D-Flow: Multi-modality Flow Matching for D-peptide Design

Fang Wu*
Stanford University
CA, USA

Tinson Xu*
The University of Chicago
IL, USA

Shuting Jin*
Wuhan University of Technology
Hubei, China

Xiangru Tang
Yale University
CT, USA

Zerui Xu
The University of Chicago
IL, USA

James Zou
Stanford University
CA, USA

Brian Hie†
Stanford University
CA, USA

Abstract

Proteins play crucial roles in biological processes, with therapeutic peptides emerging as promising pharmaceutical agents. They allow new possibilities to leverage target binding sites that were previously undruggable. While deep learning (DL) has advanced peptide discovery, generating D-proteins composed of D-amino acids remains challenging due to the scarcity of natural examples. This paper proposes D-Flow, a full-atom flow-based framework for de novo D-peptide design. D-Flow is conditioned on receptor binding and utilizes a comprehensive representation of peptide structure, incorporating backbone frames, side-chain angles, and discrete amino acid types. A mirror-image algorithm is implemented to address the lack of training data for D-proteins, which converts L-receptors' chirality. Furthermore, we enhance D-Flow's capacity by integrating large protein language models (PLMs) with structural awareness through a lightweight structural adapter. A two-stage training pipeline and a controlling toolkit also enable D-Flow to transition from general protein design to targeted binder design while preserving pretraining knowledge. Extensive experimental results on the PepMerge benchmark demonstrate D-Flow's effectiveness, particularly in developing peptides with entire D-residues. This approach represents a significant advancement in computational D-peptide design, offering unique opportunities for bioorthogonal and stable molecular tools and diagnostics. The code is available in https://github.com/smiles724/PeptideDesign

1. Introduction

Proteins are the building blocks of life and make essential contributions to nearly all fundamental biological processes in the cell. They fold into specific 3D conformations to perform distinct functionalities [27, 37]. Remarkably, therapeutic peptides, compromising a limited number of well-ordered residues, are single-chain proteins and an irregular class of pharmaceutical agents [69]. Peptide drugs occupy a unique chemical and pharmacological space between small and large molecules [42]. Due to this specialty, increasing efforts have adopted deep learning (DL) algorithms to facilitate their discovery [36, 41]. Notably, unbound peptide chains typically have high free energy and entropy, resulting in unstable conformations. In contrast, they can elicit pharmacological effects upon binding to specific receptors, forming stable complexes where both the receptor and ligand adopt equilibrium structures [31].

Diffusion or flow models [30, 44, 60], a recent generative family, are revolutionizing the field of image generation [17], text-to-image generation [50], image segmentation [7], and also finding rapid applications in other domains such as video generation [29], and 3D generation [59]. Notable initiatives [71, 75] have explored the potential of diffusion or flow models in the context of drug discovery. These models have been applied to various protein representations, including carbon-alpha only structures [67], torsion angles [78], and the SE(3) backbone frame representation [80], and diverse scenes such as molecular design [28], antibody engineering [46, 48], *de novo* protein design [79], and peptide discovery [55] as well.

Despite emerging interest in developing 3D diffusion or flow to peptides, little attention has been paid to D-protein generation. It is worth mentioning that though every cell in the human body contains proteins, their cornerstone, 20 categories of amino acids, can exist in two stereoisomers: L (levo) and D (dextro), which are non-

---
*Equal contributions.
†Corresponding author: brianhie@stanford.edu

superimposable mirror images of each other [38], and this chirality characteristic is determined by the orientation of the functional groups around the alpha carbon [24]. D-proteins are protein molecules whose polypeptide chains consist of D-amino acids and the achiral amino acid glycine. They can form specific heterochiral protein-protein interactions (PPIs) with natural L-protein targets and possess remarkable potential as therapeutics, and diagnostics because of their high bioorthogonality and stability [63]. However, in nature, the vast majority of peptides and proteins are made from L-amino acids, which poses a great challenge for DL mechanisms, as no training data is accessible.

To overcome this obstacle, we introduce a mirror-imaged multi-modality model, named D-Flow (see Fig. 1), to produce D-peptides. D-Flow characterizes the peptide sequence and structure using rigid backbone frames within the SE(3) manifold, the side-chain angles on the high-dimensional torus, and the discrete amino acid types on the categorical space [10]. Each modality has an analytical flow, and they jointly capture the distribution of full-atom peptide structure conditioned on the receptor. During the inference phase, a novel mirror-image algorithm is adopted to transform the L-receptor and convert its natural chirality. Moreover, as datasets containing complete 3D structures of peptide complexes are orders of magnitude smaller than sequence-only databases [77], the scarcity of high-quality peptide-receptor pairs inevitably constrains the performance of DL approaches. As a resolution, we propose to empower the flow model with large protein language models (PLMs) and a two-stage training pipeline. Specifically, a structural surgery is implemented on PLMs using a lightweight structural adapter [84]. Consequently, PLMs are endowed with structural awareness and can facilitate the regression of residue categories given intermediate-state conformation information. Secondly, D-Flow embraces a controlling toolkit [82] to achieve awareness of target proteins and transfer from the general protein design task to the binder design task with the minimum loss of the pretraining knowledge. Comprehensive experiments on the PepMerge benchmark [41] verify the effectiveness of D-Flow, particularly, for its breakthrough in generating pure D-peptides. Besides, it also outpasses existing DL baselines in generating L-peptides by a large margin. Our study opens up avenues for systematically exploring the DL-based mirror-image protein universe, paving the way for a wide range of design applications targeting L-proteins.

2. Related Works

2.1. Diffusion and Flow for Protein Design

Generative models have displayed incredible promise in designing novel protein structures for custom functions, with much of this progress driven by advancements in diffusion models [17, 30, 60]. Among distinct protein representations, the frame representation has achieved state-of-the-art performance in tasks like *de novo* backbone design and motif scaffolds [67], as exemplified by RFdiffusion [71]. Recently, flow methods [44, 64] appear as an alternative to diffusion models, offering a deterministic approach by eliminating stochasticity from the sampling process. They have been extended to Riemannian manifolds [12, 44] and possess an attractive connection to optimal transport (OT). The linear interpolation schedule of flow results in more direct sampling trajectories, reducing the number of integration steps needed [44]. In the computer vision field, flow matching (FM) has already demonstrated comparable results to diffusion models with more computational efficiency [53].

The past few years have witnessed a growth in applying diffusion and flow models [35, 46, 48, 76] for antibody design, a special protein family in the immune system to resist antigens, mainly focusing on inpainting complementarity-determining regions (CDRs) at the interface between the antigen and the framework. Lately, non-trivial efforts have been put into developing DL algorithms for peptides. Li et al. [41] encompasses the dynamics of side-chain angles. Contrarily, Lin et al. [42] relies on RDE [47] to complete the solution to full-atom design. Kong et al. [36] presents a full-atom diffusion on the latent geometric space learned by VAE. Nevertheless, none considers the extension of sequence models (*e.g.*, PLMs) and multi-stage training paradigms to alleviate the shortage of available binding structures.

