

HARNESSING 100,000 COMPUTERS AROUND THE WORLD TO PROBE KINASE ACTIVATION, INHIBITION, AND RESISTANCE



John D. Chodera

MSKCC Computational and Systems Biology Program

<http://www.choderalab.org>

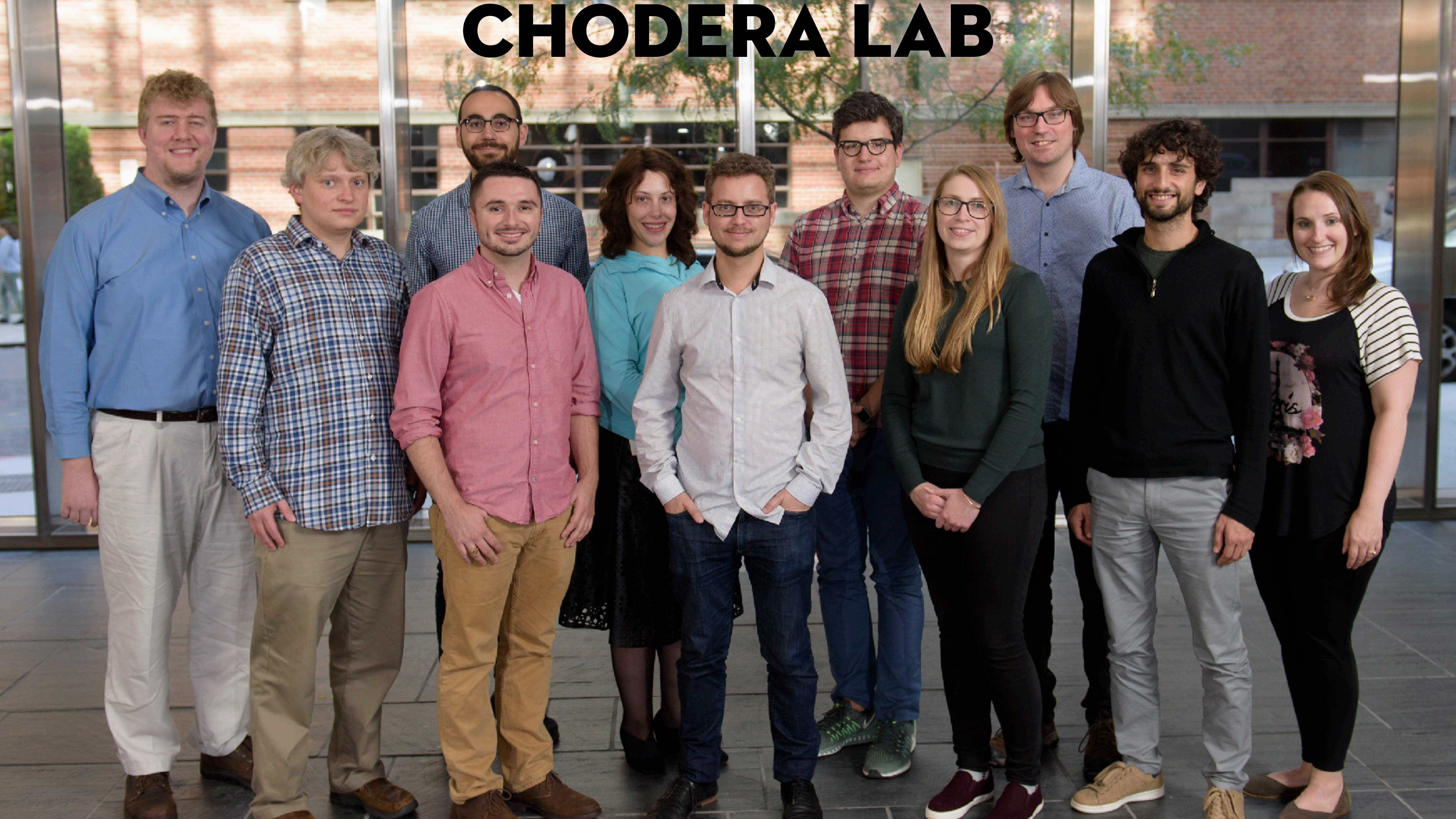
DISCLOSURES:

- Scientific Advisory Board, OpenEye Scientific

All funding: <http://choderalab.org/funding>

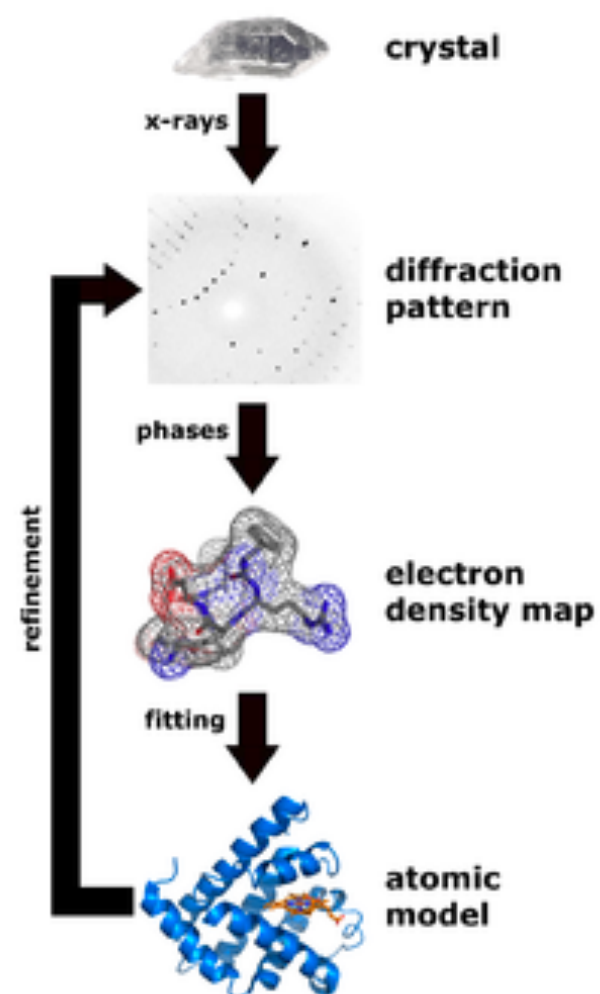
15 Oct 2018 - SKI Talk - New York, NY

CHODERA LAB



STRUCTURAL BIOLOGY HAS PROVIDED US WITH ATOMISTIC INSIGHT INTO BIOLOGY

X-ray crystallography

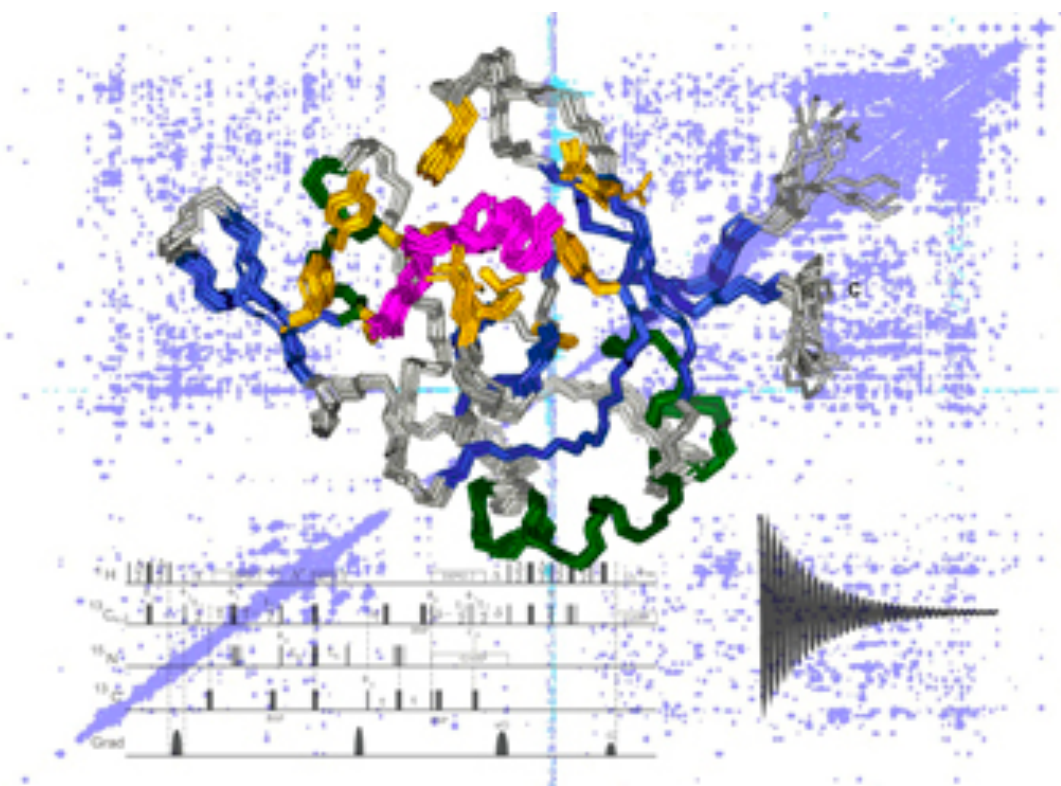


https://en.wikipedia.org/wiki/X-ray_crystallography



Brookhaven National Laboratory - NSLS II

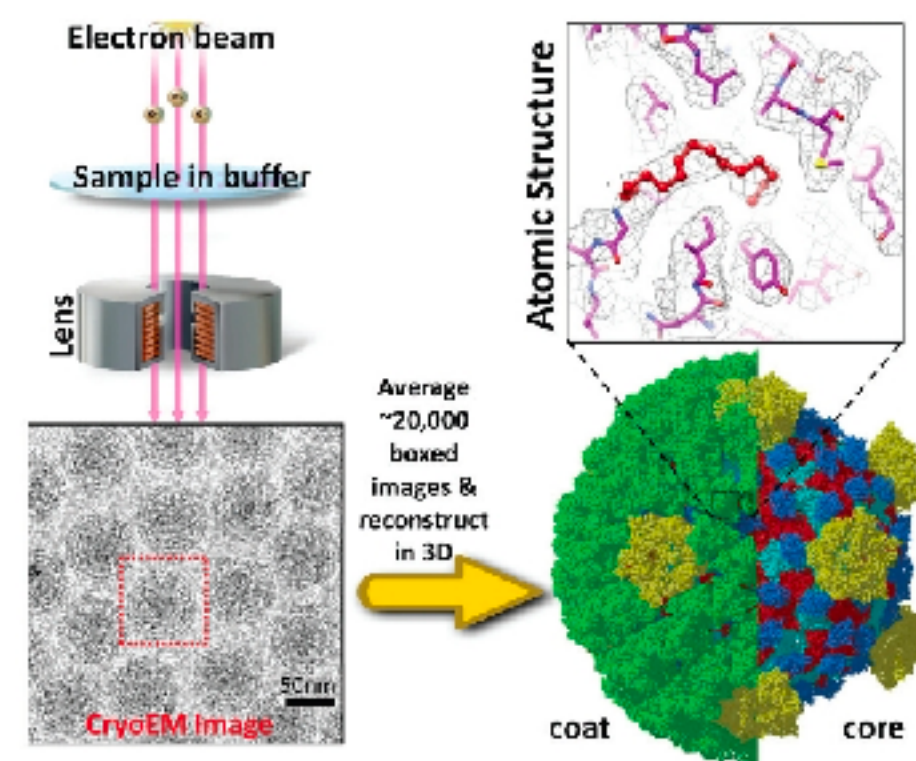
NMR



<http://events.embo.org/11-nmr/>



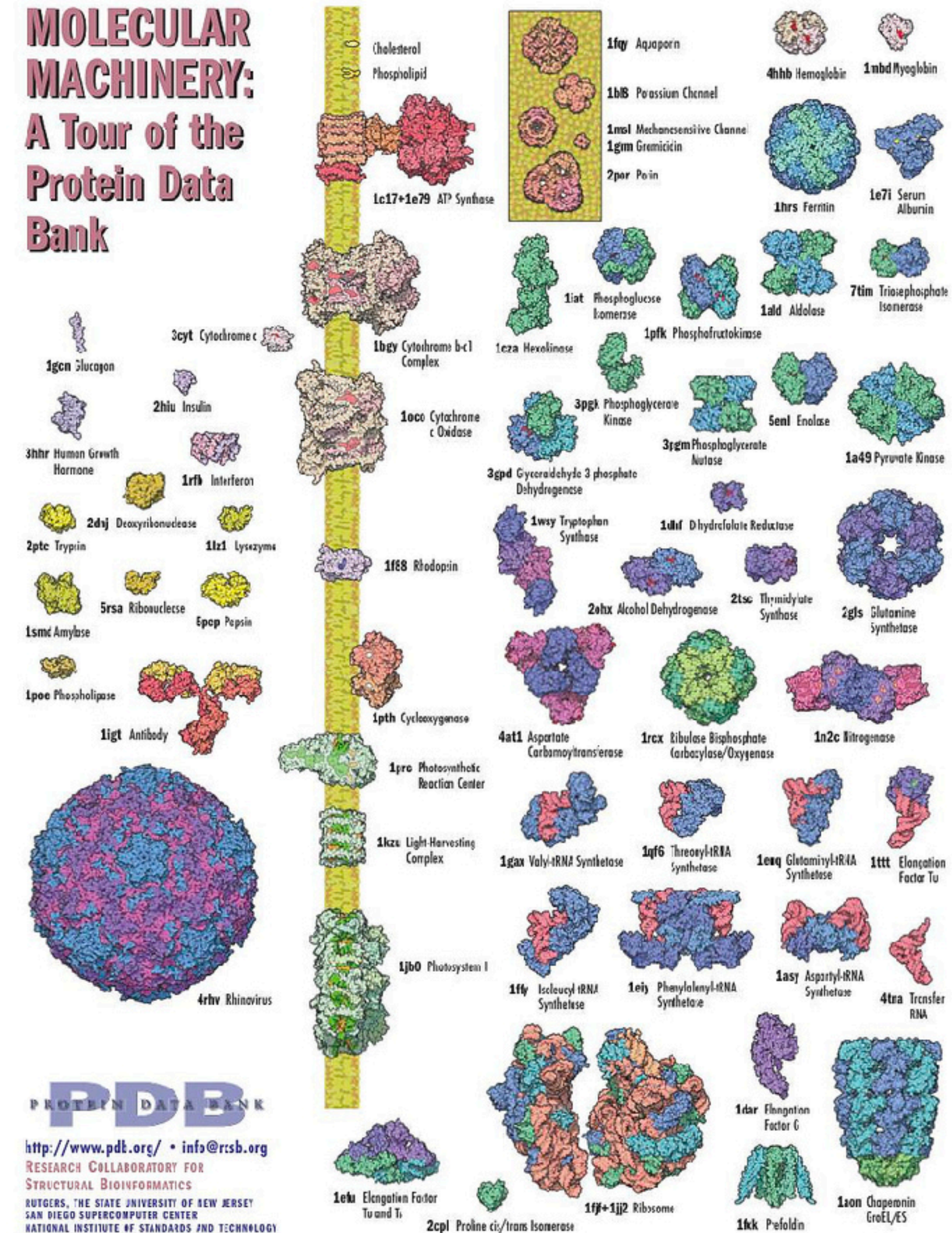
cryo-EM



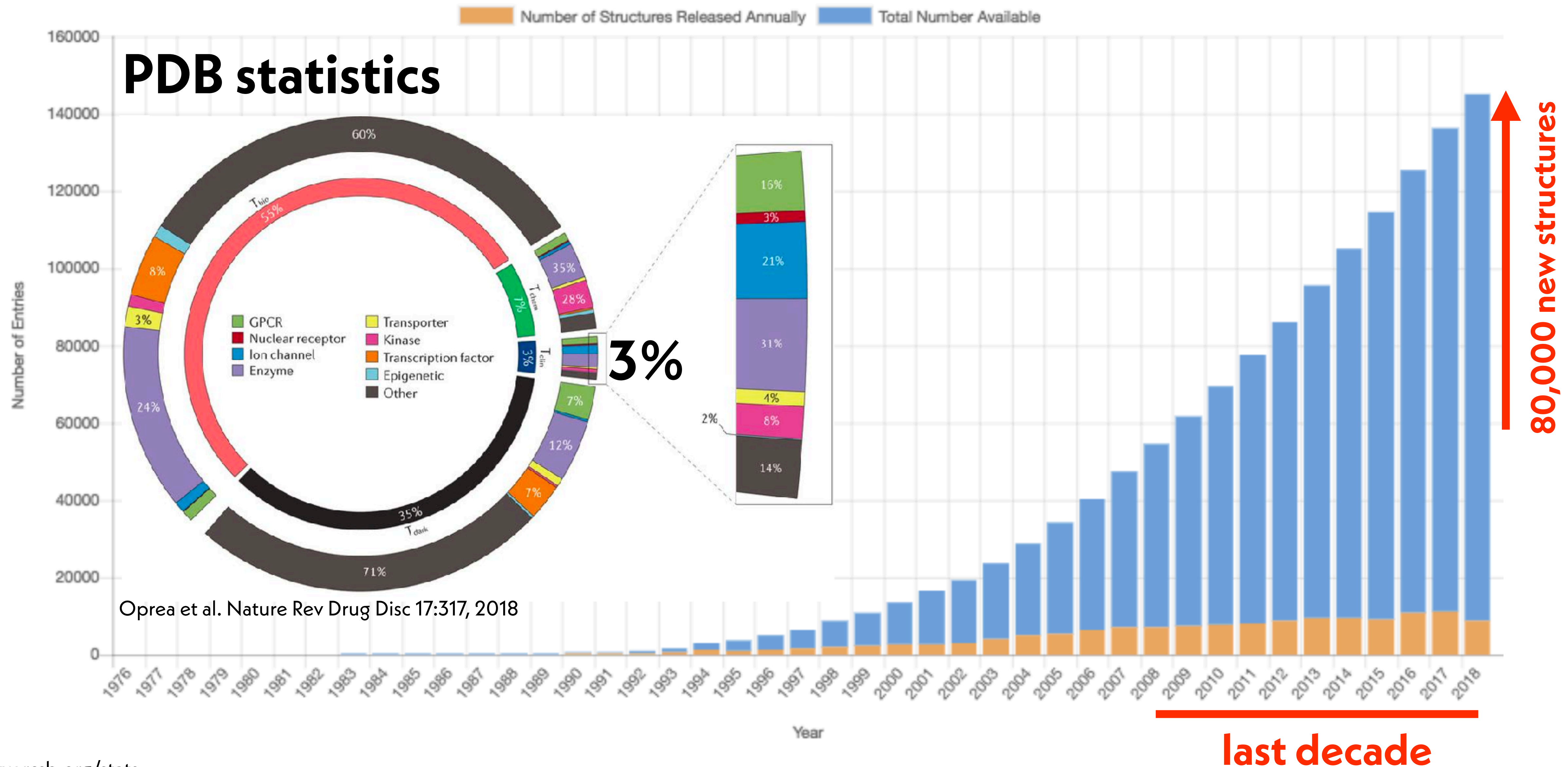
<https://www.americanlaboratory.com/914-Application-Notes/>



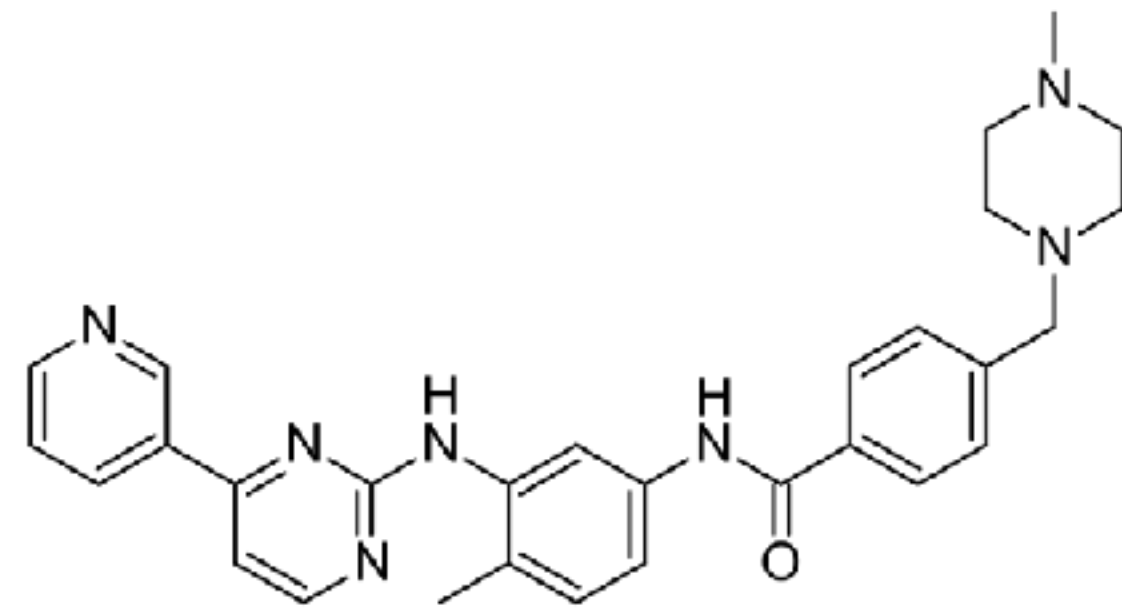
MOLECULAR MACHINERY: A Tour of the Protein Data Bank



THE LAST DECADE HAS PRODUCED AN ENORMOUS NUMBER OF BIOMOLECULAR STRUCTURES

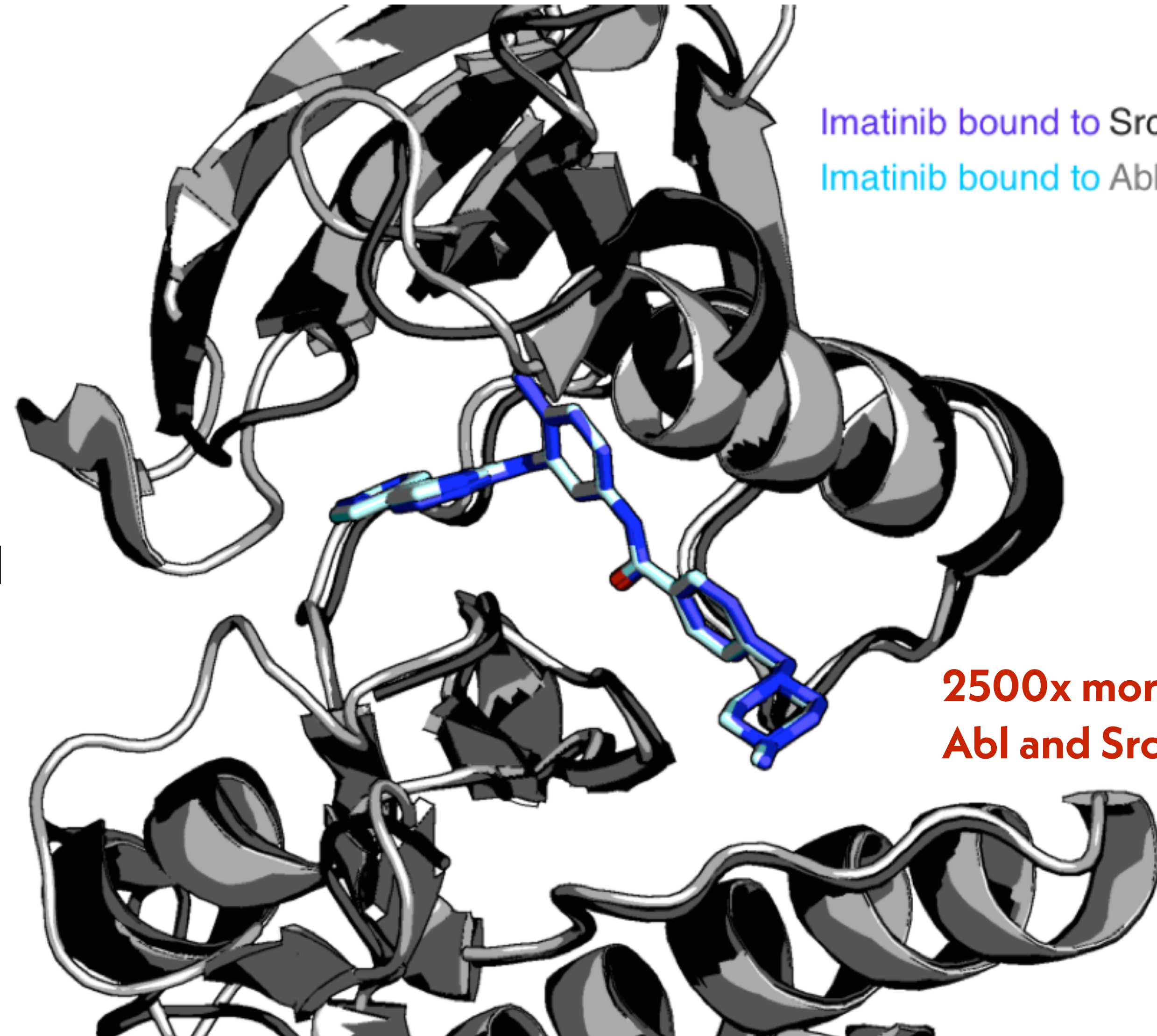


STRUCTURE TELL ONLY PART OF THE STORY: DYNAMICS IS KEY TO UNDERSTANDING FUNCTION AND DISEASE



imatinib

approved by USFDA in 2001

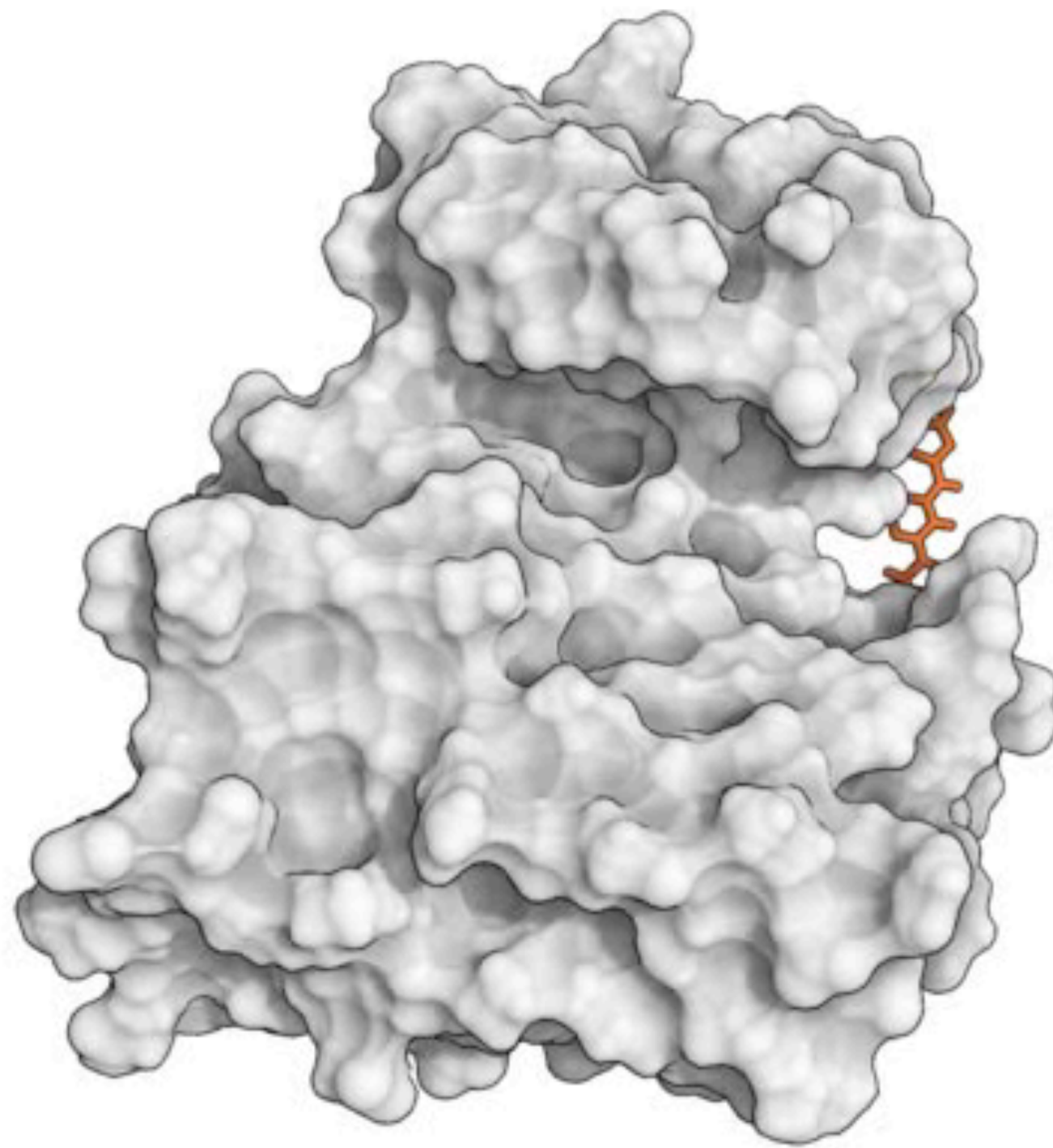


Imatinib bound to Src

Imatinib bound to Abl

2500x more selective for Abl over Src
Abl and Src share 54% sequence identity

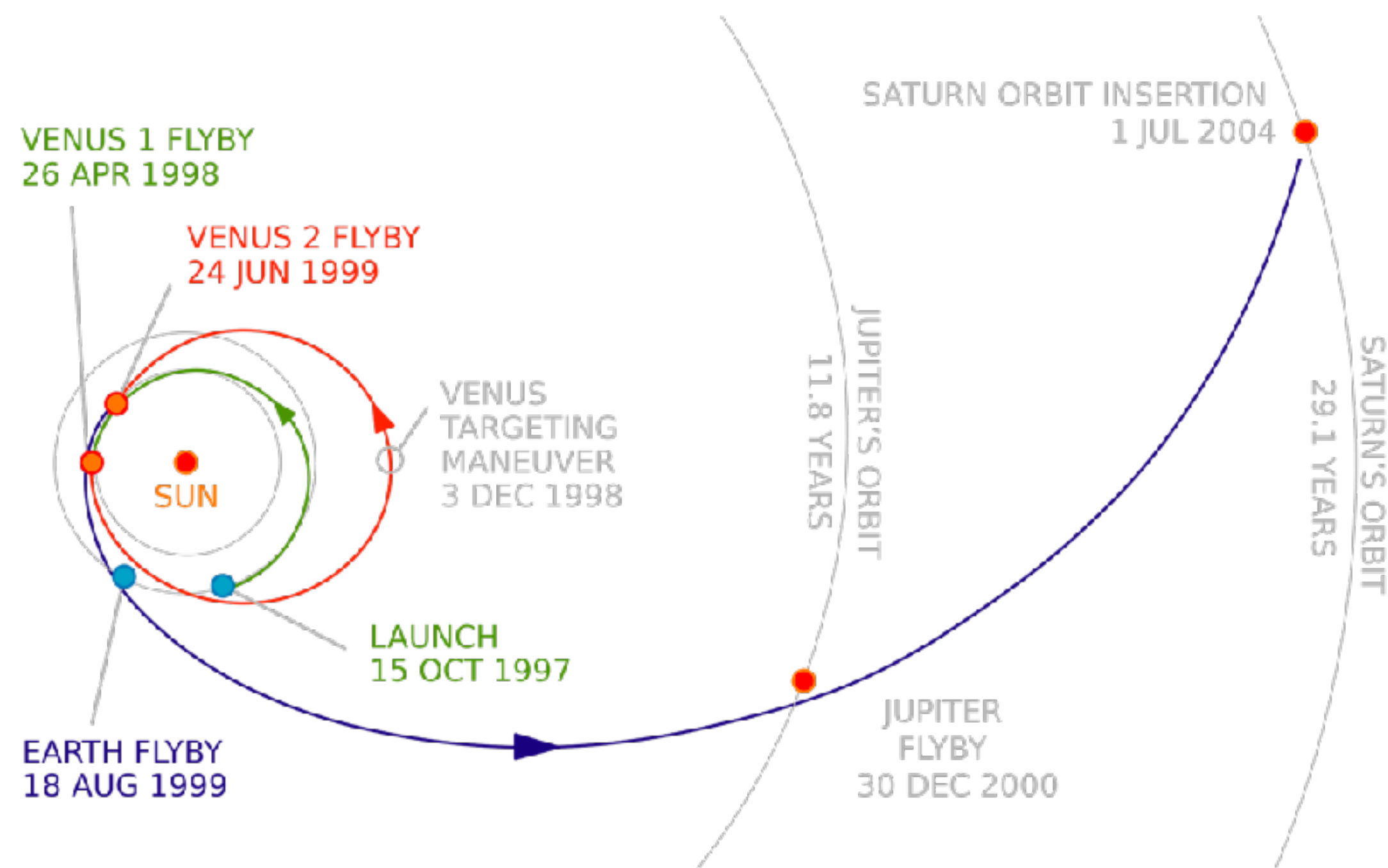
MOLECULAR MODELING AND SIMULATION CAN BREATHE LIFE INTO STATIC STRUCTURES



$$E_{total} = \underbrace{\sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]}_{\text{Bonded}} + \underbrace{\sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]}_{\text{Non-bonded}}$$

