Novelty Assessment Report

Paper: Towards All-Atom Foundation Models for Biomolecular Binding Affinity Prediction
PDF URL: https://openreview.net/pdf?id=o0Qfsq1fK8
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-01

Abstract

Biomolecular interactions play a critical role in biological processes. While recent breakthroughs like AlphaFold 3 have enabled accurate modeling of biomolecular complex structures, predicting binding affinity remains challenging mainly due to limited high-quality data. Recent methods are often specialized for specific types of biomolecular interactions, limiting their generalizability. In this work, we repurpose AlphaFold 3 for representation learning to predict binding affinity, a non-trivial task that requires shifting from generative structure prediction to encoding observed geometry, simplifying the heavily conditioned trunk module, and designing a framework to jointly capture sequence and structural information. To address these challenges, we introduce the **Atom-level Diffusion Transformer (ADiT)**, which takes sequence and structure as inputs, employs a unified tokenization scheme, integrates diffusion transformers, and removes dependencies on multiple sequence alignments and templates. We pre-train three ADiT variants on the PDB dataset with a denoising objective and evaluate them across protein-ligand, drug-target, protein-protein, and antibody-antigen interactions. The model achieves state-of-the-art or competitive performance across benchmarks, scales effectively with model size, and successfully identifies wet-lab validated affinity-enhancing antibody mutations, establishing a generalizable framework for biomolecular interactions. We plan to release the code upon acceptance.

Core Task Landscape

This paper addresses: **Biomolecular Binding Affinity Prediction**

A total of **50 papers** were analyzed and organized into a taxonomy with **20 categories**.

Taxonomy Overview

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

- **Deep Learning Architectures for Structure-Based Prediction**
- **Sequence-Based and Hybrid Feature Methods**
- **Specialized Learning Strategies and Objectives**
- **Mathematical and Topological Descriptor Methods**
- **Classical and Physics-Based Computational Methods**
- **Specialized Biomolecular Interaction Contexts**
- **Machine Learning with Classical Descriptors**
- **Reviews, Surveys, and Methodological Assessments**
- **Experimental and Biophysical Studies**
- **Uncategorized or Insufficient Information**

Complete Taxonomy Tree

- Biomolecular Binding Affinity Prediction Survey Taxonomy
- Deep Learning Architectures for Structure-Based Prediction
  - Convolutional Neural Network Approaches (4 papers)
  - [7] Development and evaluation of a deep learning model for proteinâ□□ligand binding affinity prediction (Stepniewska-Dziubinska, 2018) View paper
  - [14] DeepAtom: a framework for protein-ligand binding affinity prediction (Yanjun Li, 2019) View paper
  - [21] Deep learning in drug design: protein-ligand binding affinity prediction (M. Rezaei, 2020) View paper
  - [28] Onionnet: a multiple-layer intermolecular-contact-based convolutional neural network for proteinâ□□ligand binding affinity prediction (ZHENG Liangzhen, 2019) View paper
  - Graph Neural Network Methods
  - Interaction-Aware Graph Networks (4 papers)
    - [10] Structure-aware Interactive Graph Neural Networks for the Prediction of Protein-Ligand Binding Affinity (Li Shuangli, 2021) View paper
    - [30] Proteinâ□□ligand binding affinity prediction with edge awareness and supervised attention (Yuliang Gu, 2023) View paper
    - [44] Interaction-Based Inductive Bias in Graph Neural Networks: Enhancing Protein-Ligand Binding Affinity Predictions From 3D Structures (Ziduo Yang, 2024) View paper
    - [49] Structure-Aware Graph Attention Diffusion Network for Proteinâ□□Ligand Binding Affinity Prediction (Mei Li, 2023) View paper
  - Multi-Scale and Hierarchical Graph Architectures (2 papers)
    - [9] MGraphDTA: deep multiscale graph neural network for explainable drugâ□□target binding affinity prediction (Ziduo Yang, 2022) View paper
    - [36] NHGNN-DTA: a node-adaptive hybrid graph neural network for interpretable drugâ□□target binding affinity prediction (Haohuai He, 2023) View paper

