

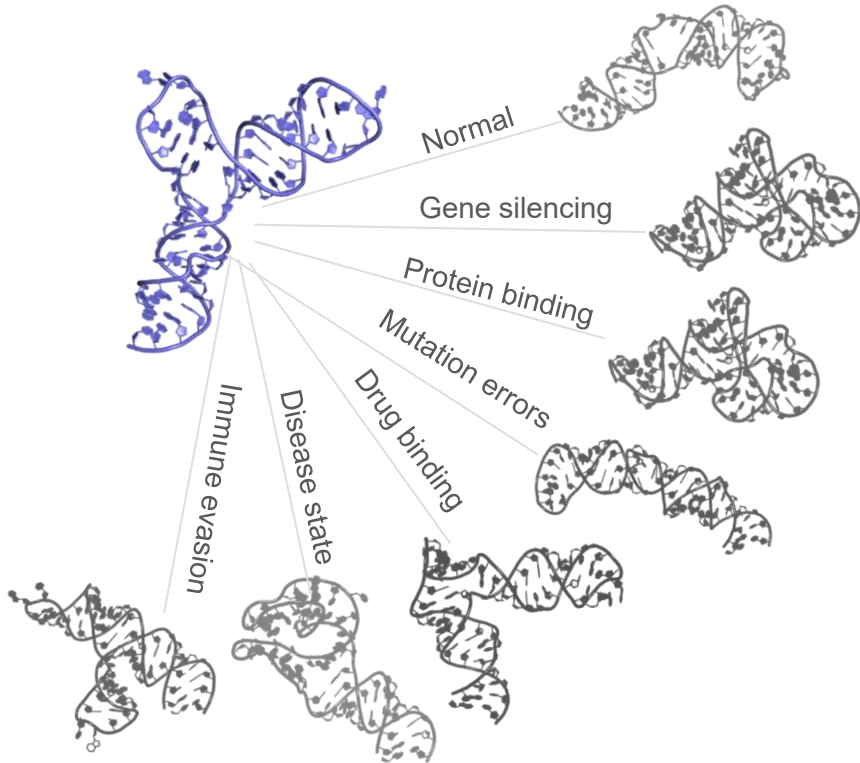
base4

D Y N A M I C S

The Foundational Data Layer for BioAI

No AI Model Can Be Predictive of Biology and Outcomes Without Capturing Dynamics

This is a 0 or 1 problem at the atomic level that no one else is solving



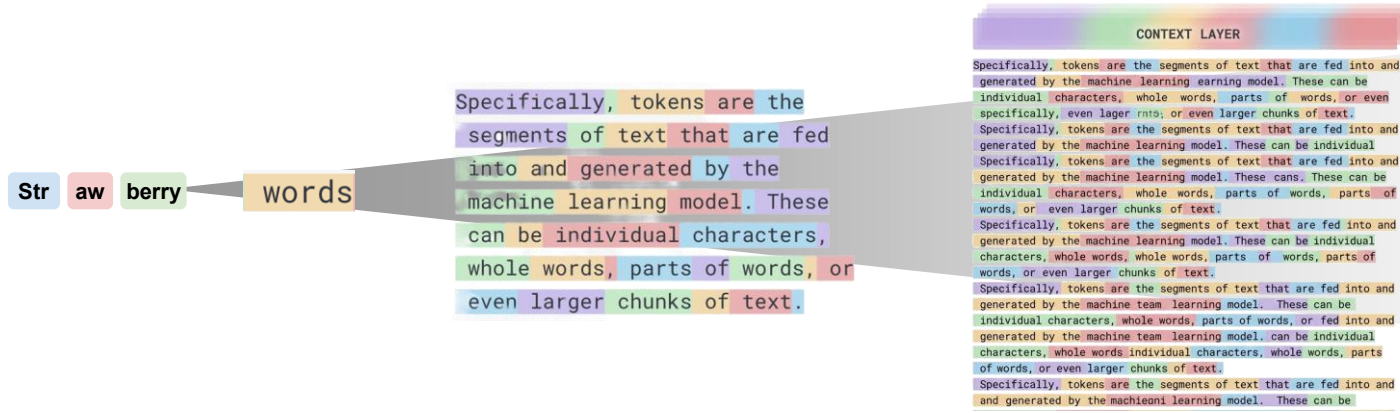
SHAPE DETERMINES FUNCTION

“In reality, all macromolecules dynamically alternate between conformational states to carry out their biological functions.”