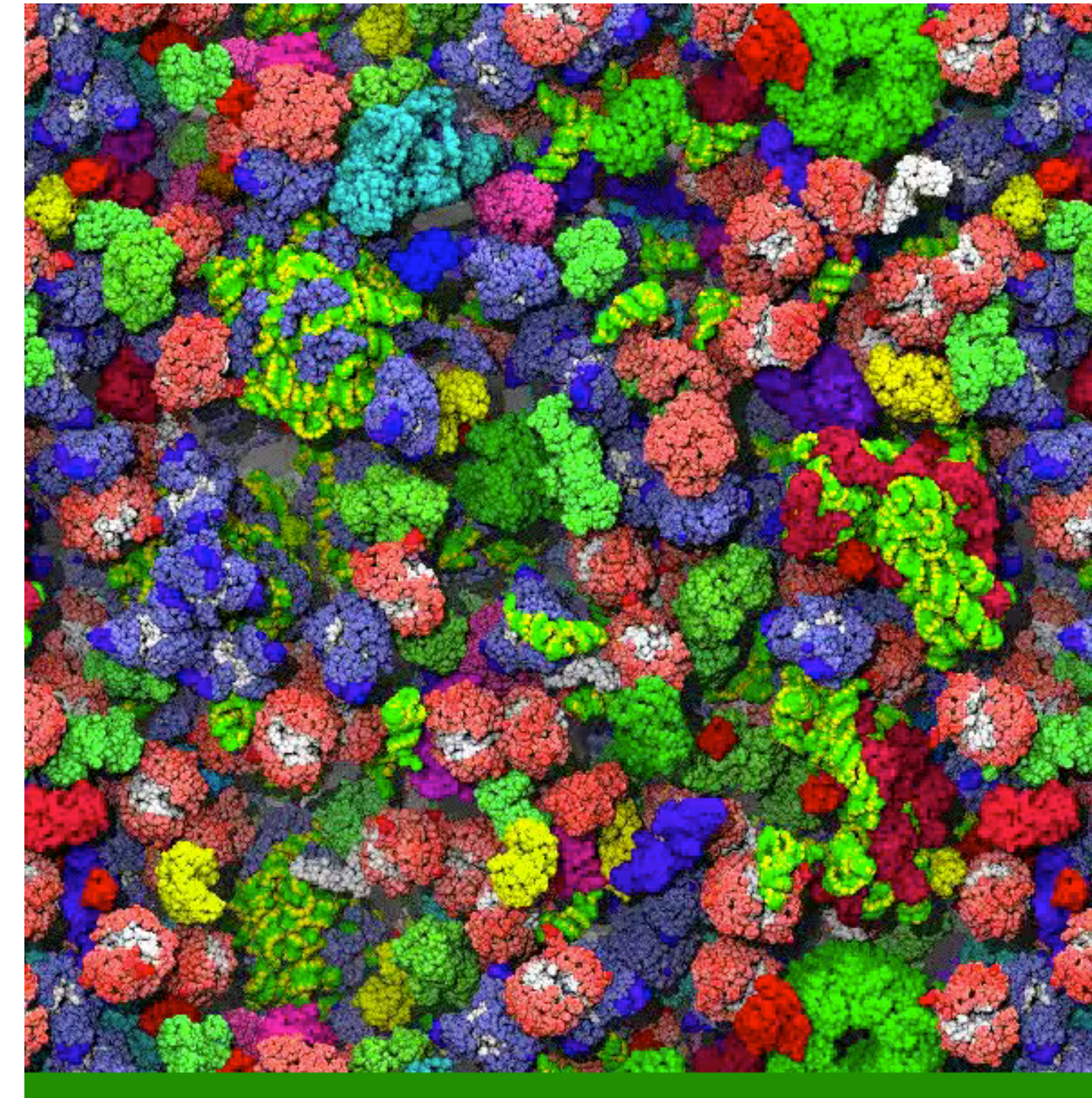
ATOMISTIC MOLECULAR SIMULATION POSSESSES UNIQUE CHALLENGES

spacecraft trajectory



<https://solarsystem.nasa.gov/resources/11776/cassini-trajectory/>

interior of a cell

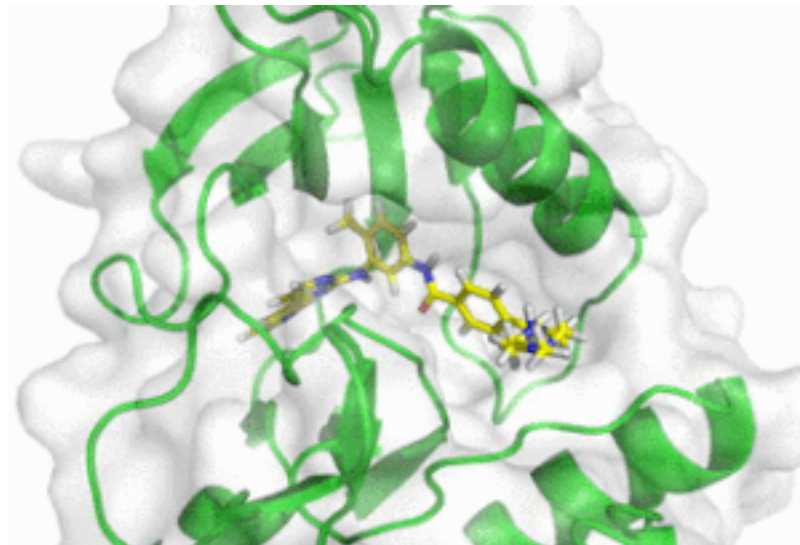


McGuffee and Elcock. PLoS Comput Biol. 2010 Mar 5;6(3):e1000694.e

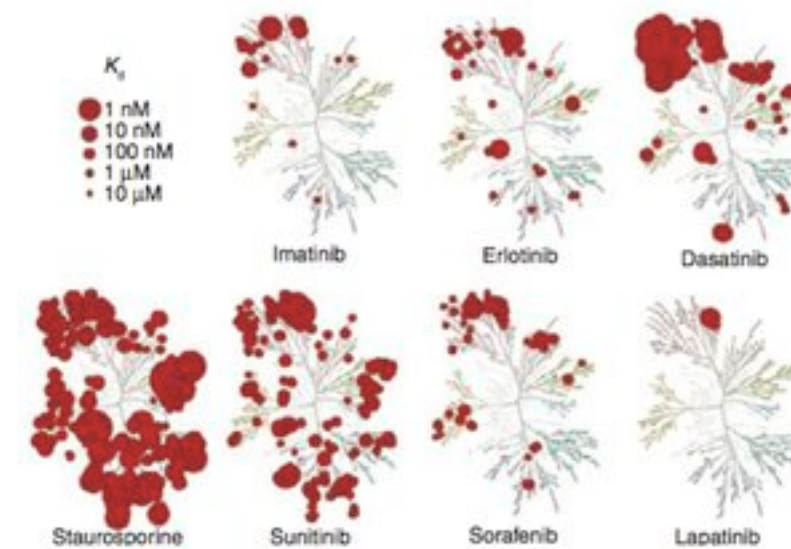
Making sense of stochastic biomolecular behavior requires the development of new technologies

CHODERA LAB

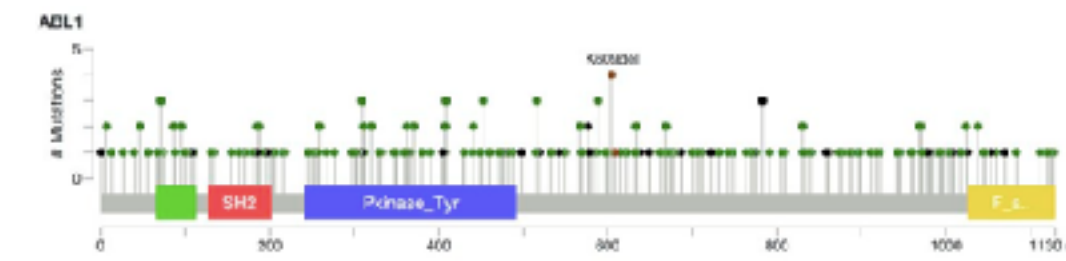
How can biophysical modeling play a major role in the era of cancer genomics?



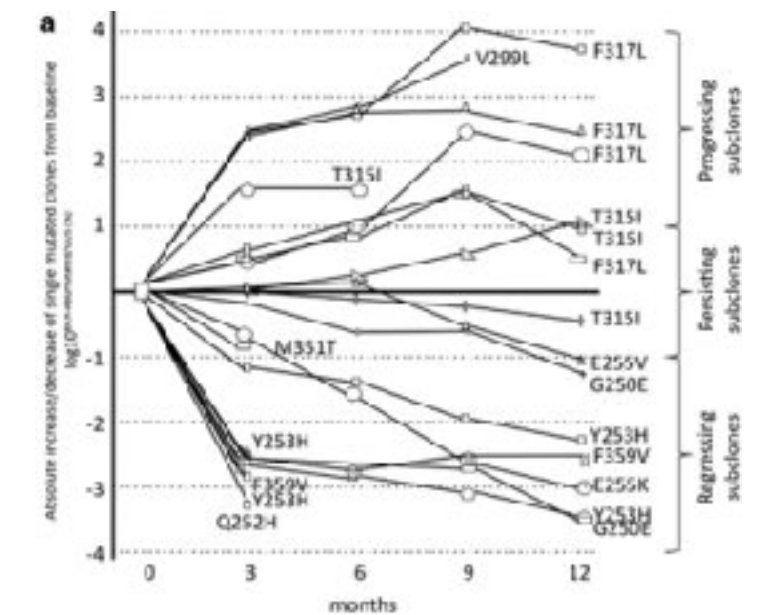
**SELECTIVE INHIBITOR DESIGN:
TARGETS/ANTITARGETS**



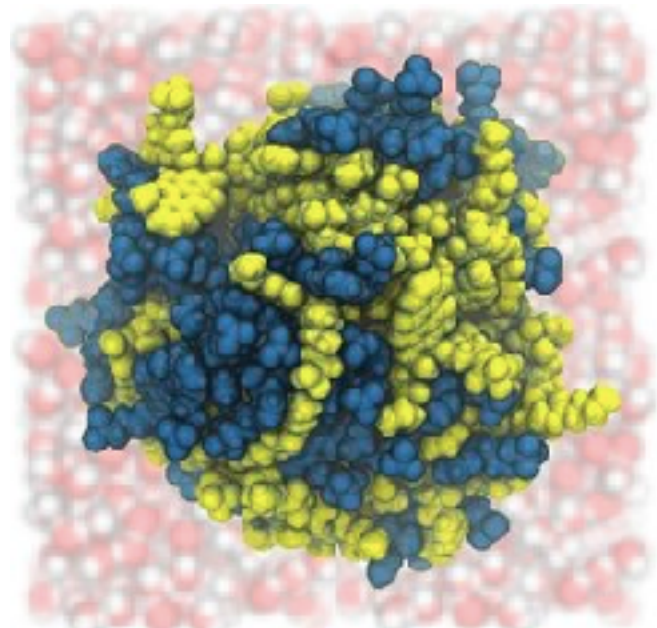
**KINASE INHIBITOR
SELECTIVITY**



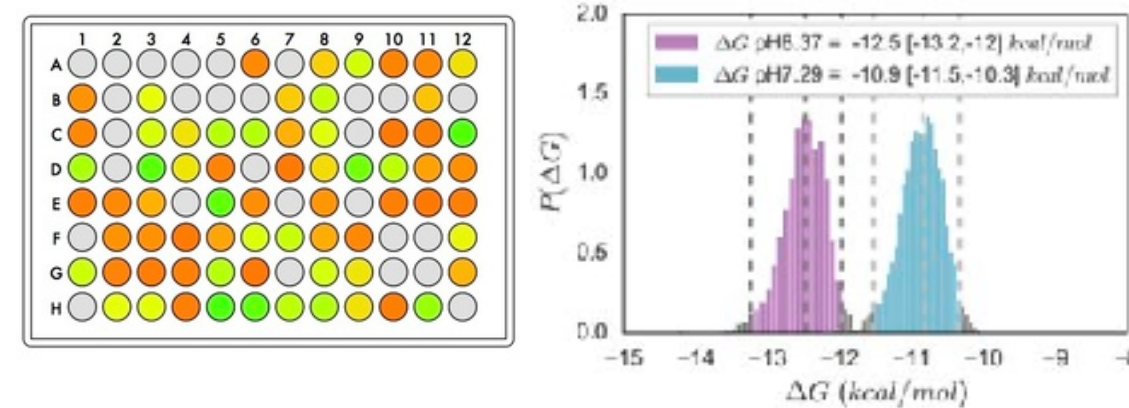
**PREDICTING DRUG
SENSITIVITY/RESISTANCE**



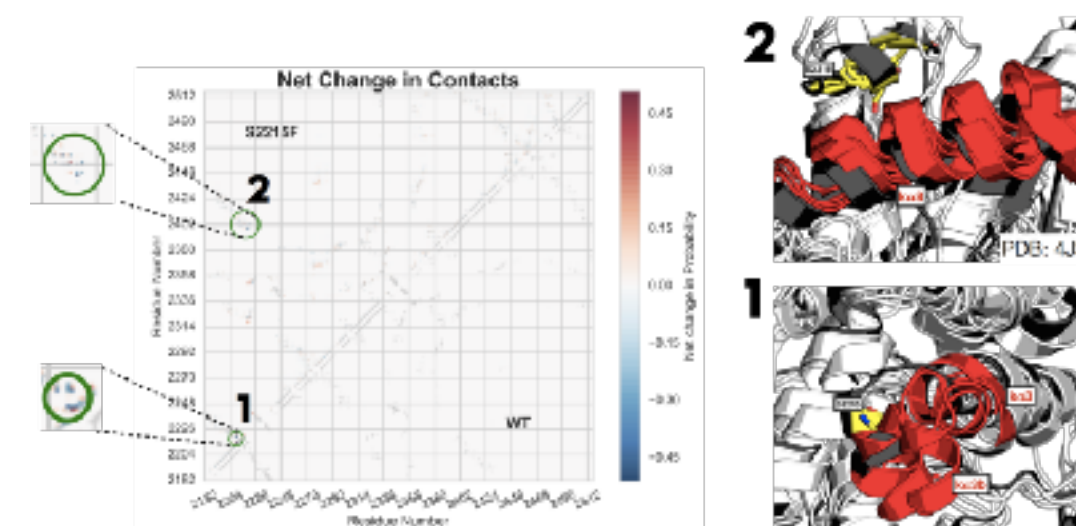
**ANTICIPATING
DRUG RESISTANCE**



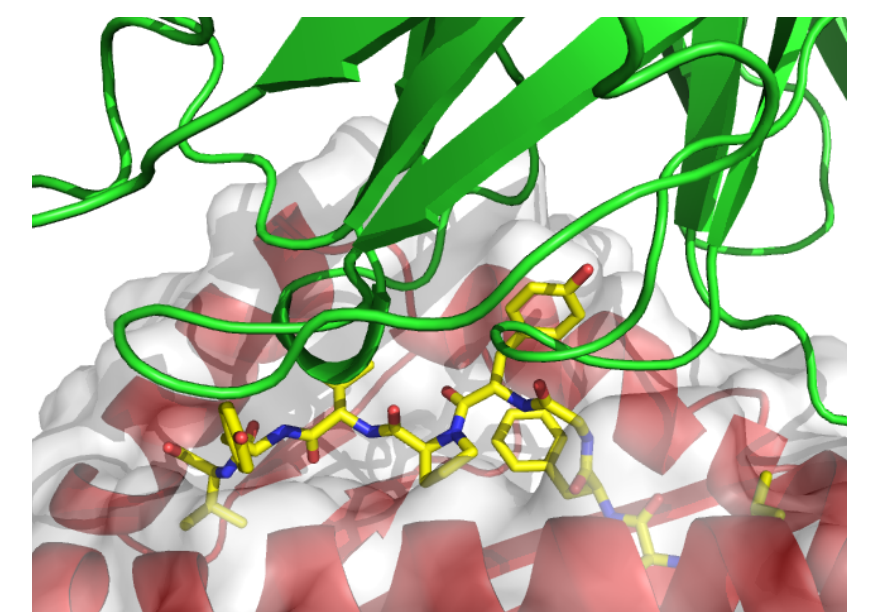
**NOVEL DRUG DELIVERY
MODALITIES**



**AUTOMATED BIOPHYSICAL
ASSAYS AND INFERENCE**



**MECHANISMS OF
ONCOGENIC ACTIVATION**



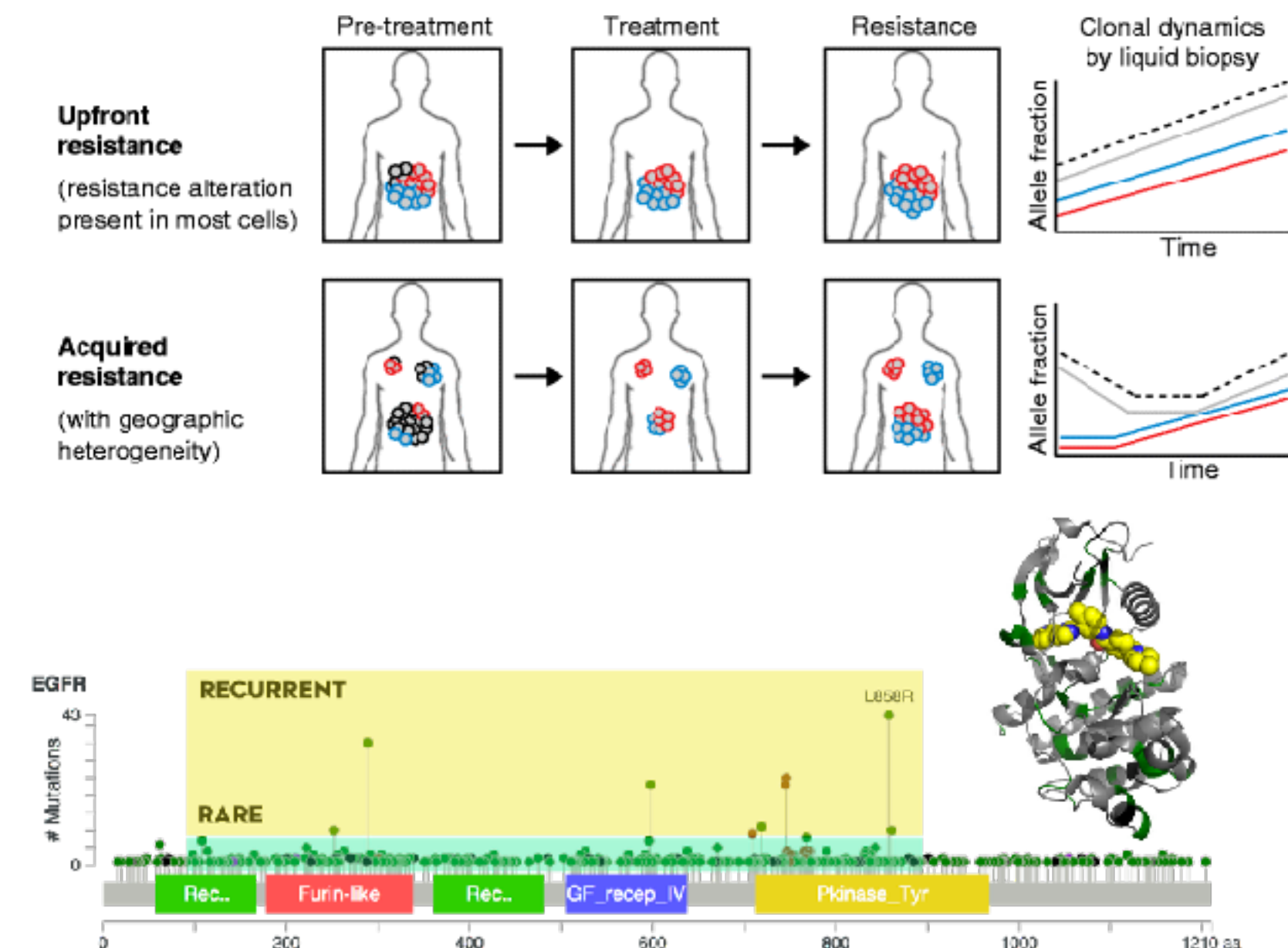
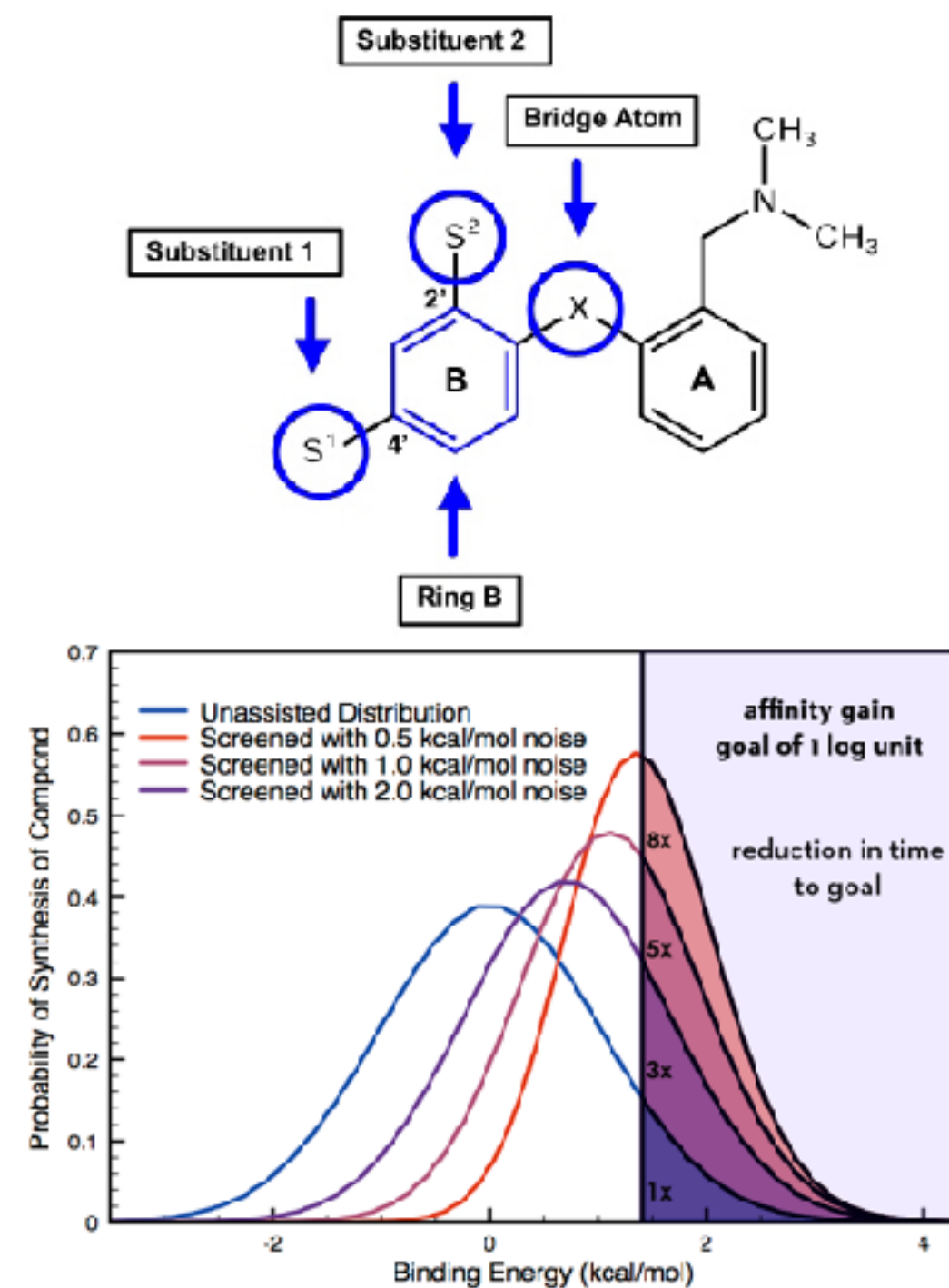
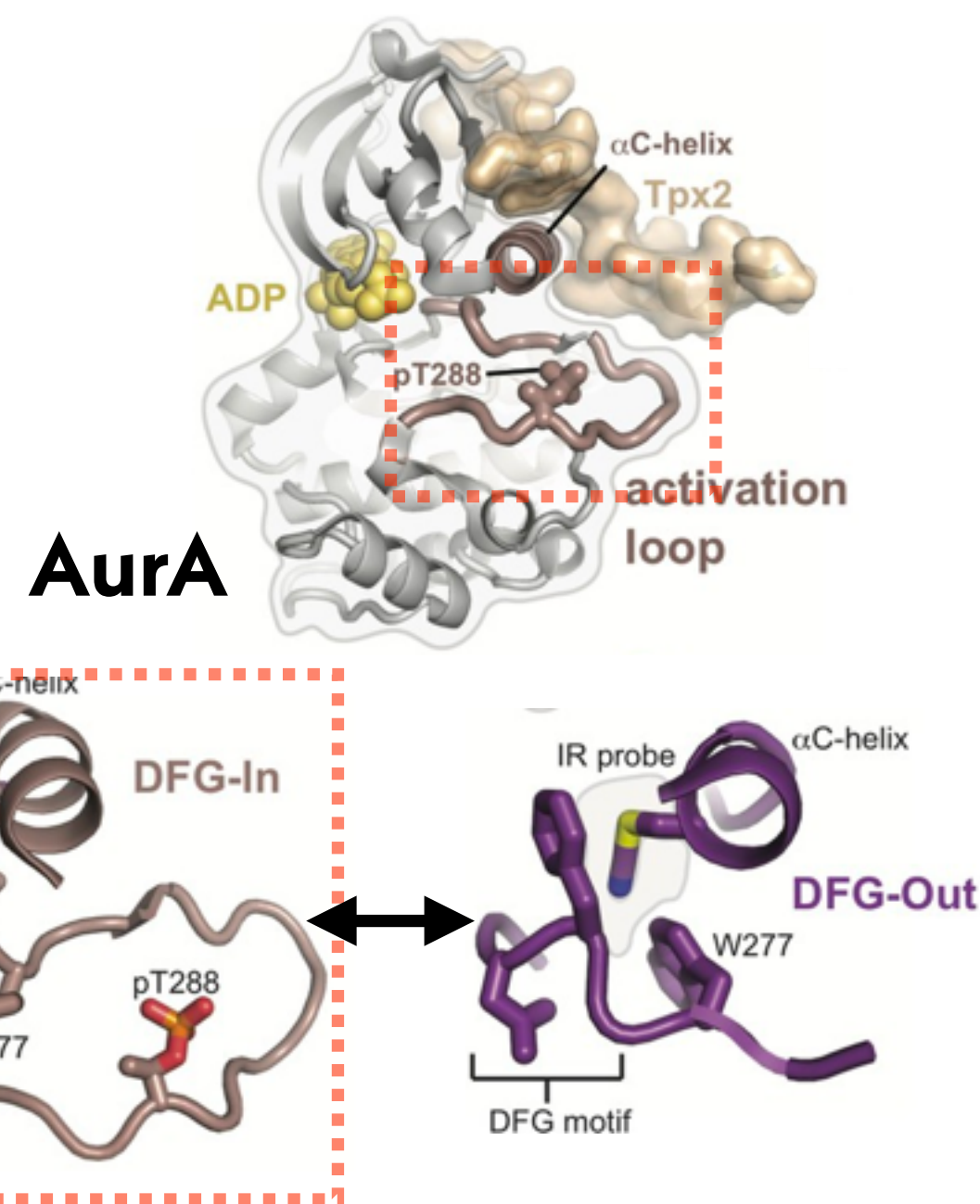
**CANCER
IMMUNOTHERAPY**

NEW APPROACHES TO MOLECULAR SIMULATIONS ARE A POWERFUL WAY TO ADDRESS QUESTIONS IN CANCER RESEARCH AND THERAPY

Understand molecular mechanisms of function and disease

Improve drug discovery success rates

Predict clinical response and emergence of drug resistance



Chodera lab papers:
 Cell Chem Biology, in press.
 eLife 7:e32766, 2018
 Nature 559:125, 2018
 Nature Chem Biol 13:494, 2017; 13:402, 2017
 J Clin Invest 126:3527, 2016

