

## Guiding Diffusion Models for Antibody Sequence and Structure Co-design with Developability Properties

Amelia Villegas-Morcillo <sup>\*</sup>, Jana M. Weber <sup>†</sup> and Marcel J. T. Reinders <sup>‡</sup>

Department of Intelligent Systems, Delft University of Technology, Delft 2628 CD, The Netherlands



(Received 25 May 2024; accepted 9 August 2024; published 12 September 2024)

Recent advances in deep generative methods have allowed antibody sequence and structure co-design. This study addresses the challenge of tailoring the highly variable complementarity-determining regions (CDRs) in antibodies to fulfill developability requirements. We introduce a guidance approach that integrates property information into the antibody design process using diffusion probabilistic models. This approach allows us to simultaneously design CDRs conditioned on antigen structures while considering critical properties like solubility and folding stability. Our property-guided diffusion model offers versatility by accommodating diverse property constraints, presenting a promising avenue for computational antibody design in therapeutic applications.

DOI: [10.1103/PRXLife.2.033012](https://doi.org/10.1103/PRXLife.2.033012)

### I. INTRODUCTION

Antibodies are Y-shaped proteins produced by the immune system in response to pathogens called antigens [1]. Antibodies are composed of two heavy and two light chains, with a constant and a variable region [Fig. 1(a)]. The variable region in the antibody constitutes the paratope, which interacts with the antigen's epitope. Within the variable domains, there are six *complementarity-determining regions* (CDRs), denoted as H1, H2, H3, L1, L2, and L3. These regions show high variability in both sequence and structure across antibodies (especially CDR-H3) and determine the specificity of an antibody for a particular antigen.

Antibody engineering involves the refinement of the CDRs to enhance functionality or certain properties. From a therapeutic perspective, there is a significant interest in the *in silico* design of CDRs capable of binding to specific antigens. Traditional approaches rely on energy-based optimization by minimizing the Rosetta energy function over multiple sequence samples [2], which is computationally intensive and time-consuming. To overcome these limitations, deep generative language models have been proposed to discover CDR amino acid sequences [3,4]. Subsequently, leveraging deep-learning-based structure modeling methods, the antibody three-dimensional (3D) structure can be predicted from the designed CDR sequence [5,6]. Recent advancements in deep generative methods offer enhanced performance by

co-designing both the sequence and the structure of CDRs simultaneously [7–9]. One notable advantage of co-design over sequence-based approaches is the capability to condition on both the antigen epitope and antibody framework structures during generation, which has proven useful for affinity optimization [8,9].

Next to the antigen-targeting performance of antibodies, measured as binding affinity and specificity, their *developability* properties are essential for therapeutic developments. These include factors such as solubility, aggregation propensity, thermal stability, and immunogenicity. These properties are vital to ensure that the antibody can be manufactured and is suitable for clinical applications [10,11]. Some aspects of the antibody developability depend on amino-acid-related attributes, such as hydrophobicity or electrostatic charges. A subset of these attributes is integrated into the therapeutic antibody profiling [12] to filter out candidate antibodies with poor developability. Other aspects of developability depend on the interplay of antibody sequence and structure. Some machine-learning approaches leverage this information to predict general developability [13]. In addition, various prediction tools have been proposed for specific developability properties, such as for solubility and aggregation [14], for immunogenicity risks [15], or for binding affinity and stability [16]. It is worth noting that, while existing antibody design methods partially optimize for antigen-targeting properties, the integration of developability parameters remains an open and crucial challenge.

Consequently, in this study, we employ deep generative models for antibody design to generate *de novo* sequences and structures for the CDR loops. Beyond conditioning on the antigen structure, our approach involves guiding the model to produce candidate antibodies with favorable developability attributes. We propose a property-guided denoising diffusion probabilistic model (DDPM). DDPMs [17,18] have demonstrated their capability to generate realistic protein sequences and structures [19,20], including antibodies [8]. As outlined

<sup>\*</sup>Contact author: A.O.VillegasMorcillo@tudelft.nl

<sup>†</sup>Contact author: J.M.Weber@tudelft.nl

<sup>‡</sup>Contact author: M.J.T.Reinders@tudelft.nl

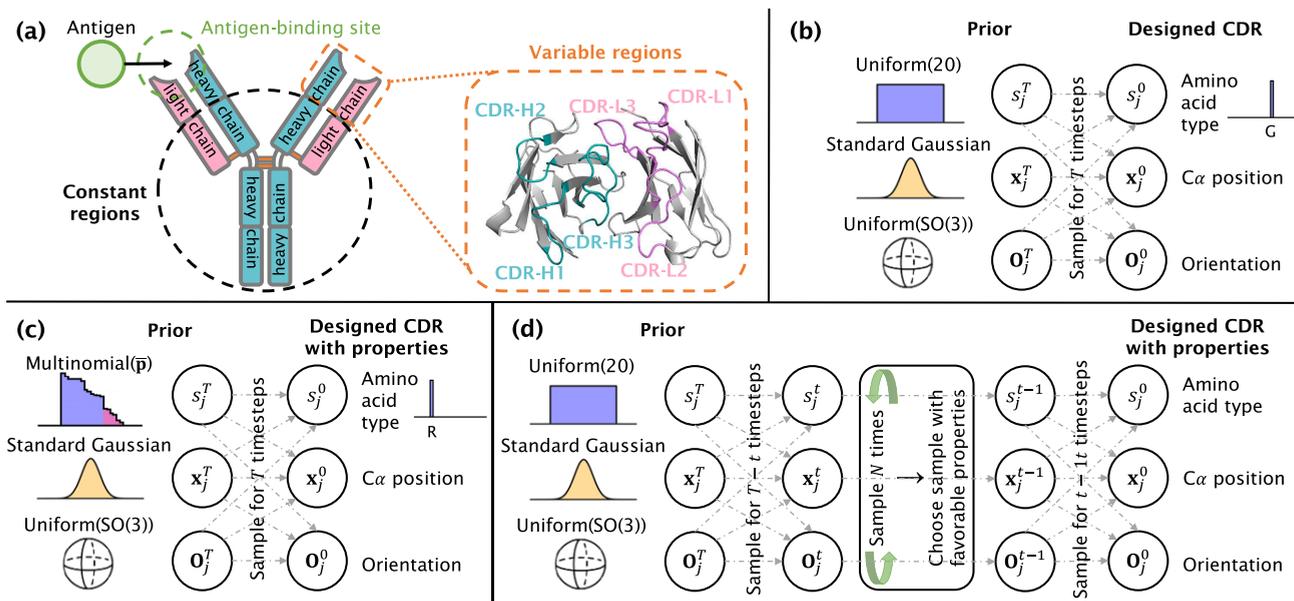


FIG. 1. (a) Illustration of an antibody, featuring the two heavy (in blue) and two light (in pink) chains. The variable regions in the enlarged area encompass the antigen-binding site including the six CDR loops. (b)–(d) Visualization of the generative diffusion process, showcasing the prior distributions for each modality and the designed CDR, for (b) the property-unguided mode, (c) the property-aware prior approach, and (d) the sampling by property approach. (Note: The neural network parametrization is omitted from the figures but is present before sampling at each generation timestep  $t$ .)

in Ref. [20], DDPMs are well suited to protein design due to their ability to generate highly diverse outputs, guide each generation step towards specific design objectives, and explicitly model protein 3D structures using rotationally equivariant networks [19,21]. These capabilities are essential for co-designing antibody CDRs using property information.