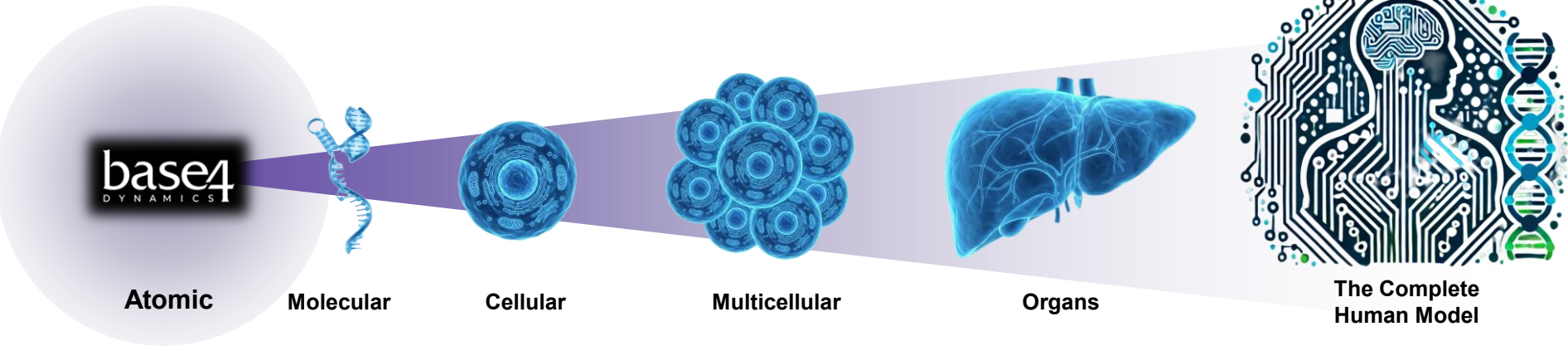
— Ganser, Kelly, Herschlag & Al-Hashimi,
Nature Reviews Molecular Cell Biology, 2019
Cited 600+ times

Base4 is Uniquely Scaling the Atomic Data that Unlocks AI for Biology

Without atomic insights, BioAI lacks the fundamental causation data to power its models



-  Gemini
-  Claude
-  Grok
-  GPT-4

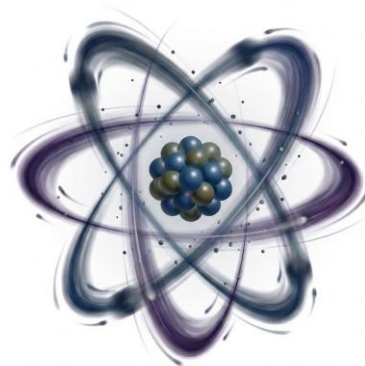


Introducing Atomic Tokens: The Fundamental Units of the Molecular Language

Biology is dynamic. BioAI must be too.

