# Rhizome

Autonomous Atomistic Discovery.

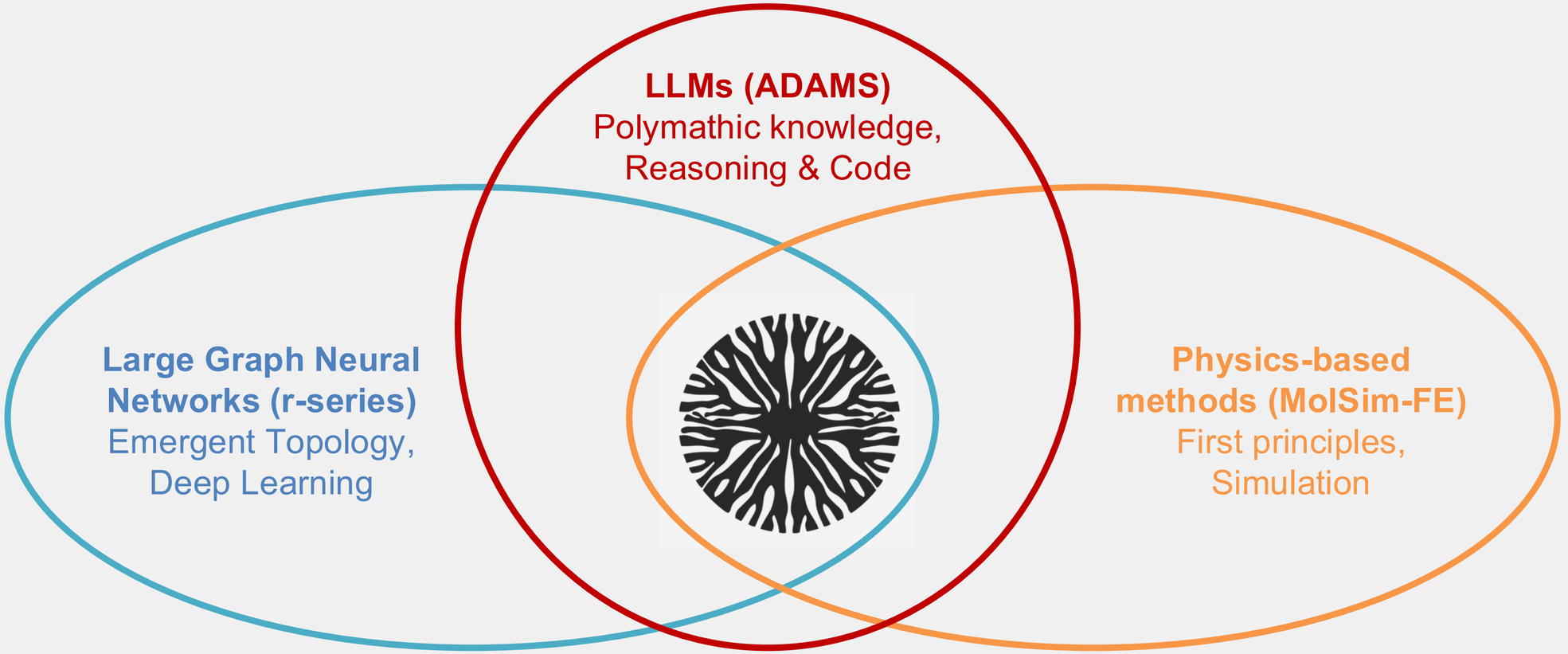
# Vision

Spearhead the convergence of

**Natural Language Intelligence** and

**Atomistic Intelligence**.

# Architecture of Autonomy



# Focus

First, we are hyper-focused on **Small Molecule Drug Discovery**.

# Focus

First, we are hyper-focused on **Small Molecule Drug Discovery**.

Small molecules are the **source code of modern medicine**.

If we can **engineer atoms to survive the human body**, we can **engineer them to power the world**.

# Steps to Autonomous Drug Discovery

- **Build the foundational primitives**:

  Develop frontier systems purpose-built for small molecule drug discovery.