Here, we use *guidance* to control the diffusion model output as opposed to *conditioning*. Although these two terms are often used interchangeably in the literature, there is a key difference between them [22]. In guidance-based approaches, the diffusion model is trained to generate generic proteins, and the controlling (i.e., inclusion of property information) is performed during sampling. In contrast, conditioning involves training the diffusion model to accept specific information as input, requiring retraining or fine-tuning if another type of information is needed for controlling the output. In our particular case, the diffusion model for CDR co-design is conditioned on the rest of the antibody-antigen structure during both training and generative sampling, while we guide it using developability properties exclusively during the sampling process. This eliminates the need for retraining the diffusion model.

As the central contribution of our work, we introduce two distinct approaches to guide the generative diffusion process: one incorporates a *property-aware prior*, while the other involves *sampling by property*. Notably, our proposed solutions are adaptable to any property or set of properties that can be computed or predicted based on the intermediate designs at the sequence or structure level. We observe that by imposing property constraints, our model yields antibodies with more favorable developability profiles while preserving their structural integrity compared to the reference antibodies [44].

## II. RELATED WORK

### A. Generative models for antibodies

In the wake of advancements in deep generative modeling for proteins [19,20,23–25], there has been a surge in the development of antibody-specific models. Some of these models focused solely on generating antibody sequences, often for tasks like affinity maturation [3] or CDR infilling [4,26,27]. Saka *et al.* [3] employed a long short-term memory model to discover antibody sequences with high affinity. Meanwhile, Shuai *et al.* [4] proposed an antibody language model to redesign CDR sequences using bidirectional context and conditioning on chain type and species. Melnyk *et al.* [26] repurposed a general pretrained language model for antibody CDR design. Additionally, Gruver *et al.* [27] introduced a method for controllable categorical diffusion using a masked language model-style denoising approach. In contrast to these methods, Eguchi *et al.* [28] delved into antibody structure generation utilizing variational autoencoders to encode-decode information from the backbone torsion angles and pairwise distances.

As both CDR sequences and structures are typically unknown, more recent models have adopted a co-design approach where antibody sequence and structure are generated concurrently. Co-design emerges from the necessity to generate optimal sequences with corresponding structures, without relying on *a priori* knowledge of the structure since it is rarely available [7]. However, one key challenge lies in the scarcity of antibody structure data, which is limited to several thousand examples in the SABDab database [29], hindering the learning ability of deep models to predict the relationship between sequence and structural data [30]. Within this co-design regime, some methods [7,30] adopted iterative approaches to

generate CDR sequences and structures through alternating steps. Jin *et al.* [7] employed a graph neural network (GNN) to co-design CDRs in the heavy chain given the rest of the framework region. In contrast, Gao *et al.* [30] combined an antibody pretrained language model with a hierarchical message-passing network (HMPN), incorporating the antigen epitope as input data. In subsequent works, Jin *et al.* [31] also used an HMPN model to generate CDR-H3 structures and sequences that bind to a specified antigen epitope structure. Kong *et al.* [32] additionally included the light chain of the antibody as input to an E(3)-equivariant GNN framework for heavy-chain CDR generation. In their follow-up work, Kong *et al.* [9] integrated the epitope-docking step in their model, generating a full-atom representation of the antibody structure alongside the paratope sequence. In alignment with these strategies, Luo *et al.* [8] proposed a diffusion-based generative model for CDR sequence-structure co-design (in both the heavy and light chains), incorporating information from the antigen epitope and the antibody framework while modeling amino acid orientations in SO(3). More recently, Martinkus *et al.* [33] demonstrated *in vitro* the expressibility and binding of antibodies generated by diffusion models. While some of these methods, such as the method in Ref. [7], incorporate specific property-guided optimization for virus neutralization, none of them offer a platform for general therapeutic antibody sequence-structure co-design that integrates information beyond antigen-targeting performance.

### B. Conditional and guided generation using diffusion models

DDPMs learn stochastic processes to generate data by denoising samples from a prior distribution [17,18]. For image synthesis, Song *et al.* [18] and Dhariwal and Nichol [34] used the gradients of an image classifier pretrained on noisy inputs from various timesteps of diffusion to control the generation process, and they compared it to a conditionally trained diffusion model. For protein design, Lianza *et al.* [35] employed sequence-based potentials, such as user-defined charge composition, isoelectric points, or hydrophobicity, as functions to control the generation process. However, these methods often require additional training with conditioning parameters in a fine-tuning stage. In contrast, we adopt gradient-free guidance approaches for integrating property information into diffusion models for antibody generation without the need for retraining. Bypassing the retraining phase provides flexibility, enabling a single diffusion model to generate samples that fulfill a broader range of criteria. In this context, our *sampling by property* approach shares more similarities with concurrent works utilizing sequential Monte Carlo (SMC) methods for particle filtering to address inpainting problems such as the motif-scaffolding problem in proteins [21,36]. More specifically, Wu *et al.* [36] introduce a precise method for conditional sampling. They use a diffusion model trained without conditioning as the sample generator (prior) and a classifier trained on clean data as the target predictor (likelihood distribution). By employing a method called *twisted* SMC, they simulate a group of weighted trajectories (particles) in the diffusion process. These trajectories are guided by proposals from the classifier (samples given the target) and weighted according to specific schemes.

## III. METHODS

### A. Diffusion model for antibody design

Our work builds upon an existing method for antibody sequence and structure co-design using diffusion models. Specifically, we used the DiffAb model [8], which enables the joint generation of CDR sequences and structures while conditioning on the antibody framework and bound antigen. The model requires three inputs: amino acid types denoted as  $s_i \in \{\text{ACDEFGHIKLMNPQRSTVWY}\}$ ,  $C_\alpha$  atom positions denoted as  $\mathbf{x}_i \in \mathbb{R}^3$ , and amino acid orientations denoted as  $\mathbf{O}_i \in \text{SO}(3)$ , where  $i$  is the position of the amino acid in the sequence. We generate one CDR loop at a time, denoted as  $\mathcal{R} = \{(s_j, \mathbf{x}_j, \mathbf{O}_j) | j = l + 1, \dots, l + m\}$ , given the rest of the antibody-antigen complex  $\mathcal{C} = \{(s_i, \mathbf{x}_i, \mathbf{O}_i) | i \neq j\}$ .

The forward diffusion process ( $t = 0, \dots, T$ ) gradually introduces noise into each modality through different distributions  $q$  towards the prior distributions. For the amino acid types,  $q(s_j^t | s_j^0)$  follows a multinomial distribution; for the  $C_\alpha$  positions,  $q(\mathbf{x}_j^t | \mathbf{x}_j^0)$  is modeled as Gaussian; and for the amino acid orientations,  $q(\mathbf{O}_j^t | \mathbf{O}_j^0)$  is an isotropic Gaussian. Starting from the prior distributions, the generative diffusion process ( $t = T, \dots, 0$ ) transforms each modality toward the data distribution, as depicted in Fig. 1(b). In this process, parametric models  $p_\theta$  are employed to approximate the posterior distributions at each generation timestep. Different neural networks are used for the three modalities, with a shared encoder and separate decoders. A summary of the main equations can be found in Table I. For an in-depth understanding of the diffusion process, neural network architectures, and training of the models, we refer readers to Ref. [8].

