we can perform sampling from the PLM latent space by solving the ODE with a specified number of steps.

Additionally, we train an extra decoder to map the sampled PLM embeddings back to discrete sequences (details are provided in the Appendix).

3.3 FLOW MATCHING WITH COMBINED STRUCTURE AND PLM EMBEDDINGS

By integrating these two modalities, we obtain the final multi-modal CFM objective, which is a weighted sum of the individual flow matching objectives:

$$\mathcal{L}_{\text{CFM}} = \mathbb{E}_t(\lambda^{\text{SE}(3)} \mathcal{L}_{\text{SE}(3)} + \lambda^{\text{PLM}} \mathcal{L}_{\text{PLM}}) \quad (6)$$

To reduce chain breaks and steric clashes in the generated structures, we additionally incorporate a pairwise atomic distance-based local loss $\mathcal{L}_{2\text{D}}$ as an auxiliary loss term. Furthermore, to ensure semantic consistency between the predicted PLM embeddings and the original PLM embeddings, we also include a cosine similarity loss \mathcal{L}_{cos} . Additionally, we incorporate a cross-entropy loss \mathcal{L}_{Dec} from the jointly trained decoder. Therefore, the total loss is the weighted sum of all loss terms:

$$\mathcal{L} = \mathcal{L}_{\text{CFM}} + \lambda^{2\text{D}} \mathcal{L}_{2\text{D}} + \lambda^{\text{cos}} \mathcal{L}_{\text{cos}} + \lambda^{\text{Dec}} \mathcal{L}_{\text{Dec}} \quad (7)$$

Additional details regarding model architecture, training configurations, and inference procedures are provided in the Appendix.

4 EXPERIMENTS

In this section, we evaluate the capability of IgFlow-LM in both unconditional design of antibody variable domains and conditional design of CDR loops. IgFlow-LM is trained on antibody structures from SabDab (Dunbar et al., 2014). To prevent data leakage, we partition the dataset based on the germline families of the antibodies, ensuring that no identical heavy-light chain germline family combinations appear in both the training and test sets.

4.1 UNCONDITIONAL DESIGN OF ANTIBODY VARIABLE DOMAINS

In this section, we evaluate the capability of IgFlow-LM in the unconditional generation of paired heavy and light chain variable domains. We adopt the same combinations of heavy and light chain lengths as used in IgDiff (Cutting et al., 2025). For each length combination, we sample 8 structures, resulting in a total of 2000 generated structures.

We adopted the novelty and diversity metrics proposed in IgDiff to evaluate the feasibility of IgFlow-LM for unconditional antibody design. Novelty is determined by calculating the RMSD of the HCDR3 loop relative to its nearest match, serving to quantify the degree of distinctness between the generated structures and known structures within the training set. Specifically, we first employed ANARCI (Dunbar & Deane, 2016) to annotate the sequences generated by IgFlow-LM, thereby

identifying the boundaries of all CDR regions. Subsequently, the nearest match was identified using the TM-score (Zhang & Skolnick, 2004) among entries with identical HCDR3 lengths. Diversity measures the model’s capacity to yield multiple conformationally distinct solutions under identical design conditions. We calculated the pairwise RMSD between all CDR loops of the same type and length, following the same RMSD calculation methodology utilized for the novelty assessment.

Figure 2 (left) shows the HCDR3 RMSD distribution of antibodies generated by IgFlow-LM (1.7 Å) and the HCDR3 RMSD distribution of antibodies in the test set with the same HCDR3 loop lengths (0.91 Å). We additionally display the HCDR3 RMSD distribution of IgFlow-LM-generated antibodies relative to the test set (1.05 Å). The distributions indicate that IgFlow-LM is capable of generating novel HCDR3 loops based on the shape of their distribution

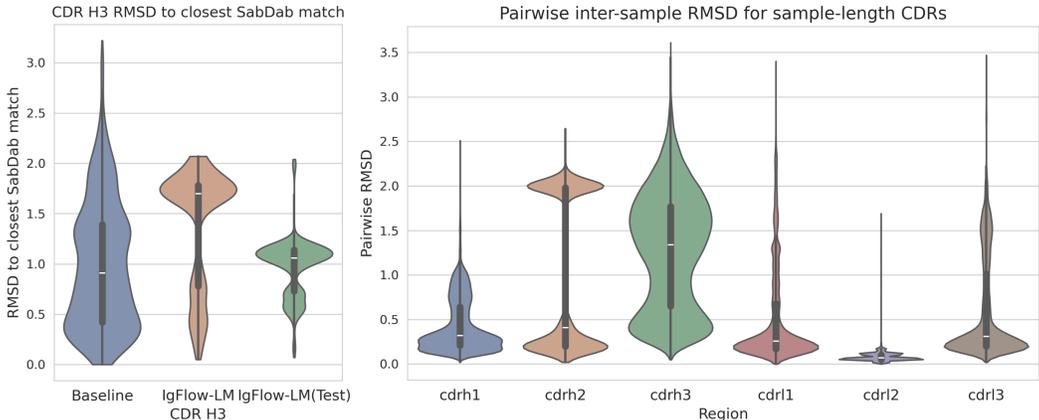


Figure 2: Novelty and diversity metrics computed from 2000 sampled structures. Left: Comparison of novelty between HCDR3 structures generated by IgFlow-LM and randomly selected samples from the training set. For each generated structure, the displayed RMSD corresponds to the structural deviation from its closest match in the training set (SAbDab) as determined by the TM-score, considering only structures with matching HCDR3 loop lengths. Right: Pairwise RMSD values among CDR loops of identical length generated by IgFlow-LM, illustrating the conformational diversity within each CDR type.

Figure 2 (right) shows the distribution of CDR regions across antibody structures generated by IgFlow-LM. We observe that HCDR3 exhibits the highest diversity, which is consistent with findings reported in previous studies (Nagaraj et al., 2024; Cutting et al., 2025; D’Angelo et al., 2018). Additionally, we note that LCDR2 demonstrates the lowest diversity, likely due to its relatively lower mutability compared to other CDR regions (Sankar et al., 2020; Teixeira et al., 2022). In summary, these results demonstrate that IgFlow-LM is capable of designing diverse CDR conformations across different loop regions while adhering to the general structural properties characteristic of each CDR type.

We next analyze the design capabilities of IgFlow-LM. For this evaluation, we select IgDiff (Cutting et al., 2025) and IgFlow (Nagaraj et al., 2024) as baseline models. We first assess the self-consistency of the model. To mitigate potential biases introduced by any single folding algorithm, we employ three independent structure prediction tools—ABodyBuilder2 (ABB2) (Abanades et al., 2023), Ig-Fold (Ruffolo et al., 2023), and ESMFold (Lin et al., 2023)—to fold the generated sequences (see Appendix for details on the folding models). For each folding model, we compute the root-mean-square deviation (RMSD) between the predicted structure and the original sampled structure, which we refer to as the self-consistency RMSD (scRMSD). It should be noted that IgDiff is not capable of simultaneous sequence–structure co-design; therefore, we utilize AbMPNN (Dreyer et al., 2023) to design heavy and light chain sequences for IgDiff, followed by structural folding of the designed sequences.

The results are presented in Table 1. As shown, IgFlow-LM achieves substantially higher success rates across the six CDR regions—0.82, 0.88, and 0.76—compared to IgFlow (0.63, 0.68, and 0.61), although it remains slightly below IgDiff (0.92, 0.96, and 0.80). We attribute IgDiff’s superior

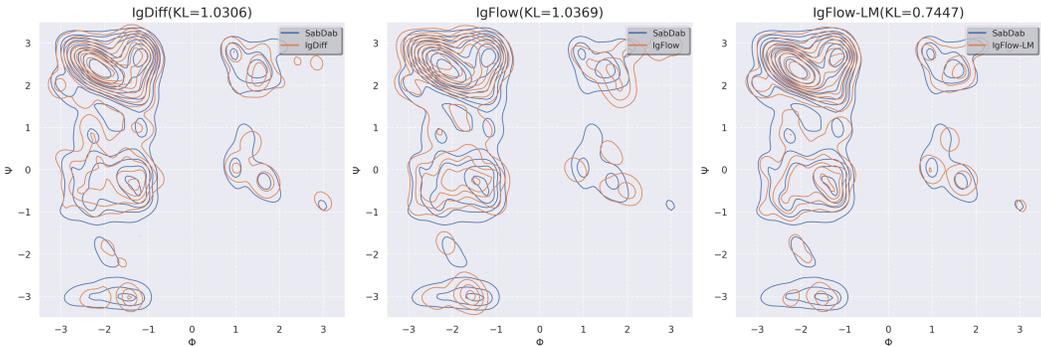


Figure 3: Ramachandran plots (Goodman et al., 1970) comparing the backbone dihedral angle distributions of antibody variable domains generated by IgFlow and IgFlow-LM against those in the SABdab training set. The KL divergence value in each subplot’s title indicates the Kullback–Leibler divergence between the generated structures and the SABdab reference distribution.

performance to the fact that it was trained on structures folded using ABB2, leading to a closer alignment between its training distribution and that of antibody-specific folding models. Consequently, IgDiff exhibits a noticeable drop in scRMSD when evaluated with the general-purpose protein folding model ESMFold. Nevertheless, IgFlow-LM demonstrates strong design capabilities overall.