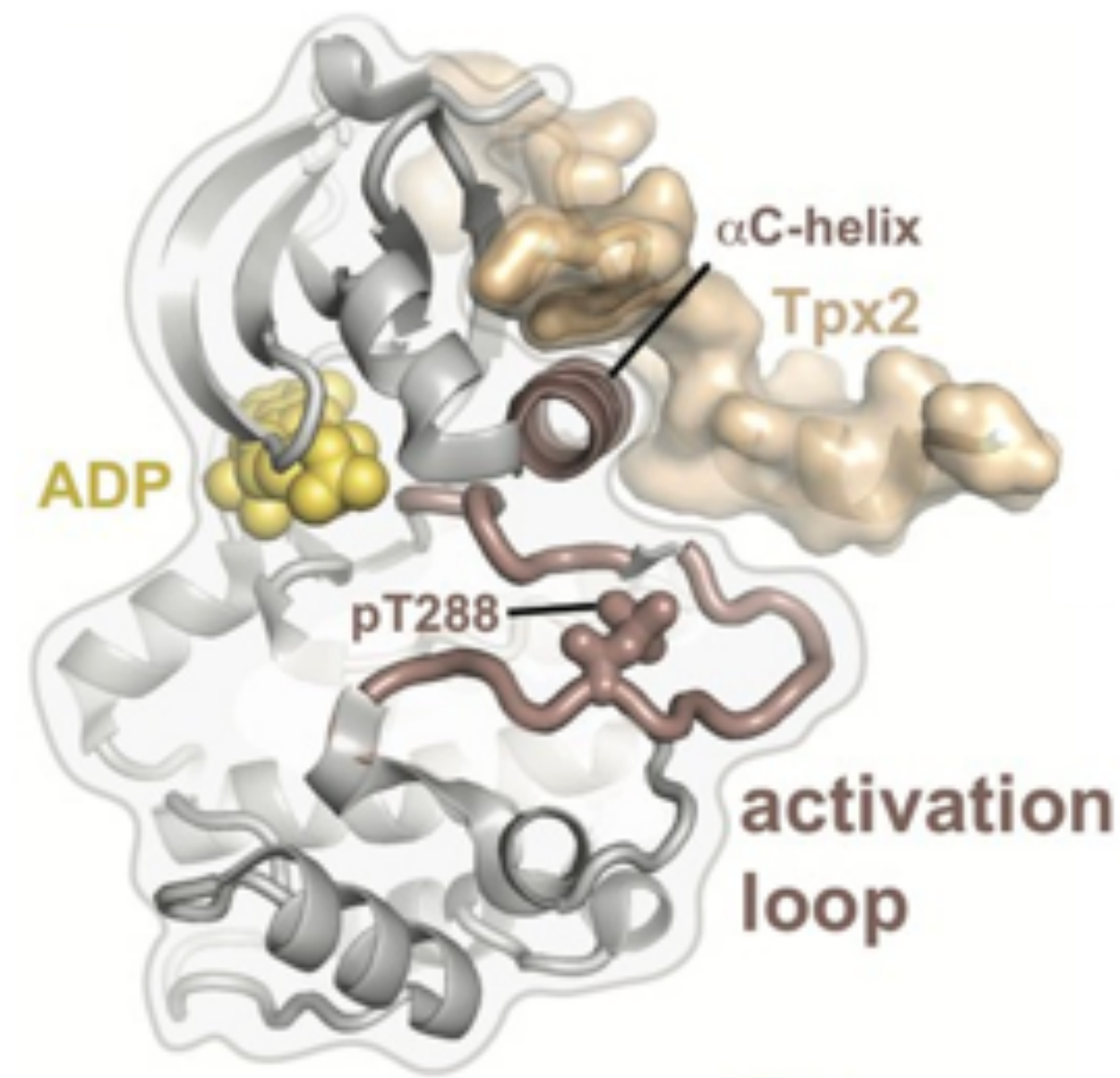
Chodera lab papers:
 Cell Chem Biol, in press
 Nature Materials 17:361, 2018
 J Chem Eng Data 62:1559, 2017
 J Comput Aid Mol Des 30:945, 2016; 29:1073, 2015
 J Phys Chem B 122:5466, 2018; 22:5579, 2018; 119:12912, 2015

Chodera lab papers:
 Communications Biology 1:70, 2018
 Biochemistry 57:4675, 2018
 J Clin Invest 126:3527, 2016

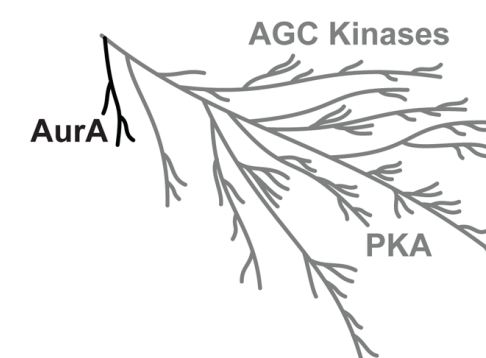
BIOPHYSICAL MODELING CAN REVEAL KEY INSIGHTS INTO THE MOLECULAR MECHANISM OF FUNCTION AND DISEASE

Non-canonical activation in AurA

essential for mitosis
multiple conformations
relevant to regulation
implicated in multiple cancers

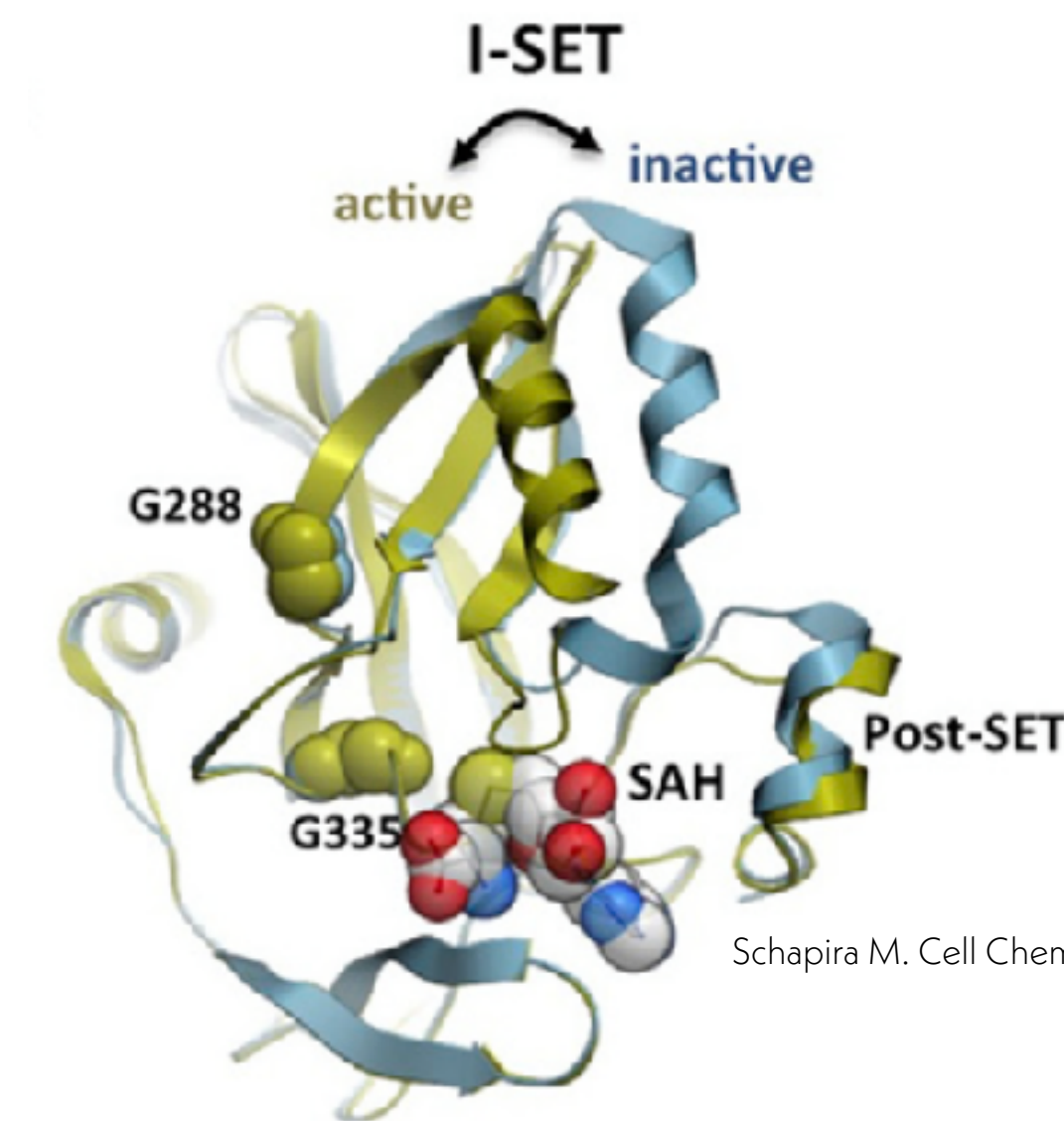


AurA

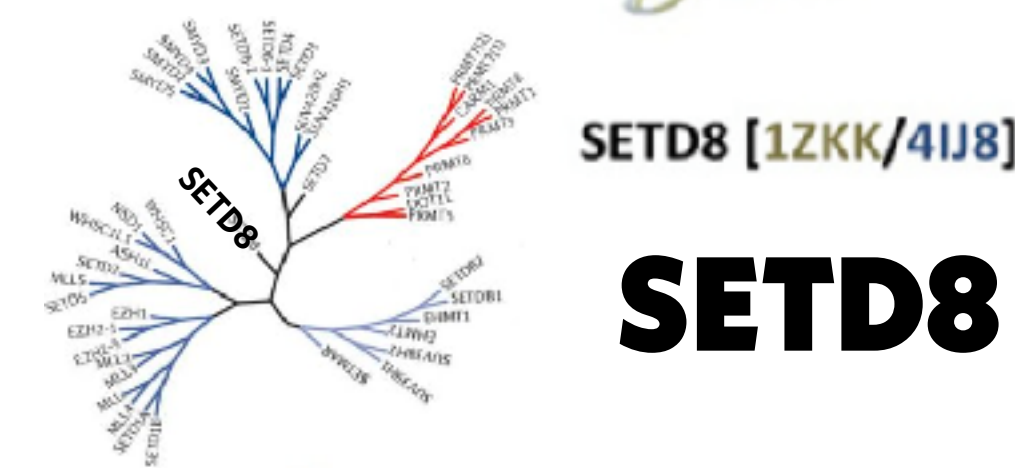


eLife 7:e32766, 2018
Nature Chem Biol 13:402, 2017

Functional motions in SETD8



Schapira M. Cell Chem Biol 23:1067-1076, 2016.



SETD8

methylates histone H4K20
flexible to accommodate
various substrates
epigenetic cancer target

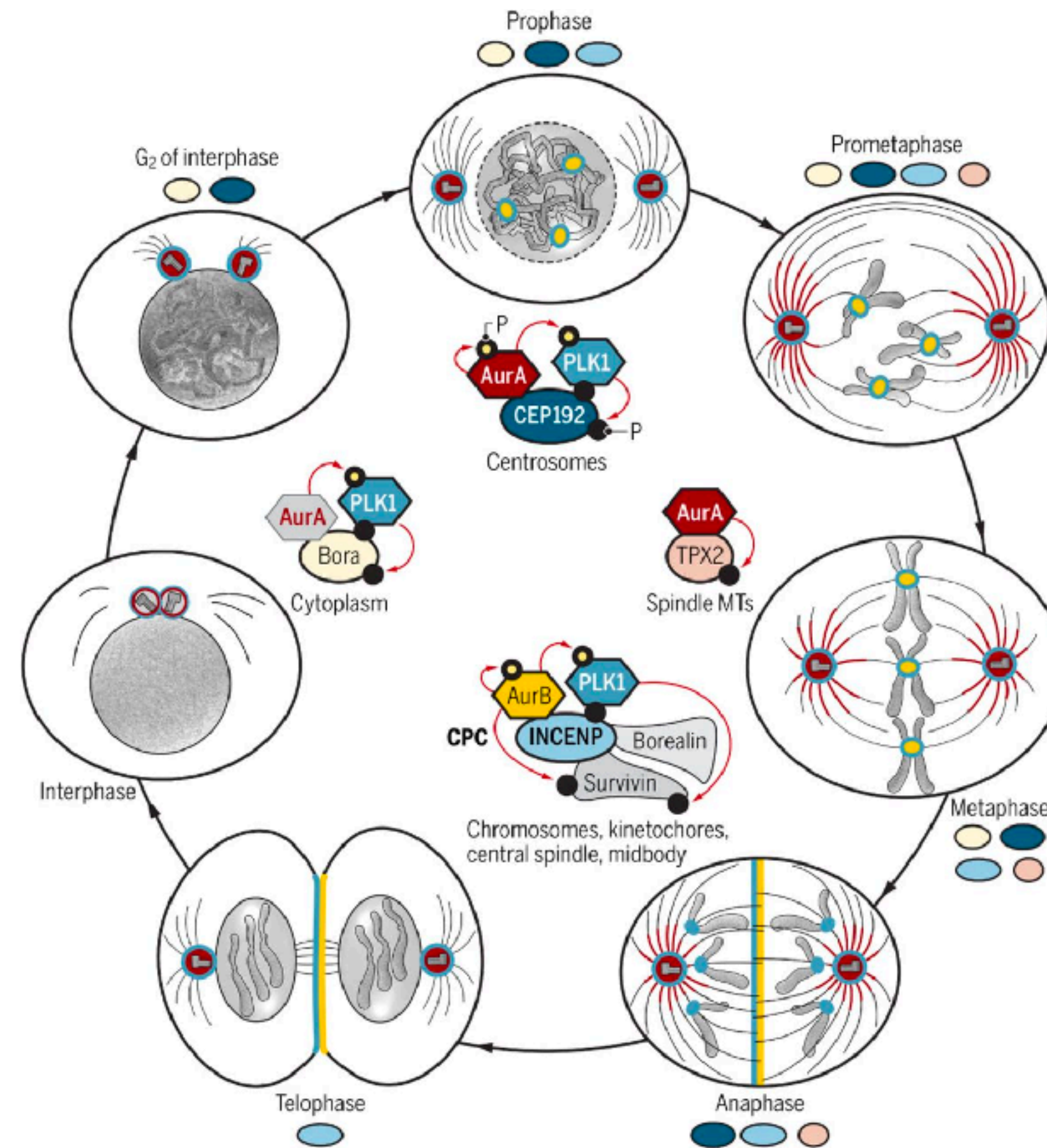
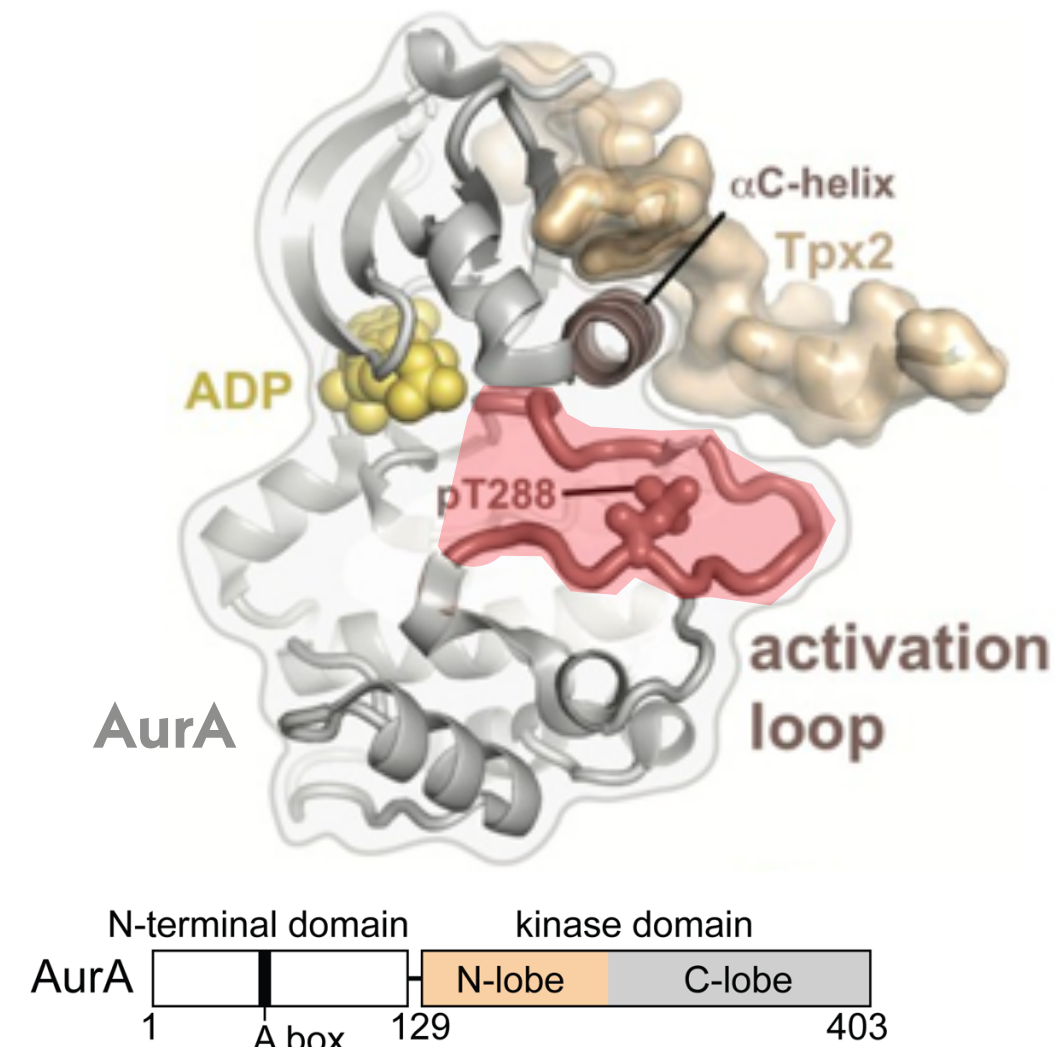
NICK LEVINSON
UMN

MINKUI LUO
SKI

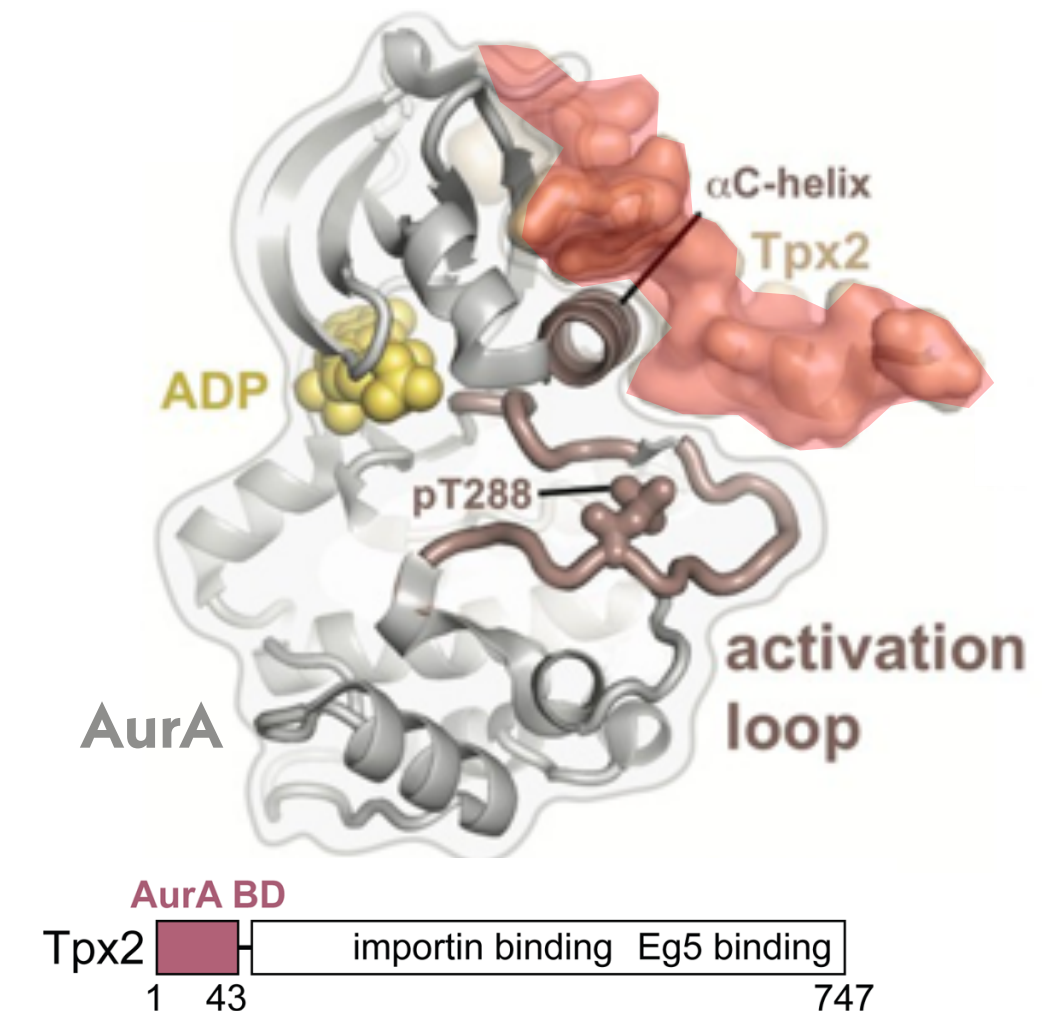
bioRxiv: <https://doi.org/10.1101/438994>

HUMAN KINASE AURORA A EXISTS IN TWO DISTINCTLY REGULATED POOLS

centrosomal pool
activated by phosphorylation



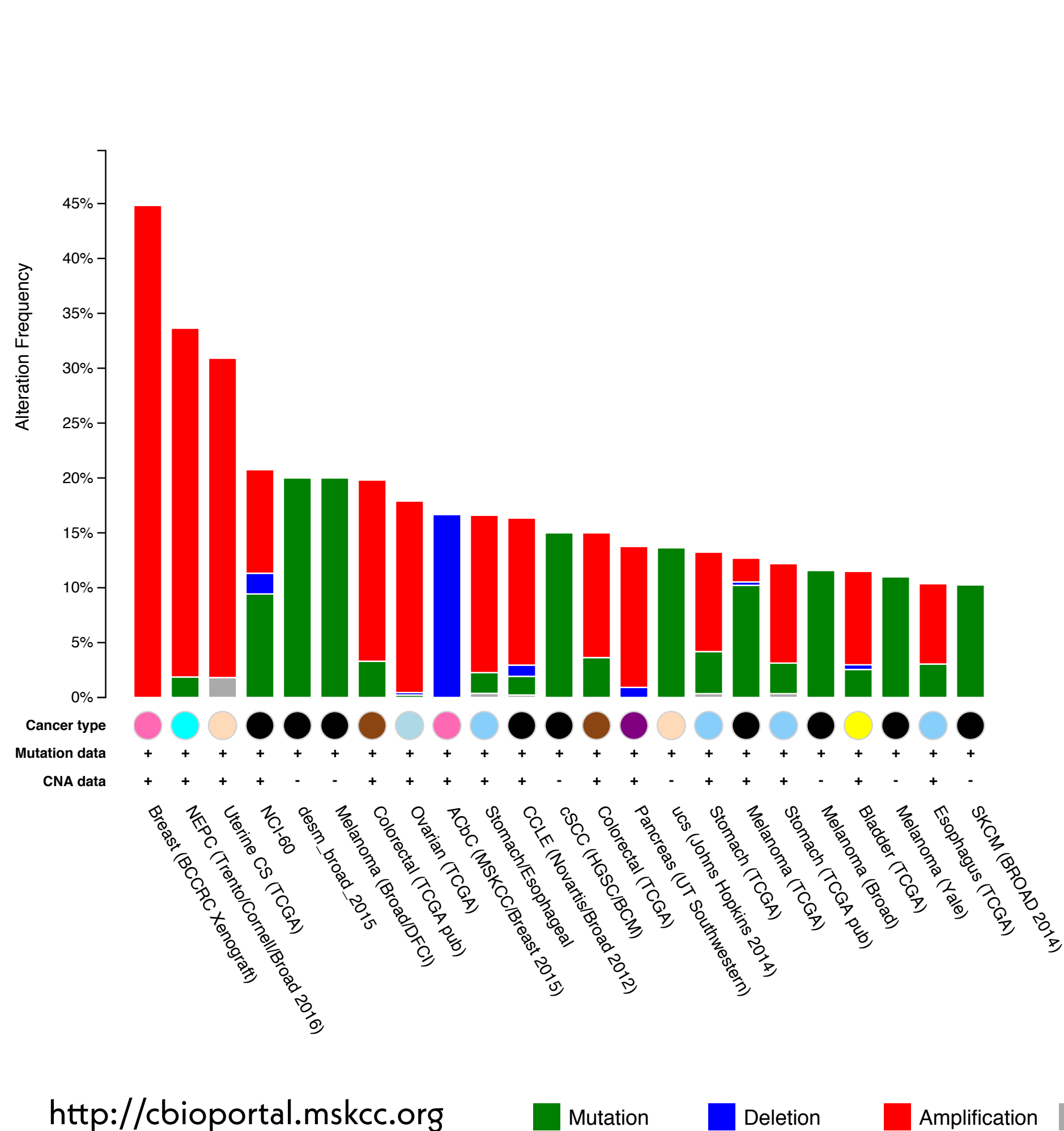
spindle microtubule-associated pool
kept dephosphorylated by PP6
activated by Tpx2 binding



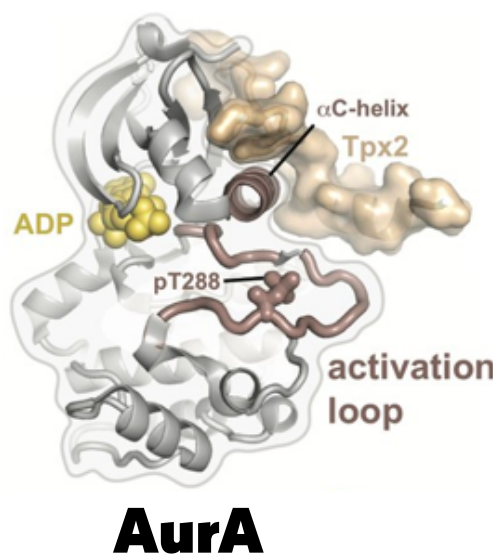
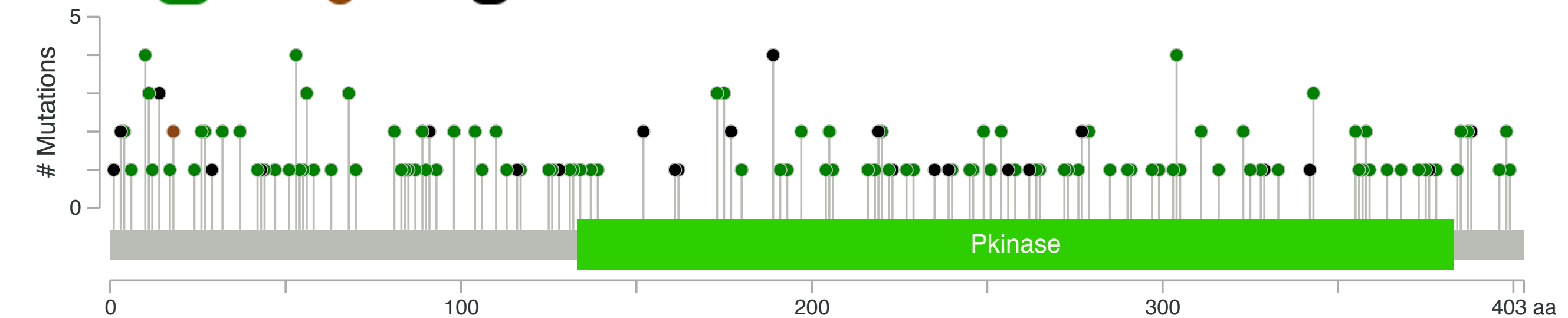
How do these functionally distinct activation pathways differ mechanistically?

NICK LEVINSON
UMN

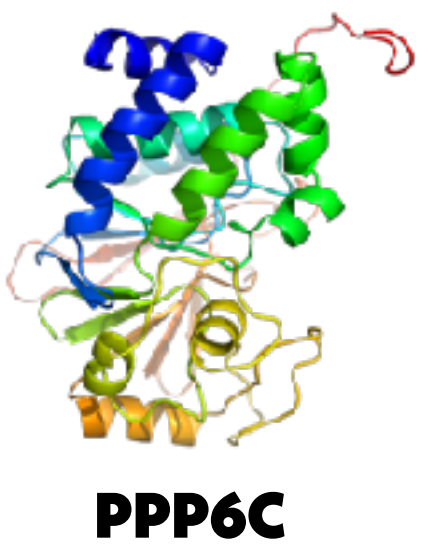
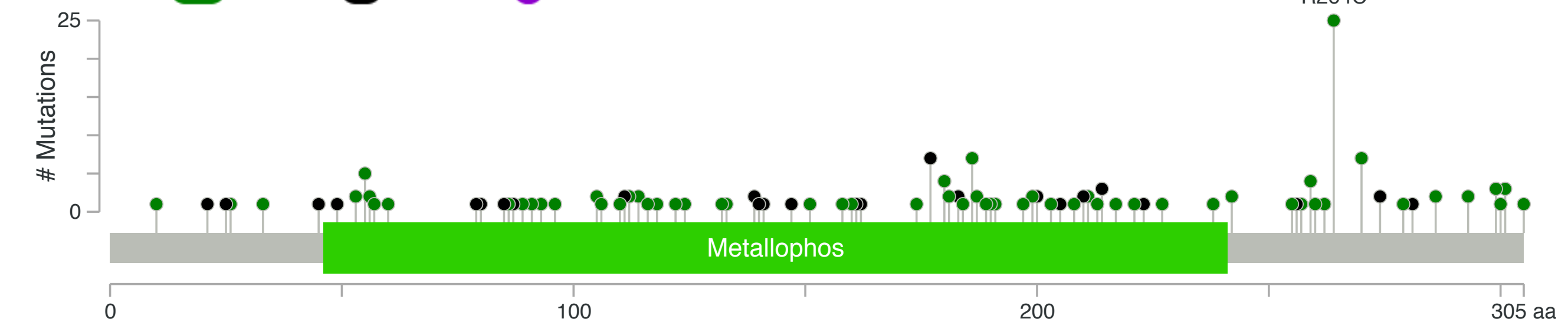
AURORA A AND ITS PHOSPHATASE PP6 ARE MUTATED OR AMPLIFIED IN MANY CANCERS



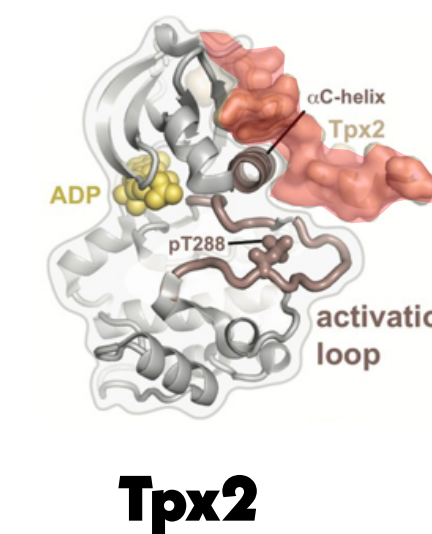
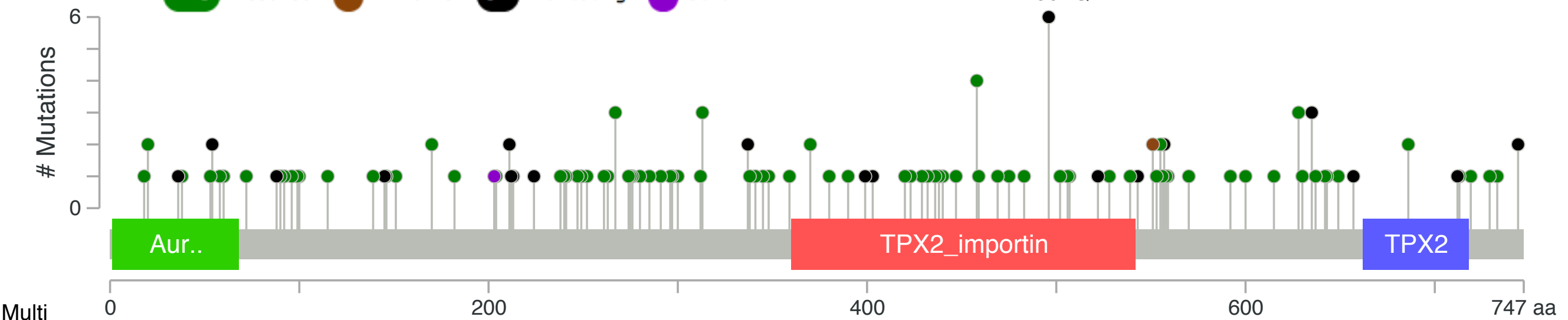
AURKA 183 Missense 1 In-frame 33 Truncating



