

SURFACE-BASED PEPTIDE DESIGN WITH MULTI-MODAL FLOW MATCHING

Anonymous authors

Paper under double-blind review

ABSTRACT

Therapeutic peptides show promise in targeting previously undruggable binding sites, with recent advancements in deep generative models enabling full-atom peptide co-design for specific protein receptors. However, the critical role of molecular surfaces in protein-protein interactions (PPIs) has been underexplored. To bridge this gap, we propose an *omni-design* peptides generation paradigm, called SurfFlow, a novel surface-based generative algorithm that enables comprehensive co-design of sequence, structure, and surface for peptides. SurfFlow employs a multi-modality conditional flow matching (CFM) architecture to learn distributions of surface geometries and biochemical properties, enhancing peptide binding accuracy. Evaluated on the comprehensive PepMerge benchmark, SurfFlow consistently outperforms full-atom baselines across all metrics. These results highlight the advantages of considering molecular surfaces in *de novo* peptide discovery and demonstrate the potential of integrating multiple protein modalities for more effective therapeutic peptide discovery. Anonymous codes are available at <https://anonymous.4open.science/r/SurfFlow-880B/>.

1 INTRODUCTION

Peptides, short-chain proteins composed of roughly 2 to 50 amino acids linked by peptide bonds, play critical roles in various biological processes, including cell signaling, enzymatic catalysis, and immune responses (Wang et al., 2022). They are essential mediators in pharmacology due to their ability to bind cell surface receptors with high affinity and specificity, inducing intracellular effects with low toxicity, minimal immunogenicity, and ease of delivery, as well as being readily synthesized and modified. (Muttenthaler et al., 2021). Conventional simulation or searching methods of peptide discovery rely on frequent calculations of physical energy functions, a process hindered by the vast design space of peptides (Bhardwaj et al., 2016). This has spurred a growing demand for computational approaches that facilitate *in silico* peptide design and analysis.

Recent progress in diffusion probabilistic (Ho et al., 2020) and flow-based (Lipman et al., 2022) models have illuminated tremendous promise in molecular design (Guan et al., 2023), antibody engineering (Luo et al., 2022; Martinkus et al., 2024), *de novo* protein design (Yim et al., 2023; Wu et al., 2024) as well as peptide discovery (Ramasubramanian et al., 2024). Peptides, when unbound, often exist in high-energy, high-entropy states with unstable conformations, only becoming functional upon binding to target receptors. Therefore, peptide design must be explicitly conditioned on binding pockets Vanhee et al. (2011). Besides, as residues interact with each other through non-covalent forces formed by side-chain groups, increasing efforts have been made to capture protein-peptide interactions through full-atom geometries (Kong et al., 2024; Lin et al., 2024; Li et al., 2024), pushing sequence-structure co-design beyond just backbones.

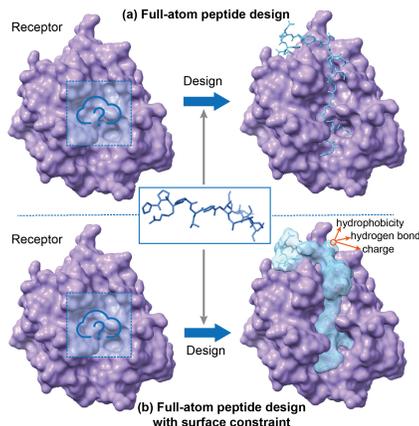


Figure 1: Comparison of full-atom peptide design with and without the surface constraint.

054 Despite advances, growing attention is given to molecular surfaces in protein-protein interactions
055 (PPIs), as these interactions are largely dictated by how complementarily the surfaces of interacting
056 proteins fit together (Kastritis & Bonvin, 2013; Song et al., 2024). The shape, electrostatic potential,
057 and hydrophobicity of molecular surfaces are key determinants of the interaction’s strength and
058 specificity (Jones & Thornton, 1996; Lee et al., 2023; Wu & Li, 2024b). The surface geometry
059 such as protrusions, grooves, and clefts enables lock-and-key or induced-fit mechanisms essential
060 for specific binding. The molecular surfaces act as a fundamental interface that dictates how proteins
061 recognize and bind to each other, making it as essential as side chains in PPIs. For these reasons,
062 it is vital to simultaneously consider all molecular modalities – sequence, structure, and surface –
063 during peptide generation (see Fig. 1), enhancing the consistency across aspects in what we term
064 *omni-design*.

065 Toward this goal, we propose SurfFlow, a surface-based generative algorithm built upon a multi-
066 modality Conditional Flow Matching (CFM) architecture (Lipman et al., 2022; Albergo & Vanden-
067 Eijnden, 2022; Albergo et al., 2023). In SurfFlow, CFM is applied not only to internal geometries
068 such as backbone and side-chain angles but also to the molecular surface, represented by surface
069 point positions and unit norm vectors as a rigid frame in $SE(3)$. Additionally, omni-design requires
070 incorporating biochemical property constraints, as complementary surface geometries alone do not
071 guarantee successful binding-accurate placement of charges, polarity, or hydrophobicity at the bind-
072 ing interface is also necessary (Gainza et al., 2023). To achieve this, SurfFlow learns the trans-
073 formation from prior distributions to the expected biochemical property distributions. Since surface
074 features such as hydrophobicity are categorical, we apply the Discrete Flow Models (DFMs) (Camp-
075 bell et al., 2024) to discrete data space using Continuous-Time Markov Chains (CTMC). Finally,
076 recognizing that key peptide characteristics like cyclicality and disulfide bonds influence stability and
077 binding affinity (Buckton et al., 2021), we include these factors as additional conditions to enhance
078 the capacity and generalization of SurfFlow. We evaluate SurfFlow on the comprehensive peptide
079 design benchmark PepMerge (Li et al., 2024), and experiments demonstrate that it consistently out-
080 performs full-atom baselines across all metrics, highlighting the advantages of considering surfaces
081 for *de novo* peptide design.

082 2 PRELIMINARY AND BACKGROUND

083 **Proteins and Molecular Surfaces.** A protein is a biomolecule consisting of multiple amino acid
084 residues, each defined by its type, backbone frame, and side-chain torsion angles (Fisher, 2001).
085 The type of the i -th residue, denoted by $a_i \in \{1 \dots 20\}$, is determined by its side-chain R group.
086 The rigid frame of each residue is constructed from the coordinates of four backbone heavy atoms
087 N, C α , C, and O, with C α positioned at the origin. This frame is represented by a position vector
088 $\mathbf{x}_i \in \mathbb{R}^3$ and a rotation matrix $O_i \in SO(3)$ (Jumper et al., 2021). Unlike the backbone, the side-
089 chain conformation is more flexible, involving up to four rotatable torsion angles between side-chain
090 atoms, denoted by $\chi_i \in [0, 2\pi)^4$. Additionally, the backbone torsion angle $\varphi_i \in [0, 2\pi)$ affects the
091 position of the oxygen atom.
092

093 We further consider the molecular surface, which is computed by moving a probe of a certain radius
094 (approximately 1 Å) along the protein to calculate the Solvent Accessible Surface (SAS) and Solvent
095 Excluded Surface (SES). The probe’s coordinates define the surface as an oriented point cloud $Q =$
096 $\{q_i : 1 \leq i \leq m\}$. Each surface point q_i has associated attributes $(\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \Upsilon_i^s)$, where $\mathbf{x}_i^s \in$
097 \mathbb{R}^3 represents its 3D coordinates, and $\mathbf{n}_i^s \in \mathbb{R}^3$ is the corresponding unit normal vector. $\tau_i^s \in$
098 \mathbb{R}^{ψ_τ} and $\Upsilon_i^s \in \mathbb{R}^{\psi_\tau}$ capture its continuous and categorical physicochemical properties, such as
099 hydrophobicity, hydrogen bonding, and charge (Gainza et al., 2020; Song et al., 2024; Wu & Li,
100 2024b).
101

102 This work focuses on designing peptides based on target proteins. Formally, given a peptide C^{pep}
103 with n_{pep} residues and a target protein C^{rec} with n_{rec} residues, we aim to model the conditional
104 joint distribution $p(C^{\text{pep}} | C^{\text{rec}})$. The receptor can be sufficiently and succinctly parameterized as
105 $C^{\text{rec}} = \{(a_i, O_i, \mathbf{x}_i, \chi_i)\}_{i=1}^{n_{\text{rec}}}$, where $\chi_i[0] = \varphi_i$ and $\chi_i \in [0, 2\pi)^5$. As for the ligand peptide, the
106 surface is also portrayed, resulting in $C_{\text{pep}} = \{(a_j, O_j, \mathbf{x}_j, \chi_j)\}_{j=1}^{n_{\text{pep}}} \cup \{(\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \Upsilon_i^s)\}_{i=1}^m$ with
107 $m \gg n_{\text{pep}}$. Practically, software like PyMol (DeLano et al., 2002) or MSMS (Robinson et al., 2014)
can be utilized to compute the raw molecular surface of a protein.

Probability Path and Flow. Let $\mathbb{P}(\mathcal{M})$ be the space of probability distributions on a manifold \mathcal{M} with a Riemannian metric g . A probability path $p_t : [0, 1] \times \mathcal{M} \rightarrow \mathbb{P}(\mathcal{M})$ is an interpolation in the probability space between two distributions $p_0, p_1 \in \mathbb{P}(\mathcal{M})$ indexed by a continuous parameter t .