    - Contrastive and Self-Supervised Graph Learning (2 papers)
        - [23] Contrastive pre-training and 3D convolution neural network for RNA and small molecule binding affinity prediction (Saisai Sun, 2024) View paper
        - [41] GraphCL-DTA: A Graph Contrastive Learning With Molecular Semantics for Drug-Target Binding Affinity Prediction (Xinxing Yang, 2023) View paper
    - Transformer and Attention-Based Architectures (2 papers)
    - [29] CAPLA: improved prediction of proteinâ□ligand binding affinity by a deep learning approach based on a cross-attention mechanism (Zhi Jin, 2023) View paper
    - [33] Distance plus attention for binding affinity prediction (Julia Rahman, 2024) View paper
    - Diffusion Transformer and Foundation Models ★ (2 papers)
    - [0] Towards All-Atom Foundation Models for Biomolecular Binding Affinity Prediction (Anon et al., 2026) View paper
    - [1] Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction (Saro Passaro, 2025) View paper
- Sequence-Based and Hybrid Feature Methods
    - Sequence-Only Neural Network Models (2 papers)
    - [3] MHCflurry: open-source class I MHC binding affinity prediction (TJ O'Donnell, 2018) View paper
    - [4] DeepDTAF: a deep learning method to predict protein-ligand binding affinity (Kaili Wang, 2021) View paper
    - Hybrid Feature Integration (3 papers)
    - [11] Improved proteinâ□ligand binding affinity prediction with structure-based deep fusion inference (Jones, 2021) View paper
    - [40] Learning from the ligand: using ligand-based features to improve binding affinity prediction (Charlotte M. Deane, 2020) View paper
    - [43] DataDTA: a multi-feature and dual-interaction aggregation framework for drugâ□target binding affinity prediction (Yan Zhu, 2023) View paper
- Specialized Learning Strategies and Objectives (4 papers)
    - [13] Multi-task bioassay pre-training for protein-ligand binding affinity prediction (Jiaxian Yan, 2023) View paper
    - [18] Drug-target binding affinity prediction model based on multi-scale diffusion and interactive learning (Zhiqin Zhu, 2024) View paper
    - [27] PLANET: A Multi-Objective Graph Neural Network Model for Proteinâ□Ligand Binding Affinity Prediction (Xiangying Zhang, 2023) View paper
    - [37] Harnessing pre-trained models for accurate prediction of protein-ligand binding affinity. (Jiashan Li, 2025) View paper
- Mathematical and Topological Descriptor Methods (2 papers)
    - [17] Persistent spectralâ□based machine learning (PerSpect ML) for protein-ligand binding affinity prediction (Zhenyu Meng, 2021) View paper
    - [50] Persistent Path-Spectral (PPS) Based Machine Learning for Protein-Ligand Binding Affinity Prediction (Ran Liu, 2023) View paper
- Classical and Physics-Based Computational Methods
    - Free Energy and Molecular Dynamics Simulations (2 papers)
    - [2] Calculation of protein-ligand binding affinities (Michael K Gilson, 2007) View paper
    - [47] Binding energy distribution analysis method (BEDAM) for estimation of Proteinâ□ Ligand binding affinities (Emilio Gallicchio, 2010) View paper
    - Empirical Scoring Functions and Docking (2 papers)
    - [32] Assessment of programs for ligand binding affinity prediction (Ryangguk Kim, 2008) View paper
    - [38] Computational methods for calculation of ligand-binding affinity (Walter Filgueira de Azevedo, 2008) View paper
    - Force Field Evaluation and Parameterization (1 papers)
    - [25] Current State of Open Source Force Fields in Proteinâ□Ligand Binding Affinity Predictions (David F. Hahn, 2024) View paper
- Specialized Biomolecular Interaction Contexts
    - Protein-Protein Interaction Affinity (3 papers)
    - [45] Computational prediction of proteinâ□protein binding affinities (Till Siebenmorgen, 2020) View paper
    - [46] Proteinâ□protein binding affinity prediction from amino acid sequence (Kumar Yugandhar, 2014) View paper
    - [48] Systematic investigation of machine learning on limited data: A study on predicting protein-protein binding strength (Feifan Zheng, 2024) View paper
    - RNA-Small Molecule Interaction Affinity (1 papers)
    - [24] Reliable method for predicting the binding affinity of RNA-small molecule interactions using machine learning (Sowmya R Krishnan, 2024) View paper
    - Drug Resistance and Mutation Effects (1 papers)
    - [19] Computational studies of proteinâ□drug binding affinity changes upon mutations in the drug target (Ran Friedman, 2022) View paper
- Machine Learning with Classical Descriptors (2 papers)
    - [39] Hybrid Quantum Neural Network Approaches to Proteinâ□Ligand Binding Affinity Prediction (Maria Avramouli, 2024) View paper
    - [42] Pred-binding: large-scale proteinâ□ligand binding affinity prediction (Tao Weiyang, 2016) View paper
- Reviews, Surveys, and Methodological Assessments (7 papers)
    - [5] Biomolecular interaction prediction: the era of AI (Haoping Wang, 2025) View paper
    - [8] Structure-based, deep-learning models for protein-ligand binding affinity prediction (Debby D. Wang, 2024) View paper
    - [15] Recent advances in computational and experimental protein-ligand affinity determination techniques (Visvaldas Kairys, 2024) View paper
    - [16] Prediction of proteinâ□ligand binding affinity via deep learning models (Huiwen Wang, 2024) View paper
    - [26] Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: a review (Rocco Meli, 2022) View paper
    - [31] Modern machineâ□learning for binding affinity estimation of proteinâ□ligand complexes: Progress, opportunities, and challenges (Tobias Harren, 2024) View paper
    - [35] Recent advances in machine learning predictions of protein-ligand binding affinities (Jian Jiang, 2026) View paper
- Experimental and Biophysical Studies (3 papers)
    - [6] Computational analysis of protein-ligand interaction by targeting a cell cycle restrainer (Rahul Singh, 2023) View paper
    - [22] Regulation of protein-ligand binding affinity by hydrogen bond pairing (Deliang Chen, 2016) View paper

