IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

YuelDesign: AI Diffusion Models for Protein Flexibility

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-04-18 · yueldesign · ai drug design · diffusion models · protein flexibility · computational biology · ligand binding · generative ai · structure-based drug design

Original article: https://intuitionlabs.ai/articles/yueldesign-ai-drug-design-protein-flexibility



In this article

1. Executive Summary
2. 1. Introduction
3. 1.1. Challenges of Modern Drug Discovery
4. 2. Background: Protein Flexibility and AI in Drug Design
5. 2.1. Protein Flexibility in Biology and Drug Discovery
6. 2.2. AI and Diffusion Models in Molecular Generation
7. 3. The YuelDesign Framework
8. 3.1. YuelDesign, YuelPocket, and YuelBond: UVA's AI Suite
9. 3.2. Data and Training
10. 3.3. Neural Architecture: The E3former Backbone
11. 3.4. Dual Diffusion Generators: EDM and D3PM
12. 4. Results and Analysis
13. 4.1. Chemical Properties of Generated Molecules
14. 4.2. Chemical Diversity and Functional Groups
15. 4.3. Pocket Conformation and Binding Geometry
16. 4.4. Docking Scores and Binding Affinity Metrics
17. 4.5. Case Study: Cyclin-Dependent Kinase 2 (CDK2)
18. 4.6. Structural Evolution During Generation
19. 5. Comparative Perspectives
20. 5.1. Rigid vs. Flexible Pocket Models
21. 5.2. Other AI Generative Approaches
22. 5.3. Advantages and Limitations
23. 6. Discussion and Future Directions
24. 6.1. Implications for Drug Discovery
25. 6.2. Future Work and Enhancements
26. 6.3. Considerations and Cautions
27. 7. Conclusion
28. References

Executive Summary

Drug discovery remains a costly, time-consuming, and high-failure endeavor. Protein flexibility complicates predictions of how targets will accommodate new drug molecules (www.eurekalert.org). Traditional methods typically treat the protein binding site as a rigid "lock," while in reality the protein "lock" is **constantly jiggling** and changing shape as a ligand approaches (www.eurekalert.org). A new suite of AI tools developed at the University of Virginia (UVA) – led by Dr. Nikolay V. Dokholyan – aims to overcome this limitation by **co-designing** drug ligands *and* their protein pockets simultaneously using advanced diffusion models (www.eurekalert.org) (pmc.ncbi.nlm.nih.gov). The centerpiece, **YuelDesign**, generates new small-molecule

drug candidates that are well-matched to their protein targets even as those targets flex and shift (www.eurekalert.org) (pmc.ncbi.nlm.nih.gov). Companion tools **YuelPocket** and **YuelBond** respectively identify optimal binding sites and reconstruct chemical bonds (Wang and Dokholyan).

In validation studies reported in *Science Advances* (April 2026) and news releases (pmc.ncbi.nlm.nih.gov) (www.eurekalert.org), YuelDesign-produced molecules showed higher drug-likeness (QED scores) and fewer problematic large rings than the compared diffusion-based methods, with broadly comparable synthetic accessibility scores, while maintaining chemical validity (pmc.ncbi.nlm.nih.gov). Crucially, YuelDesign's co-generation of ligand and pocket yielded **better docking scores and more native-like binding interactions** than rigid-pocket methods (pmc.ncbi.nlm.nih.gov). For example, in the case of the cancer target **CDK2**, only YuelDesign captured the native breakage of a key salt bridge (between Lys-33 and Asp-145) that allows tight binding (med.virginia.edu) (pmc.ncbi.nlm.nih.gov). In short, YuelDesign represents a significant advance in AI-driven drug design by honoring the biological reality of protein flexibility. Its development on UVA's AI drug design platform suggests a promising new direction for more efficient and effective lead discovery.

1. Introduction

1.1. Challenges of Modern Drug Discovery

Clinical attrition spans human testing, rather than late-stage trials alone. The BIO, Informa Pharma Intelligence and QLS Advisors analysis reported an overall likelihood of approval from Phase I of **7.9%** across all developmental candidates in its 2011–2020 dataset (report). A key reason for these dismal statistics is the difficulty of predicting how a drug molecule will **bind** to its protein target in the complex, dynamic environment of the body. A compound may look promising in silico (on the computer) by docking to a static structure of the protein, but when tested biologically it might bind too weakly or at the wrong place, yielding no therapeutic benefit or causing side effects (www.eurekalert.org).

At the molecular level, the challenge arises because **proteins are not rigid**. Classical models like the "lock-and-key" hypothesis assume the protein binding site is fixed, but empirical studies have long shown that most proteins can adopt multiple conformations. In many cases ligand binding induces significant rearrangements in the binding pocket (the *induced-fit* phenomenon) (www.eurekalert.org) (pmc.ncbi.nlm.nih.gov). For example, seminal reviews note that a *"large number of pharmaceutically relevant targets"* are highly flexible, raising serious doubt about treating them as rigid during in silico design (www.nature.com). Even modest side-chain movements or loop shifts within the binding site—often ≤1–2 Å RMSD—can dramatically change binding affinity (pmc.ncbi.nlm.nih.gov) (med.virginia.edu). Nonetheless, most computational drug-design methods to date have continued to assume fixed pockets, because explicitly modeling protein motion has been computationally expensive and difficult.

Meanwhile, **artificial intelligence (AI)** and machine learning have rapidly advanced drug design. Generative models can now propose novel molecules de novo, optimizing for properties like drug-likeness or receptor binding. Early generations of these models were "structure-free" – e.g. language models or graph generators operating on SMILES strings (pmc.ncbi.nlm.nih.gov). More recently, **3D generative models** have emerged, treating molecules as point clouds or graphs and using techniques like equivariant neural networks or diffusion

processes to propose 3D structures. Examples include diffusion-based approaches such as *DiffSBDD* (structure-based drug design) (pmc.ncbi.nlm.nih.gov), *DiffSMol* and *DiffBP* (pmc.ncbi.nlm.nih.gov), which can generate ligand coordinates in the context of a protein pocket. However, these pioneering methods largely **treat the protein pocket as static**, optimizing the ligand against a fixed receptor geometry (pmc.ncbi.nlm.nih.gov).

The emergence of generative AI models (like diffusion models) in drug discovery is well-timed with advances in protein structure prediction (e.g. AlphaFold) and ML-based scoring, but a persistent blind spot remains: most models ignore receptor flexibility. This gap motivated the UVA team to develop **YuelDesign**, a diffusion-based framework that explicitly **models pocket dynamics during ligand generation** (pmc.ncbi.nlm.nih.gov). By "designing the key while the lock is moving," the UVA approach aims to produce drug candidates that are better fit to realistic, flexible targets (www.eurekalert.org).