### B. Antibody design guided on properties

We develop two distinct strategies for integrating property information into the generative process of the trained diffusion model, DiffAb [8]. The first strategy, *property-aware prior*, involves replacing the noninformative uniform distribution for amino acid types (see Table I) with a multinomial distribution informed by a specific property of interest. The second strategy, *sampling by property*, involves sampling multiple times at each generation step and subsequently selecting the sample with the most optimal property values.

#### 1. Property-aware prior

The prior distribution for the amino acid types follows a uniform distribution across 20 classes representing the 20 types of amino acids. In this approach, we propose starting the generation of amino acid types (at  $t = T$ ) by sampling from a property-aware prior in the following form:

$$\begin{aligned} s_j^T &\sim \text{Multinomial}(\mathbf{p}) \\ &= (1 - b) \cdot \text{Uniform}(20) + b \cdot \text{Multinomial}(\mathbf{p}). \end{aligned} \quad (1)$$

Here,  $\mathbf{p} = [p_1, \dots, p_{20}]$ , where  $p_k$  represents the probability of the amino acid type  $k$  given a property of interest. The uniform and multinomial components are weighted by a constant  $b$  that can be adjusted based on the application requirements. This approach is depicted in Fig. 1(c). Given the prior in Eq. (1), the posterior probabilities at each generation timestep

TABLE I. Summary of equations defining the forward and generative diffusion processes, the prior distributions, and the training objectives of the three different modalities used in DiffAb [8]. Here,  $F_\theta(\cdot)$ ,  $G_\theta(\cdot)$ , and  $H_\theta(\cdot)$  are the neural network models for the amino acid types,  $C_\alpha$  atom positions, and orientations, respectively;  $D_{\text{KL}}$  is the Kullback-Leibler divergence;  $\mathcal{N}(\cdot)$  is the Gaussian distribution;  $\epsilon_j$  is the standard Gaussian noise added to the input to obtain the noisy version  $\mathbf{x}'_j = \sqrt{\bar{\alpha}^t} \mathbf{x}_j^0 + \sqrt{1 - \bar{\alpha}^t} \epsilon_j$ ;  $\mathcal{IG}_{\text{SO}(3)}(\cdot)$  is the isotropic Gaussian distribution in  $\text{SO}(3)$ ; and  $\text{ScaleRot}(\cdot)$  is the rotation scaling function defined in Ref. [8, Appendix B.3].

Feature modality		Equation
Amino acid type	Forward:	$q(s'_j   s_j^0) = \text{Multinomial}[\bar{\alpha}^t \cdot \text{onehot}(s'_j) + (1 - \bar{\alpha}^t) \cdot \frac{\mathbf{1}}{20}]$
	Prior:	$s_j^T \sim \text{Uniform}(20)$
	Generative:	$p_\theta(s_j^{t-1}   \mathcal{R}^t, C) = \text{Multinomial}\{F_\theta(\mathcal{R}^t, C)[j]\}$
	Objective:	$L_{\text{type}}^t = \mathbb{E}\{\frac{1}{m} \sum_j D_{\text{KL}}[q(s_j^{t-1}   s_j^0)    p_\theta(s_j^{t-1}   \mathcal{R}^t, C)]\}$
$C_\alpha$ position	Forward:	$q(\mathbf{x}'_j   \mathbf{x}_j^0) = \mathcal{N}[\mathbf{x}'_j   \sqrt{\bar{\alpha}^t} \cdot \mathbf{x}_j^0, (1 - \bar{\alpha}^t) \mathbf{I}]$
	Prior:	$\mathbf{x}_j^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_3)$
	Generative:	$p_\theta(\mathbf{x}'_j^{t-1}   \mathcal{R}^t, C) = \mathcal{N}[\mathbf{x}'_j^{t-1}   \mu_p(\mathcal{R}^t, C), (1 - \bar{\alpha}^t) \mathbf{I}]$ $\mu_p(\mathcal{R}^t, C) = \frac{1}{\sqrt{\alpha^t}} (\mathbf{x}'_j - \frac{1 - \alpha^t}{\sqrt{1 - \bar{\alpha}^t}} G_\theta(\mathcal{R}^t, C)[j])$
	Objective:	$L_{\text{pos}}^t = \mathbb{E}[\frac{1}{m} \sum_j \ \epsilon_j - G_\theta(\mathcal{R}^t, C)[j]\ ^2]$
Orientation in $\text{SO}(3)$	Forward:	$q(\mathbf{O}'_j   \mathbf{O}_j^0) = \mathcal{IG}_{\text{SO}(3)}[\mathbf{O}'_j   \text{ScaleRot}(\sqrt{\bar{\alpha}^t}, \mathbf{O}_j^0), 1 - \bar{\alpha}^t]$
	Prior:	$\mathbf{O}_j^T \sim \text{Uniform}[\text{SO}(3)]$
	Generative:	$p_\theta(\mathbf{O}'_j^{t-1}   \mathcal{R}^t, C) = \mathcal{IG}_{\text{SO}(3)}\{\mathbf{O}'_j^{t-1}   H_\theta(\mathcal{R}^t, C)[j], 1 - \bar{\alpha}^t\}$
	Objective:	$L_{\text{ori}}^t = \mathbb{E}[\frac{1}{m} \sum_j \ \mathbf{O}_j^0\ ^T H_\theta(\mathcal{R}^t, C)[j] - \mathbf{I}\ _F^2]$

$t$  are defined as

$$q(s_j^{t-1} | s_j^t, s_j^0) = \text{Multinomial}\{[\alpha^t \cdot \text{onehot}(s_j^t) + (1 - \alpha^t) \bar{\mathbf{p}}] \odot [\bar{\alpha}^{t-1} \cdot \text{onehot}(s_j^0) + (1 - \bar{\alpha}^{t-1}) \bar{\mathbf{p}}]\}. \quad (2)$$

Here,  $\bar{\alpha}^t = \prod_{\tau=1}^t \alpha^\tau = \prod_{\tau=1}^t (1 - \beta^\tau)$ , with  $\beta^t$  denoting the cosine variance schedule (ranging from 0 to 1 as  $t$  increases from 0 to  $T$ ).  $s_j^0$  is the amino acid type approximated by the neural network model during the generative diffusion process. This posterior enforces resampling to rely more on the property-aware prior at the start of the process ( $t \rightarrow T$ ) and more on the previously sampled amino acid  $s_j^t$  towards the end of the process ( $t \rightarrow 0$ ). Note that we need to divide the posterior probabilities by their sum to ensure they add up to 1.

Although this approach can accommodate any amino acid-related property, in this study, we focused on the hydropathy score [37] as a proxy for solubility and aggregation. We use this score as it effectively describes the hydrophilic or hydrophobic nature of amino acids on a single scale. Consequently, negative values indicate greater hydrophilicity, while positive hydropathy values indicate greater hydrophobicity. In the Supplemental Material, Fig. S1 [38] presents the hydropathy score for each amino acid type and its translation to probabilities (for different values of  $b$ ). Here, hydrophilic amino acids (i.e., with low hydropathy scores) are assigned higher probabilities, and vice versa.