- [34] Evaluating the carcinogenic potential of trazodone hydrochloride via duplex DNA targeting: Molecular interaction, binding mechanism and affinity assessment via â□¦ (P Sharma, 2025) View paper

- Uncategorized or Insufficient Information (2 papers)
  - [12] Prediction of protein-ligand binding affinity with deep learning (Yuxiao Wang, 2023) View paper
  - [20] Protein-Ligand Binding Affinity Prediction (Pratibha Halyal, 2025) View paper

Narrative

Core task: biomolecular binding affinity prediction. The field has evolved from classical physics-based computational methods and mathematical descriptors toward a rich ecosystem of machine learning approaches. At the top level, the taxonomy distinguishes deep learning architectures for structure-based prediction, sequence-based and hybrid feature methods, specialized learning strategies, mathematical and topological descriptors, classical computational techniques, and domain-specific interaction contexts (e.g., protein–protein, protein–ligand, or nucleic acid binding). Deep learning architectures have become particularly prominent, with branches exploring graph neural networks, convolutional models, attention mechanisms, and more recently diffusion-based and foundation model paradigms. Sequence-based methods often leverage pre-trained language models or hybrid encodings that combine sequence and structural information, while specialized learning strategies address challenges such as limited data, multi-task learning, and contrastive objectives. Mathematical and topological approaches (e.g., persistent homology) offer interpretable geometric features, and classical methods remain relevant for benchmarking and physics-informed priors.

Recent years have seen growing interest in large-scale pre-training and foundation models that can generalize across diverse biomolecular contexts. All-Atom Foundation Models[0] exemplifies this trend by learning representations at the atomic level, aiming for broad applicability to various binding prediction tasks. This work sits within the diffusion transformer and foundation model branch, closely related to efforts like Boltz-2[1], which also explores generative and predictive modeling of biomolecular structures. In contrast, many other deep learning methods focus on task-specific architectures—such as graph convolutions for protein–ligand complexes (e.g., MGraphDTA[9], Structure-Aware Interactive[10]) or attention-based fusion strategies (DeepDTAF[4], Deep Fusion Inference[11])—that excel in narrower settings but may require retraining for new interaction types. The shift toward foundation models reflects an ambition to unify disparate prediction tasks under a single learned prior, though open questions remain about how well such models capture fine-grained energetic details compared to specialized or physics-based approaches.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction

Authors: Saro Passaro, Gabriele Corso, Jeremy Wohlwend, Mateo Reveiz, Stephan Thaler, et al. (16 authors total) | **Year/Venue**: 2025 • bioRxiv | **URL**: View paper

Abstract

Accurately modeling biomolecular interactions is a central challenge in modern biology. While recent advances, such as AlphaFold3 and Boltz-1, have substantially improved our ability to predict biomolecular complex structures, these models still fall short in predicting binding affinity, a critical property underlying molecular function and therapeutic efficacy. Here, we present Boltz-2, a new structural biology foundation model that exhibits strong performance for both structure and affinity pr...

Relationship Analysis

Both papers belong to the Diffusion Transformer and Foundation Models category, employing large-scale transformer architectures for biomolecular interaction modeling. They overlap in addressing binding affinity prediction across multiple interaction types (protein-ligand, protein-protein) using foundation model approaches with diffusion-based components. However, the original paper (ADiT) focuses on adapting AlphaFold 3's architecture for representation learning with a denoising pre-training objective, while Boltz-2 emphasizes controllability features (experimental method conditioning, distance constraints) and claims to approach free-energy perturbation (FEP) method performance for small molecule-protein binding affinity prediction.

