Novelty Assessment Report

Paper: La-Proteina: Atomistic Protein Generation via Partially Latent Flow Matching
PDF URL: https://openreview.net/pdf?id=RDerF20JYT
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-27

Abstract

Recently, many generative models for de novo protein structure design have emerged. Yet, only few tackle the difficult task of directly generating fully atomistic structures jointly with the underlying amino acid sequence. This is challenging, for instance, because the model must reason over side chains that change in length during generation. We introduce La-Proteina for atomistic protein design based on a novel partially latent protein representation: coarse backbone structure is modeled explicitly, while sequence and atomistic details are captured via per-residue latent variables of fixed dimensionality, thereby effectively side-stepping challenges of explicit side-chain representations. Flow matching in this partially latent space then models the joint distribution over sequences and full-atom structures. La-Proteina achieves state-of-the-art performance on multiple generation benchmarks, including all-atom co-designability, diversity, and structural validity, as confirmed through detailed structural analyses and evaluations. Notably, La-Proteina also surpasses previous models in atomistic motif scaffolding performance, unlocking critical atomistic structure-conditioned protein design tasks. Moreover, La-Proteina is able to generate co-designable proteins of up to 800 residues, a regime where most baselines collapse and fail to produce valid samples, demonstrating La-Proteina's scalability and robustness.

Core Task Landscape

This paper addresses: **Atomistic Protein Structure and Sequence Co-Design**

A total of **50 papers** were analyzed and organized into a taxonomy with **13 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Generative Model Architectures for Co-Design**
- **Functional and Context-Conditioned Protein Design**
- **Inverse Folding and Sequence Optimization**
- **Representation Learning and Pretraining for Protein Design**
- **Evaluation, Benchmarking, and Design Principles**
- **Reviews, Perspectives, and Historical Foundations**

Complete Taxonomy Tree

- Atomistic Protein Structure and Sequence Co-Design Survey Taxonomy
- Generative Model Architectures for Co-Design
  - Diffusion-Based Co-Design Models (4 papers)
  - [5] Multistate and functional protein design using RoseTTAFold sequence space diffusion (Jacob Merle Gershon, 2024) View paper
  - [15] Joint protein sequence-structure co-design via Equivariant diffusion (R Vinod, 2022) View paper
  - [21] Pre-Training Protein Encoder via Siamese Sequence-Structure Diffusion Trajectory Prediction (Zhang, 2023) View paper
  - [26] Multimodal diffusion for joint design of protein sequence and structure. (Shaowen Zhu, 2025) View paper
  - Flow Matching and Continuous Normalizing Flow Co-Design ★ (3 papers)
  - [0] La-Proteina: Atomistic Protein Generation via Partially Latent Flow Matching (Anon et al., 2026) View paper
  - [3] Co-design protein sequence and structure in discrete space via generative flow (Sen Yang, 2025) View paper
  - [40] ProteinZen: Combining Latent and SE (3) Flow Matching for All-Atom Protein Generation (AJ Li, 2024) View paper
  - Autoregressive and Language Model-Based Co-Design (4 papers)
  - [1] Atomic context-conditioned protein sequence design using LigandMPNN (Justas Dauparas, 2025) View paper
  - [19] A high-level programming language for generative protein design (Brian Hie, 2022) View paper
  - [44] Towards Protein Sequence & Structure Co-Design with Multi-Modal Language Models (SZ Lu, 2025) View paper
  - [45] RamaNet: Computational de novo helical protein backbone design using a long short-term memory generative adversarial neural network (Sari Sabban, 2019) View paper
- Functional and Context-Conditioned Protein Design
  - Ligand-Binding and Pocket Design (4 papers)
  - [2] Full-atom protein pocket design via iterative refinement (Zhang Zai-xi, 2023) View paper
  - [4] A defined structural unit enables de novo design of small-moleculeâbinding proteins (Nicholas F. Polizzi, 2020) View paper
  - [14] Context-aware geometric deep learning for protein sequence design (Lucien F. Krapp, 2023) View paper
  - [30] Protein-ligand co-design: a case for improving binding affinity between Type II NADH: quinone oxidoreductase and quinones (Vladimir Porokhin, 2025) View paper
  - Antibody and Antigen-Specific Design (6 papers)
  - [7] Fast and Accurate Antibody Sequence Design via Structure Retrieval (Zhang Xing-yi, 2025) View paper
  - [9] Antigen-Specific Antibody Design and Optimization with Diffusion-Based Generative Models for Protein Structures (Shitong Luo, 2022) View paper