Table 1: Success rates of 2000 unconditional samples generated by IgFlow-LM and IgFlow, evaluated using two scRMSD metrics across two different folding models: ABB2 and ESMFold. Joint CDR scRMSD is defined as the fraction of samples where the RMSD of all six CDR loops between the refolded and sampled structures is below 2Å. CDR H3 scRMSD measures the fraction of samples where the RMSD of the HCDR3 loop between the refolded and sampled structures is below 2Å.

Folding Model		IgDiff	IgFlow	IgFlow-LM
ABB2	Joint CDR scRMSD	0.92	0.63	0.82
	CDR H3 scRMSD	0.93	0.69	0.84
IgFold	Joint CDR scRMSD	0.96	0.68	0.88
	CDR H3 scRMSD	0.97	0.73	0.89
ESMFold	Joint CDR scRMSD	0.77	0.61	0.77
	CDR H3 scRMSD	0.86	0.69	0.81

We then further visualize the backbone dihedral angle distributions of the structures generated by IgFlow-LM and IgFlow. The results are shown in Figure 3. It can be observed that IgFlow-LM generates antibody structures that closely adhere to the allowed distributions of backbone dihedral angles. Moreover, IgFlow-LM covers a broader range of low-density regions compared to IgFlow. Additionally, IgFlow-LM achieves a Kullback–Leibler (KL) divergence of 0.4309, which is significantly lower than that of IgDiff (1.0306) and IgFlow (1.0369). This indicates that the antibody conformations generated by IgFlow-LM—benefiting from its joint modeling with a protein language model—are more consistent with naturally observed antibody structures. We further report the deviations of the designed structures from ideal bond lengths and bond angles for both IgFlow-LM and IgFlow in Figure 6.

To further investigate the capability of IgFlow-LM in designing stable antibody structures, we employed PyRosetta (Chaudhury et al., 2010) to compute the interfacial binding free energy (ΔG) between the heavy and light chains of antibodies designed by IgDiff, IgFlow, and IgFlow-LM, serving as a metric for structural stability. Notably, in addition to side-chain repacking performed via PyRosetta, all designed antibody structures underwent a relaxation procedure to relieve steric clashes and refine local geometry. To mitigate stochastic variability, we computed ΔG ten times for each antibody and reported the average value as its final ΔG . As shown in Table 4, antibodies designed by IgFlow-LM exhibit the lowest average ΔG , indicating stronger interfacial stability. Moreover,

Moreover, only the CDR regions are subject to design, while the remaining portions of the variable domain are held constant.

First, we evaluated the CDR-loop design capabilities of RFantibody (Bennett et al., 2025), DiffAb (Luo et al., 2022), IgFlow, and IgFlow-LM using scRMSD. It should be noted that none of the aforementioned models incorporate antigen information when modeling CDR regions. We selected 100 antibody structures from the test set constructed above, which collectively span the common lengths of antibody variable regions and include non-canonical CDR loop lengths. For each reference structure, we generated 20 designed structures. Subsequently, we folded all designed sequences using a folding model and computed the RMSD between each predicted structure and the corresponding original sampled structure.

The results for the design of the full CDR regions and the CDR H3 region are presented in Tables 3 and 7, respectively. It can be observed that IgFlow-LM exhibits consistent performance regardless of the folding model employed. In contrast, the other baseline models show significant performance discrepancies between the general protein folding model, ESMFold, and the two antibody-specific folding models. This suggests that IgFlow-LM, through its synergistic modeling with the PLM, demonstrates greater stability compared to the baseline models.

Table 2: Comparison of self-consistency RMSD (scRMSD) success rates across the six CDR regions for RFantibody, DiffAb, IgFlow, and IgFlow-LM in conditional CDR design tasks. Higher scRMSD success rates indicate better designability in the given region. "All" refers to the proportion of samples where all six CDRs loops are independently considered designable.

Folding Model		All	HCDR1	HCDR2	HCDR3	LCDR1	LCDR2	LCDR3
ABB2	RFantibody	0.84	0.99	0.99	0.86	0.97	1.00	0.99
	DiffAb	0.90	0.99	0.99	0.90	0.99	1.00	1.00
	IgFlow	0.88	0.99	0.99	0.89	0.99	1.00	0.99
	IgFlow-LM	0.88	1.00	0.99	0.88	0.99	1.00	1.00
IgFold	RFantibody	0.89	1.00	0.99	0.90	0.99	1.00	1.00
	DiffAb	0.92	0.99	0.99	0.93	0.99	1.00	1.00
	IgFlow	0.89	0.99	0.99	0.90	1.00	1.00	0.99
	IgFlow-LM	0.89	1.00	0.99	0.90	1.00	1.00	1.00
ESMFold	RFantibody	0.92	1.00	0.99	0.95	0.97	1.00	0.99
	DiffAb	0.86	0.99	0.99	0.89	0.96	1.00	0.99
	IgFlow	0.83	0.99	0.99	0.86	0.96	1.00	0.99
	IgFlow-LM	0.86	1.00	0.99	0.87	0.97	1.00	0.99

To further validate the hypothesis regarding the enhanced stability of IgFlow-LM, we employed Cochran’s Q test to assess the overall variance of each generative model across different folding backends, followed by pairwise McNemar’s tests with Holm correction. We evaluated the performance metrics for all models in the HCDR3 design task. The results are summarized in Table 8. For all four generative models, Cochran’s Q test indicated no overall significant difference among the folding backends (all p-values = 1.0). regarding the IgFlow-LM model, only the comparison between ESMFold and IgFold revealed a significant difference after correction (p = 0.0044), whereas ABB2 showed no significant difference with either. The success rates consistently remained within a narrow range (84.7%–87.5%, with overlapping 95% confidence intervals), demonstrating the high robustness of IgFlow-LM across different folding models. In contrast, the other generative models (IgFlow, DiffAb, and RFantibody) exhibited multiple significant pairwise differences and greater variability across folding models, suggesting lower consistency in structural fidelity.

Subsequently, we employed PyRosetta to calculate the heavy-light chain interface ΔG and the Relax RMSD for the CDR regions redesigned by RFantibody, DiffAb, IgFlow, and IgFlow-LM. The evaluation protocol included side-chain repacking (with the exception of DiffAb) followed by structural relaxation. The results are summarized in Tables 9 and 10. Notably, IgFlow-LM exhibited the lowest interface ΔG and the minimal structural deviation upon relaxation (Relax RMSD) among all evaluated models.

To further investigate the impact of the PLM on the diversity of CDR regions designed by IgFlow-LM, we compared the amino acid diversity of sequences generated during conditional CDR loop design against those from DiffAb and IgFlow. The HCDR3 diversity results are presented in Figure 5. It can be observed that IgFlow-LM exhibits superior overall diversity compared to DiffAb and IgFlow, and more closely resembles the distribution of native antibodies. Furthermore, we calculated the positional information entropy for the HCDR3 sequences designed by DiffAb, IgFlow, and IgFlow-LM, yielding values of 2.1435, 1.4333, and 1.9758, respectively. Collectively, these results indicate that, aided by the PLM, IgFlow-LM is capable of designing CDR loop sequences with high diversity.

Additionally, we used AbNatiV2 to compute humanization scores for the CDR regions redesigned by RFantibody, DiffAb, IgFlow, and IgFlow-LM. The results are shown in Tables 11 to 13. IgFlow-LM again demonstrates the highest overall humanization score.

4.3 ABLATION STUDY

To further investigate the impact of integrating a PLM into the generative framework on antibody design performance, we conducted an ablation study in which the PLM probability flow is removed. Specifically, during inference, we perform ODE evolution only on the SE(3) backbone frames, while omitting the ODE integration over the PLM latent flow. In both cases, the same decoder is used to map the final generated PLM embeddings to amino acid sequences. The variant without PLM latent flow evolution is denoted as IgFlow-LM (w/o PLM latent Flow).

As shown in Table 3, disabled the evolution of the ODE of the PLM probability flow leads to a noticeable degradation in design performance. This indicates that when relying solely on structural signals, the model tends to generate conformations that are structurally plausible but may yield sequences that are less natural or evolutionarily inconsistent. The results suggest that the PLM probability flow plays an important role in preserving sequence plausibility and diversity during the generation process.

Table 3: Success rates of 2000 unconditional samples generated by IgFlow-LM and IgFlow-LM (w/o PLM latent Flow), evaluated using two scRMSD metrics across two different folding models: ABB2 and ESMFold. Joint CDR scRMSD is defined as the fraction of samples where the RMSD of all six CDR loops between the refolded and sampled structures is below 2Å. CDR H3 scRMSD measures the fraction of samples where the RMSD of the HCDR3 loop between the refolded and sampled structures is below 2Å.