- **Build agentic infrastructure:**

  Agents that harness these primitives to autonomously reason, plan, and adapt in a continual learning loop

# Steps to Autonomous Drug Discovery

- **Build the foundational primitives**:

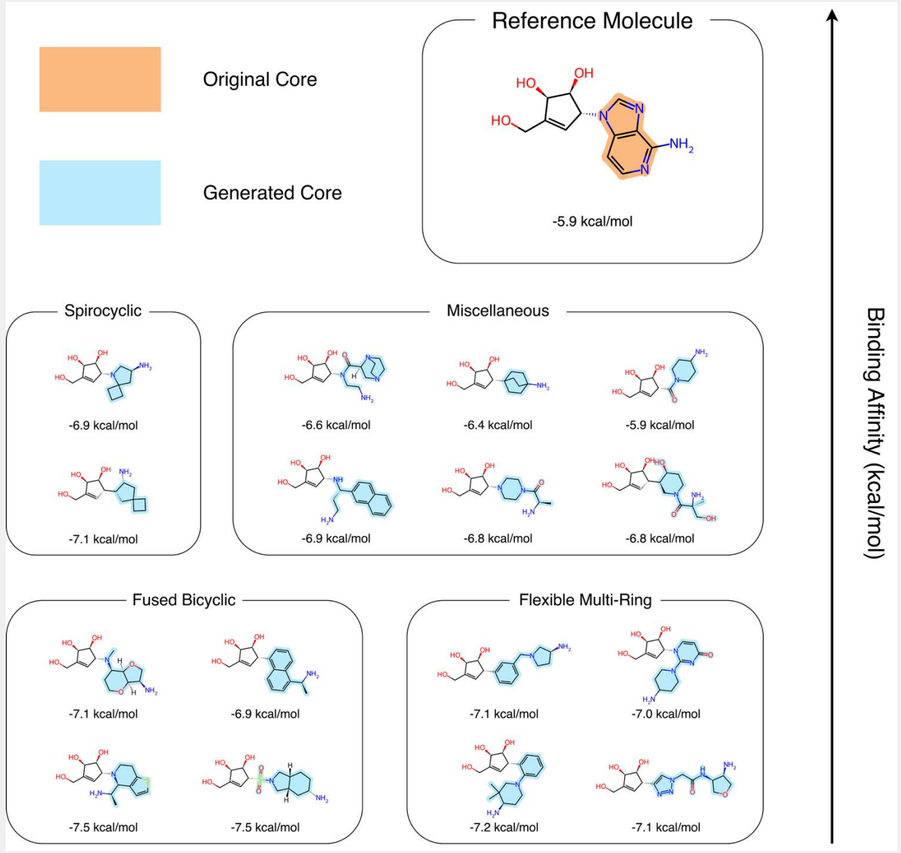
  Develop frontier systems purpose-built for small molecule drug discovery.

- **Build agentic infrastructure:**

  Agents that harness these primitives to autonomously reason, plan, and adapt in a continual learning loop