- [16] GeoAB: towards realistic antibody design and reliable affinity maturation (Yufei Huang, 2024) View paper
- [18] Repurposing alphafold3-like protein folding models for antibody sequence and structure co-design (N Yang, 2025) View paper
- [20] Incorporating pre-training paradigm for antibody sequence-structure co-design (GAO Kaiyuan, 2022) View paper
- [24] Guiding diffusion models for antibody sequence and structure co-design with developability properties (Amelia Villegas-Morcillo, 2024) View paper
- Motif-Conditioned and Functional Geometry-Guided Design (3 papers)
- [6] Joint design of protein sequence and structure based on motifs (Song, 2023) View paper
- [12] Functional Geometry Guided Protein Sequence and Backbone Structure Co-Design (Song, 2023) View paper
- [31] ApexGen: Simultaneous design of peptide binder sequence and structure for target proteins (Xiaoqiong Xia, 2025) View paper
- Protein-Protein Interface Co-Design (3 papers)
- [10] Dynamics-inspired Structure Hallucination for Protein-protein Interaction Modeling (F Wu, 2025) View paper
- [46] Predicting the tolerated sequences for proteins and protein interfaces using RosettaBackrub flexible backbone design (Colin A. Smith, 2011) View paper
- [47] Prediction of protein-protein interface sequence diversity using flexible backbone computational protein design (Elisabeth L. Humphris, 2008) View paper

- Inverse Folding and Sequence Optimization
  - Fixed Backbone Inverse Folding (2 papers)
  - [11] DS-ProGen: A Dual-Structure Deep Language Model for Functional Protein Design (Li Yanting, 2025) View paper
  - [48] All-Atom Protein Sequence Design Based on Geometric Deep Learning. (Jiale Liu, 2024) View paper
  - Flexible Backbone Sequence Design (6 papers)
  - [22] Comparison of Rosetta flexibleâ□□backbone computational protein design methods on binding interactions (Amanda L. Loshbaugh, 2020) View paper
  - [27] RosettaBackrubâ□□a web server for flexible backbone protein structure modeling and design (F. Lauck, 2010) View paper
  - [29] Coupling backbone flexibility and amino acid sequence selection in protein design (Alyce Su, 1997) View paper
  - [34] Coupling protein side-chain and backbone flexibility improves the re-design of protein-ligand specificity (Noah Ollikainen, 2015) View paper
  - [37] Recapitulation of protein family divergence using flexible backbone protein design (Christopher T. Saunders, 2005) View paper
  - [43] A structural homology approach for computational protein design with flexible backbone (David Simoncini, 2018) View paper
  - Sequence Sampling and Optimization Strategies (4 papers)
  - [38] Automatic protein design with all atom force-fields by exact and heuristic optimization (L Wernisch, 2000) View paper
  - [41] Design of proteins by parallel tempering in the sequence space (Spiwok, 2025) View paper
  - [42] Relaxed Sequence Sampling for Diverse Protein Design (Ban, 2025) View paper
  - [50] Multi-objective optimization for designing structurally similar proteins with diverse sequences (Ryo Akiba, 2025) View paper
- Representation Learning and Pretraining for Protein Design (3 papers)
  - [8] Learning Complete Protein Representation by Dynamically Coupling of Sequence and Structure (Bozhen Hu, 2024) View paper
  - [23] Implicit modeling of the conformational landscape and sequence allows scoring and generation of stable proteins (Yehlin Cho, 2024) View paper
  - [39] FlexRibbon: Joint Sequence and Structure Pretraining for Protein Modeling (Jianwei Zhu, 2025) View paper
- Evaluation, Benchmarking, and Design Principles
  - Design Principles and Structural Constraints (4 papers)
  - [17] BC-Design: A Biochemistry-Aware Framework for Inverse Protein Design (Xiangru Tang, 2025) View paper
  - [33] High-resolution protein design with backbone freedom (Pehr B. Harbury, 1998) View paper
  - [35] Investigating the impacts of sidechains on de-novo protein design (Cooper Svajda, 2025) View paper
  - [36] DivPro: diverse protein sequence design with direct structure recovery guidance (Xinyi Zhou, 2025) View paper
- Reviews, Perspectives, and Historical Foundations (5 papers)
  - [13] Advances in protein structure prediction and design (Brian Kuhlman, 2019) View paper
  - [25] De novo protein design: fully automated sequence selection (Bassil I. Dahiyat, 1997) View paper
  - [28] The role of ai-driven de novo protein design in the exploration of the protein functional universe (Guohao Zhang, 2025) View paper
  - [32] Protein Manufacture: Protein Design Assisted by Machine Learning from Backbone to Sequence (Man Xu, 2024) View paper
  - [49] A Model-Centric Review of Deep Learning for Protein Design (Kyro, 2025) View paper