Folding Model		IgFlow-LM (w/o PLM latent Flow)	IgFlow-LM
ABB2	Joint CDR scRMSD	0.59	0.64
	CDR H3 scRMSD	0.60	0.72
IgFold	Joint CDR scRMSD	0.61	0.69
	CDR H3 scRMSD	0.61	0.75
ESMFold	Joint CDR scRMSD	0.55	0.61
	CDR H3 scRMSD	0.58	0.71

5 CONCLUSION

In this work, we integrate a probabilistic flow from a protein language model with an SE(3) flow-matching architecture to design antibody variable domains under both conditional and unconditional settings. We find that IgFlow-LM is capable of generating realistic antibodies consistent with its training distribution and produces structures that better conform to physical constraints. Under conditional design, IgFlow-LM outperforms baseline models in overall structural quality and also demonstrates superior sequence diversity compared to all other baselines. Future work will involve training IgFlow-LM on a larger number of experimentally determined and synthetic antibody structures, exploring improved strategies for integrating protein language models with structure design, and investigating antigen-aware co-design of antibody sequences and structures.

REFERENCES

- 486
487
488 Brennan Abanades, Wing Ki Wong, Fergus Boyles, Guy Georges, Alexander Bujotzek, and Char-
489 lotte M Deane. Immunebuilder: Deep-learning models for predicting the structures of immune
490 proteins. *Communications Biology*, 6(1):575, 2023.
- 491
492 Rebecca F Alford, Andrew Leaver-Fay, Jeliuzko R Jeliuzkov, Matthew J O’Meara, Frank P DiMaio,
493 Hahnbeom Park, Maxim V Shapovalov, P Douglas Renfrew, Vikram K Mulligan, Kalli Kappel,
494 et al. The rosetta all-atom energy function for macromolecular modeling and design. *Journal of*
495 *chemical theory and computation*, 13(6):3031–3048, 2017.
- 496
497 Heli Ben-Hamu, Samuel Cohen, Joey Bose, Brandon Amos, Aditya Grover, Maximilian Nickel,
498 Ricky TQ Chen, and Yaron Lipman. Matching normalizing flows and probability paths on mani-
499 folds. *arXiv preprint arXiv:2207.04711*, 2022.
- 500
501 Nathaniel R Bennett, Brian Coventry, Inna Goreshnik, Buwei Huang, Aza Allen, Dionne Vafeados,
502 Ying Po Peng, Justas Dauparas, Minkyung Baek, Lance Stewart, et al. Improving de novo protein
503 binder design with deep learning. *Nature Communications*, 14(1):2625, 2023.
- 504
505 Nathaniel R Bennett, Joseph L Watson, Robert J Ragotte, Andrew J Borst, DéJenaé L See, Connor
506 Weidle, Riti Biswas, Yutong Yu, Ellen L Shrock, Russell Ault, et al. Atomically accurate de novo
507 design of antibodies with rdiffusion. *Nature*, pp. 1–11, 2025.
- 508
509 Andrew Campbell, Jason Yim, Regina Barzilay, Tom Rainforth, and Tommi Jaakkola. Generative
510 flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design.
511 *arXiv preprint arXiv:2402.04997*, 2024.
- 512
513 Sidhartha Chaudhury, Sergey Lyskov, and Jeffrey J Gray. Pyrosetta: a script-based interface for im-
514 plementing molecular modeling algorithms using rosetta. *Bioinformatics*, 26(5):689–691, 2010.
- 515
516 Ricky TQ Chen and Yaron Lipman. Flow matching on general geometries. *arXiv preprint*
517 *arXiv:2302.03660*, 2023.
- 518
519 Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary
520 differential equations. *Advances in neural information processing systems*, 31, 2018.
- 521
522 Tianlai Chen, Pranay Vure, Rishab Pulugurta, and Pranam Chatterjee. Amp-diffusion: Integrating
523 latent diffusion with protein language models for antimicrobial peptide generation. *bioRxiv*, pp.
524 2024–03, 2024.
- 525
526 Daniel Cutting, Frédéric A Dreyer, David Errington, Constantin Schneider, and Charlotte M Deane.
527 De novo antibody design with se (3) diffusion. *Journal of Computational Biology*, 32(4):351–361,
528 2025.
- 529
530 Valentin De Bortoli, Emile Mathieu, Michael Hutchinson, James Thornton, Yee Whye Teh, and
531 Arnaud Doucet. Riemannian score-based generative modelling. *Advances in neural information*
532 *processing systems*, 35:2406–2422, 2022.
- 533
534 Xinqiang Ding, Zhengting Zou, and Charles L Brooks III. Deciphering protein evolution and fitness
535 landscapes with latent space models. *Nature communications*, 10(1):5644, 2019.
- 536
537 Jinhui Dong, Jessica A Finn, Peter A Larsen, Timothy PL Smith, and James E Crowe Jr. Structural
538 diversity of ultralong cdrh3s in seven bovine antibody heavy chains. *Frontiers in immunology*,
539 10:558, 2019.
- 540
541 Frédéric A Dreyer, Daniel Cutting, Constantin Schneider, Henry Kenlay, and Charlotte M Deane. In-
542 verse folding for antibody sequence design using deep learning. *arXiv preprint arXiv:2310.19513*,
543 2023.
- 544
545 James Dunbar and Charlotte M Deane. Anarci: antigen receptor numbering and receptor classifica-
546 tion. *Bioinformatics*, 32(2):298–300, 2016.
- 547
548 James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, Jiye
549 Shi, and Charlotte M Deane. Sabdab: the structural antibody database. *Nucleic acids research*,
42(D1):D1140–D1146, 2014.

- 540 Sara D’Angelo, Fortunato Ferrara, Leslie Naranjo, M Frank Erasmus, Peter Hrabec, and Andrew RM
541 Bradbury. Many routes to an antibody heavy-chain cdr3: necessary, yet insufficient, for specific
542 binding. *Frontiers in immunology*, 9:395, 2018.
- 543 Murray Goodman, Antonio S Verdini, Nam S Choi, and Yukio Masuda. Polypeptide stereochem-
544 istry. *Top. Stereochem*, 5:69–166, 1970.
- 546 Brian L Hie, Varun R Shanker, Duo Xu, Theodora UJ Bruun, Payton A Weidenbacher, Shaogeng
547 Tang, Wesley Wu, John E Pak, and Peter S Kim. Efficient evolution of human antibodies from
548 general protein language models. *Nature biotechnology*, 42(2):275–283, 2024.
- 549 Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in*
550 *neural information processing systems*, 33:6840–6851, 2020.
- 552 Emiel Hooeboom, Didrik Nielsen, Priyank Jaini, Patrick Forré, and Max Welling. Argmax flows
553 and multinomial diffusion: Learning categorical distributions. *Advances in neural information*
554 *processing systems*, 34:12454–12465, 2021.
- 555 Emiel Hooeboom, Victor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffu-
556 sion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–
557 8887. PMLR, 2022.
- 559 Vincent Hu, Di Wu, Yuki Asano, Pascal Mettes, Basura Fernando, Björn Ommer, and Cees Snoek.
560 Flow matching for conditional text generation in a few sampling steps. In *Proceedings of the 18th*
561 *Conference of the European Chapter of the Association for Computational Linguistics (Volume*
562 *2: Short Papers)*, pp. 380–392, 2024.
- 563 John Ingraham, Vikas Garg, Regina Barzilay, and Tommi Jaakkola. Generative models for graph-
564 based protein design. *Advances in neural information processing systems*, 32, 2019.
- 565 Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Bonaventure FP
566 Dossou, Chanakya Ajit Ekbote, Jie Fu, Tianyu Zhang, Michael Kilgour, Dinghuai Zhang, et al.
567 Biological sequence design with gflownets. In *International Conference on Machine Learning*,
568 pp. 9786–9801. PMLR, 2022.
- 570 Wengong Jin, Jeremy Wohlwend, Regina Barzilay, and Tommi Jaakkola. Iterative refinement graph
571 neural network for antibody sequence-structure co-design. *arXiv preprint arXiv:2110.04624*,
572 2021.
- 573 Bowen Jing, Stephan Eismann, Patricia Suriana, Raphael JL Townshend, and Ron Dror. Learning
574 from protein structure with geometric vector perceptrons. *arXiv preprint arXiv:2009.01411*, 2020.
- 575 John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger,
576 Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate
577 protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.
- 579 Henry Kenlay, Frédéric A Dreyer, Aleksandr Kovaltsuk, Dom Miketa, Douglas Pires, and Char-
580 lotte M Deane. Large scale paired antibody language models. *PLOS Computational Biology*, 20
581 (12):e1012646, 2024.
- 582 Asif Khan, Alexander I Cowen-Rivers, Antoine Grosnit, Derrick-Goh-Xin Deik, Philippe A Robert,
583 Victor Greiff, Eva Smorodina, Puneet Rawat, Rahmad Akbar, Kamil Dreczkowski, et al. Toward
584 real-world automated antibody design with combinatorial bayesian optimization. *Cell Reports*
585 *Methods*, 3(1), 2023.
- 587 Jonas Köhler, Leon Klein, and Frank Noé. Equivariant flows: exact likelihood generative learning
588 for symmetric densities. In *International conference on machine learning*, pp. 5361–5370. PMLR,
589 2020.
- 590 Xiangzhe Kong, Wenbing Huang, and Yang Liu. End-to-end full-atom antibody design. *arXiv*
591 *preprint arXiv:2302.00203*, 2023a.
- 592 Xiangzhe Kong, Wenbing Huang, and Yang Liu. End-to-end full-atom antibody design. *arXiv*
593 *preprint arXiv:2302.00203*, 2023b.