# Product Progress: r1, Foundational Model

## The GPT-1 Moment for Small Molecule Discovery is here.




The first model to learn the **topology of small molecules**

# Product Progress: r1, Foundational Model

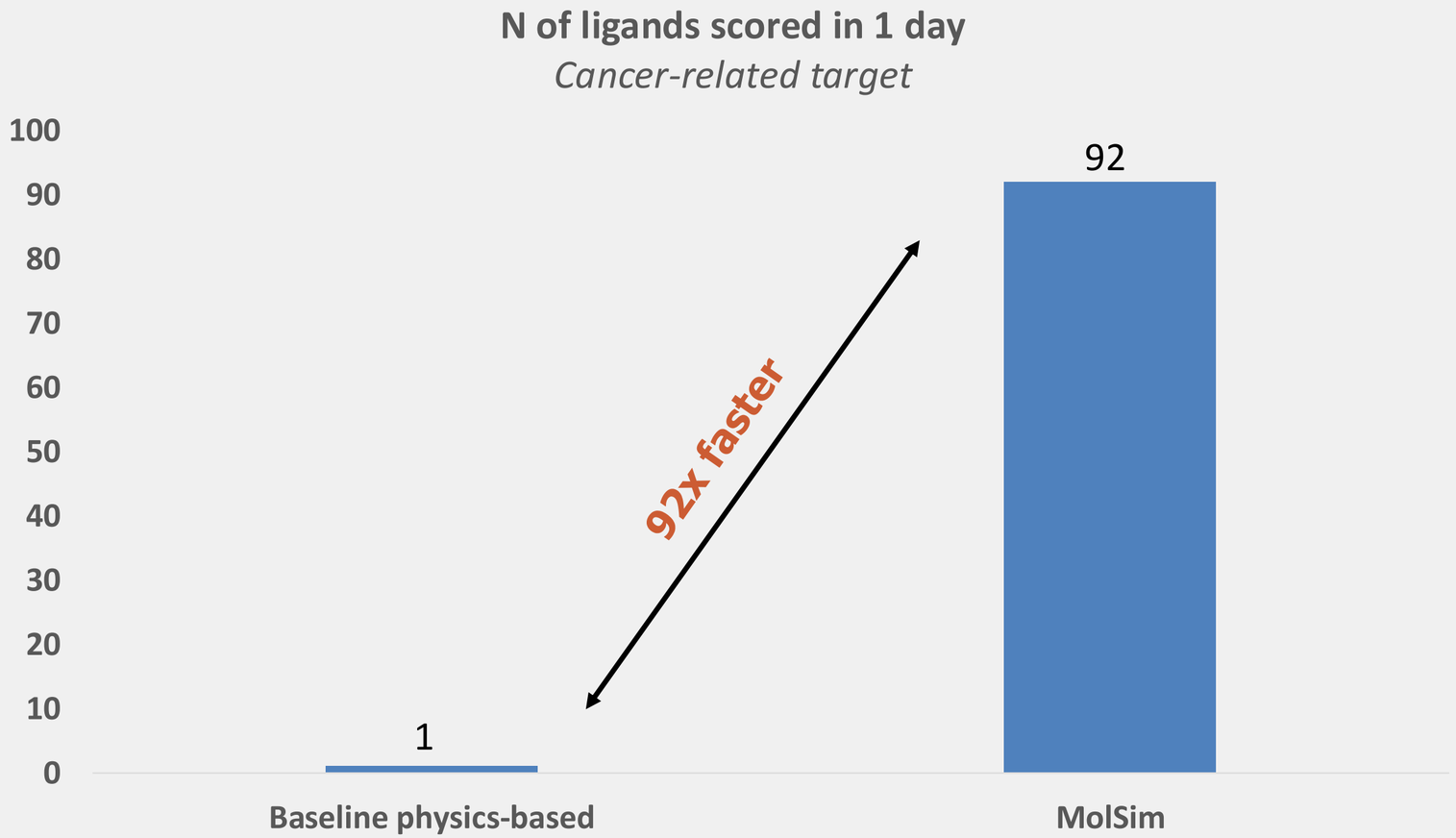
| Generative Model | Architecture | Fragment-based | Regional Generation | Control over Molecular Structure | Structural Diversity |
|---|---|---|---|---|---|
| **r1 (Rhizome)** | **GNN** | ✓ | ✓ | ✓ | ✓ |
| SynFlowNet (Valence Labs) | GNN | ✓ | Not Guaranteed | ✗ | ✓ |
| SAFE-GPT (Academia) | Transformer | ✓ | ✓ | ✗ | ✗ |
| GenMol (NVidia) | Transformer | ✓ | Not Guaranteed | ✗ | ✗ |
| Enki (Variational AI) | VAE | ✗ | ✗ | ✗ | Unknown |
| MoLFormer (IBM) | Transformer | ✗ | ✗ | ✗ | Unknown |
| COATI (Terray) | E(3)-GNN | ✗ | ✗ | ✗ | Unknown |
| Chemistry42 (Insilico) | GAN, LLMs | ✗ | ✗ | ✗ | Unknown |