Narrative

Core task: atomistic protein structure and sequence co-design. The field has evolved into a rich landscape organized around several complementary themes. Generative Model Architectures for Co-Design explores modern neural frameworks—ranging from diffusion models and flow matching to equivariant networks—that jointly generate backbone geometry and amino acid sequences. Functional and Context-Conditioned Protein Design addresses the challenge of designing proteins that bind specific ligands, recognize antigens, or satisfy functional constraints, often requiring careful modeling of binding pockets and interaction interfaces. Inverse Folding and Sequence Optimization focuses on methods that take a fixed or flexible backbone and optimize sequences to stabilize the fold, including classical energy-based approaches and newer learning-based techniques. Representation Learning and Pretraining for Protein Design examines how large-scale pretraining on structural databases can inform downstream design tasks, while Evaluation, Benchmarking, and Design Principles provides the experimental and computational standards needed to validate designed proteins. Finally, Reviews, Perspectives, and Historical Foundations situate recent advances within the broader trajectory of computational protein engineering.

Within Generative Model Architectures, a particularly active line of work centers on flow-based methods that model continuous transformations of structure and sequence. LaProteina[0] exemplifies this direction by employing flow matching techniques for co-design, sitting alongside other flow-based approaches such as Discrete Generative Flow[3] and ProteinZen[40], which explore discrete and continuous normalizing flows respectively. These methods contrast with diffusion-based frameworks like RoseTTAFold Diffusion[5] and Equivariant Diffusion Codesign[15], which rely on iterative denoising rather than deterministic flow trajectories. A key trade-off across these branches involves balancing the expressiveness of generative architectures with computational efficiency and the ability to incorporate physical constraints. LaProteina[0] contributes to this conversation by demonstrating how flow matching can achieve competitive design quality while offering more direct control over the generative process, positioning it as a methodologically distinct yet complementary alternative to diffusion and energy-based paradigms.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. Co-design protein sequence and structure in discrete space via generative flow

Authors: Sen Yang, Lingli Ju, Peng Cheng, Jianglin Zhou, Yamin Cai, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Abstract Motivation Generative models have demonstrated considerable promise in de novo protein design. Traditional approaches typically focus on either sequence or structure in isolation, limiting the capacity to explore the intricate sequenceâ□□structure landscape and achieve optimal designs. However, joint protein sequence and structure co-design remains a largely underexplored challenge. Results We present CoFlow, a discrete model for protein co-design from scratch or given constraints. CoFl...