- 594 Tim Kucera, Matteo Togninalli, and Laetitia Meng-Papaxanthos. Conditional generative modeling
595 for de novo protein design with hierarchical functions. *Bioinformatics*, 38(13):3454–3461, 2022.
596
- 597 Xiang Li, John Thickstun, Ishaan Gulrajani, Percy S Liang, and Tatsunori B Hashimoto. Diffusion-
598 lm improves controllable text generation. *Advances in neural information processing systems*, 35:
599 4328–4343, 2022.
- 600 Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin,
601 Robert Verkuil, Ori Kabeli, Yaniv Shmueli, et al. Evolutionary-scale prediction of atomic-level
602 protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.
603
- 604 Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching
605 for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.
606
- 607 Sidney Lyayuga Lisanza, Jacob Merle Gershon, Samuel WK Tipps, Jeremiah Nelson Sims, Lucas
608 Arnoldt, Samuel J Hendel, Miriam K Simma, Ge Liu, Muna Yase, Hongwei Wu, et al. Multistate
609 and functional protein design using rosettafold sequence space diffusion. *Nature biotechnology*,
610 pp. 1–11, 2024.
- 611 Amy X Lu, Wilson Yan, Sarah A Robinson, Kevin K Yang, Vladimir Gligorijevic, Kyunghyun Cho,
612 Richard Bonneau, Pieter Abbeel, and Nathan Frey. Generating all-atom protein structure from
613 sequence-only training data. *bioRxiv*, pp. 2024–12, 2024.
- 614 Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific
615 antibody design and optimization with diffusion-based generative models for protein structures.
616 *Advances in Neural Information Processing Systems*, 35:9754–9767, 2022.
617
- 618 Zhenjie Luo, Aoyun Geng, Leyi Wei, Quan Zou, Feifei Cui, and Zilong Zhang. Cpl-diff: A diffusion
619 model for de novo design of functional peptide sequences with fixed length. *Advanced Science*,
620 12(20):2412926, 2025.
- 621 Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine
622 learning research*, 9(Nov):2579–2605, 2008.
623
- 624 Sai Pooja Mahajan, Jeffrey A Ruffolo, Rahel Frick, and Jeffrey J Gray. Hallucinating structure-
625 conditioned antibody libraries for target-specific binders. *Frontiers in immunology*, 13:999034,
626 2022.
- 627 Sanjay Nagaraj, Amir Shanehsazzadeh, Hyun Park, Jonathan King, and Simon Levine. Igflow: Flow
628 matching for de novo antibody design. *Advances in Neural Information Processing Systems*, 2024.
629
- 630 Jo Erika T Narciso, Iris Diana C Uy, April B Cabang, Jenina Faye C Chavez, Juan Lorenzo B Pablo,
631 Gisela P Padilla-Concepcion, and Eduardo A Padlan. Analysis of the antibody structure based on
632 high-resolution crystallographic studies. *New biotechnology*, 28(5):435–447, 2011.
633
- 634 Erik Nijkamp, Jeffrey A Ruffolo, Eli N Weinstein, Nikhil Naik, and Ali Madani. Progen2: exploring
635 the boundaries of protein language models. *Cell systems*, 14(11):968–978, 2023.
- 636 Tobias H Olsen, Fergus Boyles, and Charlotte M Deane. Observed antibody space: A diverse
637 database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein
638 Science*, 31(1):141–146, 2022.
639
- 640 Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lip-
641 man, and Ricky TQ Chen. Multisample flow matching: Straightening flows with minibatch cou-
642 plings. *arXiv preprint arXiv:2304.14772*, 2023.
- 643 Aubin Ramon, Niccolo Frassetto, Haowen Zhao, Xing Xu, Matthew Greenig, Shimobi Onuoha, and
644 Pietro Sormanni. Deep learning assessment of nativeness and pairing likelihood for antibody and
645 nanobody design with abnativ2. *bioRxiv*, pp. 2025–10, 2025.
646
- 647 Roshan Rao, Joshua Meier, Tom Sercu, Sergey Ovchinnikov, and Alexander Rives. Transformer
protein language models are unsupervised structure learners. *Biorxiv*, pp. 2020–12, 2020.

- 648 Matthew IJ Raybould, Claire Marks, Konrad Krawczyk, Bruck Taddese, Jaroslaw Nowak, Alan P
649 Lewis, Alexander Bujotzek, Jiye Shi, and Charlotte M Deane. Five computational developability
650 guidelines for therapeutic antibody profiling. *Proceedings of the National Academy of Sciences*,
651 116(10):4025–4030, 2019.
- 652 William H Robinson. Sequencing the functional antibody repertoire—diagnostic and therapeutic
653 discovery. *Nature Reviews Rheumatology*, 11(3):171–182, 2015.
- 654 Jeffrey A Ruffolo, Lee-Shin Chu, Sai Pooja Mahajan, and Jeffrey J Gray. Fast, accurate antibody
655 structure prediction from deep learning on massive set of natural antibodies. *Nature communica-*
656 *tions*, 14(1):2389, 2023.
- 657 Jeffrey A Ruffolo, Aadyot Bhatnagar, Joel Beazer, Stephen Nayfach, Jordan Russ, Emily Hill, Riffat
658 Hussain, Joseph Gallagher, and Ali Madani. Adapting protein language models for structure-
659 conditioned design. *BioRxiv*, pp. 2024–08, 2024.
- 660 Kannan Sankar, Kam Hon Hoi, and Isidro Hötzel. Dynamics of heavy chain junctional length biases
661 in antibody repertoires. *Communications biology*, 3(1):207, 2020.
- 662 Samuel Schmitz, Emily A Schmitz, James E Crowe Jr, and Jens Meiler. The human antibody se-
663 quence space and structural design of the v, j regions, and cdrh3 with rosetta. In *MAbs*, volume 14,
664 pp. 2068212. Taylor & Francis, 2022.
- 665 Harry W Schroeder Jr and Lisa Cavacini. Structure and function of immunoglobulins. *Journal of*
666 *allergy and clinical immunology*, 125(2):S41–S52, 2010.
- 667 Amir Shanehsazzadeh, Julian Alverio, George Kasun, Simon Levine, Jibrán A Khan, Chelsea
668 Chung, Nicolas Diaz, Breanna K Luton, Ysis Tarter, Cailen McCloskey, et al. In vitro validated
669 antibody design against multiple therapeutic antigens using generative inverse folding. *BioRxiv*,
670 pp. 2023–2012, 2023.
- 671 Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised
672 learning using nonequilibrium thermodynamics. In *International conference on machine learn-*
673 *ing*, pp. 2256–2265. pmlr, 2015.
- 674 Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv*
675 *preprint arXiv:2010.02502*, 2020a.
- 676 Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben
677 Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint*
678 *arXiv:2011.13456*, 2020b.
- 679 Andre AR Teixeira, Sara D’Angelo, M Frank Erasmus, Camila Leal-Lopes, Fortunato Ferrara,
680 Laura P Spector, Leslie Naranjo, Esteban Molina, Tamara Max, Ashley DeAgüero, et al. Si-
681 multaneous affinity maturation and developability enhancement using natural liability-free cdrs.
682 In *MAbs*, volume 14, pp. 2115200. Taylor & Francis, 2022.
- 683 Brian L Trippe, Jason Yim, Doug Tischer, David Baker, Tamara Broderick, Regina Barzilay, and
684 Tommi Jaakkola. Diffusion probabilistic modeling of protein backbones in 3d for the motif-
685 scaffolding problem. *arXiv preprint arXiv:2206.04119*, 2022.
- 686 Yuko Tsuchiya and Kenji Mizuguchi. The diversity of h 3 loops determines the antigen-binding
687 tendencies of antibody cdr loops. *Protein Science*, 25(4):815–825, 2016.
- 688 Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural compu-*
689 *tation*, 23(7):1661–1674, 2011.
- 690 Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eise-
691 nach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of
692 protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.
- 693 James C Whisstock and Arthur M Lesk. Prediction of protein function from protein sequence and
694 structure. *Quarterly reviews of biophysics*, 36(3):307–340, 2003.

702 Jason Yim, Andrew Campbell, Andrew YK Foong, Michael Gastegger, José Jiménez-Luna, Sarah
703 Lewis, Victor Garcia Satorras, Bastiaan S Veeling, Regina Barzilay, Tommi Jaakkola, et al. Fast
704 protein backbone generation with se (3) flow matching. *arXiv preprint arXiv:2310.05297*, 2023a.
705

706 Jason Yim, Brian L Trippe, Valentin De Bortoli, Emile Mathieu, Arnaud Doucet, Regina Barzilay,
707 and Tommi Jaakkola. Se (3) diffusion model with application to protein backbone generation.
708 *arXiv preprint arXiv:2302.02277*, 2023b.