2.2. D-peptide Technology

Proteins are built from chiral molecules, specifically amino acids that exist in two mirror-image forms: L (Levorotatory) and D (Dextrorotatory) [16]. Natural ribosomes exclusively use L-amino acids to synthesize proteins, resulting in L-framework proteins throughout nature. While engineered ribosomes [26], post-translational modification systems [34], and non-ribosomal peptide synthetases [51] can incorporate some D-amino acids into L-protein chains, proteins made entirely of D-amino acids have never been found in nature [38] and must be chemically synthesized in the laboratory. Despite the synthetic challenges, D-proteins are valuable research tools. They fold into mirror images of their L-counterparts and offer unique opportunities for studying fundamental protein mechanisms, developing stable molecular binding agents, and even exploring the possibility of mirror-image biological systems [19, 49]. To identify D-proteins capable of binding to a target L-protein, mirror-image peptide phage display methods have been developed [9, 11, 54, 58, 85], which involves screening a phage library of L-peptides with a target D-protein. However, it remains challenging to precisely target a specific surface region of the target protein and confirm the presence of
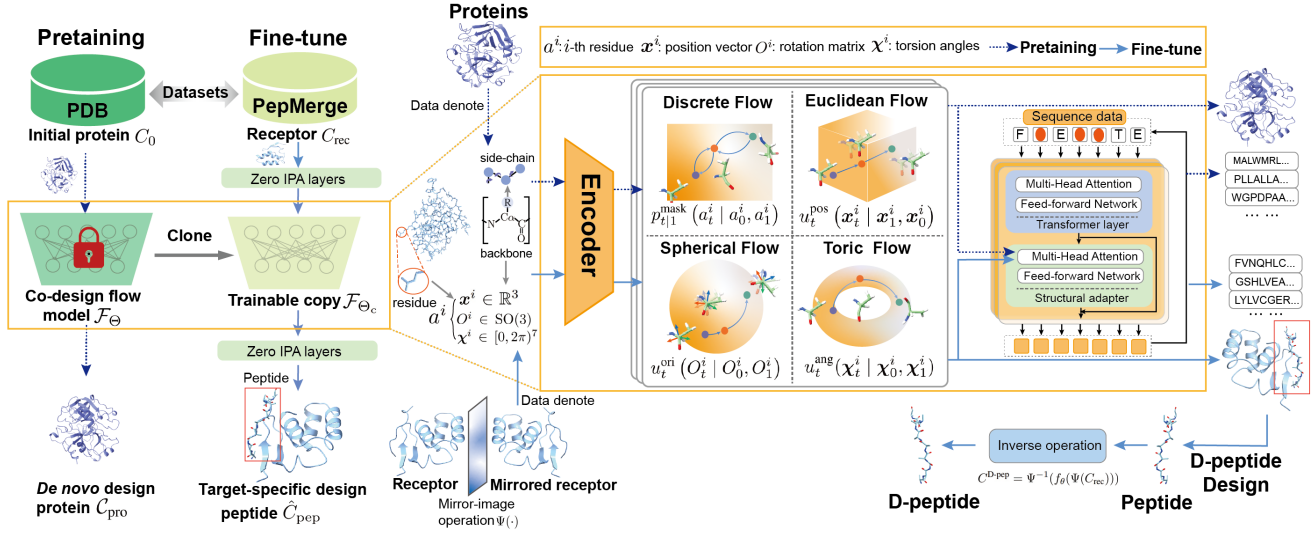
Figure 1. Illustration of D-Flow.

valid binders within the initial random library [63]. More importantly, no preceding works have explored the possibility of DL co-design methods [46, 80] for D-protein generation, not to mention the incorporation and verification of this mirror-image technique.

3. Preliminary and Background

Proteins. A protein is a biomolecule consisting of n amino acid residues, each defined by its type, backbone frame, and side-chain torsion angles [22]. The type of the i-th residue, denoted by $a^i \in \mathcal{A}$, is determined by its side-chain R group, where $\mathcal{A} = \{1 \dots 20\}$. The rigid frame of each residue is constructed from the coordinates of four backbone heavy atoms N, Cα, C, and O, with Cα positioned at the origin. This frame is represented by a position vector $\boldsymbol{x}^i \in \mathbb{R}^3$ and a rotation matrix $O^i \in \mathrm{SO}(3)$ [32]. Additionally, seven torsion angles $\boldsymbol{\chi}^i \in [0, 2\pi)^7$ are considered and three of them belong to the backbone, where $\boldsymbol{\chi}^i[0]$ is the angle around C − N bond, $\boldsymbol{\chi}^i[1]$ is the angle around N − Cα bond, and $\boldsymbol{\chi}^i[2]$ affects the position of the oxygen atom. Unlike the backbone, the side-chain conformation is more flexible, involving up to four rotatable torsion angles between side-chain atoms, denoted by $\boldsymbol{\chi}^i[3:] \in [0, 2\pi)^4$.

Flow on Riemannian Manifolds. Flow matching (FM) [1, 2, 44, 64] is a simulation-free method for learning a continuous normalizing flow (CNF). On a manifold $\mathcal{M}$ with a Riemannian metric g, the CNF $\Phi : \mathcal{M} \to \mathcal{M}$ is defined by a one-parameter diffeomorphism that integrates along a time-dependent vector field $u_t \in \mathcal{U}$, where $u_t(x) \in \mathcal{T}_x\mathcal{M}$ falls at the tangent space of the manifold at $x \in \mathcal{M}$.

With an initial condition of $\phi_0(x) = x$, the time-dependent flow $\phi_t : \mathcal{M} \to \mathcal{M}$ and the final diffeomorphis-mis are attained by setting $\Phi(x) = \phi_1(x)$ and solving the ordinary differential equation (ODE) $\frac{\mathrm{d}\phi_t}{\mathrm{d}t}(x) = u_t(\phi_t(x))$ on $\mathcal{M}$ over $t \in [0, 1]$. Remarkably, $\phi_t(x)$ transports the point x along the vector field $u_t(x)$ from time 0 up to time t, satisfying another ODE: $\mathrm{d}x = u_t(x)\mathrm{d}t$.

Denote $\mathbb{P}(\mathcal{M})$ as the space of probability distributions on $\mathcal{M}$. Then the flow can transform a simple prior density $p_0 \in \mathbb{P}(\mathcal{M})$ towards the data distribution $p_1 \in \mathbb{P}(\mathcal{M})$ via the push-forward equation $p_t = [\phi_t]_\# p_0$, and its density is $p_t(x) = [\phi_t]_\# p_0(x) = p_0\left(\phi_t^{-1}(x)\right) \det\left[\frac{\partial \phi_t^{-1}}{\partial x}(x)\right]$. This sequence of probability distributions $\{p_t : t \in [0, 1]\}$ is referred to the probability path, which interpolates $\mathbb{P}(\mathcal{M})$ between p_0 and p_1.

The density p_t is characterized by the Fokker-Planck equation: $\frac{\mathrm{d}p_t}{\mathrm{d}t} = -\mathrm{div}(u_t p_t)$, also known as the continuity equation. Under these conditions, u_t is said to be the probability flow for p_t, and p_t is said to be the marginal probability path generated by u_t. Although u_t is intractable in general, it can be learned efficiently by decomposing the target probability path p_t as a mixture of tractable conditional probability paths $p_t(x|x_1)$, which satisfy $p_0(x|x_1) = p_0(x)$ and $p_1(x|x_1) = \delta_{x_1}(\cdot)$. Consequently, the desired unconditional p_t is an average of the conditional probability paths with respect to the data distribution: $p(x) = \int p_t(x|x_1)p(x_1)dx_1$

The key insight of conditional Riemannian flow matching (CRFM) is to fit a vector field $v_\theta \in \mathcal{U}$ with parameters θ to the conditional vector field $u_t(x|x_1)$, which produces $p_t(x|x_1)$. Its objective falls at the tangent space $\mathcal{T}_x\mathcal{M}$ as:

$$\mathcal{L}_{\mathrm{CRFM}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}[0,1], p_1(x_1), p_t(x|x_1)} \|v_\theta(x, t) - u_t(x|x_1)\|_g^2,$$
(1)

This loss can be parameterized by defining the conditional flow $x_t = \phi_t(x_0|x_1)$, where ϕ_t is the solution to $\frac{d\phi_t}{dt}(x) = u_t(\phi_t(x_0|x_1)|x_1)$ with an initial condition of $\phi_0(x_0|x_1) = x_0$. Finally, the loss can be written as $\mathcal{L}_{\text{CRFM}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}[0,1], x_1 \sim p_1(x_1), x_0 \sim p_0(x_0)} \| v_\theta(x,t) - \dot{x}_t \|_g^2$. During the inference period, samples can be generated by solving the ODE related to the neural vector field v_θ to push $x_0 \in \mathcal{M}$ from the source density p_0 to the data distribution p_1 in time efficiently.

4. Methods

4.1. Flow for Full-atom Protein Generation

Discrete Flow for 1D Amino Acid Sequence. Though flow models succeeds in continuous spatial signals like images [30, 57], they falter when applied to discrete sequential data. Two recent lines adapt diffusion or flow to the discrete setting: one embeds the discrete data in continuous space and adopts the continuous diffusion or flow [18, 41, 70], and the other designs the transformation process immediately over categorical state spaces [6, 10, 25]. In this work, we investigate both approaches and observe a significant advantage of discrete flow over the continuous version.