Str **aw** **berry**

Words/Fragments
Units of meaning

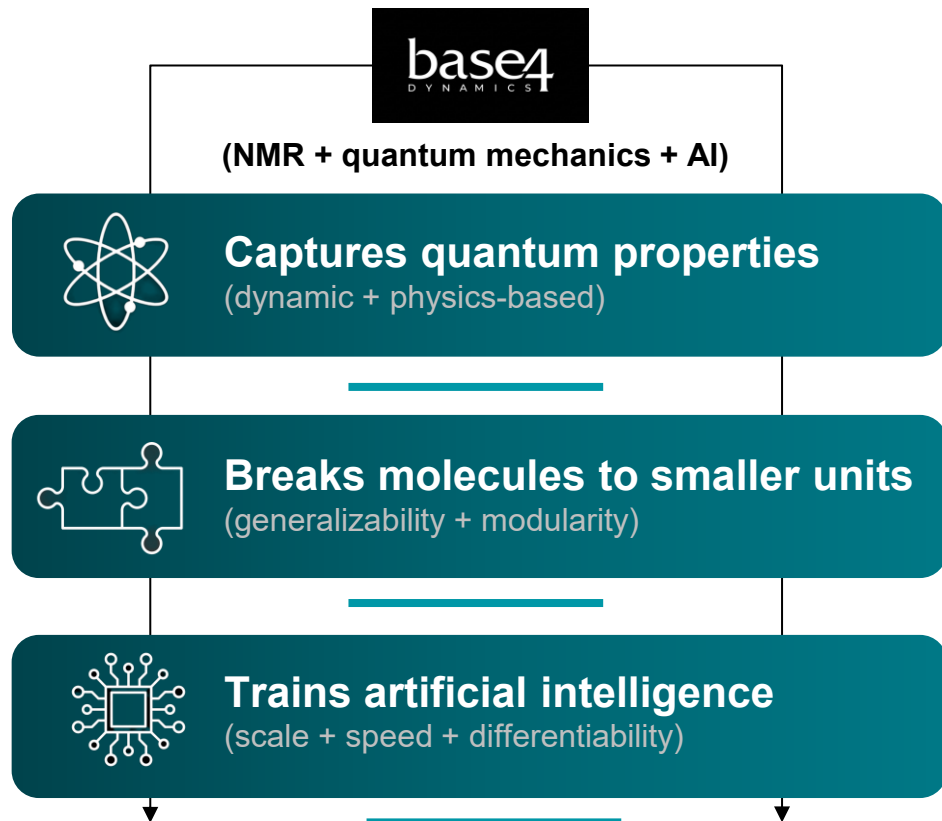


Atomic Tokens
Units of function

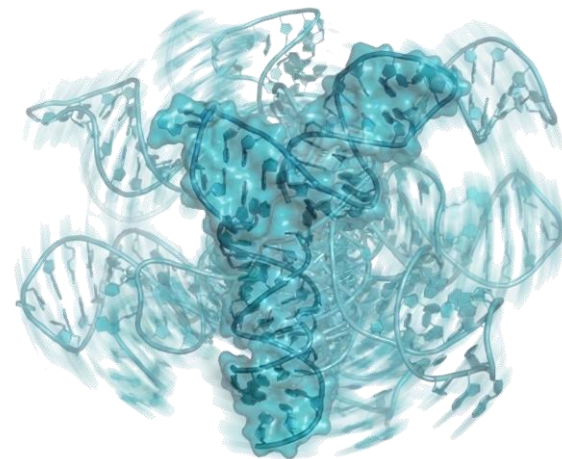
Atomic Tokens: Base4's ground truth data that captures how each atom moves and changes properties through time