709 Shu Zhang, Xinyi Yang, Yihao Feng, Can Qin, Chia-Chih Chen, Ning Yu, Zeyuan Chen, Huan
710 Wang, Silvio Savarese, Stefano Ermon, et al. Hive: Harnessing human feedback for instructional
711 visual editing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern
712 Recognition*, pp. 9026–9036, 2024.

713 Yang Zhang and Jeffrey Skolnick. Scoring function for automated assessment of protein structure
714 template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4):702–710, 2004.
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755

A MODEL ARCHITECTURE DETAILS

A.1 REPRESENTATIONS

SE(3)-Flow Matching. In three-dimensional structural modeling, the spatial conformation of a protein is determined by the rotations and translations of its rigid groups. These transformations are commonly described using two standard mathematical groups: SO(3) and SE(3).

SO(3) (the three-dimensional special orthogonal group) represents the set of all three-dimensional rotations and consists of all 3×3 orthogonal matrices with determinant 1. It encompasses only rotational transformations, excluding translations, and preserves both distances and angles. In protein modeling, SO(3) is typically used to describe the rotation of local residue coordinate frames. SE(3) (the three-dimensional special Euclidean group) represents the set of all three-dimensional rigid-body transformations, formed by combining rotations from SO(3) with translations in \mathbb{R}^3 . It includes both rotational and translational components and is used to describe the full rigid-body pose of atomic groups or residues in space. Modern structure prediction models, such as AlphaFold and RoseTTAFold, operate directly in SE(3) space.

In IgFlow-LM, we utilize the SE(3)-equivariant flow matching framework introduced in FrameFlow Yim et al. (2023a) to model the backbone structure of antibodies. For backbone parameterization, we directly adopt the framework used in AlphaFold2. According to the description in AlphaFold2, each residue in an antibody can be represented by a rigid-body frame consisting of a rotation matrix $r_i \in \text{SO}(3)$ and a translation vector $x_i \in \mathbb{R}^3$, which together form a group element Hu et al. (2024). The backbone of an antibody with N residues can then be represented as the sequence of these frames $\{T^{(1)}, \dots, T^{(N)}\} \in \text{SE}(3)^N$. It is important to note that, following the convention in FrameFlow, we represent protein structures within the zero-center-of-mass (CoM) subspace of $\mathbb{R}^{N \times 3}$ by subtracting the CoM from all data points. This step is crucial for ensuring that the distribution of generated frames remains SE(3)-equivariant.

Based on the work by Yim et al. (2023b), we can independently model SO(3) and \mathbb{R}^3 during both training and sampling, meaning that these two components can be modeled separately. Following the setup in FrameFlow, the prior distribution for SO(3) is chosen as a uniform distribution over SO(3), while the prior distribution for \mathbb{R}^3 is selected as an isotropic Gaussian distribution $\mathcal{N}(0, I_3)$. Therefore, the prior distribution for a single frame can be set as $p_0(T_0^{(i)}) = \mathcal{U}(\text{SO}(3)) \otimes \mathcal{N}(0, I_3)$. The conditional flow $T_t = \psi_t(T_0|T_1)$ is then defined as the geodesic between T_0 and T_1 :

$$T_t = \exp_{T_0}(\log_{T_0}(T_1)) \quad (8)$$

where \exp_T denotes the exponential map at point T , and \log_T represents the logarithmic map at point T . By applying the logarithmic map to obtain the tangent vector from T_0 to T_1 , scaling it by the factor t , and then transforming it back to the manifold using the exponential map, the resulting point lies on the geodesic. According to the work by Yim et al. (2023b), Equation (1) can be simplified into the following forms:

$$\text{Translations}(\mathbb{R}^3) : x_t = (1 - t)x_0 + tx_1 \quad (9)$$

$$\text{Rotations}(\text{SO}(3)) : \exp_{r_0}(t \log_{r_0}(r_1)) \quad (10)$$

The exponential map \exp_{r_0} can be directly computed using the Rodrigues formula, while the logarithmic map \log_{x_0} is calculated based on the results from Yim et al. (2023b). Therefore, the overall objective can be written as:

$$\mathcal{L}_{\text{SE}(3)} = \mathbb{E}_{t, p_0(\mathbf{T}_0), p_1(\mathbf{T}_1)} \left[\sum_{n=1}^N \{ \|v_x^{(n)}(\mathbf{T}_t, t) - x_t^{(n)}\|_{\mathbb{R}^3}^2 + \|v_r^{(n)}(\mathbf{T}_t, t) - r_t^{(n)}\|_{\text{SO}(3)}^2 \} \right] \quad (11)$$

where N represents the number of residues in the antibody. This objective approximates the CFM target mentioned in our manuscript. The vector fields $v_x^{(n)}(\mathbf{T}_t, t)$ and $v_r^{(n)}(\mathbf{T}_t, t)$ can be fitted using SE(3)-equivariant neural networks. And we calculate the derivatives of $x_t^{(n)}$ and $r_t^{(n)}$ with respect to time t :

$$x_t^{(n)} = \frac{x_1^{(n)} - x_t^{(n)}}{1 - t}, r_t^{(n)} = \frac{\log_{r_t^{(n)}}(r_1^{(n)})}{1 - t} \quad (12)$$

We define $\{\hat{x}_1^{(n)}, \hat{r}_1^{(n)}\}_{n=1}^N$ as the interpolated frames \mathbf{T}_t at time t , which are predictions based on the original frame \mathbf{T}_1 . According to the work by Yim et al. (2023b), we reparameterize the objective as:

$$\mathcal{L}_{\text{SE}(3)} = \mathbb{E}_{t, p_0(\mathbf{T}_0), p_1(\mathbf{T}_1)} \left[\frac{1}{(1-t)^2} \sum_{n=1}^N \{ \|\hat{x}_1^{(n)}(\mathbf{T}_t, t) - x_1^{(n)}\|_{\mathbb{R}^3}^2 + \|\log_{r_t^{(n)}}(\hat{r}_1^{(n)}(\mathbf{T}_t, t)) - \log_{r_t^{(n)}}(r_1^{(n)})\|_{\text{SO}(3)}^2 \} \right] \quad (13)$$

PLM-based Probability Flows Matching for PLM Embeddings. PLMs convert discrete amino acid sequences into continuous numerical representations (Ding et al., 2019), allowing us to directly define a vector field between the prior distribution $p_0(\mathbf{x}_0) = p_{\text{noise}}(\mathbf{x}_0)$ and the target distribution $p_1(\mathbf{x}_1) = p(\mathbf{x})$ based on the FM principle:

$$\frac{d\mathbf{x}_t}{dt} = v_\theta(\mathbf{x}_t, t) \quad (14)$$

Here, we choose an isotropic Gaussian distribution $\mathcal{N}(0, I_D)$ as the prior distribution for PLM embeddings, where D denotes the dimensionality of the PLM embeddings. The corresponding flow can be defined as a conditional Gaussian distribution (Song et al., 2020b):

$$p_t(\mathbf{x}_t) = \mathbb{E}_{\mathbf{x}_t \sim \mathcal{N}(\mu_t, \sigma_t)} p(\mathbf{x}_1 | \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_1 | \mu_t(\mathbf{x}_t), \sigma_t(\mathbf{x}_1)) \quad (15)$$

This form is widely used in existing generative models (Ho et al., 2020). Similarly, we can define the following flow matching objective for the vector field $u_t(\mathbf{x}_t)$:

$$\mathcal{L}_{\text{FM}} = \mathbb{E}_{t, p_t(\mathbf{x}_t)} \|v_\theta(\mathbf{x}_t, t) - u_t(\mathbf{x}_t)\|^2 \quad (16)$$

Empirically, we construct a conditional vector field $u_t(\mathbf{x}_t | \mathbf{x}_1, \mathbf{x}_0)$:

$$u_t(\mathbf{x}_t | \mathbf{x}_1, \mathbf{x}_0) = (1-t)\mathbf{x}_0 + t\mathbf{x}_1 \quad (17)$$

This form intuitively constructs a geodesic (i.e., a straight line segment) between two points in Euclidean space, which aligns with some existing flow matching losses used for image generation (Chen & Lipman, 2023). At this point, we reparameterize the original flow matching objective as follows:

$$\mathcal{L}_{\text{PLM}} = \mathbb{E}_{t, p_1(\mathbf{x}_1), p_0(\mathbf{x}_0)} \|u_t(\mathbf{x}_t | \mathbf{x}_1, \mathbf{x}_0) - v_\theta(\mathbf{x}_t, t)\|^2 \quad (18)$$

Decoder for Mapping PLM Embeddings Back to Discrete Sequence Space. After generating new PLM embeddings using IgFlow-LM, we require a decoder to map these embeddings back to the discrete sequence space, thereby obtaining complete new antibody sequence-structure pairs. The decoder maps the generated PLM embeddings $\hat{\mathbf{x}}_1 = [\hat{x}_1^{(1)}, \dots, \hat{x}_1^{(L)}]$ to a discrete sequence:

$$\hat{y}^{(i)} = \mathcal{L}_{\mathbb{D}} = \text{Softmax}(\mathbf{W}_0 \cdot f_\phi(\hat{\mathbf{x}}_1^{(i)}) + \mathbf{b}_0) \quad (19)$$

where f_ϕ represents a multi-layer perceptron (MLP), \mathbf{W}_0 and \mathbf{b}_0 are the projection matrix and bias term of the output layer, respectively, and $\hat{y}^{(i)}$ denotes the predicted amino acid probability distribution at position i . We optimize this decoder using cross-entropy loss:

$$\mathcal{L}_{\mathbb{D}} = -\frac{1}{L} \sum_{l=1}^L \log \hat{p}_\phi(a_l | \hat{\mathbf{x}}_1^{(l)}) \quad (20)$$

where $\hat{p}_\phi(a_l | \hat{\mathbf{x}}_1^{(l)})$ represents the decoder’s predicted probability distribution for the l -th amino acid, and a_l denotes the true value of the l -th amino acid.