In the following sections, we examine the YuelDesign framework in depth. We review its architectural innovations, how it was trained and tested on structural data, and the results demonstrating its benefits. We compare YuelDesign to other generative methods, analyze its performance data (including metrics like QED, synthetic accessibility, and docking scores), and discuss illustrative case studies. Finally, we consider the broader implications for drug discovery, challenges ahead, and future directions.

2. Background: Protein Flexibility and AI in Drug Design

2.1. Protein Flexibility in Biology and Drug Discovery

Biomolecules are inherently dynamic: they wiggle, breathe, and switch conformations. Classic biochemistry recognized this decades ago. The induced-fit concept, first articulated by Daniel E. Koshland in 1958, posited that *"the binding of a substrate (ligand) to an enzyme induces a change in the structure of the enzyme that facilitates catalysis"* (pmc.ncbi.nlm.nih.gov). In other words, proteins often adjust their shape to accommodate ligands. Subsequent experimental evidence has documented many proteins with "tense" and "relaxed" states (e.g. oxygenated vs deoxygenated hemoglobin) and shows that ligand binding can stabilize one conformation over another (pmc.ncbi.nlm.nih.gov).

Modern structural biology confirms that most proteins have innate flexibility. Cozzini *et al.* (2008) note that *"proteins have an intrinsic ability to undergo functionally relevant conformational transitions under native state conditions"* (pmc.ncbi.nlm.nih.gov). Such motions can occur over a wide range of time scales, from picoseconds to seconds, and involve small side-chain rotations up to large domain movements (pmc.ncbi.nlm.nih.gov). For pharmacological targets, these shifts are often critical. For example, many kinases (common cancer drug targets) adopt distinct "DFG-in" and "DFG-out" active/inactive states upon ligand binding; nuclear hormone receptors rely on repositioning of an entire helix; G-protein coupled receptors switch between active/inactive rotamer ensembles, etc. These flexibilities determine whether a potential drug fits

snugly or is repelled. Brylinski and Skolnick (2008) analyzed 521 apo–holo structures and found that while the **backbones** of proteins typically move <1 Å on average, side-chain conformations can shift enough to alter binding-site shape (pmc.ncbi.nlm.nih.gov).

In practical drug discovery, ignoring protein flexibility can mislead lead optimization. Compounds that seem to dock well to a rigid protein model may perform poorly in vitro if the protein actually adopts a different conformation. Conversely, some ligands only bind effectively by *inducing* a pocket change. Capturing this interplay (often called ensemble docking or flexible docking) traditionally demanded expensive molecular dynamics or sampling methods. Only a few docking programs incorporate limited side-chain rotamer sampling or explicit induced-fit algorithms (pmc.ncbi.nlm.nih.gov). Thus, the "rigid receptor hypothesis" remains a simplifying assumption in most in silico pipelines (www.nature.com).

The consequence is like designing a key for a lock that **never** moves. Dokholyan of UVA illustrates this metaphor: a drug must be designed as the protein "jiggles and changes shape" to realistically bind (www.eurekalert.org). If the model treats the protein as a frozen statue, the fit is often unrealistic and promising candidates will fail in experiments (www.eurekalert.org). As one industry summary put it, current AI tools *"overcome limitations of existing options"* by treating proteins as flexible rather than frozen snapshots (www.eurekalert.org). In light of this, there is a clear opportunity: generative models that **co-evolve** the protein pocket and ligand can better emulate induced-fit binding.

2.2. AI and Diffusion Models in Molecular Generation

Over the last decade, AI methods for molecular design have matured rapidly. Early approaches used genetic algorithms or Bayesian optimization for targeted modifications. Recently, deep generative networks have been used for de novo drug design. These approaches fall broadly into two classes: **2D/graph generation** and **3D structure generation** (pmc.ncbi.nlm.nih.gov).

- **2D Generative Models:** These include language models (e.g. RNNs over SMILES or SELFIES strings) and graph-based VAEs or GANs. Pioneering work (GraphVAE, MolGAN, etc. in 2018–2020) could generate novel molecular graphs with desirable property distributions (pmc.ncbi.nlm.nih.gov). While powerful, 2D models require separate geometry (conformer) prediction and typically do not account for target protein context.
- **3D Equivariant Models:** More recently, researchers have developed generative models that output full 3D coordinates of molecules (and sometimes bound complexes). Diffusion probabilistic models – which iteratively "denoise" random configurations into realistic ones – have become prominent. For example, DiffSBDD (Schneuing *et al.*, 2024) uses an equivariant graph transformer to generate ligand geometry within a protein pocket, but assumes the pocket is fixed (pmc.ncbi.nlm.nih.gov). Other 3D diffusion-based methods include Pocket-based Molecular Diffusion (PMDM), DiffSMol, DiffBP, and structure-conditioned methods like TargetDiff (pmc.ncbi.nlm.nih.gov).

Beyond diffusion, other frameworks such as normalizing flows (GraphNVP/MoFlow) or reinforcement learning are also used for molecules. But diffusion models have shown exceptional success at capturing high-dimensional continuous structures (pmc.ncbi.nlm.nih.gov). Crucially, standard diffusion frameworks generate molecules **given** a fixed pocket or binding pose, rather than learning the pocket's response.

In parallel, major advances like AlphaFold have revolutionized protein structure prediction, enabling the modeling of many targets. For example, AlphaFold 3 (May 2024) now predicts not only protein structures but also interactions with small molecules, effectively modeling complexes (time.com). However, even AlphaFold's interaction predictions operate on static snapshots; they do not inherently optimize a ligand's chemistry. In short, while AI is transforming structural biology (AlphaFold) and molecule generation (diffusion models), **jointly modeling both in a dynamic binding scenario** is largely unexplored.

The UVA YuelDesign work addresses this gap. By combining a graph Transformer backbone with dual diffusion processes, it **simultaneously learns pocket and ligand structures** (pmc.ncbi.nlm.nih.gov). This co-generation explicitly captures induced-fit effects: as the ligand is formed, the protein pocket can flex in tandem, leading to more realistic binding configurations. Such methods represent a new frontier in AI-driven drug design, promising to accelerate lead discovery by bridging the protein flexibility that has historically been overlooked.