Our Proprietary Platform Generates Atomic Data at Scale to Resolve Dynamics

Built over 15 years, now scaled to power predictive BioAI



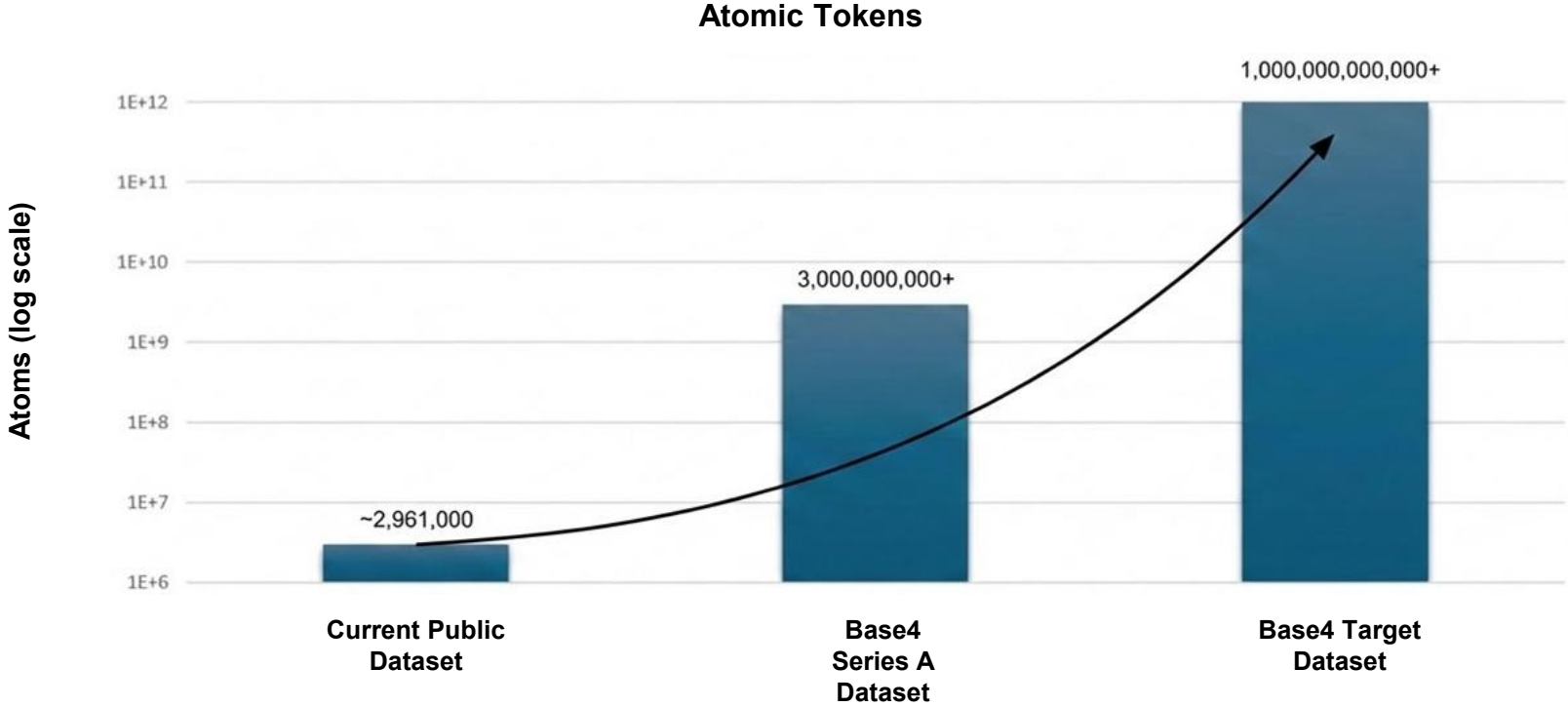
Atomic Dynamics Data



In 15 years, no one else has been able to achieve accurate dynamics data generation

Scalable Production Will Expand Available Atomic Behavior Data by Over 300,000x

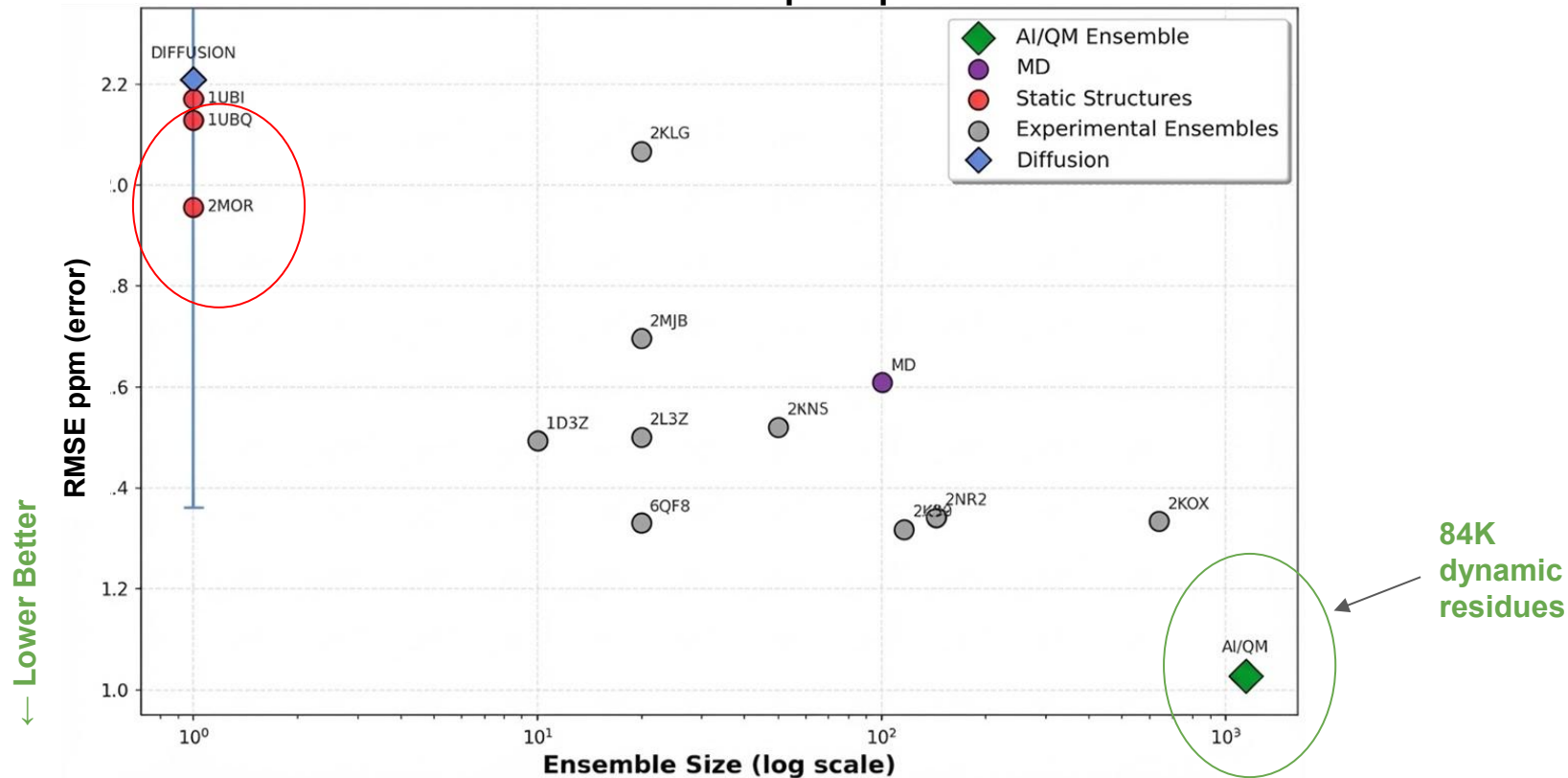
For comparison, LLMs took off when tokens went from billions to trillions