# Product Progress: MolSim-FE

## MolSim-FE: Accelerated Physics Simulations with near wet-lab accuracy

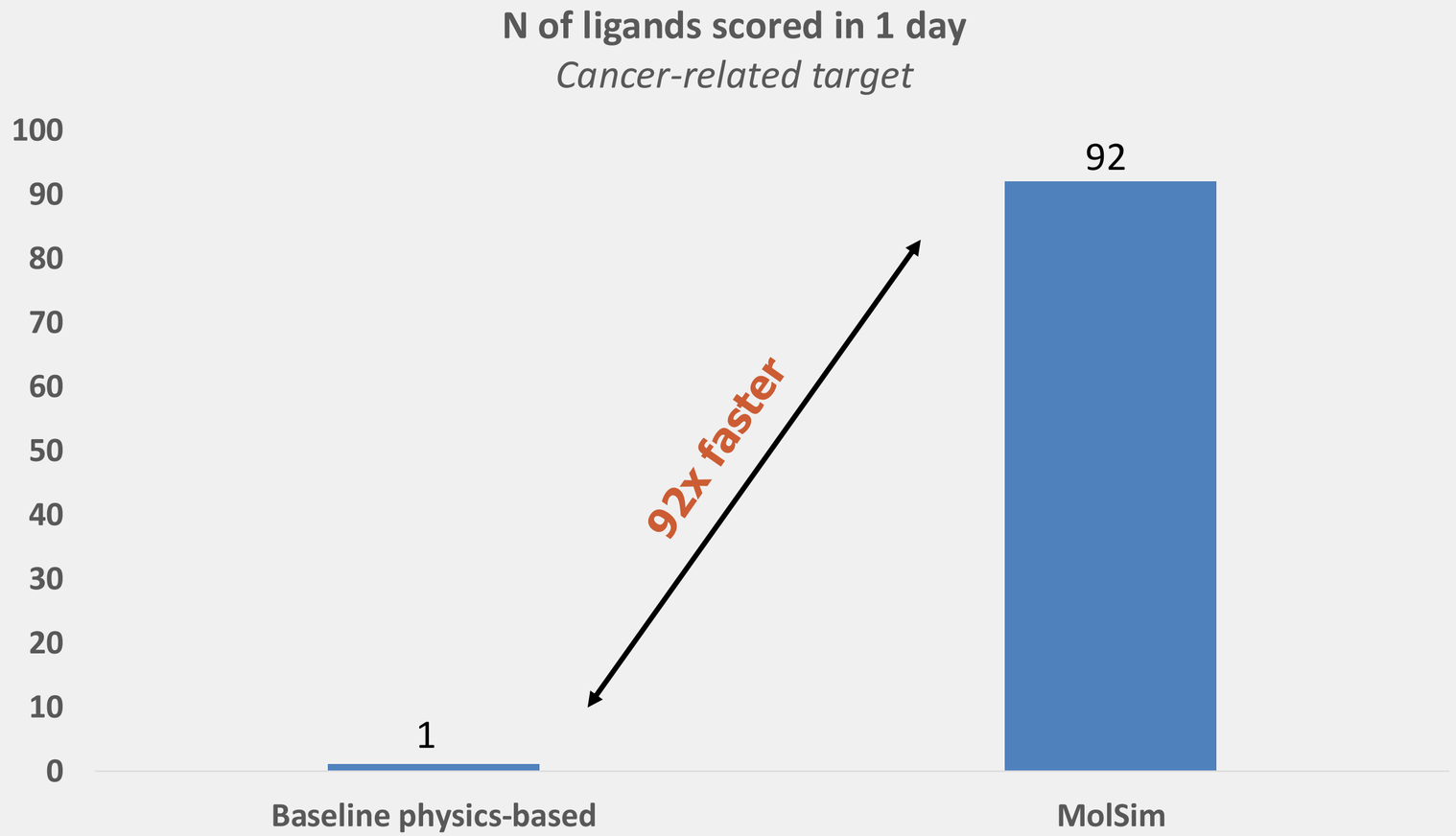
# Product Progress: MolSim-FE

## MolSim-FE: Accelerated Physics Simulations with near wet-lab accuracy



# Product Progress: MolSim-FE

## MolSim-FE: Accelerated Physics Simulations with near wet-lab accuracy




Further R&D will make it **1000x faster** than a standard pipeline.

# Steps to Autonomous Drug Discovery

- **Build the right low-level primitives**:
  Develop frontier AI systems purpose-built for small molecule drug discovery.

- **Build agentic infrastructure:**
  Agents that harness these primitives to autonomously reason, plan, and adapt in a continual learning loop.

# Product Progress: ADAMS

# Product Progress: ADAMS

## First end-to-end production autonomous run

## Coming 2026

# Next 12-months of growth

Grow drug development collaborations with biotechs and surpass $500K in ACV.

Secure a landmark drug discovery partnership with a global pharma leader.