Relationship Analysis

Both papers belong to the Flow Matching and Continuous Normalizing Flow Co-Design category, employing flow-based generative models for joint protein sequence-structure generation. They overlap in addressing atomistic co-design through flow matching frameworks, but differ fundamentally in their approach: La-Proteina uses a partially latent continuous representation with separate flows for backbone and per-residue latent variables in continuous space, while CoFlow operates entirely in discrete space using discrete flows integrated with a multi-modal masked language model for tokenized sequence and structure representations.

---

2. ProteinZen: Combining Latent and SE (3) Flow Matching for All-Atom Protein Generation

Authors: AJ Li, T Kortemme | **Year/Venue**: 2024 | **URL**: View paper

Abstract

â□¦ and continuous aspects of protein structure: amino acid â□¦ The second is a purely atomistic representation ({a(i,j)}, {h(i,j)}â□¦ Overall, we find that metrics under the co-design evaluation â□¦

Relationship Analysis

Both papers belong to the Flow Matching and Continuous Normalizing Flow Co-Design category, employing flow matching techniques for joint sequence-structure generation in continuous spaces. They overlap in using partially latent representations where backbone coordinates are modeled explicitly while sequence and side-chain details are captured in latent variables, and both train VAEs followed by flow matching models. The key difference is that ProteinZen uses SE(3) flow matching on backbone frames (rotation and translation separately) combined with Euclidean flow matching on latents, while La-Proteina uses a unified flow matching approach on α-carbon coordinates and latents with separate interpolation times, and La-Proteina demonstrates superior scalability (up to 800 residues) and performance on motif scaffolding tasks.

Contributions Analysis

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: La-Proteina: Partially Latent Flow Matching Framework for Atomistic Protein Design

Description: The authors propose La-Proteina, a generative model that uses a partially latent representation where the α-carbon backbone is modeled explicitly and sequence plus side-chain details are encoded in fixed-size per-residue latent variables. Flow matching in this hybrid space jointly models the distribution over sequences and full-atom structures.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Branching Flows: Discrete, Continuous, and Manifold Flow Matching with Splits and Deletions

URL: View paper

Brief Assessment

Branching Flows[54] focuses on variable-length sequence generation through branching/deletion mechanisms across discrete and continuous spaces, not on the specific partially latent representation design (explicit α-carbon backbone with fixed-size per-residue latents for side-chains) that defines La-Proteina's contribution.

---

2. A Variational Perspective on Generative Protein Fitness Optimization

URL: View paper

Brief Assessment

Variational Fitness Optimization[55] focuses on protein fitness optimization through variational inference in latent space for discovering high-fitness variants, not on joint generation of protein sequences and full-atom structures using partially latent flow matching.

---

3. Co-design protein sequence and structure in discrete space via generative flow

URL: View paper

Brief Assessment

Discrete Generative Flow[3] operates entirely in discrete space for sequence-structure co-design, whereas La-Proteina uses a partially latent continuous representation with explicit α-carbon modeling and flow matching in hybrid space.

---

4. All-atom protein design via SE (3) flow matching with ProteinZen

URL: View paper

Brief Assessment

ProteinZen Flow Matching[57] uses SE(3) flow matching on oriented rigid bodies for all-atom generation, not a partially latent representation. The candidate decomposes residues into rigid bodies in continuous space, while the original encodes sequence and side-chains into fixed-size latent variables per residue.

---

5. ProteinZen: Combining Latent and SE (3) Flow Matching for All-Atom Protein Generation

URL: View paper

Brief Assessment

ProteinZen[40] uses a different architectural approach with SE(3) flow matching on backbone frames combined with latent representations, whereas La-Proteina explicitly models α-carbon coordinates with per-residue latent variables for side chains and sequences. The technical implementations and design philosophies differ substantially.