A flow on \mathcal{M} is defined by a one-parameter diffeomorphism $\Phi : \mathcal{M} \rightarrow \mathcal{M}$, which is the result of integrating instantaneous deformations described by a time-varying vector field $u_t \in \mathcal{U}$. $u_t(x) \in \mathcal{T}_x \mathcal{M}$ is the gradient vector of the path p_t on x at time t . By solving the following Ordinary Differential Equation (ODE) on \mathcal{M} over $t \in [0, 1]$ with an initial condition of $\phi_0(x) = x$: $\frac{d\phi_t}{dt}(x) = u_t(\phi_t(x))$. We acquire the time-dependent flow $\phi_t : \mathcal{M} \rightarrow \mathcal{M}$ and the final diffeomorphism by setting $\Phi(x) = \phi_1(x)$. Notably, $\phi_t(x)$ is also the solution of another ODE: $dx = u_t(x)dt$, which transports the point x along the vector field $u_t(x)$ from time 0 up to time t . Given a source density p_0 , $\phi_t(x)$ induces a push-forward operation $p_t = [\phi_t]_{\#} p_0$. It reshapes the point density $x \sim p_0$ to a more complicated one along $u_t(x)$, and the change-of-variable operator $\#$ is defined by $[\phi_t]_{\#} p_0(x) = p_0(\phi_t^{-1}(x)) \det \left[\frac{\partial \phi_t^{-1}}{\partial x}(x) \right]$. The time-varying density p_t is characterized by the Fokker-Planck equation: $\frac{dp_t}{dt} = -\text{div}(u_t p_t)$, 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 . Flow Matching (FM) (Lipman et al., 2022; Albergo & Vanden-Eijnden, 2022; Albergo et al., 2023) trains a Continuous Normalizing Flow (CNF) by fitting a vector field $v_\theta \in \mathcal{U}$ with parameters θ to a target vector field u_t that produce a probability path p_t . Its objective falls at the tangent space as:

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

As u_t is intractable, an alternative is to construct the conditional probability path $p_t(x|x_1)$ with a conditional vector field $u_t(x|x_1)$. The objective becomes: $\mathcal{L}_{\text{CRFM}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}[0,1], x_1 \sim p_1(x_1), x \sim p_t(x|x_1)} \|v_\theta(x, t) - u_t(x|x_1)\|_g^2$. Riemannian FM and Conditional Riemannian FM (CRFM) objectives are proven to share the same gradients (Lipman et al., 2022; Tong et al., 2023; Chen & Lipman, 2023). During inference, one can solve the ODE related to the neural vector field v_θ to push $x_0 \in \mathcal{M}$ from the source distribution p_0 to the data distribution p_1 in time.

Continuous-Time Markov Chains. CTMC (Norris, 1998) is a class of continuous-time discrete stochastic processes and is closely linked to probability flows. Suppose a categorical variable x has S states and its trajectory x_t over time $t \in [0, 1]$ follows a CTMC, x_t alternates between resting in its current state and periodically jumping to another randomly chosen state. The frequency and destination of the jumps are determined by the rate matrix $R_t \in \mathbb{R}^{S \times S}$ with the constraint its off-diagonal elements are non-negative. The probability x_t will jump to a different state j is $R_t(x_t, j) dt$ for the next infinitesimal time step dt . The transition probability is written as:

$$p_{t+dt|t}(j | x_t) = \begin{cases} R_t(x_t, j) dt & \text{for } j \neq x_t \\ 1 + R_t(x_t, x_t) dt & \text{for } j = x_t \end{cases} = \delta\{x_t, j\} + R_t(x_t, j) dt, \quad (2)$$

where $\delta\{i, j\}$ is the Kronecker delta. $\delta\{i, j\}$ is 1 when $i = j$ and is otherwise 0. $R_t(x_t, x_t) := -\sum_{k \neq x_t} R_t(x_t, k)$ in order for $p_{t+dt|t}(\cdot | i)$ to sum to 1. Using compact notation, $p_{t+dt|t}$ is therefore a categorical distribution with probabilities $\delta\{x_t, \cdot\} + R_t(x_t, \cdot) dt$ denoted as $\text{Cat}(\delta\{x_t, j\} + R_t(x_t, j) dt)$. Namely, $j \sim p_{t+dt|t}(j | x_t) \iff j \sim \text{Cat}(\delta\{x_t, j\} + R_t(x_t, j) dt)$. In practice, we need to simulate the sequence trajectory with finite time intervals Δt . A trajectory can be simulated with Euler steps (Sun et al., 2022).

$$x_{t+\Delta t} \sim \text{Cat}(\delta\{x_t, x_{t+\Delta t}\} + R_t(x_t, x_{t+\Delta t}) \Delta t), \quad (3)$$

where the variable x starts from an initial sample $x_0 \sim p_0$ at time $t = 0$. The rate matrix R_t along with an initial distribution p_0 together define the CTMC. With the marginal distribution at time t as $p_t(x_t)$, the Kolmogorov equation allows us to relate the rate matrix R_t to the change in $p_t(x_t)$:

$$\partial_t p_t(x_t) = \underbrace{\sum_{j \neq x_t} R_t(j, x_t) p_t(j)}_{\text{incoming}} - \underbrace{\sum_{j \neq x_t} R_t(x_t, j) p_t(x_t)}_{\text{outgoing}}. \quad (4)$$

The difference between the incoming and outgoing probability mass is the time derivative of the marginal $\partial_t p_t(x_t)$. Subsequently, we attain $\partial_t p_t = R_t^\top p_t$ where the marginals are treated as probability mass vectors: $p_t \in [0, 1]^S$ and defines an ODE in a vector space. The probability path p_t is said to be generated by R_t if $\partial_t p_t = R_t^\top p_t$ for $\forall t \in [0, 1]$ (Campbell et al., 2024).

162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215

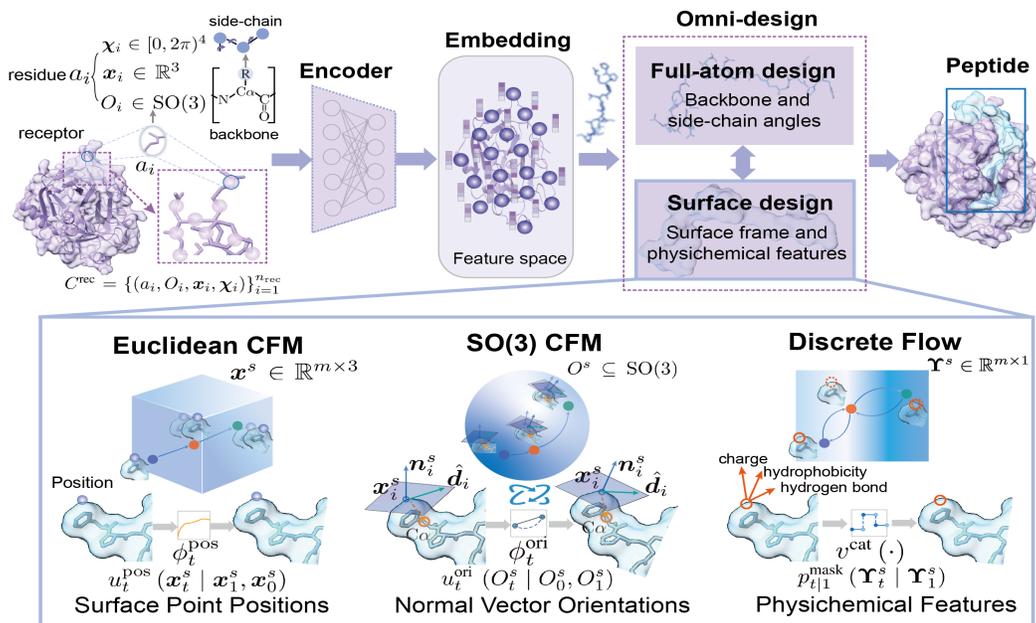


Figure 2: Workflow of SurfFlow for our peptide omni-design, which considers the multi-modality consistency among sequence, structure, and molecular surface during the generation process.

3 METHOD

A molecular surface encapsulates both the 3D geometry of a protein in Euclidean space and its biochemical attributes, such as hydrophobicity and charge (Gainza et al., 2020). The interplay between surface shape and these biochemical properties is essential for defining a protein’s function. Given a target receptor with specific geometric and biochemical constraints, SurfFlow concurrently generates the peptide’s internal structure and external surface. Moreover, it can also account for key factors such as cyclicity and disulfide bonds (see Fig. 2).

3.1 FLOW MATCHING FOR SURFACE GENERATION

A CFM framework is employed to learn the conditional peptide distribution based on its target protein $p(C^{\text{pep}} | C^{\text{rec}})$. This joint probability is empirically decomposed into the product of probabilities of the internal structure elements and the external surface elements:

$$p(C^{\text{pep}} | C^{\text{rec}}) \propto p\left(\{(a_j, O_j, \mathbf{x}_j, \chi_j)\}_{j=1}^{n_{\text{pep}}} | C^{\text{rec}}\right) p\left(\{(\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \Upsilon_i^s)\}_{i=1}^m | C^{\text{rec}}\right), \quad (5)$$

where $p\left(\{(\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \Upsilon_i^s)\}_{i=1}^m | C^{\text{rec}}\right)$ is further separated as the product of four basic elements $p\left(\{\mathbf{x}_i^s\}_{i=1}^m | C^{\text{rec}}\right) p\left(\{\mathbf{n}_i^s\}_{i=1}^m | C^{\text{rec}}\right) p\left(\{\tau_i^s\}_{i=1}^m | C^{\text{rec}}\right) p\left(\{\Upsilon_i^s\}_{i=1}^m | C^{\text{rec}}\right)$. The construction of different probability flows on the surface point’s position $p(\mathbf{x}_i^s | C^{\text{rec}})$, orientation $p(\mathbf{n}_i^s | C^{\text{rec}})$, continuous properties $p(\tau_i^s | C^{\text{rec}})$, and categorical properties $p(\Upsilon_i^s | C^{\text{rec}})$ is elaborated as follows.