Contributions Analysis

Overall novelty summary. The paper introduces ADiT, an atom-level diffusion transformer that repurposes AlphaFold 3's architecture for binding affinity prediction across multiple biomolecular interaction types. It resides in the 'Diffusion Transformer and Foundation Models' leaf, which contains only two papers in the entire 50-paper taxonomy. This sparse population suggests the work occupies an emerging research direction where large-scale foundation models are being adapted from generative structure prediction to affinity estimation tasks, rather than the more crowded graph neural network or convolutional branches.

The taxonomy reveals that most structure-based deep learning methods cluster in graph neural network subcategories (interaction-aware, multi-scale, contrastive) and convolutional approaches, which together account for roughly a dozen papers. Transformer and attention-based architectures form a smaller adjacent branch with two papers, while sequence-based and hybrid feature methods constitute another major direction. ADiT diverges from these by combining diffusion processes with transformer blocks at atomic resolution, positioning itself closer to generative modeling paradigms than to task-specific graph or grid-based encoders.

Among 26 candidates examined, the pre-training and fine-tuning framework (Contribution 2) encountered two refutable candidates, indicating that denoising objectives on PDB data have precedent in the limited search scope. The core ADiT architecture (Contribution 1) and the AlphaFold 3 adaptation strategy (Contribution 3) each examined 10 candidates with zero refutations, suggesting these specific design choices—unified tokenization, removal of MSA dependencies, and the shift from generative to representation learning—appear less directly overlapping with prior work in the top-26 semantic matches.

Based on this limited search of 26 candidates, the work appears to introduce novel architectural adaptations in a sparsely populated research area, though the pre-training strategy shows some overlap with existing foundation model efforts. The analysis does not cover the full literature landscape, and a broader search might reveal additional related work in generative modeling or protein language model domains.

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

Contribution 1: Atom-level Diffusion Transformer (ADiT) for biomolecular binding affinity prediction

Description: The authors propose ADiT, a unified foundation model that accepts both sequence and structure inputs for diverse biomolecular interactions. The model uses a unified tokenization scheme for proteins and molecules, incorporates diffusion transformers for multi-level representation learning, and eliminates the need for MSAs and templates unlike AlphaFold 3.

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. High Performance Binding Affinity Prediction with a Transformer-Based Surrogate Model

URL: View paper

Brief Assessment

Transformer-Based Surrogate[72] focuses on SMILES-based transformers for molecular docking surrogates, not atom-level diffusion transformers for structure-based binding affinity prediction. The candidate operates on SMILES representations without structural inputs, while ADiT processes all-atom 3D structures with diffusion transformers.

---

2. Generative Models in Protein Engineering: A Comprehensive Survey

URL: View paper

Brief Assessment

Protein Engineering Survey[74] is a comprehensive survey paper that reviews existing generative models in protein engineering. It does not present novel methods but rather categorizes and discusses existing approaches including diffusion models for protein design. The survey does not claim to introduce ADiT or similar atom-level diffusion transformers for binding affinity prediction.

---

3. ProtT-Affinity: Sequence-Based Protein-Protein Binding Affinity Prediction Using ProtT5 Embeddings

URL: View paper

Brief Assessment

ProtT-Affinity[71] is a sequence-only model using ProtT5 embeddings with a lightweight transformer, while the original paper proposes an atom-level diffusion transformer that accepts both sequence and structure inputs with unified tokenization and eliminates MSA dependencies.

---

4. ProteinReDiff: Complex-based ligand-binding proteins redesign by equivariant diffusion-based generative models

URL: View paper

Brief Assessment

ProteinReDiff[68] focuses on protein redesign using equivariant diffusion models for generating high-affinity ligand-binding proteins, not on binding affinity prediction or representation learning from existing structures.

---

5. Diffbp: Generative diffusion of 3d molecules for target protein binding

URL: View paper

Brief Assessment

DiffBP[53] focuses on generative diffusion for 3D molecule generation targeting protein binding, not on binding affinity prediction using diffusion transformers. The candidate is a generative model for structure prediction, while the original paper adapts AlphaFold 3 for representation learning and affinity prediction.