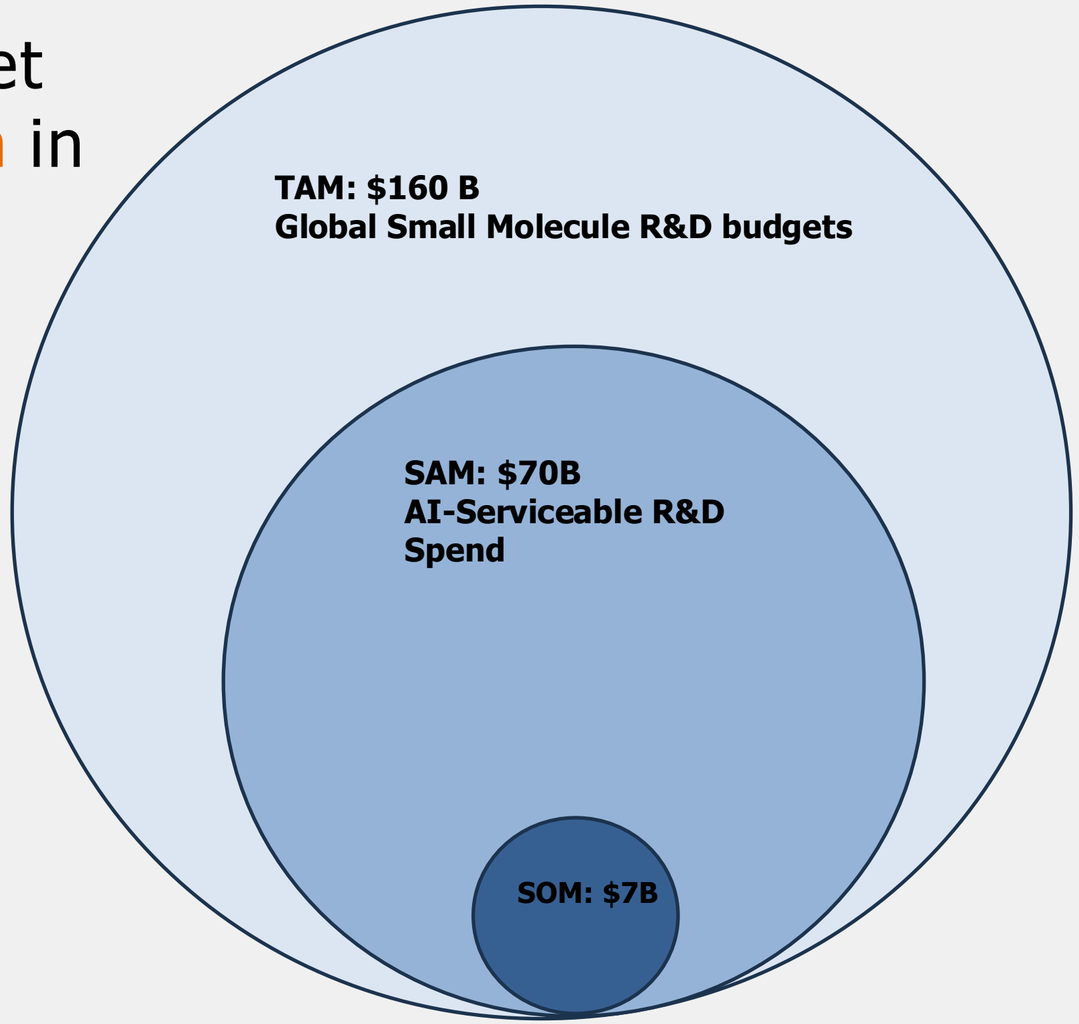
Secure multiple landmark grants with leading foundations and governmental institutions.

Launch successive generations of r-series of foundational models and MolSim-FE.

Discover first family of lead drug candidates through end-to-end Autonomous Drug Design.

# Market Funnel

Serviceable-Obtainable-Market (SOM): **$7 Billion** in annual revenue*




*See Appendix for more details.

**Sources: EvaluatePharma ’23, IQVIA ’23, JPM Global Pharma ’23, Deloitte R&D ROI ’23, BCG AI in Drug Discovery ’24, McKinsey AI Value in Biopharma ’23.**