---

6. ProtFlow: Fast Protein Sequence Design via Flow Matching on Compressed Protein Language Model Embeddings

URL: View paper

Brief Assessment

ProtFlow[52] operates on compressed embeddings from protein language models for sequence design only, while the original paper jointly generates full-atom structures and sequences using a partially latent representation where α-carbon backbones are explicit and side-chain details are in latent space. These are fundamentally different architectural approaches for different design tasks.

---

7. Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design

URL: View paper

Brief Assessment

Discrete Flow Codesign[51] focuses on discrete sequence generation using continuous time Markov chains for flow matching, while the original paper proposes a partially latent representation where backbone is explicit and side-chains are encoded in latent variables. These are fundamentally different architectural approaches to protein design.

---

8. Sequence-Augmented SE(3)-Flow Matching For Conditional Protein Backbone Generation

URL: View paper

Brief Assessment

Conditional Backbone Generation[58] focuses on sequence-conditioned backbone generation using SE(3)-flow matching, while La-Proteina addresses full atomistic protein design (backbone + side chains + sequence) using a partially latent representation. These are distinct technical approaches to different protein design problems.

---

9. Consistent Synthetic Sequences Unlock Structural Diversity in Fully Atomistic De Novo Protein Design

URL: View paper

Brief Assessment

Consistent Synthetic Sequences[56] focuses on creating high-quality training datasets by aligning synthetic sequences with structures, then retraining existing models (including La-Proteina). It does not claim to introduce the partially latent flow matching framework itself.

---

10. Sequence-augmented SE (3)-flow matching for conditional protein generation

URL: View paper

Brief Assessment

Sequence Augmented Flow[53] focuses on sequence-conditioned backbone generation using SE(3)-flow matching, while La-Proteina addresses full-atom structure generation with a partially latent representation. The candidate does not challenge the novelty of La-Proteina's approach to atomistic protein design with explicit α-carbon modeling and latent side-chain encoding.

Contribution 2: State-of-the-Art Performance on Unconditional Atomistic Protein Generation

Description: The authors demonstrate that La-Proteina achieves state-of-the-art results on unconditional atomistic protein generation benchmarks, outperforming existing methods in all-atom co-designability, diversity, and structural validity metrics.

This contribution was assessed against **6 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Geometric-aware models for protein design

URL: View paper

Brief Assessment

Geometric Aware Models[62] is a thesis focused on joint representation learning and protein co-design methods, not on unconditional atomistic protein generation benchmarks. The candidate does not address the specific metrics (all-atom co-designability, diversity, structural validity) evaluated in the original paper's unconditional generation task.

---

2. Full-Atom Peptide Design with Geometric Latent Diffusion

URL: View paper

Brief Assessment

Peptide Latent Diffusion[60] focuses on target-specific peptide design with binding site conditioning, not unconditional atomistic protein generation. The candidate addresses sequence-structure co-design for peptides given binding sites, which is a fundamentally different task from unconditional generation benchmarks.

---

3. Joint Design of Protein Surface and Backbone Using a Diffusion Bridge Model

URL: View paper

Brief Assessment

Diffusion Bridge Model[59] focuses on conditional protein-peptide design with receptor surface guidance, not unconditional atomistic protein generation benchmarks measuring co-designability, diversity, and structural validity.

---

4. Towards Protein Sequence & Structure Co-Design with Multi-Modal Language Models

URL: View paper

Brief Assessment

Multimodal Language Models[44] focuses on sequence-structure co-design using masked language models (ESM3) with sampling strategies, not on unconditional atomistic protein generation with flow matching architectures that achieve state-of-the-art co-designability metrics.

---

5. Generating Novel, Designable, and Diverse Protein Structures by Equivariantly Diffusing Oriented Residue Clouds

URL: View paper

Brief Assessment

Oriented Residue Clouds[61] focuses on backbone generation using Cα coordinates with diffusion models, not full atomistic generation including side chains and sequences. The candidate does not address all-atom co-designability metrics that are central to the original paper's contribution.