3. The YuelDesign Framework

3.1. YuelDesign, YuelPocket, and YuelBond: UVA's AI Suite

UVA's AI drug design platform now comprises a **suite of tools** developed by Nikolay Dokholyan's group. The primary components are:

- **YuelDesign**: A novel generative model using **equivariant diffusion** to design small-molecule ligands that fit flexibly-defined protein pockets. It is the centerpiece of the platform (www.eurekalert.org) (pmc.ncbi.nlm.nih.gov).
- **YuelPocket**: A graph-neural-network model that **predicts ligand-binding pockets** on arbitrary protein structures. It can operate even on predicted structures (e.g. AlphaFold models) to highlight likely druggable sites (med.virginia.edu).
- **YuelBond**: A multimodal graph neural network for **bond reconstruction**, including bond-order reassignment and reconstruction from generated molecular geometries (Wang and Dokholyan).

These components work in concert. In a typical workflow, YuelPocket first analyzes a target protein to define a binding region. YuelDesign then generates candidate drug molecules within that pocket, adapting to its flexibility. YuelBond can then reconstruct bonds in the generated molecular structures (Wang and Dokholyan). Together, the pipeline "transforms how new drugs are created" by automating pocket identification and flexible-design in one framework (med.virginia.edu).

Component	Purpose	Methodology (Key Technology)	Reference
YuelDesign	Generates novel ligand structures for a given protein pocket, modeling pocket flexibility	E(3)-equivariant Transformer (E3former) + dual diffusion (EDM for coordinates, D3PM for atom types) (pmc.ncbi.nlm.nih.gov)	UVA News (www.eurekalert.org), SciAdv (pmc.ncbi.nlm.nih.gov)

Component	Purpose	Methodology (Key Technology)	Reference
YuelPocket	Identifies optimal ligand-binding sites on a protein (including predicted structures)	Graph Neural Network recognizing pocket geometry (med.virginia.edu)	UVA News (med.virginia.edu)
YuelBond	Reconstructs bonds in generated molecular structures	Multimodal graph neural network for bond reconstruction and bond-order reassignment	Wang and Dokholyan (PubMed)

Table 1: Summary of the UVA AI drug-design tools. The suite (YuelDesign, YuelPocket, YuelBond) operates together to predict binding sites and co-design ligand–protein complexes. Citations indicate sources for each tool's description.

3.2. Data and Training

YuelDesign was trained on high-quality protein–ligand complexes from the **Binding MOAD** database (pmc.ncbi.nlm.nih.gov). Binding MOAD is a curated repository of ~41,409 crystal structures of protein–ligand complexes, of which roughly 15,223 have quantitative binding affinity data (pmc.ncbi.nlm.nih.gov). From this dataset, relevant binding pockets were extracted. In pre-processing, the authors defined the binding site as all protein residues having at least one atom within 6 Å of the ligand (pmc.ncbi.nlm.nih.gov). All pocket and ligand atoms (with 3D coordinates) were treated together, building a joint coordinate set. Each atom was annotated by a one-hot type feature (C, N, O, etc.) and whether it belonged to the protein or ligand (pmc.ncbi.nlm.nih.gov).

Crucially, the dataset was split (80% train, 20% test) with care to avoid information leakage. Similar proteins (>30% sequence identity) or similar ligands (Tanimoto >0.85) were placed in the same split (pmc.ncbi.nlm.nih.gov). This *homology filtering* ensured the model is tested on genuinely novel pockets and chemistries, avoiding trivial memorization. In total, YuelDesign trained on tens of thousands of complexes sampled from MOAD, spanning diverse targets.

3.3. Neural Architecture: The E3former Backbone

At the heart of YuelDesign is **E3former**, an equivariant Transformer adapted from AlphaFold's Evoformer block (pmc.ncbi.nlm.nih.gov). E3former operates on a combined sequence of protein-pocket and ligand atoms. Each atom is a "token" with associated features (atom type, residue context, etc.), and each pair of atoms has relational features (e.g. distance, whether they are in the same residue) (pmc.ncbi.nlm.nih.gov). This lets E3former reason over both chemical identity and spatial geometry simultaneously.

Key features of E3former include:

- **Equivariance**: The model maintains *E(3) equivariance* to rotations and translations, meaning predictions are consistent regardless of how the complex is oriented. This is implemented via specialized attention and coordinate update modules that use geometric vectors between atoms (pmc.ncbi.nlm.nih.gov).

- **Coordinate head**: Unlike AlphaFold's Evoformer (which outputs distance distributions), E3former directly outputs 3D coordinate displacements for atoms. A final "coordinate head" aggregates direction vectors between atoms, ensuring each predicted shift respects equivariance (pmc.ncbi.nlm.nih.gov).
- **Single-sequence input**: Because the system does not rely on evolutionary MSA data, E3former uses single-chain attention separately for protein and ligand, combined via "triangle" attention/multiplication modules that propagate spatial context (pmc.ncbi.nlm.nih.gov).
- **Dual diffusion interface**: After embedding, the E3former outputs feed into two diffusion processes: one for coordinates (EDM) and one for atom types (D3PM). The model thus learns to refine both geometry and chemistry together (pmc.ncbi.nlm.nih.gov).

This design allows E3former to handle up to ~400 atoms at a time (full-atom protein side chains + ligand), a much larger dimensionality than prior ligand-only generation models (which were ~15–50 atoms). By leveraging triangle self-attention, the network captures long-range spatial relations across the complex.

3.4. Dual Diffusion Generators: EDM and D3PM

YuelDesign employs **two coupled diffusion models**:

- **EDM (Elucidated Diffusion Model)** for **continuous 3D coordinates** (pmc.ncbi.nlm.nih.gov). EDM treats each atom's 3D position as a continuous variable. During generation, atoms start from random noise in space and are iteratively "denoised" to converge on a realistic geometry for both the protein pocket and ligand. Noise injection and denoising steps respect the geometry; thanks to E3former's coordinate head, updates are equivariant.
- **D3PM (Discrete Denoising Diffusion Probabilistic Model)** for **categorical atom types** (pmc.ncbi.nlm.nih.gov). D3PM handles the discrete nature of atom/element identity. Initially, ligand atom types are randomized (like a one-hot bitvector with random category), and a Markovian diffusion process gradually selects realistic atom identities. In conjunction with coordinate diffusion, the model learns how geometry and chemistry co-evolve.

These dual processes run in tandem within a single framework. At each diffusion step, E3former processes a partially-formed complex and outputs both coordinate updates and atom-type distributions. The training objective enforces that at step 0 (final output), the pocket and ligand match the ground truth. In practice, the model is trained end-to-end to regress corrupt inputs to the correct atomic configuration.