## 2. Sampling by property

Given that not all properties can be defined solely at the amino acid level, we introduce a second strategy incorporating properties associated with the antibody sequence and structure. By leveraging the stochastic nature of DDPMs, in this approach we sample  $N$  times at each generation timestep and

then select the sample with the most desirable property value, as shown in Fig. 1(d). For instance, we choose the minimum if we aim to minimize the property value. When multiple properties are considered, we opt for the sample with the minimum sum of all property values (known as the Pareto optimal solution [39]). Alternatively, one could consider using a weighted sum of properties to allow for different importance ratios among properties. In a more flexible version, we convert the  $N$  property values into probabilities through the softmax function and then sample the next timestep from this distribution. Here, the assumption is that all  $\mathcal{R}$  samples generated at the same step in the process are equally valid in terms of  $(\mathbf{s}, \mathbf{x}, \mathbf{O})$ .

For the sampling by property approach, we guide our model using both the hydropathy score and the folding energy ( $\Delta G$ ). To compute the difference in folding energy, we employ the  $\Delta \Delta G$  predictor from Ref. [16], which relies on a graph convolutional network (GCN) model to predict the energy difference between the reference and generated antibodies. To obtain predicted  $\Delta \Delta G$  values for the  $n$ -sampled CDR, we feed the model with the amino acid sequence and  $C_\alpha$  atom positions at the current generation timestep ( $s^{t-1}, \mathbf{x}^{t-1}$ ), as well as those from the previous timestep ( $s^t, \mathbf{x}^t$ ). The pseudocode for the sampling by property approach can be found in Algorithm 1.

## C. Benchmark dataset and trained model

To benchmark our guided approaches, we employ the test set described in Ref. [8], which comprises 19 antibody-antigen complexes sourced from the SABDab database [29]. The CDR-H3 sequences of the test antibodies share a maximum of 50% sequence identity with each other and with the training data. The test set includes protein antigens from various pathogens, including influenza and SARS-CoV-2.

ALGORITHM 1. Generative diffusion process with the sampling by property approach combining two properties (hydropathy score and predicted  $\Delta\Delta G$ ).

---



---

```

1:  $s_j^T \sim \text{Uniform}(20)$ ,  $\mathbf{x}_j^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_3)$ ,  $\mathbf{O}_j^T \sim \text{Uniform}(\text{SO}(3))$ ,  $\forall j = l + 1, \dots, l + m$ 
2: for  $t = T : 0$  do
3:    $\forall j$ ,  $s_j^0 = F_\theta(\mathcal{R}^t, C)[j]$ ,  $\tilde{\epsilon}_j^{t-1} = G_\theta(\mathcal{R}^t, C)[j]$ ,  $\tilde{\mathbf{O}}_j^{t-1} = H_\theta(\mathcal{R}^t, C)[j]$ 
                                     // run neural network models
4:   for  $n = 1 : N$  do
5:      $\forall j$ ,  $s_j^{t-1}[n] \sim q(s_j^{t-1}|s_j^t, \tilde{s}_j^0)$  // sample amino acid type from posterior [Eq. (2)]
6:      $\mathbf{z}[n] \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  // sample noise for  $C_\alpha$  position
7:      $\forall j$ ,  $\mathbf{x}_j^{t-1}[n] = \frac{1}{\sqrt{\alpha^t}}(\mathbf{x}_j^t - \frac{1-\alpha^t}{\sqrt{1-\alpha^t}}\tilde{\epsilon}_j^{t-1}) + \sigma^t \mathbf{z}[n]$  // compute  $C_\alpha$  position
8:      $\mathbf{E}[n] \sim \mathcal{IG}_{\text{SO}(3)}(\mathbf{I}, \sigma^2)$  // sample noise for orientation
9:      $\forall j$ ,  $\mathbf{O}_j^{t-1}[n] = \mathbf{E}[n]\tilde{\mathbf{O}}_j^{t-1}$  // compute orientation
10:     $h[n] = \sum_{j=l+1}^{l+m} f_{\text{hydro}}(s_j^{t-1}[n])$  // compute hydropathy score [37]
11:     $g[n] = \text{GCN}_\phi(\{s^{t-1}[n], \mathbf{x}^{t-1}[n]\}, (\mathbf{s}^t, \mathbf{x}^t))$  // run  $\Delta\Delta G$  predictor  $\text{GCN}_\phi$  [16]
12:   end for
13:    $\hat{n} \leftarrow \arg \min_n ([h[1], \dots, h[n]] + [g[1], \dots, g[n]])$  // select sample based on properties
14:    $\{s^{t-1}, \mathbf{x}^{t-1}, \mathbf{O}^{t-1}\} \leftarrow \{s^{t-1}[\hat{n}], \mathbf{x}^{t-1}[\hat{n}], \mathbf{O}^{t-1}[\hat{n}]\}$ 
15: end for
16: return  $\forall j$ ,  $\{s_j^0, \mathbf{x}_j^0, \mathbf{O}_j^0\}$ 

```

---



---

To guide the generation diffusion process, we leverage the `codesign_single` model from DiffAb, which has been trained to generate all CDRs, one at a time randomly selected for each training sample. Using this model, we design single CDRs from random values given the rest of the antibody-antigen complex.

#### D. Evaluation

For each test complex, we generate 100 designs for each of the six CDRs through  $T = 100$  timesteps of generation, each one maintaining the same length as the reference test CDR. We evaluate the designs using the following metrics.

(i) The AAR (amino acid recovery) measures the sequence identity between the reference and generated CDR sequences.

(ii) The root-mean-square deviation (RMSD) computes the  $C_\alpha$  atom distance between the reference and generated CDR structures.

(iii) The hydropathy score averages the hydropathy values over the generated CDR sequences. Note that negative scores indicate hydrophilicity, while positive scores indicate hydrophobicity.

(iv) The predicted  $\Delta\Delta G$  [16] measures the difference in folding energy ( $\Delta G$ ) between the reference and generated CDRs, considering atoms (N,  $C_\alpha$ , C, O) after reconstructing the backbone structure from  $C_\alpha$  atom positions and orientations (see Ref. [8]).

We use the predicted  $\Delta\Delta G$  as it is computationally more efficient and has moderate to high correlation with experimental measures of antibody energy upon mutations [16]. For AAR, higher values are preferable, while lower values are desired for RMSD, hydropathy score, and predicted  $\Delta\Delta G$ . Note that we aim to generate CDRs with improved property values (low hydropathy score and predicted  $\Delta\Delta G$ ) without deteriorating the structural integrity (we expect slight deviations in AAR and RMSD to the reference CDR).

## IV. RESULTS AND DISCUSSION

### A. Guidance on properties is effective

We assess our property-aware prior approach using the hydropathy score, with  $b = 0.8$  (see Fig. S2 [38] for the impact of  $b$  on the hydropathy score of the final designs). The sampling by property approach is tested for both the hydropathy score and  $\Delta\Delta G$ . As suggested by Fig. S3 [38], we select the sample with minimum predicted  $\Delta\Delta G$  in  $N = 20$  samples at each generation timestep. In Fig. S3 [38], we empirically show that selecting the most optimal sample (i.e., the one with minimum predicted  $\Delta\Delta G$ ) at every timestep produces the best results. Additionally, the metrics do not improve when increasing the number of samples  $N$  at every timestep. These findings contradict the results in Ref. [36], where improvements are observed when increasing the number of particles. Also, they used all particles at every timestep to resample the next set, whereas we select the most optimal one, resulting in a single guided trajectory. This approach also facilitates the joint optimization of multiple properties by selecting the Pareto optimal solution. In addition to analyzing individual properties, we test the combination of both properties in two ways: sampling by  $\Delta\Delta G$  with a hydropathy-aware prior, and jointly sampling by  $\Delta\Delta G$  and hydropathy score (selecting the minimum of the unweighted sum, see Algorithm 1).