---

6. Consistent Synthetic Sequences Unlock Structural Diversity in Fully Atomistic De Novo Protein Design

URL: View paper

Prior Art Analysis

Consistent Synthetic Sequences[56] demonstrates that by retraining La-Proteina on their improved dataset, they achieve superior performance metrics compared to the original La-Proteina results. Specifically, they report improvements of +54% in structural diversity

and +27% in co-designability over the baseline La-Proteina model, indicating that the original state-of-the-art claims were surpassed by subsequent work using better training data.

Evidence

Evidence 1 - **Rationale**: This pair shows that Consistent Synthetic Sequences[56] achieved higher performance metrics than the original La-Proteina by retraining on improved data, demonstrating that the original state-of-the-art claims were subsequently exceeded. - **Original**: la-proteina achieves state-of-the-art performance on key generation benchmarks, including all-atom co-designability, diversity, and structural validity. - **Candidate**: By retraining la-proteina, which models discrete residue type and side chain structure in a continuous latent space, on this dataset, we achieve new state-of-the-art results, with improvements of +54% in structural diversity and +27% in co-designability.

Evidence 2 - **Rationale**: This evidence pair demonstrates that the candidate paper identifies limitations in training data quality that affected prior models' performance, and by addressing these limitations, they surpassed the original state-of-the-art results claimed by La-Proteina. - **Original**: main contributions. (i) we propose la-proteina, a partially latent flow matching framework designed for the joint generation of protein sequence and fully atomistic structure. (ii) la-proteina achieves state-of-the-art performance in unconditional protein generation. - **Candidate**: high-quality training datasets are crucial for the development of effective protein design models, but existing synthetic datasets often include unfavorable sequence-structure pairs, impairing generative model performance. we leverage proteinmpnn, whose sequences are experimentally favorable as well...

Contribution 3: Atomistic Motif Scaffolding for Indexed and Unindexed Tasks

Description: The authors successfully apply La-Proteina to atomistic motif scaffolding tasks, including both indexed (where motif residue positions are specified) and unindexed (where positions are unknown) setups, as well as all-atom and tip-atom scaffolding variants, outperforming existing baselines.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Atomic details of biomineralization proteins inspiring protein design and reengineering for functional biominerals

URL: View paper

Brief Assessment

Biomineralization Proteins[68] focuses on structural studies of biomineralization proteins and their interactions with minerals, not on computational protein design or motif scaffolding tasks.

---

2. Illuminating the universe of enzyme catalysis in the era of artificial intelligence

URL: View paper

Brief Assessment

Enzyme Catalysis AI[67] discusses motif scaffolding in the context of enzyme design but does not present a computational method or framework for atomistic motif scaffolding tasks. The candidate focuses on enzyme catalysis understanding rather than developing generative models for protein design.

---

3. Structure-guided utilization of lignocellulose for catalysis, energy, and biomaterials

URL: View paper

Brief Assessment

Lignocellulose Utilization[65] focuses on biomaterial processing and catalysis applications, not protein design or atomistic motif scaffolding tasks.

---

4. Computational enzyme design by catalytic motif scaffolding

URL: View paper

Brief Assessment

Catalytic Motif Scaffolding[69] focuses on enzyme design by scaffolding catalytic arrays (active sites) in de novo proteins, not on general atomistic motif scaffolding tasks with indexed/unindexed setups as evaluated in the original paper.

---

5. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease

URL: View paper

Brief Assessment

Cas9 PAM Recognition[70] focuses on structural biology of CRISPR-Cas9 protein-DNA interactions, not computational protein design or motif scaffolding methods.

---

6. Physics-based approach to extend a de novo TIM barrel with rationally designed helix-loop-helix motifs.

URL: View paper

Brief Assessment

TIM Barrel Extension[71] focuses on extending a specific de novo TIM barrel protein using physics-based Rosetta design for helix-loop-helix motifs, not on general atomistic motif scaffolding tasks or structure-conditioned protein design frameworks.