D3PM predicts ligand atom types, while EDM generates continuous coordinates; bond reconstruction follows denoising. Critically, the model **gives the protein pocket freedom to move** during the generation – instead of anchoring the pocket, E3former can reorient side chains as needed. This co-generation of pocket and ligand is the key innovation: rather than treating pocket coordinates as fixed context, **the pocket itself is a trainable output** (pmc.ncbi.nlm.nih.gov).

4. Results and Analysis

The performance of YuelDesign was evaluated extensively on the held-out test set of protein–ligand complexes, using both intrinsic quality metrics and docking simulations. All reported comparisons utilize the same benchmarks, running YuelDesign side-by-side with two representative rigid-pocket diffusion methods: **DiffSBDD** (Schneuing *et al.*, Nat. Comput. Sci. 2024) and **PMDM** (Huang *et al.*, *Nature Communications* **15**, 2657 (2024), DOI: 10.1038/s41467-024-46569-1). Below we summarize key findings, drawing primarily on the *Science Advances* paper's figures and text (pmc.ncbi.nlm.nih.gov).

4.1. Chemical Properties of Generated Molecules

A core question is whether YuelDesign can produce *drug-like* molecules. The authors evaluated generated ligands on metrics including **Quantitative Estimate of Drug-likeness** (QED) (pmc.ncbi.nlm.nih.gov), **Lipinski's Rule-of-Five** compliance, **synthetic accessibility score (SAS)** (pmc.ncbi.nlm.nih.gov), presence of large (macro) rings, and basic validity/connectivity (single component, no errors) (pmc.ncbi.nlm.nih.gov). These were compared across ligand sizes (number of heavy atoms) and against DiffSBDD and PMDM as baselines (pmc.ncbi.nlm.nih.gov).

Several trends emerged (see Table 2). Notably, **YuelDesign outperformed the others on key "desirability" metrics**. Its molecules maintained **higher connectivity** (nearly always a single bonded component) and a **much lower rate of problematic large rings** compared to DiffSBDD and especially PMDM (pmc.ncbi.nlm.nih.gov). For example, as molecule size increases, PMDM-produced structures often contained multiple disjoint fragments or excessive macrocycles, whereas YuelDesign kept nearly all atoms connected and formed fewer rings >6 atoms (pmc.ncbi.nlm.nih.gov).

In drug-likeness, YuelDesign consistently achieved **higher QED scores** than the other methods across all size bins (pmc.ncbi.nlm.nih.gov). Lipinski RO5 compliance (measuring whether MW, logP, H-bond donors/acceptors are in drug-like ranges) was also better for YuelDesign, especially on smaller molecules (pmc.ncbi.nlm.nih.gov). All methods saw QED and RO5 scores decline for very large molecules (as expected), but YuelDesign's decline was more gradual. Synthetic accessibility (SAS) scores – lower being easier – were roughly comparable across methods, indicating no major synthetic burden added by YuelDesign (pmc.ncbi.nlm.nih.gov). All three models produced chemically valid structures nearly 100% of the time once connectivity was assured (pmc.ncbi.nlm.nih.gov).

<table> <thead><tr><th>Metric</th><th>YuelDesign</th><th>DiffSBDD</th><th>PMDM</th> <th>Source/Notes</th></tr></thead> <tbody> <tr><td>Connectivity (single component)</td><td>High (near 1.0)</td><td>Medium (declines with size)</td><td>Low (significant fragmentation at large sizes)</td> <td>Higher is better (pmc.ncbi.nlm.nih.gov)</td></tr> <tr><td>Large-ring rate (>6 atoms)</td><td>Lowest</td> <td>Moderate</td><td>Highest</td><td>Lower rate indicates fewer synthetic issues (pmc.ncbi.nlm.nih.gov) </td></tr> <tr><td>QED (drug-likeness)</td><td>Highest across sizes</td><td>Lower</td><td>Lower</td> <td>YuelDesign maintains higher QED at all sizes (pmc.ncbi.nlm.nih.gov)</td></tr> <tr><td>Lipinski's RO5 compliance</td><td>Best for small molecules</td><td>Worse (small); declines similarly at large sizes</td> <td>Worse (small); declines similarly</td><td>YuelDesign outperforms at small sizes (pmc.ncbi.nlm.nih.gov) </td></tr> <tr><td>Synthetic Accessibility (SAS)</td><td>Comparable to others</td><td>Comparable</td>

<td>Comparable</td><td>All methods yield similar SAS distributions (pmc.ncbi.nlm.nih.gov)</td></tr> <tr> <td>Validity (chemical rules)</td><td>~100%</td><td>~100%</td><td>~100%</td><td>All ensure valid molecules (except PMDM slight dip at size 35) (pmc.ncbi.nlm.nih.gov)</td></tr> </tbody> </table>

Table 2: Comparison of generative model output (molecules for target Pteridine reductase 1, PDB 3JQA) on key quality metrics. "Connectivity" measures whether all atoms form a single bonded component. YuelDesign consistently avoids large rings and retains high QED, as described by the authors (pmc.ncbi.nlm.nih.gov).

The authors attribute YuelDesign's superior performance to its modeling of pocket flexibility. For instance, Figure 2B (from the paper) showed that YuelDesign's ligand designs triggered pocket adjustments that improved binding interactions (e.g. additional π–π stacking) compared to the rigid methods (pmc.ncbi.nlm.nih.gov). In practical terms, YuelDesign could generate more structurally complex but still drug-like molecules. In summary, YuelDesign's outputs exhibit favorable physicochemical profiles required for drug candidates, often better than rigid-pocket diffusion methods (pmc.ncbi.nlm.nih.gov).

4.2. Chemical Diversity and Functional Groups

Diversity of chemical functionality is another critical factor. An ideal generative model should explore varied scaffolds and functional group combinations, rather than collapsing to a narrow chemical class. The authors analyzed the frequency of major functional groups (aromatic rings, ethers, amines, halogens, etc.) in the generated molecules versus the native ligands from crystal structures (pmc.ncbi.nlm.nih.gov).

Overall, YuelDesign's molecules **mirrored the native ligand distribution** quite closely. As shown in their Figure 3B, alcohol groups appeared in ~85% of both YuelDesign ligands and native complexes, with amines present in ~90% of Yuel designs vs ~70% of natives (pmc.ncbi.nlm.nih.gov). This suggests YuelDesign is not simply inventing exotic chemistries; it learns typical drug-like functional patterns. A slight discrepancy was enrichment of small rings (e.g. cyclopropane, epoxide) in Yuel designs (pmc.ncbi.nlm.nih.gov). The paper notes this as a known quirk: when sampling from Gaussian noise, very small cycles can form easily without violating distance constraints, even though real drugs avoid them. This indicates a limitation of current diffusion formulations (there is no hard "no 3-member ring" rule).