PPP6C 150 Missense 35 Truncating 1 Other

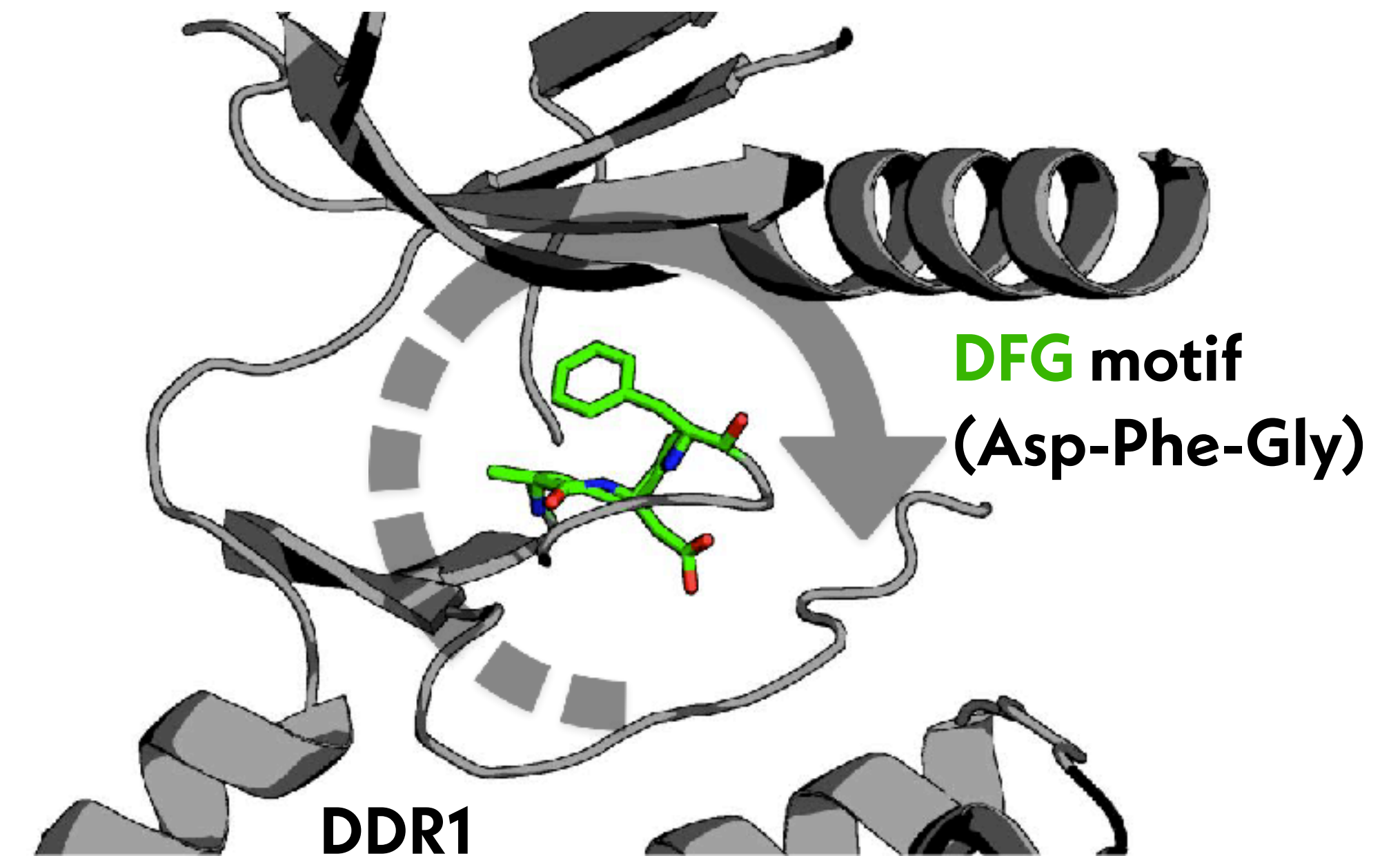
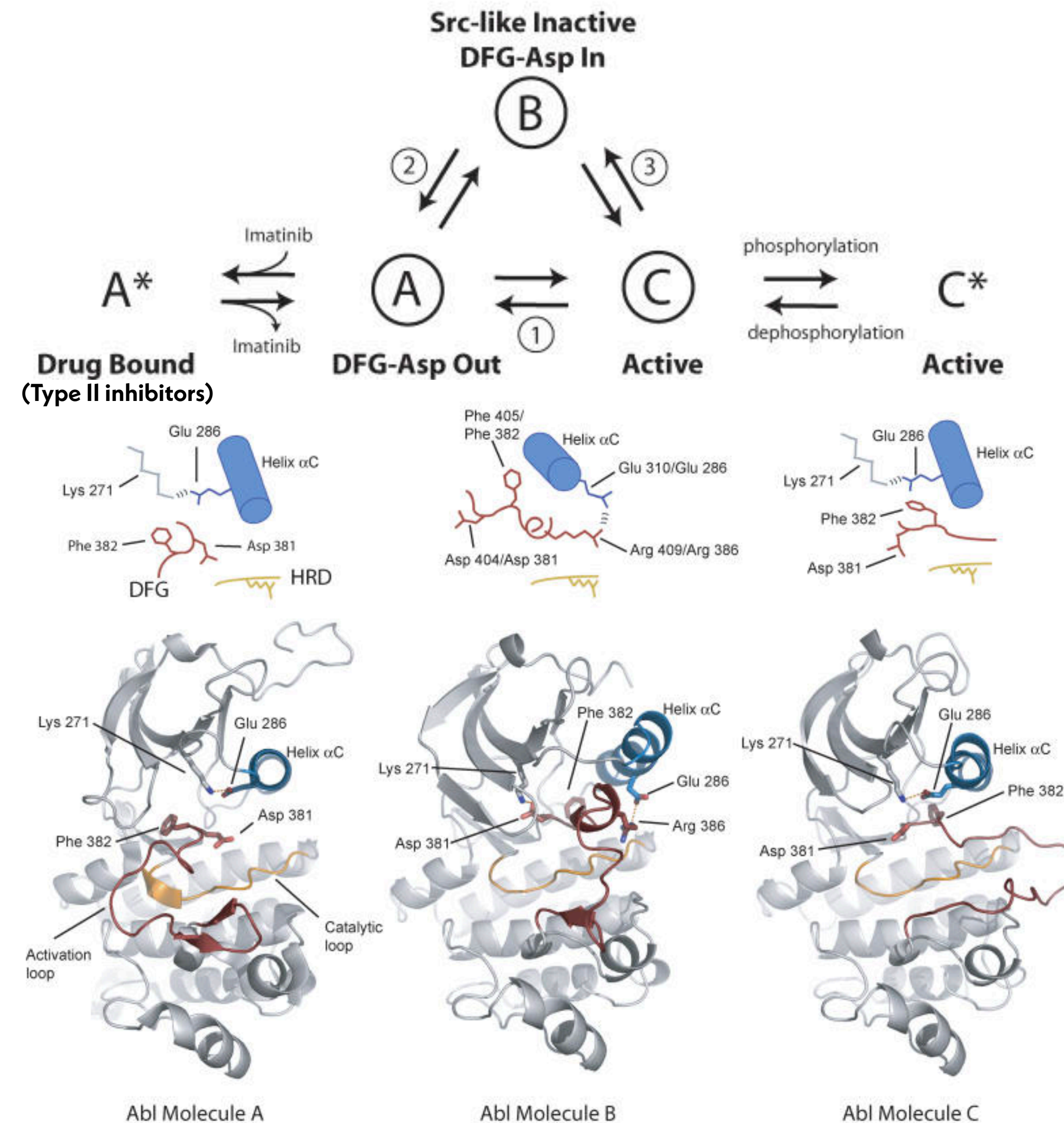


TPX2 143 Missense 1 In-frame 37 Truncating 1 Other



Several AurA inhibitors entered clinical trials in the late 2000s, but clinical response was disappointing.
Could this be due to failure to be selective for distinctly regulated pools of AurA?

CANONICAL KINASE ACTIVATION INVOLVES FLIP OF **DFG** MOTIF ASPARTATE



Hanson*, Georgiou*, Miller, Rest, Chodera, Seeliger. Cell Chemical Biology, in press.

**Does Tpx2 activate Aurora A by inducing
a DFG out-to-in transition?**

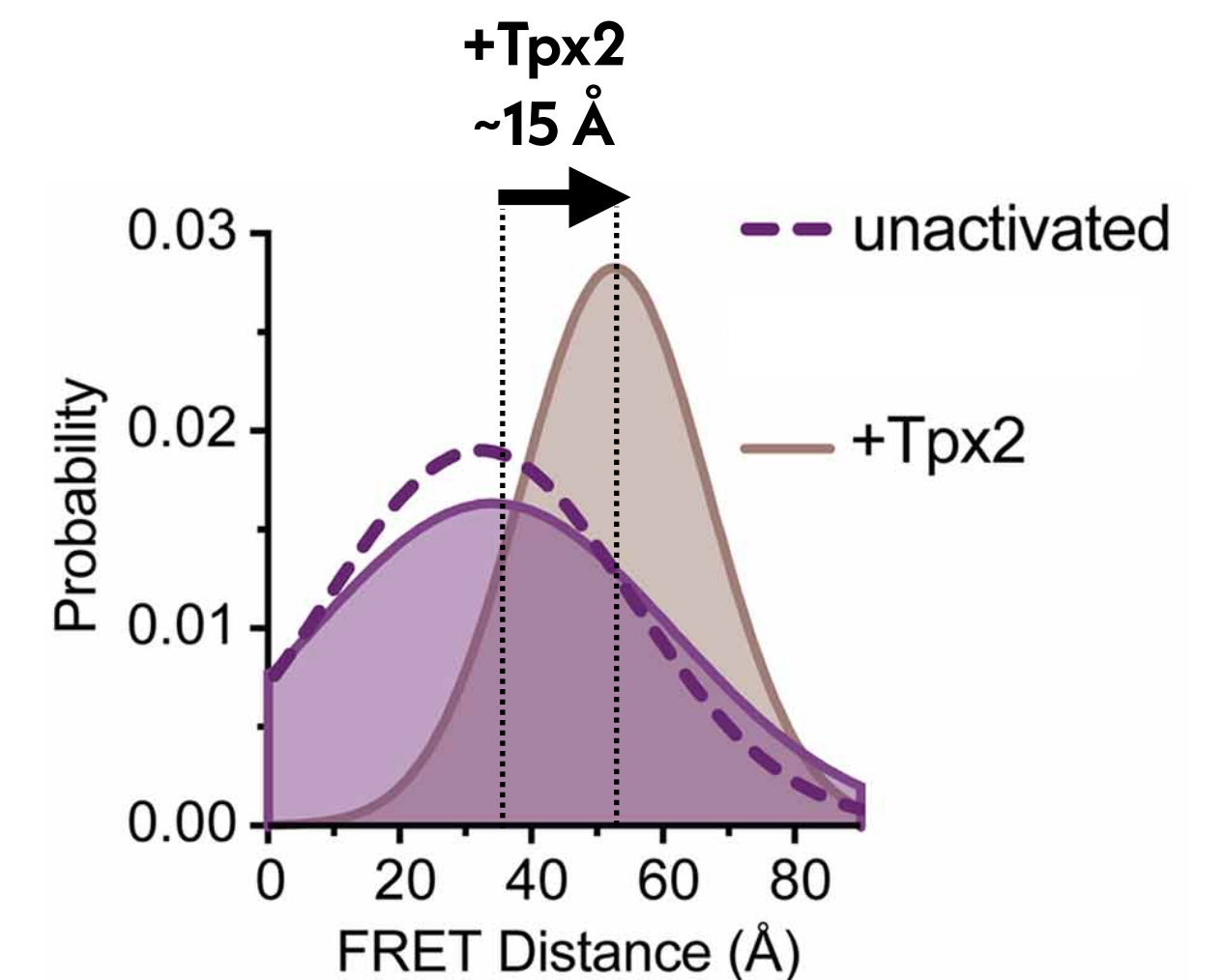
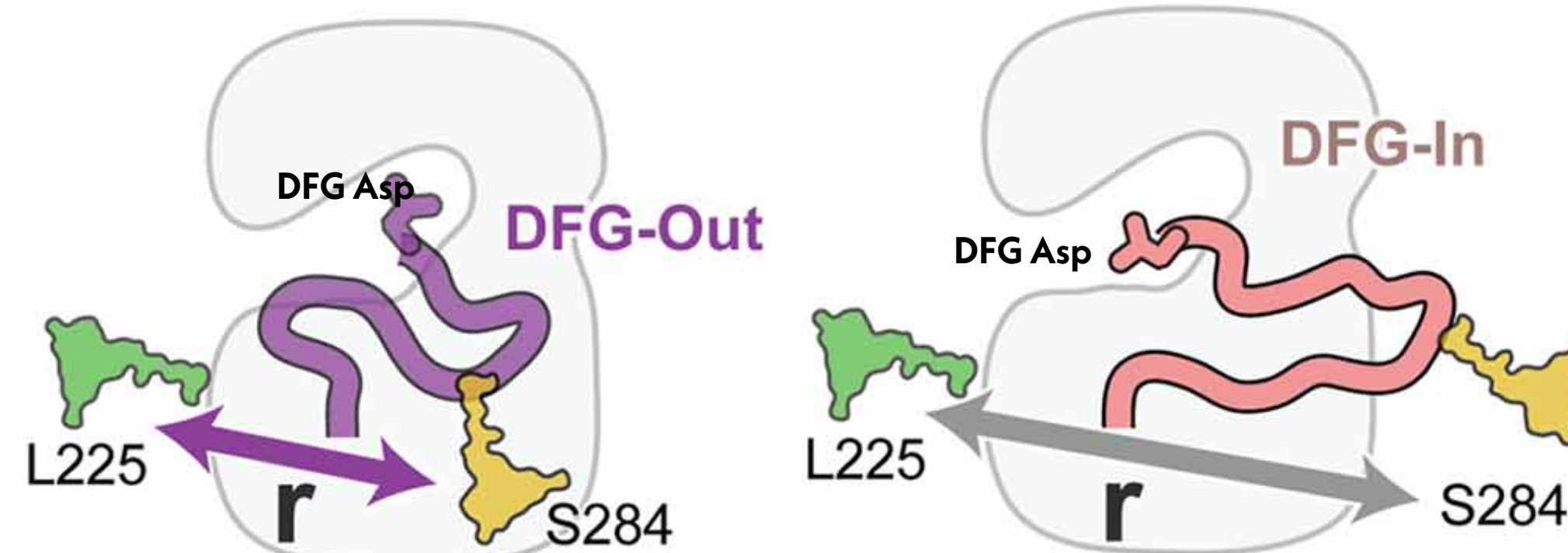
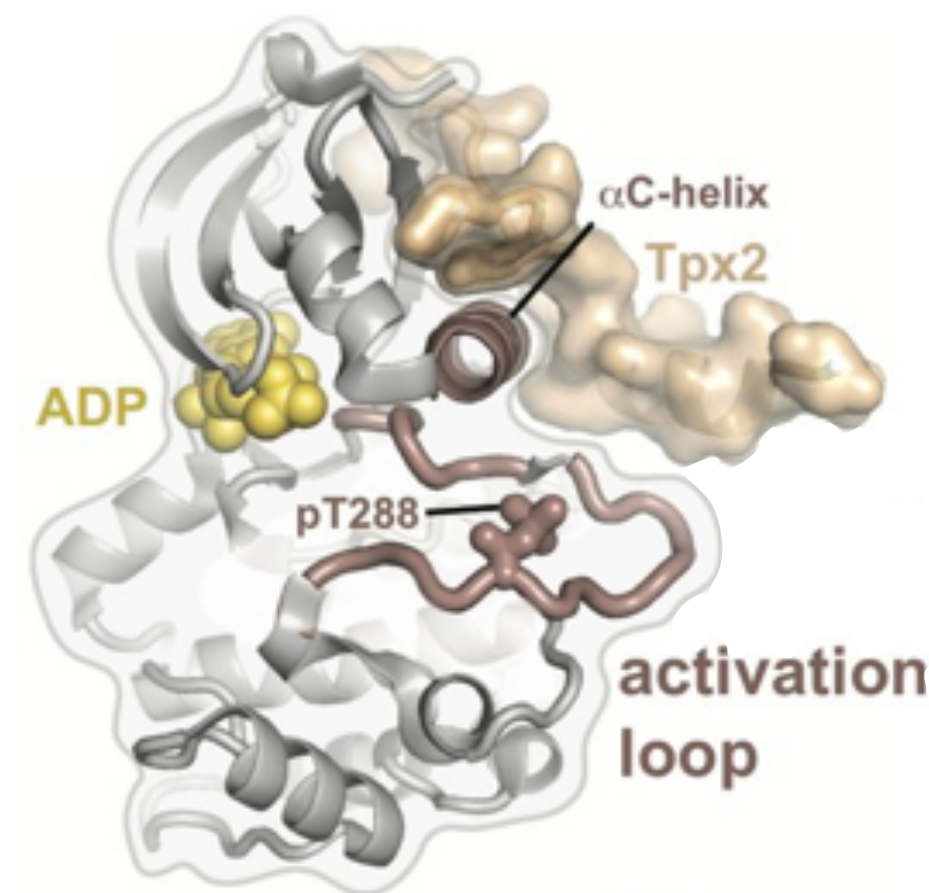
SONYA HANSON



BINDING OF THE ALLOSTERIC ACTIVATOR PROTEIN TPX2 INDUCES A NANOMETER-SCALE MOVEMENT OF THE ACTIVATION LOOP

FRET labeling

(fluorescence resonance energy transfer)

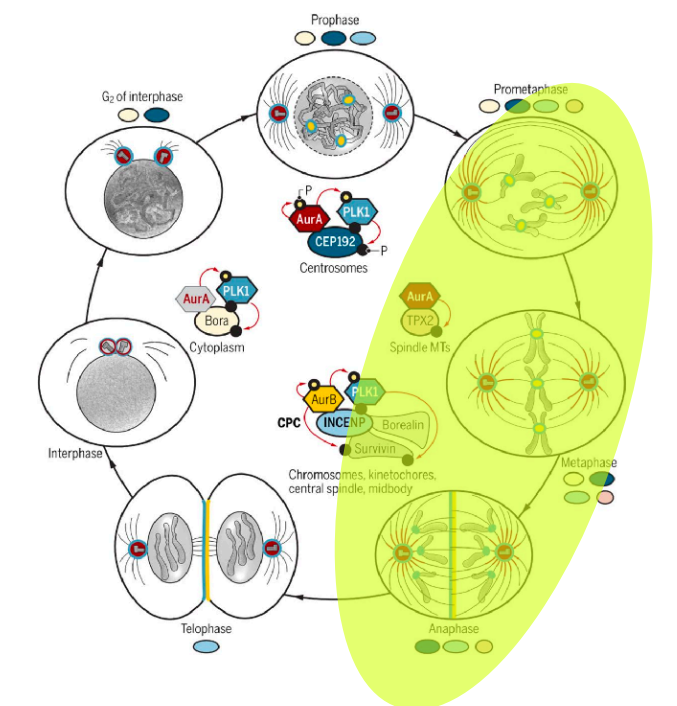


**Tpx2 binding shifts DFG motif from out to in.
But there's a problem...**

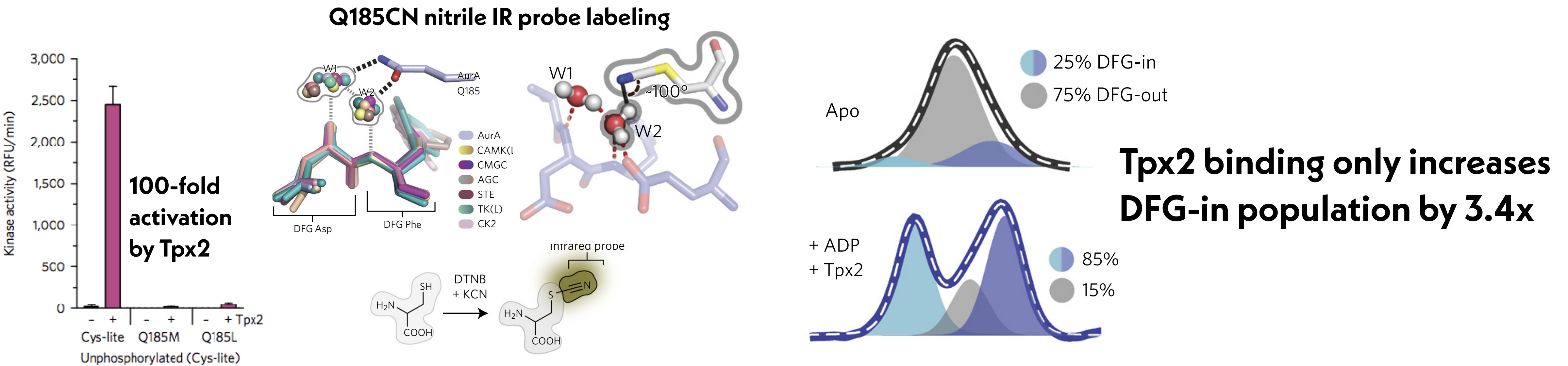


NICK LEVINSON
UMN

Cyphers, Ruff, Behr, Chodera, and Levinson. Nature Chemical Biology 13:402, 2017



DFG POPULATION SHIFT IS INSUFFICIENT TO EXPLAIN 100-FOLD ACTIVATION BY TPX2



**Where does the rest of the 100x activation by Tpx2 come from?
Can molecular simulations shed some light on this?**



NICK LEVINSON
UMN

MOLECULAR SIMULATIONS STRUGGLE TO ACCESS KINASE CONFORMATIONAL TIMESCALES

simulation timestep	1 femtosecond	10^{-15} seconds
interesting kinase events	1 microsecond	10^{-6} seconds
number of timesteps	1 billion timesteps	

MOLECULAR SIMULATIONS STRUGGLE TO ACCESS KINASE CONFORMATIONAL TIMESCALES

\$50K
9 TFLOP/S



years on a single CPU



CONSUMER-GRADE GRAPHICS PROCESSORS CAN BREAK THROUGH THE SPEED BARRIER

\$50K
9 TFLOP/S



years on a single CPU



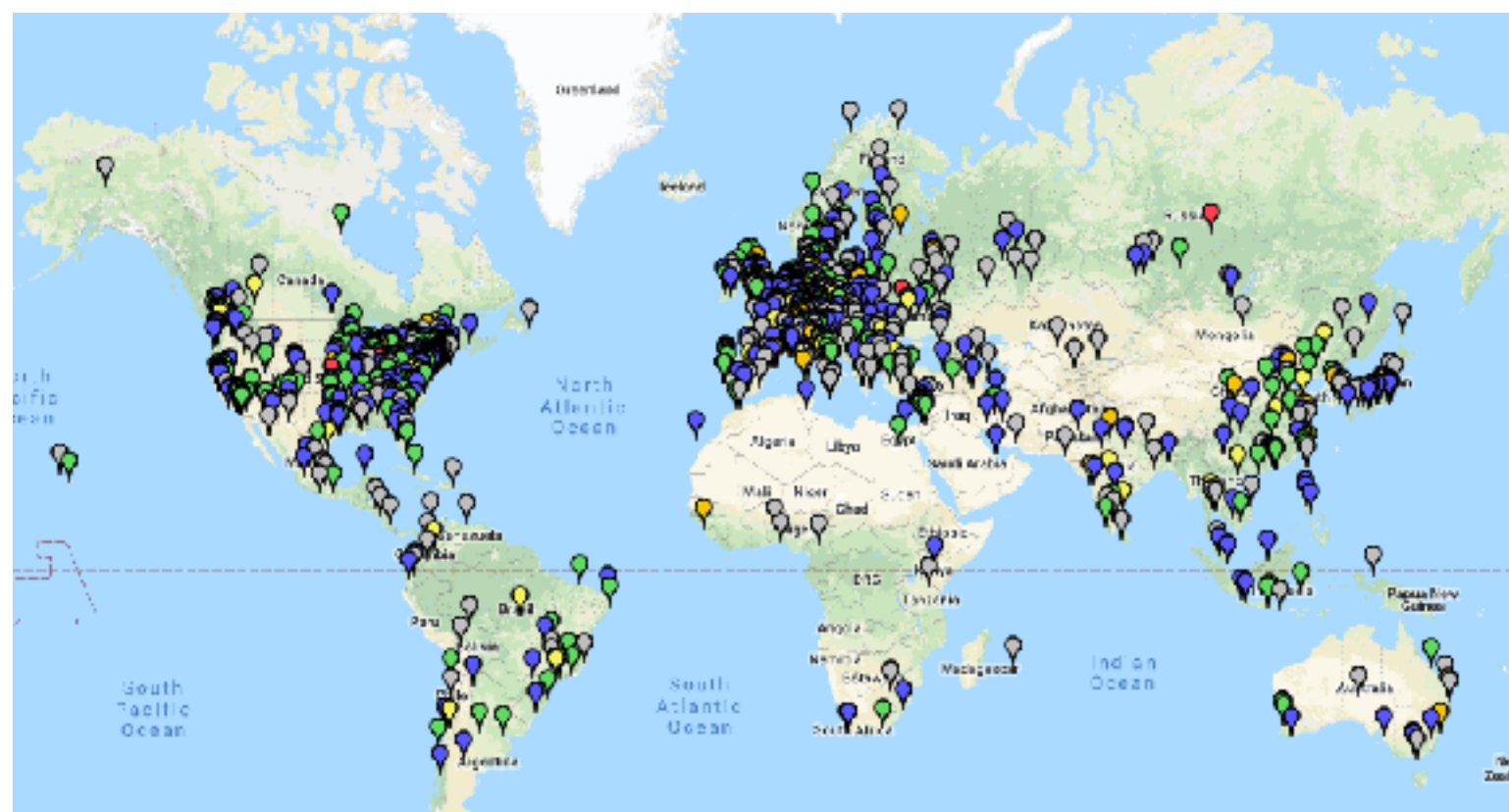
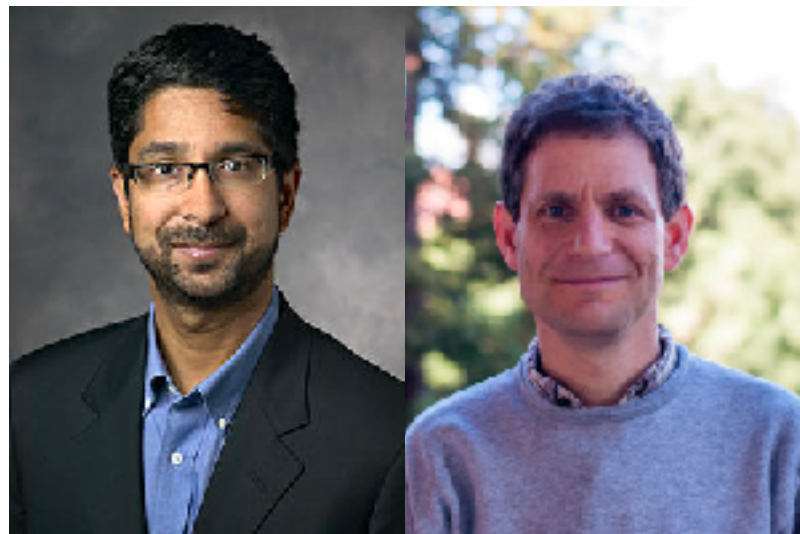
\$500
9 TFLOP/S

NVIDIA GTX 1080
100x more power/\$ than CPUs

WE BUILT THE OPEN-SOURCE **OPENMM**, THE FIRST GPU-ACCELERATED BIOMOLECULAR SIMULATION CODE

NVIDIA GTX-1080 (\$500)
9 TFLOP/S SINGLE PRECISION

VIJAY PANDE **PETER EASTMAN**
STANFORD/A16Z **STANFORD**



OpenMM speedup (GTX-1080) over 12-core Xeon X5650 CPU for DHFR

method	natoms	gromacs CPU	OpenMM GPU	speedup
GB/SA	2,489	2.54 ns/day	789 ns/day	311 x
RF	23,558	18.8 ns/day	572 ns/day	30.4 x
PME	23,558	6.96 ns/day	337 ns/day	48.4 x

<http://openmm.org> OpenMM 7.1.0 development snapshot benchmark
gromacs benchmarks from <http://biowulf.nih.gov/apps/gromacs-gpu.html>