Position. Euclidean CFM is utilized to generate surface point positions $\mathbf{x}^s \in \mathbb{R}^{m \times 3}$. Following common practice (Yim et al., 2023; Lin et al., 2024; Li et al., 2024), we adopt the standard isotropic Gaussian $\mathcal{N}(0, \mathbf{I}_3)$ as the prior, with the target distribution being $p(\mathbf{x}^s | C^{\text{rec}})$. The conditional flow is defined as a linear interpolation between sampled noise $\mathbf{x}_0^s \sim \mathcal{N}(0, \mathbf{I}_3)$ and data points $\mathbf{x}_1^s \sim p(\mathbf{x}^s | C^{\text{rec}})$. This linear interpolation ensures a straight trajectory, contributing to training and sampling efficiency by following the shortest path between two points in Euclidean space (Liu et al., 2022). The conditional vector field u_t^{pos} is obtained by taking the time derivative of the linear flow ϕ_t^{pos} using Independent Coupling techniques:

$$\phi_t^{\text{pos}}(\mathbf{x}_0^s, \mathbf{x}_1^s) = t\mathbf{x}_1^s + (1-t)\mathbf{x}_0^s, \quad (6)$$

$$u_t^{\text{pos}}(\mathbf{x}_t^s | \mathbf{x}_1^s, \mathbf{x}_0^s) = \mathbf{x}_1^s - \mathbf{x}_0^s = \frac{\mathbf{x}_1^s - \mathbf{x}_t^s}{1-t}. \quad (7)$$

We use a time-dependent translation-invariant surface network $v^{\text{pos}}(\cdot)$ to predict the conditional vector field based on the current interpolant \mathbf{x}_t and the timestep t . The CFM objective of the surface point cloud position is formulated as:

$$\mathcal{L}_{\text{pos}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}(0,1), p(\mathbf{x}_1^s), p(\mathbf{x}_0^s), p(\mathbf{x}_i^s | \mathbf{x}_0^s, \mathbf{x}_1^s)} \|v^{\text{pos}}(\mathbf{x}_t^s, t, C^{\text{rec}}) - (\mathbf{x}_1^s - \mathbf{x}_0^s)\|_2^2, \quad (8)$$

where $\mathcal{U}(0, 1)$ is a uniform distribution on $[0, 1]$. During generation, we first sample from the prior $\mathbf{x}_0^s \sim \mathcal{N}(0, \mathbf{I}_3)$ and solve the probability flow with the learned predictor $v^{\text{pos}}(\cdot)$ using the N -step forward Euler method to get the position of residue j with $t = \{0, \dots, \frac{N-1}{N}\}$:

$$\mathbf{x}_{t+\frac{1}{N}}^s = \mathbf{x}_t^s + \frac{1}{N} v^{\text{pos}}(\mathbf{x}_t^s, t, C^{\text{rec}}). \quad (9)$$

Normal Vector Orientations. The normal vector \mathbf{n}_i^s is a unit vector perpendicular to the tangent plane of the surface at \mathbf{x}_i^s . It reveals essential geometric information of the surface orientation and curvature, directly tied to the protein’s functional (Song et al., 2024; Wu & Li, 2024b). Convex regions, for example, might be more accessible for binding, while concave regions might be better suited for pockets or clefts involved in substrate binding (Laskowski et al., 1996). To capture its orientation, we construct a set of rotation matrices $O^s \subseteq \text{SO}(3)$ with respect to the global frame, whose element is defined by $O_i^s = (\mathbf{n}_i^s, \hat{\mathbf{d}}_i, \mathbf{n}_i^s \times \hat{\mathbf{d}}_i) \in \text{SO}(3)$. Here, $\hat{\mathbf{d}}_i$ is a unit vector orthogonal to \mathbf{n}_i^s and is acquired by normalizing the cross product between \mathbf{n}_i^s and the direction pointing from the surface point \mathbf{x}_i^s to its nearest C_α coordinate. This frame O_i^s effectively describes the geometric relationship between the protein surface and the underlying backbone structure.

The 3D rotation group $\text{SO}(3)$ is a smooth Riemannian manifold, with its tangent space, $\mathfrak{so}(3)$, forming a Lie algebra consisting of skew-symmetric matrices. Elements of $\mathfrak{so}(3)$ can be viewed as infinitesimal rotations around specific axes and represented as rotation vectors in \mathbb{R}^3 (Blanco-Claraco, 2021). In line with prior work (Li et al., 2024; Lin et al., 2024), we adopt the uniform distribution over $\text{SO}(3)$ as the prior $p(O_0^s)$. Just as FM in Euclidean space is based on the shortest path between two points, we extend this idea to $\text{SO}(3)$ by using geodesics, which define the minimal rotational distance between two orientations (Lee, 2018). These geodesics provide a natural framework for interpolating and evolving rotations while respecting the geometry of the manifold (Bose et al., 2023; Yim et al., 2023). The conditional flow ϕ^{ori} and vector field u_t^{ori} are constructed by geodesic interpolation between $O_0^s \subseteq \mathcal{U}(\text{SO}(3))$ and $O_1^s \in p(O^s | C^{\text{rec}})$, with the geodesic distance decreasing linearly over time:

$$\phi_t^{\text{ori}}(O_0^s, O_1^s) = \exp_{O_0^s} \left(t \log_{O_0^s}(O_1^s) \right), \quad (10)$$

$$u_t^{\text{ori}}(O_t^s | O_0^s, O_1^s) = \frac{\log_{O_t^s}(O_1^s)}{1-t}, \quad (11)$$

where $\exp(\cdot)$ and $\log(\cdot)$ are the exponential and logarithm maps on $\text{SO}(3)$ that can be computed efficiently using Rodrigues’ formula (Li et al., 2024). A rotation-equivariant surface network $v^{\text{ori}}(\cdot)$ is applied to predict the vector field u_t^{ori} , represented as rotation vectors. The CFM objective on $\text{SO}(3)$ is formulated as:

$$\mathcal{L}_{\text{ori}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}(0,1), p(O_1^s), p(O_0^s), p(O_t^s | O_0^s, O_1^s)} \left\| v^{\text{ori}}(O_t^s, t, C^{\text{rec}}) - \frac{\log_{O_t^s}(O_1^s)}{1-t} \right\|_{\text{SO}(3)}^2, \quad (12)$$

where the vector field $v^{\text{ori}}(\cdot)$ is defined in the tangent space $\mathfrak{so}(3)$ of $\text{SO}(3)$, with the norm $|\cdot|^2$ derived from the canonical metric on $\text{SO}(3)$. In the inference phase, the process is initialized at $O_0^s \sim \mathcal{U}(\text{SO}(3))$ and proceeds by following the geodesic in $\text{SO}(3)$, taking small steps over time t :

$$O_{t+\frac{1}{N}} = \exp_{O_t^s} \left(\frac{1}{N} v^{\text{ori}}(O_t^s, t, C^{\text{rec}}) \right). \quad (13)$$

Physicochemical Features. FM is typically applied to continuous spaces. However, certain biochemical characteristics take on discrete, categorical values. For example, each surface point can

be classified into three categories based on its hydrogen bonding potential: donor, acceptor, or neutral. This challenge also arises in protein generation tasks that focus solely on structure (Luo et al., 2022), where residue types follow a categorical distribution. To address this, previous studies (Li et al., 2024; Lin et al., 2024) have employed soft one-hot encoding to map categorical distributions to a probability simplex or directly applied FM to multinomial distributions. However, this straightforward approach may result in suboptimal performance for protein co-design. Recently, more advanced FM methods tailored for discrete spaces have been proposed (Campbell et al., 2024; Gat et al., 2024; Stark et al., 2024) to overcome these limitations.

Discrete flow models (DFMs) are generative algorithms designed to operate in discrete spaces by simulating a probability flow that transitions from noise to data. DFMs trace a trajectory of discrete variables that align the noise-to-data flow, allowing the generation of new samples. Building on the work of Campbell et al. (2024); Gat et al. (2024), we implement a DFM using CTMCs for the biochemical properties $\mathbf{Y}^s \in \mathbb{R}^{m \times 1}$. Specifically, we train a neural network $v^{\text{cat}}(\cdot)$ to approximate the true denoising distribution $p_{1|t}(\mathbf{Y}_t^s | \mathbf{Y}_1^s)$. This is done using a cross-entropy loss, where the model learns to predict the clean data point \mathbf{Y}_1^s when given a noisy input $\mathbf{Y}_t^s \sim p_{t|1}(\mathbf{Y}_t^s | \mathbf{Y}_1^s)$ as:

$$\mathcal{L}_{\text{cat}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}(0,1), p(\mathbf{r}_1^s), p_{t|1}(\mathbf{r}_t^s | \mathbf{r}_1^s)} [\log v^{\text{cat}}(\mathbf{Y}_t^s, t, C^{\text{rec}})]. \quad (14)$$

Here, rather than a linear interpolation towards \mathbf{Y}_1^s from a uniform prior $p_{t|1}^{\text{unif}}(\mathbf{Y}_t^s | \mathbf{Y}_1^s) = \text{Cat}(t\delta\{\mathbf{Y}_1^s, \mathbf{Y}_t^s\} + (1-t)\frac{1}{S})$, we adopt an artificially introduced mask state \mathbf{M} and the conditional path becomes (Campbell et al., 2024):

$$p_{t|1}^{\text{mask}}(\mathbf{Y}_t^s | \mathbf{Y}_1^s) = \text{Cat}(t\delta\{\mathbf{Y}_1^s, \mathbf{Y}_t^s\} + (1-t)\delta\{\mathbf{M}, \mathbf{Y}_t^s\}). \quad (15)$$

Notably, $\mathcal{L}^{\text{cat}}(\cdot)$ has a strong relation to the model loglikelihood and the Evidence Lower Bound (ELBO) used to train diffusion models (Campbell et al., 2024). It also does not depend on $R_t(\mathbf{Y}_t^s, j | \mathbf{Y}_1^s)$. There are many choices for $R_t(\mathbf{Y}_t^s, j | \mathbf{Y}_1^s)$ that all generate the same $p_{t|1}(\mathbf{Y}_t^s | \mathbf{Y}_1^s)$. At inference time, we pick the rate matrix for $\mathbf{Y}_t^s \neq j$ as:

$$R_t^*(\mathbf{Y}_t^s, j | \mathbf{Y}_1^s) := \frac{\text{ReLU}(\partial_t p_{t|1}(j | \mathbf{Y}_1^s) - \partial_t p_{t|1}(\mathbf{Y}_t^s | \mathbf{Y}_1^s))}{S \cdot p_{t|1}(\mathbf{Y}_t^s | \mathbf{Y}_1^s)}, \quad (16)$$

where $\text{ReLU}(a) = \max(a, 0)$ and $\partial_t p_{t|1}$ can be found by differentiating our explicit form for $p_{t|1}$ in Equ. 15. This choice of R_t^* assumes $p_{t|1}(\mathbf{Y}_t^s | \mathbf{Y}_1^s) > 0$.