Nevertheless, the main takeaway is that **YuelDesign captures a broad chemical vocabulary similar to actual drugs** (pmc.ncbi.nlm.nih.gov). Its ability to generate a variety of heterocycles, aromatic rings, polar groups, etc., aligns well with medicinal chemistry norms. One implication is that YuelDesign is not "mode-collapsing" to a few chemotypes; it produces **diverse structures with balanced functional content** (pmc.ncbi.nlm.nih.gov). This diversity is important for exploring chemical space around a target, suggesting YuelDesign could uncover novel scaffolds not represented in training data.

4.3. Pocket Conformation and Binding Geometry

Since YuelDesign explicitly models the protein pocket, it is crucial to evaluate the **predicted pocket conformations**. In their analysis, the authors computed the root-mean-square deviation (RMSD) of the generated pocket atoms versus the original crystal pocket for each test complex. They found a **median RMSD of ~1.8 Å** (pmc.ncbi.nlm.nih.gov). This amount of deviation is biologically reasonable: it indicates that YuelDesign typically made modest side-chain adjustments (often <2 Å) rather than wholesale rearrangements.

By comparison, purely flexible-backbone docking methods often expect sub-2 Å shifts as normal (MedusaDock and AutoDock Vina, which sample rotamers, operate on similar scales (pmc.ncbi.nlm.nih.gov)). As one reference notes, about 80% of apo–holo pairs differ by ≤1 Å (pmc.ncbi.nlm.nih.gov), so a median of 1.8 Å suggests YuelDesign is capturing realistic, sometimes slightly larger, induced fits.

Figure 4A from the paper plots the distribution of pocket RMSDs. About 50% of pockets had <2 Å deviation, and only a small tail exceeded 3–4 Å. The authors interpret this as “reasonable”: it shows the model is not wildly distorting pockets, yet is allowing more flexibility than rigid docking. Indeed, they highlight one case (Pteridine reductase 1, PDB 3JQA) where YuelDesign adjusted certain side chains to form new polar contacts while preserving key π–π stacking interactions (pmc.ncbi.nlm.nih.gov). In other words, YuelDesign found a conformation that better complements its ligand, unlike static methods.

These pocket adjustments had real impact on binding metrics. YuelDesign’s ability to slightly reorient residues translated to stronger predicted affinity. As discussed below, YuelDesign’s ligands achieved better docking scores against the “moving” pockets they generated, in contrast to DiffSBDD/PMDM which used unaltered pockets.

4.4. Docking Scores and Binding Affinity Metrics

To quantify binding compatibility, the study computed scores from two widely-used docking programs: **AutoDock Vina** and **MedusaDock** (pmc.ncbi.nlm.nih.gov). They took each designed ligand and docked it into its protein pocket. (For YuelDesign, the pocket is as generated; for DiffSBDD/PMDM, the pocket is fixed to the original structure.)

The results were clear: **YuelDesign-generated ligands had significantly better docking scores**. Figure 4D-E of the paper show the distributions. In both Vina’s and Medusa’s scoring, YuelDesign ligands achieved on average *lower* (more favorable) energies than those from DiffSBDD or PMDM. The text states: *“Molecules generated by YuelDesign consistently achieved better scores than those from the other two methods”* (pmc.ncbi.nlm.nih.gov). This indicates that allowing pocket flexibility enabled discovering ligand conformations and poses that fit more tightly.

Further, a “redocking” analysis (taking the designed ligand, re-docking into its designed pocket using MedusaDock) reinforced the effect. YuelDesign compounds resulted in *lower RMSD* to their original docked pose compared to DiffSBDD/PMDM ligands (pmc.ncbi.nlm.nih.gov), indicating greater consistency between generated and redocked poses.

No designed method exactly recapitulated native ligand chemistries: the Tanimoto similarity between any designed ligand and the original crystal ligand (e.g. DX4 in 3JQA) was low for all models (pmc.ncbi.nlm.nih.gov). But YuelDesign did edge out the others slightly; it managed to generate molecules that were at least somewhat closer analogs to the real binders. Importantly, YuelDesign’s improvement in docking scores was not merely due to copying known chemistry, but due to **intrinsically better complementarity with the pocket**.

Taken together, these computational analyses suggest **predicted docking energies comparable to native ligands** (Wang et al.). They do not establish experimental potency, binding stability, efficacy, or clinical benefits.

4.5. Case Study: Cyclin-Dependent Kinase 2 (CDK2)

A compelling demonstration of YuelDesign's unique capabilities is provided by the analysis on **CDK2** – a well-known kinase involved in cell cycle regulation and cancer therapy. CDK2 is famous for undergoing a salt-bridge rearrangement upon inhibitor binding: in the apo (unbound) form, Lysine-33 (K33) and Aspartate-145 (D145) form an internal salt bridge; in the holo (ligand-bound) state, that bridge breaks and D145 often forms hydrogen bonds with the inhibitor (pmc.ncbi.nlm.nih.gov). This motion is "functionally relevant": many CDK2 inhibitors rely on engaging D145 for potency, and they can only bind when the K33–D145 bridge is disrupted (pmc.ncbi.nlm.nih.gov).

The YuelDesign study exploited this as a test. They compared how often each method's designed ligands could approach D145 closely enough (<3 Å) to break the salt bridge. The metrics were striking: about **half** of YuelDesign's ligands had a minimum noncarbon-atom distance to D145 under 3 Å, whereas DiffSBDD and PMDM rarely did (pmc.ncbi.nlm.nih.gov). In practical terms, YuelDesign produced many ligands that inserted polar groups into the pocket such that D145 reoriented into binding range.

DiffSBDD and PMDM keep the K33–D145 distance fixed, whereas YuelDesign permits side-chain rearrangements (Wang et al.).

In YuelDesign complexes, one can visually see that the model "broke" the salt bridge and let D145 swing toward the ligand's polar moieties (often forming H-bonds or π interactions) (pmc.ncbi.nlm.nih.gov). This effect closely mimics known CDK2 inhibitors (like those in PDB entries 2FVD, 1DI8, 1KE8, 1PYE) which all show D145 interacting with the bound drug (pmc.ncbi.nlm.nih.gov). In short, **only YuelDesign could generate ligands that resembled the active, D145-binding state of CDK2**. As a UVA news release summarized, *"when designing molecules for… CDK2, only YuelDesign could capture the critical structural changes that happen when a drug binds."* (med.virginia.edu).