A.2 MODEL ARCHITECTURE

IgFlow-LM is a modified version of the FrameFlow architecture (Yim et al., 2023a), adapted for sequence-structure co-design of antibodies. We set the hidden dimension of the IPA to 16 and increased the number of IPA modules from 6 to 10. Additionally, we increased the number of transformer layers per block from 2 to 4, thereby enhancing the overall depth of the network to better accommodate the multimodal task of jointly generating PLM embeddings and antibody structures.

Dropout was applied to each transformer layer with a dropout rate of 0.1, randomly setting 0.1 of the neurons to zero during training to mitigate overfitting.

Encoder. The model’s encoder takes as input the three-dimensional structure of the target protein and outputs node embeddings and edge (residue-pair) embeddings. For node (residue) embeddings, we use a combination of the following features:

- Interpolated PLM embeddings.
- Positional encoding of residuals.
- Chain index encoding.
- Timestep embedding.

In the case of conditional CDR region design for antibodies, additional features derived from non-CDR regions are incorporated:

- Residue type embeddings.
- Atomic coordinates, including both backbone and side-chain atoms.
- The dihedral angles of the backbone, encoded using sine-based positional embeddings.
- Side-chain torsion angles, similarly encoded with sine-based embeddings.

Each feature is first processed through separate MLPs. The resulting representations are concatenated and further transformed via an additional MLP to produce the final node embeddings.

For edge (residue-pair) embeddings, we employ a mixture of the following features:

- Cross-concatenation of the two node embeddings.
- Relative positional encoding between residues.
- Positional bins of the interpolated translation vector between residue pairs.

Additionally, for conditional CDR design in antibodies, the following non-CDR-derived pairwise features are included:

- Residue type pair embeddings, represented by a learnable 20×20 dimensional embedding matrix.
- Relative sequence position, modeled via learnable embeddings of the sequence distance between residue pairs.
- Euclidean distance between the two residues.
- Relative orientation between residues, characterized by computing inter-residue backbone dihedral angles and encoded using sine-based embeddings.

Similarly, these edge features are individually encoded through dedicated MLPs, concatenated, and passed through another MLP to generate the final edge embeddings.

Similarly, these features are encoded using different MLPs and are concatenated features and transformed by another MLP to form the final edge embedding. The encoder is SE(3)-invariant, meaning that its output will be the same regardless of any global rigid transformation. Here we set the hidden dimension of residue embedding as 256, and the hidden dimension of pair embedding hidden dimension as 128.

Protein Language Model. For the choice of PLM, we select IgBert (Kenlay et al., 2024), which is specifically trained on the Observed Antibody Space (OAS) Olsen et al. (2022). IgBert can accept both heavy chain and light chain sequences as input and output the corresponding latent space embeddings. IgBert appends three special tokens—start, end, and chain separation—to the input sequences. The output is an embedding of size $\mathbb{R}^{(L+3) \times 1024}$, where L is the length of the antibody sequence, and 1024 is the dimensionality of IgBert. Since these special tokens may affect the alignment between sequences and structures, we remove the embeddings corresponding to these three special tokens before using them for model training.

We also employ self-conditioning based on distances and backbone dihedral angles. Specifically, during training, the predicted translation vectors and backbone coordinates, obtained with gradient computation disabled, are used as part of the model input. The translation vectors are then converted into discrete distance bin features, and backbone dihedral angles are computed from the backbone coordinates, both of which are incorporated as part of the initial encoding features.

Furthermore, we find that incorporating additional equivariant features—such as backbone dihedral angles and residue centrality—as well as invariant features—including relative sequence position and secondary structure information—is beneficial in the conditional CDR generation setting. In other words, during conditional CDR generation, information from the non-CDR regions within the entire variable domain is provided as additional conditioning input to the model.

A.3 TRAINING DETAILS

Dataset. We trained IgFlow-LM on the Structural Antibody Database (SAbDab) (Dunbar et al., 2014). The antibody structures in this database are all experimentally resolved. We preprocessed each structure by retaining only those with a resolution below 4Å and trimming their variable regions. To avoid potential issues arising from unresolved residues in the structural data, we aligned sequences using the variable-region sequences provided by the SAbDab database as the reference. For residues lacking resolved coordinates, we assigned zero-valued coordinate placeholders and excluded these unresolved residues from loss computation.

To prevent data leakage, we partitioned the dataset based on the germline families of the antibodies, ensuring that no identical heavy-light chain germline family combinations appear in both the training and test sets. For antibodies lacking family annotations for either the heavy or light chain, we first attempted to assign their germline families using ANARCI. Only those sequences for which family assignment could not be determined were included in the training set. After this procedure, the final dataset consists of 726 test samples and 9,850 training samples. The germline family combinations reserved for the test set are as follows: IGHV1-IGKV10, IGHV14-IGKV6, IGHV1-IGKV12, IGHV5-IGKV3, IGHV14-IGKV14, IGHV4-IGLV2, IGHV3-IGLV6, IGHV2-IGKV5.

Loss Function. The flow matching-based loss comprises two components: the SE(3) flow-related loss and the PLM probability flow-related loss. The SE(3) flow-related loss includes terms for the predicted rotation matrices and translation vectors relative to their ground-truth values. The translation loss is computed as the L2 distance between the model-predicted and true translation vectors. The rotation loss is defined as the L2 loss on SO(3), measuring the discrepancy between predicted and ground-truth rotation matrices. Both losses are normalized by the number of valid residues during averaging. Their respective weights are set to 2.0 and 1.0. Additionally, we incorporate two structural auxiliary losses: a backbone atomic position loss and a local proximity loss based on pairwise interatomic distances, with weights set to 0.2 and 0.2, respectively. The PLM probability flow-related loss is defined as the L2 distance between the model-predicted and true PLM latent embeddings, again averaged over the number of valid residues. This term is assigned a weight of 0.02. To further regularize the PLM embedding predictions, we include a cosine similarity loss as an auxiliary objective, with a weight of 0.25.

The decoder is trained jointly with the flow matching model to strengthen the alignment between the generated embeddings and the sequence decoding process. The decoder is optimized using the cross-entropy loss, with a weight of 1.0.

We used the AdamW optimizer with a learning rate of , a batch size of 32, and weight decay set to 0.01. A cosine learning rate schedule was employed, with a warm-up phase for the first 500 optimization steps, followed by a gradual decay to over 30,000 optimization steps. Training was conducted on four A100 PCIe 40G GPUs, totaling 30,000 optimization steps. Additionally, during training, we randomly decided whether to use self-conditioning for model optimization with a probability of 0.5.

The code for IgFlow-LM can be accessed at <https://anonymous.4open.science/r/IgFlow-LM-8848>.

B APPENDIX

Table 4: Interfacial ΔG between heavy and light chains and RMSD between pre- and post-relaxation structures for antibodies designed by IgDiff, IgFlow, and IgFlow-LM calculated using PyRosetta. We report the means and standard deviations for all metrics.

Model	ΔG	Relax RMSD
IgDiff	-57.2654 \pm 9.3825	1.04 \pm 0.2675
IgFlow	-53.2978 \pm 9.8673	1.56 \pm 0.3358
IgFlow-LM	-58.1383 \pm 7.9444	0.93 \pm 0.2158

Table 5: Humaneness scores of antibody sequences designed by IgFlow and IgFlow-LM evaluated using AbNatiV2. "Heavy-Light" denotes the humaneness score for the heavy-light chain pairing. All scores range from 0 to 1, where higher scores indicate a closer resemblance to natural human antibodies.

Model	Heavy-Light	CDR H1	CDR H2	CDR H3
IgDiff	0.8486 \pm 0.0611	0.7755 \pm 0.1779	0.6329 \pm 0.2591	0.7588 \pm 0.1810
IgFlow	0.8664 \pm 0.0524	0.7179 \pm 0.2007	0.5635 \pm 0.1774	0.8295 \pm 0.1444
IgFlow-LM	0.8658 \pm 0.0802	0.8052 \pm 0.1940	0.7210 \pm 0.2467	0.7122 \pm 0.1699

Table 6: Humaneness scores of antibody sequences designed by IgFlow and IgFlow-LM evaluated using AbNatiV2. All scores range from 0 to 1, where higher scores indicate a closer resemblance to natural human antibodies.