We Dramatically Beats Static Structures and Experimental Ensembles

In addition to better accuracy, our data enables >20,000x speedup and differentiable modeling

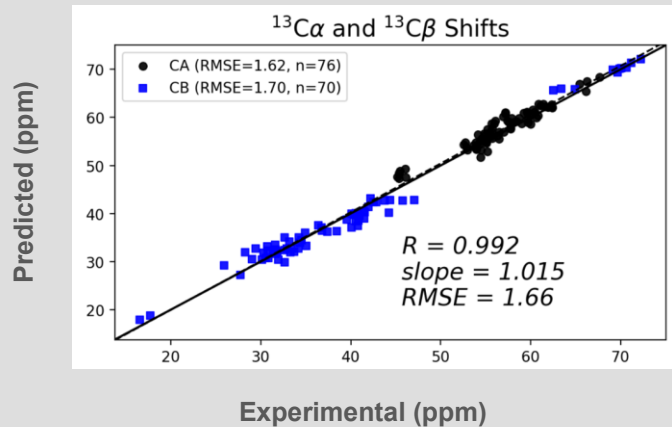
Demonstration on ubiquitin protein



We Dramatically Beats Static Structures and Experimental Ensembles

In addition to better accuracy, our data enables >20,000x speedup and differentiable modeling

Classical Quantum Mechanics



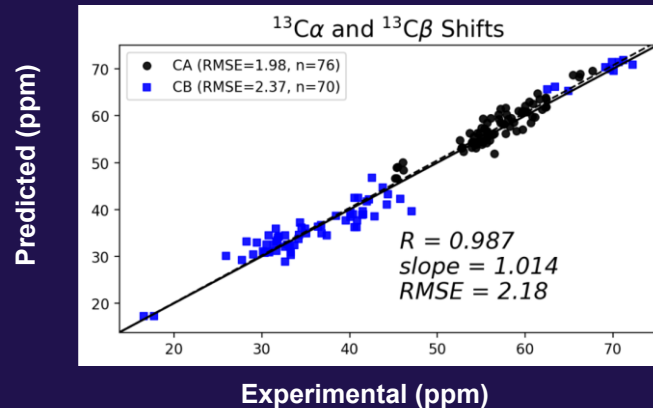
One iteration of one billion
atomic tokens requires

~75 years

Achieves
speedup
while
preserving
accuracy



Base4



One iteration of one billion
atomic tokens requires

~1 day

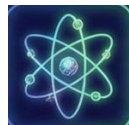
A Proprietary Physics-Grounded Data Moat for Molecular Behavior