Overall Training Loss. Combining all modalities, the final FM objective is for conditional peptide generation obtained as the weighted sum of three loss functions in Equ. 8, 12, and 14 as well as several additional loss. It can be written as:

$$\mathcal{L}_{\text{CFM}} = \lambda_{\text{pos}} \mathcal{L}_{\text{pos}} + \lambda_{\text{ori}} \mathcal{L}_{\text{ori}} + \lambda_{\text{cat}} \mathcal{L}_{\text{cat}} + \lambda_{\text{con}} \mathcal{L}_{\text{con}} + \lambda_{\text{str}} \mathcal{L}_{\text{str}}, \quad (17)$$

where λ_* are the hyperparameters to control the impact of different loss components. \mathcal{L}_{con} is the loss function for continuous biochemical properties, and \mathcal{L}_{str} is the FM objective for modeling the factorized distribution of residues' positions $\{a_j\}_{j=1}^{n_{\text{pep}}}$, orientations $\{O_j\}_{j=1}^{n_{\text{pep}}}$, amino acid types $\{x_j\}_{j=1}^{n_{\text{pep}}}$, and side-chain torsion angles $\{\chi_j\}_{j=1}^{n_{\text{pep}}}$ as discussed in Equ. 5 and Appendix C. The network details to parameterize the generation procedure is illustrated in Appendix A.

3.2 FLOW MATCHING WITH CONDITIONS

Inspired by the success of controllable image generation (Zhang et al., 2023), we propose peptide design conditioned on key factors c , such as sequence length n_{pep} , cyclicality, and the presence of disulfide bonds. Our objective becomes $p(C^{\text{pep}} | C^{\text{rec}}, c)$, allowing flow models to incorporate additional conditions. Accordingly, the vector field networks are adapted to $v^{\text{pos}}(x_t^s, t, C^{\text{rec}}, c)$, $v^{\text{ori}}(O_t^s, t, C^{\text{rec}}, c)$, and $v^{\text{cat}}(\mathbf{Y}_t^s, t, C^{\text{rec}}, c)$. In practice, we focus on two primary goals: (1) **Cyclic peptides** offer enhanced stability by constraining the backbone, thus reducing conformational flexibility and increasing resistance to enzymatic degradation. This structure form improves binding affinity due to more defined and stable conformations. (2) **Disulfide bonds**, covalent interactions between cysteine residues, assist in proper folding and structural stabilization. These bonds protect peptides from oxidative damage and proteolytic enzymes, enhancing their resistance to harsh

Table 1: Evaluation of methods in the sequence-structure co-design task. The **best** and suboptimal results are labeled **boldly** and underlined.

	Geometry				Energy		Design	
	AAR % \uparrow	RMSD \AA \downarrow	SSR % \uparrow	BSR % \uparrow	Stb. % \uparrow	Aff. % \uparrow	Des. % \uparrow	Div. \uparrow
RFdiffusion (Watson et al., 2023)	40.14	4.17	63.86	26.71	26.82	16.53	78.52	0.38
ProteinGen (Lisanza et al., 2023)	45.82	4.35	29.15	24.62	23.48	13.47	71.82	0.54
Diffusion (Luo et al., 2022)	47.04	3.28	74.89	49.83	15.34	17.13	48.54	<u>0.57</u>
PepFlow (Li et al., 2024)	<u>51.25</u>	<u>2.07</u>	<u>83.46</u>	<u>86.89</u>	18.15	21.37	65.22	0.42
SurfFlow	54.07	1.96	85.11	87.38	<u>22.46</u>	22.51	<u>73.60</u>	0.61

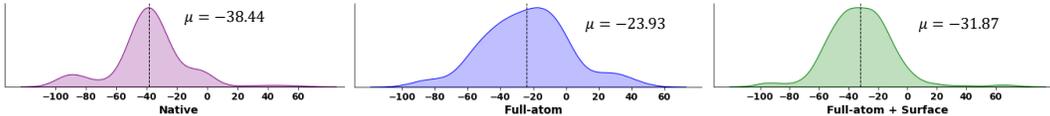


Figure 3: Binding energy distributions of designed and native peptides, where the lower is better.

industrial conditions. Additionally, disulfide bonds can improve biological activity by creating conformational constraints, further enhancing the therapeutic potential of peptide-based drugs.

4 EXPERIMENTS

We comprehensively evaluate SurfFlow on unconditioned and conditioned sequence-structure co-design tasks and the side-chain packing problem. For benchmarking, we use the PepMerge dataset (Li et al., 2024) derived from PepBDB (Wen et al., 2019) and Q-BioLip (Wei et al., 2024). Following the methodology of Li et al. (2024), we cluster the peptide-protein complexes based on 40% sequence identity using mmseqs2 (Steinegger & Söding, 2017), after filtering out duplicates and applying empirical criteria (*e.g.*, resolution $< 4\text{\AA}$, peptide length between 3 and 25). This process yields 8,365 non-redundant complexes across 292 clusters. To ensure a fair and direct comparison, we employ the same test set as Li et al. (2024), consisting of 10 clusters and 158 complexes. More experimental details and additional results are elaborated on Appendix B.

4.1 UNCONDITIONED SEQUENCE-STRUCTURE CO-DESIGN

Baselines. Two lines of state-of-the-art protein design approaches are chosen as baselines. The first kind ignores the side-chain conformations, including RFDiffusion (Watson et al., 2023) and ProteinGen (Lisanza et al., 2023). RFDiffusion produces protein backbones and sequences are later forecast by ProteinMPNN (Dauparas et al., 2022). ProteinGen improves RFDiffusion by jointly sampling backbones and corresponding sequences. The other kind considers a full-atom style protein generation, including Diffusion (Luo et al., 2022) and PepFlow (Li et al., 2024).

Evaluation Metrics. Generated peptides are evaluated from three key aspects. (1) **Geometry:** Designed peptides should closely resemble native sequences and structures. We use the amino acid recovery rate (**AAR**) to quantify sequence identity between generated peptides and ground truth. Structural similarity is assessed through the root-mean-square deviation (**RMSD**) of C_{α} atoms after aligning the complexes. Secondary-structure similarity ratio (**SSR**) measures the proportion of shared secondary structures, while the binding site ratio (**BSR**) compares the overlap between the binding sites of the generated and native peptides on the target protein. (2) **Energy:** Our goal is to design high-affinity peptide binders that enhance the stability of protein-peptide complexes. **Affinity** is defined as the percentage of generated peptides with higher binding affinities (lower binding energies) than the native peptide, while **Stability** indicates the proportion of complexes with lower total energy than the native state. Energy calculations are performed using Rosetta (Alford et al., 2017). (3) **Design: Designability** reflects the consistency between designed sequences and structures. It is measured by the fraction of sequences that can fold into structures similar to their corresponding generated forms, with C_{α} RMSD $< 2\text{\AA}$ as the threshold. We use ESMFold (Lin

378
379
380
381
382
383
384
385
386
387
388
389
390
391

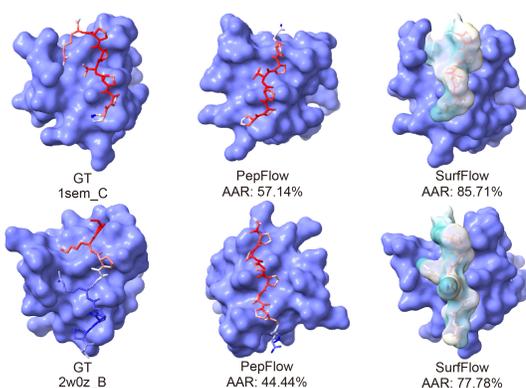


Figure 4: Peptide designed by different algorithms and references.

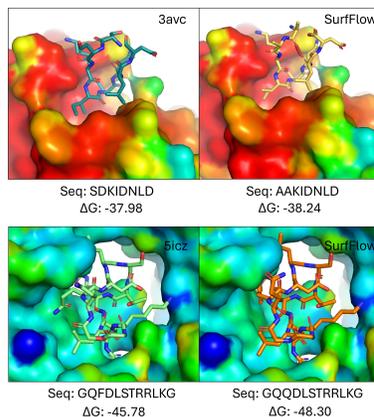


Figure 5: Peptide design with the cyclic condition.

392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412

et al., 2022) to refold the sequences. **Diversity**, measured as the average of one minus the pairwise TM-Score (Zhang & Skolnick, 2004), indicates structural dissimilarity among designed peptides.

Results. Table 1 illustrates that our SurfFlow generates significantly more diversified and consistent peptides with better binding energy and closer resembles compared to the baselines. Specifically, SurfFlow achieves the state-of-the-art AAR of 54.07% and RMSD of 1.96 Å with improvements of 5.51% and 5.31% over the full-atom PepFlow. Besides, it also owns a stronger capacity of designing peptides with more accurate binding ratio of 87.38% and higher affinity proportion of 22.51% (see Fig. 3). These statistics confirm the benefits of explicitly modeling the surface geometry and biochemical constraints. It is worth mentioning that the decoupled approach, RFDiffusion, attains better Stability (26.82% v.s. 22.46%) and Designability (78.52% v.s. 73.60%), as it is trained in the entire PDB and are prone towards structures with more stable motifs (Li et al., 2024). Figure 4 presents two examples of designed peptides by full-atom PepFlow and surface-based SurfFlow and Appendix D.2 shows how the positions of surface point clouds evolves over the entire time period. Evidently, SurfFlow generates peptides with topologically similar geometries, irrespective of their native length. Notably, the produced peptides share comparable side-chain compositions and conformations, enabling efficient interactions with the target protein at the appropriate binding site.