# Founding Team

**Xhuliano, CEO**
**Founder**

**Yiwen (PhD), CSO**
**Head of Chemistry**

**Gregory (PhD), CTO**
**Head of Physics**

**Our team comes from the top institutions across AI and Chemistry.**

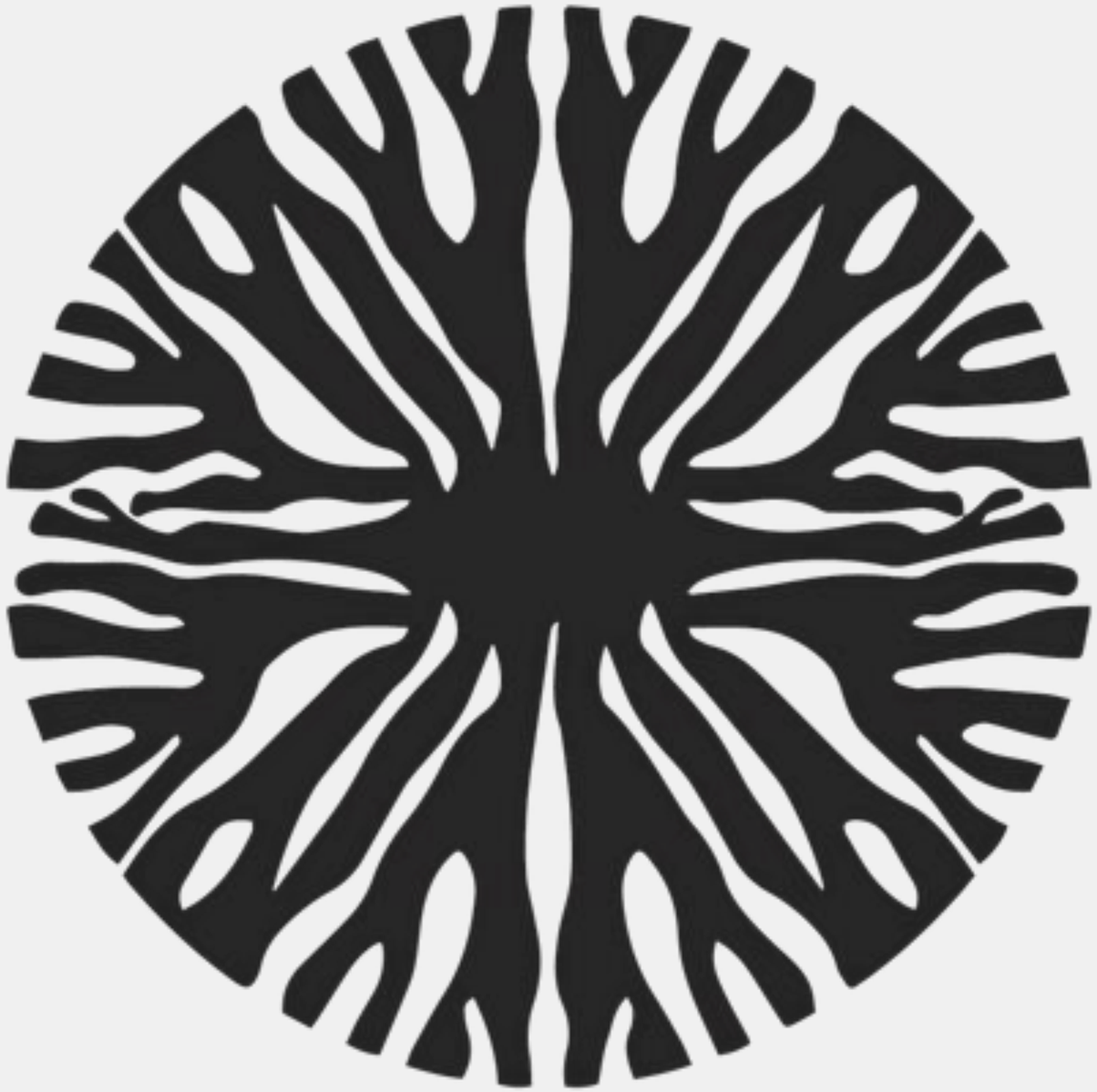
# Scientific Advisory Board

**John Proudfoot, PhD**

- Co-inventor of Nevirapine.
- Co-inventor of 50+ patents and patent applications.
- Director in Medical Chemistry at Boehringer Ingelheim responsible for implementation of new technologies such as fragment-based lead identification.

**Mark Esposito, PhD**

- Founder, CSO Kayothera
- Received his Ph.D. in molecular biology from Princeton while working in the laboratory of Yibin Kang.
- Kayothera has 3 preclinical candidates.



Autonomous Atomistic Discovery.

# Appendix: Supplementary Information

Eroom’s Law: The inflation-adjusted **cost of developing a new drug doubles every nine years**.

---

Patent Cliffs: **$200B+** in **revenue at risk** (2024-2030).

---

Target Drought: **90%** of human proteins remain "undruggable".

---

Late-Stage Failures: **60% fail** in Phase 2/3 after **$100M+ investment**.

---

# Appendix: Why now?

The explosion in NL intelligence + our tools will render most computational chemists' workflows obsolete.

No one has applied our novel model architectures and AI solutions at scale.

Remarkable progress in compute, algos and data in last 2 years make most AIDD pre-2022 moot.

Pharmaceutical labs being brought back onshore pose a huge increase in lab and labor costs.

Large % of world class deep-tech talent from academic & national labs will migrate to industry.

# Appendix: FAQs

- **Q: Difference with others in the autonomous scientist space?:**
  - A: Most are too general. These solutions try to solve everything and at the end will solve nothing. E.g. “Autonomous Scientists for Biology”. Their “advanced AI” is running a deep reasoning, “o-N” model, from OpenAI that sifts through papers. This is not the way to creating new drugs and curing disease. Our work is hyper specific. We will automate away all the in-silico work currently done by humans for small molecule drug discovery.

- **Q: More specifically, what about companies such as FutureHouse or XtalPi?:**
  - A: FutureHouse is a non-profit that takes a top-down approach to this problem. Our approach is a bottoms-up approach. LLM scaffolding and orchestration come secondary to developing frontier AI models that solve problems that general purpose LLMs will never be able to do such as molecular generation, molecular simulation, etc.
  - A: As it relates to XtalPi, we believe we are the West’s answer to this question.

# Appendix: FAQs continued

- **Q: Why would your team win?:**
  - A: Our team brings together world-class interdisciplinary expertise across engineering, artificial intelligence, pharmacology, and chemistry. We move with speed and clarity, guided by a bold vision that sets us apart from the competition.

- **Q: Why is wet-lab validation needed?:**
  - A: Until we have perfect simulations, we will need to rigorously test our models in the real world. This will develop a reputation and trust, so we become the de facto partner for $n^{th}+1$ drug development program.