---

6. DTITR: End-to-end drug-target binding affinity prediction with transformers

URL: View paper

Brief Assessment

DTITR[69] focuses on drug-target binding affinity prediction using 1D sequential data with standard transformers and cross-attention mechanisms, not diffusion transformers or all-atom structural modeling. The architectural approaches and input modalities differ fundamentally from ADiT.

---

7. MolSculptor: an adaptive diffusion-evolution framework enabling generative drug design for multi-target affinity and selectivity

URL: View paper

Brief Assessment

MolSculptor[70] focuses on generative drug design using diffusion models for multi-target inhibitor optimization, not on binding affinity prediction across diverse biomolecular interactions. The candidate employs diffusion transformers for molecular generation in latent space, whereas the original paper uses diffusion transformers for representation learning from protein-ligand structures to predict binding affinity.

---

8. A unified conditional diffusion framework for dual protein targets-based bioactive molecule generation

URL: View paper

Brief Assessment

Dual Targets Conditional[66] focuses on molecule generation for dual protein targets using diffusion models, not binding affinity prediction. The candidate addresses a fundamentally different task (generative design vs. discriminative prediction).

---

9. AptaDiff: de novo design and optimization of aptamers based on diffusion models

URL: View paper

Brief Assessment

AptaDiff[67] focuses on aptamer design using diffusion models for nucleic acid sequences, not on atom-level protein-ligand binding affinity prediction or biomolecular complex representation learning. The technical domains and applications are fundamentally different.

---

10. General Binding Affinity Guidance for Diffusion Models in Structure-Based Drug Design

URL: View paper

Brief Assessment

Binding Affinity Guidance[73] focuses on guiding diffusion models during ligand generation for structure-based drug design, not on building a foundation model for binding affinity prediction across diverse biomolecular interactions. The candidate uses diffusion for generation, while the original uses diffusion transformers for representation learning and affinity prediction.

Contribution 2: Pre-training and fine-tuning framework with denoising objective on PDB dataset

Description: The authors develop a two-stage training approach where ADiT models are first pre-trained on large-scale PDB data using a denoising self-supervised objective, then fine-tuned for downstream binding affinity prediction tasks across multiple interaction types.

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. Full-Atom Peptide Design with Geometric Latent Diffusion

URL: View paper

Brief Assessment

Geometric Latent Diffusion[63] focuses on peptide design using a variational autoencoder with latent diffusion, not a pre-training and fine-tuning framework for binding affinity prediction across multiple interaction types.

---

2. Protein A-like peptide generation based on generalized diffusion model

URL: View paper

Brief Assessment

Protein A-like Peptide[64] focuses on peptide generation using diffusion models, not on binding affinity prediction tasks. The candidate's limited context mentions PDB and denoising diffusion but does not demonstrate a similar pre-training framework for affinity prediction.

---

3. Joint design of protein surface and backbone using a diffusion bridge model

URL: View paper

Brief Assessment

Diffusion Bridge Model[61] focuses on joint surface-structure design for protein-protein interactions using diffusion bridge models, not on binding affinity prediction tasks. The candidate does not employ a pre-training and fine-tuning framework for affinity prediction.

---

4. Pre-training Protein Models with Molecular Dynamics Simulations for Drug Binding

URL: View paper

Prior Art Analysis

Molecular Dynamics Pretraining[65] demonstrates prior work that uses a denoising pre-training approach on protein structures for drug binding tasks. The candidate paper explicitly describes pre-training protein models using a denoising generative task on MD trajectories, followed by fine-tuning for binding affinity prediction. Both papers employ denoising objectives during pre-training and evaluate on binding affinity prediction tasks, with the candidate using PDB-derived data for pre-training. The candidate's framework predates the original paper and establishes that denoising pre-training on structural data for binding prediction was already explored.

Evidence

Evidence 1 - **Rationale**: Both papers use PDB-derived data for pre-training. The candidate uses PDBBind (which is derived from PDB) for their pre-training dataset, showing prior use of PDB data for this purpose. - **Original**: for pre-training, we curate datasets from the protein data bank (pdb) (berman et al., 2000), resulting in 433,297 protein single chains, 481,382 protein-protein interaction examples, and 427,947 proteinligand interaction examples. - **Candidate**: in regards to the pre-training data collection, we selected sixty-four protein-ligand pairs in pdbbind and run their md simulations.