413
414

4.2 CONDITIONED SEQUENCE-STRUCTURE CO-DESIGN

415
416
417
418
419
420
421
422
423
424
425

There are several key characteristics of effective peptides widely recognized in the biological community. Cyclic peptides, for instance, are polypeptide chains with a circular bond sequence. Many occur naturally and exhibit antimicrobial or toxic properties, while others are laboratory-synthesized (Jensen, 2009). These cyclic peptides typically demonstrate high resistance to digestion, making them attractive to researchers developing novel oral medications like antibiotics and immunosuppressants (Craik, 2006). Additionally, evidence indicates that disulfide patterns are crucial in the folding and structural stabilization of peptides. The deliberate introduction of disulfide bridges into natural or engineered peptides can often enhance their biological activities, specificities, and stabilities (Annis et al., 1997). Given these insights, we propose a conditional co-design challenge. This challenge involves calculating the proportions of cyclic peptides and peptides containing disulfide bonds within generated peptide sets for evaluation purposes.

426
427
428
429
430
431

Results. Table 2 documents the results with several Key findings. Without conditional constraints, neither full-atom deep generative models nor our omni-design method met the requirements for cyclic peptides or disulfide bridges. Even with the incorporation of molecular surfaces, only 2-4% of the 189 designed peptides in the test samples exhibited cyclic peptides or contained disulfide bonds. In contrast, When trained with additional conditions and prompted to generate peptides with specific properties, SurfFlow significantly increases the proportion of peptides with desired characteristics, with cyclic peptides from 2.67% to 8.02% and disulfide bonds from 4.27% to 9.10%. Figure 5 visualizes two

Table 2: Proportions of cyclic peptides and peptides with disulfide bonds designed by different mechanisms and in the original PepMerge dataset.

Metrics	PepFlow	SurfFlow (w/o c)	SurfFlow	PepMerge
Cyclic	2.13%	2.67%	8.02%	15.50%
Disulfide Bond%	3.21%	4.27%	9.10%	18.18%

Table 3: Evaluation results of the side-chain packing task.

	MSE $^{\circ}$ \downarrow				Correct % \uparrow
	χ_1	χ_2	χ_3	χ_4	
Rosetta	38.31	43.23	53.61	71.67	57.03
SCWRL4	30.06	40.40	49.71	53.79	60.54
DLPacker	22.44	35.65	58.53	61.70	60.91
AttnPacker	19.04	28.49	40.16	60.04	61.46
DiffPack	17.92	26.08	36.20	67.82	62.58
PepFlow	17.38	24.71	33.63	58.49	62.79
SurfFlow	17.13	23.86	31.97	55.08	63.02

examples with the cyclic condition. Interestingly, in the case of 3AVC, where the native peptide is non-cyclic, SurfFlow generated a novel cyclic peptide with a lower binding energy ($\Delta G = -37.98$ v.s. $\Delta G = -38.24$). For 5ICZ, which originally had a cyclic peptide, SurfFlow designed a new type of cyclic peptide with a better binding energy ($\Delta G = -45.78$ v.s. $\Delta G = -48.30$).

4.3 SIDE-CHAIN PACKING

Side-chain packing is a critical task in protein structure modeling, focusing on the prediction of peptide side-chain angles. Our approach generates 64 distinct side-chain conformations for each peptide using multiple models, employing a partial sampling strategy to efficiently recover the most probable side-chain angles. This method allows us to navigate the conformational space effectively while reducing computational overhead and maintaining high accuracy.

Baselines. We compare SurfFlow against several established approaches. These include energy-based methods such as RosettaPacker (Leman et al., 2020) and SCWRL4 (Krivov et al., 2009), which rely on physical energy minimization techniques for side-chain positioning. Additionally, we evaluate against learning-based models including DLPacker (Misiura et al., 2022), AttnPacker (McPartlon & Xu, 2023), and DiffPack (Zhang et al., 2024), which leverage various DL strategies like attention mechanisms and diffusion to predict side-chain configurations based on learned representations of protein structure.

Metrics. Two primary metrics are employed for evaluation. First, we calculate the Mean Absolute Error (MAE) of four key torsion angles: χ_1 , χ_2 , χ_3 , and χ_4 . Given the inherent flexibility of side chains and the importance of small deviations in structural biology, we also report the proportion of predictions that fall within a 20° deviation from the ground truth. This additional metric captures the practical accuracy of side-chain prediction, emphasizing how well the models perform within biologically relevant error margins.

Results. Table 3 reports the results, it can be observed that the incorporation of molecular surfaces contributes to more accurate predictions of all four side-chain angles compared to full-atom models and other baselines. SurfFlow attains the highest correct ratio of 63.02%. The enhanced accuracy suggests that surface information provides crucial context for predicting side-chain configurations, likely by better representing the local environment and potential interactions that influence side-chain positioning.

5 RELATED WORKS

Protein Design with Generative DL. Generative models have made significant strides in protein design, particularly in applications like engineering enzyme active sites (Yeh et al., 2023). These

486 methods generally fall into three categories: sequence design, structure design, and co-design. In
487 sequence design, protein sequences are crafted through techniques such as oracle-guided directed
488 evolution (Jain et al., 2022) or by leveraging protein language models (Madani et al., 2020; Verkuil
489 et al., 2022). Another common strategy, called fix-backbone sequence design, involves generating
490 sequences that fit a predefined backbone structure (Ingraham et al., 2019; Jing et al., 2020; Hsu et al.,
491 2022; Gao et al., 2022b;a; Zheng et al., 2023b). Given the importance of 3D structural information
492 in proteins, some approaches focus on first generating protein backbone structures (Anand & Achim,
493 2022), which are then paired with sequence prediction models like ProteinMPNN (Dauparas et al.,
494 2022) to determine the matching sequence. Co-design methods, which generate sequence-structure
495 pairs simultaneously, are especially useful for antibody design (Jin et al., 2021; Kong et al., 2022;
496 Wu & Li, 2024a). Recent studies also emphasize side-chain interactions in conditional protein
497 generation, enabling full atomic detail (Martinkus et al., 2024; Krishna et al., 2024). However, none
498 of these methods have addressed the simultaneous generation of protein sequence, structure, and
499 surfaces.

500 From a technical standpoint, diffusion and flow-based models have become popular for designing
501 novel and diverse proteins (Ingraham et al., 2023; Lin & AlQuraishi, 2023). Some studies focus on
502 sequence generation alone using discrete diffusion models (Alamdari et al., 2023; Frey et al., 2023;
503 Gruver et al., 2024; Yi et al., 2024). These models are also applied to generate protein structures
504 in 3D or SE(3) space (Trippe et al., 2022; Anand & Achim, 2022; Bose et al., 2023; Yim et al.,
505 2023; Wu et al., 2024). Among them, RFDiffusion (Watson et al., 2023) has seen notable success,
506 with wet-lab validation of generated proteins. However, these methods often require a separate
507 model for sequence generation. In co-design, earlier efforts include ProteinGenerator (Lisanza et al.,
508 2023), which uses Euclidean diffusion over one-hot encoded amino acids while predicting structure
509 at each step using RosettaFold (Baek et al., 2021). Protpardelle (Chu et al., 2024) applies Euclidean
510 diffusion to structure while iteratively predicting the sequence. Multiflow (Campbell et al., 2024)
511 introduces a DFM model over protein sequences, offering flexible conditioning during inference.
512 Early co-design methods by (Luo et al., 2022; Shi et al., 2022) focused on designing CDR loops
513 in antibodies. Co-design for peptides is also gaining traction, with models like PepFlow (Li et al.,
514 2024) and PPIFlow (Lin et al., 2024) excelling in full-atom peptide design using multi-modality
515 FM. Finally, PepGLAD (Kong et al., 2024) explored peptide structure and sequence diffusion, but
516 no code has been made available.

517 **Molecular Surface Modeling.** The properties of a protein’s molecular surface are crucial in de-
518 termining the nature and strength of its interactions with other molecules. This surface is shaped
519 by van der Waals (vdW) radii (Connolly, 1983) and is often represented as meshes created from
520 signed distance functions. MaSIF (Gainza et al., 2020) was a pioneering effort in applying mesh-
521 based geometric deep learning to abstract internal protein folds and study protein interactions. Later,
522 Sverrisson et al. (2021) simplified the process by representing molecular surfaces as point clouds,
523 assigning atom types to each point to reduce pre-computation costs. Other key works have inte-
524 grated protein surface data with structural information in a multimodal approach (Somnath et al.,
525 2021), using advanced pretraining techniques (Wu & Li, 2024b) and implicit neural representations
526 (INRs)(Park et al., 2019) for self-supervised learning Lee et al. (2023) and dynamic structure mod-
527 eling Sun et al. (2024). Despite these developments, protein design based on surface characteristics
528 remains relatively unexplored. However, recent progress, such as Gainza et al. (2023)’s extension
529 of MaSIF for *de novo* binder design, and SurfPro (Song et al., 2024), which eliminates the need for
530 handcrafted feature extraction, has begun to address this gap by generating functional proteins di-
531 rectly from surface data. SurfFlow stands out as the first approach to generate all protein modalities
532 simultaneously using flow-based algorithms.