Model	CDR L1	CDR L2	CDR L3
IgDiff	0.6834 \pm 0.3131	0.7517 \pm 0.4110	0.6501 \pm 0.2782
IgFlow	0.5182 \pm 0.3487	0.7200 \pm 0.4305	0.7341 \pm 0.1941
IgFlow-LM	0.7452 \pm 0.3304	0.8192 \pm 0.2986	0.8088 \pm 0.2161

Table 7: Comparison of HCDR3 scRMSD success rates among RFantibody, DiffAb, IgFlow, and IgFlow-LM for the conditional CDR H3 design task. A higher scRMSD success rate indicates superior designability of the specified region.

Folding Model	HCDR3	
ABB2	RFantibody	0.88
	DiffAb	0.88
	IgFlow	0.89
	IgFlow-LM	0.87
IgFold	RFantibody	0.91
	DiffAb	0.90
	IgFlow	0.90
	IgFlow-LM	0.88
ESMFold	RFantibody	0.94
	DiffAb	0.87
	IgFlow	0.85
	IgFlow-LM	0.86

Table 8: Structural accuracy ($\text{scRMSD} \leq 2\text{\AA}$) of designed HCDR3 regions across generative models and folding backends. Values are reported as success rate with 95% confidence interval. Statistical significance is evaluated using McNemar’s test with Holm correction; * indicates $p_{\text{adj}} < 0.05$.

Model	ESMFold	ABB2	IgFold
IgFlow-LM	0.847 (0.831–0.862)	0.858 (0.841–0.872)	0.875 (0.860–0.889)*
IgFlow	0.853 (0.837–0.868)*	0.900 (0.886–0.912)*	0.902 (0.889–0.915)*
DiffAb	0.877 (0.862–0.891)*	0.884 (0.870–0.898)*	0.902 (0.889–0.915)*
RFantibody	0.940 (0.929–0.950)*	0.878 (0.863–0.892)*	0.915 (0.902–0.926)*

Table 9: Heavy-light chain interface ΔG and Relax RMSD calculated using PyRosetta for IgDiff, IgFlow, and IgFlow-LM in the 6-CDR redesign task. The RMSD reflects the structural deviation between the antibody structures before and after relaxation. We report the mean and standard deviation for all metrics.

Model	ΔG	Relax RMSD
RFantibody	-55.1967±8.8900	1.12±0.3438
DiffAb	-56.0825±8.7548	1.10±0.3822
IgFlow	-57.0922±8.5538	1.11±0.4084
IgFlow-LM	-57.1340±8.2845	0.96±0.3221

Table 10: Heavy-light chain interface ΔG and Relax RMSD calculated using PyRosetta for IgDiff, IgFlow, and IgFlow-LM in the HCDR3 redesign task. The RMSD reflects the structural deviation between the antibody structures before and after relaxation. We report the mean and standard deviation for all metrics.

Model	ΔG	Relax RMSD
RFantibody	-55.0949±9.1402	1.24±0.3460
DiffAb	-54.4597±8.4309	1.01±0.4162
IgFlow	-55.1652±8.8481	1.05±0.4147
IgFlow-LM	-55.2103±8.8527	0.91±0.3110

1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133

Table 11: Humanness scores calculated using AbNatiV2 for antibody sequences generated by RFantibody, DiffAb, IgFlow, and IgFlow-LM in the 6-CDR redesign task. Scores range from 0 to 1, with higher values indicating a closer resemblance to native human antibodies.

Model	CDR H1	CDR H2	CDR H3
RFantibody	0.4609±0.1877	0.4036±0.2419	0.3402±0.1748
DiffAb	0.7858±0.2048	0.5045±0.2240	0.6923±0.2025
IgFlow	0.8220±0.1769	0.7584±0.1716	0.7313±0.1802
IgFlow-LM	0.8501±0.1894	0.7896±0.2033	0.7024±0.1863

Table 12: Humanness scores calculated using AbNatiV2 for antibody sequences generated by RFantibody, DiffAb, IgFlow, and IgFlow-LM in the 6-CDR redesign task. Scores range from 0 to 1, with higher values indicating a closer resemblance to native human antibodies.

Model	CDR L1	CDR L2	CDR L3
RFantibody	0.5331±0.2803	0.4641±0.3192	0.6012±0.2326
DiffAb	0.6396±0.3685	0.7389±0.2396	0.6636±0.2040
IgFlow	0.8580±0.3178	0.8586±0.1833	0.8451±0.1541
IgFlow-LM	0.8182±0.3492	0.9134±0.1771	0.8355±0.1651

Table 13: Humanness scores calculated using AbNatiV2 for antibody sequences generated by RFantibody, DiffAb, IgFlow, and IgFlow-LM in the CDR H3 redesign task. Scores range from 0 to 1, with higher values indicating a closer resemblance to native human antibodies.

Model	CDR H3
RFantibody	0.3545±0.2003
DiffAb	0.6593±0.2074
IgFlow	0.7705±0.1680
IgFlow-LM	0.6937±0.2014

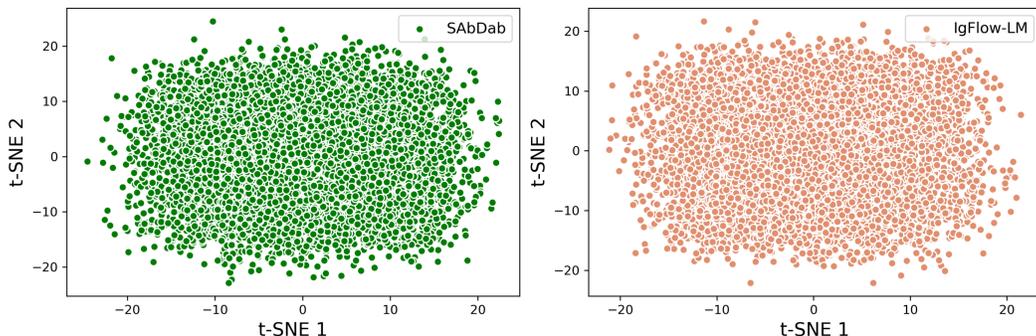


Figure 5: Visualization of IgBert embedding vectors generated by IgFlow-LM and reduced in dimension via t-SNE, alongside visualization of IgBert embedding vectors derived from SabDab antibody sequences.

1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187

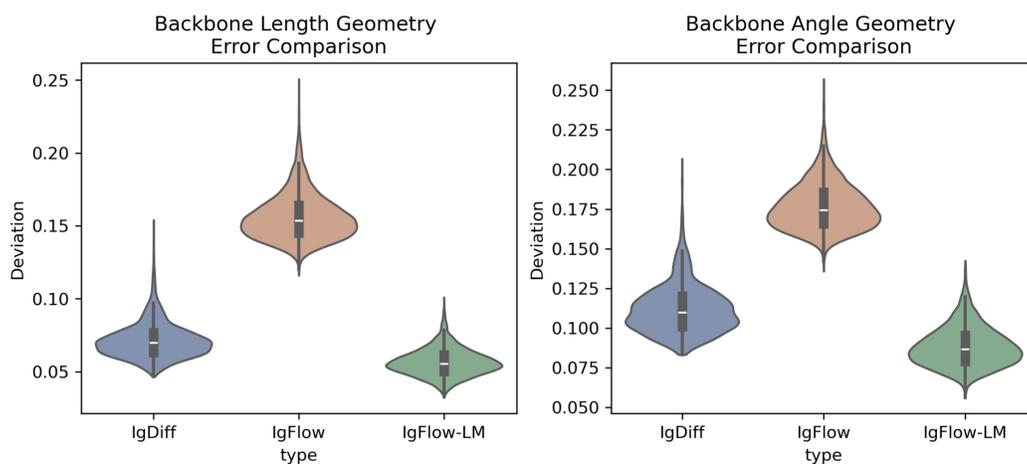


Figure 6: Deviations of bond lengths and dihedral angles from their ideal values in structures generated by IgDiff, IgFlow, and IgFlow-LM. The ideal bond length is 1.329 Å; the ideal cosine values for the dihedral angles CA-C-N and C-N-CA are -0.4415 and -0.5255, respectively.