An exclusive, large-scale dataset unavailable anywhere else



PROPRIETARY

Novel dynamics data



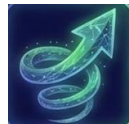
DYNAMICS + PHYSICS GROUNDED

Atomic level foundation



COMBINATORIAL

Fragmentation enables combinatorial tokens



COMPOUNDING

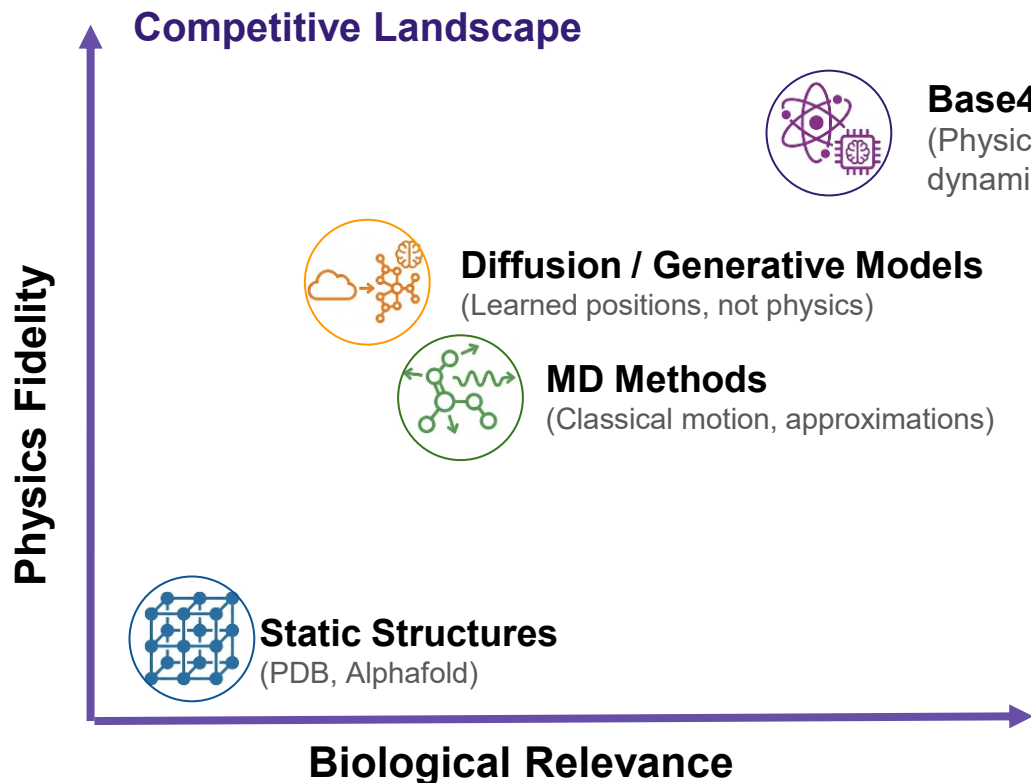
Modularity allows for exponential growth

Biology depends on how molecules behave

We are generating the only molecular level data that can be the training foundation for AI models of living systems

Moving From Static Snapshots and Approximations to **Atomic and Dynamic**

The only physics-grounded data engine capturing molecular behavior at scale



QM data is high content per data point. It can factor in change, novelty, and quantum phenomena, not just generalized motion.

NO ONE IS PROVIDING THIS DATA

Data Platform Drives Expanding Revenue Model as We Move Up the BioAI Stack

A data infrastructure company foundational to BioAI

Step 1: Dynamics Data Licensing

- Broad license for foundation models (Deepmind, Isomorphic, Xaira, Lilly...)
- Target-focused for pharma (GPCR, protein-protein, kinase conformation...)
- Interaction / chemistry focused (medicinal chemistry, enzyme engineering, covalent drug developers...)