- **Q: Why focus on just in-silico, can the wet lab be automated?:**
  - A: A ChatGPT moment in physical intelligence will need to occur. We believe this is 3-10 years away and requires billions in resources. When it occurs, we will immediately incorporate advanced humanoid robotics in wet-lab experiments and create the first end to end pipeline to go from target to Investigational New Drug (IND) in <3 months.

# Appendix: FAQs continued

- **Q: Why stop at being a foundry for bringing drugs to IND?:**
  - A: *We get the drug to orbit; drug developers flies the mission to market.* Phase II/III are administratively expensive and would take away from our focus which is perfecting the drug design pipeline.

- **Q: Why can’t LLMs understand molecules?:**
  - A: LLMs model the underlying latent structure of language while our models understand the underlying latent topology of molecular structures.

- **Q: How about materials?:**
  - A: Once we solve drugs, materials will be a translation task.

# Appendix: FAQs continued

- **Q: How did you get to a $7B SOM?:**
  - Top-Down Analysis
    - → 30 % of discovery dollars are already externalized (CROs, CDMOs, SaaS)
      - → $70B × 0.30 = $21B.
    - → A category winner typically captures ~⅓ of the outsourced slice.
      - → $21 B × 0.33 ≈ $7 B SOM.

  - Bottom-Up Analysis
    - 20 Big Pharma (top-20 by R&D) with $1B avg. annual discovery + pre-clinical budgets.
    - ~230 mid/late-stage biotechs with $90M annual discovery + pre-clinical budgets.
    - ~370 early-stage biotechs with $25M annual discovery + pre-clinical budgets.
    - ~1,100 micro-biotechs with $18M annual discovery + pre-clinical budgets.
      - → $70B in SAM.
      - → 50% of this can be addressed by our workflows, leading to $35B in addressable workflows.
        - 7-year adoption target of 45% of the 1700+ logos and 45% in wallet share.
        - → $35B x .45 x .45 ≈ $7B SOM.

# Appendix: MolSim case study

Cancer-related target:

| Protocol | Number of true binders in 10% best scored | Time to score with 1 GPU |
|---|---|---|
| Docking | 1/20 | Seconds |
| **MolSim-FE (Rhizome)** | **12/20** | **Minutes** |
| Baseline Physics (Academia/Industry) | 17/20 | Days |

**Our method lies on the pareto frontier of speed vs accuracy.**

In 6-months, we will push the pareto frontier to be as fast as docking and as accurate as StatMech.

# Appendix: Fragment-constrained Drug Design

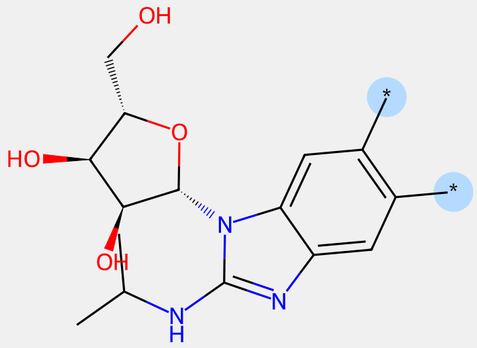
In real-world drug discovery scenarios, fragment-constrained drug design is preferred over blind de novo design because it limits the vast chemical space to synthetically and biologically plausible regions, enabling efficient target-specific property optimization.

- **Preservation of Validated Scaffold.** During hit-to-lead and lead optimization stages, molecular scaffolds with proven biological activity need to be preserved (identified during the initial hit discovery range). These stages focus on multi-property optimization without reducing the potency of the drug candidate.
- **Regional Modifications.** With structure-activity relationships established via high-throughput experimentation, virtual screening and active learning, specific regions of the molecular candidates are usually identified for further modifications.
- **Synthetic Accessibility.** Modifications of the synthetic route are also greatly constrained if the generated molecules only exhibit limited changes in pre-determined regions. Since the synthetic route of the initial hit molecule is already known, only a small number of individual steps require change.
- **Risk Aversion and Efficiency.** Exploration within a limited search space is much more efficient and lower-risk than pure de novo design, as problematic molecules (higher toxicity, etc.) can be easily avoided.