Figure 2 illustrates the performance metrics for intermediate CDR-H3 designs at every 10 timesteps of generation, comparing the unguided generation from the original DiffAb model [Unguided], which serves as our baseline to all our guided models:

(i) hydropathy-aware prior with  $b = 0.8$  [Hydro. ( $b = 0.8$ )],

(ii) sampling by hydropathy (minimum over 20 samples) [Hydro. (20 Min)],

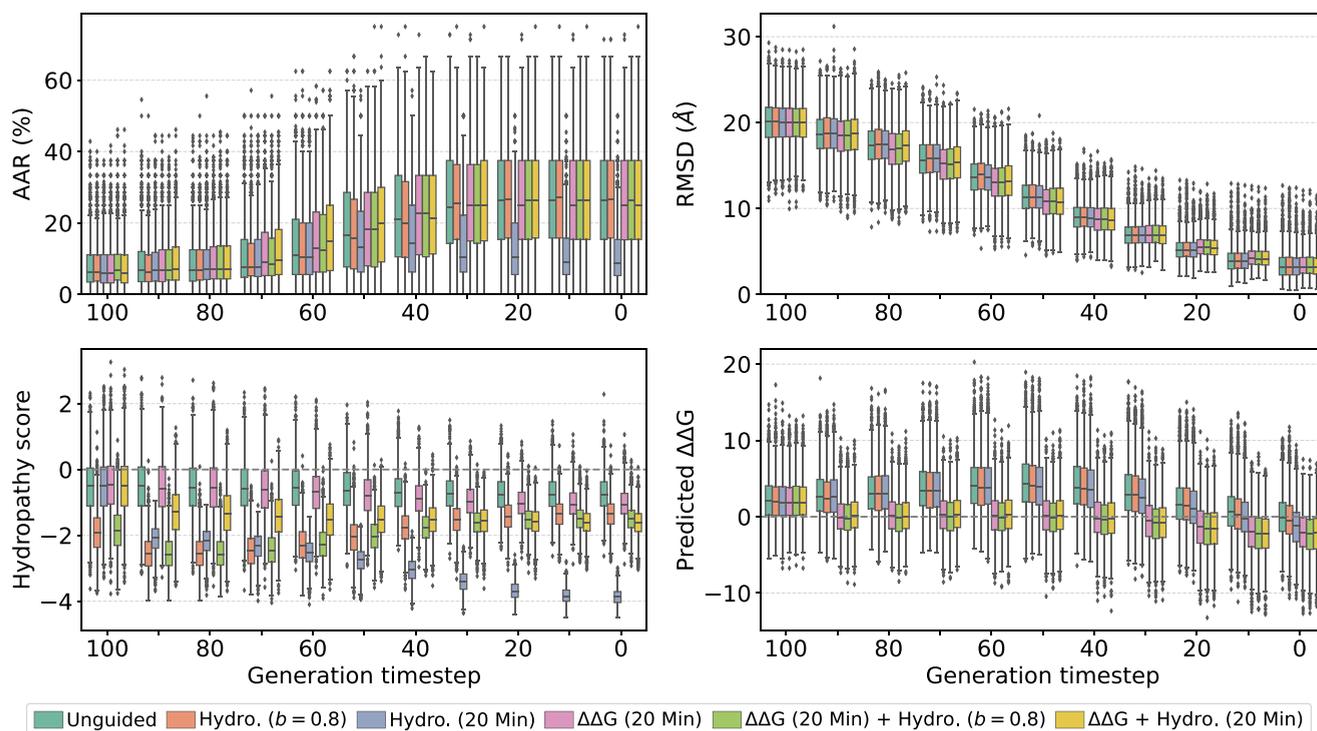


FIG. 2. Per-timestep metrics on the 19 test complexes (design CDR-H3). The boxplots represent the distribution of metric values (AAR, RMSD, hydropathy score, and predicted  $\Delta\Delta G$ ) over 100 designed CDRs for each test complex. Here, we compare the unguided mode with different property-guided models: hydropathy-aware prior, sampling by hydropathy or  $\Delta\Delta G$ , and combinations of both.

(iii) sampling by  $\Delta\Delta G$  (minimum over 20 samples) [ $\Delta\Delta G$  (20 Min)],

(iv) sampling by  $\Delta\Delta G$  with a hydropathy-aware prior [ $\Delta\Delta G$  (20 Min) + Hydro. ( $b = 0.8$ )], and

(v) jointly sampling by  $\Delta\Delta G$  and hydropathy [ $\Delta\Delta G$  + Hydro. (20 Min)].

As observed in Fig. 2, sampling by hydropathy results in a larger change in the hydropathy score and the AAR, compared to the unguided mode, than using a hydropathy-aware prior, even with a high value of  $b$ . This indicates that the most substantial differences between the generated CDR sequences and the reference CDR occur when sampling by hydropathy. Furthermore, we note that, in comparison to the unguided mode, exclusive sampling by  $\Delta\Delta G$  improves the hydropathy score, whereas exclusive sampling by hydropathy does the same for the predicted  $\Delta\Delta G$ . When both properties are combined, the most favorable outcomes are achieved, with the majority of designs exhibiting hydropathy scores and predicted  $\Delta\Delta G$  values below 0. This implies that the designed CDRs have more hydrophilic profiles and improved energy relative to the reference CDRs. These results are further supported by Mann-Whitney statistical tests, revealing significant differences in the final metric distributions (at  $t = 0$ ) across different models for the entire test set (see Fig. S4 [38]). Meanwhile, the values of AAR and RMSD are consistent across models, which is desirable to avoid significant deviations from the reference CDR. While the sequence similarities within guided designs deviate from the unguided ones (as expected, Fig. S5 [38]), the RMSD values remain close (Fig. S6 [38]).

These observations apply to all other CDRs as well. Figure S7 [38] displays per-timestep metrics and Table S1 [38] includes the performance metrics for the final designs of all CDRs. We can see that AAR decreases and RMSD increases slightly after guidance, which is expected. Furthermore, each CDR exhibits some variations in response to the guidance. For CDR-H2 and L2, sampling by hydropathy results in significantly lower AAR, leading to better predicted  $\Delta\Delta G$  compared to sampling by  $\Delta\Delta G$ .

## B. Amino acid composition changes with hydropathy guidance

Hydropathy guidance aims to design CDR sequences containing hydrophilic amino acid types (with lower hydropathy values) without deteriorating the target binding affinities. This effect is illustrated in the amino acid compositions in Fig. 3. We note that sampling by hydropathy causes the most significant shift in the final amino acid distribution toward arginine (R) and aspartic acid (D), two of the most hydrophilic amino acids. Furthermore, this approach eliminates most of the hydrophobic amino acids. Using the hydropathy-aware prior, the effect is not as strong, primarily because the model relies less on the prior towards the end of the generation process. Sampling by  $\Delta\Delta G$ , whether alone or in combination with hydropathy, also increases the number of hydrophilic amino acids, such as tyrosine (Y). These results, along with Table S1 of the Supplemental Material [38], indicate a correlation between both properties. Improved hydropathy profiles (more negative and thus more hydrophilic) lead to greater exposure of solvent-accessible surface areas within the antibody

