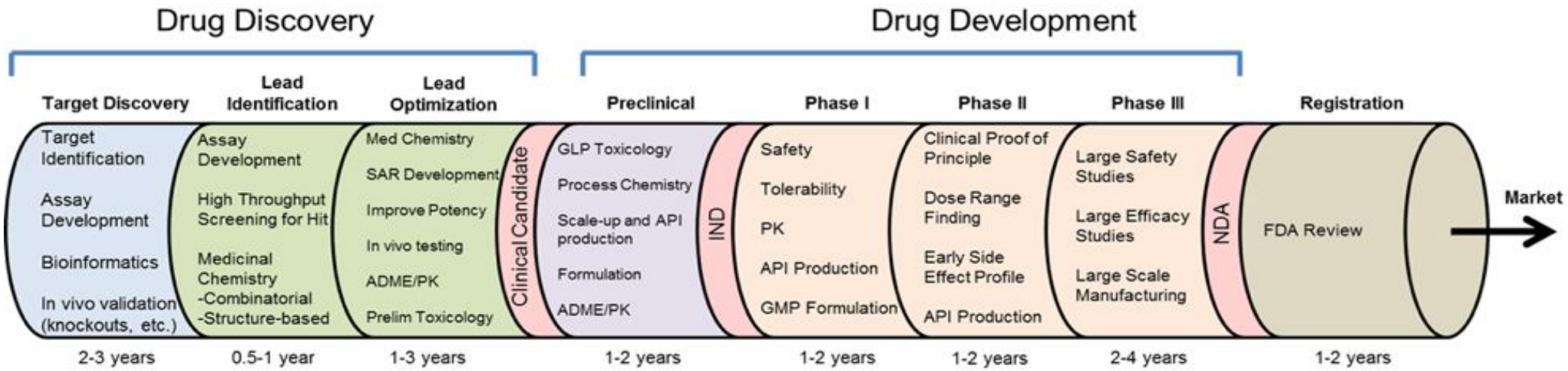
Schrödinger

New Dimensions: How Computation is Accelerating Structure-Based Drug Design

National Academies of Sciences, Engineering, and Medicine

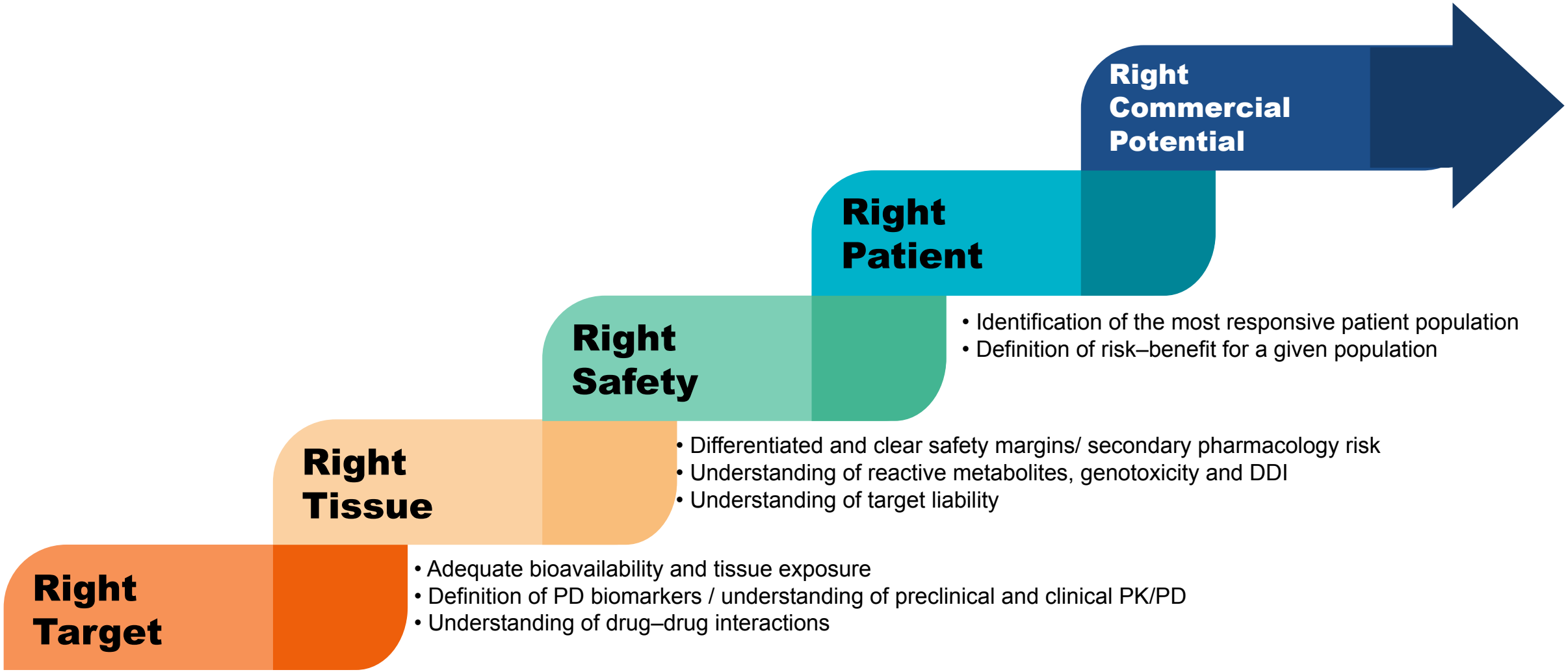
February 2023

Drug Discovery and Development



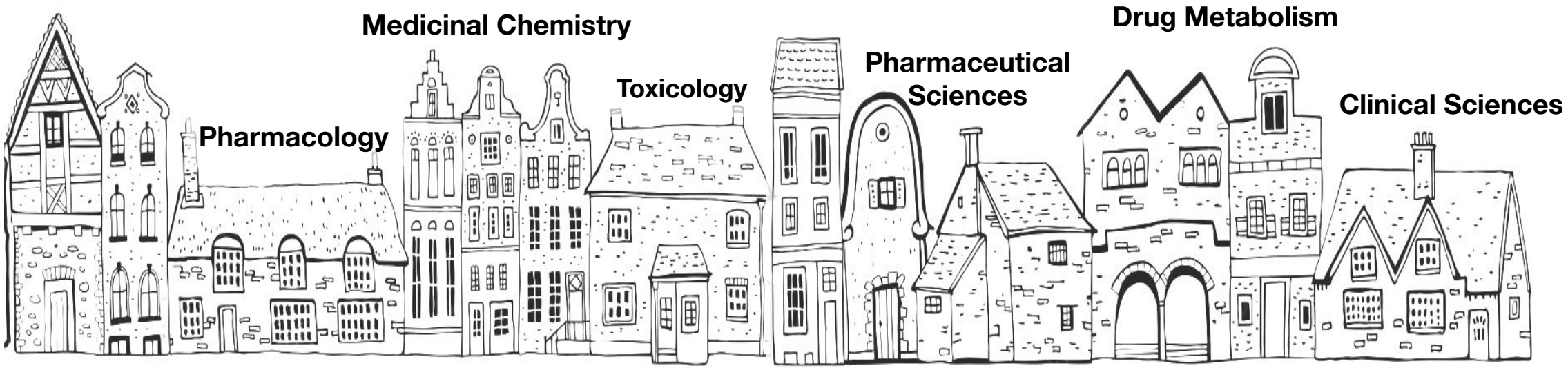
- The process takes between 10-17 years for each drug (patent life is 20 years from date of grant)
- Requires the synthesis of ~10,000 molecules to get 10 molecules in clinical trials, and 1 new drug

5R Framework for Drug Development

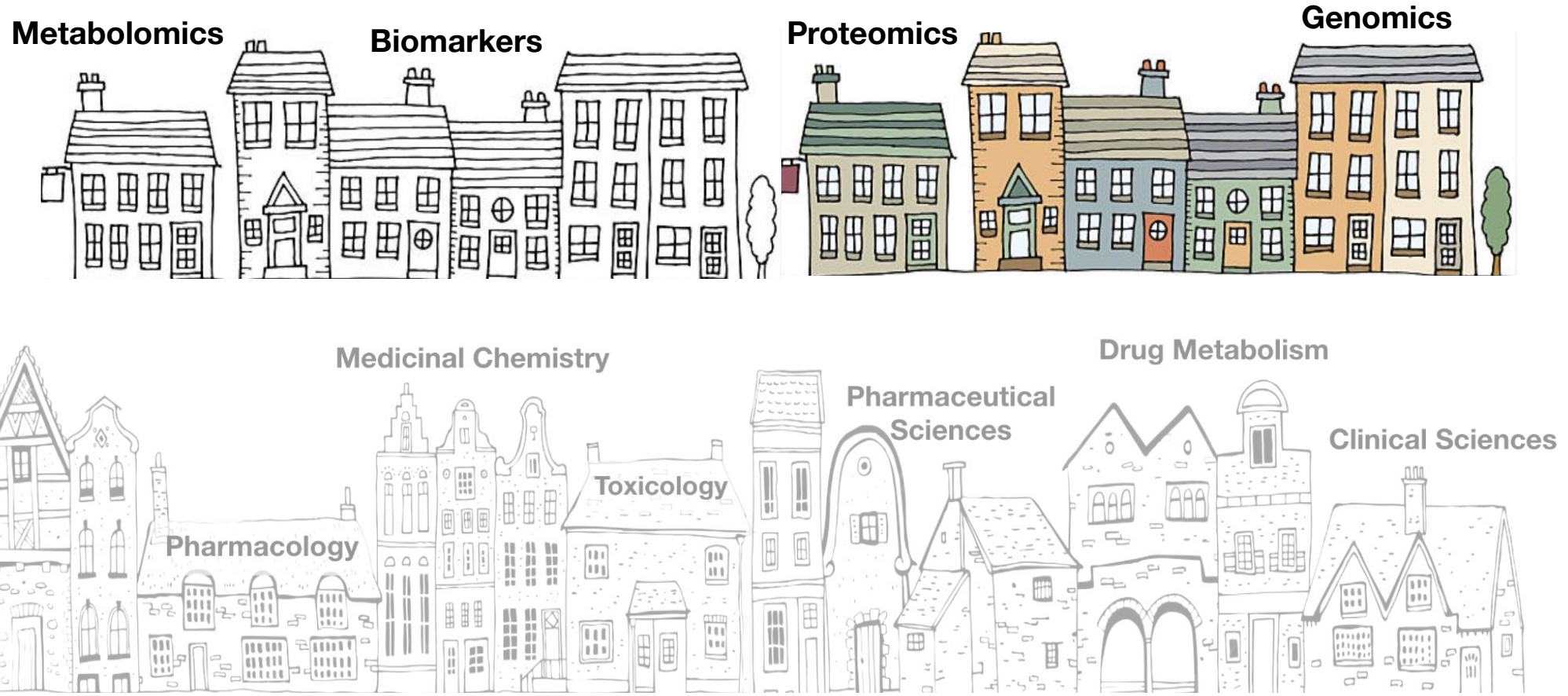


Nat Rev Drug Discov (2018) 17, 167-181 *Morgan P,*

It Takes a Village!



Growing Impact of Omics & Precision Sciences



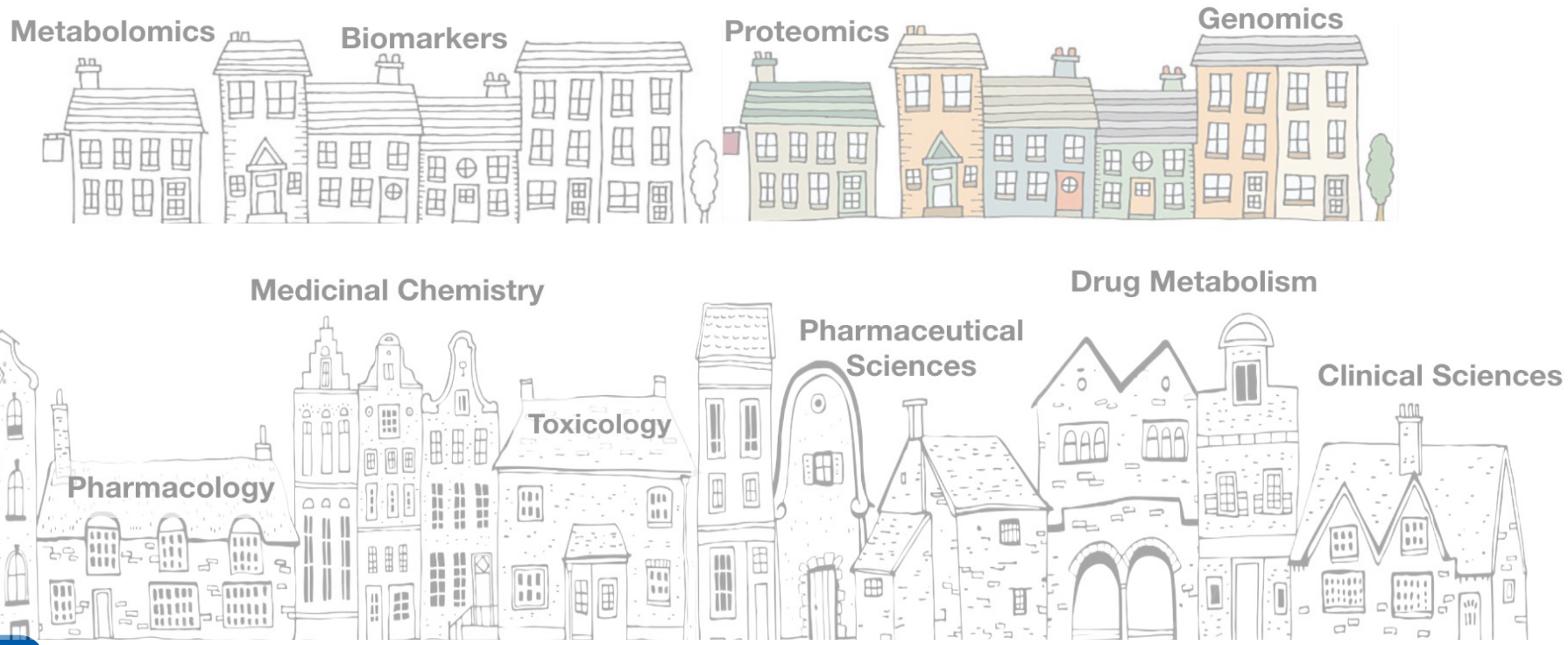
Growing Impact of Computational Sciences