OpenMM

<http://openmm.org>

downloads **150k total**



OpenMMTools

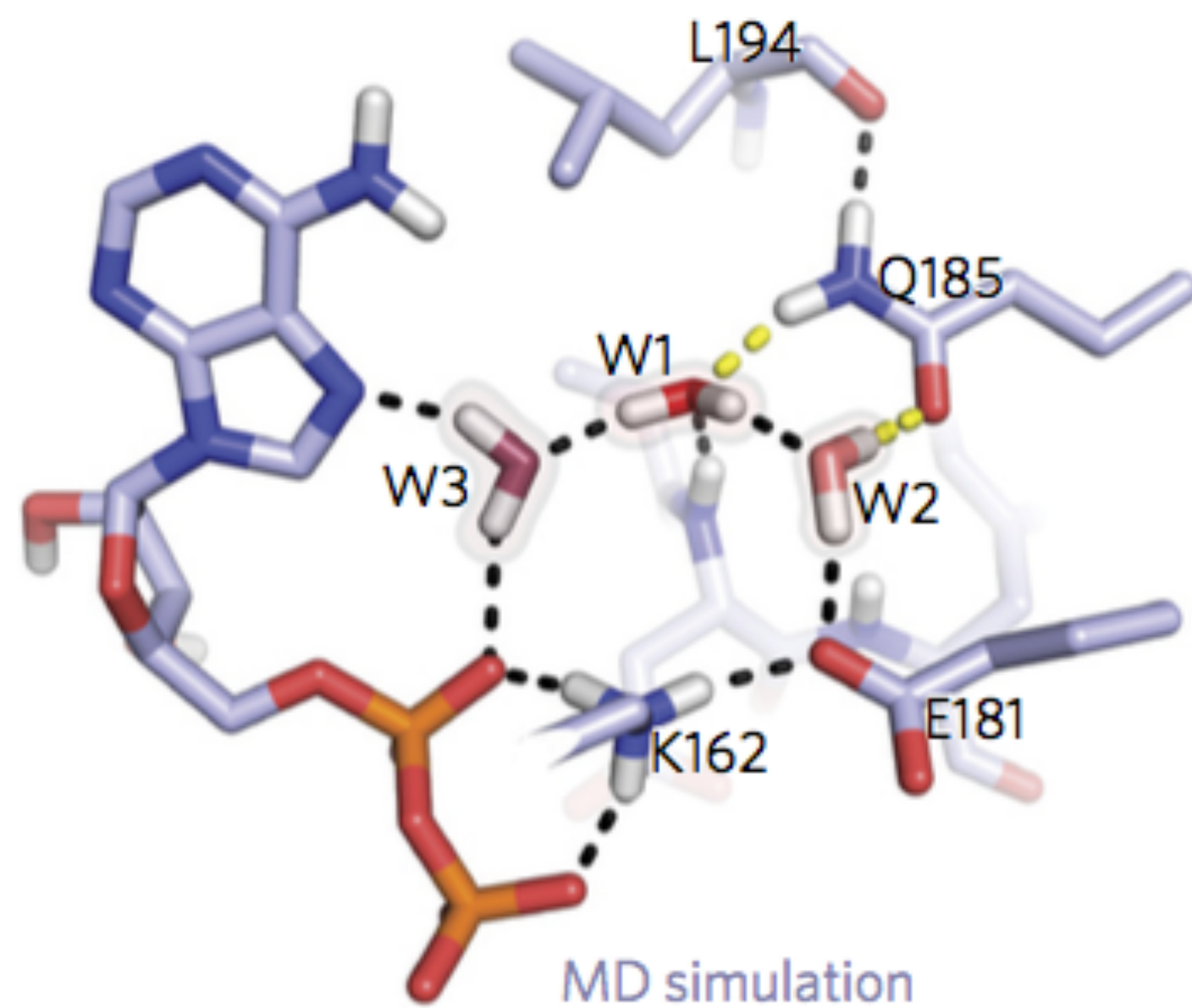
<http://github.com/coderalab/openmmtools>

downloads **62k total**

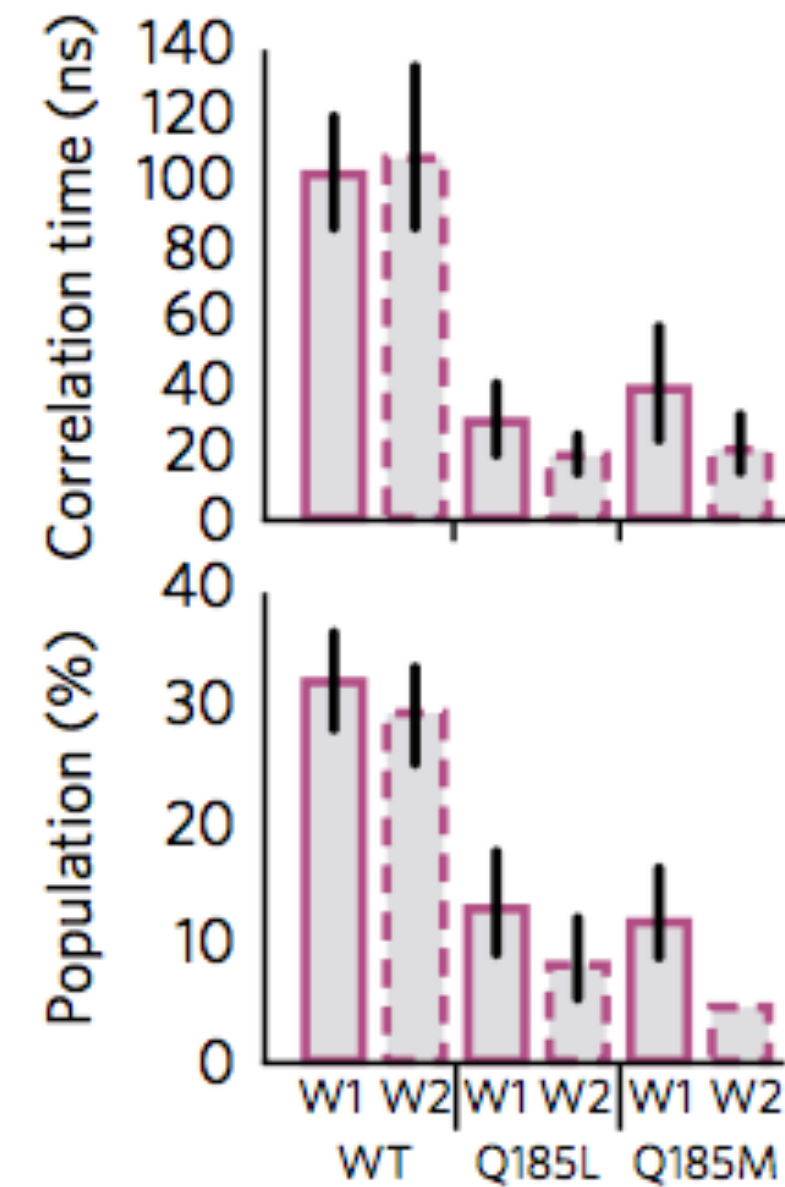
Eastman, Swails, **Chodera**, McGibbon, Zhao, Beauchamp, Wang, Simonett, Harrigan, Stern, Wiewiora, Brooks, Pande. PLoS Comput Biol 13:e1005659, 2017

Eastman, Friedrichs, **Chodera**, Radmer, Bruns, Ku, Beauchamp, Lande, Wang, Shukla, Tye, Houston, Stich, Klein, Shirts, Pande. J Chem Theor Comput 9:461, 2013

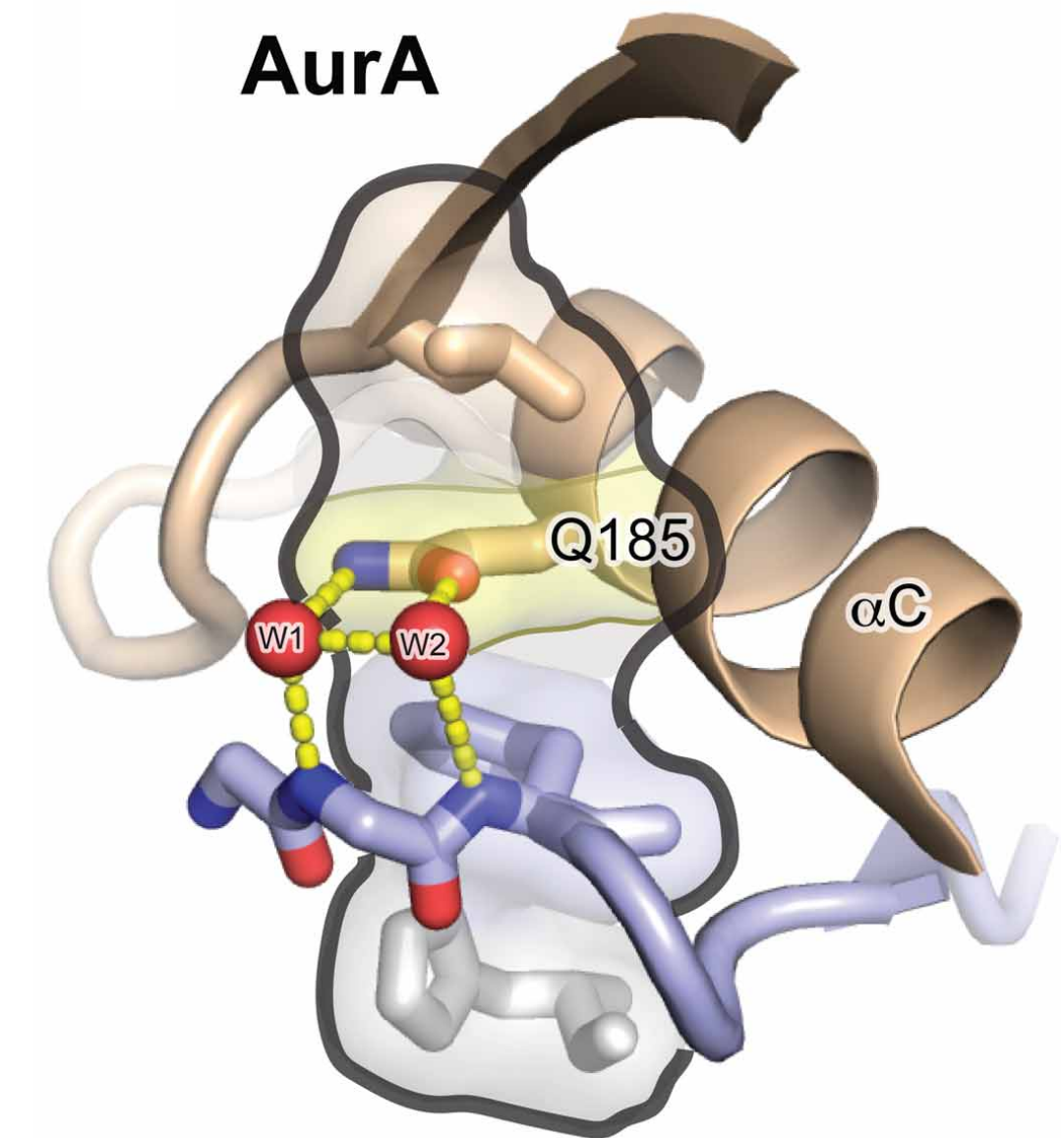
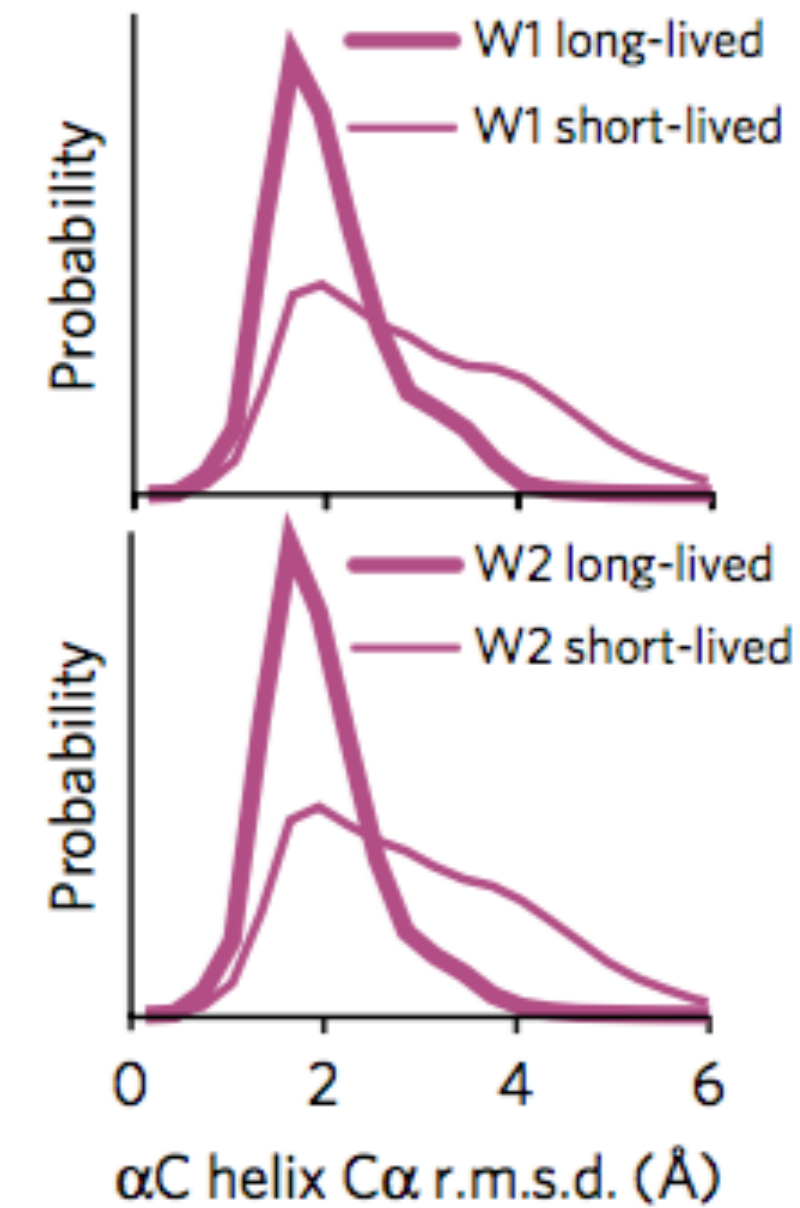
A LONG-LIVED ALLOSTERIC WATER NETWORK IS KEY TO TPX2 ACTIVATION



ordered waters



α C helix



Ultra-stable ordered water network is a key structural feature that explains the remainder of 100-fold allosteric activation of AurA by Tpx2



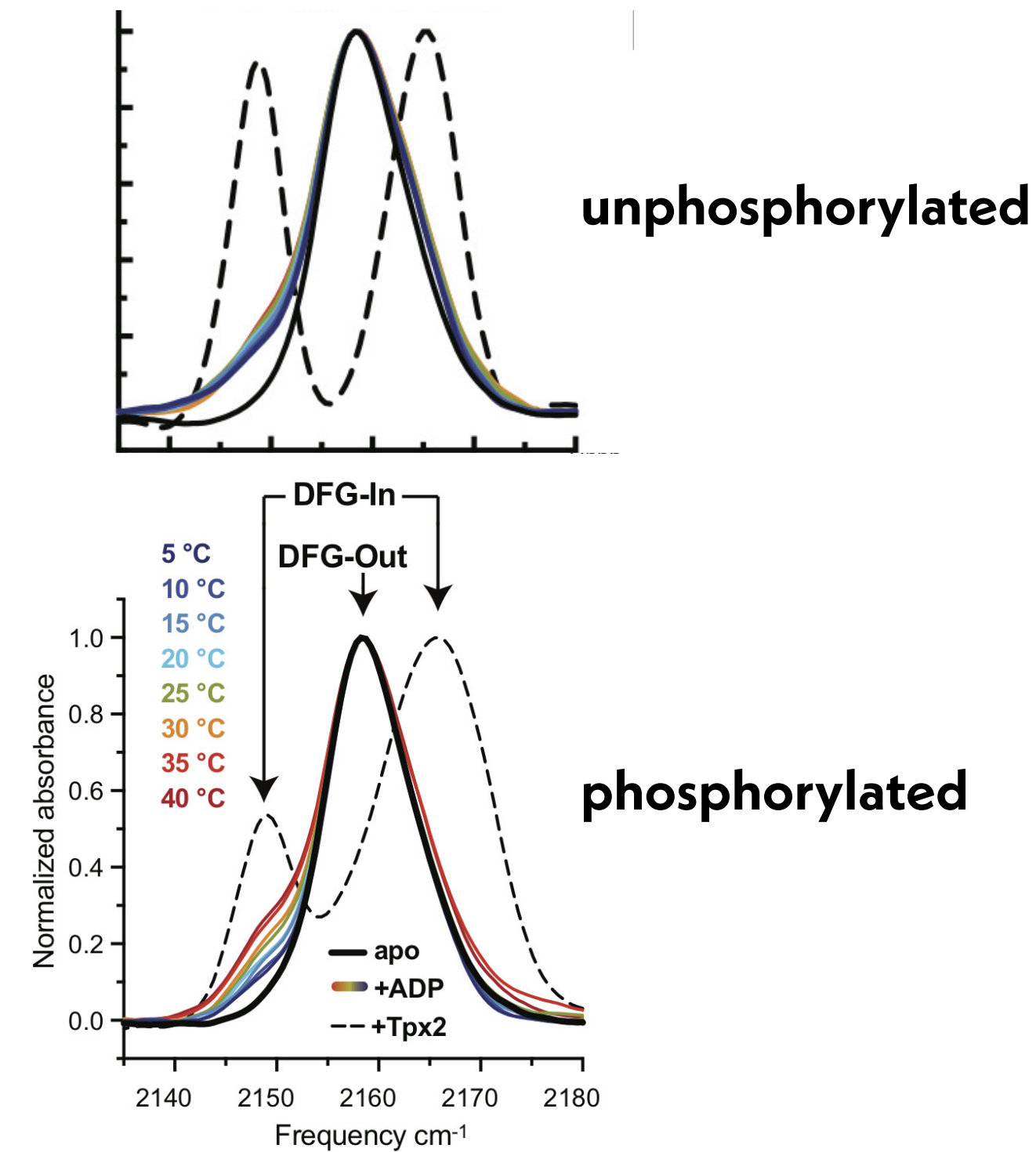
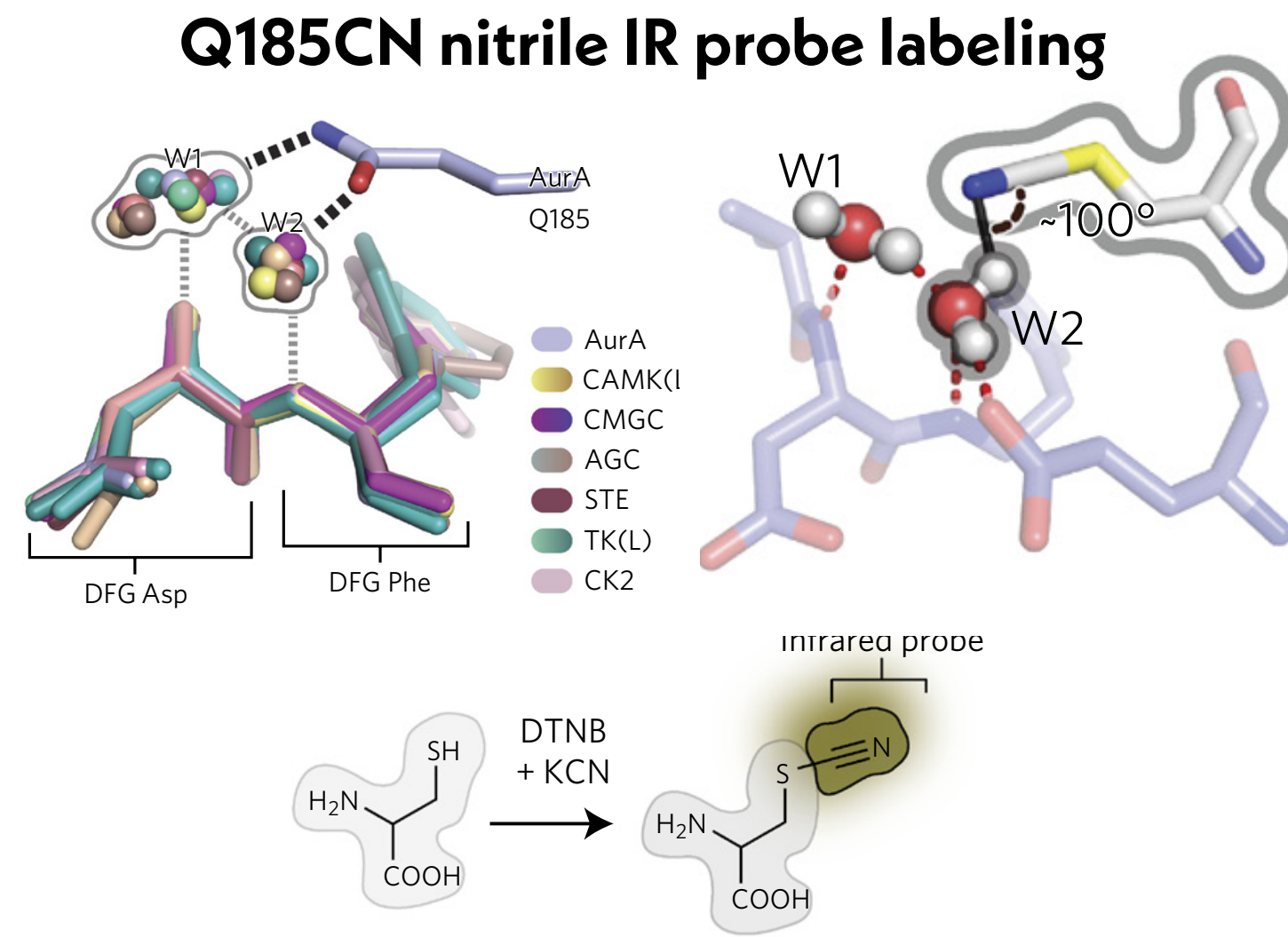
NICK LEVINSON
UMN

Cyphers, Ruff, Behr, Chodera, and Levinson. Nature Chemical Biology 13:402, 2017



JULIE BEHR

DOES PHOSPHORYLATION ALSO CAUSE A DFG-OUT TO -IN POPULATION SHIFT?



Phosphorylation of T288 does not shift the DFG population!



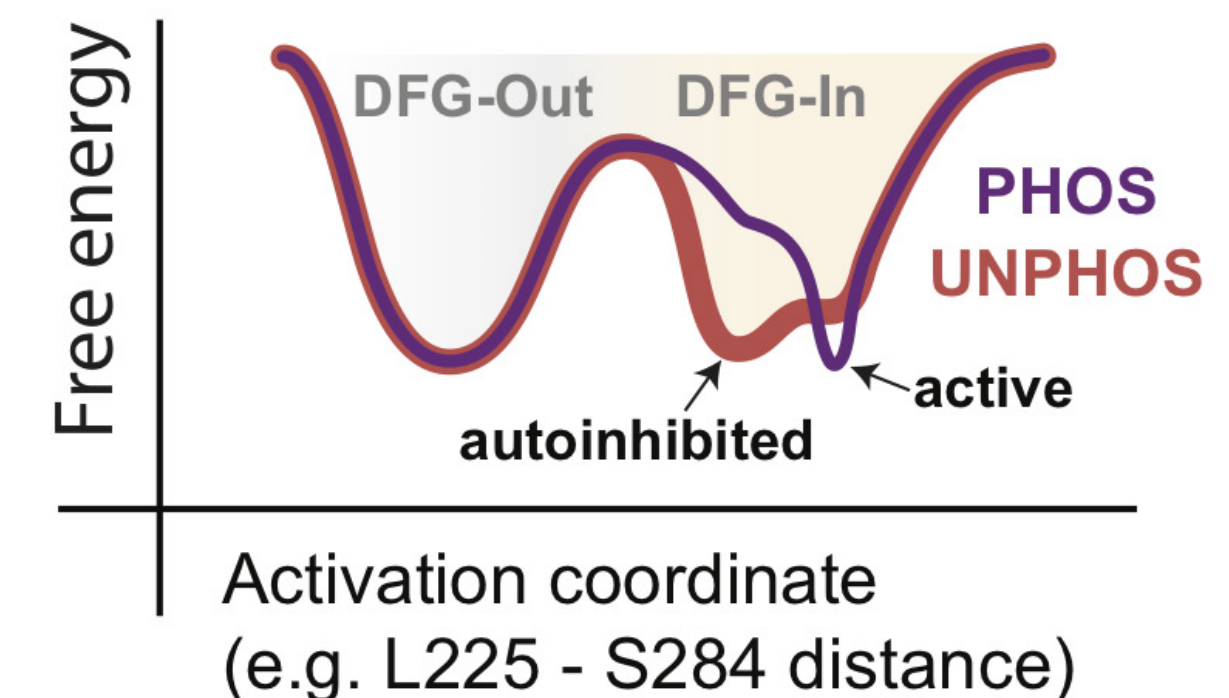
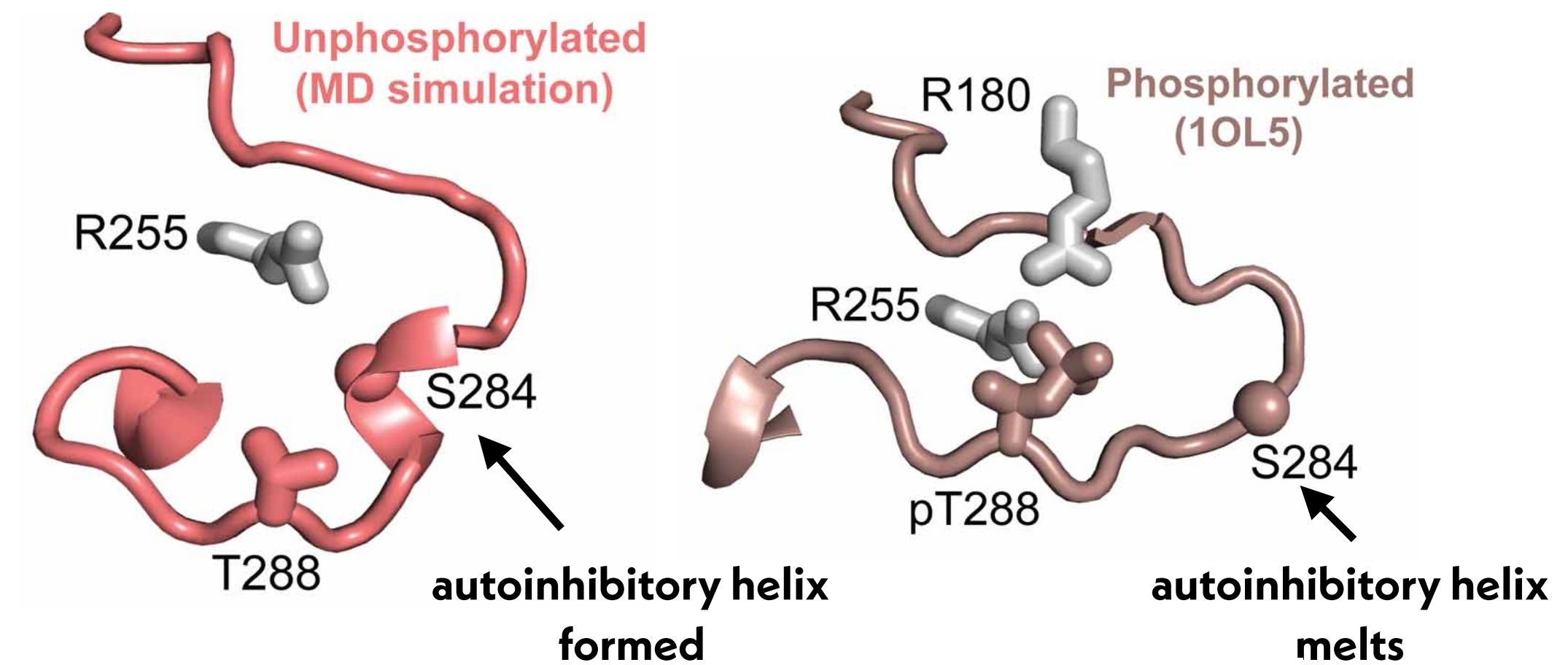
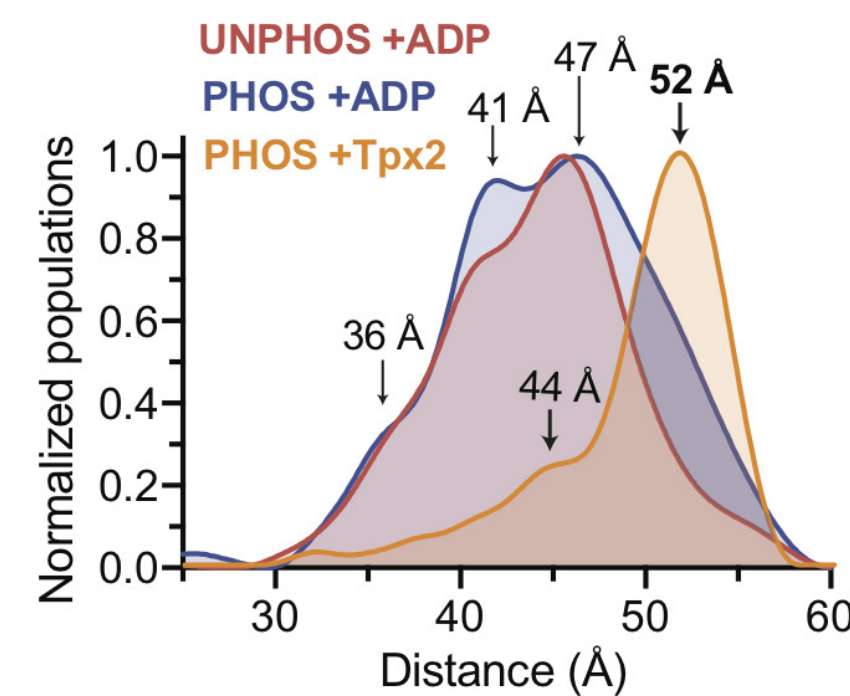
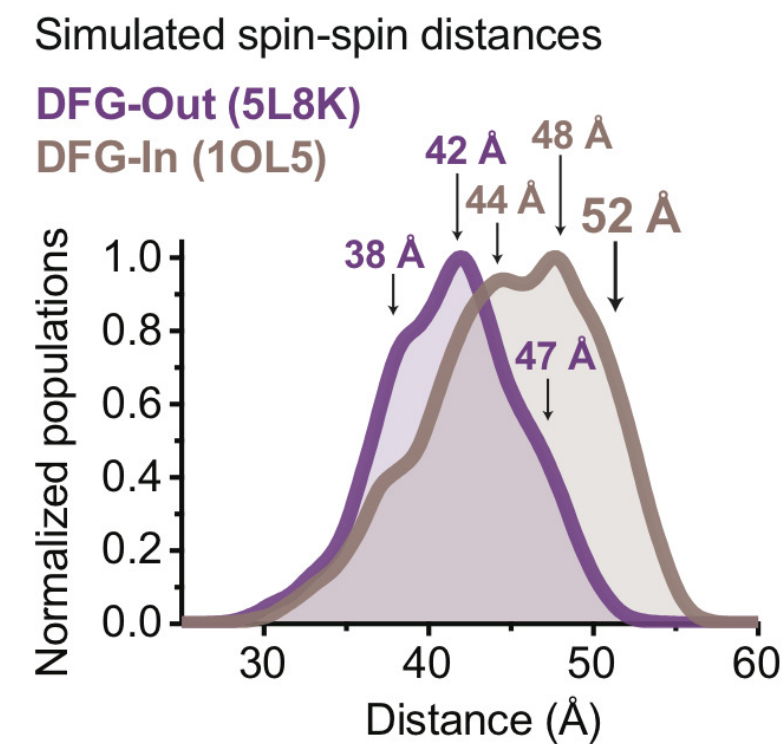
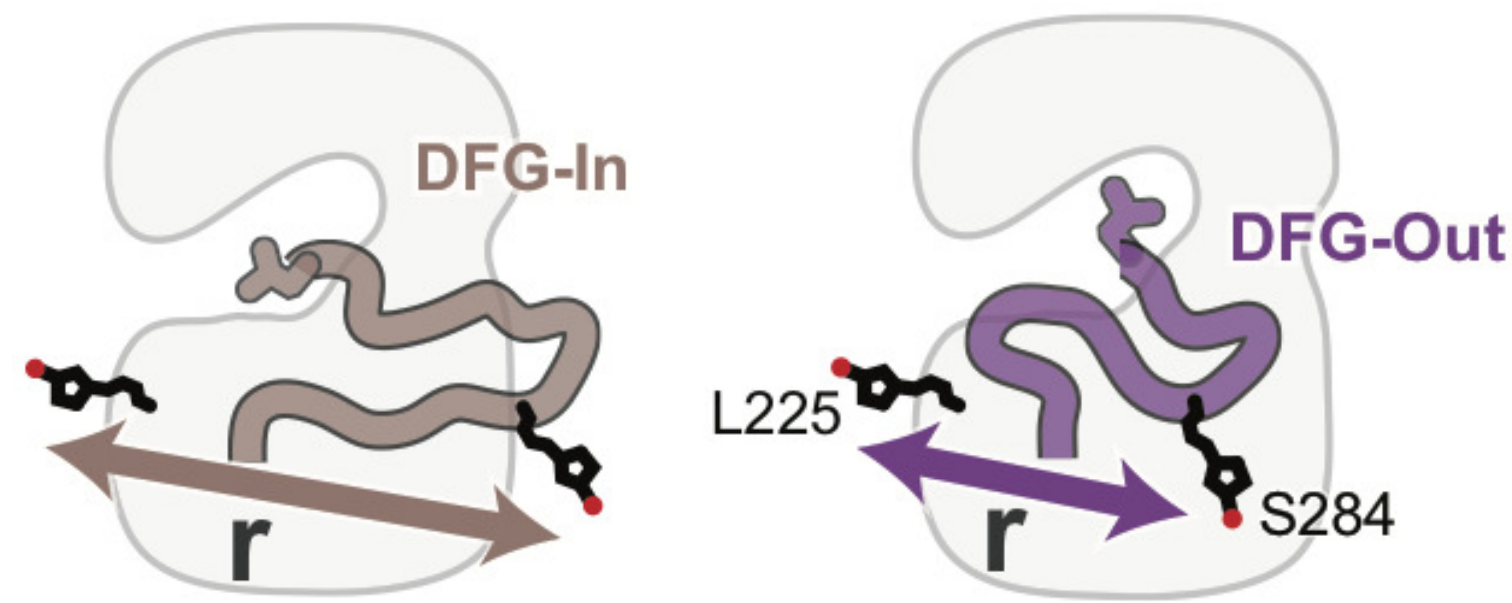
NICK LEVINSON
UMN

Ruff, Muretta, Thompson, Lake, Cyphers, Albanese, Hanson, Behr, Thomas, **Chodera**, and Levinson. eLife 7:e32766, 2018

PHOSPHORYLATION DOES NOT CAUSE A LARGE DFG-OUT TO -IN POPULATION SHIFT

EPR labels for DEER

(double electron-electron resonance)



Phosphorylation does **not** alter DFG-out/in equilibrium,
but remodels the DFG-in population!