533 6 CONCLUSION

534 This work presents SurfFlow, a novel deep generative model that produces all protein modalities –
535 sequence, structure, and surface – concurrently. We apply SurfFlow to solving a specific peptide
536 design challenge and integrate some key characteristics like cyclicity and disulfide bonds into the
537 generation process. Empirical results prove the reasonability and promise of considering molecular
538 surfaces for protein discovery. Limitation and future work is elucidated in Appendix E.
539

540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593

REFERENCES

- Sarah Alamdari, Nitya Thakkar, Rianne van den Berg, Alex Xijie Lu, Nicolo Fusi, Ava Pardis Amini, and Kevin K Yang. Protein generation with evolutionary diffusion: sequence is all you need. *bioRxiv*, pp. 2023–09, 2023.
- Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*, 2022.
- 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.
- Rebecca F Alford, Andrew Leaver-Fay, Jeliasko R Jeliaskov, 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.
- Namrata Anand and Tudor Achim. Protein structure and sequence generation with equivariant denoising diffusion probabilistic models. *arXiv preprint arXiv:2205.15019*, 2022.
- Ioana Annis, Balazs Hargittai, and George Barany. Disulfide bond formation in peptides. *Methods in enzymology*, 289:198–221, 1997.
- Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, 2021.
- Gaurav Bhardwaj, Vikram Khipple Mulligan, Christopher D Bahl, Jason M Gilmore, Peta J Harvey, Olivier Cheneval, Garry W Buchko, Surya VSRK Pulavarti, Quentin Kaas, Alexander Eletsky, et al. Accurate de novo design of hyperstable constrained peptides. *Nature*, 538(7625):329–335, 2016.
- José Luis Blanco-Claraco. A tutorial on $se(3)$ transformation parameterizations and on-manifold optimization. *arXiv preprint arXiv:2103.15980*, 2021.
- Avishek Joey Bose, Tara Akhound-Sadegh, Kilian Fatras, Guillaume Hugué, Jarrid Rector-Brooks, Cheng-Hao Liu, Andrei Cristian Nica, Maksym Korablyov, Michael Bronstein, and Alexander Tong. $Se(3)$ -stochastic flow matching for protein backbone generation. *arXiv preprint arXiv:2310.02391*, 2023.
- Laura K Buckton, Marwa N Rahimi, and Shelli R McAlpine. Cyclic peptides as drugs for intracellular targets: the next frontier in peptide therapeutic development. *Chemistry—A European Journal*, 27(5):1487–1513, 2021.
- 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.
- Ricky TQ Chen and Yaron Lipman. Riemannian flow matching on general geometries. *arXiv preprint arXiv:2302.03660*, 2023.
- Alexander E Chu, Jinho Kim, Lucy Cheng, Gina El Nesr, Minkai Xu, Richard W Shuai, and Po-Ssu Huang. An all-atom protein generative model. *Proceedings of the National Academy of Sciences*, 121(27):e2311500121, 2024.
- Michael L Connolly. Analytical molecular surface calculation. *Journal of applied crystallography*, 16(5):548–558, 1983.
- David J Craik. Seamless proteins tie up their loose ends. *Science*, 311(5767):1563–1564, 2006.
- 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 proteinmpnn. *Science*, 378(6615):49–56, 2022.

594 Warren L DeLano et al. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein*
595 *Crystallogr*, 40(1):82–92, 2002.

596

597 Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances*
598 *in neural information processing systems*, 34:8780–8794, 2021.

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

601

602 Nathan C Frey, Daniel Berenberg, Karina Zadorozhny, Joseph Kleinhenz, Julien Lafrance-Vanasse,
603 Isidro Hotzel, Yan Wu, Stephen Ra, Richard Bonneau, Kyunghyun Cho, et al. Protein discovery
604 with discrete walk-jump sampling. *arXiv preprint arXiv:2306.12360*, 2023.

605 Pablo Gainza, Freyr Sverrisson, Frederico Monti, Emanuele Rodola, Davide Boscaini, Michael M
606 Bronstein, and Bruno E Correia. Deciphering interaction fingerprints from protein molecular
607 surfaces using geometric deep learning. *Nature Methods*, 17(2):184–192, 2020.

608 Pablo Gainza, Sarah Wehrle, Alexandra Van Hall-Beauvais, Anthony Marchand, Andreas Scheck,
609 Zander Hartevelde, Stephen Buckley, Dongchun Ni, Shuguang Tan, Freyr Sverrisson, et al. De
610 novo design of protein interactions with learned surface fingerprints. *Nature*, 617(7959):176–
611 184, 2023.

612

613 Zhangyang Gao, Cheng Tan, Pablo Chacon, and Stan Z Li. Pifold: Toward effective and efficient
614 protein inverse folding. *arXiv preprint arXiv:2209.12643*, 2022a.

615 Zhangyang Gao, Cheng Tan, and Stan Z Li. Alphadesign: A graph protein design method and
616 benchmark on alphafolddb. *arXiv preprint arXiv:2202.01079*, 2022b.

617

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

620 Nate Gruver, Samuel Stanton, Nathan Frey, Tim GJ Rudner, Isidro Hotzel, Julien Lafrance-Vanasse,
621 Arvind Rajpal, Kyunghyun Cho, and Andrew G Wilson. Protein design with guided discrete
622 diffusion. *Advances in neural information processing systems*, 36, 2024.

623 Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, and Jianzhu Ma. 3d equiv-
624 ariant diffusion for target-aware molecule generation and affinity prediction. *arXiv preprint*
625 *arXiv:2303.03543*, 2023.

626

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

629 Chloe Hsu, Robert Verkuil, Jason Liu, Zeming Lin, Brian Hie, Tom Sercu, Adam Lerer, and Alexan-
630 der Rives. Learning inverse folding from millions of predicted structures. *bioRxiv*, 2022.

631

632 John Ingraham, Vikas Garg, Regina Barzilay, and Tommi Jaakkola. Generative models for graph-
633 based protein design. *Advances in neural information processing systems*, 32, 2019.

634 John B Ingraham, Max Baranov, Zak Costello, Karl W Barber, Wujie Wang, Ahmed Ismail, Vincent
635 Frappier, Dana M Lord, Christopher Ng-Thow-Hing, Erik R Van Vlack, et al. Illuminating protein
636 space with a programmable generative model. *Nature*, 623(7989):1070–1078, 2023.

637

638 Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Bonaventure FP
639 Dossou, Chanakya Ajit Ekbote, Jie Fu, Tianyu Zhang, Michael Kilgour, Dinghuai Zhang, et al.
640 Biological sequence design with gflownets. In *International Conference on Machine Learning*,
641 pp. 9786–9801. PMLR, 2022.

642 Knud J Jensen. Peptide and protein design for biopharmaceutical applications. 2009.

643 Wengong Jin, Jeremy Wohlwend, Regina Barzilay, and Tommi Jaakkola. Iterative refinement graph
644 neural network for antibody sequence-structure co-design. *arXiv preprint arXiv:2110.04624*,
645 2021.

646

647 Bowen Jing, Stephan Eismann, Patricia Suriana, Raphael JL Townshend, and Ron Dror. Learning
from protein structure with geometric vector perceptrons. *arXiv preprint arXiv:2009.01411*, 2020.

648 Susan Jones and Janet M Thornton. Principles of protein-protein interactions. *Proceedings of the*
649 *National Academy of Sciences*, 93(1):13–20, 1996.

650

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

654 Panagiotis L Kastiris and Alexandre MJJ Bonvin. On the binding affinity of macromolecular in-
655 teractions: daring to ask why proteins interact. *Journal of The Royal Society Interface*, 10(79):
656 20120835, 2013.

657

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

660 Xiangzhe Kong, Wenbing Huang, and Yang Liu. Full-atom peptide design with geometric latent
661 diffusion. *arXiv preprint arXiv:2402.13555*, 2024.

662

663 Rohith Krishna, Jue Wang, Woody Ahern, Pascal Sturmfels, Preetham Venkatesh, Indrek Kalvet,
664 Gyu Rie Lee, Felix S Morey-Burrows, Ivan Anishchenko, Ian R Humphreys, et al. Generalized
665 biomolecular modeling and design with rosettafold all-atom. *Science*, 384(6693):ead12528, 2024.

666 Georgii G Krivov, Maxim V Shapovalov, and Roland L Dunbrack Jr. Improved prediction of protein
667 side-chain conformations with scwrl4. *Proteins: Structure, Function, and Bioinformatics*, 77(4):
668 778–795, 2009.

669 Roman A Laskowski, Nicholas M Luscombe, Mark B Swindells, and Janet M Thornton. Protein
670 clefts in molecular recognition and function. *Protein science: a publication of the Protein Society*,
671 5(12):2438, 1996.

672

673 John M Lee. *Introduction to Riemannian manifolds*, volume 2. Springer, 2018.

674

675 Youhan Lee, Hasun Yu, Jaemyung Lee, and Jaehoon Kim. Pre-training sequence, structure, and
676 surface features for comprehensive protein representation learning. In *The Twelfth International*
677 *Conference on Learning Representations*, 2023.

678 Julia Koehler Leman, Brian D Weitzner, Steven M Lewis, Jared Adolf-Bryfogle, Nawsad Alam, Re-
679becca F Alford, Melanie Aprahamian, David Baker, Kyle A Barlow, Patrick Barth, et al. Macro-
680 molecular modeling and design in rosetta: recent methods and frameworks. *Nature methods*, 17
681 (7):665–680, 2020.

682 Jiahua Li, Chaoran Cheng, Zuofan Wu, Ruihan Guo, Shitong Luo, Zhizhou Ren, Jian Peng, and
683 Jianzhu Ma. Full-atom peptide design based on multi-modal flow matching. *arXiv preprint*
684 *arXiv:2406.00735*, 2024.

685

686 Haitao Lin, Odin Zhang, Huifeng Zhao, Dejun Jiang, Lirong Wu, Zicheng Liu, Yufei Huang, and
687 Stan Z Li. Ppflow: Target-aware peptide design with torsional flow matching. *bioRxiv*, pp. 2024–
688 03, 2024.

689 Yeqing Lin and Mohammed AlQuraishi. Generating novel, designable, and diverse protein struc-
690 tures by equivariantly diffusing oriented residue clouds. *arXiv preprint arXiv:2301.12485*, 2023.

691

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

695 Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching
696 for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

697

698 Sidney Lyayuga Lisanza, Jake Merle Gershon, Sam Tipps, Lucas Arnoldt, Samuel Hendel,
699 Jeremiah Nelson Sims, Xinting Li, and David Baker. Joint generation of protein sequence and
700 structure with rosettafold sequence space diffusion. *bioRxiv*, pp. 2023–05, 2023.

701

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and
transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

702 Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific
703 antibody design and optimization with diffusion-based generative models. *bioRxiv*, 2022.
704

705 Ali Madani, Bryan McCann, Nikhil Naik, Nitish Shirish Keskar, Namrata Anand, Raphael R Eguchi,
706 Po-Ssu Huang, and Richard Socher. Progen: Language modeling for protein generation. *arXiv*
707 *preprint arXiv:2004.03497*, 2020.

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

712 Matthew McPartlon and Jinbo Xu. An end-to-end deep learning method for protein side-chain pack-
713 ing and inverse folding. *Proceedings of the National Academy of Sciences*, 120(23):e2216438120,
714 2023.

715 Mikita Misiura, Raghav Shroff, Ross Thyer, and Anatoly B Kolomeisky. Dlpacker: Deep learning
716 for prediction of amino acid side chain conformations in proteins. *Proteins: Structure, Function,*
717 *and Bioinformatics*, 90(6):1278–1290, 2022.