Computational Chemistry

Protein Structure Prediction

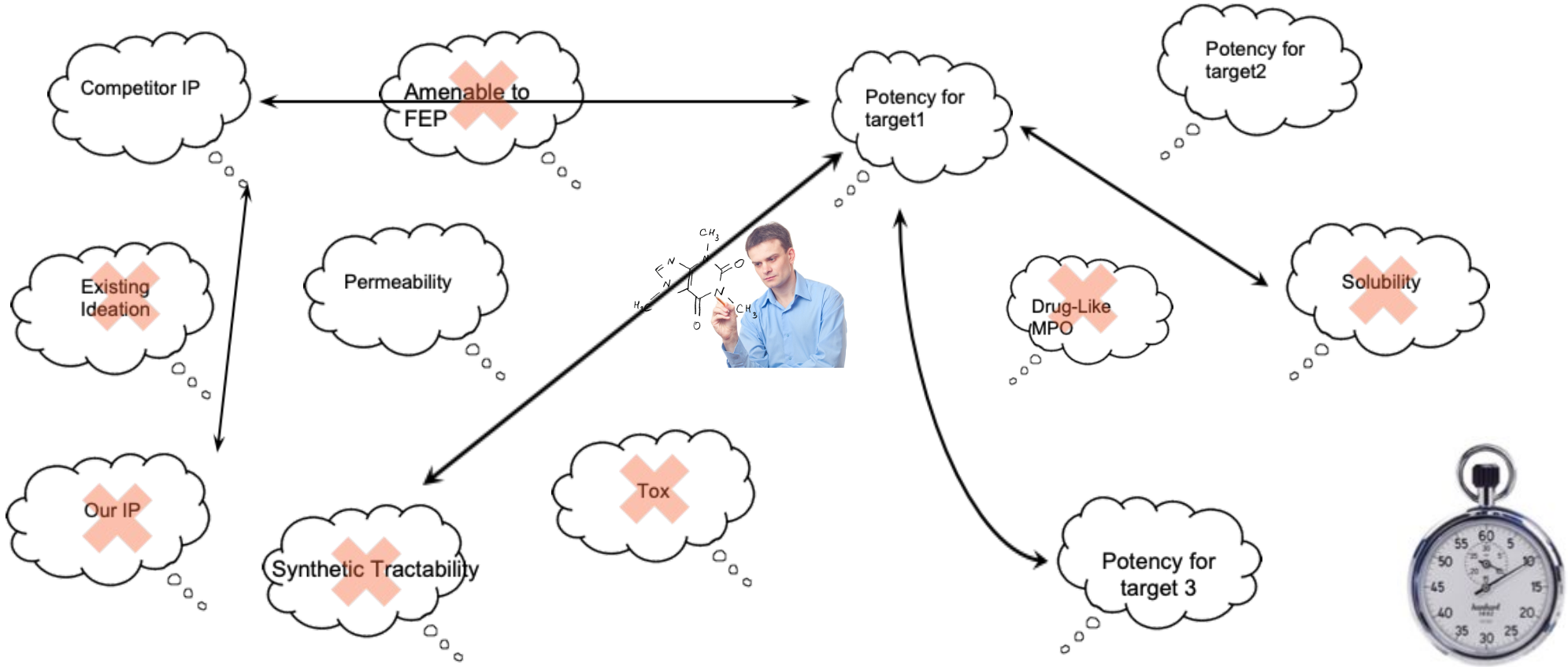
Computational Biology / Transcriptomics

Whole Genome Sequencing



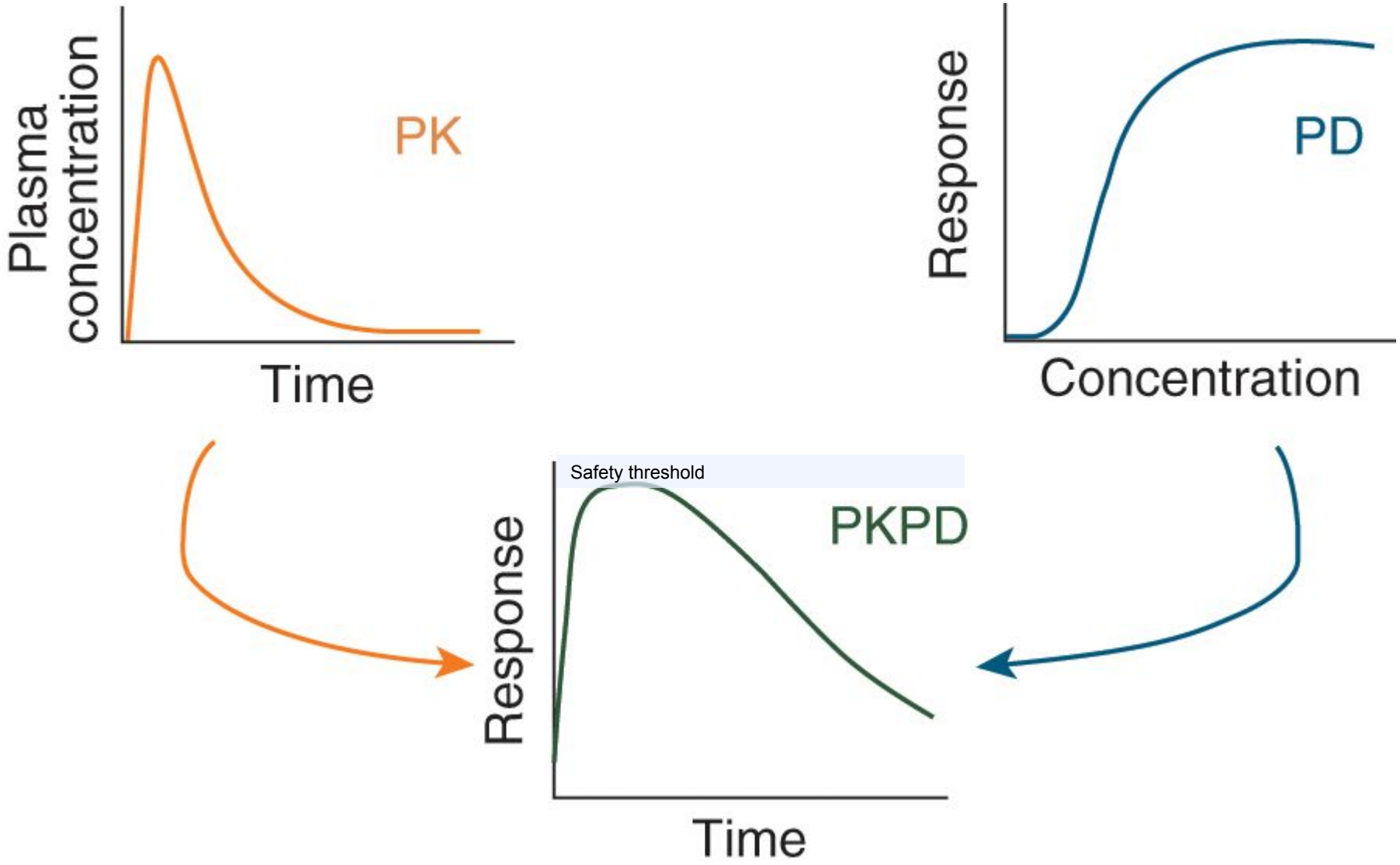
Drug Discovery is a Complex Multi-Parameter Optimization Problem

- Chemical space is vast
- Can high performance computing help us design more complete molecules faster?



Kutchukian PS et al. (2012) Inside the Mind of a Medicinal Chemist: The Role of Human Bias in Compound Prioritization during Drug Discovery. PLoS ONE 7(11): e48476. https://doi.org/10.1371/journal.pone.0048476

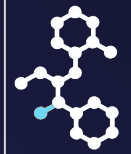
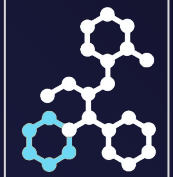
How will your compound will behave in vivo?



Designing Drugs Is Extremely Hard!

Lengthy, Capital-intensive, and Subject to High Failure Rates

Need to identify a molecule that balances a large number of anti-correlated properties:

						
• Potency	✓	×	✓	×	✓	✓
• Selectivity	×	✓	✓	✓	×	✓
• Solubility	×	×	×	✓	✓	×
• Bioavailability	×	×	×	×	×	×
• Clearance / Half-life	×	×	×	×	×	×
• Permeability	×	×	×	×	×	×
• Drug-drug interactions	×	×	×	×	×	×
• Synthesizability	×	×	×	×	×	×

*Paul, S.M., Mytelka, D.S., Dunwiddie, C.T. et al. "How to improve R&D productivity: the pharmaceutical industry's grand challenge." *Nature reviews Drug discovery*

Potential Solutions to Accurately Predicting Drug Properties

Decades-long Challenge – Two Major Approaches

1 Knowledge-based machine learning (often referred to as AI)

If AI can:

- beat humans at chess and Go
- recognize faces in photos
- autonomously drive cars

Can it be used to design drugs?

2 Rigorous first principles physics-based modeling

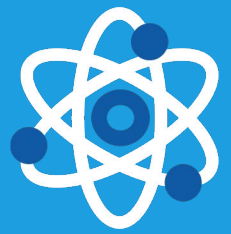
Requires deep understanding of the physics underlying highly complex molecular interactions

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \qquad i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t)$$

Physics & Machine Learning Are Complementary

Machine Learning / Artificial Intelligence

- ✓ Effective at interpolation
- ✓ Fast
- ✓ Can handle very large datasets
- ✘ Requires very large training set
- ✘ Cannot extrapolate



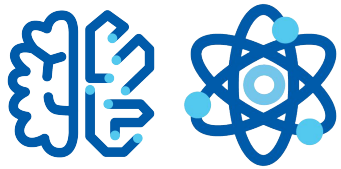
Physics-based Methods

- ✓ No training set required
- ✓ Can extrapolate into novel chemical space
- ✓ Accurate
- ✘ Slow

Physics & Machine Learning Are Complementary

Machine Learning / Artificial Intelligence

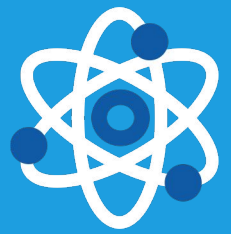
- ✓ Effective at interpolation
- ✓ Fast
- ✓ Can handle very large datasets
- ✘ Requires very large training set
- ✘ Cannot extrapolate



Machine Learning + Physics

Training set for ML generated using physics

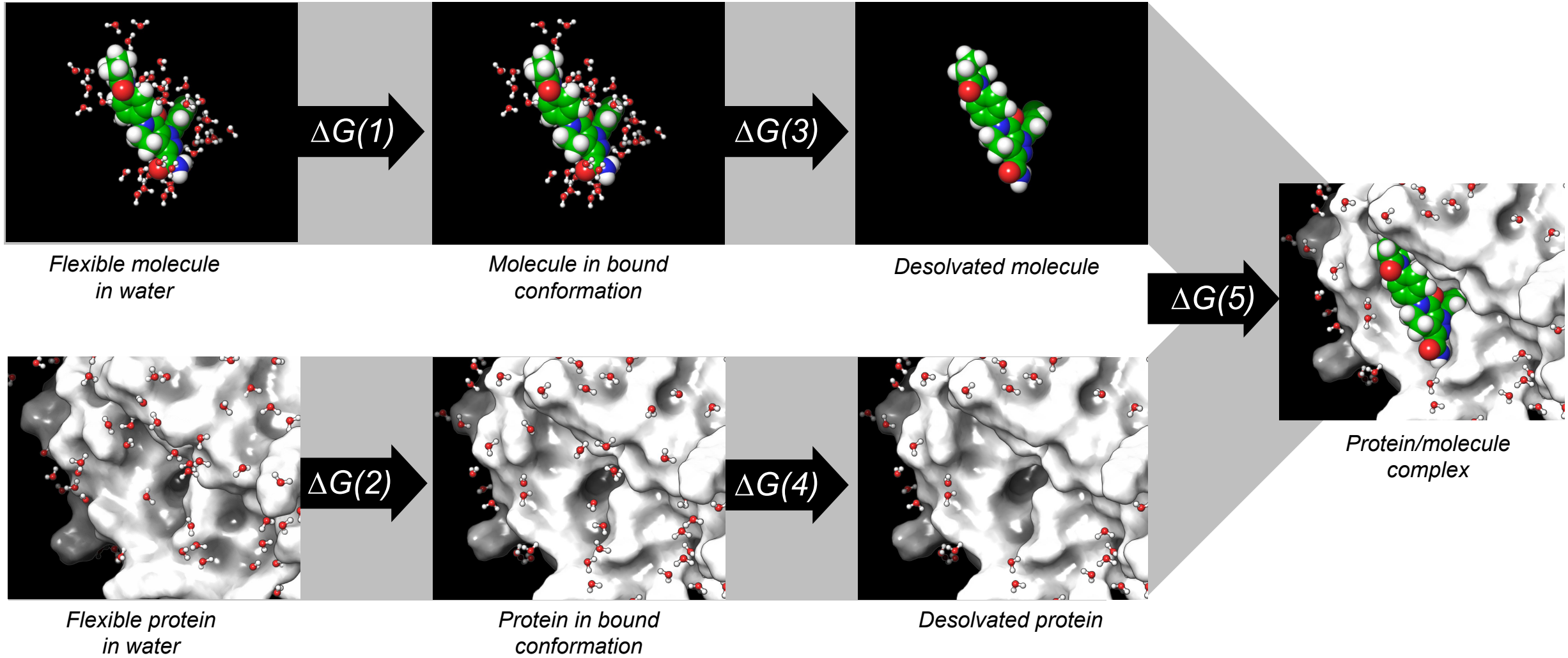
- ✓ Fast
- ✓ Accurate
- ✓ Can handle very large datasets
- ✓ Can extrapolate into novel chemical space



Physics-based Methods

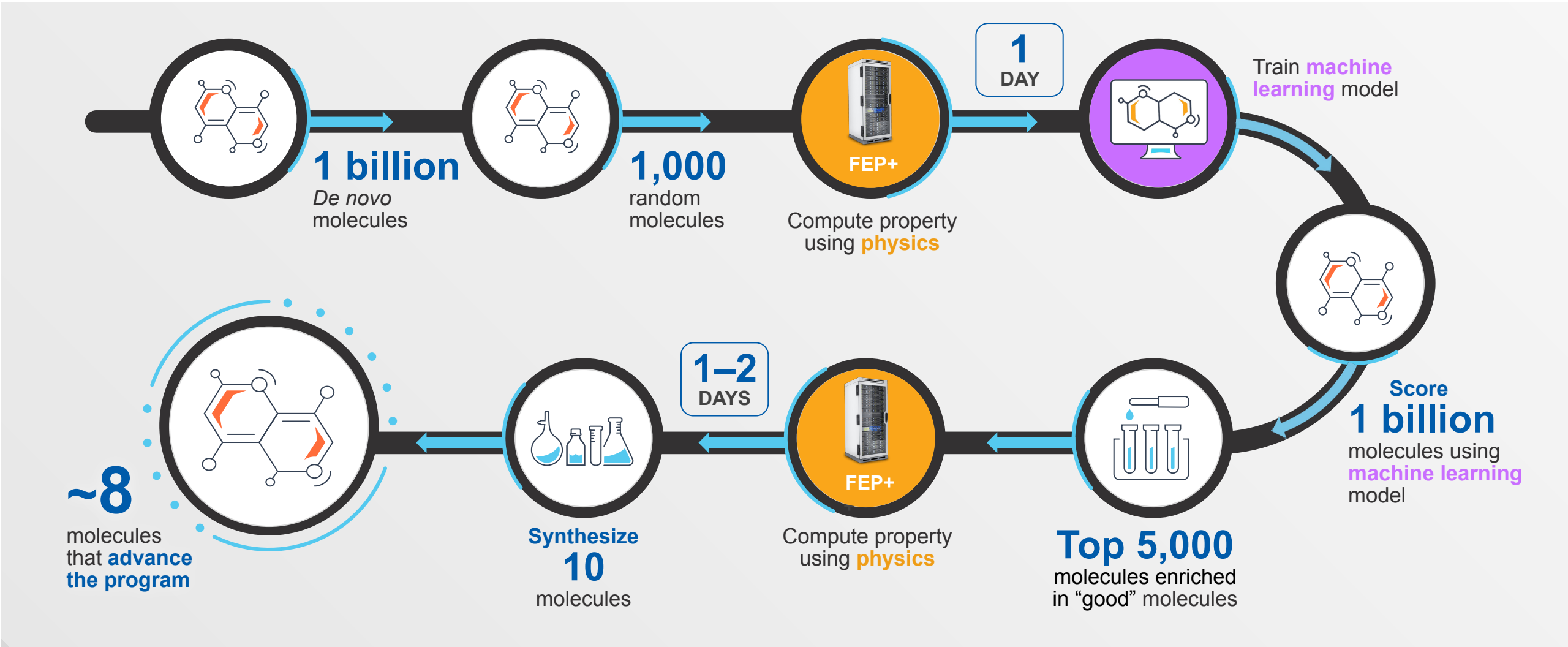
- ✓ No training set required
- ✓ Can extrapolate into novel chemical space
- ✓ Accurate
- ✘ Slow