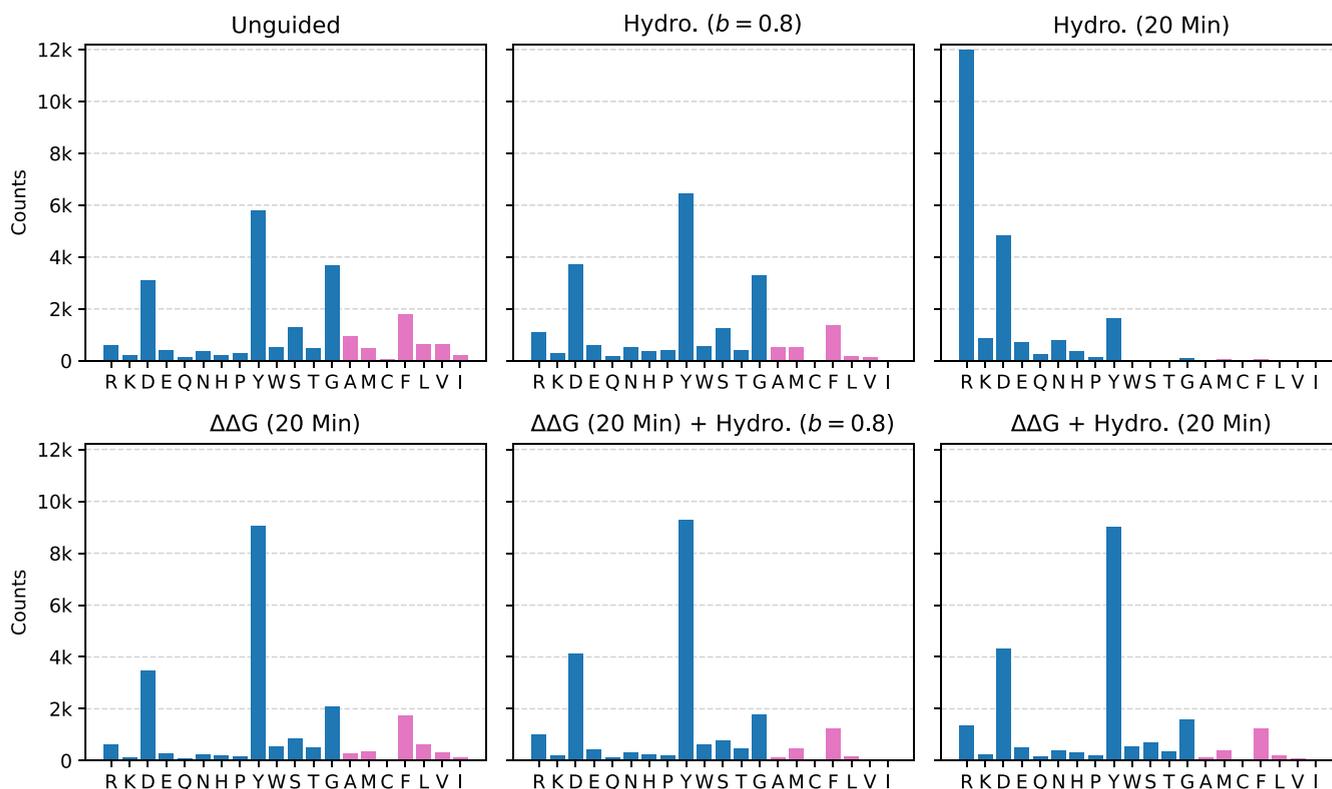


FIG. 3. Amino acid composition for the 19 test complexes (100 CDR-H3 designs each). Amino acid types are ordered by ascending hydropathy score, while counts are colored by negative (blue) or positive (pink) hydropathy. Note that negative scores indicate hydrophilicity, while positive scores indicate hydrophobicity.

CDRs, facilitating the accessibility to antigens. While exposed hydrophobic amino acids generally promote protein-protein interaction [40], in the context of antibodies, extensive hydrophobic patches can result in aggregation with both self and other molecules [41], primarily due to nonspecific interactions [42].

### C. Energy distributions shift with $\Delta\Delta G$ guidance, even after relaxation

The objective of  $\Delta\Delta G$  guidance is to generate CDR loops with enhanced folding stability, potentially improving

antibody-antigen binding. Figure 4(a) shows the relationship between the predicted  $\Delta\Delta G$  and the hydropathy score for the final CDR-H3 designs, revealing the positive correlation between these two properties. Considering that lower values are desirable for both properties, we calculated the Pareto frontiers for the three approaches. Notably, we observe that the three frontiers are clearly separated, with the guided approaches exhibiting a trend towards the lowest values. Thus, they outperform a naive filter on top-scoring samples from the unguided model. The most favorable Pareto solutions are obtained when jointly sampling by  $\Delta\Delta G$  and hydropathy. This behavior is also observed for all other test complexes

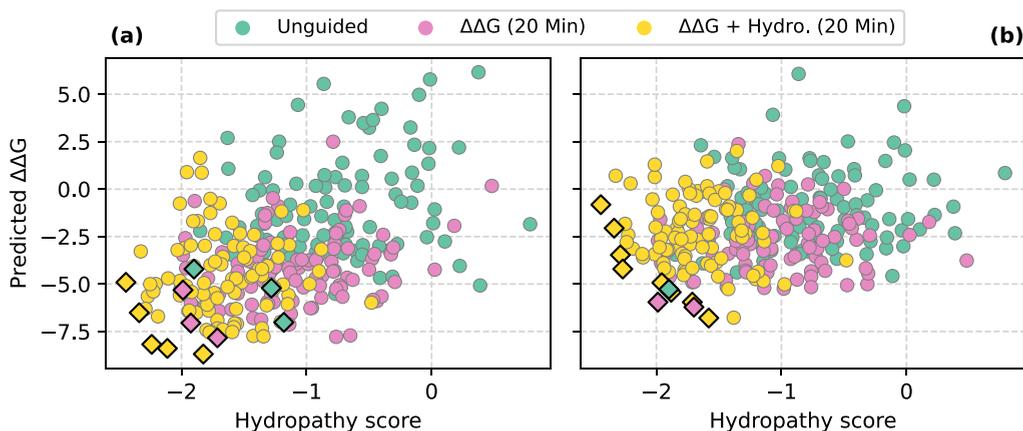


FIG. 4. Distribution of hydropathy scores and predicted  $\Delta\Delta G$  for test complex 7chf\_A\_B\_R (design CDR-H3), (a) before and (b) after Rosetta relaxation. The highlighted points (diamond markers) correspond to the Pareto optimal solutions.

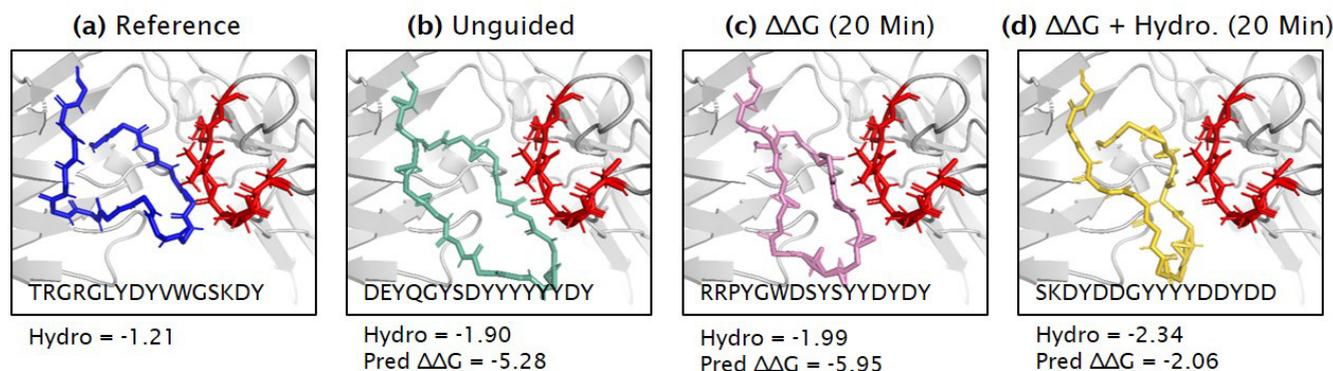


FIG. 5. Sequence-structure designs from the Pareto frontier over hydropathy score and predicted  $\Delta\Delta G$  after Rosetta relaxation (test complex 7chf\_A\_B\_R, design CDR-H3). The antigen epitope is displayed in red.