This case study is highly illustrative: it shows that YuelDesign isn't merely rearranging for the sake of novelty, but is capturing real, functional pocket flexibility. It outperforms rigid methods exactly on a hard drug-design challenge. The success with CDK2 suggests YuelDesign holds promise for other targets where induced fit is key (e.g. many kinases, GPCRs, etc.).

4.6. Structural Evolution During Generation

Beyond final results, the authors also examined **how** molecules assemble through the diffusion process. By inspecting intermediate steps, they observed a logical sequence: atom types stabilize mostly in the later steps, while geometric refinement is more continuous (pmc.ncbi.nlm.nih.gov). Initially, atoms float with transient bond distances; mid-way through denoising (around step 25) bond lengths begin to fix, and by step 75 most chemistry is set (pmc.ncbi.nlm.nih.gov). This hierarchical evolution indicates the model first lays out a rough scaffold (forming general shape) and then goes back to tweak precise bond types and lengths. Such insights (Figure 6 in the paper) confirm that the co-diffusion of geometry and types is proceeding sensibly and could inform further model optimization (e.g. untangling heavy vs. light atom updates).

While detailed, these process observations reinforce confidence that YuelDesign can converge from noise to valid molecular complexes. In sum, the results section demonstrates that YuelDesign (on multiple metrics and case studies) outperforms prior diffusion-based design methods by leveraging protein flexibility.

5. Comparative Perspectives

YuelDesign represents a significant step, but it must be viewed in context of other advances and parallel approaches.

5.1. Rigid vs. Flexible Pocket Models

As discussed, most contemporary structure-based design methods assume a fixed pocket. DiffSBDD (Schneuing *et al.*, 2024) is a recent example that uses equivariant diffusion to place ligands in a pocket, but it **freezes** the pocket coordinates (pmc.ncbi.nlm.nih.gov). Similarly, other diffusion frameworks (e.g. DiffSMol (pmc.ncbi.nlm.nih.gov), DiffBP (pmc.ncbi.nlm.nih.gov)) do not alter the pocket. These are faster (lower dimensionality) but fail to capture induced-fit. In contrast, YuelDesign intentionally generates pocket atoms, allowing side chains to rotate or shift (pmc.ncbi.nlm.nih.gov).

Beyond diffusion, traditional docking tools like MedusaDock or AutoDock Vina incorporate limited flexibility by sampling side-chain rotamers (pmc.ncbi.nlm.nih.gov), but even these are far more constrained than YuelDesign's joint generation. Standard scoring methods (MedusaScore (pmc.ncbi.nlm.nih.gov), empirical force fields) assume the protein position. In fact, Dokholyan's own prior work on MedusaDock (2013, 2019) sought to improve docking by adding backbone flexibility, but it remained a local sampling process (pmc.ncbi.nlm.nih.gov). YuelDesign instead uses learned co-evolution, a paradigm shift.

A recent Nature Computational Science brief (Schneuing *et al.*, 2024) introduced an equivariant diffusion model for fixed-pocket drug design, reporting modest gains in docking scores and chemical quality over baselines. YuelDesign builds on this wave by directly addressing the missing flexibility dimension. As the authors note, rigidity imposes a strict barrier (the unbroken K33–D145 salt bridge in CDK2) that simply cannot be overcome without pocket motion (pmc.ncbi.nlm.nih.gov). Thus, YuelDesign widens the search space and captures more biology.

5.2. Other AI Generative Approaches

Outside diffusion, other AI-driven design tools exist. **Generative adversarial networks (GANs)** and **variational autoencoders (VAEs)** have been used to suggest molecules from learned latent spaces (pmc.ncbi.nlm.nih.gov), but these typically ignore 3D entirely. **Reinforcement learning** frameworks craft compounds to optimize scoring functions, but they often struggle with 3D geometry as input. **Language models** can propose SMILES strings of potential drugs, but again lack explicit structure context.

Notably, the Institute for Protein Design (Baker Lab) published *RFdiffusion* (2022-23) for de novo protein design – a diffusion model that generates protein backbones, followed by sequence design with ProteinMPNN in the reported workflow (Watson et al., 2023). RFdiffusion (Nature 2023) demonstrated the power of diffusion

in bio-molecular design, but it was focused on entire protein creation, not small molecules for binding pockets. Insilico Medicine (a company) and others have reported diffusion-based small-molecule generators, but details are often proprietary or 2D-based.

A recent review (Buvailo 2026) highlighted diffusion's promise in drug discovery, noting that *"diffusion models show promise in small molecule design"* by improving diversity and novelty (www.sciencedirect.com). Drug Discovery Today (Vol. 30, 2025) also features a keynote on "Unraveling the potential of diffusion models in small-molecule generation" (DruggDiscoveryToday 2025), suggesting broad academic interest. However, few of these specifically tackle pocket flexibility. The 2024 Nature Communications paper by Huang *et al.* describes **PMDM**, the comparator discussed above, as a conditional model for generating molecules that fit specified targets (Huang et al.). Related flexible-pocket approaches include **Apo2Mol**, which uses diffusion to generate ligand structures and corresponding holo pocket conformations from apo inputs (Apo2Mol), and **FlexSBDD**, which uses flow matching to model flexible protein–ligand complexes (FlexSBDD).

In industry, AI platforms like **Exscientia** and **Insilico** deploy various deep-learning tools for drug design, but details on how they handle pocket flexibility are often undisclosed. Amazon's new **Bio Discovery** platform (AWS) focuses on streamlining antibody and protein engineering workflows, emphasizing "lab-in-the-loop" for rapid screening (www.techradar.com). While AWS's tool is aimed at broader biological R&D, it underscores that companies see the value in integrating advanced models with wet labs. Similarly, DeepMind's recent AlphaFold 3 aims to model interactions, reflecting interest in bridging protein–ligand dynamics (time.com). YuelDesign fits into this landscape as a provider of *novel* small-molecule candidates that respect realistic binding modes, complementing risk-averse static approaches.

5.3. Advantages and Limitations

YuelDesign's **advantages** are clear from the data: improved molecular properties and docking for flexible targets (pmc.ncbi.nlm.nih.gov). It can co-create pocket-side changes, enabling designs that rigid methods miss (as in CDK2) (pmc.ncbi.nlm.nih.gov) (med.virginia.edu). The generated ligands exhibit similar functional- group diversity to real binders (pmc.ncbi.nlm.nih.gov), indicating it learns to produce chemically plausible drugs. The use of an equivariant backbone (E3former) is also novel and robust, adapting AlphaFold's innovations to ligand design (pmc.ncbi.nlm.nih.gov).