Following Campbell et al. [10] and Gat et al. [25], we define FM in a discrete setting and consider a Continuous-Time discrete Markov Chain (CTMC) paradigm. The categorical variable a_t^i jumps between states in the amino acid type space $\mathcal{A}$ depending on a continuous time $t \in [0,1]$. u_t, called the *probability velocity* as reminiscent of the velocity field in CNF, is defined as the rate of probability change of the current sample a_t^i in each of its $\|\mathcal{A}\| = 20$ categories. Thus, each category of the residue type $a_t^i \sim p_t$ is updated independently with the Euler step as $a_{t+\frac{1}{N}}^i \sim \delta_{a_t^i}(\cdot) + \frac{1}{N} u_t(a^i, \cdot)$. For a sufficiently large number of timesteps $N \to \infty$ and any potential state $z \in \mathcal{A}$, the probability velocity u_t is required to satisfy

$$\sum_{a^i \in \mathcal{A}} u_t(a^i, z) = 0 \text{ and } u_t(a^i, z) \geq 0, \quad (2)$$

for $\forall i \in [n]$ and $a^i \neq z$. Then the unconditional probability velocity u_t can be constructed as the marginalization of conditional one $u_t(a^i, z \mid a_0^i, a_1^i)$. As a consequence, we attain the marginal velocity written as $u_t(a^i, z) = \sum_{a_0^i, a_1^i \in \mathcal{A}} u_t(a^i, z \mid a_0^i, a_1^i) p_t(a_0^i, a_1^i \mid z)$, which generates the probability path $p_t(a^i)$. Generally, the conditional probability paths can be represented as a convex sum of m conditional probabilities, namely, $p_t(a^i \mid a_0^i, a_1^i) = \sum_{j=1}^k \kappa_t^j p^j(a^i \mid a_0^i, a_1^i)$, where the schedulers κ_t^j are collectively non-negative and satisfy $\sum_j \kappa_t^j = 1$. A simple yet useful instance of conditional probability paths is $p_t(a^i \mid a_0^i, a_1^i) = \kappa_t \delta_{a_1^i}(a^i) + (1 - \kappa_t)\delta_{a_0^i}(a^i)$, where the scheduler κ_t satisfies $\kappa_0 = 0$, $\kappa_1 = 1$ and monotonically increases in t (*i.e.*, $\dot{\kappa}_t \geq 0$). This results in

$p_0(a^i|a_0^i, a_1^i) = \delta_{a_0^i}(a^i)$ and $p_1(a^i|a_0^i, a_1^i) = \delta_{a_1^i}(a^i)$. Subsequently, we get the marginal probability velocity:

$$u_t(a^i, z) = \frac{\dot{\kappa}_t}{1 - \kappa_t} \left[p_{1|t}(a^i \mid z) - \delta_z(a^i) \right], \quad (3)$$

where $p_{1|t}(a^i \mid z) = \sum_{a_0^i, a_1^i} \delta_{a_1^i}(a^i) p_t(a_0^i, a_1^i \mid z)$ is used as the notation for the probability denoiser. The training goal is to learn this denoiser $p_{1|t}$ by minimizing the cross-entropy (CE) loss:

$$\mathcal{L}_{\text{aa}}(\theta) = \mathbb{E}_{i \in [n], t \sim \mathcal{U}(0,1), p_0(a^i), p_1(a^i), p_t(a^i|a_0^i, a_1^i)} \left[\log p_{1|t}(a_1^i \mid a_t^i) \right]. \quad (4)$$

where $\mathcal{U}(0,1)$ is a uniform time distribution on $[0,1]$, and $v^{\text{aa}}(a_t^i, t)$ is the neural network with parameters θ to approximate the true denoising distribution $p_{1|t}$. Given a noisy input $a_t^i \sim p_{t|1}(a^i \mid a_1^i, a_0^i)$, the model learns to predict the clean data point a_1^i. Here, rather than a linear interpolation towards a_1^i from a prior of uniform categorical distribution $p_{t|1}^{\text{unif}}(a_t^i \mid a_0^i, a_1^i) = \text{Cat}\left(t\delta_{a_1^i}(\cdot) + (1 - t)\frac{1}{\|\mathcal{A}\|} \right)$, we adopt an artificially introduced mask state M and the conditional path becomes [10]:

$$p_{t|1}^{\text{mask}}(a_t^i \mid a_0^i, a_1^i) = \text{Cat}\left(t\delta_{a_1^i}(\cdot) + (1 - t)\delta_{\text{M}}(\cdot) \right). \quad (5)$$

Multi-modality FM for 3D Protein Structures. Apart from the sequential amino acid type, we follow common practice [10, 41, 42, 79] and construct different probability flows containing Euclidean, spherical, and toric FMs for positions $x \in \mathbb{R}^{n \times 3}$, orientations $O \subseteq \text{SO}(3)$, and torsion angles $\chi \in [0, 2\pi)^{n \times 7}$, respectively. To be specific, a vanilla Gaussian FM on Euclidean manifolds with the prior $\mathcal{N}(0, I_3)$ is leveraged to generate x^i. As for 3D rotation group $\text{SO}(3)$, it is also a smooth Riemannian manifold with its tangent space $\mathfrak{so}(3)$ forming a Lie algebra consisting of skew-symmetric matrice. We establish flows based on the geodesics in the context of $\text{SO}(3)$ [39] and select the uniform distribution over $\text{SO}(3)$ as the prior $p(O_0^i)$. For torsion angles $\chi^i \in [0, 2\pi)^7$, each of them can be represented as a point on the unit circle $\mathbb{S}^1$. Thus, χ^i lies on the 7-dimensional hypertorus $\mathbb{T}^7 = (\mathbb{S}^1)^7$ as the Cartesian product of all seven unit circles. This flat torus $\mathbb{T}^7$ can be viewed as the quotient space $(\mathbb{R}^7 \mod 2\pi\mathbb{Z})^7$ that inherits the Riemannian metric from Euclidean space, where the uniform distribution on $[0, 2\pi)^7$ is utilized as the prior. The conditional flows are prescribed as:

$$\phi_t^{\text{pos}}(x_0^i, x_1^i) = tx_1^i + (1 - t)x_0^i, \quad x_0^i \sim \mathcal{N}(0, I_3), \quad (6)$$

$$\phi_t^{\text{ori}}(O_0^i, O_1^i) = \exp_{O_0^i}\left(t \log_{O_0^i}(O_1^i) \right), \quad O_0^i \sim \mathcal{U}(\text{SO}(3)), \quad (7)$$

$$\phi_t^{\text{ang}}(\chi_0^i, \chi_1^i) = \left[t\chi_1^i + (1 - t)\chi_0^i \right] \mod 2\pi, \quad (8)$$

$$\chi_0^i \sim \mathcal{U}([0, 2\pi)^7), \quad (9)$$

where $\exp(\cdot)$ and $\log(\cdot)$ are the exponential and logarithm maps on $SO(3)$ that can be computed efficiently using Rodrigues' formula [41, 79]. Subsequently, the conditional vector fields u_t^{pos}, u_t^{ori}, and u_t^{ang} can be obtained by taking the time derivative of linear flows ϕ_t^{pos}, ϕ_t^{ori}, and ϕ_t^{ang} using Independent Coupling techniques:

$$u_t^{\text{pos}}\left(\boldsymbol{x}_t^i \mid \boldsymbol{x}_1^i, \boldsymbol{x}_0^i\right) = \boldsymbol{x}_1^i - \boldsymbol{x}_0^i = \frac{\boldsymbol{x}_1^i - \boldsymbol{x}_t^i}{1-t}. \quad (10)$$

$$u_t^{\text{ori}}\left(O_t^i \mid O_0^i, O_1^i\right) = \frac{\log_{O_t^i}\left(O_1^i\right)}{1-t}, \quad (11)$$

$$u_t^{\text{ang}}(\boldsymbol{\chi}_t^i \mid \boldsymbol{\chi}_0^i, \boldsymbol{\chi}_1^i) = \left(\frac{\boldsymbol{\chi}_1^i - \boldsymbol{\chi}_t^i}{1-t} + \pi\right) \mod 2\pi - \pi, \quad (12)$$