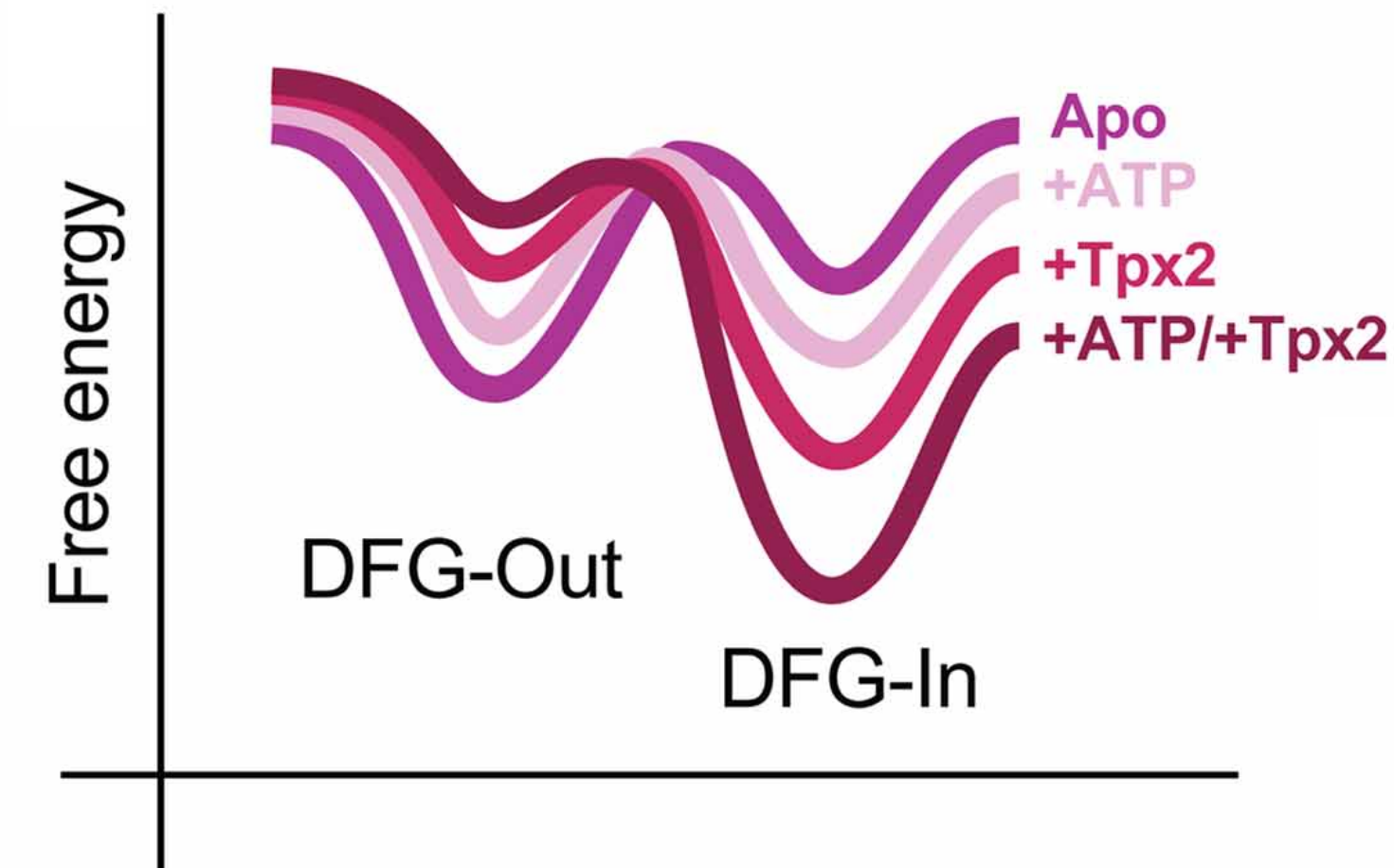
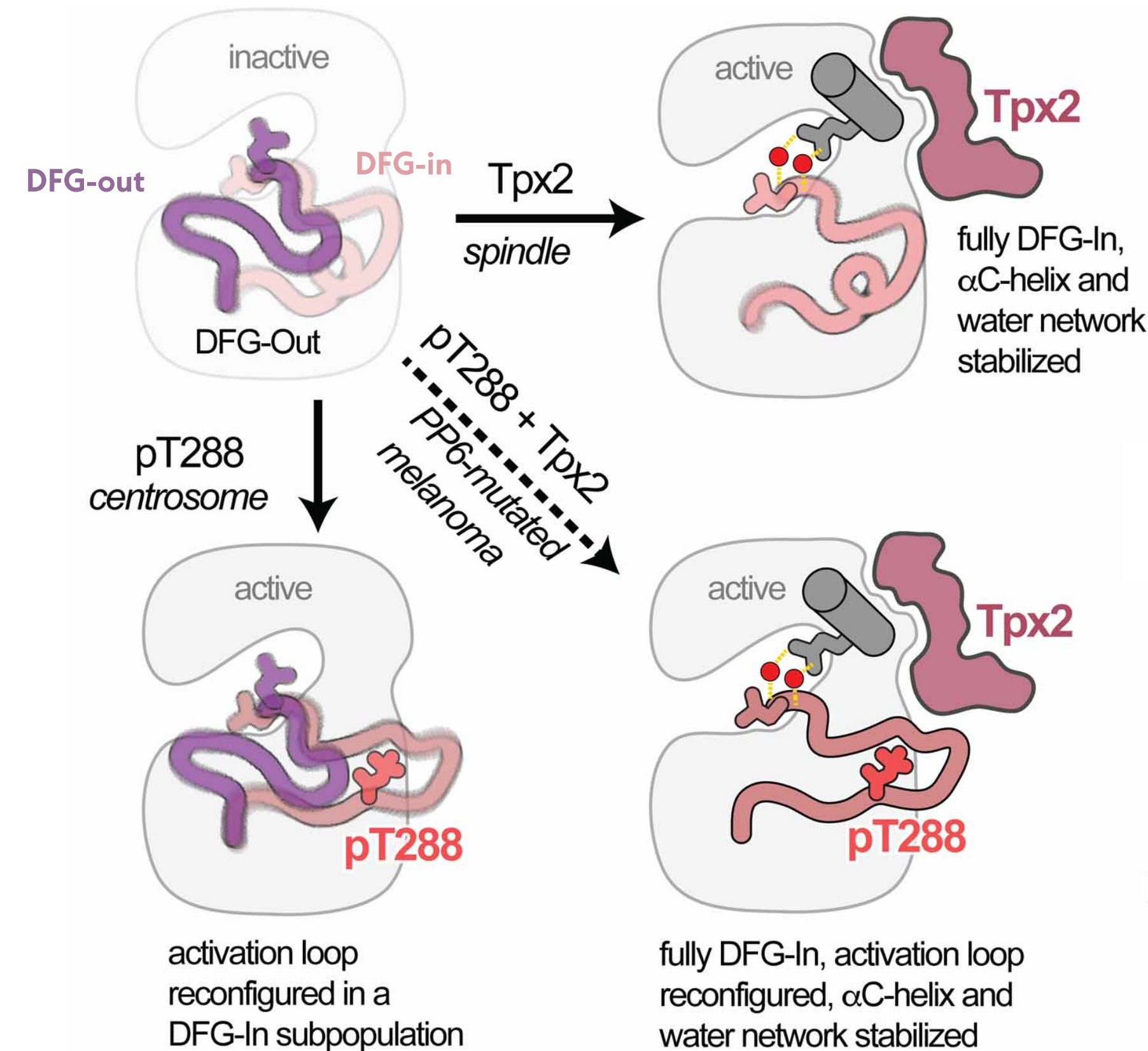
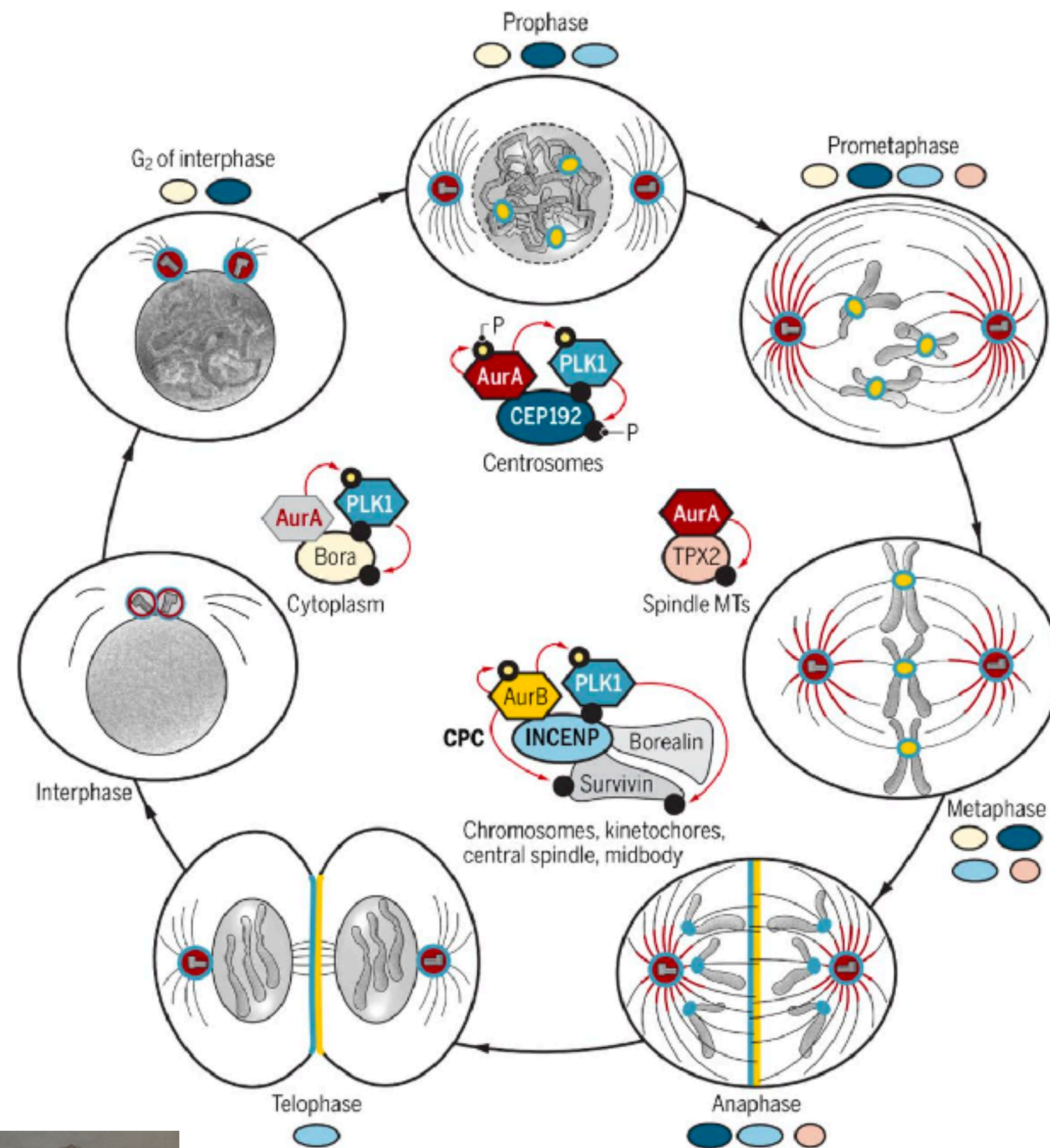
NICK LEVINSON

UMN

Ruff, Muretta, Thompson, Lake, Cyphers, Albanese, Hanson, Behr, Thomas, Chodera, and Levinson. eLife 7:e32766, 2018

STEVEN ALBANESE

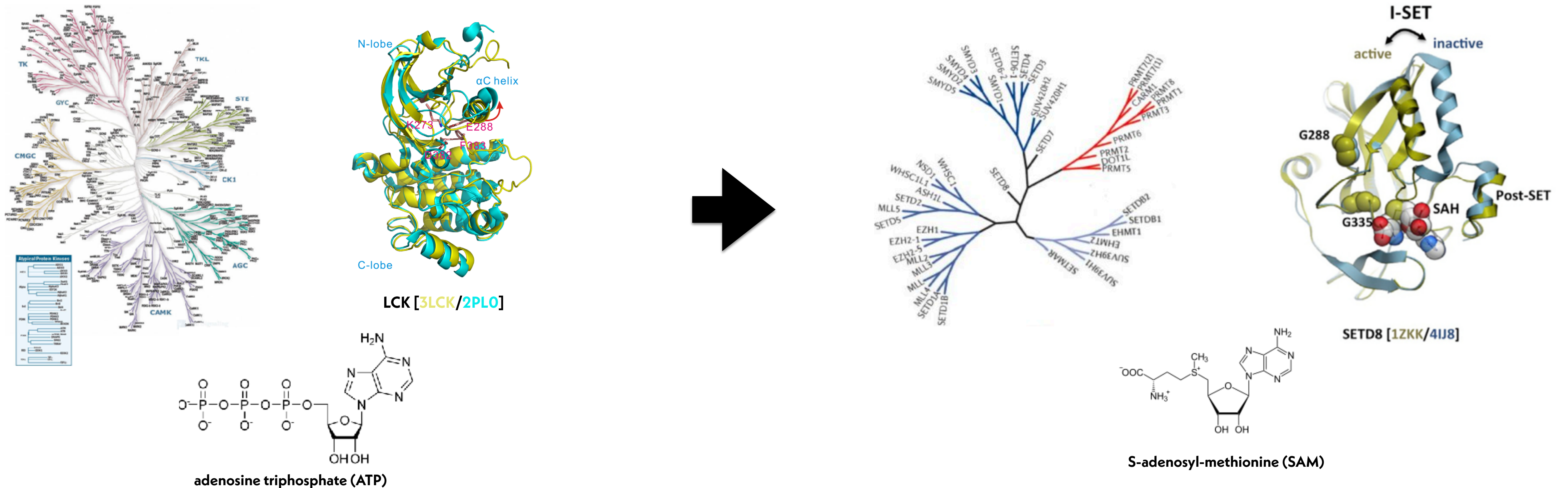
AN INTEGRATED MODEL EXPLAINS TWO DISTINCT MECHANISMS OF AURORA ACTIVATION



NICK LEVINSON
UMN

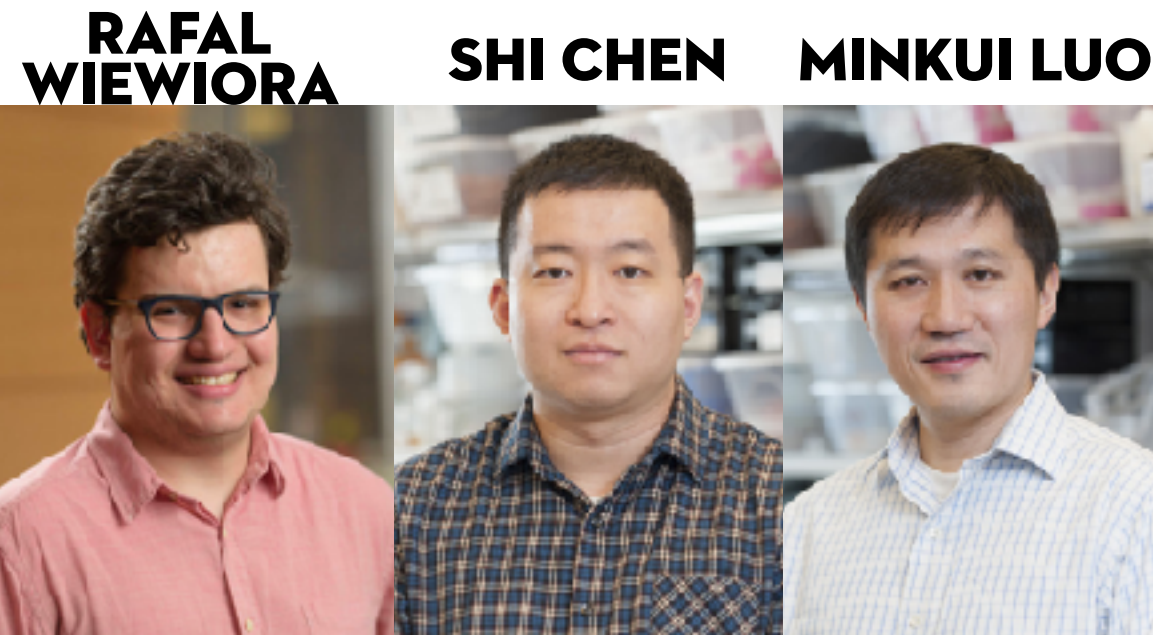
Will lead to new insights into how to selectively target separate pools of Aurora A with small molecules

PROTEIN LYSINE METHYLTRANSFERASES (PKMTs) ARE DYNAMICALLY HETEROGENEOUS EPIGENETIC CANCER TARGETS



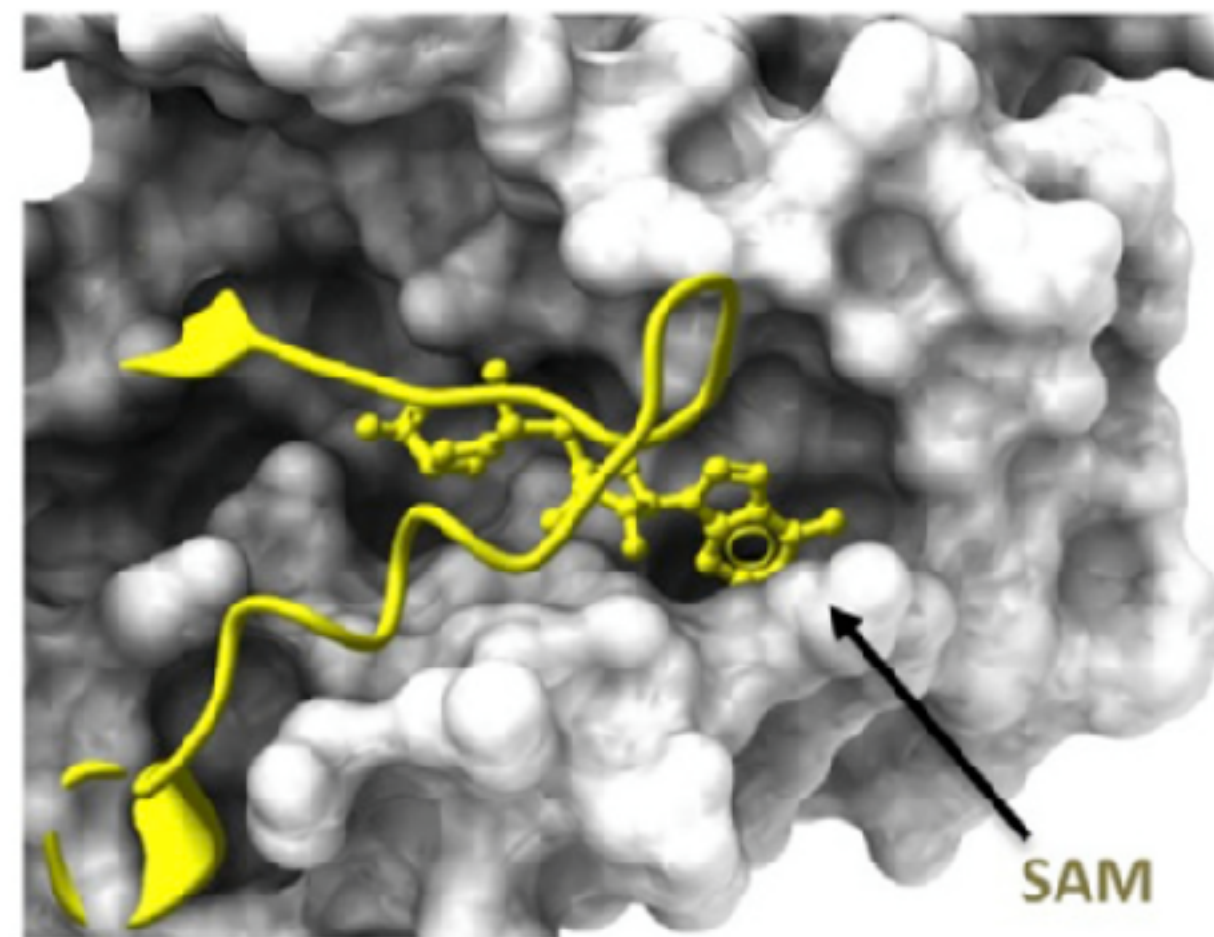
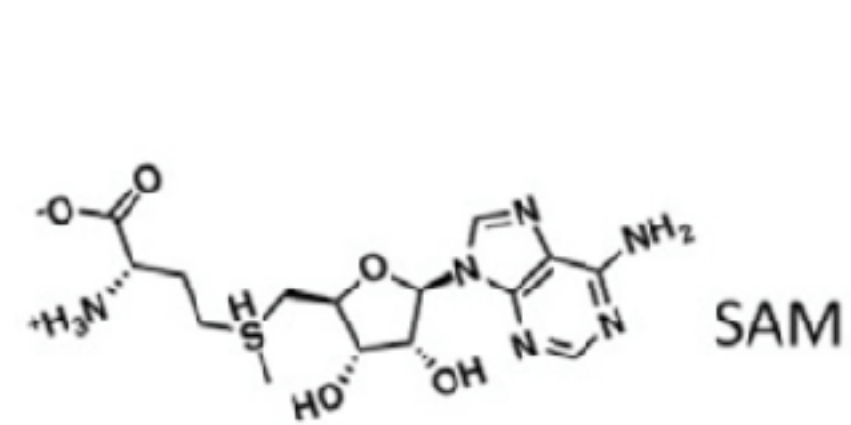
Can we use molecular simulation to learn about PKMT function and opportunities to exploit functional dynamics for selective inhibition?

Close collaboration between **Rafal Wiewiora** (Chodera lab) and **Shi Chen** (Luo Lab) both co-first authors

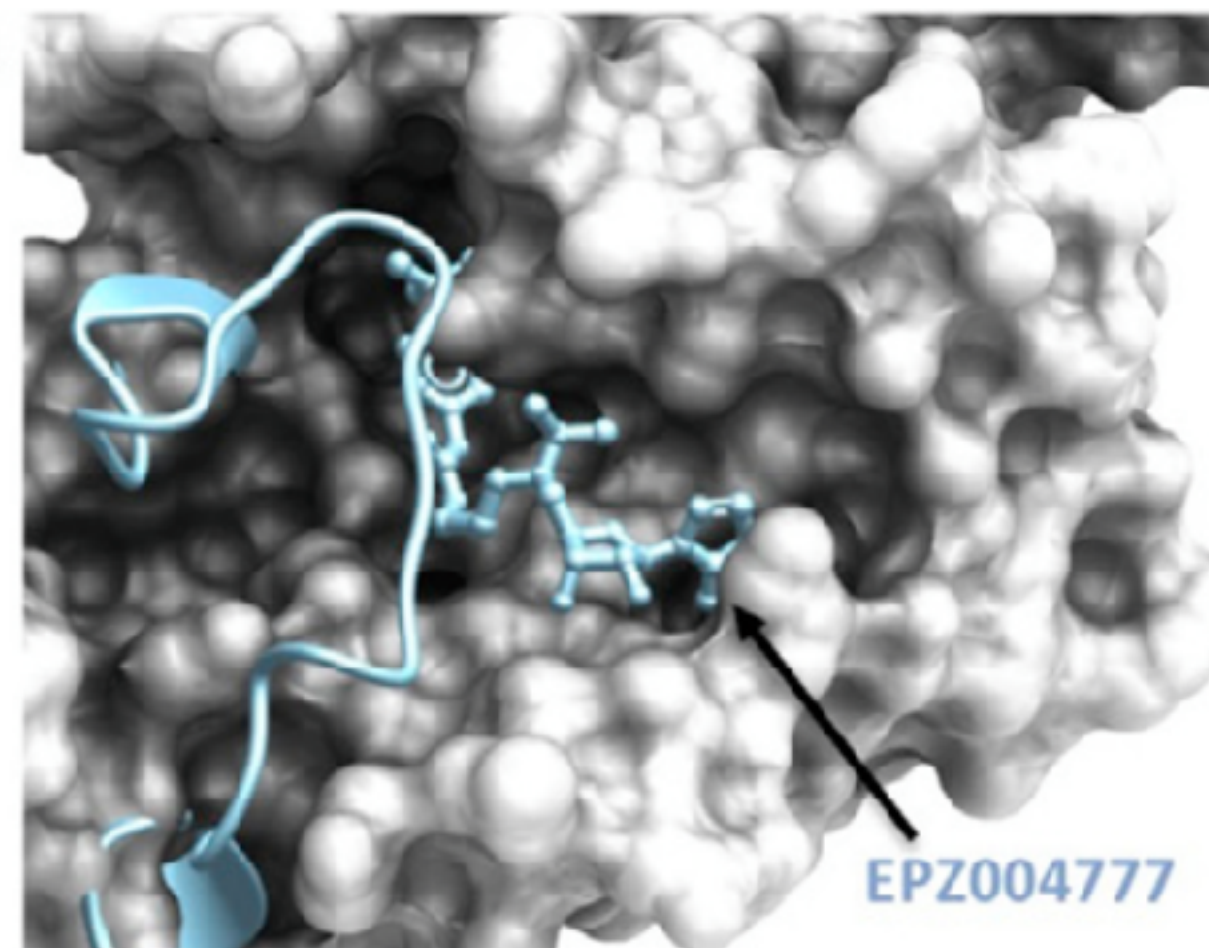
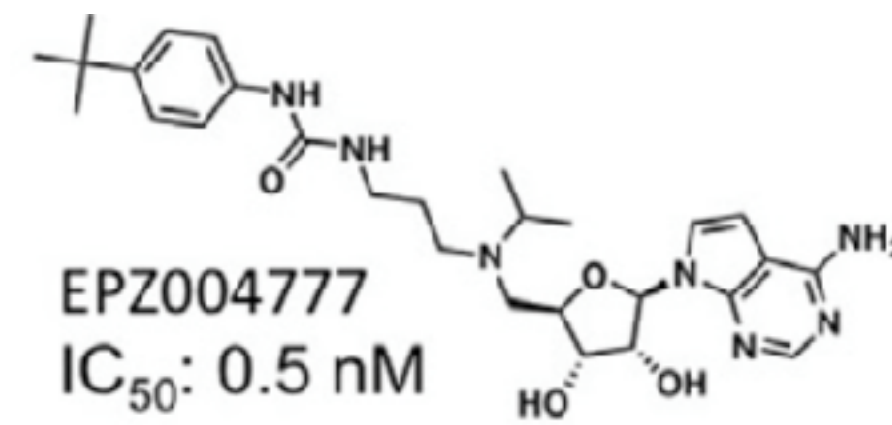


Kinases:
Manning et al. Science 298:1912, 2002
Xu et al. PLoS One 6:e22644, 2011
PKMTs:
Tian et al. Current Cancer Drug Targets 13:558, 2013
Schapira. Cell Chem Biol 23:1067, 2016

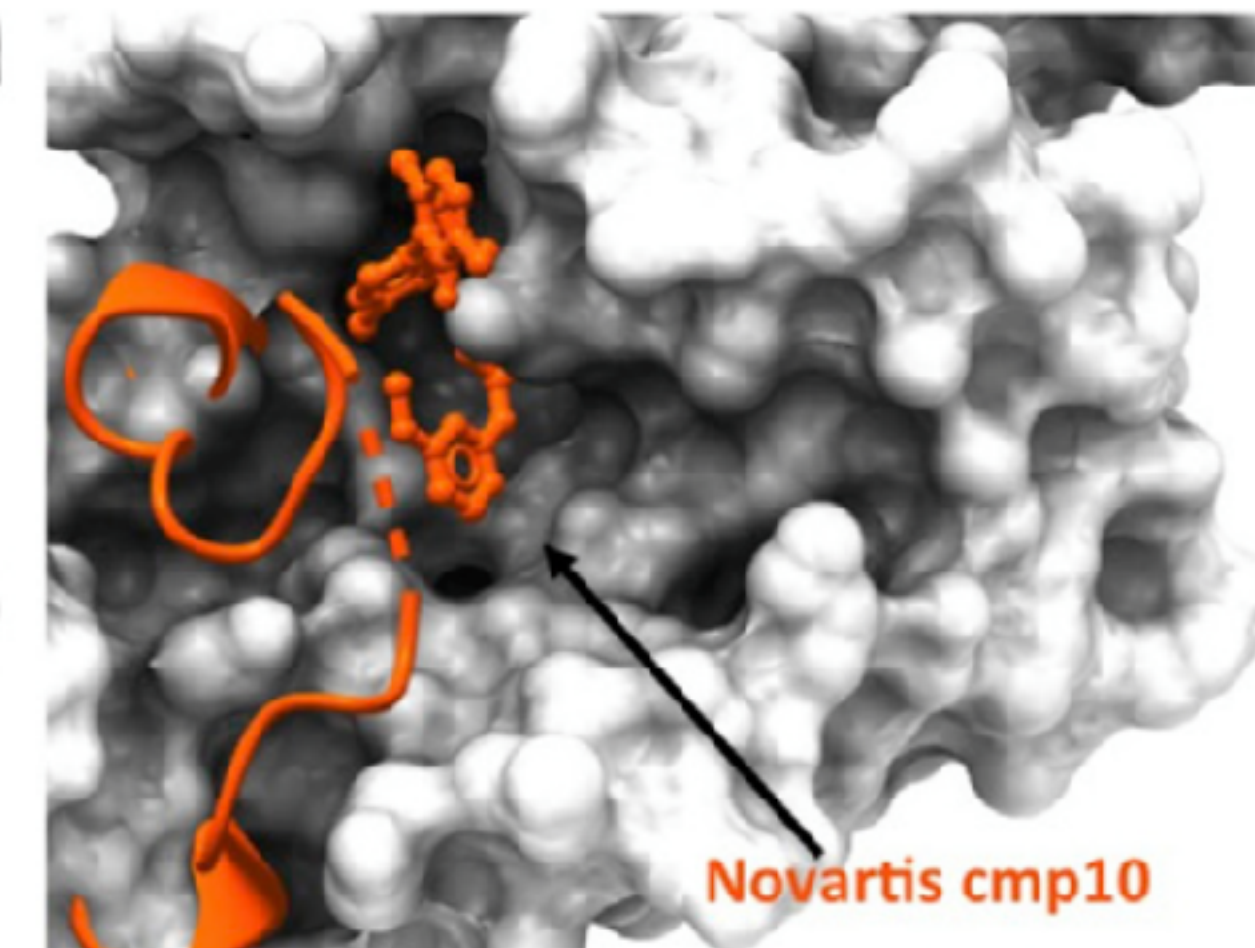
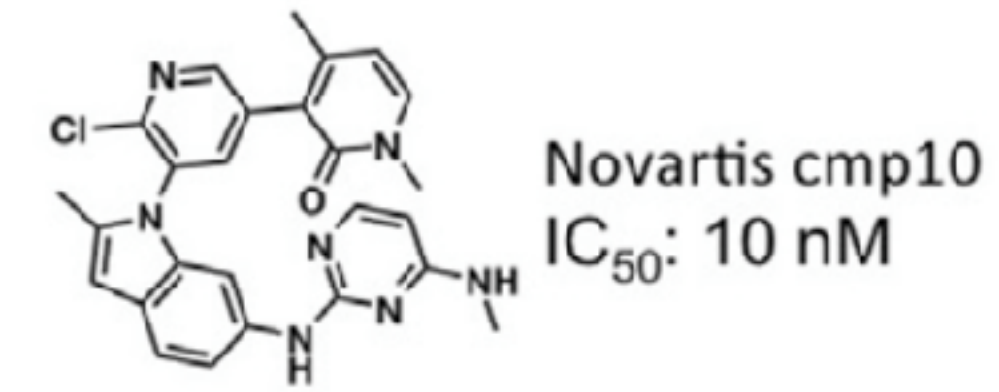
CONFORMATIONAL HETEROGENEITY IS BOTH A CHALLENGE AND OPPORTUNITY FOR CHEMICAL PROBE DESIGN