(see Fig. S8 [38]). The empirical run-time comparisons for the test complex in Fig. 4 can be found in Table S2 of the Supplemental Material [38].

We then reconstruct the side-chain atoms using Rosetta [43], resulting in a refined structure. Figure 4(b) shows that even though the three Pareto frontiers become closer after relaxation, the distributions of the guided models are still nearer to the lower-left part than the unguided mode. Furthermore, compared to prerelaxation, we attain a larger number of Pareto optimal solutions for the combined sampling by  $\Delta\Delta G$  and hydropathy. The performance metrics for both the pre- and postrelaxed designs are in Table S3 of the Supplemental Material [38]. For this example, we also visualize the resulting CDR-H3 structures in relation to the antigen epitope. We select those designs that are present in both Pareto frontiers, before and after relaxation. As observed in Fig. 5, different CDR sequences lead to similar structures compared to the reference, but exhibit improved hydropathy and predicted  $\Delta\Delta G$  values.

## V. CONCLUSION

We successfully developed two methodologically distinct strategies for guiding diffusion models in the field of antibody design. Most notably, we can guide the generative process toward novel CDR designs with desired properties. An advantage of our approaches is their pure integration into the generative diffusion process, eliminating the need for retraining the models. While we assess our approaches

using two specific properties, hydropathy (accounting for the hydrophilic or hydrophobic nature of amino acids) and folding energy, our methodological framework can seamlessly accommodate any desired property derived from the amino acid sequence, the structure, or both. We also demonstrate that our guided approaches enable the optimization of multiple properties at once, leading to a better set of Pareto optimal solutions. While empirical results support our modeling choices, exploring a mathematical foundation remains of interest to better understand the validity and biases in the designs introduced by our sampling approach, compared to similar approaches [21,36]. Finally, while our computational models and metrics provide valuable insights into the potential improvements of the designed CDRs, they are limited in capturing the complexity of real biological systems. Therefore, experimental validation in a wet laboratory is necessary to verify that the predicted changes in solubility, aggregation, and folding/binding energy values translate to actual improvements in practice.

The code used in this research is available at GitHub [44].

## ACKNOWLEDGMENTS

The authors thank Dr. Chirag Raman, Mahdi Naderibeni, Colm Seale, Gabriel Vogel, and Paolo Sortino for their valuable input and discussions regarding diffusion models; the Delft Bioinformatics Lab for their feedback on the project; and the reviewers of the Machine Learning in Structural Biology (MLSB) workshop who helped strengthen the paper.

- [1] C. A. Janeway *et al.*, *Immunobiology: The Immune System in Health and Disease* (Taylor & Francis Group UK: Garland Science, New York, 2001).
- [2] J. Adolf-Bryfogle, O. Kalyuzhnyi, M. Kubitz, B. D. Weitzner, X. Hu, Y. Adachi, W. R. Schief, and R. L. Dunbrack, Jr., RosettaAntibodyDesign (RABD): A general framework for computational antibody design, *PLoS Comput. Biol.* **14**, e1006112 (2018).
- [3] K. Saka, T. Kakuzaki, S. Metsugi, D. Kashiwagi, K. Yoshida, M. Wada, H. Tsunoda, and R. Teramoto, Antibody de-

sign using LSTM based deep generative model from phage display library for affinity maturation, *Sci. Rep.* **11**, 5852 (2021).

- [4] R. W. Shuai, J. A. Ruffolo, and J. J. Gray, IgLM: Infilling language modeling for antibody sequence design, *Cell Syst.* **14**, 979 (2023).
- [5] J. A. Ruffolo, L.-S. Chu, S. P. Mahajan, and J. J. Gray, Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies, *Nat. Commun.* **14**, 2389 (2023).