---

7. Molecule generation for target protein binding with hierarchical consistency diffusion model

URL: View paper

Brief Assessment

Target Binding Generation[64] focuses on molecule generation for target protein binding using hierarchical diffusion models, not on atomistic motif scaffolding for protein structure design tasks.

---

8. Protein engineering for improving and diversifying natural product biosynthesis

URL: View paper

Brief Assessment

Natural Product Engineering[66] focuses on protein engineering for natural product biosynthesis, not atomistic motif scaffolding for structure-conditioned protein design tasks.

---

9. De novo protein designâ□□From new structures to programmable functions

URL: View paper

Brief Assessment

Programmable Functions Design[63] discusses motif scaffolding conceptually but does not provide technical details about atomistic methods, indexed/unindexed setups, or comparative performance that would refute the original paper's novelty claims.

---

10. Identification of (S)-selective transaminases for the asymmetric synthesis of bulky chiral amines

URL: View paper

Brief Assessment

Transaminase Identification[72] focuses on enzyme engineering and transaminase development for asymmetric synthesis of bulky chiral amines, not on atomistic motif scaffolding for protein structure generation or design tasks.

Appendix: Text Similarity Detection

Textual similarity detection checked 25 papers and found 2 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design

Detected in: Contribution: contribution_1

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] La-Proteina: Atomistic Protein Generation via Partially Latent Flow Matching View paper
- [1] Atomic context-conditioned protein sequence design using LigandMPNN View paper
- [2] Full-atom protein pocket design via iterative refinement View paper
- [3] Co-design protein sequence and structure in discrete space via generative flow View paper
- [4] A defined structural unit enables de novo design of small-moleculeâ□□binding proteins View paper
- [5] Multistate and functional protein design using RoseTTAFold sequence space diffusion View paper
- [6] Joint design of protein sequence and structure based on motifs View paper
- [7] Fast and Accurate Antibody Sequence Design via Structure Retrieval View paper
- [8] Learning Complete Protein Representation by Dynamically Coupling of Sequence and Structure View paper
- [9] Antigen-Specific Antibody Design and Optimization with Diffusion-Based Generative Models for Protein Structures View paper
- [10] Dynamics-inspired Structure Hallucination for Protein-protein Interaction Modeling View paper
- [11] DS-ProGen: A Dual-Structure Deep Language Model for Functional Protein Design View paper
- [12] Functional Geometry Guided Protein Sequence and Backbone Structure Co-Design View paper
- [13] Advances in protein structure prediction and design View paper
- [14] Context-aware geometric deep learning for protein sequence design View paper
- [15] Joint protein sequence-structure co-design via Equivariant diffusion View paper
- [16] GeoAB: towards realistic antibody design and reliable affinity maturation View paper
- [17] BC-Design: A Biochemistry-Aware Framework for Inverse Protein Design View paper
- [18] Repurposing alphafold3-like protein folding models for antibody sequence and structure co-design View paper
- [19] A high-level programming language for generative protein design View paper
- [20] Incorporating pre-training paradigm for antibody sequence-structure co-design View paper
- [21] Pre-Training Protein Encoder via Siamese Sequence-Structure Diffusion Trajectory Prediction View paper
- [22] Comparison of Rosetta flexibleâ□□backbone computational protein design methods on binding interactions View paper
- [23] Implicit modeling of the conformational landscape and sequence allows scoring and generation of stable proteins View paper
- [24] Guiding diffusion models for antibody sequence and structure co-design with developability properties View paper
- [25] De novo protein design: fully automated sequence selection View paper
- [26] Multimodal diffusion for joint design of protein sequence and structure. View paper
- [27] RosettaBackrubâ□□a web server for flexible backbone protein structure modeling and design View paper
- [28] The role of ai-driven de novo protein design in the exploration of the protein functional universe View paper
- [29] Coupling backbone flexibility and amino acid sequence selection in protein design View paper
- [30] Protein-ligand co-design: a case for improving binding affinity between Type II NADH: quinone oxidoreductase and quinones View paper
- [31] ApexGen: Simultaneous design of peptide binder sequence and structure for target proteins View paper
- [32] Protein Manufacture: Protein Design Assisted by Machine Learning from Backbone to Sequence View paper
- [33] High-resolution protein design with backbone freedom View paper
- [34] Coupling protein side-chain and backbone flexibility improves the re-design of protein-ligand specificity View paper
- [35] Investigating the impacts of sidechains on de-novo protein design View paper
- [36] DivPro: diverse protein sequence design with direct structure recovery guidance View paper
- [37] Recapitulation of protein family divergence using flexible backbone protein design View paper
- [38] Automatic protein design with all atom force-fields by exact and heuristic optimization View paper
- [39] FlexRibbon: Joint Sequence and Structure Pretraining for Protein Modeling View paper
- [40] ProteinZen: Combining Latent and SE (3) Flow Matching for All-Atom Protein Generation View paper
- [41] Design of proteins by parallel tempering in the sequence space View paper
- [42] Relaxed Sequence Sampling for Diverse Protein Design View paper
- [43] A structural homology approach for computational protein design with flexible backbone View paper
- [44] Towards Protein Sequence & Structure Co-Design with Multi-Modal Language Models View paper
- [45] RamaNet: Computational de novo helical protein backbone design using a long short-term memory generative adversarial neural network View paper
- [46] Predicting the tolerated sequences for proteins and protein interfaces using RosettaBackrub flexible backbone design View paper
- [47] Prediction of protein-protein interface sequence diversity using flexible backbone computational protein design View paper
- [48] All-Atom Protein Sequence Design Based on Geometric Deep Learning. View paper