718 Markus Muttenthaler, Glenn F King, David J Adams, and Paul F Alewood. Trends in peptide drug
719 discovery. *Nature reviews Drug discovery*, 20(4):309–325, 2021.

720 James R Norris. *Markov chains*. Number 2. Cambridge university press, 1998.

721

722 Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove.
723 Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings*
724 *of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 165–174, 2019.

725

726 Vishva Saravanan Ramasubramanian, Soham Choudhuri, and Bhaswar Ghosh. A hybrid diffusion
727 model for stable, affinity-driven, receptor-aware peptide generation. *bioRxiv*, pp. 2024–03, 2024.

728

729 Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. Hierarchical text-
730 conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, 1(2):3, 2022.

731

732 Emma C Robinson, Saad Jbabdi, Matthew F Glasser, Jesper Andersson, Gregory C Burgess,
733 Michael P Harms, Stephen M Smith, David C Van Essen, and Mark Jenkinson. Msm: a new
734 flexible framework for multimodal surface matching. *Neuroimage*, 100:414–426, 2014.

735

736 Victor Garcia Satorras, Emiel Hooeboom, and Max Welling. E (n) equivariant graph neural net-
737 works. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

738

739 Chence Shi, Chuanrui Wang, Jiarui Lu, Bozitao Zhong, and Jian Tang. Protein sequence and struc-
740 ture co-design with equivariant translation. *arXiv preprint arXiv:2210.08761*, 2022.

741

742 Vignesh Ram Somnath, Charlotte Bunne, and Andreas Krause. Multi-scale representation learning
743 on proteins. *Advances in Neural Information Processing Systems*, 34:25244–25255, 2021.

744

745 Zhenqiao Song, Tinglin Huang, Lei Li, and Wengong Jin. Surfpro: Functional protein design based
746 on continuous surface. *arXiv preprint arXiv:2405.06693*, 2024.

747

748 Hannes Stark, Bowen Jing, Chenyu Wang, Gabriele Corso, Bonnie Berger, Regina Barzilay, and
749 Tommi Jaakkola. Dirichlet flow matching with applications to dna sequence design. *arXiv*
750 *preprint arXiv:2402.05841*, 2024.

751

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

754

755 Daiwen Sun, He Huang, Yao Li, Xinqi Gong, and Qiwei Ye. Dsr: dynamical surface representation
as implicit neural networks for protein. *Advances in Neural Information Processing Systems*, 36,
2024.

Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, and Hanjun Dai. Score-based continuous-time
discrete diffusion models. *arXiv preprint arXiv:2211.16750*, 2022.

756 Freyr Sverrisson, Jean Feydy, Bruno E Correia, and Michael M Bronstein. Fast end-to-end learning
757 on protein surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern*
758 *Recognition*, pp. 15272–15281, 2021.

759

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

763

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

767

768 Peter Vanhee, Almer M van der Sloot, Erik Verschueren, Luis Serrano, Frederic Rousseau, and Joost
769 Schymkowitz. Computational design of peptide ligands. *Trends in biotechnology*, 29(5):231–239,
2011.

770

771 Robert Verkuil, Ori Kabeli, Yilun Du, Basile IM Wicky, Lukas F Milles, Justas Dauparas, David
772 Baker, Sergey Ovchinnikov, Tom Sercu, and Alexander Rives. Language models generalize be-
773 yond natural proteins. *BioRxiv*, pp. 2022–12, 2022.

774

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

777

778 Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eise-
779 nach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of
780 protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

781

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

784

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

787

788 Fang Wu and Stan Z Li. A hierarchical training paradigm for antibody structure-sequence co-design.
Advances in Neural Information Processing Systems, 36, 2024a.

789

790 Fang Wu and Stan Z Li. Surface-vqmae: Vector-quantized masked auto-encoders on molecular
791 surfaces. In *International Conference on Machine Learning*, pp. 53619–53634. PMLR, 2024b.

792

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

795

796 Andy Hsien-Wei Yeh, Christoffer Norn, Yakov Kipnis, Doug Tischer, Samuel J Pellock, Declan
797 Evans, Pengchen Ma, Gyu Rie Lee, Jason Z Zhang, Ivan Anishchenko, et al. De novo design of
798 luciferases using deep learning. *Nature*, 614(7949):774–780, 2023.

799

800 Kai Yi, Bingxin Zhou, Yiqing Shen, Pietro Liò, and Yuguang Wang. Graph denoising diffusion for
inverse protein folding. *Advances in Neural Information Processing Systems*, 36, 2024.

801

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

805

806 Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image
807 diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*,
808 pp. 3836–3847, 2023.

809

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.

810 Yangtian Zhang, Zuobai Zhang, Bozитай Zhong, Sanchit Misra, and Jian Tang. Diffpack: A torsional
 811 diffusion model for autoregressive protein side-chain packing. *Advances in Neural Information*
 812 *Processing Systems*, 36, 2024.

813 Qinqing Zheng, Matt Le, Neta Shaul, Yaron Lipman, Aditya Grover, and Ricky TQ Chen. Guided
 814 flows for generative modeling and decision making. *arXiv preprint arXiv:2311.13443*, 2023a.

815 Zaixiang Zheng, Yifan Deng, Dongyu Xue, Yi Zhou, Fei Ye, and Quanquan Gu. Structure-informed
 816 language models are protein designers. In *International conference on machine learning*, pp.
 817 42317–42338. PMLR, 2023b.

821 A PARAMETERIZATION WITH NETWORKS

822
 823 In order to model the joint distribution of the peptide based on its target protein $p(C^{\text{pep}} | C^{\text{rec}})$,
 824 we adopt an encoder-decoder framework to generate peptides. The encoder extracts the geometric
 825 and biochemical features of the receptor C^{rec} as the condition of the generation process, while the
 826 decoder regresses the vector fields of our multi-modality flow matching architecture.

827 **Encoder.** We first utilize a time-independent equivariant geometric encoder to capture the context
 828 information of the receptor. Specifically, it takes the sequence and structure of the target protein C_{rec}
 829 and computes the hidden residue representations $\mathbf{h}_{\text{rec}} \in \mathbb{R}^{n_{\text{rec}} \times \psi_{\text{rec}}}$ and the residue-pair embeddings
 830 $\mathbf{z}_{\text{rec}} \in \mathbb{R}^{n_{\text{rec}} \times \psi_{\text{pair}}}$.

831
 832 **Decoder.** The decoder receives $(\mathbf{h}_{\text{rec}}, \mathbf{z}_{\text{rec}})$ and is time-dependent. It consists of two geometric
 833 networks: one is the 6-layer Invariant Point Attention (IPA) module (Jumper et al., 2021) to regress
 834 the vector fields of the internal structures $\{(a_j, O_j, \mathbf{x}_j, \chi_j)\}_{j=1}^{n_{\text{pep}}}$, and the other is an variant of
 835 equivariant graph neural networks (EGNN) (Satorras et al., 2021) to regress the vector fields of
 836 the surface geometry $\{(\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \mathbf{Y}_i^s)\}_{i=1}^m$. To be specific, we first sample a random timestep
 837 $t \sim \mathcal{U}(0, 1)$ to construct the time-dependent vector fields for every modality of the peptide C_{pep} ,
 838 containing sequence, structure, and surface. Both IPA and EGNN take the timestamp t , the interplant
 839 state of peptide’s internal structure $(a_t, O_t, \mathbf{x}_t, \chi_t)$ as well as receptor’s information $(\mathbf{h}_{\text{rec}}, \mathbf{z}_{\text{rec}})$ as
 840 input, and the interplant state of peptide’s surface $(\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \mathbf{Y}_i^s)$ is also forwarded into EGNN.
 841 Subsequently, IPA recovers the internal structure of original peptide $(\hat{a}_1, \hat{O}_1, \hat{\mathbf{x}}_1, \hat{\chi}_1)$, while the
 842 3-layer EGNN recovers the surface $(\hat{\mathbf{x}}_1^s, \hat{\mathbf{n}}_1^s, \hat{\tau}_1^s, \hat{\mathbf{Y}}_1^s)$. Moreover, two additional losses containing
 843 the backbone position loss $\mathcal{L}_{\text{bb}}(\theta)$ and the torsion angle loss $\mathcal{L}_{\text{tor}}(\theta)$ (Li et al., 2024) is imposed for
 844 extra constraint.

845
 846 **Equivariance.** The joint distribution $p(C^{\text{pep}} | C^{\text{rec}})$ must satisfy the roto-translational equivari-
 847 ance to ensure the generalization. That is, for any translation vector $\epsilon \in \mathbb{R}^3$ and for any orthogonal
 848 matrix $O \in \mathbb{R}^{3 \times 3}$, it should satisfy:

$$849 \quad p(OC^{\text{pep}} + \epsilon | OC^{\text{rec}} + \epsilon), \quad (18)$$

850
 851 where $OC^{\text{pep}} + \epsilon = \{(a_j, O_j, O\mathbf{x}_j + \epsilon, \chi_j)\}_{j=1}^{n_{\text{pep}}} \cup \{(O\mathbf{x}_i^s + \epsilon, \mathbf{n}_i^s, \tau_i^s, \mathbf{Y}_i^s)\}_{i=1}^m$. Following the
 852 standard operation called zero-mass-center (Yim et al., 2023; Lin et al., 2024), we subtract the mass
 853 center of the receptor from all inputs’ coordinates to achieve the invariance to translation, which
 854 also improves the training stability. Moreover, it can be proven that when the prior distributions
 855 $p(a_0)$, $p(O_0)$, $p(\mathbf{x}_0)$, $p(\chi_0)$, $p(\mathbf{x}_0^s)$, $p(\mathbf{n}_0^s)$, $p(\tau_0^s)$, and $p(\mathbf{Y}_0^s)$ are SE(3)-invariant, while the vector
 856 fields $v^{\text{cat}}(\cdot)$ and $v^{\text{con}}(\cdot)$ are SE(3)-invariant, and the vector field $v^{\text{ori}}(\cdot)$ is SO(3)-equivariant and
 857 T(3)-invariant, and the vector field $v^{\text{pos}}(\cdot)$ is SE(3)-equivariant, then the density $p(C^{\text{pep}} | C^{\text{rec}})$
 858 generated by the ODE sampling process is SE(3)-equivariant. Notably, the choice of IPA and EGNN
 859 guarantees the equivariance and invariance requirement of those vector fields.