DOT1L active

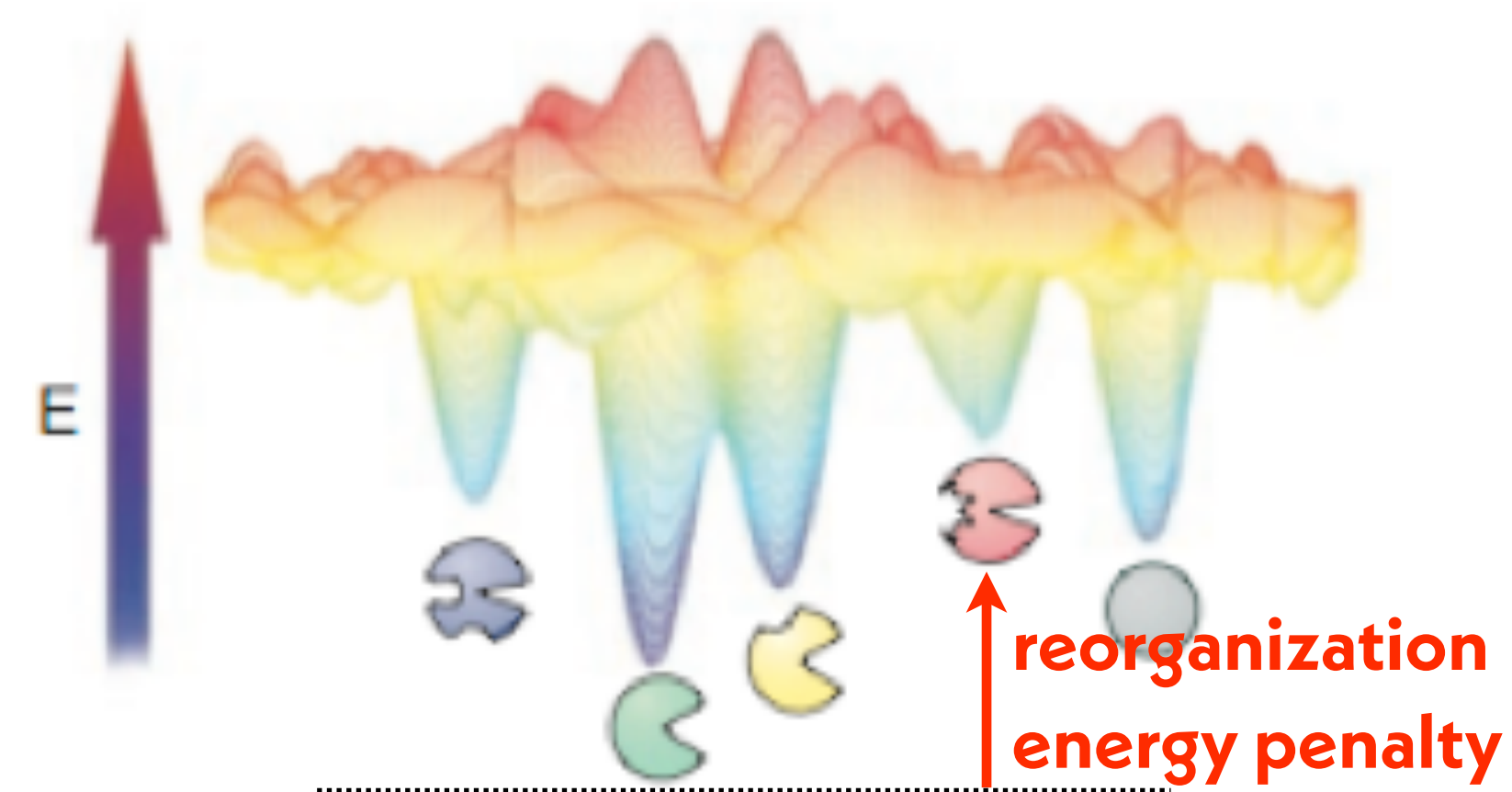
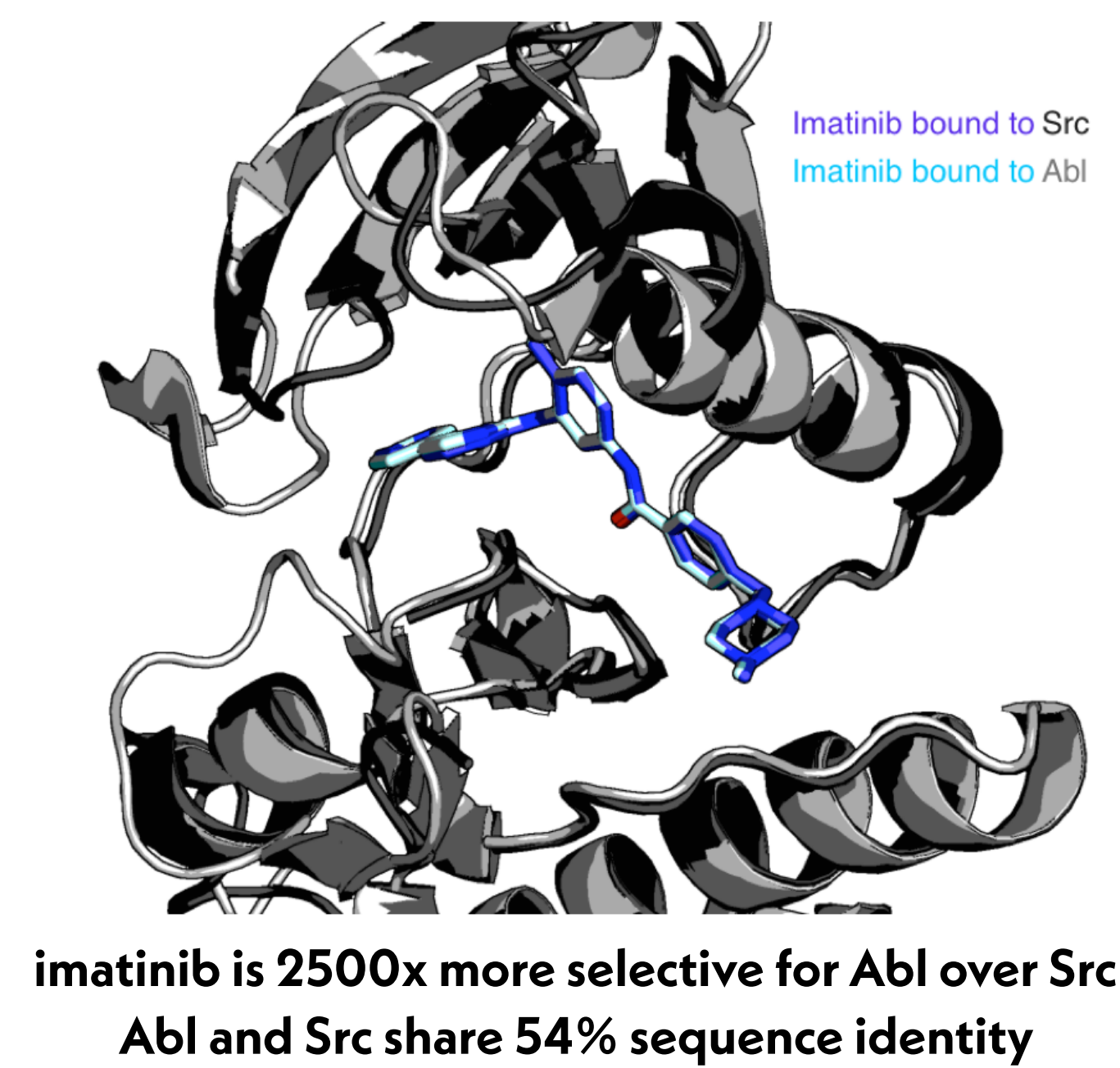


DOT1L inactive

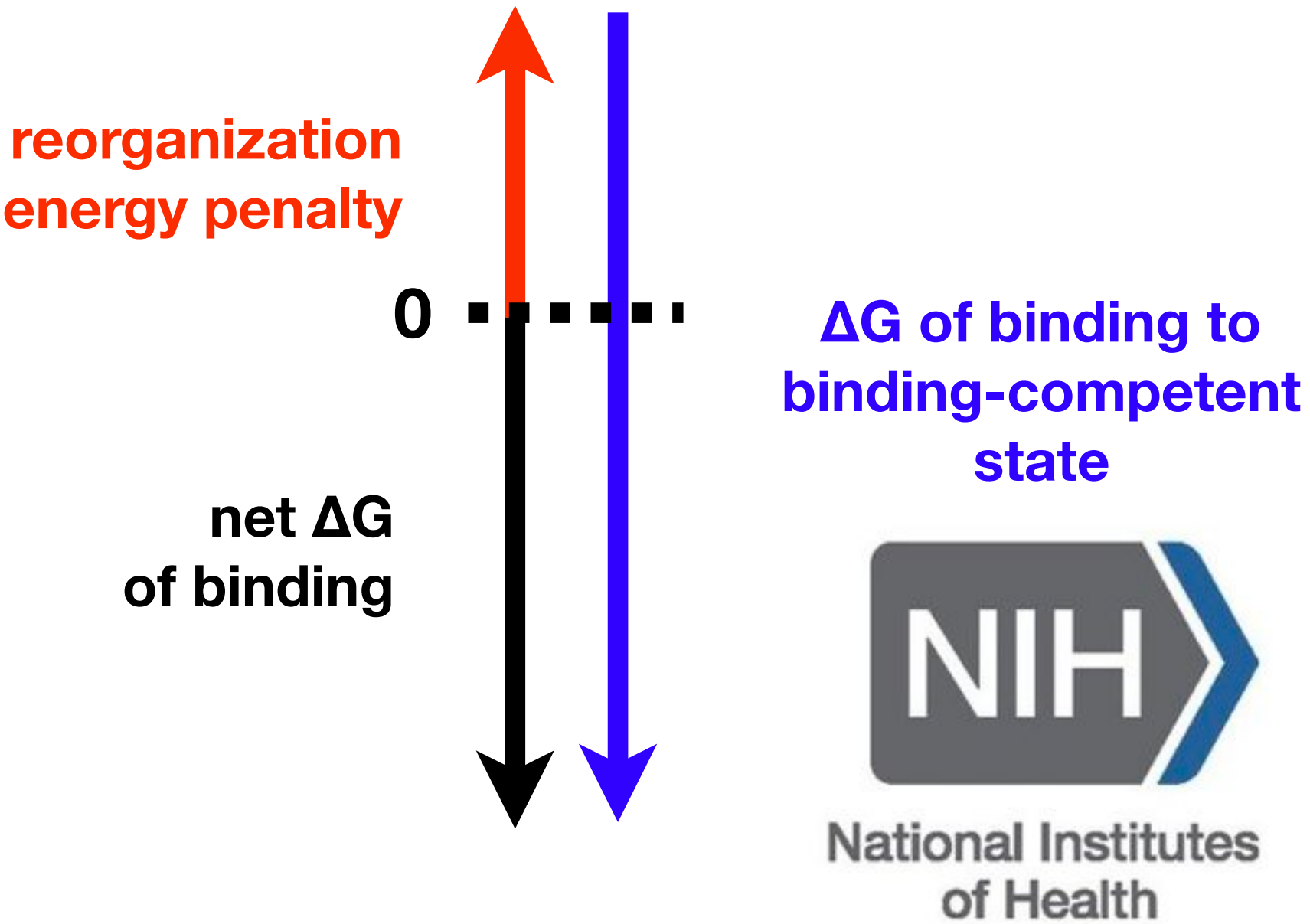


DOT1L inactive

IN KINASES, SOME KINASE INHIBITORS EXPLOIT DIFFERENCES IN ENERGIES OF BINDING-COMPETENT CONFORMATIONS TO ACHIEVE SELECTIVITY

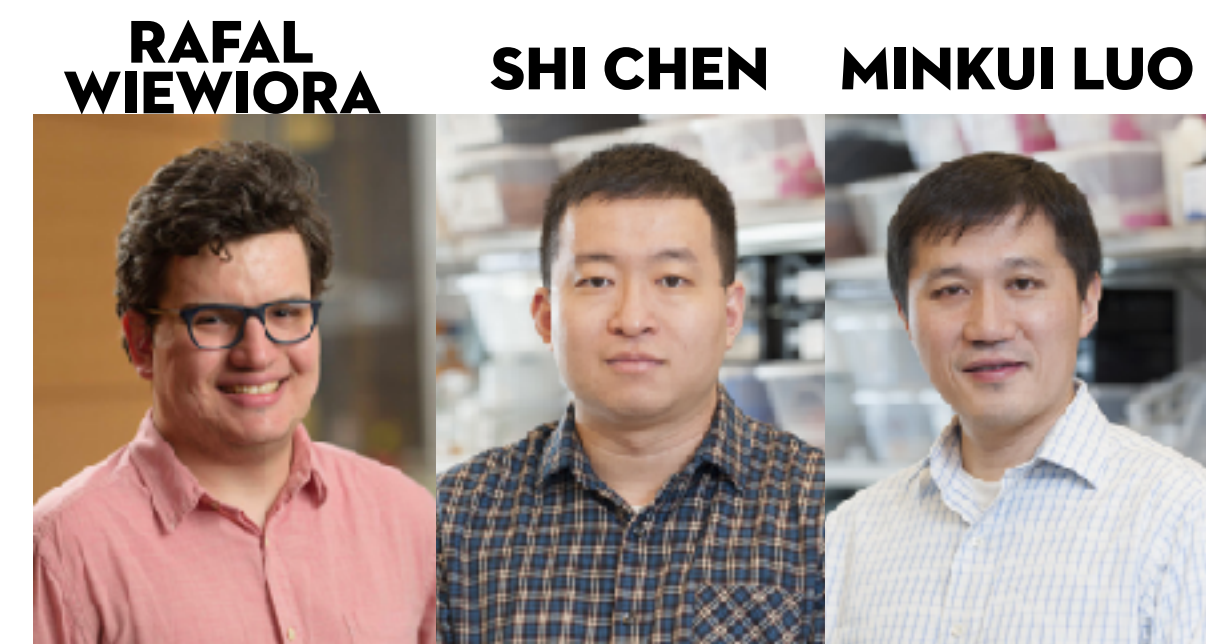
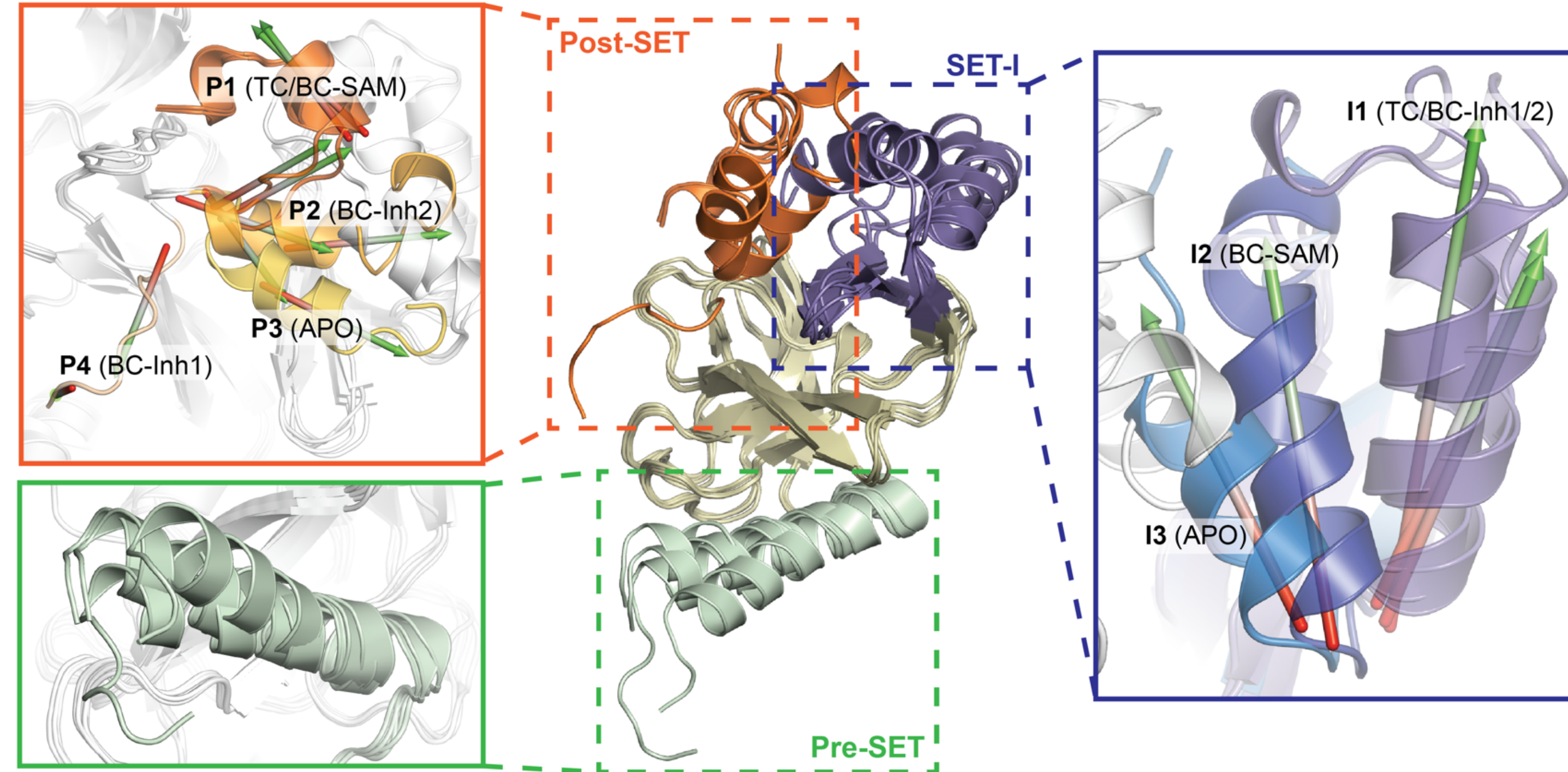
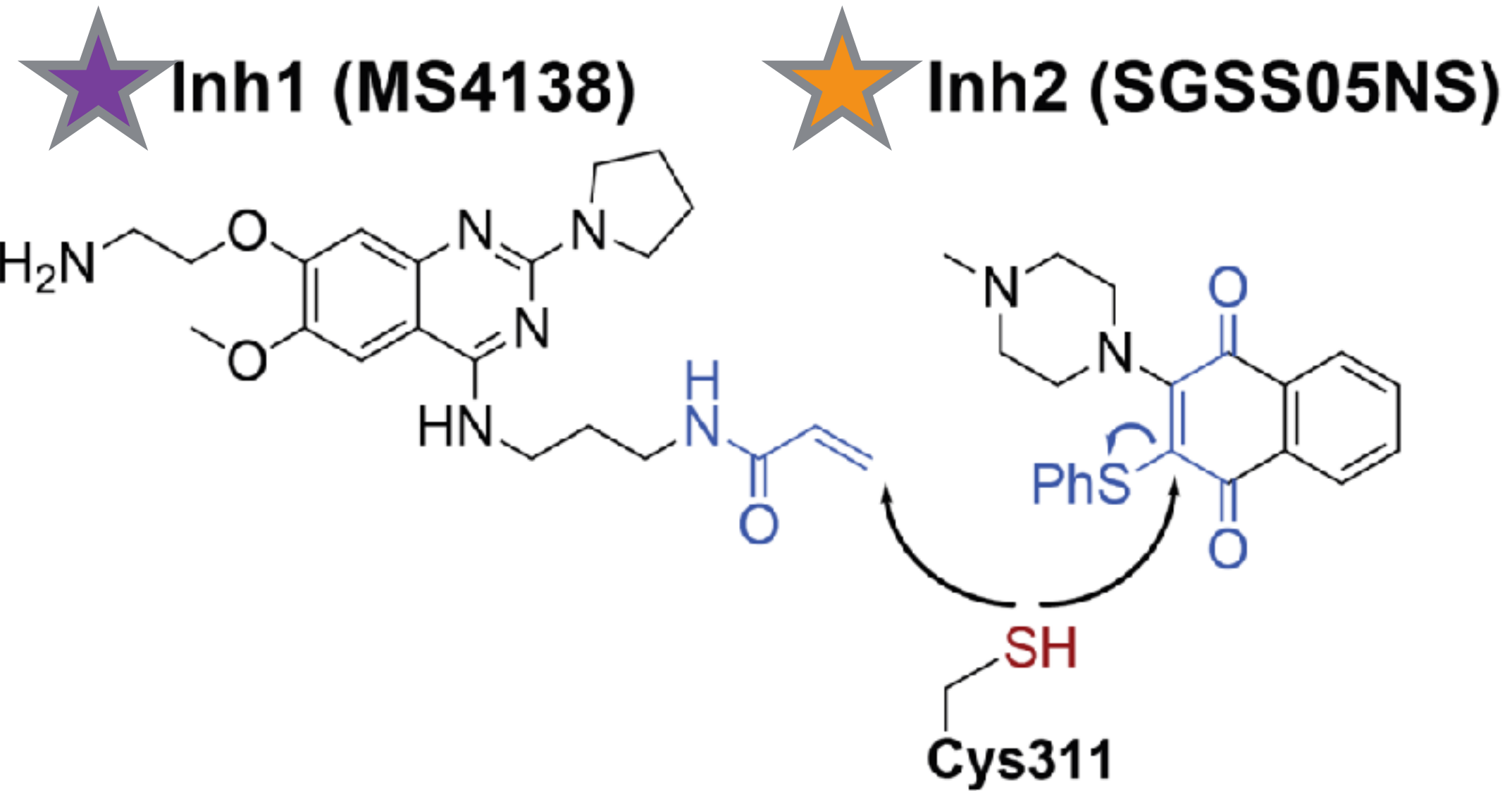
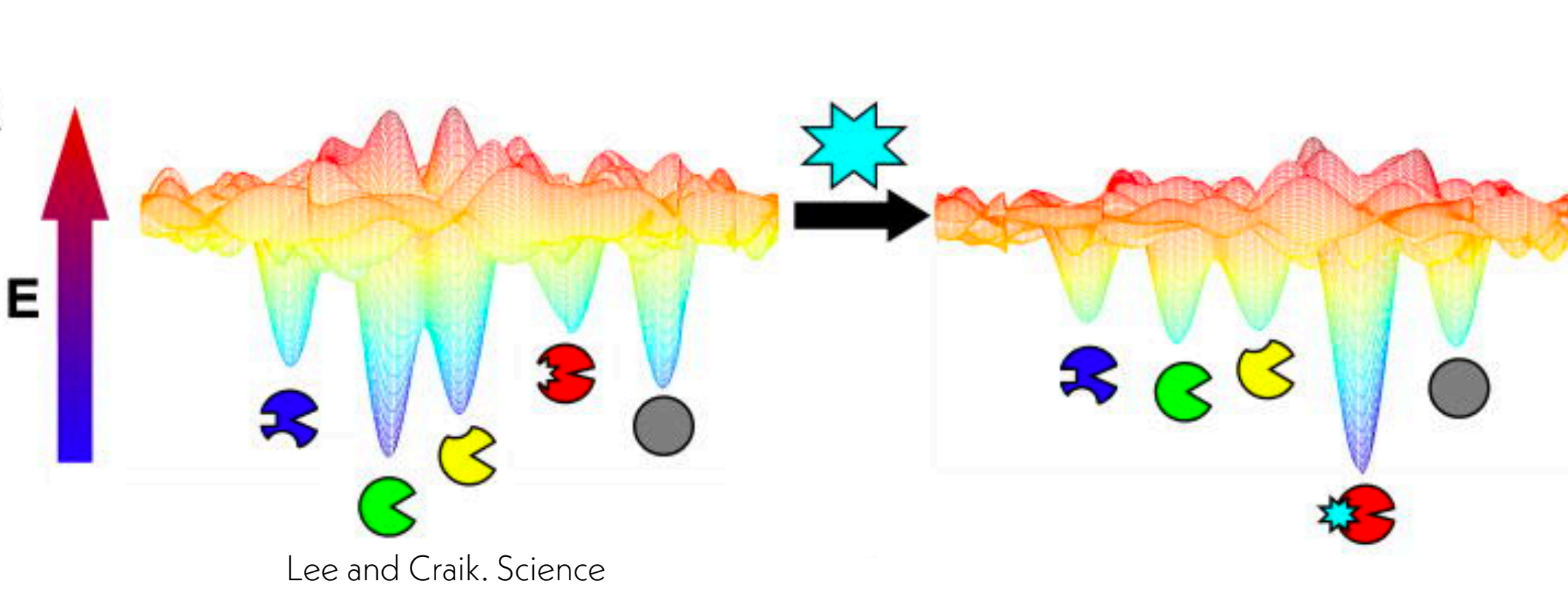


Lee and Craik. Science 324:213, 2009

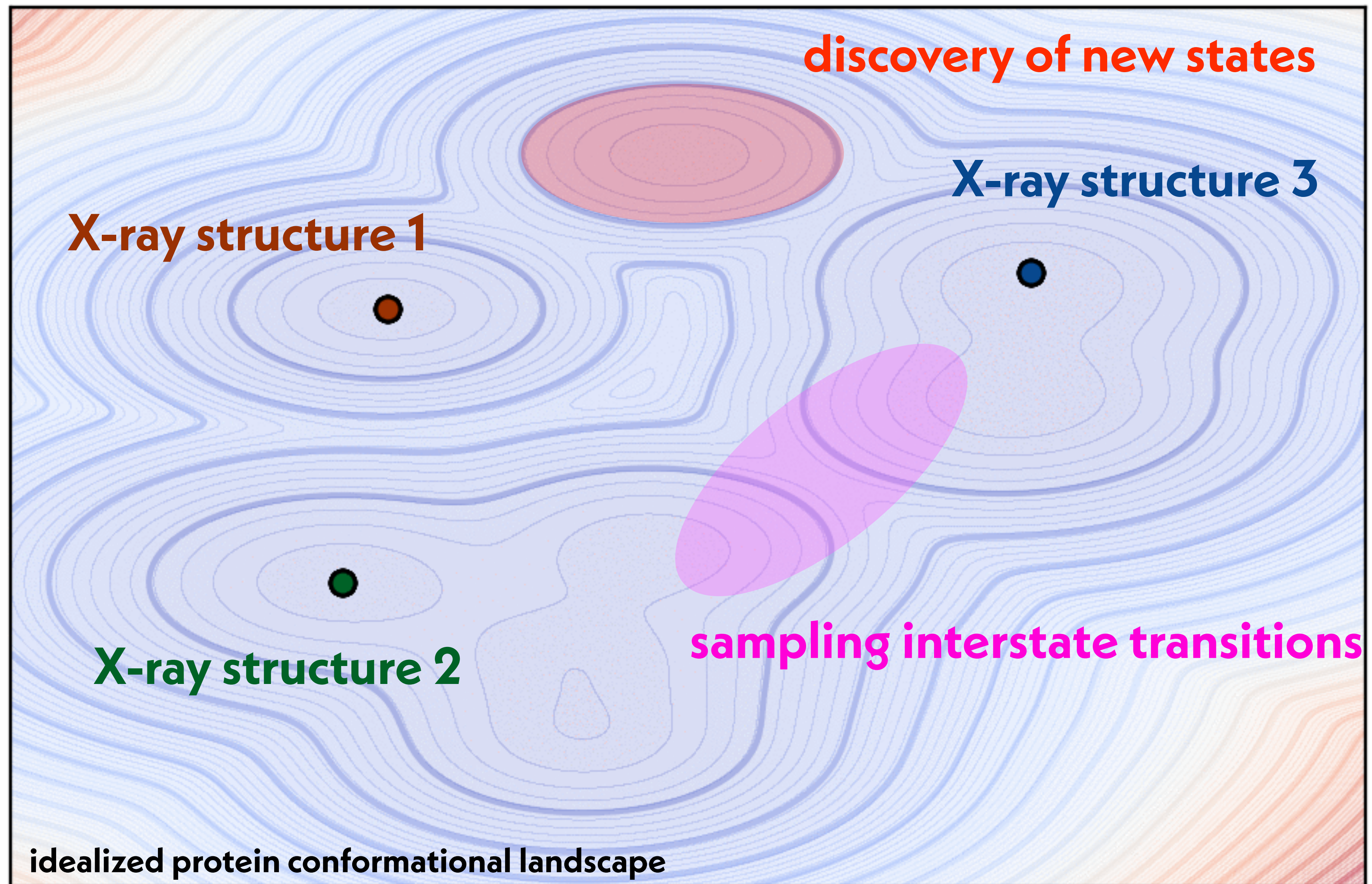


How can we understand the relative energetics of SETD8 conformations, as well as find druggable conformations we haven't seen before?

COVALENT LIGANDS CAN TRAP HIDDEN CONFORMATIONS OF SETD8 TO ALLOW VISUALIZATION VIA X-RAY CRYSTALLOGRAPHY



SIMULATIONS SEEDED FROM X-RAY STRUCTURES CAN IDENTIFY HIDDEN CONFORMATIONS AND CHARACTERIZE FUNCTIONAL DYNAMICS



FOLDING@HOME PROVIDES IMMENSE COMPUTATIONAL RESOURCES FOR PROBING BIOMOLECULAR DYNAMICS

<http://foldingathome.org>

THE FOLDING@HOME
CONSORTIUM (FAHC)

WHERE IT ALL BEGAN

The journey started in the year 2000, at Stanford University. The Pande Lab, directed by professor Vijay Pande, founded the Folding@home project. Ever since, the team has been researching protein folding, computational drug design and other types of molecular dynamics.

LEARN MORE



OS Type	Native TFLOPS*	x86 TFLOPS*	Active CPUs	Active Cores	Total CPUs
Windows	857	857	67,467	187,104	5,857,235
Mac OS X	91	91	8,083	85,382	217,033
Linux	87	87	6,383	26,457	882,200
NVIDIA GPU	1	2	4	4	348,371
ATI GPU	10,243	21,613	7,178	7,178	426,335
NVIDIA Fermi GPU	36,065	76,097	21,570	21,587	624,822
Total	47,344	98,747	110,685	327,712	8,355,996

1927698 people have non-anonymously contributed to Folding@home.



GREG BOWMAN
WUSTL



VINCENT VOELZ
TEMPLE UNIVERSITY



JOHN CHODERA
MSKCC

~100 PFLOP/S OF AGGREGATE COMPUTATIONAL POWER!