Rigorous Simulation of Full Thermodynamic Cycle Required to Accurately Predict Binding Affinity of a Molecule to a Protein



$$\Delta G_{bind} = \Delta G(1) + \Delta G(2) + \Delta G(3) + \Delta G(4) + \Delta G(5)$$

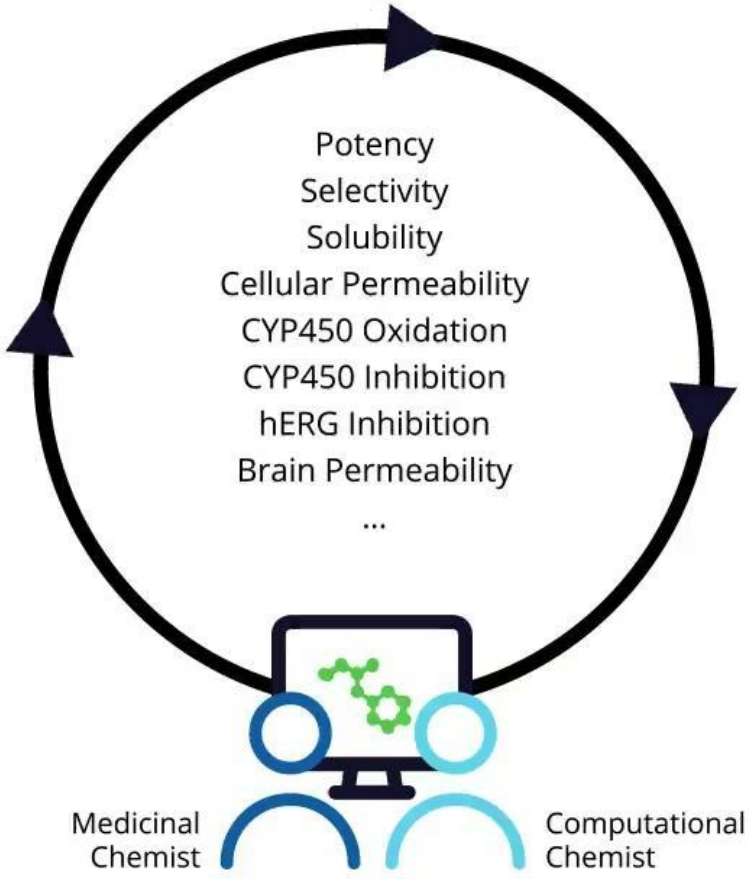
Combining Accuracy of Physics with Speed of Machine Learning Enables Ultra-Large Scale Exploration of Chemical Space

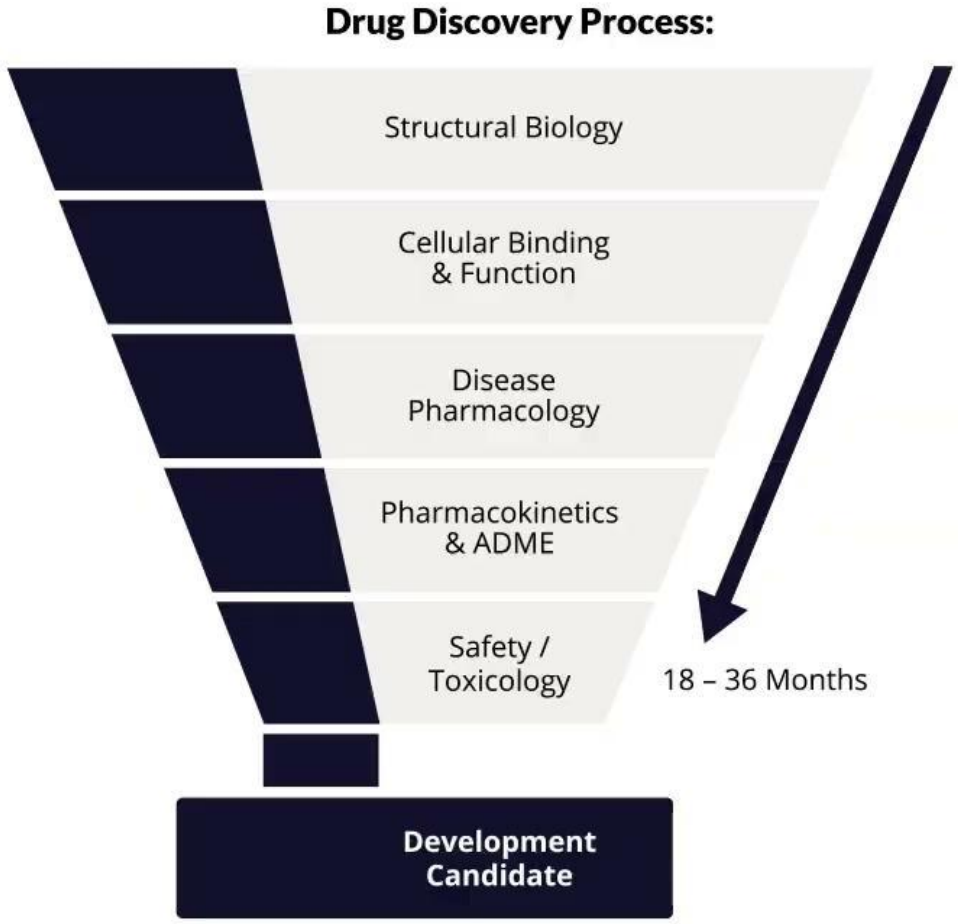


Incorporating Physics-based Compound Design into Drug Discovery

Upto **100 billion** idea molecules scored

100s of molecules synthesized and tested





Demonstrated Benefits of Integrating ML+ Physics-based methods

Reduces time and cost, and increases quality vs. traditional drug design

Traditional Drug Design

- **Manual** molecule design
- **~5,000** molecules synthesized and tested over **~4 – 6 years**

Hit Discovery → **Hit-to-Lead** → **Lead Optimization** → **!**

Drug development candidate with **property issues**

Schrödinger's Physics-Based Platform

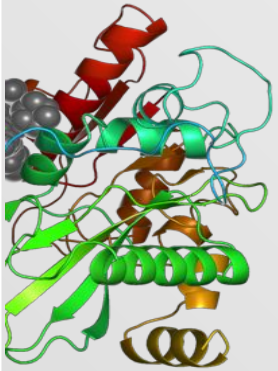
Hit Discovery → **Hit-to-Lead** → **Lead Optimization** → ✓

Drug development candidate with **Optimal property profile**

- **Billions** of molecules tested in computational assays
- **<1,000** molecules synthesized and tested over **~1.5 – 3 years**

Why is now the time for broader adoption of structure-based drug design?

External trends



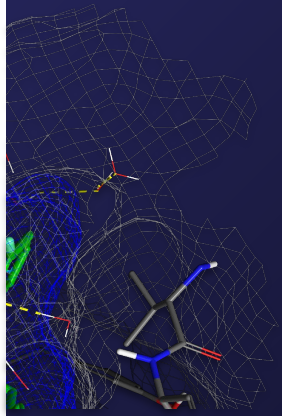
Structural Biology Advances

Exponential growth of experimental and predicted protein structures from x-ray, cryo-EM and AlphaFold

Computing Power & Speed

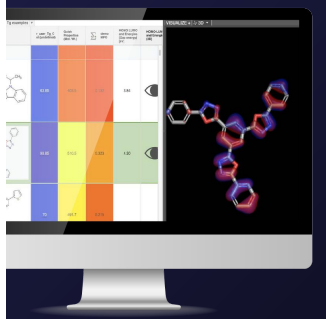
Massive power of GPU computing and burst capability on the cloud

Schrödinger Technology



Physics-Based Modeling & Machine Learning

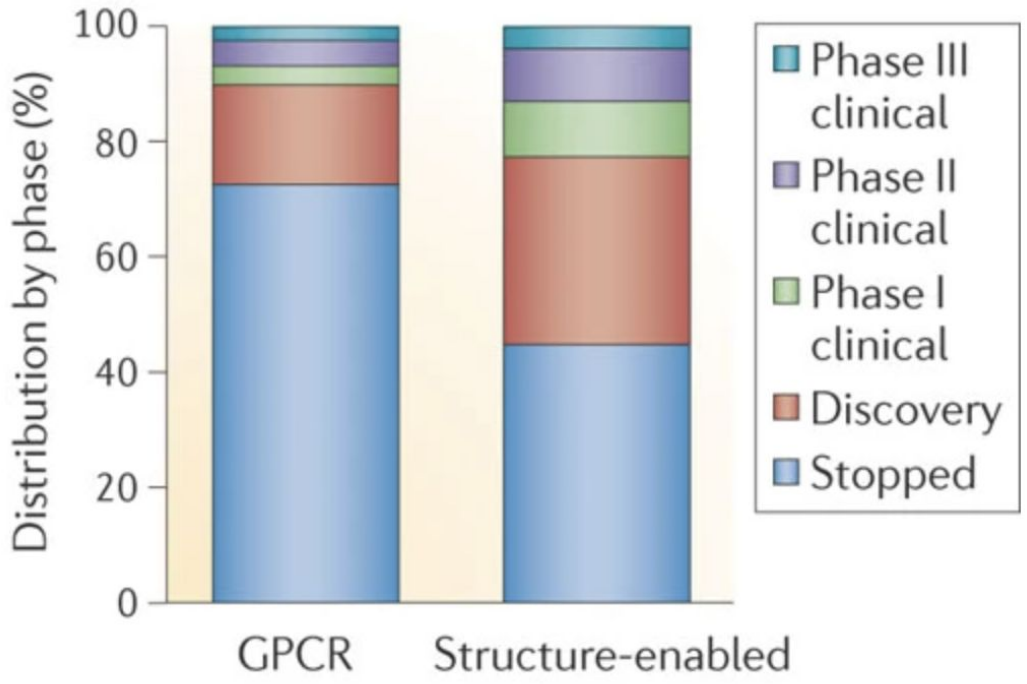
Prediction accuracy matching experimental assays, at scale



Web-based Enterprise Ideation & Analysis Informatics Platform

Cross-functional collaboration and design platform supports full project teams

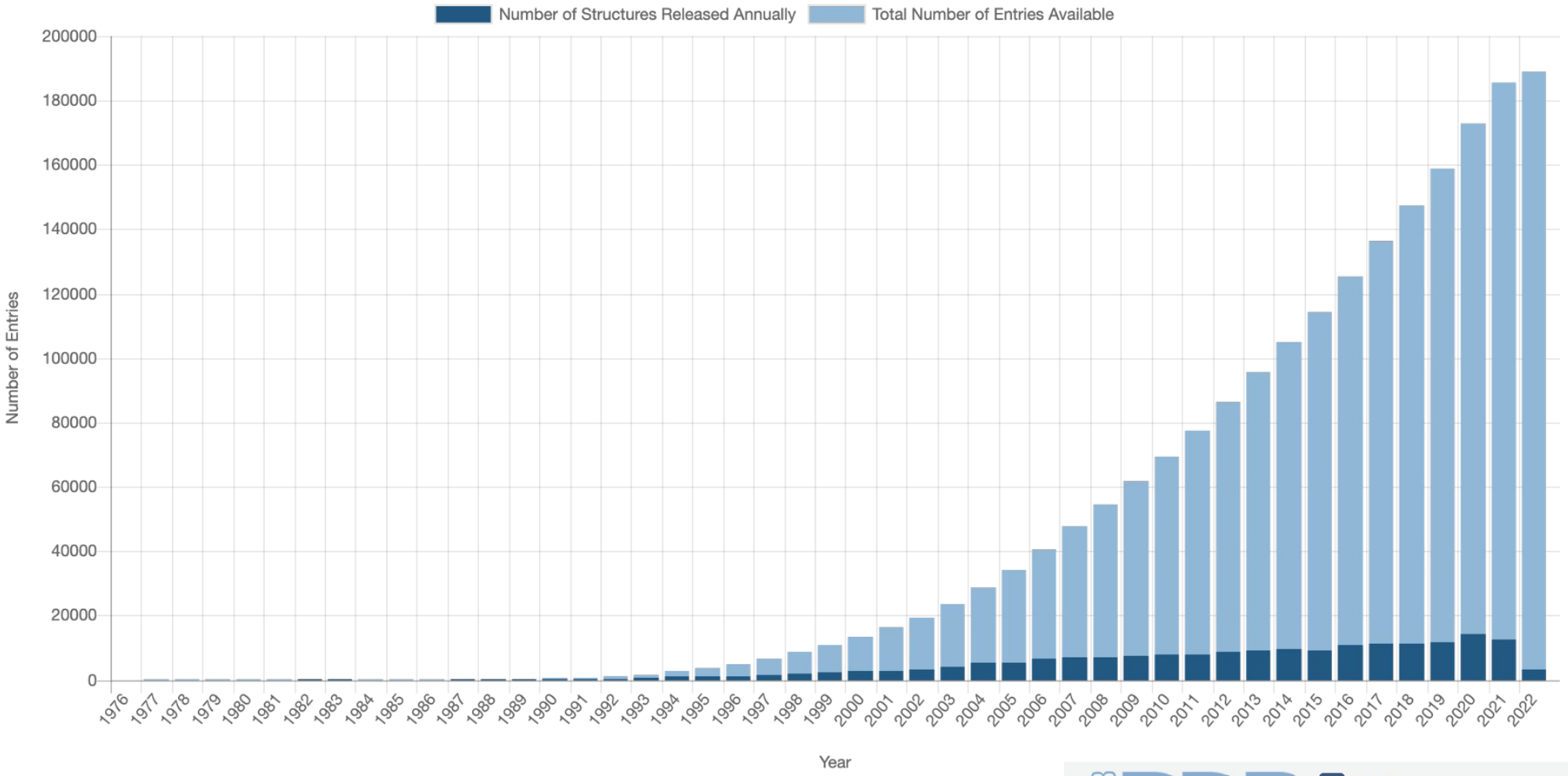
Impact of Structural Enablement on Success Rates?