- [49] A Model-Centric Review of Deep Learning for Protein Design View paper
- [50] Multi-objective optimization for designing structurally similar proteins with diverse sequences View paper
- [51] Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design View paper
- [52] ProtFlow: Fast Protein Sequence Design via Flow Matching on Compressed Protein Language Model Embeddings View paper
- [53] Sequence-augmented SE (3)-flow matching for conditional protein generation View paper
- [54] Branching Flows: Discrete, Continuous, and Manifold Flow Matching with Splits and Deletions View paper
- [55] A Variational Perspective on Generative Protein Fitness Optimization View paper
- [56] Consistent Synthetic Sequences Unlock Structural Diversity in Fully Atomistic De Novo Protein Design View paper
- [57] All-atom protein design via SE (3) flow matching with ProteinZen View paper
- [58] Sequence-Augmented SE(3)-Flow Matching For Conditional Protein Backbone Generation View paper
- [59] Joint Design of Protein Surface and Backbone Using a Diffusion Bridge Model View paper
- [60] Full-Atom Peptide Design with Geometric Latent Diffusion View paper
- [61] Generating Novel, Designable, and Diverse Protein Structures by Equivariantly Diffusing Oriented Residue Clouds View paper
- [62] Geometric-aware models for protein design View paper
- [63] De novo protein designâ□□From new structures to programmable functions View paper
- [64] Molecule generation for target protein binding with hierarchical consistency diffusion model View paper
- [65] Structure-guided utilization of lignocellulose for catalysis, energy, and biomaterials View paper
- [66] Protein engineering for improving and diversifying natural product biosynthesis View paper
- [67] Illuminating the universe of enzyme catalysis in the era of artificial intelligence View paper
- [68] Atomic details of biomineralization proteins inspiring protein design and reengineering for functional biominerals View paper
- [69] Computational enzyme design by catalytic motif scaffolding View paper
- [70] Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease View paper
- [71] Physics-based approach to extend a de novo TIM barrel with rationally designed helix-loop-helix motifs. View paper
- [72] Identification of (S)-selective transaminases for the asymmetric synthesis of bulky chiral amines View paper