Step 2: Dynamic Structural Models

- Platform access, API (by seat, usage, or enterprise)
- Drug discovery / platform deal tied to specific programs

Step 3: Dynamics Beyond Molecular Structure

Fueling

BioAI foundation modelers

Pharma drug discovery

Academic research

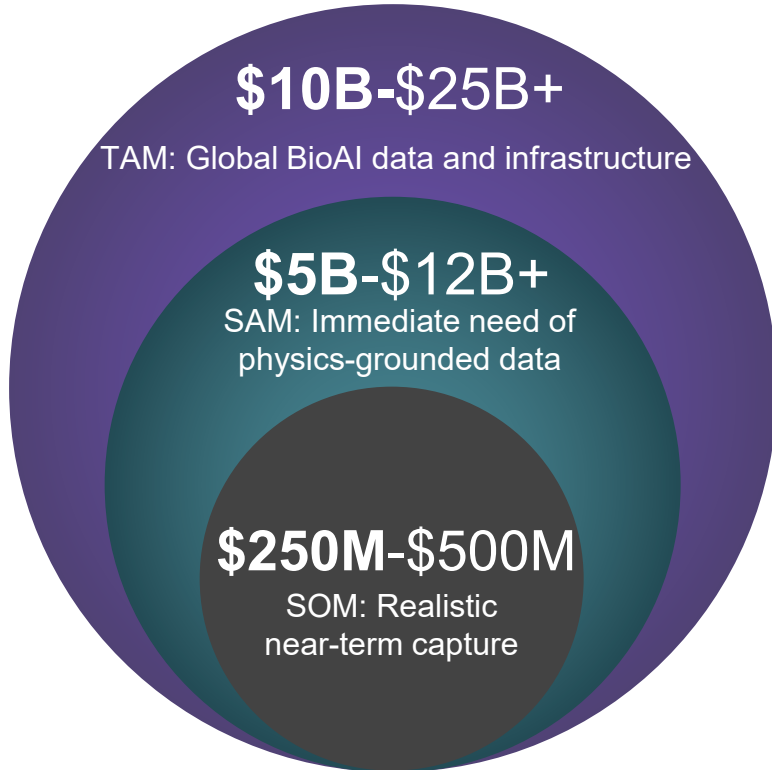
Industrial biotechs

Protein engineering

Synthetic biology

We Have Already Proven Out The Market Demand

High-value foundational data is needed today and will reshape future industry standards



Science is thirsting for dynamics data

15+ years partnering with pharma in structure-based drug discovery

Multi-million dollar collaborations leveraging small scale dynamics data with hit-ID wrapper

Customer discovery interviews with Deepmind, CZI/Biohub, Insilico Medicine + other AI-biotechs

Lack of foundational data prohibit effective incorporation of molecular level explanation for higher order observations

We Have Partnered Bespoke Atomic Level Dynamics Data for Over 15 Years

This level of data is now superlinearly scaleable to power the entire industry



Prior multi-million dollar partnerships validated early demand even when the data was not yet scaleable and was primarily for one-off projects

Now it's ready for scale and to enable all BioAI

A Team Built to Bridge Biology, Dynamics, and AI

Entrepreneurs and scientists with decades of experience in dynamics and company building



JOSHUA FAIRBANK
CEO

- Cofounder - Akebia, IPO
- Cofounder - Aerpio, IPO
- Adjunct Professor, Deep Tech - UChicago
- MBA - Chicago Booth



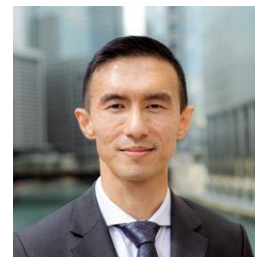
HASHIM AL-HASHIMI, PHD
Platform Architect

- Vagelos Professor of Biochemistry and Molecular Biophysics - Columbia University
- PhD (Biophysical Chemistry) - Yale
- Citations: 12094
- h-index: 64



ISAAC KIMSEY, PHD
Engine & AI

- AI development lead
- Co-lead multiple R&D partnerships
- PhD (Biochemistry) - Duke University
- 2x Nature author on RNA dynamics



YANG ZHENG
BD & Operations

- Cofounder, COO - MicroQuin (oncology)
- Cofounder, COO - Oxalo Therapeutics (renal)
- Secured multiple pharma partnerships
- MBA - Chicago Booth

Scaling The World's First Atomic Dynamics Dataset and Expanding Our Engine

Raising \$2.5M now to scale data production and enable larger round by year-end

Base4 Dynamics

\$2.5M Round



Personnel



Compute



Equipment



NMR costs

Protein dynamics dataset

RNA scalable proof of concept

Production-ready data platform

Series A-ready to scale models and commercial adoption



Thank you