Proportion in active versus stopped projects at different pipeline stages for a group of ten structure-enabled enzyme targets compared with ten G protein-coupled receptor (GPCR) targets in 2011

Source: Borshell, N., Papp, T. & Congreve, M. Valuation benefits of structure-enabled drug discovery. Nat Rev Drug Discov 10, 166 (2011). https://doi.org/10.1038/nrd3392

Structural Biology Advances over the past 50 years

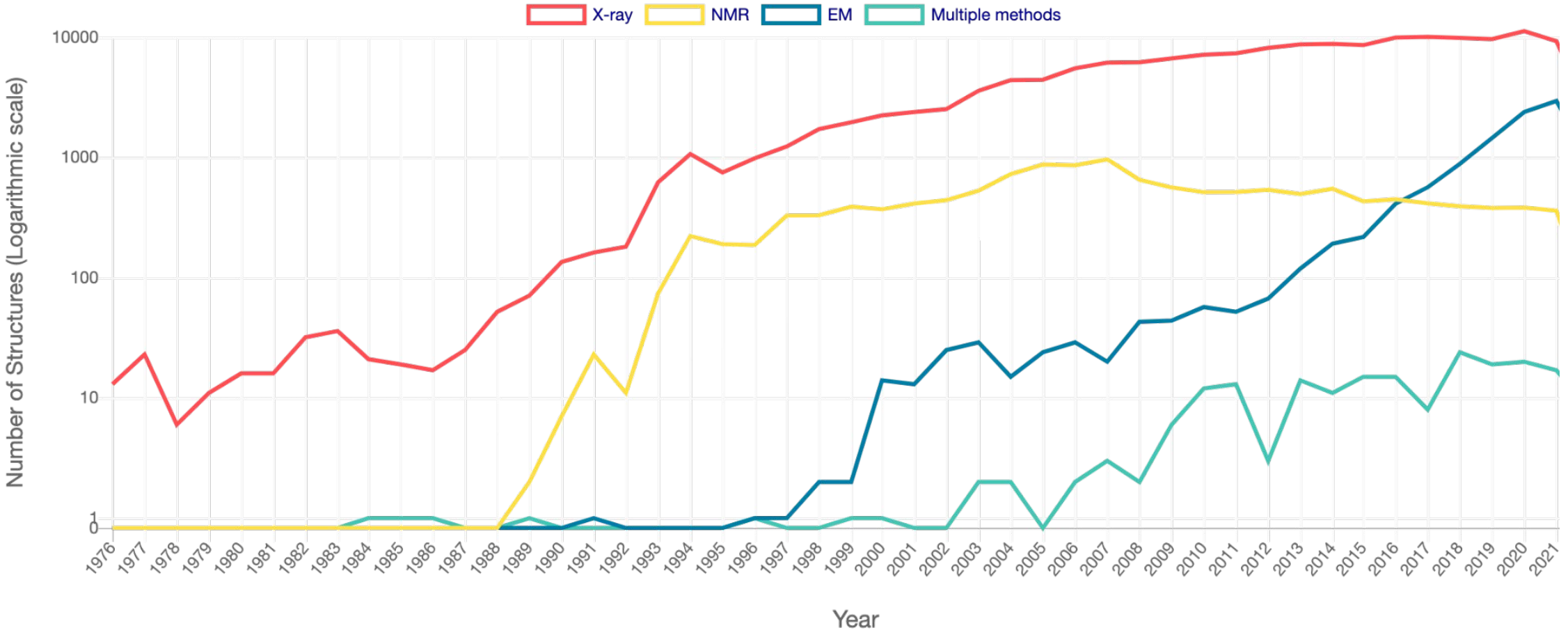


RCSB PDB PROTEIN DATA BANK

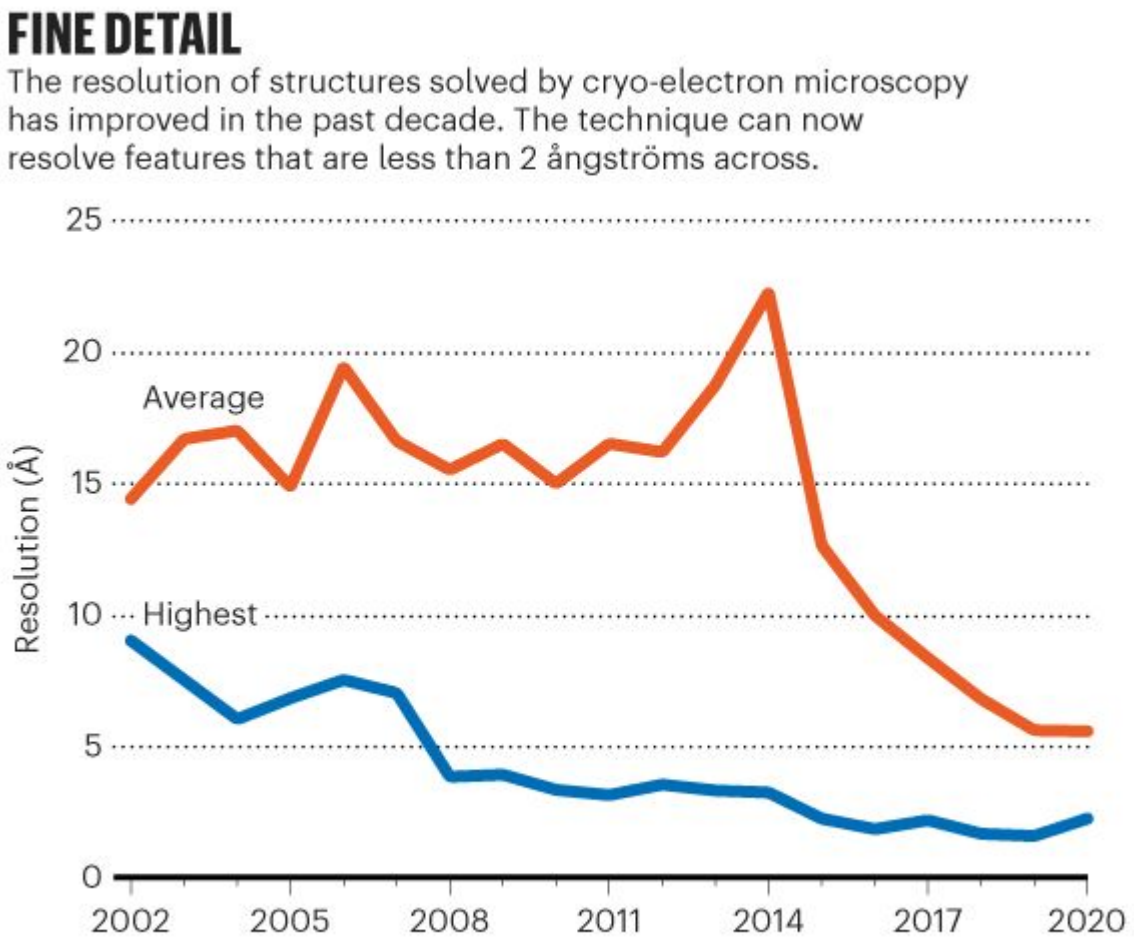
201,196 Structures from the PDB

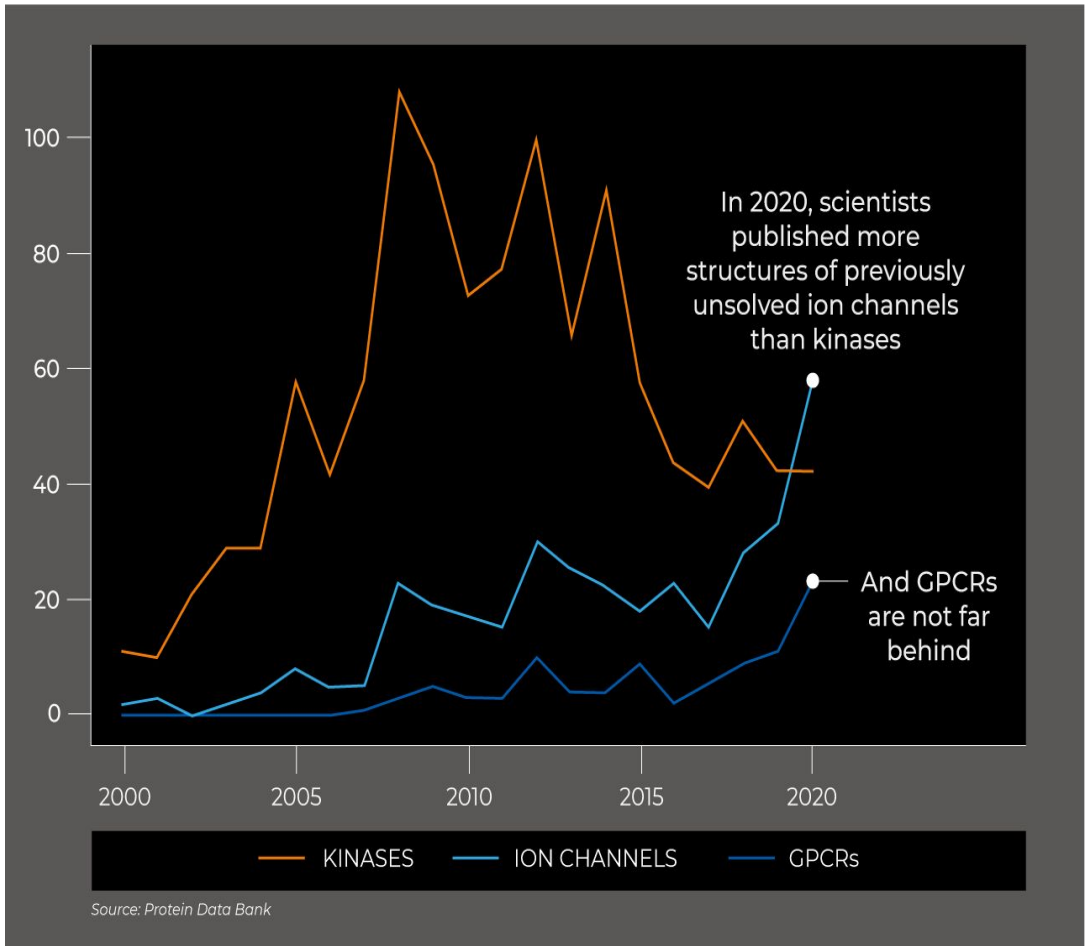
1,068,577 Computed Structure Models (CSM)

Structural Biology Advances - Growth in CryoEM Structures



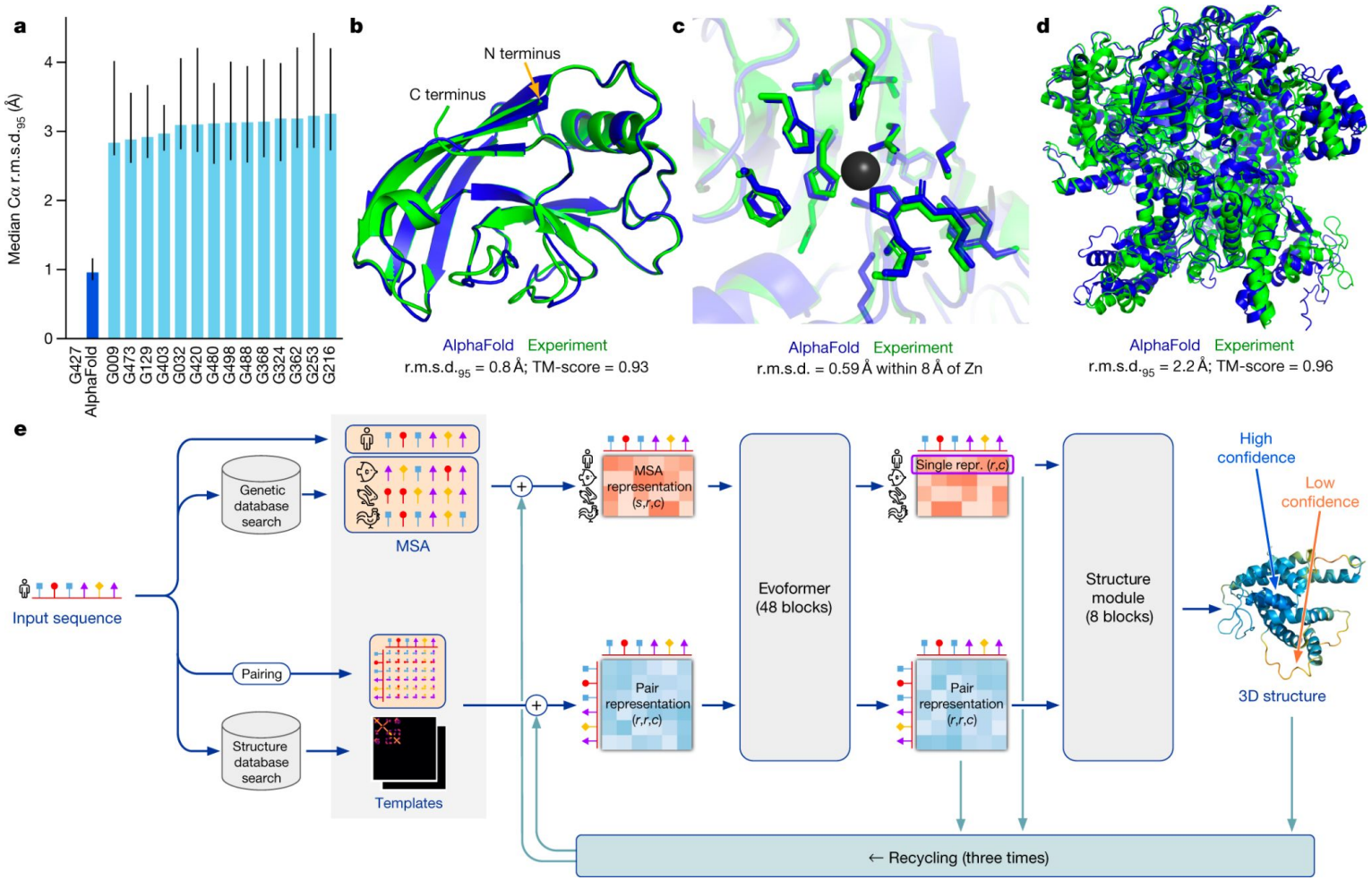
Higher Resolution Structures are Increasingly Available