However, the authors candidly note **limitations**. One issue is **strained rings**: small 3- or 4-membered rings often appear in generated molecules more than in real drugs (pmc.ncbi.nlm.nih.gov). These rings are synthetically difficult and usually avoided, but diffusion noise sometimes creates them easily. This suggests a need to integrate hard chemical rules (e.g. disallow tiny cycles) into future models or use post-filters.

Another challenge is **scalability** for larger molecules. The current model handles up to ~400 atoms (including protein side chains), but as the number of atoms grows (either from large pockets or macrocyclic ligands), the joint diffusion becomes more difficult. The authors observed a drop in validity/connectivity for extremely large ligands (pmc.ncbi.nlm.nih.gov). They attribute this to the high dimensionality of simultaneous coordinate+type diffusion in a large system, which leads to slower convergence. One proposed remedy is a **latent diffusion** approach (pmc.ncbi.nlm.nih.gov): first encode the complex into a lower-dimensional latent space, then run diffusion, which could stabilize generation for big molecules.

Lastly, like most generative methods, YuelDesign's outputs still **differ from known ligands** in detailed chemistry. The Tanimoto similarities to native ligands were low (pmc.ncbi.nlm.nih.gov). This is not surprising – exploring vast chemical space rarely finds the exact native compound – but it means YuelDesign is better for *novel leads* than for tight mimicry of an existing binder. As a result, initial candidates still need downstream validation, synthesis, and optimization. The cost and feasibility of synthesizing the proposed ligands (especially with any exotic groups) remain practical concerns. That said, the comparatively lower SAS scores suggest YuelDesign does not exclusively propose ultra-complex chemicals (pmc.ncbi.nlm.nih.gov).

In summary, YuelDesign shows strong proof-of-concept. Future work will aim to refine chemical constraints, improve large-molecule handling, and integrate user feedback or task-specific objectives (e.g. potency). On the plus side, the framework is general: with additional data or conditioning, one could tailor molecules for specific targets, multi-target profiles, or ADMET properties. The dual-diffusion, equivariant approach itself is broadly applicable beyond the reported examples.

6. Discussion and Future Directions

6.1. Implications for Drug Discovery

The YuelDesign platform has several broad implications. By explicitly incorporating protein flexibility into AI-driven design, it addresses a long-standing gap in computational drug discovery. This may translate to higher success rates in the lead optimization stage: drugs that account for induced fit upfront should have better "out of computer" performance. If YuelDesign can be scaled and validated, pharmaceutical groups could use it to generate candidate libraries tailored to their targets' unique biology.

A key implication is faster iteration. Designing ligands while the pocket flexes during computation saves time otherwise spent on separate steps: one does not need to manually generate multiple pocket conformations or run expensive MD on each design. Moreover, YuelPocket's ability to find new pockets (even on predicted structures (med.virginia.edu)) means fresh targets without known binding sites can be tackled. This broadens the scope of structure-based design to novel targets.

YuelDesign's methods are also likely to synergize with other AI advances. For example, one could feed AlphaFold-predicted protein structures into YuelPocket/YuelDesign to craft drugs for proteins lacking experimental structures. Similarly, as AlphaFold3 and other models begin to predict protein–ligand complexes, YuelDesign could take those as starting points and propose chemical optimizations. The combination of generative design with high-throughput synthesis/assay (as AWS's Bio Discovery platform envisions (www.techradar.com)) could create a more integrated "design-build-test" loop.

In the research context, these results underline that *"protein pocket flexibility is not merely a structural detail but a critical determinant of successful molecular design"* (pmc.ncbi.nlm.nih.gov). Documenting this drives home that flexible-docking has real impact on output quality. King's Law (Koshland's induced-fit theory) is reaffirmed: the lock (protein) does move, and accounting for that movement matters for drug discovery (www.eurekalert.org) (pmc.ncbi.nlm.nih.gov). Future tool developers will likely consider this lesson: even in AI, some degree of target dynamics must be built in.

Finally, adopting flexible design models can influence target selection. Currently, many drug programs avoid "difficult" targets (e.g. shallow or highly mobile pockets) because rigid design fails them. With tools like YuelDesign, previously intractable targets (for neurological or "undruggable" proteins) might become tractable, expanding the druggable genome. As Dokholyan notes, patients with cancer or neurological disorders *"desperately need better drugs targeting these wiggly proteins"*, and approaches like YuelDesign aim to break through those historical roadblocks (www.eurekalert.org).

6.2. Future Work and Enhancements

Looking forward, several avenues suggest themselves:

- **Latent Diffusion and Scaling:** As mentioned, a latent-space approach could allow YuelDesign to handle larger systems (e.g. multi-domain pockets or bigger ligands) more stably (pmc.ncbi.nlm.nih.gov). This might involve training an autoencoder for protein–ligand complexes and diffusing in a compressed space.
- **Integration with Experimental Data:** The current model trains on static X-ray structures and generic drug-likeness. Future versions could incorporate activity or binding affinity data, performing **property-conditioned generation**. For instance, docking scores or experimental IC50s could be added as training guidance so the model not only fits pockets but aims for potency. Another idea is an active learning loop: generated molecules could be tested in virtual or wet assays, and the results fed back via fine-tuning.
- **Multiple Objectives and Selectivity:** Drug development requires optimizing multiple parameters (selectivity, ADME, toxicity). Extending YuelDesign to multi-objective generation (for example, conditioning on multiple proteins to avoid off-targets) would be valuable. The diffusion framework can likely be conditioned or modified to generate compounds targeting two pockets simultaneously or avoiding certain motifs.
- **Biophysical Realism:** Incorporating more physics into the model could reduce artifacts. For example, adding an explicit potential to penalize strained rings or unrealistic bond angles could partially address the small-ring issue (pmc.ncbi.nlm.nih.gov). Similarly, integrating solvent effects or induced-fit energetics into training might align the model's geometry changes with thermodynamics.
- **Broader Biological Context:** YuelDesign focuses on the binding pocket only. In reality, protein function can be allosteric and involve larger conformational shifts. Future research could generalize the idea to modeling full-protein conformations or allosteric sites. Another direction is protein–protein interactions: could a similar diffusion model design protein inhibitors (linear peptides) accounting for interface flexibility?
- **Software Deployment:** Finally, making YuelDesign accessible to the broader community will maximize impact. Providing an easy-to-use platform or API (perhaps hosted on UVA or cloud) would allow academic labs and smaller biotech to take advantage. Integration with existing drug-discovery suites (e.g. Schrödinger or OpenEye toolkits) could accelerate adoption. Ensuring robust validation and providing interpretability (e.g. visualizing how pockets change) will build trust in the method.