Ultimately, time-dependent and equivariant neural networks $v^{\text{pos}}(\cdot)$, $v^{\text{ori}}(\cdot)$, and $v^{\text{ang}}(\cdot)$, which share the same backbone architecture as $v^{\text{aa}}(\cdot)$ but have different head predictors, are employed to approximate the conditional gradient fields u_t^{pos}, u_t^{ori}, and u_t^{ang}. The corresponding FM objectives are given below, separately:

$$\mathcal{L}_{\text{pos}}(\theta) = \mathbb{E}_{i\in[n],t\sim\mathcal{U}(0,1),p(\boldsymbol{x}_1^i),p(\boldsymbol{x}_0^i),p(\boldsymbol{x}_t^i|\boldsymbol{x}_0^i,\boldsymbol{x}_1^i)} \\ \left\| v^{\text{pos}}\left(\boldsymbol{x}_t^i, t\right) - \left(\boldsymbol{x}_1^i - \boldsymbol{x}_0^i\right) \right\|_2^2, \quad (13)$$

$$\mathcal{L}_{\text{ori}}(\theta) = \mathbb{E}_{i\in[n],t\sim\mathcal{U}(0,1),p(O_1^i),p(O_0^i),p(O_t^i|O_0^i,O_1^i)} \\ \left\| v^{\text{ori}}\left(O_t^i, t\right) - \frac{\log_{O_t^i}\left(O_1^i\right)}{1-t} \right\|_{SO(3)}^2, \quad (14)$$

$$\mathcal{L}_{\text{ang}}(\theta) = \mathbb{E}_{i\in[n],t\sim\mathcal{U}(0,1),p(\boldsymbol{\chi}_1^i),p(\boldsymbol{\chi}_0^i),p(\boldsymbol{\chi}_t^i|\boldsymbol{\chi}_0^i,\boldsymbol{\chi}_1^i)} \\ \left\| v^{\text{ang}}\left(\boldsymbol{\chi}_t^i, t\right) - (\boldsymbol{\chi}_1^i - \boldsymbol{\chi}_0^i) \right\|_2^2. \quad (15)$$

We summarize losses of all different modalities to acquire the overall FM objective as $\mathcal{L}_{\text{CFM}} = \lambda_{\text{pos}}\mathcal{L}_{\text{pos}} + \lambda_{\text{ori}}\mathcal{L}_{\text{ori}} + \lambda_{\text{ang}}\mathcal{L}_{\text{ang}} + \lambda_{\text{aa}}\mathcal{L}_{\text{aa}}$, where λ_* are the hyperparameters to control the impact of different loss components. Two additional losses are also imposed concerning the backbone atoms and distance matrix [79].

During the inference phase, we first sample from several distinct priors, *i.e.*, $\boldsymbol{x}_0^i \sim \mathcal{N}(0, \boldsymbol{I}_3)$, $O_0^i \sim \mathcal{U}(SO(3))$, and $\boldsymbol{\chi}_0^i \sim [0, 2\pi)^7$. After that, we solve the probability flow with learned predictors $v^{\text{pos}}(\cdot)$, $v^{\text{ori}}(\cdot)$, and $v^{\text{ang}}(\cdot)$ using the N-step forward Euler method to get the position, orientations, and torsion angles of i-th residue with $t = \left\{0, \ldots, \frac{N-1}{N}\right\}$:

$$\boldsymbol{x}_{t+\frac{1}{N}}^i = \boldsymbol{x}_t^i + \frac{1}{N}v^{\text{pos}}\left(\boldsymbol{x}_t^i, t\right), \quad (16)$$

$$O_{t+\frac{1}{N}}^i = \exp_{O_t^i}\left(\frac{1}{N}v^{\text{ori}}\left(O_t^i, t\right)\right), \quad (17)$$

$$\boldsymbol{\chi}_{t+\frac{1}{N}}^i = \left[\boldsymbol{\chi}_t + \frac{1}{N}v^{\text{ang}}\left(\boldsymbol{\chi}_t^i, t\right)\right] \mod 2\pi. \quad (18)$$

4.2. Parameterization with Adapter-guided Protein Language Models

To efficiently learn $(a^i, \boldsymbol{x}^i, O^i, \boldsymbol{\chi}^i)$ for every residue, we adopt and improve the FramePred architecture introduced in FrameDiff [80] and FrameFlow [79]. It effectively incorporates the Invariant Point Attention (IPA) [32] to encode spatial features and ensure the equivariance. IPA consists of multiple Transformer layers [68] to capture sequence-level attributes. In addition, considering the periodicity, torsion angles $\boldsymbol{\chi} \in [0, 2\pi)^7$ are flexibly encoded by applying multi-frequency sine and cosine transformations [41], which are also fused with the timestep embedding and residue sequence embedding into IPAs.

PLMs capture the evolutionary patterns from large-scale sequence data, and this knowledge is proven to be supportive of protein folding [43] and inverse design [84], which is evidently beneficial for our co-design task as well. Some prior studies [77] integrate this knowledge by immediately appending the geometric networks to PLMs. Drawing inspiration from LM-Design [84], we employ a lightweight structural adapter to endow PLMs with structural awareness. In our approach, if the output of IPA is denoted as $\boldsymbol{h}_{\text{IPA}} \in \mathbb{R}^{\psi_{\text{IPA}}}$, the attention at the l-th layer is computed as:

$$\boldsymbol{o} = \text{softmax}\left(\frac{\boldsymbol{h}_{\text{seq}}\boldsymbol{W}_Q \cdot \boldsymbol{h}_{\text{IPA}}^\top \boldsymbol{W}_V^\top}{\sqrt{\psi_{\text{seq}}}}\right)\boldsymbol{h}_{\text{IPA}}\boldsymbol{W}_V, \text{ where } \boldsymbol{h}_{\text{seq}} \in$$

$\mathbb{R}^{\psi_{\text{seq}}}$ is the sequential embedding of the last $(l-1)$-th layer. $\boldsymbol{W}_K$, $\boldsymbol{W}_Q$, and $\boldsymbol{W}_V$ are trainable weights for key, query, and values, separately. This module allows the usage of both sequence-based and structure-based information for protein understanding.

4.3. Controlling Flow Matching with Target Proteins

Existing *de novo* design algorithms [79, 80] are mainly targeting monomers or biomolecules without receptors. Some recent studies [41, 42, 46] have considered binder design but their performance is strictly restricted by the number of available complex structures (*e.g.*, SabDAB [20] and PepMerge [41]). To bridge the gap and exploit all crystal structures, we propose to pretrain our flow model on the vast amount of general proteins (*i.e.*, Protein Data Bank) and then fine-tune it on the target-specific design challenges.

To begin with, $\mathcal{F}_\Theta : \mathcal{P}_0 \to \mathcal{C}_{\text{pro}}$ is a co-design flow model without target awareness, which can *de novo* design a protein C_{pro} from any initial protein C_0 that is drawn from a prior distribution p_0. It is trained on general proteins to approximate $p(a, \boldsymbol{x}, O, \boldsymbol{\chi})$. To enable the awareness of $\mathcal{F}_\Theta$ to the receptor C_{rec}, we borrow ideology from the popular ControlNet architecture [82], locking (freeze) the parameters Θ of the original block and simultaneously cloning the block to a trainable copy with parameters Θ_c. The trainable copy takes a receptor protein C_{rec} as input and is connected to the locked model with zero IPAs denoted as $\mathcal{Z}(\cdot)$. Notably, both weight and bias in $\mathcal{Z}(\cdot)$ are initialized to zeros.

In practice, two instances of zero IPAs are used with parameters Θ_{z1} and Θ_{z2}, respectively:

$$\hat{C}_{\text{pep}} = \mathcal{F}_\Theta(C_0) + \mathcal{Z}_{\Theta_{z2}}\left(\mathcal{F}_{\Theta_c}\left(C_0 + \mathcal{Z}_{\Theta_{z1}}(C_{\text{rec}})\right)\right), \quad (19)$$

where $\hat{C}_{\text{pep}}$ is the designed peptide that is expected to bind with C_{rec}. In the first training step, since both the weight and bias parameters of IPAs are initialized to zero, both $\mathcal{Z}(\cdot)$ terms evaluate to zero, and $\hat{C}_{\text{pep}} = \hat{C}_{\text{pro}}$. In this way, harmful noise cannot influence the hidden states of the network layers in the trainable copy when the training starts. Zero IPA layers protect this backbone by eliminating random noise as gradients in the initial training steps.