---

5. SE (3) denoising score matching for unsupervised binding energy prediction and nanobody design

URL: View paper

Prior Art Analysis

Denoising Score Matching[62] demonstrates that denoising pre-training on PDB structures for binding-related tasks was established prior to the original paper. The candidate paper presents DSMBind, which trains an energy-based model on PDB crystal structures using SE(3) denoising score matching without binding affinity labels, then applies it to binding energy prediction tasks. This directly refutes the novelty claim of the original paper's two-stage denoising pre-training and fine-tuning framework on PDB data for binding affinity prediction.

Evidence

Evidence 1 - **Rationale**: Both papers train models on PDB-derived datasets for binding affinity prediction tasks. The candidate explicitly excludes binding affinity labels during training, demonstrating the denoising pre-training approach was already established. - **Original**: we train three versions of our adit, namely adit-s, adit-m, and adit-l, on the pdb dataset. these pre-trained models are evaluated across four different types of interactions including protein-ligand, drug-target, protein-protein, and antibodyantigen, and demonstrate state-of-the-art or competitive ... - **Candidate**: we train dsmbind on 4806 protein-ligand complexes from the refined subset of pdbbind v2020 database [26] with their binding affinity labels excluded. we removed all instances from the training set whose ligand appeared the validation and test set. we calculate the spearman correlation between the le...

Evidence 2 - **Rationale**: The candidate paper describes training an energy-based model on crystal structures using denoising score matching, establishing this approach prior to the original paper's claimed contribution. - **Original**: inspired by recent foundation models, we adopt a two-stage training approach: first pre-training our model with a denoising objective, followed by fine-tuning it for specific downstream tasks. pre-training.denoising pre-training has emerged as a widely used self-supervised learning approach for lear... - **Candidate**: our key hypothesis is that we can infer the true binding energy function (up to affine equivalence) by maximizing the likelihood of crystal structures in our training set. the motivation of our hypothesis is that a crystal structure is the lowest energy state of a protein-ligand complex. the maximum...

Evidence 3 - **Rationale**: Both papers pre-train on large PDB datasets without binding labels and then evaluate on binding affinity prediction benchmarks, showing the candidate established this framework earlier. - **Original**: we pre-train three model variants: adit-s (12m params), adit-m (35m params), and adit-l (253m params). the largest variant, adit-l, matches the number of layers and hidden dimensions used in alphafold 3. - **Candidate**: the training set of dsmbind has approximately 27000 non-redundant protein-protein complexes downloaded from pdb (with no binding affinity labels). we compare dsmbind against two physics-based models (rosetta [1] and foldx [4]), two protein-language model baselines (esm-1v [17] and esm-if [6]), and t...

---

6. Harnessing pre-trained models for accurate prediction of protein-ligand binding affinity.

URL: View paper

Brief Assessment

Cannot assess refutation as Pre-Trained Models[37] provides no full text context for comparison with the original paper's denoising pre-training framework on PDB.

Contribution 3: Adaptation strategy for converting AlphaFold 3 from generative to representation learning

Description: The authors present a non-trivial adaptation strategy that transforms AlphaFold 3's generative architecture into a representation learner by simplifying the conditioning module, leveraging the transformer-based atom and sequence-level architecture, and using large-scale pre-training to address data scarcity in affinity prediction.

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. TransVAE-DTA: Transformer and variational autoencoder network for drug-target binding affinity prediction

URL: View paper

Brief Assessment

TransVAE-DTA[58] focuses on drug-target binding affinity using transformer and variational autoencoder architectures. The candidate's full text context is insufficient (only a title fragment), making detailed comparison impossible. No evidence suggests it addresses adapting generative models like AlphaFold 3 to representation learners.

---

2. Machine learning-aided generative molecular design

URL: View paper

Brief Assessment

Generative Molecular Design[52] focuses on machine learning-aided molecular generation and design, not on adapting generative protein structure prediction models to representation learners for binding affinity tasks.

---

3. 3D molecular generative framework for interaction-guided drug design

URL: View paper

Brief Assessment

Interaction-Guided Framework[54] focuses on 3D molecular generation for drug design using interaction-aware conditioning, not on adapting generative models to representation learners for binding affinity prediction. The candidate addresses structure-based ligand design through conditional generation, while the original work addresses converting AlphaFold 3's architecture for affinity prediction tasks.

---

4. A 3D generative model for structure-based drug design

URL: View paper

Brief Assessment

Structure-Based Generative[55] focuses on 3D generative modeling for structure-based drug design using diffusion models, not on adapting generative architectures like AlphaFold 3 for representation learning and binding affinity prediction.

---

5. 3d equivariant diffusion for target-aware molecule generation and affinity prediction

URL: View paper

Brief Assessment

Equivariant Diffusion[51] focuses on generative modeling for molecule design using diffusion processes, not on adapting generative architectures to representation learners for affinity prediction.