- [6] J. Wu, F. Wu, B. Jiang, W. Liu, and P. Zhao, tFold-ab: fast and accurate antibody structure prediction without sequence homologs, in *Machine Learning for Structural Biology Workshop* (NeurIPS, New Orleans, LA, 2022).
- [7] W. Jin, J. Wohlwend, R. Barzilay, and T. S. Jaakkola, Iterative refinement graph neural network for antibody sequence-structure co-design, in *International Conference on Learning Representations* (ICLR, 2022).
- [8] S. Luo, Y. Su, X. Peng, S. Wang, J. Peng, and J. Ma, Antigen-specific antibody design and optimization with diffusion-based generative models for protein structures, in *Advances in Neural Information Processing Systems*, NeurIPS (Curran Associates Inc., New Orleans, LA, 2022), Vol. 35, pp. 9754–9767.
- [9] X. Kong, W. Huang, and Y. Liu, End-to-end full-atom antibody design, in *International Conference on Machine Learning* (JMLR.org, Honolulu, Hawaii, 2023), pp. 17409–17429.
- [10] R. Akbar, H. Bashour, P. Rawat, P. A. Robert, E. Smorodina, T.-S. Cotet, K. Flem-Karlsen, R. Frank, B. B. Mehta, M. H. Vu *et al.*, Progress and challenges for the machine learning-based design of fit-for-purpose monoclonal antibodies, *mAbs* **14**, 2008790 (2022).
- [11] J. Kim, M. McFee, Q. Fang, O. Abdin, and P. M. Kim, Computational and artificial intelligence-based methods for antibody development, *Trends Pharmacol. Sci.* **44**, 175 (2023).
- [12] M. I. J. Raybould and C. M. Deane, The therapeutic antibody profiler for computational developability assessment, in *Therapeutic Antibodies: Methods and Protocols*, edited by G. Houen (Springer, Berlin, 2022), pp. 115–125.
- [13] M. B. Swindells, C. T. Porter, M. Couch, J. Hurst, K. R. Abhinandan, J. H. Nielsen, G. Macindoe, J. Hetherington, and A. C. R. Martin, abYsis: Integrated antibody sequence and structure–management, analysis, and prediction, *J. Mol. Biol.* **429**, 356 (2017).
- [14] R. Zambrano, M. Jamroz, A. Szczasiuk, J. Pujols, S. Kmiecik, and S. Ventura, AGGRESKAN3D (A3D): Server for prediction of aggregation properties of protein structures, *Nucleic Acids Res.* **43**, W306 (2015).
- [15] D. Prihoda, J. Maamary, A. Waight, V. Juan, L. Fayadat-Dilman, D. Svozil, and D. A. Bitton, BioPhi: A platform for antibody design, humanization, and humanness evaluation based on natural antibody repertoires and deep learning, *mAbs* **14**, e2020203 (2022).
- [16] S. Shan, S. Luo, Z. Yang, J. Hong, Y. Su, F. Ding, L. Fu, C. Li, P. Chen, J. Ma *et al.*, Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization, *Proc. Natl. Acad. Sci. USA* **119**, e2122954119 (2022).
- [17] J. Ho, A. Jain, and P. Abbeel, Denoising diffusion probabilistic models, in *Advances in Neural Information Processing Systems*, Vol. 33, NeurIPS (Curran Associates Inc., Vancouver, BC, Canada, 2020), pp. 6840–6851.
- [18] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-based generative modeling through stochastic differential equations, in *International Conference on Learning Representations* (ICLR, 2021).
- [19] N. Anand and T. Achim, Protein structure and sequence generation with equivariant denoising diffusion probabilistic models, [arXiv:2205.15019](https://arxiv.org/abs/2205.15019).
- [20] J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles *et al.*, De novo design of protein structure and function with RFdiffusion, *Nature (London)* **620**, 1089 (2023).
- [21] B. L. Trippe, J. Yim, D. Tischer, D. Baker, T. Broderick, R. Barzilay, and T. S. Jaakkola, Diffusion probabilistic modeling of protein backbones in 3D for the motif-scaffolding problem, in *The Eleventh International Conference on Learning Representations* (ICLR, 2022).
- [22] A. Bansal, H.-M. Chu, A. Schwarzschild, S. Sengupta, M. Goldblum, J. Geiping, and T. Goldstein, Universal guidance for diffusion models, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada* (IEEE, Piscataway, NJ, 2023), pp. 843–852.
- [23] N. Ferruz, S. Schmidt, and B. Höcker, ProtGPT2 is a deep unsupervised language model for protein design, *Nat. Commun.* **13**, 4348 (2022).
- [24] J. B. Ingraham, M. Baranov, Z. Costello, K. W. Barber, W. Wang, A. Ismail, V. Frappier, D. M. Lord, C. Ng-Thow-Hing, E. R. Van Vlack *et al.*, Illuminating protein space with a programmable generative model, *Nature (London)* **623**, 1070 (2023).
- [25] K. E. Wu, K. K. Yang, R. van den Berg, S. Alamdari, J. Y. Zou, A. X. Lu, and A. P. Amini, Protein structure generation via folding diffusion, *Nat. Commun.* **15**, 1059 (2024).
- [26] I. Melnyk, V. Chenthamarakshan, P.-Y. Chen, P. Das, A. Dhurandhar, I. Padhi, and D. Das, Reprogramming pretrained language models for antibody sequence infilling, in *International Conference on Machine Learning* (JMLR.org, Honolulu, Hawaii, 2023), pp. 24398–24419.
- [27] N. Gruver, S. Stanton, N. Frey, T. G. J. Rudner, I. Hotzel, J. Lafrance-Vanasse, A. Rajpal, K. Cho, and A. G. Wilson, Protein design with guided discrete diffusion, in *Advances in Neural Information Processing Systems*, Vol. 36, NeurIPS (Curran Associates Inc., New Orleans, LA, 2023), pp. 12489–12517.
- [28] R. R. Eguchi, C. A. Choe, and P.-S. Huang, Ig-VAE: Generative modeling of protein structure by direct 3D coordinate generation, *PLoS Comput. Biol.* **18**, e1010271 (2022).
- [29] J. Dunbar, K. Krawczyk, J. Leem, T. Baker, A. Fuchs, G. Georges, J. Shi, and C. M. Deane, SABDab: The structural antibody database, *Nucleic Acids Res.* **42**, D1140 (2014).
- [30] K. Gao, L. Wu, J. Zhu, T. Peng, Y. Xia, L. He, S. Xie, T. Qin, H. Liu, K. He, and T.-Y. Liu, Incorporating pre-training paradigm for antibody sequence-structure co-design, [arXiv:2211.08406](https://arxiv.org/abs/2211.08406).
- [31] W. Jin, R. Barzilay, and T. Jaakkola, Antibody-antigen docking and design via hierarchical structure refinement, in *International Conference on Machine Learning* (PMLR, Baltimore, MD, 2022), pp. 10217–10227.
- [32] X. Kong, W. Huang, and Y. Liu, Conditional antibody design as 3D equivariant graph translation, in *The Eleventh International Conference on Learning Representations* (ICLR, 2022).
- [33] K. Martinkus, J. Ludwiczak, W.-C. Liang, J. Lafrance-Vanasse, I. Hotzel, A. Rajpal, Y. Wu, K. Cho, R. Bonneau, V. Gligorijevic *et al.*, AbDiffuser: Full-atom generation of in-vitro functioning antibodies, in *Advances in Neural Information Processing Systems*, NeurIPS (Curran Associates Inc., New Orleans, LA, 2023), Vol. 36, pp. 40729–40759.
- [34] P. Dhariwal and A. Nichol, Diffusion models beat GANs on image synthesis, in *Advances in Neural Information Processing Systems*, NeurIPS (Curran Associates Inc., 2021), Vol. 34, pp. 8780–8794.

- [35] S. L. Lisanza, J. M. Gershon, S. Tipps, L. Arnoldt, S. Hendel, J. N. Sims, X. Li, and D. Baker, Joint generation of protein sequence and structure with RoseTTAFold sequence space diffusion, bioRxiv, <https://doi.org/10.1101/2023.05.08.539766>.
- [36] L. Wu, B. Trippe, C. Naeseth, D. Blei, and J. P. Cunningham, Practical and asymptotically exact conditional sampling in diffusion models, in *Advances in Neural Information Processing Systems*, NeurIPS (Curran Associates Inc., New Orleans, LA, 2023), Vol. 36, pp. 31372–31403.
- [37] J. Kyte and R. F. Doolittle, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* **157**, 105 (1982).
- [38] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PRXLife.2.033012> for additional figures and tables.
- [39] H. P. Benson, Multi-objective optimization: Pareto optimal solutions, properties, in *Encyclopedia of Optimization*, edited by C. A. Floudas and P. M. Pardalos (Springer US, Boston, MA, 2009), pp. 2478–2481.
- [40] J. H. M. van Gils, D. Gogishvili, J. van Eck, R. Bouwmeester, E. van Dijk, and S. Abeln, How sticky are our proteins? Quantifying hydrophobicity of the human proteome, *Bioinf. Adv.* **2**, vbac002 (2022).
- [41] F. Waibl, M. L. Fernández-Quintero, F. S. Wedl, H. Kettenberger, G. Georges, and K. R. Liedl, Comparison of hydrophobicity scales for predicting biophysical properties of antibodies, *Front. Mol. Biosci.* **9**, 960194 (2022).
- [42] H. Ausserwöger, M. M. Schneider, T. W. Herling, P. Arosio, G. Invernizzi, T. P. J. Knowles, and N. Lorenzen, Non-specificity as the sticky problem in therapeutic antibody development, *Nat. Rev. Chem.* **6**, 844 (2022).
- [43] R. F. Alford, A. Leaver-Fay, J. R. Jeliazkov, M. J. O’Meara, F. P. DiMaio, H. Park, M. V. Shapovalov, P. Douglas Renfrew, V. K. Mulligan, K. Kappel *et al.*, The Rosetta all-atom energy function for macromolecular modeling and design, *J. Chem. Theory Comput.* **13**, 3031 (2017).
- [44] GitHub, <https://github.com/amelvim/antibody-diffusion-properties>.