4.4. D-peptide Design

Mirror-image peptides and proteins, composed of D-amino acids and the achiral glycine, have garnered significant attention as potential therapeutic and enzymatic tools due to their remarkable resistance to enzymatic digestion by natural-chirality enzymes and their exceptional biostability [81]. However, the development of mirror-image biological systems and related applications faces challenges, primarily due to the lack of effective methods for designing mirror-image (D-) proteins [38].

In this work, we generate mini D-protein binders, specifically D-peptides, using the mirror-image algorithm [54, 63]. The mirror-image operation, denoted as $\Psi(\cdot)$, transforms any given protein structure into its corresponding mirror image. $\Psi(\cdot)$ inverts the spatial configuration of the protein, effectively reflecting it across an imaginary plane while preserving the relative distances and angles between atoms. In our approach, we first apply $\Psi(\cdot)$ to the receptor protein C_{rec}, converting it into its mirror-image counterpart. Then, we pass this mirrored receptor through a flow model $f_\theta(\cdot) : \mathcal{C}_{\text{rec}} \rightarrow \mathcal{C}_{\text{pep}}$, which is trained on natural-chirality (L-) proteins, to yield the corresponding peptide. Finally, we apply the inverse operation $\Psi^{-1}(\cdot)$ to convert the resulting peptide back to the D-configuration. Formally, the generation process can be expressed as:

$$C^{\text{D-pep}} = \Psi^{-1}(f_\theta(\Psi(C_{\text{rec}}))). \quad (20)$$

5. Experiments

We execute two types of experiments to validate the effectiveness of our D-Flow approach. The first, detailed in Sec. 5.1, involves the conventional co-design challenge for L-proteins, where models generate peptides' both sequences and structures based on a given receptor binding site. The second experiment, described in Sec. 5.2, focuses on generating D-peptides. For benchmarking, we use the **PepMerge** dataset [41], derived from PepBDB [73] and Q-BioLip [72]. To ensure a fair comparison with prior work [41], we cluster peptide-protein complexes based on 40% sequence identity using MMseqs2 [61], after remov-

ing duplicates and applying empirical filters (*e.g.*, resolution < 4Å, peptide length between 3 and 25). This process results in 8,365 non-redundant complexes across 292 clusters. For consistency in comparison, we use the same test set as Li et al. [41], consisting of 10 clusters and 158 complexes. Pretraining data contain monomers between lengths 60 and 384 with resolution < 5Å downloaded from PDB [8] on August 8, 2021, ensuring no data leakage for peptides. The data is then filtered by including proteins with high secondary structure compositions only. Monomers with more than 50% loops are also removed using DSSP [33], resulting in 18,684 proteins.

5.1. Unconditioned Sequence-structure Co-design

Baselines. We select two categories of state-of-the-art protein design methods as baselines. The first category disregards side-chain conformations and includes approaches such as RFDiffusion [71] and ProteinGen [45]. RFDiffusion generates protein backbones, with sequences predicted afterward using ProteinMPNN [15], while ProteinGen improves on RFDiffusion by jointly sampling both backbones and sequences. The second category considers full-atom protein generation, containing Diffusion [46], PPIFlow [42], and PepFlow [41]. PepFlow has three variants based on whether backbones, sequences, and side-chain angles are sampled.

Metrics. Generated peptides are evaluated across three key aspects. (1) **Geometry**: Designed peptides should closely mirror native sequences and structures. We quantify sequence similarity through the amino acid recovery rate (**AAR**), which measures sequence identity between generated and ground truth peptides. Structural similarity is evaluated using the root-mean-square deviation (**RMSD**) of C_α atoms after complex alignment. The secondary-structure similarity ratio (**SSR**) captures the proportion of matching secondary structures, while the binding site ratio (**BSR**) evaluates the overlap between generated and native peptide binding sites on the target protein. (2) **Energy**: We aim to design high-affinity peptide binders that enhance protein-peptide complex stability. **Affinity** represents the percentage of generated peptides achieving higher binding affinities (lower binding energies) than the native peptide. **Stability** indicates the fraction of complexes exhibiting lower total energy than the native state. All energy calculations are performed using Rosetta [3]. (3) **Design**: **Designability** measures the consistency between designed sequences and structures, calculated as the proportion of sequences that fold into structures similar to their generated conformations (C_α RMSD < 2 Å). Sequence refolding is performed using ESMFold [43]. **Diversity**, computed as the average of one minus the pairwise TM-Score [83], quantifies structural variation among the designed peptides.

Table 1. Evaluation of methods in the traditional L-peptide sequence-structure co-design task. The **best** and <u>suboptimal</u> results are labeled boldly and underlined, respectively. Baselines with † are reproduced.

	Geometry				Energy		Design	
	AAR % ↑	RMSD Å ↓	SSR % ↑	BSR % ↑	Stb. % ↑	Aff. % ↑	Des. % ↑	Div. ↑
RFdiffusion [71]	40.14	4.17	63.86	26.71	<u>26.82</u>	16.53	**78.52**	0.38
ProteinGen [45]	45.82	4.35	29.15	24.62	23.48	13.47	71.82	0.54
Diffusion [46]	47.04	3.28	74.89	49.83	15.34	17.13	48.54	0.57
PPIFlow† [42]	48.35	–	–	–	–	–	–	–
PepFlow-Bb [41]	50.46	2.30	82.17	82.17	14.04	18.10	50.03	**0.64**
PepFlow-Seq [41]	<u>53.25</u>	2.21	<u>85.22</u>	85.19	19.20	19.39	56.04	0.50
PepFlow-Ang [41]	51.25	<u>2.07</u>	83.46	<u>86.89</u>	18.15	<u>21.37</u>	65.22	0.42
D-Flow	**58.69**	**1.63**	**89.02**	**88.47**	**26.85**	**24.31**	<u>75.14</u>	<u>0.60</u>

Main Results. As documented in Tab. 1, D-Flow generates peptide sequences with the closest resemblance to native ones, achieving a high amino acid identity of 58.69%, a 14.51% improvement over PepFlow. It excels in all geometry-related metrics, highlighting its close alignment with the binding sites of native peptides. Besides, D-Flow scores 75.14% in designability and 0.60 in diversity, maintaining a good balance between structural fidelity and variety. Moreover, D-Flow also demonstrates superior energy-based properties, achieving the best stability score at 26.85% and the best affinity score at 24.31%, which are critical for the formation of strong and stable complexes. In a nutshell, D-Flow outpasses all baselines across most key metrics, indicating its strength in consistently producing the most accurate peptide sequences and structures along with the optimal stability and affinity for their targets.

We also navigate the contribution of different components of D-Flow, containing the discrete flow matching (DFM), structural-adapted PLM, the pretraining on PDB, and the ControlNet-style transfer learning technique. Tab. 2 reports the ablation study results. It can be observed that the integration of additional unlabeled structural data yielded the most substantial improvement, increasing AAR by 5.93%. This supports our hypothesis that limited binding complex data significantly constrains generative models' *de novo* design capabilities. While incorporating 1D sequence information improved D-Flow's performance by 0.58%, this impact was less pronounced than in previous protein-related tasks [77] like ligand docking [14], ligand efficacy prediction [66], or model quality assessment [13]. We attribute this to PLMs' inherent limitations with short sequences like peptides, compared to their effectiveness with longer proteins. This gap generally arises because PLMs learn contextual relationships by observing long-range dependencies and conserved motifs – key elements that characterize protein families and evolutionary relationships [21, 56]. In short sequences, like peptides, these relationships are limited due to a lack of length and structure, making it harder for the model to capture evolutionary pat-

	DFM	PLM	Pretrain	Control	AAR%	Δ
1	-	-	-	-	50.43	
2	✓	-	-	-	51.86	+1.43
3	✓	✓	-	-	52.44	+0.58
4	✓	✓	✓	-	58.37	+5.93
5	✓	✓	✓	✓	58.69	+0.32

Table 2. Ablation studies on the effects of each module.

terns effectively. Peptides often lack secondary and tertiary structure information present in longer sequences, which is critical for understanding function and evolutionary context [4]. Those attention-based models, such as transformers [68], distribute attention weights over the input length. Shorter sequences provide fewer tokens to which attention can be allocated, which can limit the richness of the learned representation. Regarding other components, DFM demonstrated superior performance over the probability simplex mechanism [41] for amino acid type representation, improving AAR by 1.43%. The control technique provided a modest 0.32% enhancement in AAR.