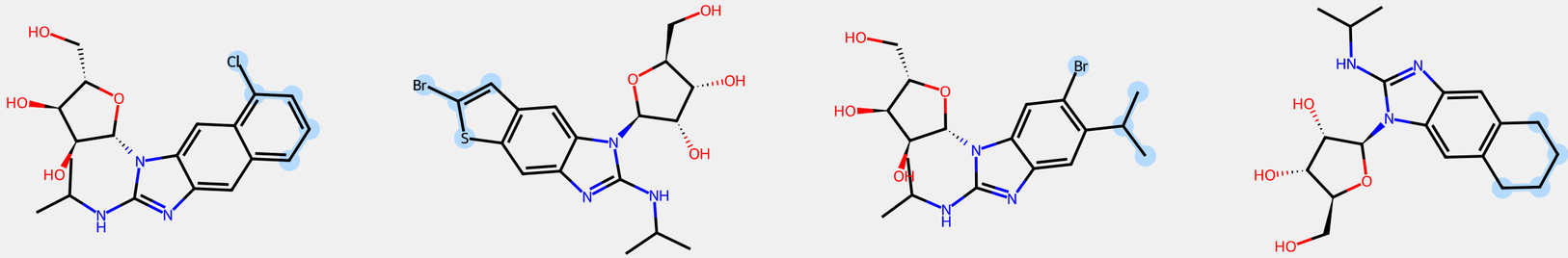
However, traditional sequence-based models (VAEs, RNNs and Transformers) do not offer control over molecular structures/fragments because molecules are not sequences by nature. **Our foundational model does all of this and more.**

# Appendix: Examples

Input



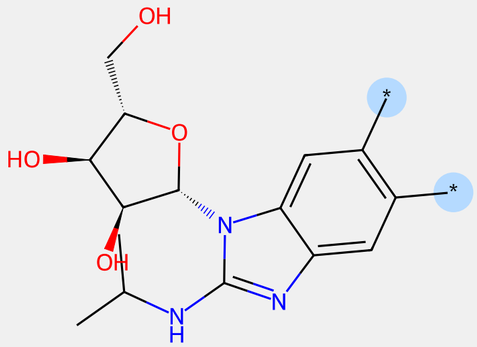

**Rhizome's r1**



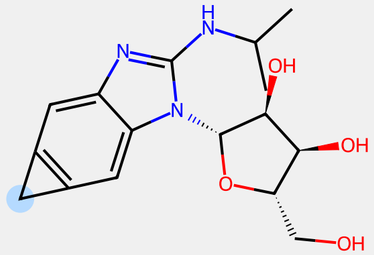

- Structural diversity. Generates linear chains, aliphatic rings, aromatic rings and aliphatic rings with heteroatoms.
- Total control over the size of the fragment to be generated.
- No unrealistic substructures. Handles aromaticity well.

# Appendix: Examples

Input



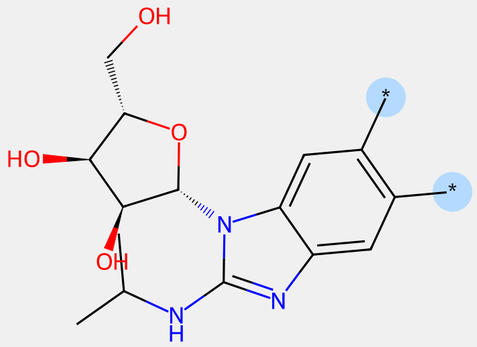

GenMol



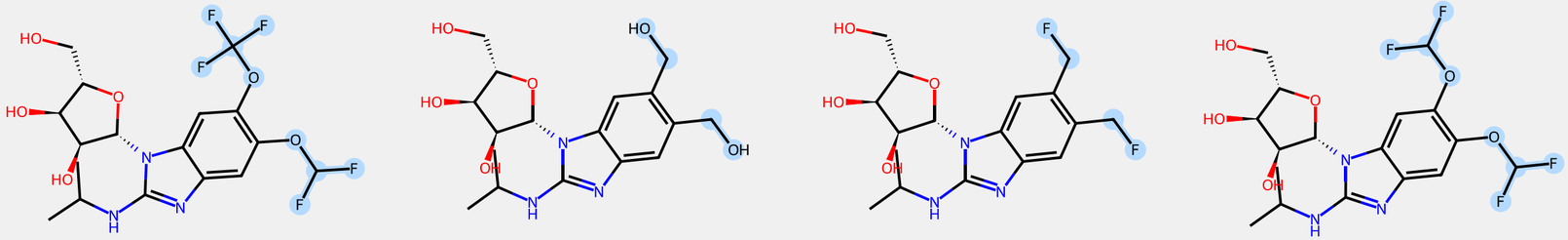

- Simply does not work. Only generate the same structure over and over again (even with multiple samples/beams).
- Clear mistake. A completely unrealistic bicyclic structure.

# Appendix: Examples

Input



SAFE-GPT



- Only generates linear chains and tends to generate symmetric ones.
- No structural diversity, strong biases leaning towards oxygens and fluorines. Does not generate novel substructures, does not generate ring structures at all.
- No control over the size of the fragment.