1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241

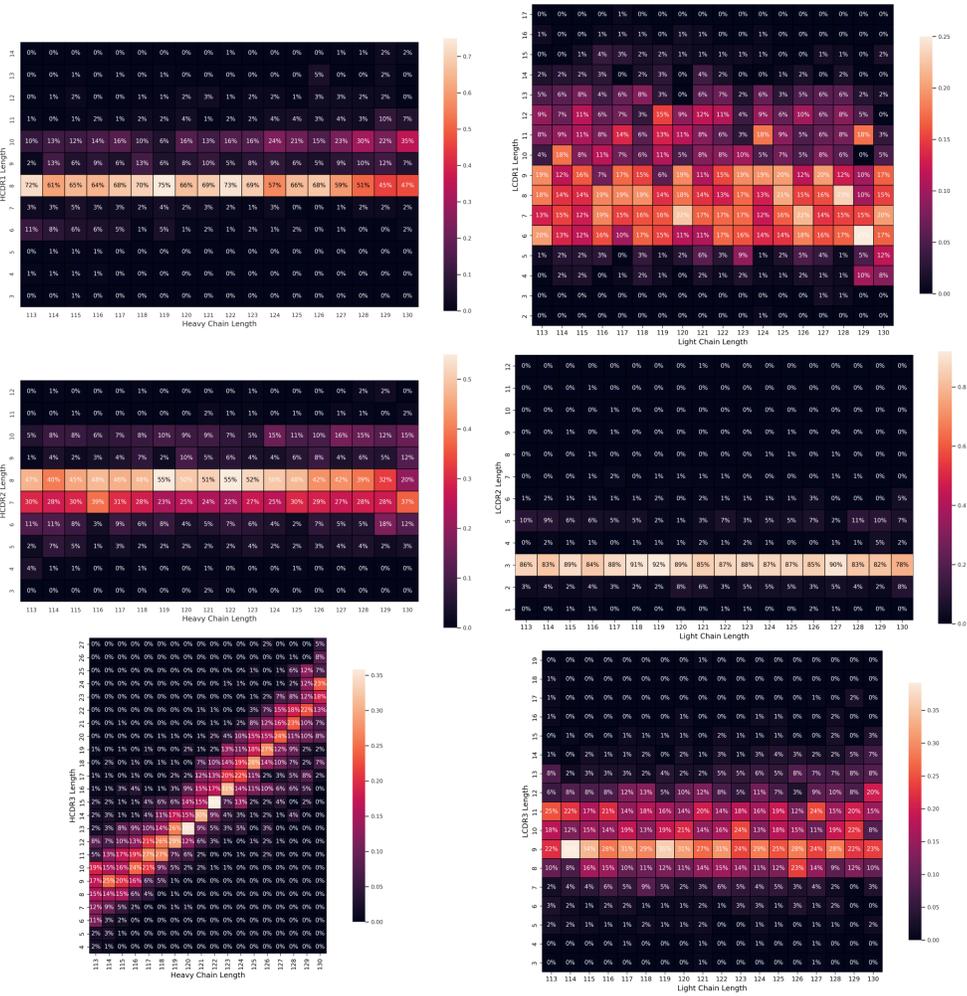


Figure 7: Distribution of CDR loop lengths across designs of specified heavy and light chain lengths. We note that HCDR length changes seem to directly correlate and possibly drive HCDR3 length, with smaller associations observed for the other chains.

1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295

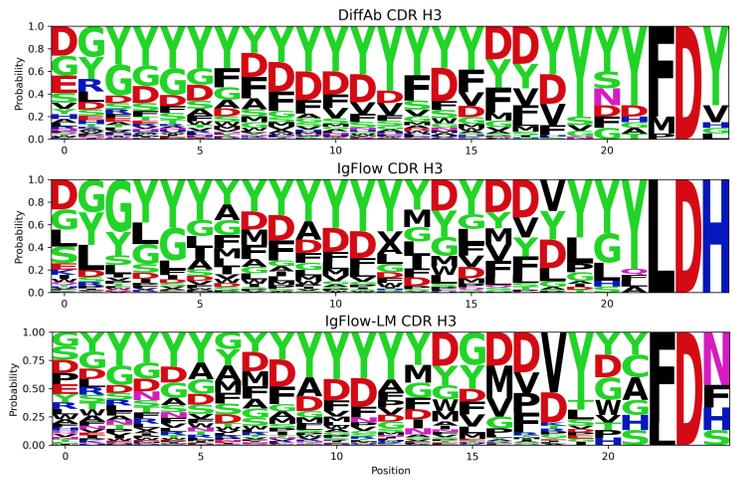


Figure 8: Sequence logo of the HCDR3 region generated by IgFlow-LM and IgFlow in the full CDR design task, depicting the frequency of amino acid occurrences at each position.

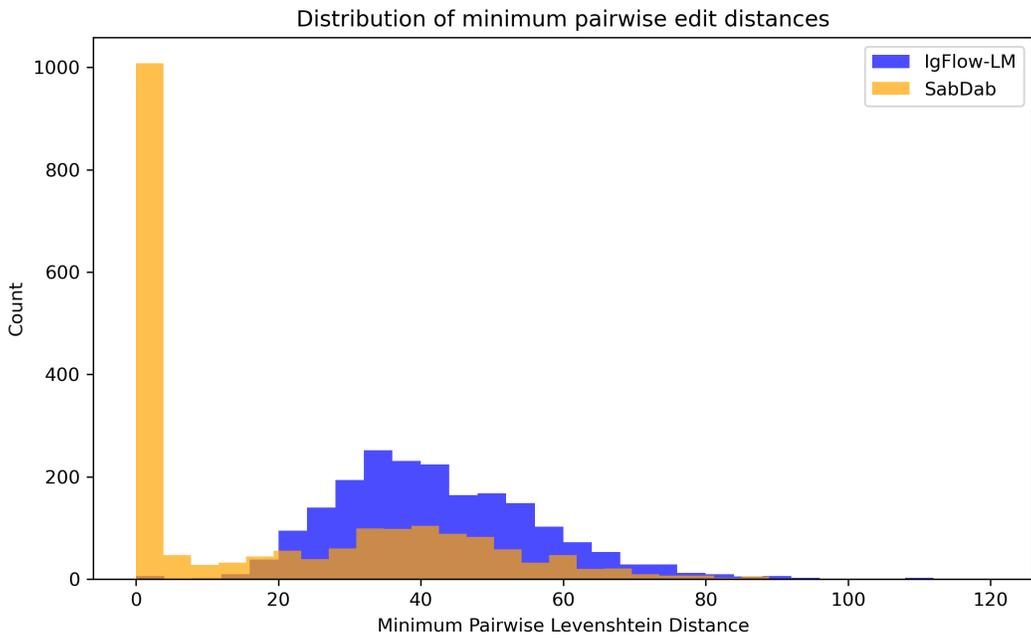


Figure 9: Histogram of the minimum pairwise Levenshtein distance between 2000 unconditioned IgFlow-LM generated structures (blue) and 2000 paired SabDab sequences (orange). SabDab contains experimentally determined antibody structures, many of which originate from similar immune responses or have undergone engineering modifications, leading to a significant degree of sequence redundancy and repetition.

1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349

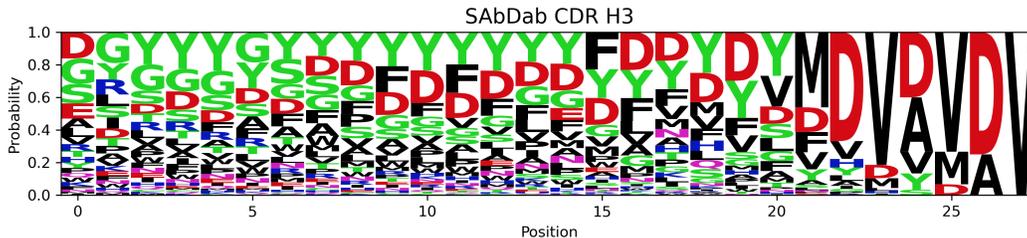


Figure 10: Distribution of CDR loop lengths across designs of specified heavy and light chain lengths. We note that HCDR length changes seem to directly correlate and possibly drive HCDR3 length, with smaller associations observed for the other chains.

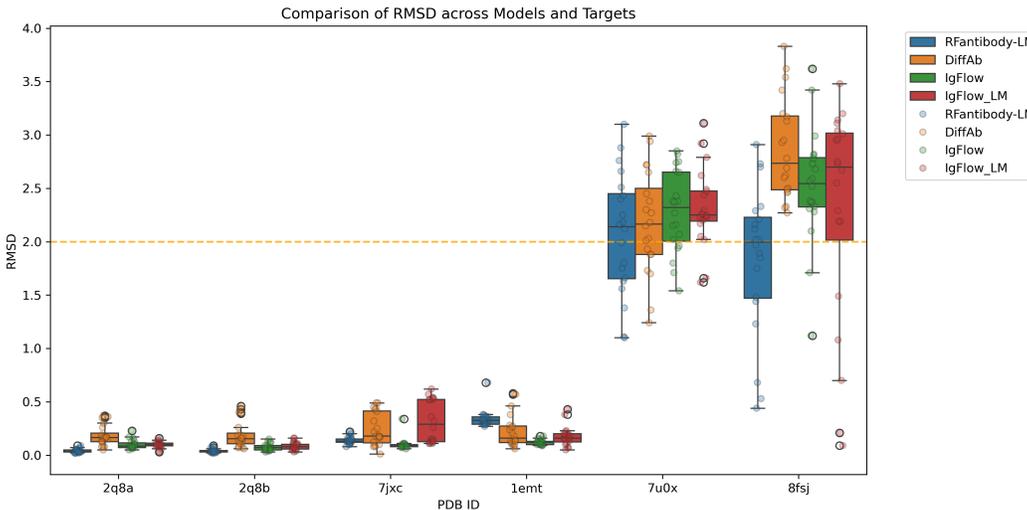


Figure 11: scRMSD distribution of RFantibody, DiffAb, IgFlow, and IgFlow-LM models when designing atypical CDR lengths in the CDR region.