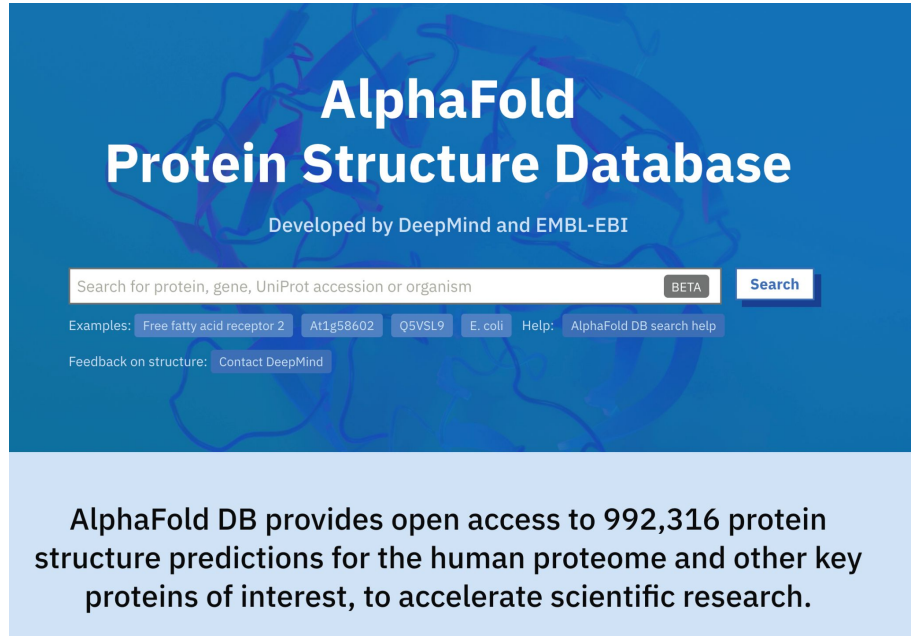




Revolutionary cryo-EM is taking over structural biology Nature 578, 201 (2020) doi: https://doi.org/10.1038/d41586-020-00341-9

Computational Models of Protein Structures: AlphaFold DB



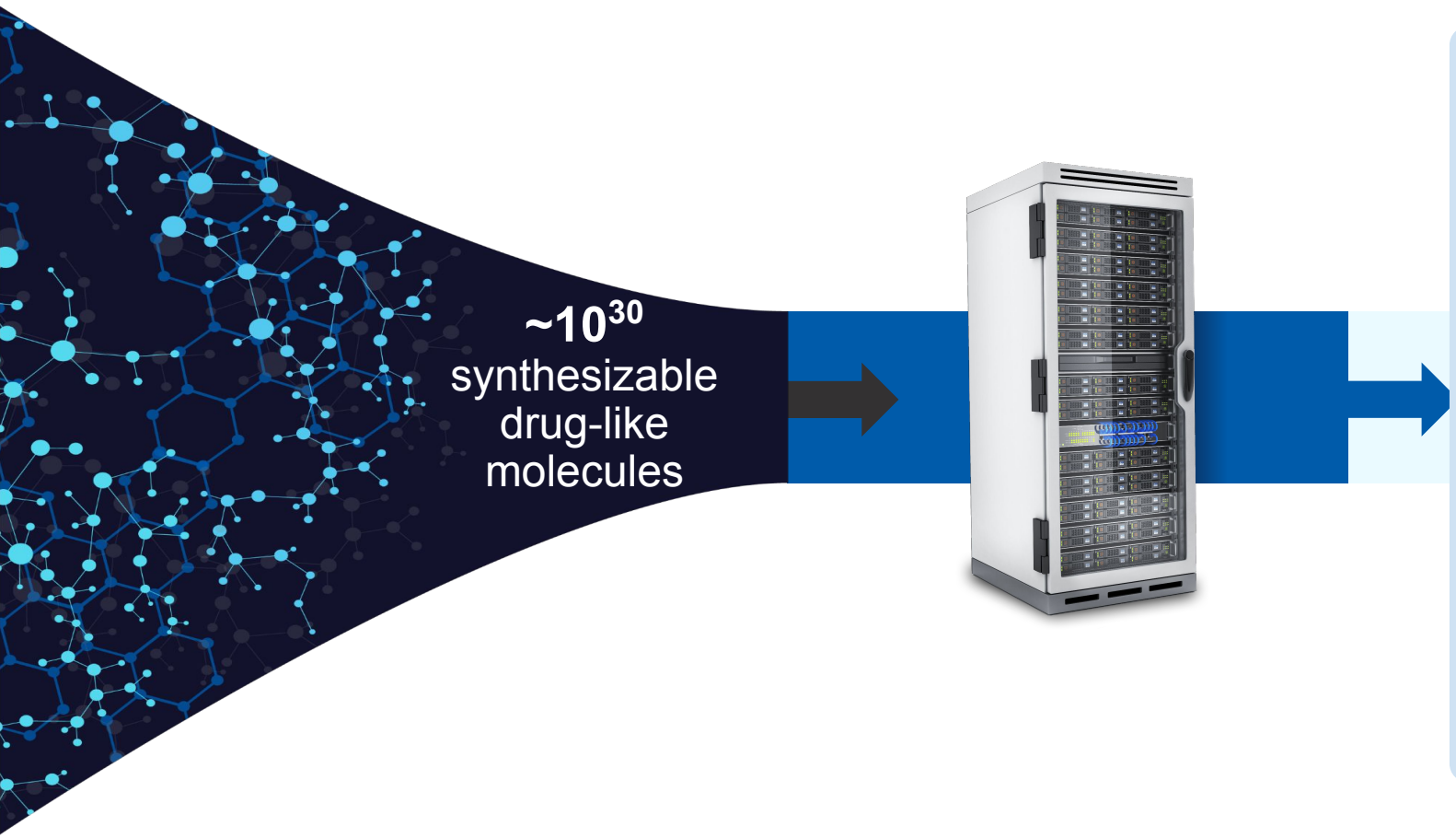


AlphaFold DB provides open access to 992,316 protein structure predictions for the human proteome and other key proteins of interest, to accelerate scientific research.

Highly accurate protein structure prediction with AlphaFold
John Jumper, et al Nature volume 596, pages 583–589 (2021)

Digital chemistry - expanding experiments through predictions

Proven to **enable precision design** of new molecules that chemists have not considered with the confidence of **accurate property predictions comparable to physical experiments**. De-risks exploration of synthetically-challenging and counter-intuitive chemistry to open new paths to discovery

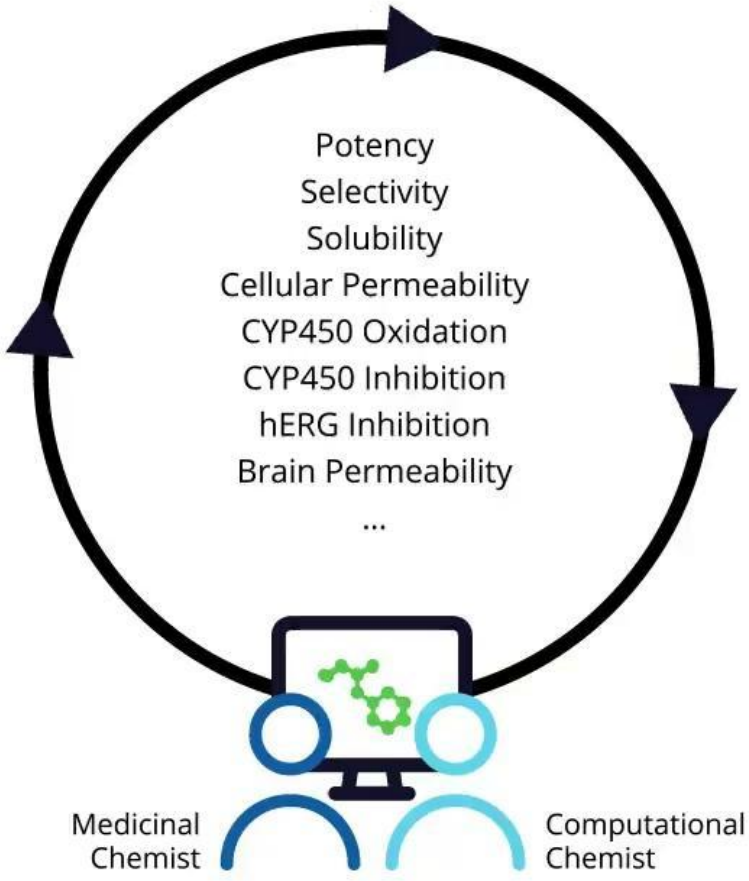


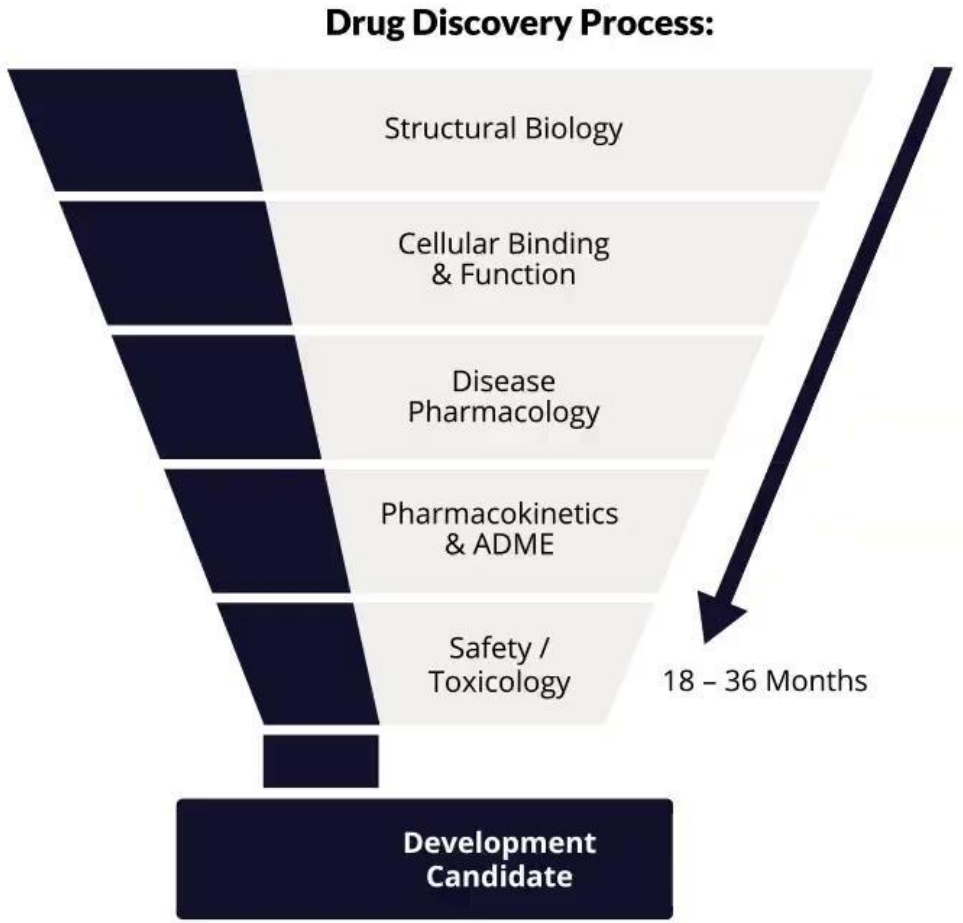
Free Energy Calculations (FEP+) as “computational assay” for binding affinity prediction allows rapid expiration of millions of R-groups (example is a Kinase p-loop above)

Incorporating Physics-based Compound Design into Drug Discovery

Upto **100 billion** idea molecules scored

100s of molecules synthesized and tested





Multi-Stage Computational Drug Discovery Solutions

Target Validation & Tractability

- CryoEM & X-ray structure optimization and ligand placement
- Computational models of protein structures
- Next generation induced fit docking
- Druggable binding site identification and tractability assessment

Hit ID

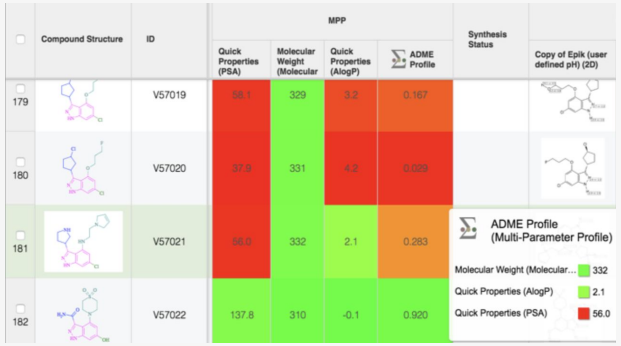
- Ultra large scale structure-based & ligand-based virtual screening (>10^9)
- ML-enhanced DNA-encoded library screening
- FEP+ enhanced fragment-based screening
- Protein FEP+ for gene-family wide selectivity optimization

Hit-to-Lead & Lead Optimization

- Physics-based computational assays combined with machine learning for large-scale synthesis aware optimization of project chemical matter (> 10^9)
- Highly accurate multi-parameter optimization of *Potency, selectivity, permeability, solubility, pKa*
- Intermediate accuracy ML models of hERG, PPB, RH, ...