5.2. D-peptide Design

Generation of Pure D-peptides. Our analysis confirms that D-Flow successfully produces *pure D-peptides* in all cases (Fig. 3). Remarkably, this achievement occurs without any D-protein training corpus, requiring only our specialized post-processing technique $\Psi(\cdot)$. To understand it, as D-Flow gradually moves the peptide's sequence and structure $(a, \boldsymbol{x}, O, \boldsymbol{\chi})$ from a prior distribution p_0 to the target data distribution p_1 at a speed of v_t during $t \in [0, 1]$, its moving trajectory is heavily dependent on the anchor fixed receptor $\mathcal{C}_{\text{rec}}$. The velocity vectors $v^{\text{pos}}(\cdot)$, $v^{\text{ori}}(\cdot)$, and $v^{\text{ang}}(\cdot)$ for Cα locations, frame orientations, and angles are computed based on the interactions between the intermediate state of peptide $\mathcal{C}_{\text{pep}}^t$ and the target protein $\mathcal{C}_{\text{rec}}$. Once a D-chirality receptor $\Psi(\mathcal{C}_{\text{rec}})$ is provided as the condition, our D-Flow $f_\theta(\cdot)$ captures the chirality nuance of this given
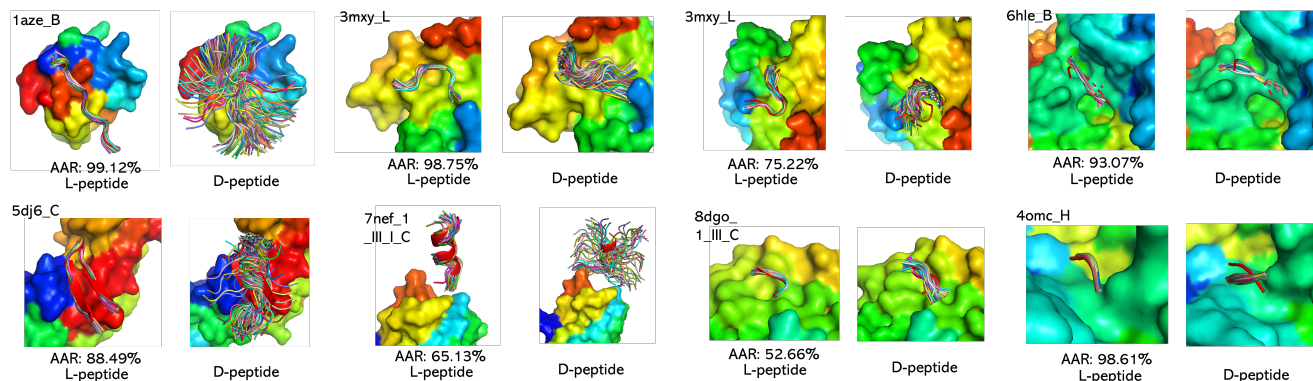
Figure 2. Comparison of L- and D-peptide examples generated by D-Flow with the same target receptor, where the red one is the ground truth L-peptide.

input. As a response, it adjusts $v^{\text{pos}}(\cdot)$, $v^{\text{ori}}(\cdot)$, and $v^{\text{ang}}(\cdot)$ to accord with this D-target protein, which leads to a product with considerable D-residues.

Model Robustness Assessment. We examined the robustness of D-Flow by testing multiple mirror-image approaches for the input receptor. Despite there being infinite possible mirror transformations, we focused on the three fundamental axes (x, y, and z). Our results demonstrate that the chirality of generated peptides remains consistently D-configured regardless of the mirroring approach used.

Conformational Distribution Analysis Fig. 2 visualizes a few examples of designed L- and D-peptides. While D-Flow successfully generates both configurations, our analysis reveals a striking contrast: L-peptides show remarkable consistency in their structural and sequence distributions, whereas D-peptides exhibit significantly higher variability, particularly in longer sequences. This phenomenon can be attributed to several fundamental factors.

Training Distribution Mismatch. Our model, trained exclusively on L-peptides, has learned to navigate L-peptide conformations' natural distribution and energy landscape. When presented with a D-receptor through the mirror-image algorithm, the model must operate in a conformational space absent from its training distribution. This distributional shift forces the model to explore the D-space with less confidence, increasing structural diversity [65]. *Loss of Evolutionary Constraints.* L-peptides in our training data reflect millions of years of evolutionary optimization, containing implicit biases about energetically favorable conformations [23, 40]. While the mirror-image transformation preserves theoretical symmetry, these evolutionary constraints don't translate perfectly to the D-space, potentially contributing to the observed variability. *Asymmetric Energy Landscape Exploration.* The FM process relies

on smooth transitions across the conformational space [44]. However, the mirror-image transformation introduces subtle numerical asymmetries in the computational representation of D-peptides. These asymmetries can affect how the model explores the energy landscape, leading to more diverse intermediate states and final structures [74]. *Stereochemical Environment Adaptation.* The significant change in the stereochemical environment during mirroring poses a unique challenge. The flow model must adapt its learned representations of molecular interactions to account for the inverted chirality [5]. This adaptation process isn't perfect, resulting in broader conformational sampling and increased structural diversity in the generated D-peptides.

These challenges align with recent findings in protein design that highlight the challenges of transferring learned molecular representations across different stereochemical spaces [62]. The energy landscape learned for L-peptides doesn't directly translate to the D-space due to fundamental differences in molecular interactions and packing arrangements. This leads to broader structural exploration and increased diversity in the generated structures [52]. This analysis suggests potential avenues for improvement, such as incorporating symmetry-aware constraints or developing specialized adaptation mechanisms for D-peptide generation, which could help stabilize the output distribution while maintaining the desired chirality.

6. Conclusion

D-peptides, being the mirror image forms, often have different properties. They are more resistant to enzymatic degradation, making them interesting for therapeutic applications, such as drug design and antimicrobial peptides [54]. This work introduces a multi-modality flow model dubbed D-Flow to generate D-peptides from scratch. D-Flow is proven to produce peptides with a considerable proportion of D-residues via the novel mirror-image algorithm.
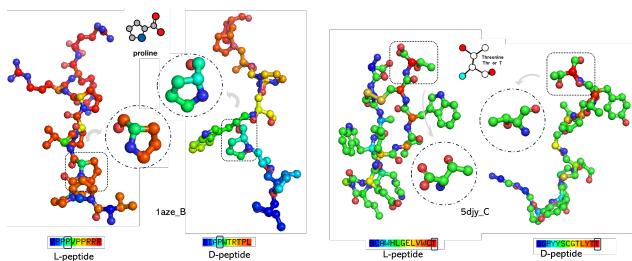
Figure 3. Perspective view of D-residues within the D-peptides, where stereogenic carbon alpha to the amino group has the D-configuration.

Moreover, it is equipped with an adapter-guided protein language model with structural awareness and leverages a ControlNet-style mechanism to bridge the gap between pre-training and fine-tuning stages.

7. Limitations and Future Works

Despite our progress, there remain significant opportunities for further improvement and exploration. This work evaluates designed proteins solely through in silico analysis. A more comprehensive validation of D-Flow's effectiveness would require experimental laboratory testing.