---

6. Diffbp: Generative diffusion of 3d molecules for target protein binding

URL: View paper

Brief Assessment

DiffBP[53] does not address adapting AlphaFold 3 or any generative model to representation learning. It proposes a novel diffusion-based generative model for molecule generation, not an adaptation strategy for existing generative architectures.

---

7. A Specialized and Enhanced Deep Generation Model for Active Molecular Design Targeting Kinases Guided by Affinity Prediction Models and Reinforcement â¦

URL: View paper

Brief Assessment

Kinase Active Design[57] focuses on de novo molecular generation for kinase inhibitors using generative models, not on adapting generative architectures like AlphaFold 3 for representation learning or binding affinity prediction tasks.

---

8. Deep generative model for drug design from protein target sequence

URL: View paper

Brief Assessment

Target Sequence Design[60] focuses on generating drug molecules directly from protein sequences using a GAN-based architecture, not on adapting AlphaFold 3's generative architecture for representation learning or binding affinity prediction.

---

9. PocketFlow is a data-and-knowledge-driven structure-based molecular generative model

URL: View paper

Brief Assessment

PocketFlow[59] is a structure-based molecular generative model for drug design, not a representation learning framework. It focuses on generating drug-like molecules inside protein pockets using autoregressive flow, which is fundamentally different from adapting AlphaFold 3's architecture for binding affinity prediction tasks.

---

10. Deep generative molecular design reshapes drug discovery

URL: View paper

Brief Assessment

Deep Generative Reshapes[56] focuses on generative models for molecular design and drug discovery, not on adapting generative architectures like AlphaFold 3 into representation learners for binding affinity prediction.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Towards All-Atom Foundation Models for Biomolecular Binding Affinity Prediction View paper
- [1] Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction View paper
- [2] Calculation of protein-ligand binding affinities View paper
- [3] MHCflurry: open-source class I MHC binding affinity prediction View paper
- [4] DeepDTAF: a deep learning method to predict protein-ligand binding affinity View paper
- [5] Biomolecular interaction prediction: the era of AI View paper
- [6] Computational analysis of protein-ligand interaction by targeting a cell cycle restrainer View paper
- [7] Development and evaluation of a deep learning model for proteinâligand binding affinity prediction View paper
- [8] Structure-based, deep-learning models for protein-ligand binding affinity prediction View paper
- [9] MGraphDTA: deep multiscale graph neural network for explainable drugâtarget binding affinity prediction View paper