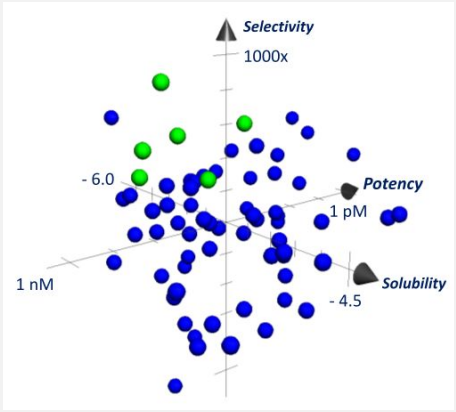
Chemical design efforts guided by large-scale multi-parametric modeling

Solving Design Challenges



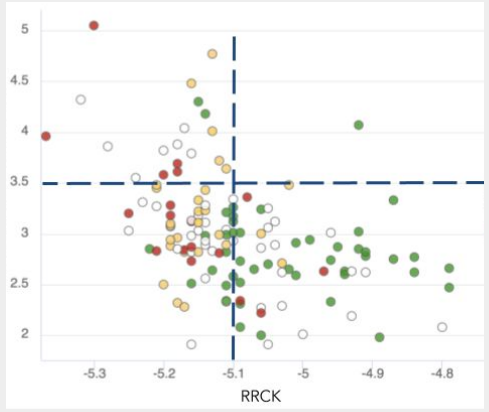
Team ideation

Scientist designs modeled by FEP+ for potency and selectivity



Iterative design and MPO

Simultaneous modeling of designs for potency, selectivity, solubility, and permeability using FEP+ and other physics-based methods



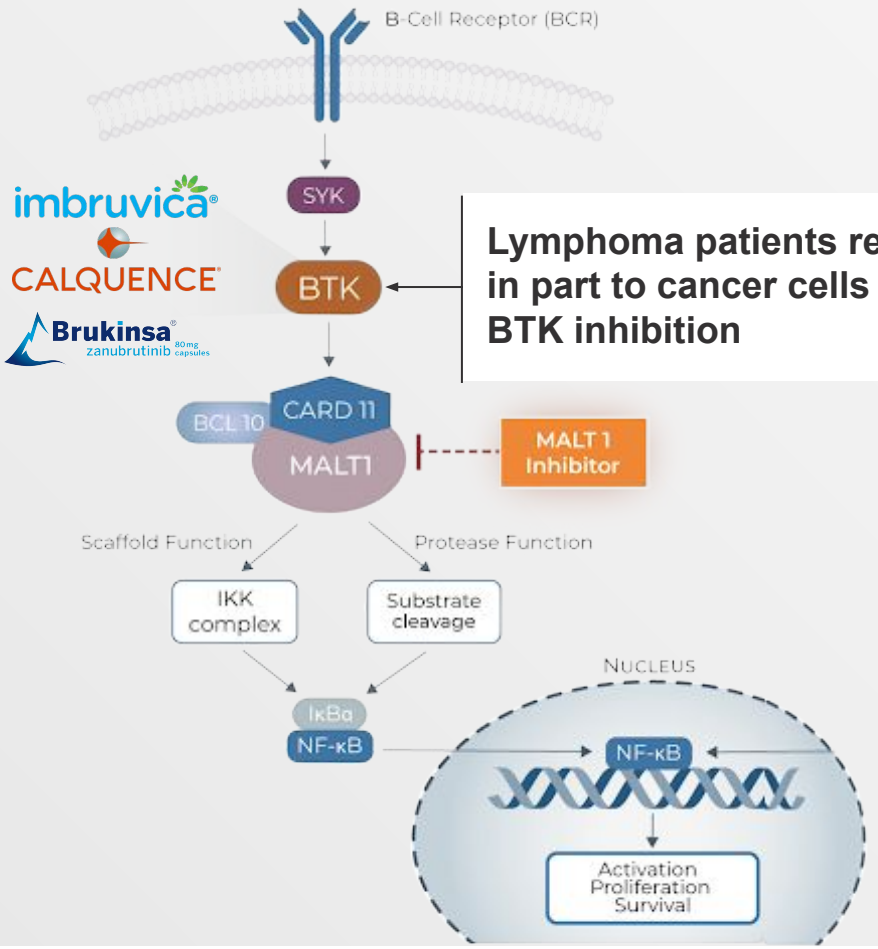
Fine-tuning of properties

FEP+ driven analogue design in a favorable chemical space

Bespoke machine learning and first principles models (pKa, BBB, ADME) based on program specific data

MALT1 inhibitors—
From launch to Development Candidate in under 2 years.

MALT1 Is a Genetically Validated Mechanism Regulating Lymphocytes



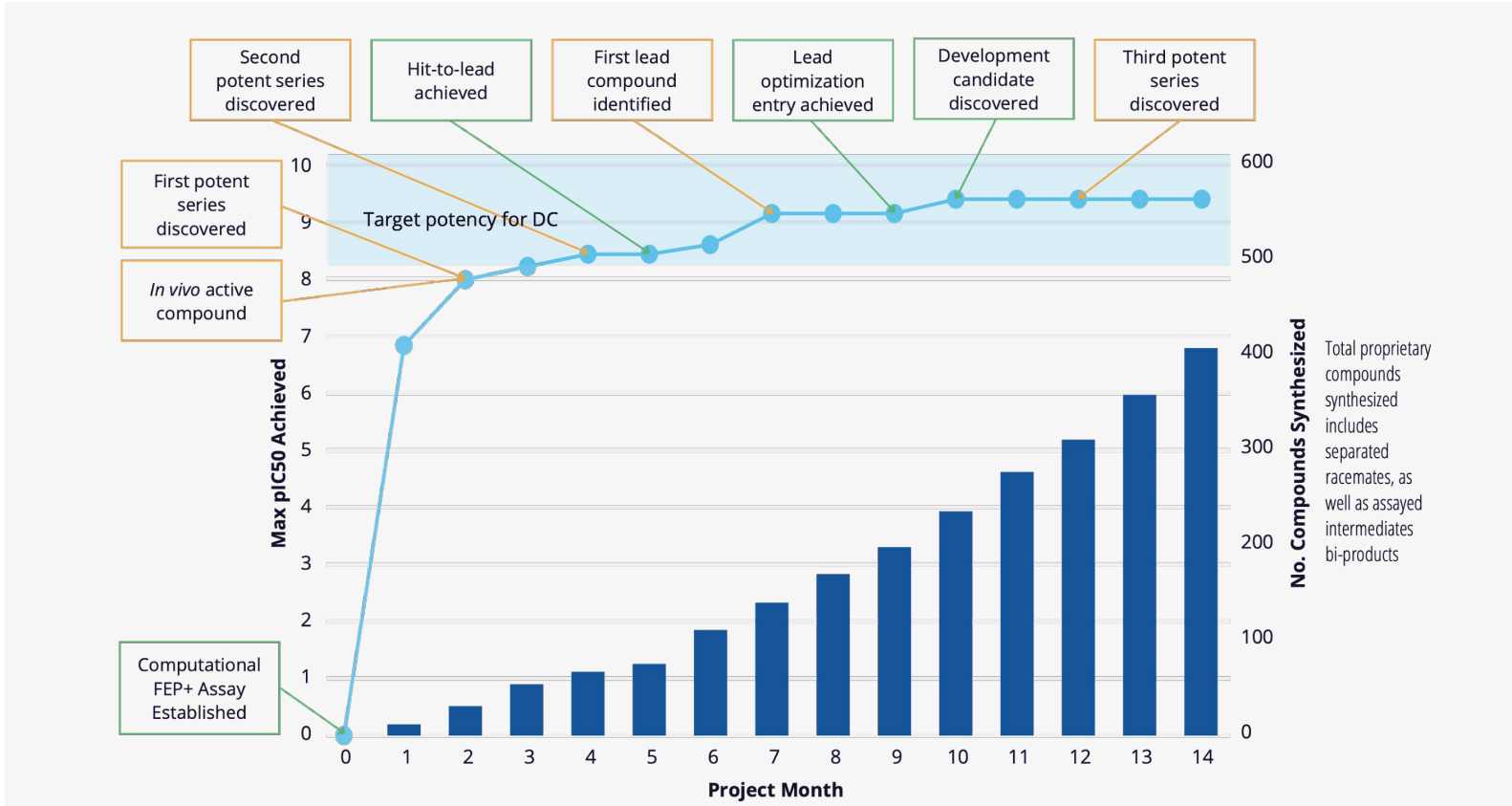
Lymphoma patients relapse due in part to cancer cells escaping BTK inhibition

Potential indications in multiple disease areas

- **B-Cell Malignancies**
- **Advanced Solid Tumors**
- **Autoimmune Diseases**

Rapid identification of several highly potent, novel series through exploration of vast chemical space

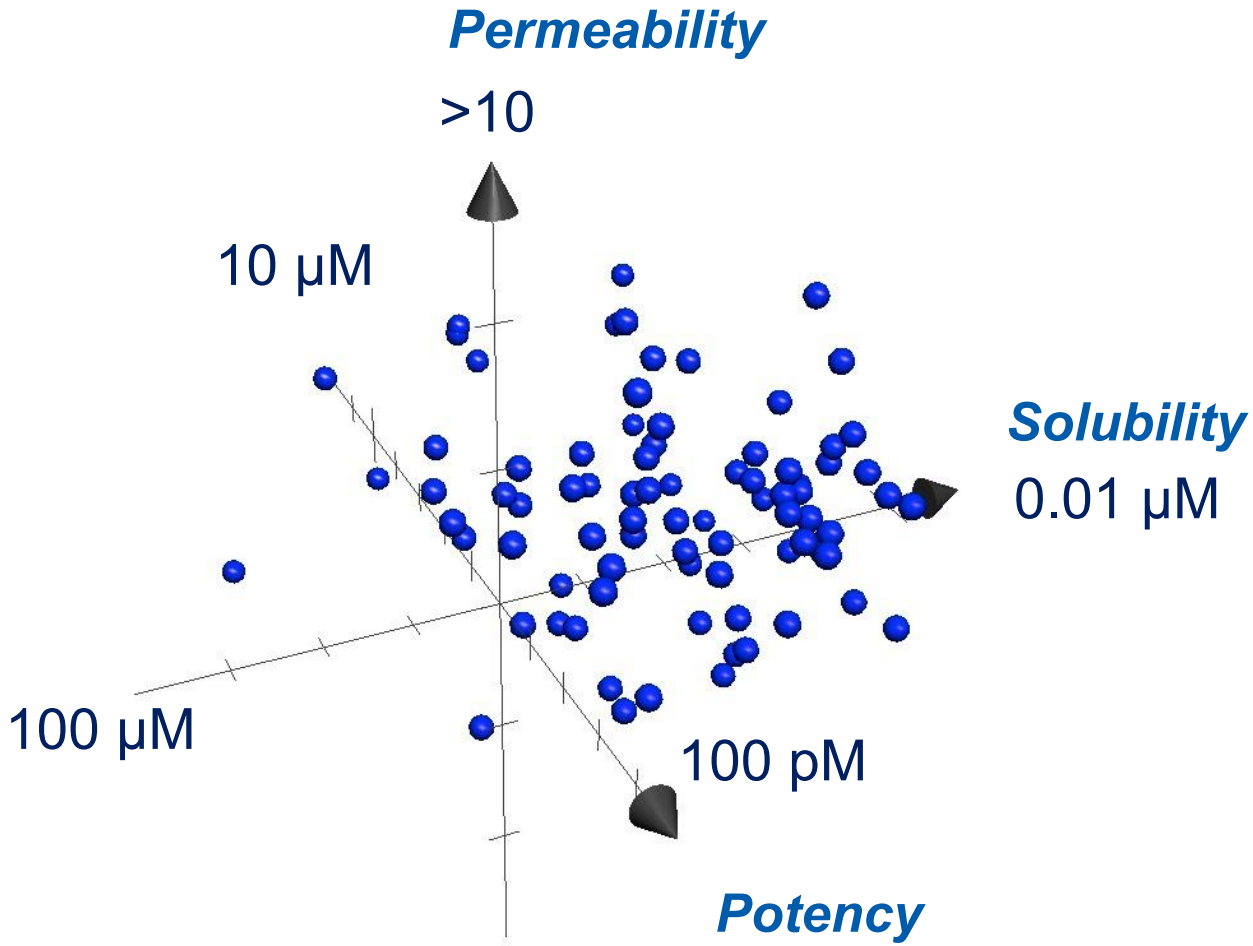
- **Highly potent, novel compounds** from two series identified quickly through exploration of billions of compounds
- ***In vivo* compounds available for biological testing** within 2 months of project start
- **Strong activity maintained** while optimizing multiple endpoints
- **Multiple additional novel, distinct potential backup series** discovered



- Billions of compounds triaged → ~12,186 modeled
- 78 compounds synthesized in the lead series in 10 months

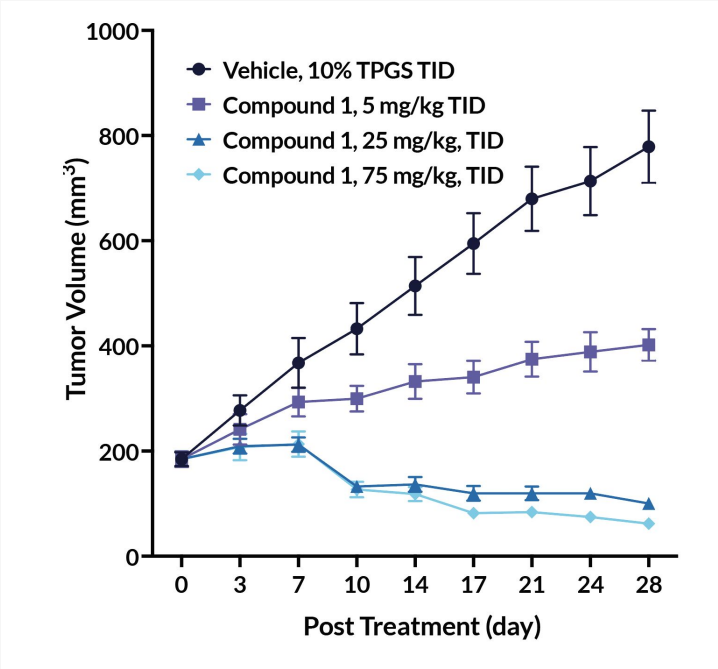
Simultaneously modeling of potency, permeability and solubility for efficient multi-parameter optimization

- **Modeling of multiple endpoints facilitated identification of potent ligands with balanced solubility and permeability**
  - Compounds synthesized based on prospective MPO predictions
- Lead molecule demonstrates very strong anti-tumor activity as **a single agent** and **in combination with BTK inhibitor** in selected lymphoma models

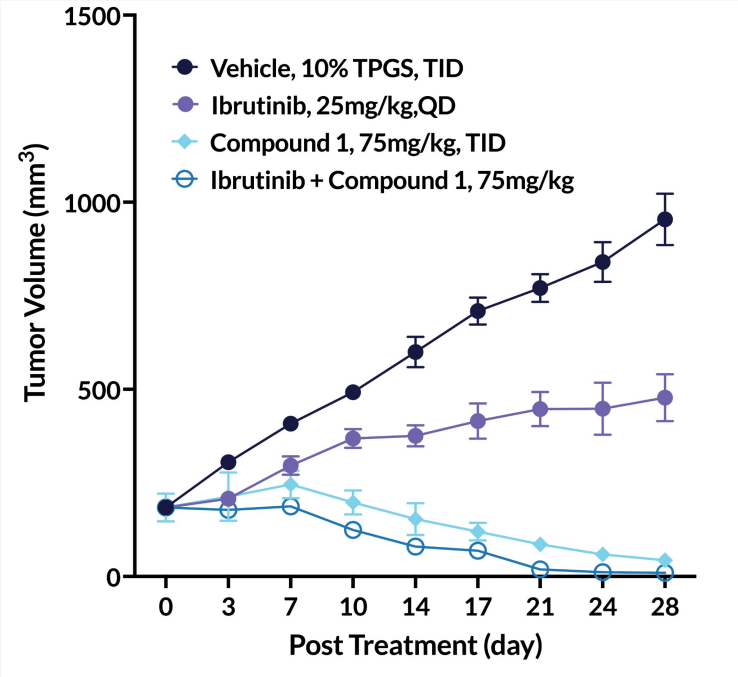


Lead candidate demonstrates strong anti-tumor activity as single agent and as combination therapy in vivo