860
 861 **Conditional Design.** Controlling the output of deep generative models such as diffusions or flows
 862 has become a hotspot in recent years. Apart from the receptor C_{rec} , we often want to create pep-
 863 tides with specific conditions, such as cyclicity and disulfide bonds. Conventionally, conditional
 controls involve classifier guidance (Dhariwal & Nichol, 2021) and classifier-free guidance (Ho

et al., 2020). Evidence indicates that with enough high-quality training data, classifier-free guidance tends to yield better results, being able to generate an almost infinite number of sample categories without the need to retrain a classifier architecture. This leads to wide usage of classifier-free guidance in modern AI systems (Ramesh et al., 2022; Zheng et al., 2023a). Due to this superiority, we adopt the classifier-free guidance to realize the insertion of peptide conditions c and transfer $p(C_{\text{pep}}|C_{\text{rec}})$ to $p(C_{\text{pep}}|C_{\text{rec}}, c)$. Specifically, we denote the null condition by \emptyset by convention and set $p(C_{\text{pep}}|C_{\text{rec}}, \emptyset) := p(C_{\text{pep}}|C_{\text{rec}})$ and $u_t(\cdot|\cdot, \emptyset) := u_t(\cdot|\cdot)$. Then taking the surface point positions for example, our training loss becomes (Zheng et al., 2023a):

$$\mathcal{L}_{\text{pos}}(\theta) = \mathbb{E}_{t \sim \mathcal{U}(0,1), b, p(\mathbf{x}_1^s), p(\mathbf{x}_0^s), p(\mathbf{x}_t^s | \mathbf{x}_0^s, \mathbf{x}_1^s)} \|w^{\text{pos}}(\mathbf{x}_t^s, t, C^{\text{rec}}|(1-b) \cdot c + b \cdot \emptyset) - (\mathbf{x}_1^s - \mathbf{x}_0^s)\|_2^2, \quad (19)$$

where $b \sim \text{Bernoulli}(p_{\text{uncond}})$ indicates the probability to use the null condition.

B EXPERIMENTAL DETAILS

B.1 TRAINING AND SAMPLING

All experiments are implemented on 4 NVIDIA A100 GPUs. Specifically, we train SurfFlow for 320K iterations and set the initial learning rate of 5e-4. A plateau scheduler is used with a factor of 0.8 and patience of 10. The minimum learning rate is 5e-6. The batch size was 32 for each distributed node. An Adam optimizer is used with a gradient clipping. A dropout ratio of 0.15 is adopted for the EGNN decoder. The weights for different loss components are set as $\lambda_{\text{pos}} = 0.2$, $\lambda_{\text{ori}} = 0.2$, $\lambda_{\text{cat}} = 1.0$, $\lambda_{\text{con}} = 1.0$, and $\lambda_{\text{str}} = 1.0$. For the encoder part, the residue embedding size is set as $\psi_{\text{rec}} = \psi_{\text{pair}} = 128$. For the decoder part, the node and edge embedding sizes are set as 128 and 64 for IPA, respectively, while as 16 and 8 for EGNN, respectively, since the number of surface points $m + m'$ are orders of magnitude larger than the number of complex’s residues $n_{\text{pep}} + n_{\text{rec}}$. We set the length of generated peptides the same as the length of their corresponding native peptides. We download the PepMerge data (Li et al., 2024) from its official repository: <https://drive.google.com/drive/folders/1bHaKDF3uCDPtfsihjZs0zmjwF6UU1uV1>.

B.2 PHYSICHEMICAL SURFACE FEATURES

Three types of surface biochemical properties are leveraged in the experiments. To be specific, free electrons and potential hydrogen bond donors (FEPH) is a categorical variable, while electrostatics and hydrophathy are continuous variables. Therefore, $\psi_{\tau} = 1$ and $\psi_{\Upsilon} = 2$. We resort to MaSIF (Gainza et al., 2020)’s scripts to acquire these surface features.

Free electrons and proton donors. The location of FEPH in the molecular surface was computed using a hydrogen bond potential as a reference. Vertices in the molecular surface whose closest atom is a polar hydrogen, a nitrogen, or an oxygen were considered potential donors or acceptors in hydrogen bonds. Then, a value from a Gaussian distribution was assigned to each vertex depending on the orientation between the heavy atoms. These initial values range from -1 (optimal position for a hydrogen bond acceptor) to +1 (optimal position for a hydrogen bond donor). Then the point is determined as an acceptor or a donor (a binary label) by whether FEPH is negative or positive.

Hydrophathy. Each vertex was assigned a hydrophathy scalar value according to the Kyte and Doolittle scale of the amino acid identity of the atom closest to the vertex. These values, in the original scale, ranged between -4.5 (hydrophilic) to +4.5 (most hydrophobic) and were then normalized to be between -1 and 1.

Poisson-Boltzmann continuum electrostatics. PDB2PQR was used to prepare protein files for electrostatic calculations and APBS (v.1.5) was used to compute Poisson-Boltzmann electrostatics for each protein. The corresponding charge at each vertex of the meshed surface was assigned using Multivalue, provided within the APBS suite. Charge values above +30 and below -30 were capped at those values and then values were normalized between -1 and 1.

C FLOW MATCHING FOR INTERNAL STRUCTURES

$\mathcal{L}_{\text{str}}(\theta)$ in Equ. 17 computes the loss for peptide’s structures $p\left(\{(a_j, O_j, \mathbf{x}_j, \chi_j)\}_{j=1}^{n_{\text{pep}}} \mid C^{\text{rec}}\right)$. Here, we provide a quick view of how to conduct full-atom CFM for residues’ positions \mathbf{x} , types a , backbone torsions O , and side-chain angles χ . To be concrete, the objective is $\phi_t(\mathbf{x}_0, \mathbf{x}_1) = t\mathbf{x}_1 + (1-t)\mathbf{x}_0$ for residues’ positions \mathbf{x} , $\exp_{O_0}(t \log_{O_0}(O_1))$ for backbone orientations O , $\phi_t(\chi_0, \chi_1) = [t\chi_1 + (1-t)\chi_0] \bmod 2\pi$ for side-chain angles χ , and $\phi_t(\tilde{a}_0, \tilde{a}_1) = t\tilde{a}_1 + (1-t)\tilde{a}_0$ for residue types a , where \tilde{a} is the representation of a using a soft one-hot encoding operation and satisfies $\text{logit}(a_i) = \tilde{a}_i \in \mathbb{R}^{20}$.

D ADDITIONAL RESULTS AND VISUALIZATION

D.1 ABLATION STUDIES

We conduct experiments to investigate the contributions of each component of our SurfFlow model. Table 4 shows that the removal of the biophysical features significantly reduces the performance with a drop of 3.76% in Designability and 4.91% in Diversity. Besides, it also indicates that inclusion of surface orientation is beneficial, which brings an improvement of 1.94% in AAR.

Table 4: Ablation studies on different components of SurfFlow.

	Geometry			Energy		Design		
	AAR % \uparrow	RMSD \AA \downarrow	SSR % \uparrow	BSR % \uparrow	Stb. % \uparrow	Aff. % \uparrow	Des. % \uparrow	Div. \uparrow
SurfFlow	54.07	1.96	85.11	87.38	22.46	22.51	73.60	0.61
w/o Position	53.26	1.99	84.79	87.15	21.30	22.38	72.09	0.60
w/o Orientation	53.04	2.00	84.60	87.04	20.79	22.46	72.36	0.60
w/o Biophysical Prop.	52.31	2.03	83.96	86.98	19.55	22.47	70.83	0.58

D.2 SURFACE POINT CLOUD EVOLUTION

We give some examples of how the surface point clouds move from time $t = 0$ to the terminal time $t = 1$. It can be seen that after time $t = 0.5$, those clouds begin to take shapes.

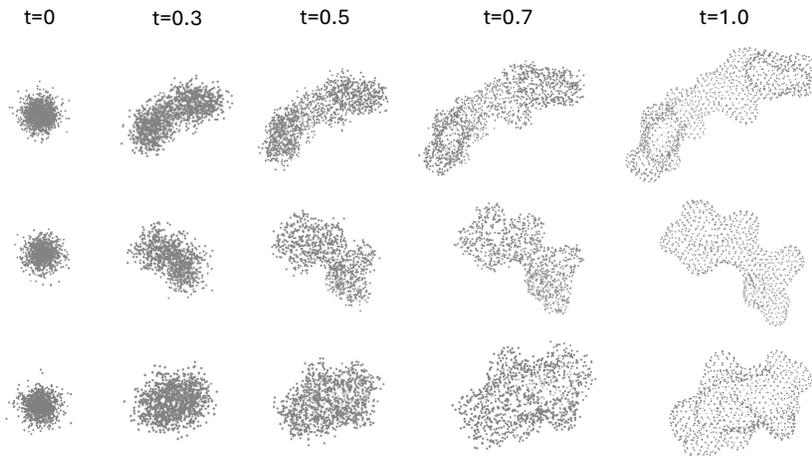


Figure 6: Visualization of the evolution of surface points’ positions over time $[0, 1]$.

972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025

E LIMITATIONS AND FUTURE WORKS

Despite the enhancement of our SurfFlow over the original full-atom design mechanism, there are still rooms for future explorations. For instance, further improvements can be expected if the surface information of the receptor’s surface information is considered and incorporated into the joint distribution modeling. Namely, our objective becomes $C_{\text{rec}} = \{a_j, O_j, \mathbf{x}_j, \chi_j\}_{j=1}^{n_{\text{rec}}} \cup \{\mathbf{x}_i^s, \mathbf{n}_i^s, \tau_i^s, \Upsilon_i^s\}_{i=1}^{m'}$, where m' is the number of receptor’s surface points. Moreover, the success of RFDiffusion implies that pretraining on regular proteins in PDB can be beneficial.