References

[1] Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*, 2022. 3

[2] Michael S Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *arXiv preprint arXiv:2303.08797*, 2023. 3

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

[4] Ethan C Alley, Grigory Khimulya, Surojit Biswas, Mohammed AlQuraishi, and George M Church. Unified rational protein engineering with sequence-based deep representation learning. *Nature methods*, 16(12):1315–1322, 2019. 7

[5] Namrata Anand and Tudor Achim. Protein structure and sequence generation with equivariant denoising diffusion probabilistic models. *arXiv preprint arXiv:2205.15019*, 2022. 8

[6] Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. Structured denoising diffusion models in discrete state-spaces. *Advances in Neural Information Processing Systems*, 34:17981–17993, 2021. 4

[7] Dmitry Baranchuk, Ivan Rubachev, Andrey Voynov, Valentin Khrulkov, and Artem Babenko. Label-efficient semantic segmentation with diffusion models. *arXiv preprint arXiv:2112.03126*, 2021. 1

[8] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000. 6

[9] Alex J Callahan, Satish Gandhesiri, Tara L Travaline, Rahi M Reja, Lia Lozano Salazar, Stephanie Hanna, Yen-Chun Lee, Kunhua Li, Olena S Tokareva, Jean-Marie Swiecicki, et al. Mirror-image ligand discovery enabled by single-shot fast-flow synthesis of d-proteins. *Nature communications*, 15(1):1813, 2024. 2

[10] Andrew Campbell, Jason Yim, Regina Barzilay, Tom Rainforth, and Tommi Jaakkola. Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design. *arXiv preprint arXiv:2402.04997*, 2024. 2, 4

[11] Hao-Nan Chang, Bei-Yuan Liu, Yun-Kun Qi, Yang Zhou, Yan-Ping Chen, Kai-Mai Pan, Wen-Wen Li, Xiu-Man Zhou, Wei-Wei Ma, Cai-Yun Fu, et al. Blocking of the pd-1/pd-l1 interaction by ad-peptide antagonist for cancer immunotherapy. *Angewandte Chemie International Edition*, 54(40):11760–11764, 2015. 2

[12] Ricky TQ Chen and Yaron Lipman. Riemannian flow matching on general geometries. *arXiv preprint arXiv:2302.03660*, 2023. 2

[13] Jianlin Cheng, Myong-Ho Choe, Arne Elofsson, Kun-Sop Han, Jie Hou, Ali HA Maghrabi, Liam J McGuffin, David Menéndez-Hurtado, Kliment Olechnovič, Torsten Schwede, et al. Estimation of model accuracy in casp13. *Proteins: Structure, Function, and Bioinformatics*, 87(12):1361–1377, 2019. 7

[14] Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. *arXiv preprint arXiv:2210.01776*, 2022. 7

[15] Justas Dauparas, Ivan Anishchenko, Nathaniel Bennett, Hua Bai, Robert J Ragotte, Lukas F Milles, Basile IM Wicky, Alexis Courbet, Rob J de Haas, Neville Bethel, et al. Robust deep learning–based protein sequence design using protein-mpnn. *Science*, 378(6615):49–56, 2022. 6

[16] Vadim A Davankov. Biological homochirality on the earth, or in the universe? a selective review. *Symmetry*, 10(12):749, 2018. 2

[17] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021. 1, 2

[18] Sander Dieleman, Laurent Sartran, Arman Roshannai, Nikolay Savinov, Yaroslav Ganin, Pierre H Richemond, Arnaud Doucet, Robin Strudel, Chris Dyer, Conor Durkan, et al. Continuous diffusion for categorical data. *arXiv preprint arXiv:2211.15089*, 2022. 4

[19] Suwei Dong, Ji-Shen Zheng, Yiming Li, Huan Wang, Gong Chen, Yongxiang Chen, Gemin Fang, Jun Guo, Chunmao He, Honggang Hu, et al. Recent advances in chemical protein synthesis: method developments and biological applications. *Science China Chemistry*, 67(4):1060–1096, 2024. 2

[20] James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, Jiye Shi, and Char-

lotte M Deane. Sabdab: the structural antibody database. *Nucleic acids research*, 42(D1):D1140–D1146, 2014. 5

[21] Ahmed Elnaggar, Michael Heinzinger, Christian Dallago, Ghalia Rihawi, Yu Wang, Llion Jones, Tom Gibbs, Tamas Feher, Christoph Angerer, Martin Steinegger, et al. Prottrans: towards cracking the language of life's code through self-supervised deep learning and high performance computing. *arXiv preprint arXiv:2007.06225*, 2020. 7

[22] Matthew Fisher. Lehninger principles of biochemistry, ; by david l. nelson and michael m. cox. *The Chemical Educator*, 6:69–70, 2001. 3

[23] Sarel J Fleishman and David Baker. Role of the biomolecular energy gap in protein design, structure, and evolution. *Cell*, 149(2):262–273, 2012. 8

[24] Noriko Fujii, Takumi Takata, Norihiko Fujii, Kenzo Aki, and Hiroaki Sakaue. D-amino acids in protein: The mirror of life as a molecular index of aging. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1866(7):840–847, 2018. 2

[25] Itai Gat, Tal Remez, Neta Shaul, Felix Kreuk, Ricky TQ Chen, Gabriel Synnaeve, Yossi Adi, and Yaron Lipman. Discrete flow matching. *arXiv preprint arXiv:2407.15595*, 2024. 4

[26] Yuki Goto, Hiroshi Murakami, and Hiroaki Suga. Initiating translation with d-amino acids. *Rna*, 14(7):1390–1398, 2008. 2

[27] Sridhar Govindarajan, Ruben Recabarren, and Richard A Goldstein. Estimating the total number of protein folds. *Proteins: Structure, Function, and Bioinformatics*, 35(4):408–414, 1999. 1

[28] Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, and Jianzhu Ma. 3d equivariant diffusion for target-aware molecule generation and affinity prediction. *arXiv preprint arXiv:2303.03543*, 2023. 1

[29] William Harvey, Saeid Naderiparizi, Vaden Masrani, Christian Weilbach, and Frank Wood. Flexible diffusion modeling of long videos. *Advances in Neural Information Processing Systems*, 35:27953–27965, 2022. 1

[30] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020. 1, 2, 4

[31] Knud J Jensen. Peptide and protein design for biopharmaceutical applications. 2009. 1

[32] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021. 3, 5

[33] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules*, 22(12):2577–2637, 1983. 6

[34] Jyumpei Kobayashi, Yasuhiro Shimizu, Yuta Mutaguchi, Katsumi Doi, and Toshihisa Ohshima. Characterization of d-amino acid aminotransferase from lactobacillus salivarius. *Journal of Molecular Catalysis B: Enzymatic*, 94:15–22, 2013. 2

[35] Xiangzhe Kong, Wenbing Huang, and Yang Liu. Conditional antibody design as 3d equivariant graph translation. *arXiv preprint arXiv:2208.06073*, 2022. 2

[36] Xiangzhe Kong, Wenbing Huang, and Yang Liu. Full-atom peptide design with geometric latent diffusion. *arXiv preprint arXiv:2402.13555*, 2024. 1, 2

[37] Houtim Lai, Longyue Wang, Ruiyuan Qian, Geyan Ye, Juhong Huang, Fandi Wu, Fang Wu, Xiangxiang Zeng, and Wei Liu. Interformer: An interaction-aware model for protein-ligand docking and affinity prediction. 2024. 1

[38] Alexander J Lander, Yi Jin, and Louis YP Luk. D-peptide and d-protein technology: recent advances, challenges, and opportunities. *ChemBioChem*, 24(4):e202200537, 2023. 2, 6

[39] John M Lee. *Introduction to Riemannian manifolds*. Springer, 2018. 4

[40] Arnon Levy and Yair Levy. Evolutionary debunking arguments meet evolutionary science. *Philosophy and Phenomenological Research*, 100(3):491–509, 2020. 8

[41] Jiahan Li, Chaoran Cheng, Zuofan Wu, Ruihan Guo, Shitong Luo, Zhizhou Ren, Jian Peng, and Jianzhu Ma. Full-atom peptide design based on multi-modal flow matching. *arXiv preprint arXiv:2406.00735*, 2024. 1, 2, 4, 5, 6, 7

[42] Haitao Lin, Odin Zhang, Huifeng Zhao, Dejun Jiang, Lirong Wu, Zicheng Liu, Yufei Huang, and Stan Z Li. Ppflow: Target-aware peptide design with torsional flow matching. *bioRxiv*, pages 2024–03, 2024. 1, 2, 4, 5, 6, 7

[43] Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Sal Candido, et al. Language models of protein sequences at the scale of evolution enable accurate structure prediction. *bioRxiv*, 2022. 5, 6

[44] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022. 1, 2, 3, 8

[45] Sidney Lyayuga Lisanza, Jake Merle Gershon, Sam Tipps, Lucas Arnoldt, Samuel Hendel, Jeremiah Nelson Sims, Xinting Li, and David Baker. Joint generation of protein sequence and structure with rosettafold sequence space diffusion. *bioRxiv*, pages 2023–05, 2023. 6, 7

[46] Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific antibody design and optimization with diffusion-based generative models. *bioRxiv*, 2022. 1, 2, 3, 5, 6, 7