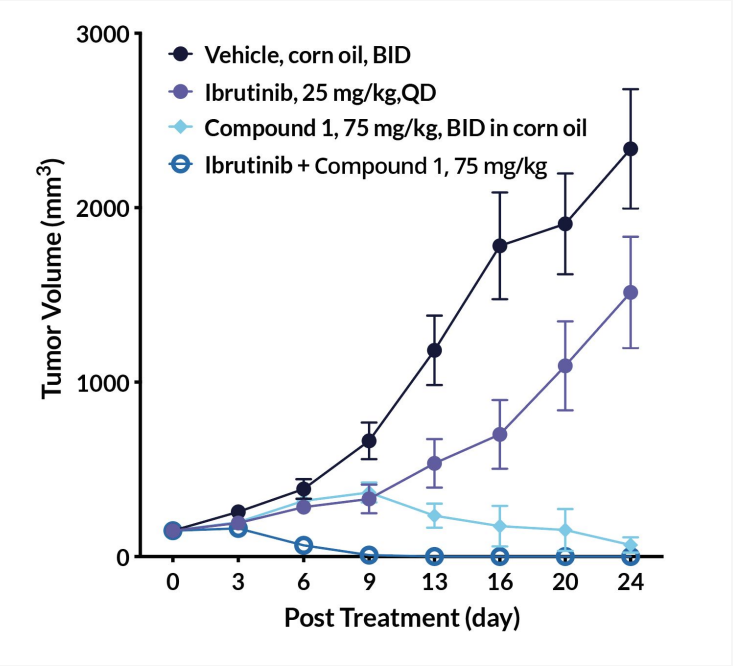
Anti-Tumor Activity in OCI-LY10 CDX model (Single agent)



Anti-Tumor Activity in OCI-LY10 CDX model (Combination with ibrutinib)



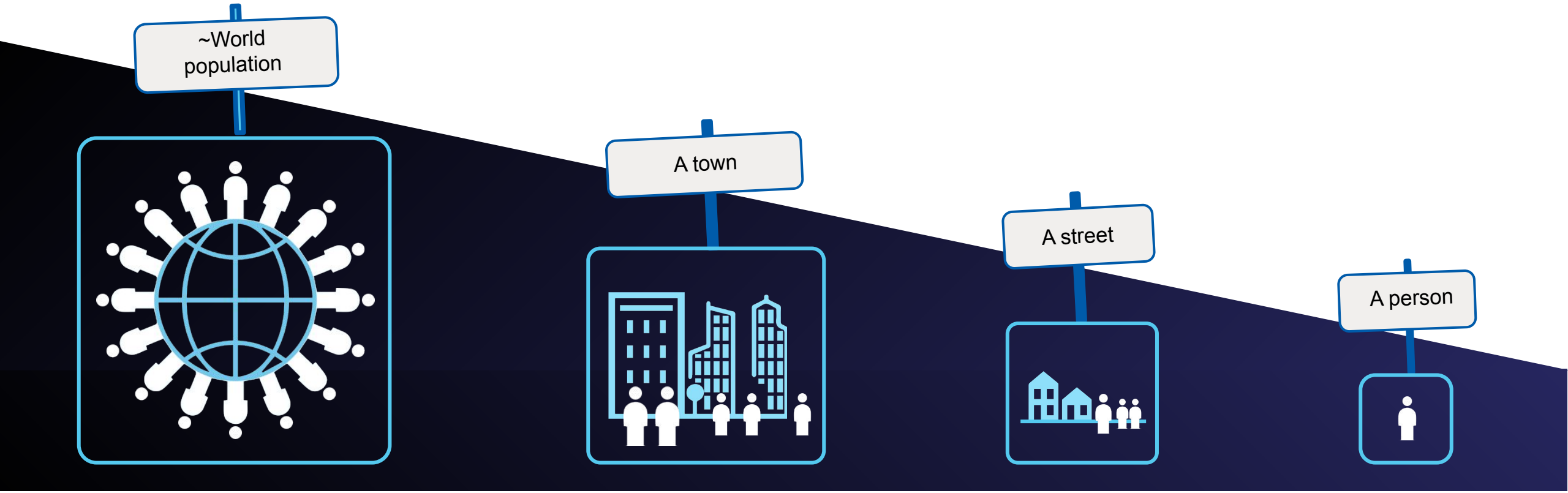
Anti-Tumor Activity in LY2298 PDX model (Combination with ibrutinib)



Yin et al., 2021 ASH Presentation

SGR-1505 is in Phase I: https://www.schrodinger.com/malt1

Computational Platform Supported Rapid MALT1 Program Progression



8.2 billion
compounds scored with machine learning

→

12,186
modelled and scored with physics

→

230
total compounds synthesized in multiple series

→

Development Candidate

~10 Months

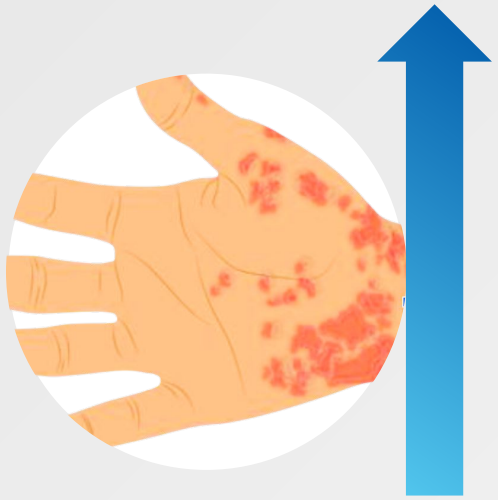
Selective TYK2 Inhibitor Collaboration

nimbus
THERAPEUTICS

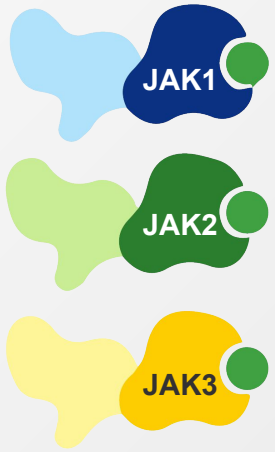
JAK/TYK Kinases Are Key Signaling Molecules in Inflammation

Including Psoriasis, Rheumatoid Arthritis, Crohn's and Lupus Disease

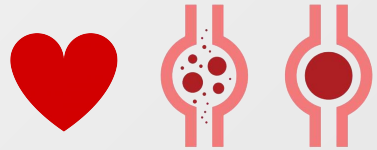
Psoriasis



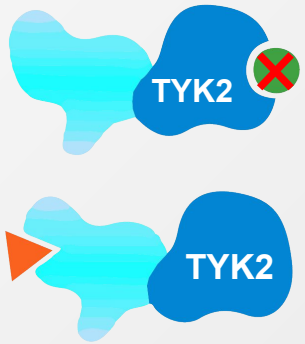
Inflammation in psoriasis is driven by increased interferon and cytokine release, which is regulated by the JAK/TYK family of proteins



Approved JAK inhibitor drugs hit multiple JAKs and have side effect safety warnings



Drug Design Goal

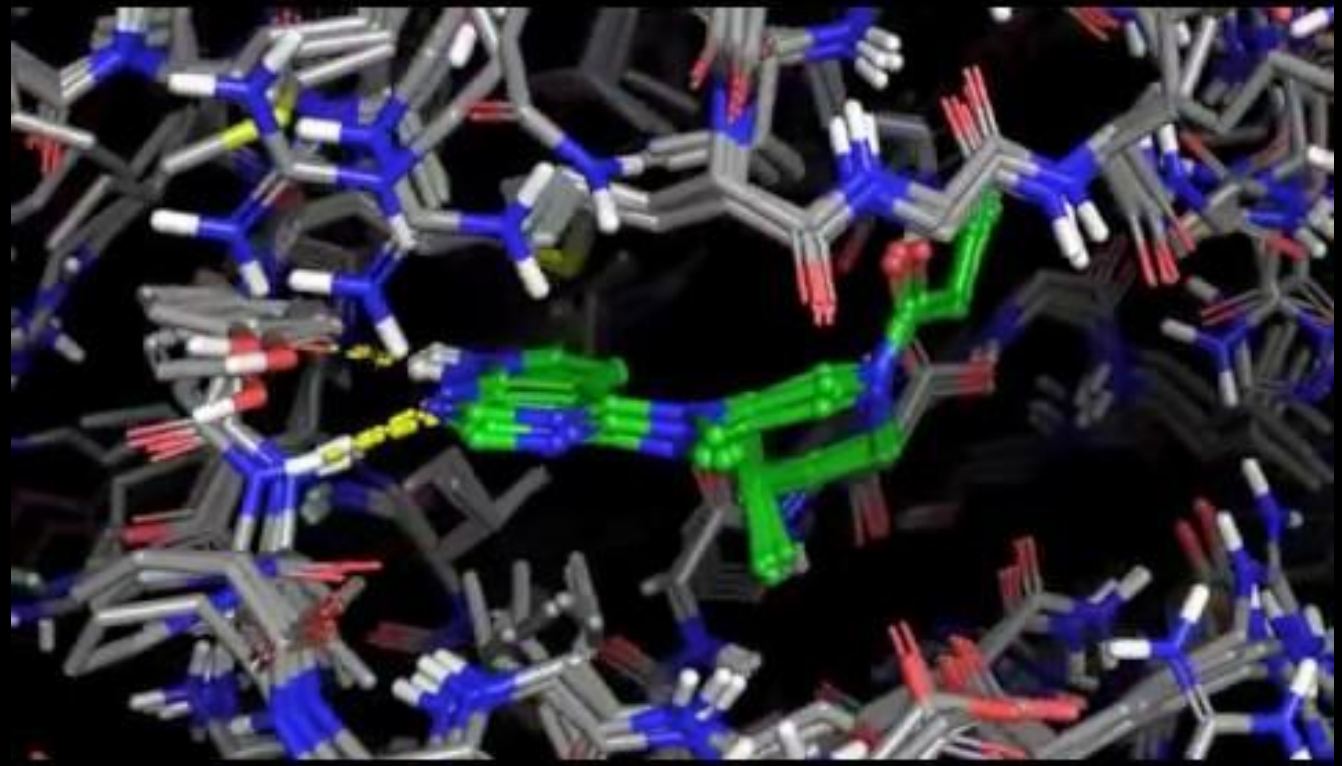


Avoid the kinase domain

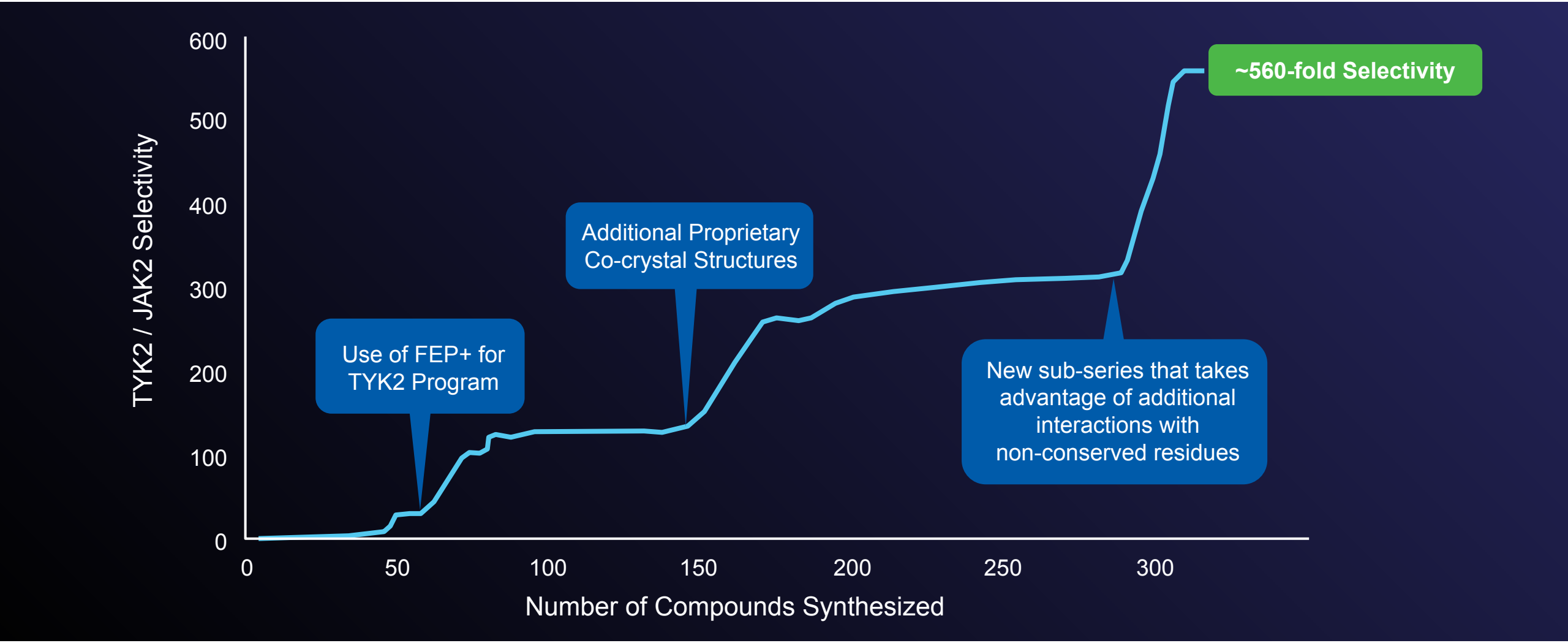
Focus on allosteric inhibitors that bind the regulatory domain and disrupt kinase activity

Selective Inhibition of TKY2 Is Highly Challenging

TYK2, JAK1, JAK2, JAK3 superimposed in complex with tofacitinib shows extremely high similarity in the active site — to achieve selectivity required accurate modeling of ligand binding to all 4 proteins



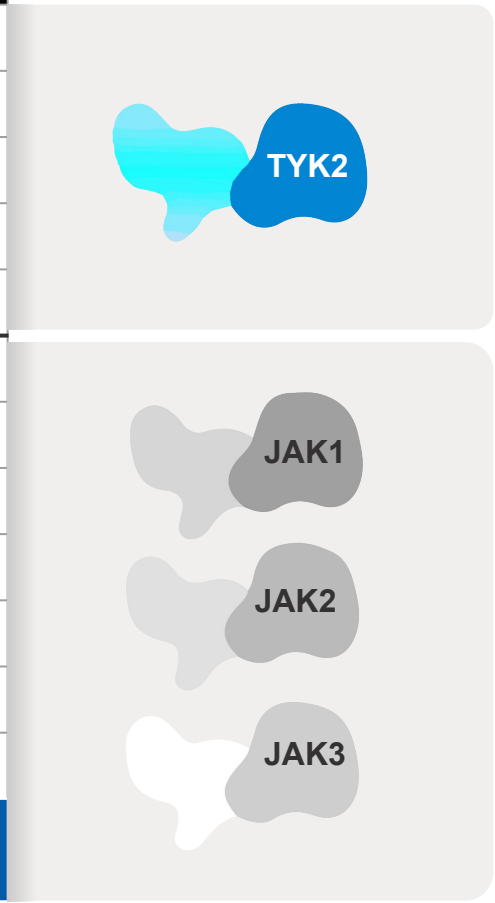
Free Energy Calculations Enabled Breakthroughs in Achieving Selectivity



NDI-034858 Shows Desired *In Vitro* Potency and Selectivity

Selective Inhibition of TYK2 vs. JAKs

In vitro potency / selectivity	NDI-034858
TYK2 JH2 Potency	0.0038 nM
TYK2 Function: IL-12 Inhibition	8.4 nM*
Interferon inhibition in human blood	51 nM*
Interferon inhibition in mouse blood	347 nM*
Interferon inhibition in rat blood	91 nM
JAK 1-3 kinase activity	>30,000 nM
JAK1 inhibition	5,000 nM
JAK2 inhibition	23,000 nM
JAK2 Function	>50,000 nM*
JAK1/3 Function	>50,000 nM*
PDE4 inhibition	>10,000 nM
hERG inhibition	>30,000 nM
87 target panel of enzymes, ion channels, receptors @10,000 nM	<50% inh for 85 targets**



- Highly potent and selective TYK2 allosteric inhibitor
- Nimbus Therapeutics' clinical data indicate compound is well-tolerated
- NDI-034858 shows desired target engagement in humans
- Phase 2b trials in moderate to severe psoriasis are ongoing*

Leit S. ACS Fall 2022 Conferences.