6.3. Considerations and Cautions

While promising, users should remember that **in silico design remains predictive**. Every generated ligand will still require experimental synthesis and testing. The high failure rate of drug candidates suggests that no computational method is foolproof. In particular, YuelDesign's tendency to create small strained rings highlights that chemists must vet suggestions. Also, the reported docking improvements are statistical; for any given target, it remains essential to sample multiple candidates and cross-check with orthogonal methods (MD simulations, free-energy calculations, etc.) before committing to synthesis.

Furthermore, AI models can sometimes encode biases from their training data. If Binding MOAD is enriched in certain target classes (e.g. enzymes, kinases) and chemistries (pyridines, purines, etc.), YuelDesign may reproduce those biases. It is encouraging that functional group distributions matched well overall (pmc.ncbi.nlm.nih.gov), but rare or novel scaffolds will need careful exploration. YuelDesign's source code is already available in the dokhlab repository and the hust220 repository, with data and code also deposited in Zenodo. These releases allow the field to test and build on the work, avoiding "black box" pitfalls.

Finally, computational expense and infrastructure are practical factors. Equivariant Transformers and diffusion models are heavy; running YuelDesign for many targets will require substantial GPU/CPU resources. However, cloud platforms (like the AWS Bio Discovery) are tailoring environments for exactly these workloads (www.techradar.com). Balancing the cost of computing vs. lab work will be an evolving consideration as adoption grows.

7. Conclusion

The YuelDesign framework represents a **paradigm shift** in structure-based drug design by unifying generative molecular design with protein flexibility. By treating the protein binding site as **dynamic**, not static, YuelDesign overcomes a key limitation of many past AI methods. In extensive benchmarks, it yielded more drug-like, diverse, and well-docked molecules than rigid-pocket diffusion models, and succeeded on challenging targets like CDK2 that require induced-fit binding (pmc.ncbi.nlm.nih.gov) (med.virginia.edu). This work uses diffusion to jointly model pocket structures and ligand conformations (pmc.ncbi.nlm.nih.gov).

UVA's Yuel design suite – also including YuelPocket and YuelBond – thus provides a compelling new toolset for AI-driven drug discovery (med.virginia.edu) (www.eurekalert.org). It builds on successes in deep learning and protein modeling to tackle a previously under-addressed problem: how to generate molecules *for moving targets*. The reported results are convincing: YuelDesign-generated compounds have high QED, maintain connectivity, and achieve binding scores on par with or better than native ligands (pmc.ncbi.nlm.nih.gov). Equally important, the method highlights that incorporating even modest pocket flexibility into generation can unlock novel design opportunities.

This development arrives at a time when AI in drug discovery is accelerating. As DeepMind's AlphaFold expands to model interactions (time.com) and industry platforms like AWS Bio Discovery promise integrated, "every-researcher" access to complex models (www.techradar.com), YuelDesign offers a way to integrate

dynamic biology directly into the earliest stages of design. It exemplifies the kind of innovation needed to drive down costs and failure rates, by producing **better-fitting keys for the locks that jostle inside our cells** (www.eurekalert.org).

Going forward, challenges remain (scalability, chemical rules, etc.), but the path is clear. We anticipate further research spurred by this work: refinements to diffusion architectures, tighter chemistry enforcement, and combinations with high-throughput lab testing. As Dokholyan and colleagues predict, such AI-driven tools *"could make a real difference for patients with cancer, neurological disorders and many other conditions where we desperately need better drugs"* (www.eurekalert.org). YuelDesign is an important step on that path, illustrating the power of flexible, AI-enhanced design in modern drug discovery.

References

All references are given inline:

- Wang *et al.*, *Science Advances* **12**(15), eaeb7045 (April 8, 2026), DOI: 10.1126/sciadv.aeb7045 (original YuelDesign publication) (pmc.ncbi.nlm.nih.gov).
- University of Virginia press releases and EurekAlert news (April 2026) (www.eurekalert.org) (med.virginia.edu).
- Schneuing *et al.*, "Structure-based drug design with equivariant diffusion models," *Nat. Comput. Sci.* **4**, 899–909 (2024) (pmc.ncbi.nlm.nih.gov).
- Brylinski & Skolnick, Proteins **70**, 363 (2008) (apo–holo RMSD analysis) (pmc.ncbi.nlm.nih.gov).
- Koshland, Angew. Chem. Int. Ed. Engl. **33**, 2375 (1995) ("lock and key" vs. "induced fit") (pmc.ncbi.nlm.nih.gov).
- Dokholyan lab publications on MedusaDock (2013, 2019) (pmc.ncbi.nlm.nih.gov).
- Cozzini *et al.*, J. Med. Chem. **51**, 6237 (2008) (review on target flexibility) (pmc.ncbi.nlm.nih.gov).
- Teague, *Nat. Rev. Drug Disc.* **2**, 527 (2003) (flexible targets in pharma) (www.nature.com).
- Amazon AWS blog and TechRadar (2026) on Bio Discovery (www.techradar.com).
- DeepMind AlphaFold3 announcement (2024) (time.com).

Note: References to figures (e.g. Fig. 2, 3, 4 in the Sci. Adv. paper) are paraphrased from those publications and cited accordingly above.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://www.eurekalert.org/news-releases/1122173
2. https://pmc.ncbi.nlm.nih.gov/articles/PMC13060587
3. https://pubmed.ncbi.nlm.nih.gov/41511953
4. https://med.virginia.edu/pharm/2026/04/02/new-ai-technology-to-speed-drug-development
5. https://go.bio.org/rs/490-EHZ-999/images/ClinicalDevelopmentSuccessRates2011_2020.pdf
6. https://www.nature.com/articles/nrd1129

7. https://pmc.ncbi.nlm.nih.gov/articles/PMC2701403
8. https://time.com/6975934/google-deepmind-alphafold-3-ai
9. https://www.nature.com/articles/s41467-024-46569-1
10. https://www.nature.com/articles/s41586-023-06415-8
11. https://www.sciencedirect.com/science/article/abs/pii/S1359644625001266
12. https://arxiv.org/abs/2511.14559
13. https://papers.neurips.cc/paper_files/paper/2024/file/60fb8cf8000f0386063fb24ead366330-Paper-Conference.pdf
14. https://www.techradar.com/pro/amazons-new-ai-bio-discovery-tool-can-provide-every-researcher-with-lab-in-the-loop-drug-discovery-40-ai-biology-models-can-filter-300-000-novel-antibody-candidates-down-to-the-top-results-for-testing-in-just-weeks
15. https://github.com/dokhlab/yuel_design
16. https://github.com/hust220/yuel_design
17. https://doi.org/10.5281/zenodo.17702010

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.