[47] Shitong Luo, Yufeng Su, Zuofan Wu, Chenpeng Su, Jian Peng, and Jianzhu Ma. Rotamer density estimator is an unsupervised learner of the effect of mutations on protein-protein interaction. *bioRxiv*, pages 2023–02, 2023. 2

[48] Karolis Martinkus, Jan Ludwiczak, Wei-Ching Liang, Julien Lafrance-Vanasse, Isidro Hotzel, Arvind Rajpal, Yan Wu, Kyunghyun Cho, Richard Bonneau, Vladimir Gligorijevic, et al. Abdiffuser: full-atom generation of in-vitro functioning antibodies. *Advances in Neural Information Processing Systems*, 36, 2024. 1, 2

[49] RC deL Milton, SCF Milton, and SBH Kent. Total chemical synthesis of a d-enzyme: the enantiomers of hiv-1 protease show reciprocal chiral substrate specificity. *Science*, 256(5062):1445–1448, 1992. 2

[50] Alex Nichol, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew, Ilya Sutskever, and Mark Chen. Glide: Towards photorealistic image generation and editing with text-guided diffusion models. *arXiv preprint arXiv:2112.10741*, 2021. 1

[51] Yasushi Ogasawara and Tohru Dairi. Peptide epimerization machineries found in microorganisms. *Frontiers in microbiology*, 9:156, 2018. 2

[52] José Nelson Onuchic and Peter G Wolynes. Theory of protein folding. *Current opinion in structural biology*, 14(1): 70–75, 2004. 8

[53] Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lipman, and Ricky TQ Chen. Multisample flow matching: Straightening flows with minibatch couplings. *arXiv preprint arXiv:2304.14772*, 2023. 2

[54] Yun-Kun Qi, Ji-Shen Zheng, and Lei Liu. Mirror-image protein and peptide drug discovery through mirror-image phage display. *Chem*, 2024. 2, 6, 8

[55] Vishva Saravanan Ramasubramanian, Soham Choudhuri, and Bhaswar Ghosh. A hybrid diffusion model for stable, affinity-driven, receptor-aware peptide generation. *bioRxiv*, pages 2024–03, 2024. 1

[56] Alexander Rives, Joshua Meier, Tom Sercu, Siddharth Goyal, Zeming Lin, Jason Liu, Demi Guo, Myle Ott, C Lawrence Zitnick, Jerry Ma, et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 118(15):e2016239118, 2021. 7

[57] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022. 4

[58] Ton NM Schumacher, Lorenz M Mayr, Daniel L Minor Jr, Michael A Milhollen, Michael W Burgess, and Peter S Kim. Identification of d-peptide ligands through mirror-image phage display. *Science*, 271(5257):1854–1857, 1996. 2

[59] J Ryan Shue, Eric Ryan Chan, Ryan Po, Zachary Ankner, Jiajun Wu, and Gordon Wetzstein. 3d neural field generation using triplane diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20875–20886, 2023. 1

[60] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems*, 32, 2019. 1, 2

[61] Martin Steinegger and Johannes Söding. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology*, 35(11):1026–1028, 2017. 6

[62] Alexey Strokach and Philip M Kim. Deep generative modeling for protein design. *Current opinion in structural biology*, 72:226–236, 2022. 8

[63] Ke Sun, Sicong Li, Bowen Zheng, Yanlei Zhu, Tongyue Wang, Mingfu Liang, Yue Yao, Kairan Zhang, Jizhong Zhang, Hongyong Li, et al. Accurate de novo design of het-

[64] Alexander Tong, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Kilian Fatras, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *arXiv preprint arXiv:2302.00482*, 2023. 2, 3

erochiral protein–protein interactions. *Cell Research*, pages 1–13, 2024. 2, 3, 6

[65] Mirko Torrisi, Gianluca Pollastri, and Quan Le. Deep learning methods in protein structure prediction. *Computational and Structural Biotechnology Journal*, 18:1301–1310, 2020. 8

[66] Raphael JL Townshend, Martin Vögele, Patricia Suriana, Alexander Derry, Alexander Powers, Yianni Laloudakis, Sidhika Balachandar, Bowen Jing, Brandon Anderson, Stephan Eismann, et al. Atom3d: Tasks on molecules in three dimensions. *arXiv preprint arXiv:2012.04035*, 2020. 7

[67] Brian L Trippe, Jason Yim, Doug Tischer, David Baker, Tamara Broderick, Regina Barzilay, and Tommi Jaakkola. Diffusion probabilistic modeling of protein backbones in 3d for the motif-scaffolding problem. *arXiv preprint arXiv:2206.04119*, 2022. 1, 2

[68] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 5, 7

[69] Lei Wang, Nanxi Wang, Wenping Zhang, Xurui Cheng, Zhibin Yan, Gang Shao, Xi Wang, Rui Wang, and Caiyun Fu. Therapeutic peptides: current applications and future directions. *Signal transduction and targeted therapy*, 7(1):48, 2022. 1

[70] Xinyou Wang, Zaixiang Zheng, Fei Ye, Dongyu Xue, Shujian Huang, and Quanquan Gu. Diffusion language models are versatile protein learners. *arXiv preprint arXiv:2402.18567*, 2024. 4

[71] Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023. 1, 2, 6, 7

[72] Hong Wei, Wenkai Wang, Zhenling Peng, and Jianyi Yang. Q-biolip: A comprehensive resource for quaternary structure-based protein–ligand interactions. *Genomics, Proteomics & Bioinformatics*, 22(1), 2024. 6

[73] Zeyu Wen, Jiahua He, Huanyu Tao, and Sheng-You Huang. Pepbdb: a comprehensive structural database of biological peptide–protein interactions. *Bioinformatics*, 35(1):175–177, 2019. 6

[74] Peter G Wolynes. Symmetry and the energy landscapes of biomolecules. *Proceedings of the National Academy of Sciences*, 93(25):14249–14255, 1996. 8

[75] Fang Wu and Stan Z Li. Diffmd: A geometric diffusion model for molecular dynamics simulations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 5321–5329, 2023. 1

[76] Fang Wu and Stan Z Li. A hierarchical training paradigm for antibody structure-sequence co-design. *Advances in Neural Information Processing Systems*, 36, 2024. 2

[77] Fang Wu, Lirong Wu, Dragomir Radev, Jinbo Xu, and Stan Z Li. Integration of pre-trained protein language models into geometric deep learning networks. *Communications Biology*, 6(1):876, 2023. 2, 5, 7

[78] Kevin E Wu, Kevin K Yang, Rianne van den Berg, Sarah Alamdari, James Y Zou, Alex X Lu, and Ava P Amini. Protein structure generation via folding diffusion. *Nature communications*, 15(1):1059, 2024. 1

[79] Jason Yim, Andrew Campbell, Andrew YK Foong, Michael Gastegger, José Jiménez-Luna, Sarah Lewis, Victor Garcia Satorras, Bastiaan S Veeling, Regina Barzilay, Tommi Jaakkola, et al. Fast protein backbone generation with se (3) flow matching. *arXiv preprint arXiv:2310.05297*, 2023. 1, 4, 5

[80] Jason Yim, Brian L Trippe, Valentin De Bortoli, Emile Mathieu, Arnaud Doucet, Regina Barzilay, and Tommi Jaakkola. Se (3) diffusion model with application to protein backbone generation. *arXiv preprint arXiv:2302.02277*, 2023. 1, 3, 5

[81] Guanwei Zhang and Ting F Zhu. Mirror-image trypsin digestion and sequencing of d-proteins. *Nature Chemistry*, 16 (4):592–598, 2024. 6

[82] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3836–3847, 2023. 2, 5

[83] Yang Zhang and Jeffrey Skolnick. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4):702–710, 2004. 6

[84] Zaixiang Zheng, Yifan Deng, Dongyu Xue, Yi Zhou, Fei Ye, and Quanquan Gu. Structure-informed language models are protein designers. In *International conference on machine learning*, pages 42317–42338. PMLR, 2023. 2, 5

[85] Xiuman Zhou, Chao Zuo, Wanqiong Li, Weiwei Shi, Xiaowen Zhou, Hongfei Wang, Shaomeng Chen, Jiangfeng Du, Guanyu Chen, Wenjie Zhai, et al. A novel d-peptide identified by mirror-image phage display blocks tigit/pvr for cancer immunotherapy. *Angewandte Chemie International Edition*, 59(35):15114–15118, 2020. 2