- [10] Structure-aware Interactive Graph Neural Networks for the Prediction of Protein-Ligand Binding Affinity View paper
- [11] Improved proteinâ□ligand binding affinity prediction with structure-based deep fusion inference View paper
- [12] Prediction of protein-ligand binding affinity with deep learning View paper
- [13] Multi-task bioassay pre-training for protein-ligand binding affinity prediction View paper
- [14] DeepAtom: a framework for protein-ligand binding affinity prediction View paper
- [15] Recent advances in computational and experimental protein-ligand affinity determination techniques View paper
- [16] Prediction of proteinâ□ligand binding affinity via deep learning models View paper
- [17] Persistent spectralâ□based machine learning (PerSpect ML) for protein-ligand binding affinity prediction View paper
- [18] Drug-target binding affinity prediction model based on multi-scale diffusion and interactive learning View paper
- [19] Computational studies of proteinâ□drug binding affinity changes upon mutations in the drug target View paper
- [20] Protein-Ligand Binding Affinity Prediction View paper
- [21] Deep learning in drug design: protein-ligand binding affinity prediction View paper
- [22] Regulation of protein-ligand binding affinity by hydrogen bond pairing View paper
- [23] Contrastive pre-training and 3D convolution neural network for RNA and small molecule binding affinity prediction View paper
- [24] Reliable method for predicting the binding affinity of RNA-small molecule interactions using machine learning View paper
- [25] Current State of Open Source Force Fields in Proteinâ□Ligand Binding Affinity Predictions View paper
- [26] Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: a review View paper
- [27] PLANET: A Multi-Objective Graph Neural Network Model for Proteinâ□Ligand Binding Affinity Prediction View paper
- [28] Onionnet: a multiple-layer intermolecular-contact-based convolutional neural network for proteinâ□ligand binding affinity prediction View paper
- [29] CAPLA: improved prediction of proteinâ□ligand binding affinity by a deep learning approach based on a cross-attention mechanism View paper
- [30] Proteinâ□ligand binding affinity prediction with edge awareness and supervised attention View paper
- [31] Modern machineâ□learning for binding affinity estimation of proteinâ□ligand complexes: Progress, opportunities, and challenges View paper
- [32] Assessment of programs for ligand binding affinity prediction View paper
- [33] Distance plus attention for binding affinity prediction View paper
- [34] Evaluating the carcinogenic potential of trazodone hydrochloride via duplex DNA targeting: Molecular interaction, binding mechanism and affinity assessment via â□¦ View paper
- [35] Recent advances in machine learning predictions of protein-ligand binding affinities View paper
- [36] NHGNN-DTA: a node-adaptive hybrid graph neural network for interpretable drugâ□target binding affinity prediction View paper
- [37] Harnessing pre-trained models for accurate prediction of protein-ligand binding affinity. View paper
- [38] Computational methods for calculation of ligand-binding affinity View paper
- [39] Hybrid Quantum Neural Network Approaches to Proteinâ□Ligand Binding Affinity Prediction View paper
- [40] Learning from the ligand: using ligand-based features to improve binding affinity prediction View paper
- [41] GraphCL-DTA: A Graph Contrastive Learning With Molecular Semantics for Drug-Target Binding Affinity Prediction View paper
- [42] Pred-binding: large-scale proteinâ□ligand binding affinity prediction View paper
- [43] DataDTA: a multi-feature and dual-interaction aggregation framework for drugâ□target binding affinity prediction View paper
- [44] Interaction-Based Inductive Bias in Graph Neural Networks: Enhancing Protein-Ligand Binding Affinity Predictions From 3D Structures View paper
- [45] Computational prediction of proteinâ□protein binding affinities View paper
- [46] Proteinâ□protein binding affinity prediction from amino acid sequence View paper
- [47] Binding energy distribution analysis method (BEDAM) for estimation of Proteinâ□ Ligand binding affinities View paper
- [48] Systematic investigation of machine learning on limited data: A study on predicting protein-protein binding strength View paper
- [49] Structure-Aware Graph Attention Diffusion Network for Proteinâ□Ligand Binding Affinity Prediction View paper
- [50] Persistent Path-Spectral (PPS) Based Machine Learning for Protein-Ligand Binding Affinity Prediction View paper
- [51] 3d equivariant diffusion for target-aware molecule generation and affinity prediction View paper
- [52] Machine learning-aided generative molecular design View paper
- [53] Diffbp: Generative diffusion of 3d molecules for target protein binding View paper
- [54] 3D molecular generative framework for interaction-guided drug design View paper
- [55] A 3D generative model for structure-based drug design View paper
- [56] Deep generative molecular design reshapes drug discovery View paper
- [57] A Specialized and Enhanced Deep Generation Model for Active Molecular Design Targeting Kinases Guided by Affinity Prediction Models and Reinforcement â□¦ View paper
- [58] TransVAE-DTA: Transformer and variational autoencoder network for drug-target binding affinity prediction View paper
- [59] PocketFlow is a data-and-knowledge-driven structure-based molecular generative model View paper
- [60] Deep generative model for drug design from protein target sequence View paper
- [61] Joint design of protein surface and backbone using a diffusion bridge model View paper
- [62] SE (3) denoising score matching for unsupervised binding energy prediction and nanobody design View paper
- [63] Full-Atom Peptide Design with Geometric Latent Diffusion View paper
- [64] Protein A-like peptide generation based on generalized diffusion model View paper
- [65] Pre-training Protein Models with Molecular Dynamics Simulations for Drug Binding View paper
- [66] A unified conditional diffusion framework for dual protein targets-based bioactive molecule generation View paper
- [67] AptaDiff: de novo design and optimization of aptamers based on diffusion models View paper
- [68] ProteinReDiff: Complex-based ligand-binding proteins redesign by equivariant diffusion-based generative models View paper
- [69] DTITR: End-to-end drug-target binding affinity prediction with transformers View paper
- [70] MolSculptor: an adaptive diffusion-evolution framework enabling generative drug design for multi-target affinity and selectivity View paper
- [71] ProtT-Affinity: Sequence-Based Protein-Protein Binding Affinity Prediction Using ProtT5 Embeddings View paper
- [72] High Performance Binding Affinity Prediction with a Transformer-Based Surrogate Model View paper
- [73] General Binding Affinity Guidance for Diffusion Models in Structure-Based Drug Design View paper
- [74] Generative Models in Protein Engineering: A Comprehensive Survey View paper