*ClinicalTrials.gov Identifiers NCT04999839 and NCT05153148.

Significant opportunity exists for new programs

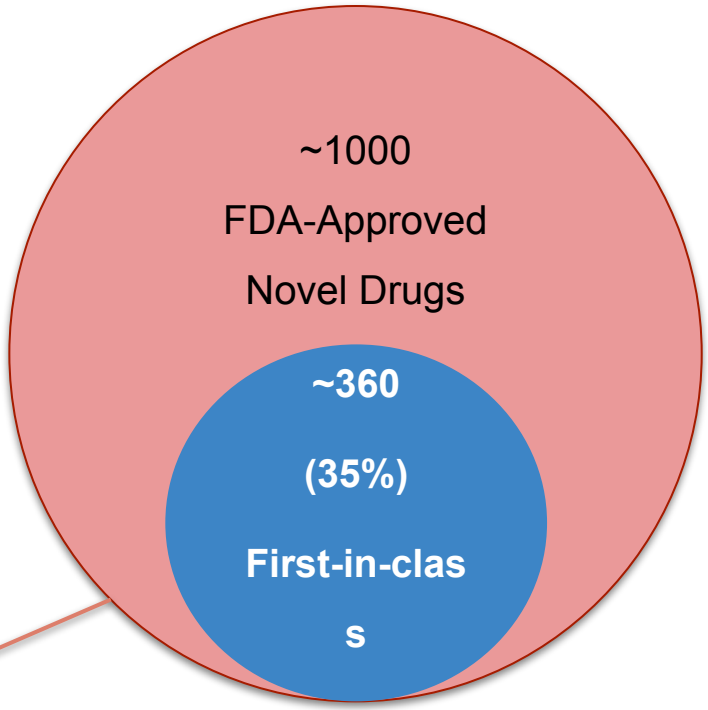
2020-2030
First in class MOAs that need tool molecules, leads and development candidates

22-40%
Druggable but not yet drugged

11%
Starting ligands available

1.4-4%
High quality drug-like compounds available

1987- 2019



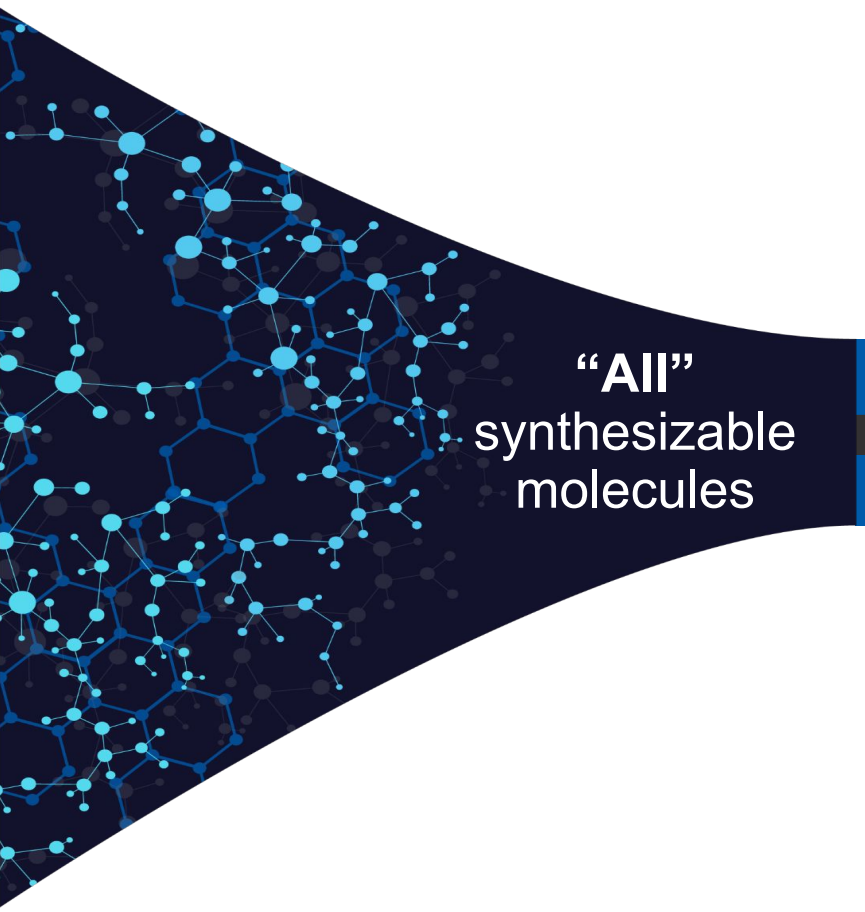
Remainder are **advance-in-class** and **addition-to-class** drugs

Santos, R., Ursu, O., Gaulton, A. *et al.* A Comprehensive map of molecular drug targets.

Nat Rev Drug Discov 16, 19–34 (2017). https://doi.org/10.1038/nrd.2016.230

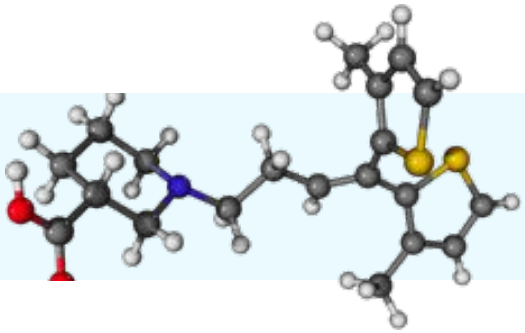
Our Vision for the Future of Drug Discovery

If we could calculate all the properties with perfect accuracy, designing drugs would have a **higher success rate**, be **faster** and **cheaper**, and would produce **higher-quality** molecules.



"All" synthesizable molecules

Select **THE** best molecule

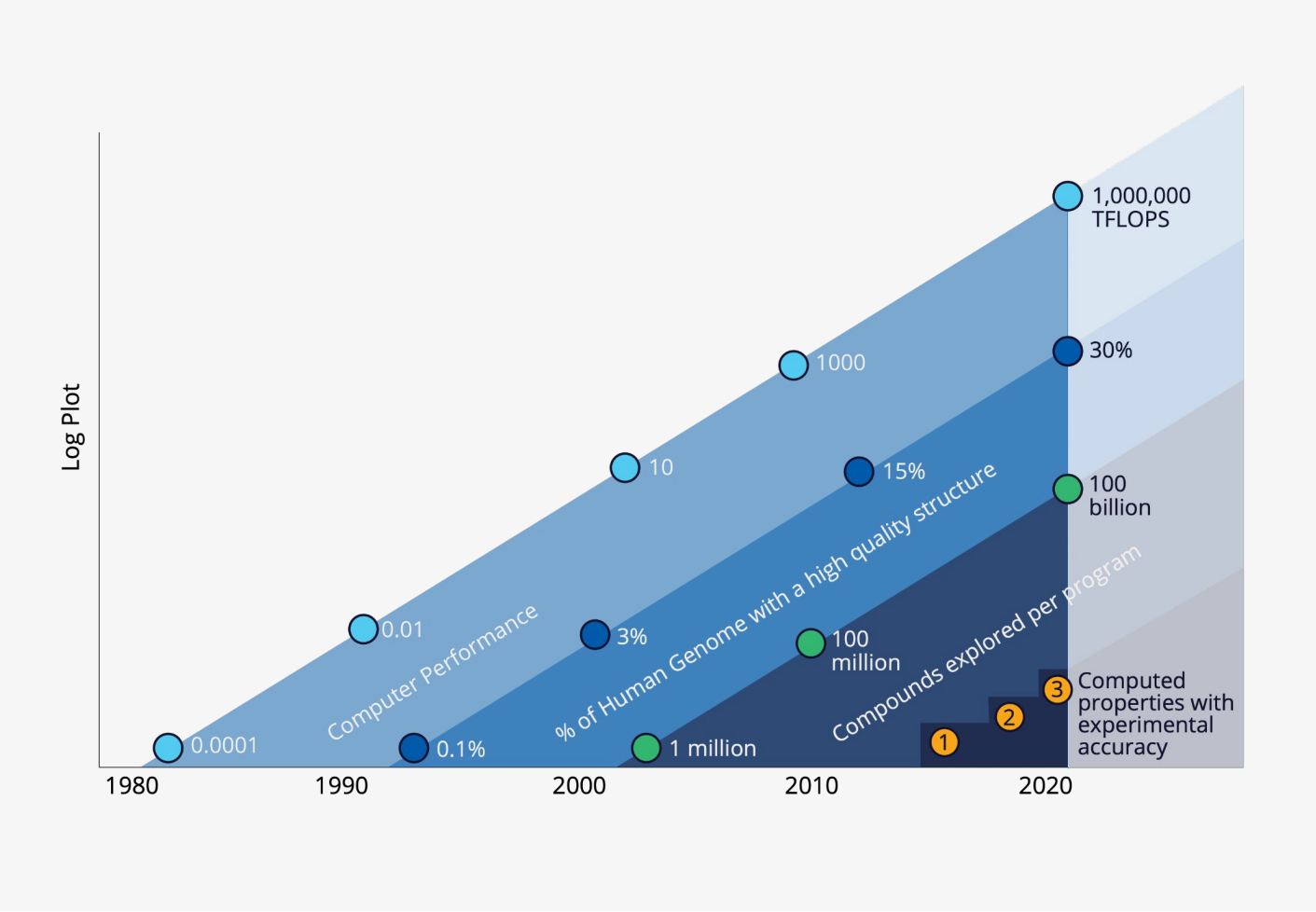


Potency	✓	Permeability	✓
Selectivity	✓	Drug-Drug Interactions	✓
Solubility	✓	Synthesizability	✓
Clearance	✓	Toxicity	✓

Bright future for digital and structure-based drug discovery

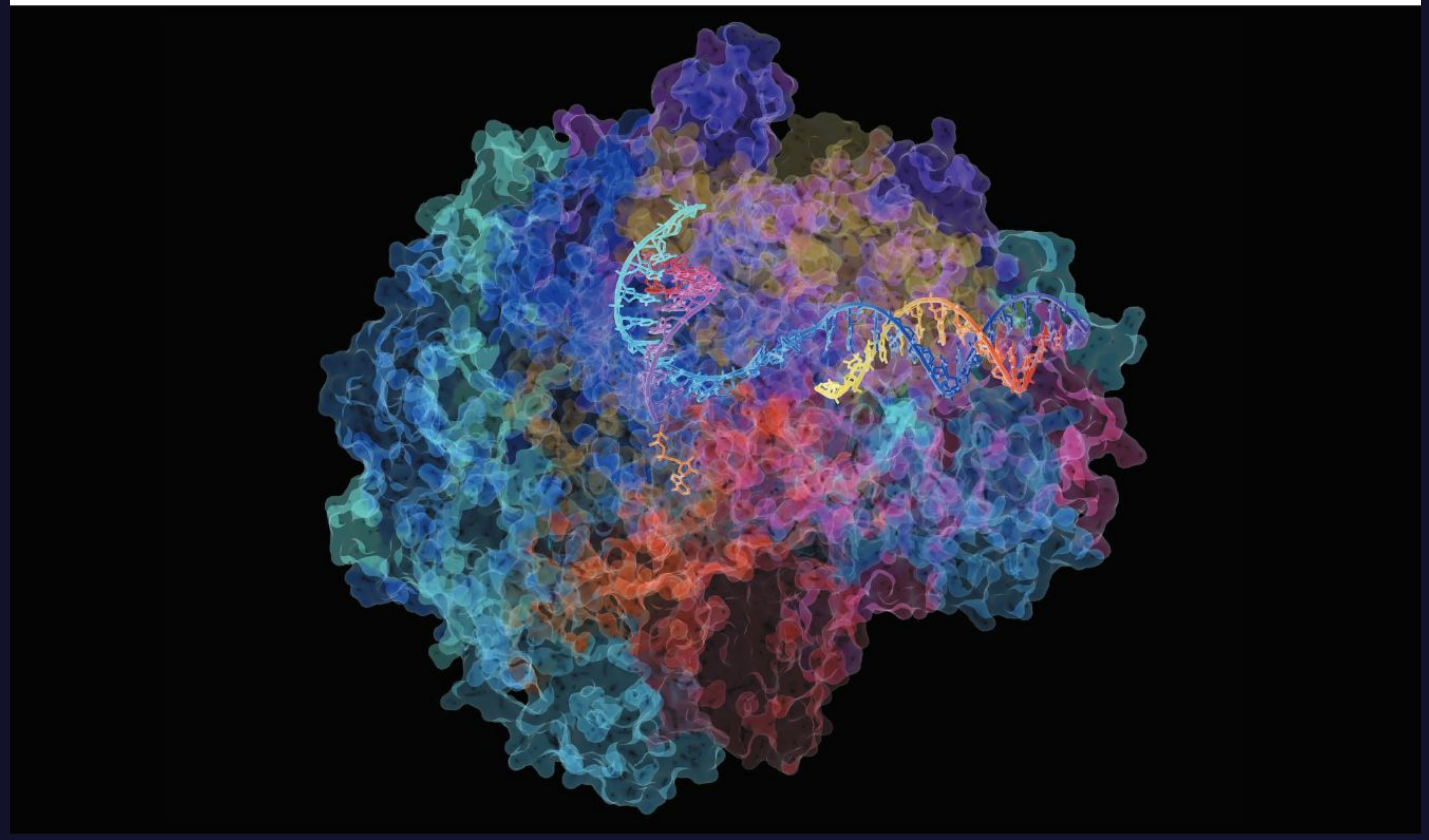
In the next decade…

- Computers **100x** faster
- High quality protein structures of **~75%** of the human genome
- Explore **trillions** of compounds per program
- **~10 molecular properties** computed with experimental accuracy
- High quality development candidates within **~1 year** of program launch



LET THE STRUCTURAL SYMPHONY BEGIN

Structural biologists are at last living the dream of visualizing macromolecules to uncover their function. But it means integrating different technologies, and that's no easy feat.



Courtesy of Dr Enrico Malito @Schrodinger
https://www.nature.com/articles/536361a

Acknowledgements

Schrödinger Development and LiveDesign Teams

Schrödinger Therapeutics Group

- MALT1 and TYK2 Discovery Teams
- SGR-1505 Development Team

Nimbus Therapeutics

- TYK2 Discovery Team
- NDI-034858 Development Team