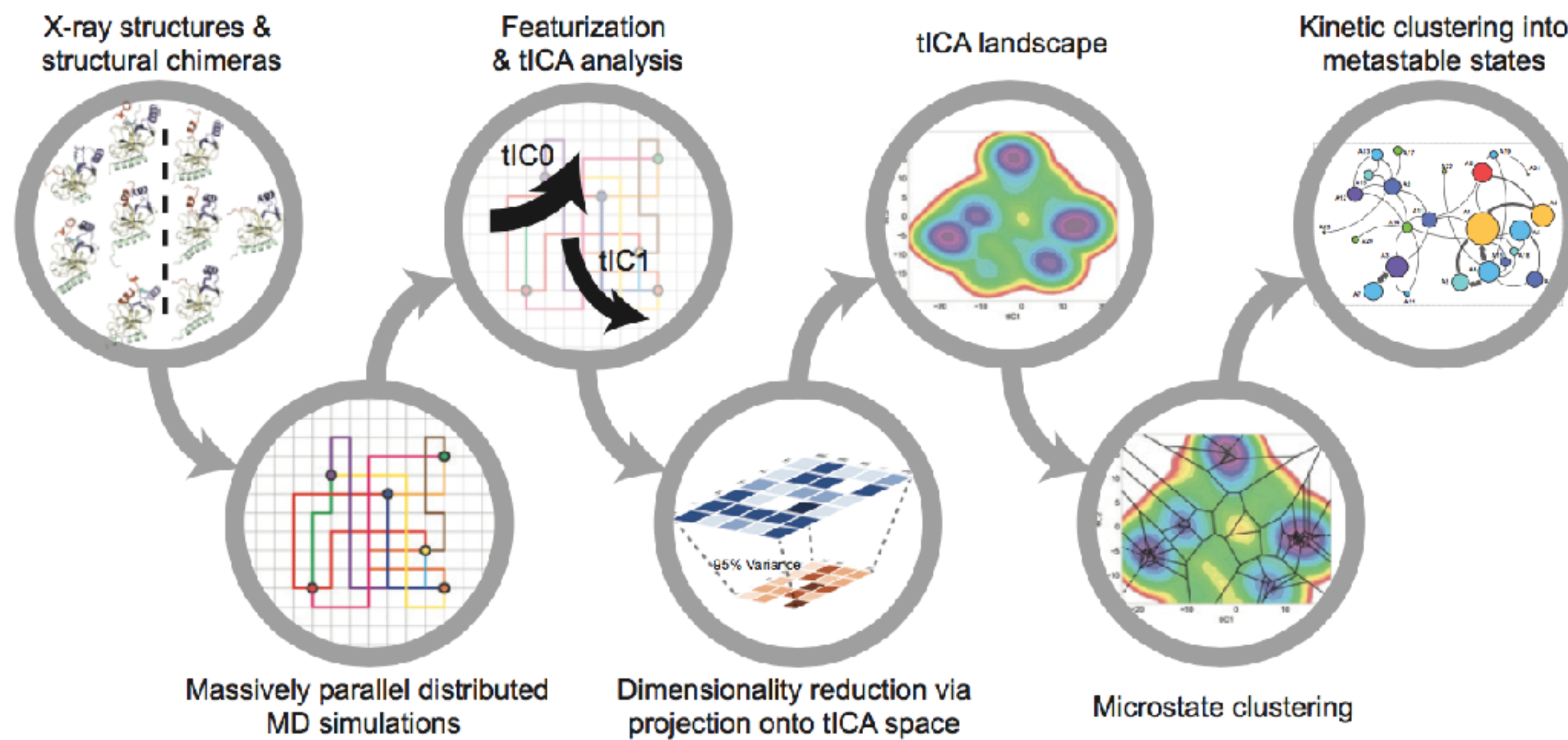
THE POWER OF CITIZEN SCIENCE

Sunway TaihuLight was the first single supercomputing installation to reach 100 PFLOP/s (at same time Folding@home did) cost **\$273M**

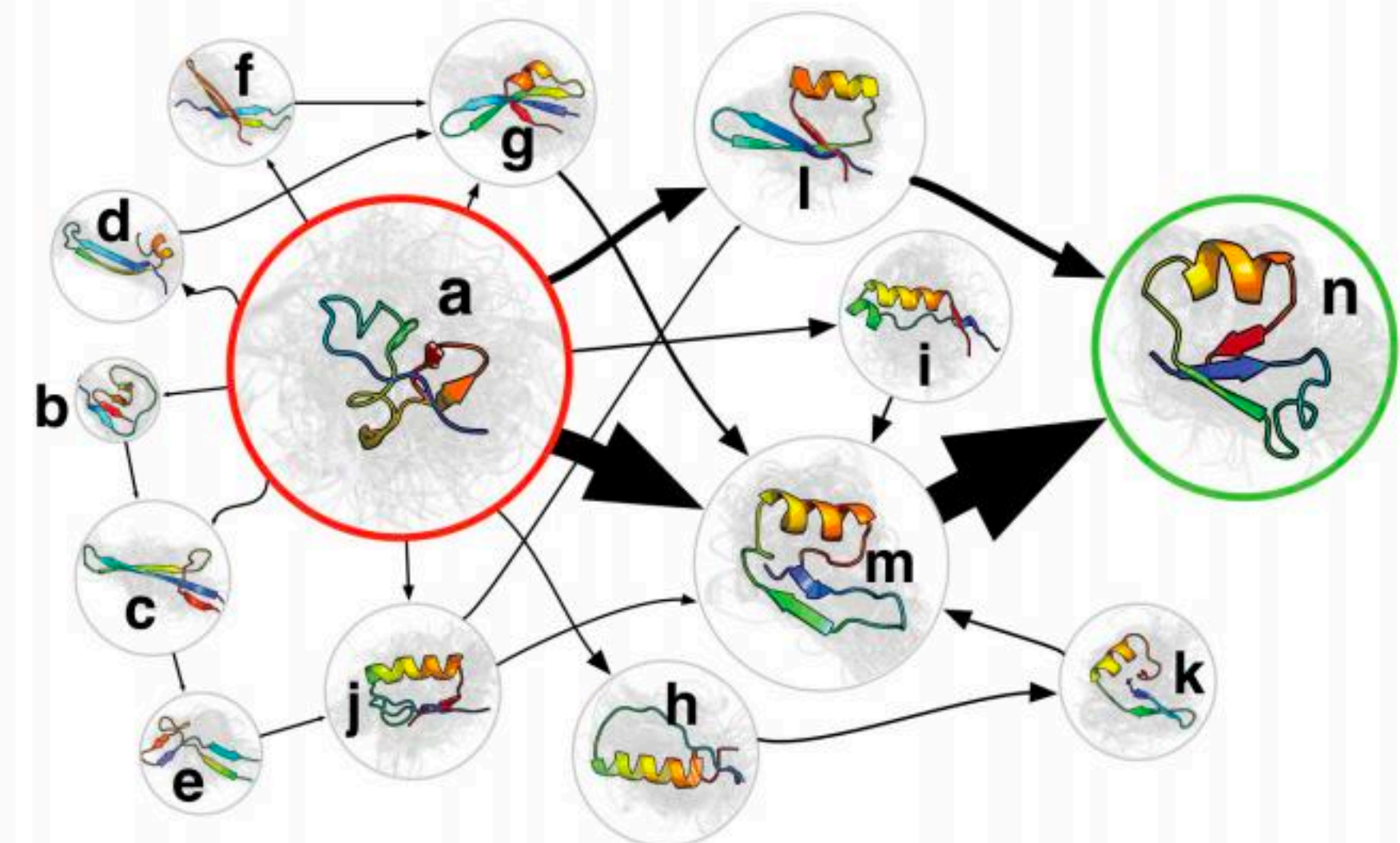


We collected **six milliseconds** of aggregate simulation data in **ten million** snapshots

WE DEVELOPED MACHINE LEARNING APPROACHES TO DISTILL MASSIVE SIMULATION DATASETS INTO SIMPLE KINETIC MODELS



first millisecond folding mechanism (NTL9)



Schwantes, Pande. J Chem Theor Comput 11:600, 2015

Pérez-Hernandez, Paul, Giorgino, De Fabritiis, Noé. J Chem Phys 139:015012, 2013

Prinz, Wu, Sarich, Keller, Fischbach, Held, **Chodera**, Schütte, Noé. J Chem Phys 134:174105, 2011

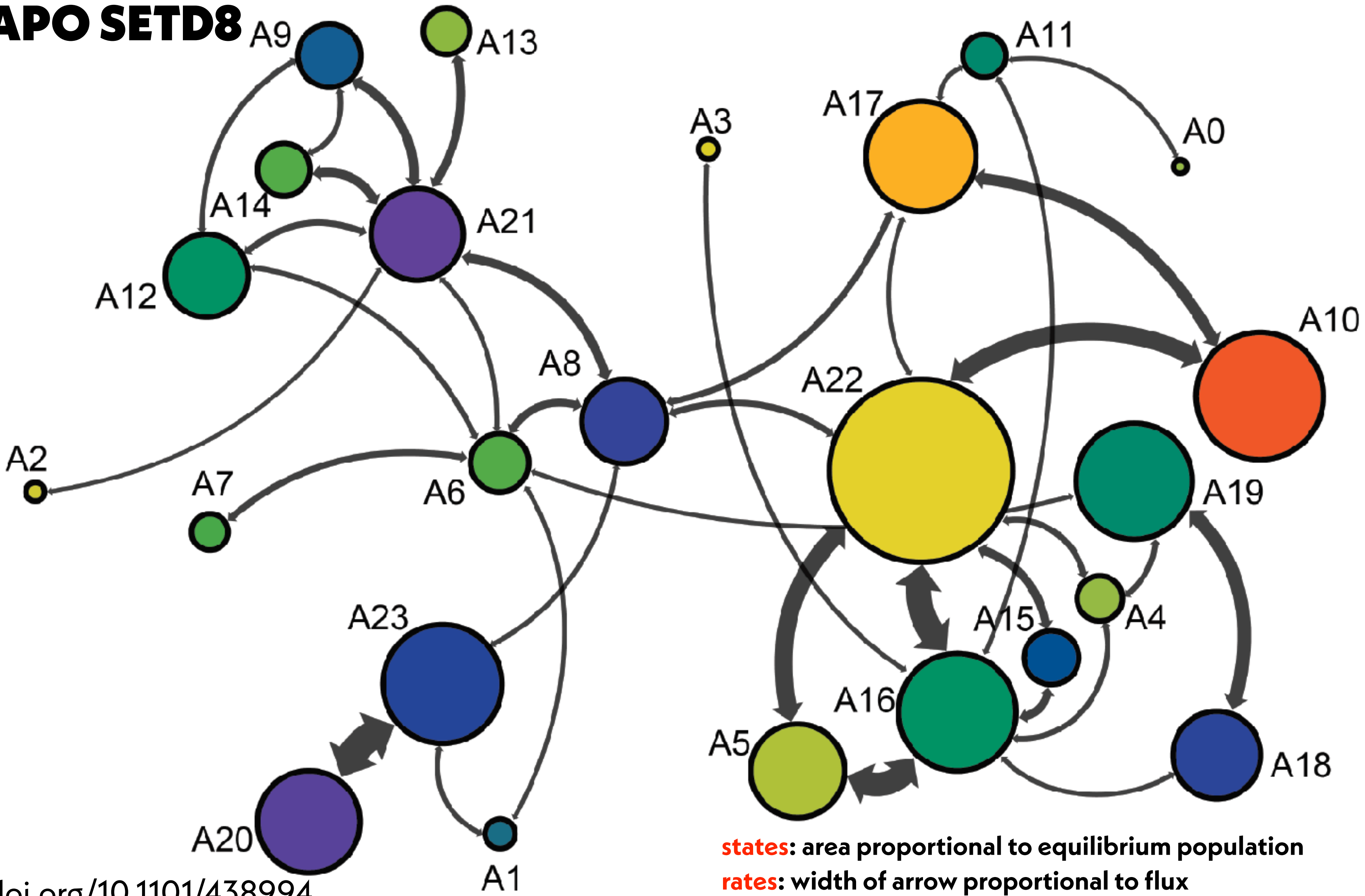
Bacallado, **Chodera**, and Pande. J Chem Phys 131:045106, 2009

Chodera*, Singhal*, Swope, Pitner, Pande, Dill. J Chem Phys 126:155101, 2007

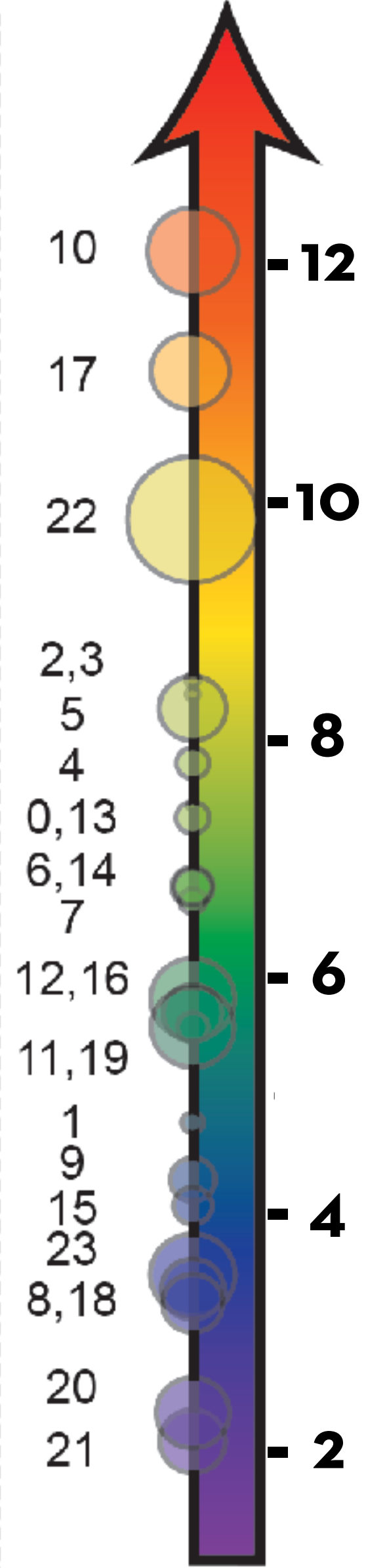
Chodera, Swope, Pitner, and Dill. Multiscale Model Simul 5:1214, 2006

**DISTILLING MILLISECONDS OF SIMULATIONS REVEALS THE
CONFORMATIONAL LANDSCAPE OF APO SETD8**

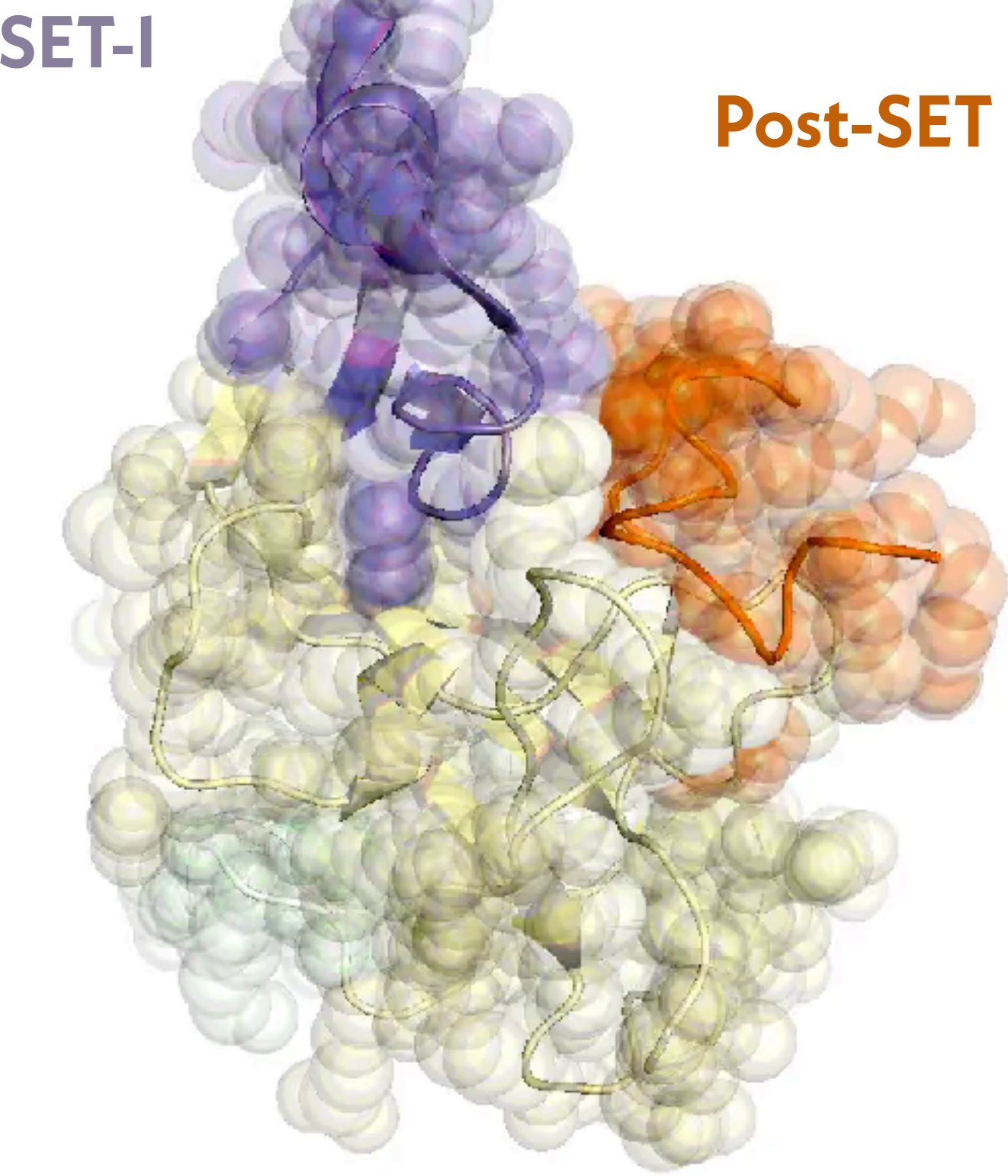
APO SETD8



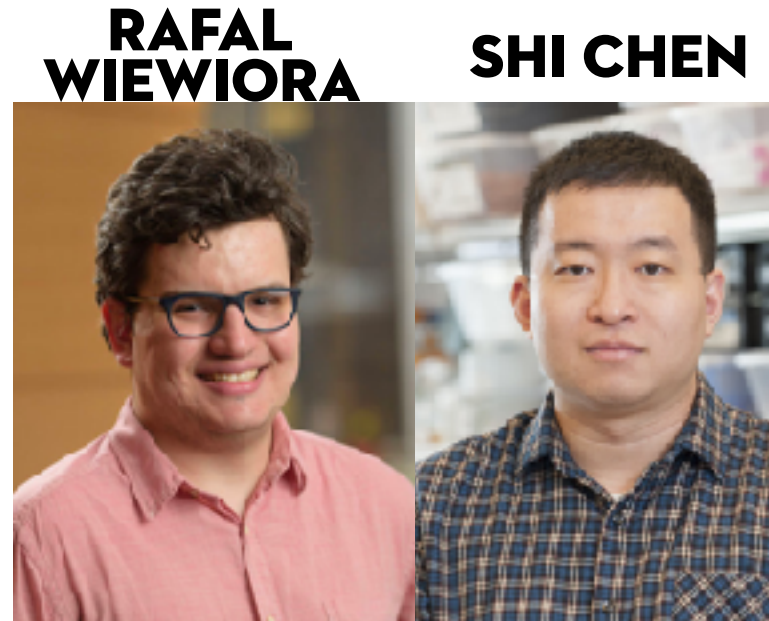
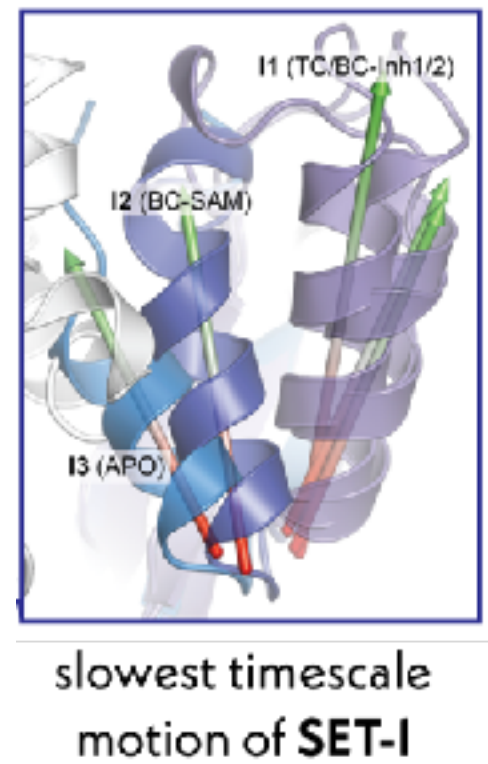
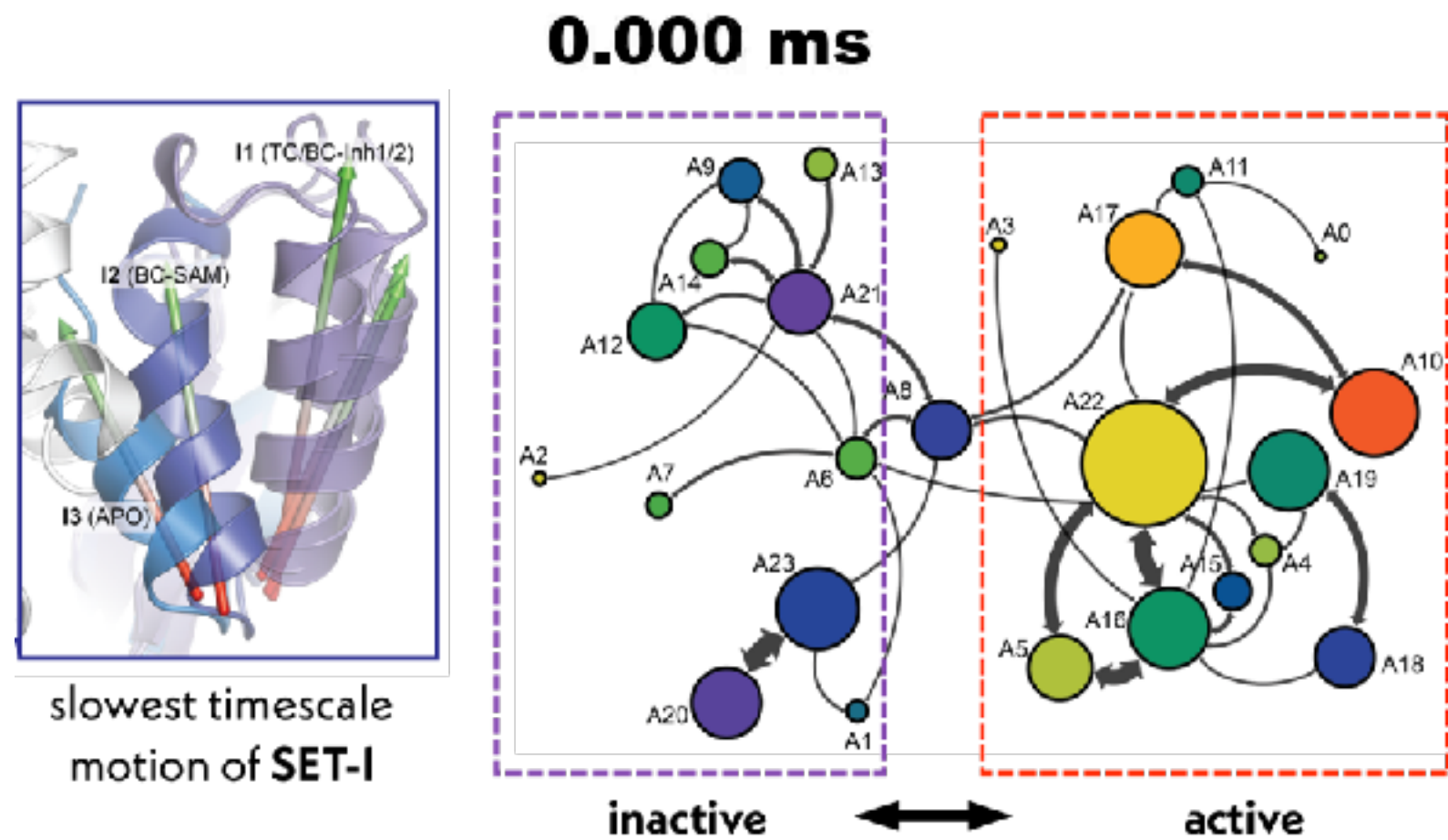
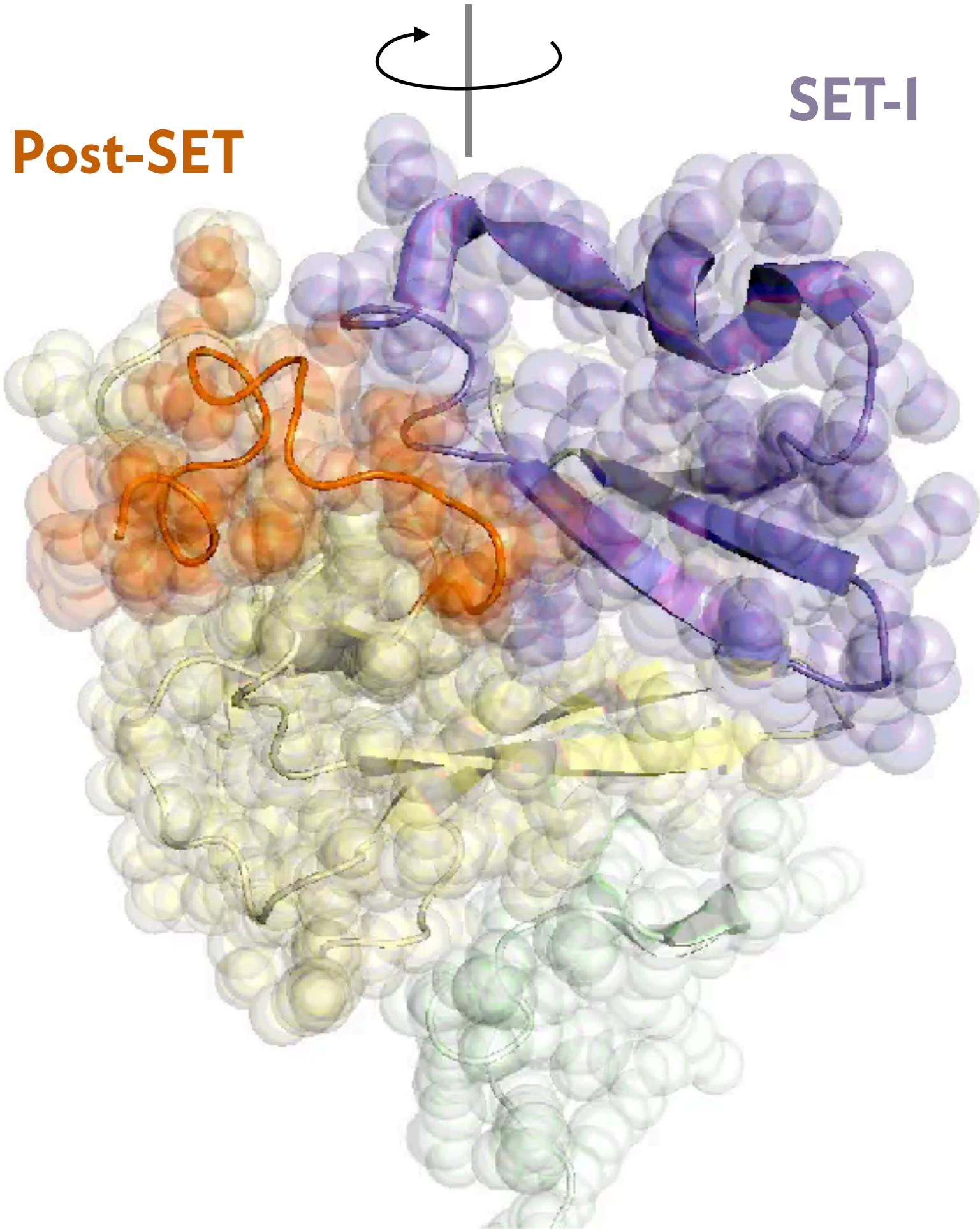
RMSD to APO (Å)



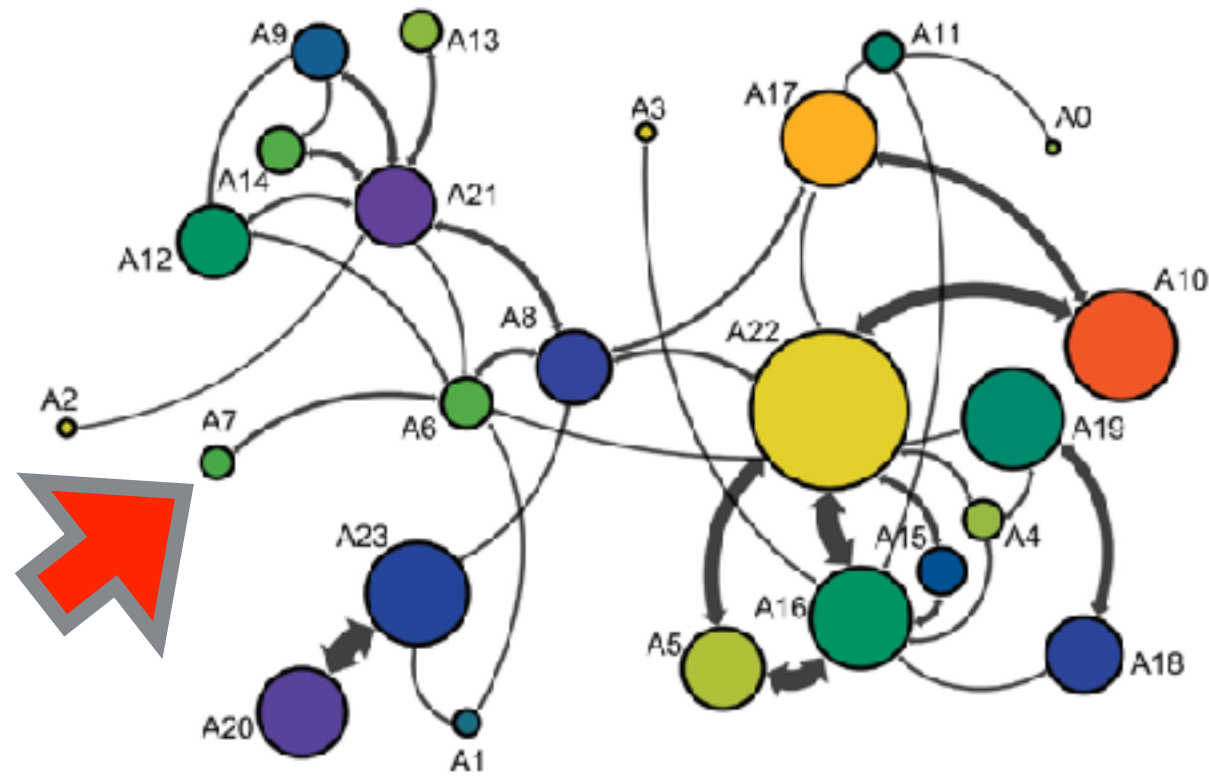
THE MARKOV STATE MODEL SUMMARIZES SLOW KINETICALLY DISTINCT CONFORMATIONAL STATES



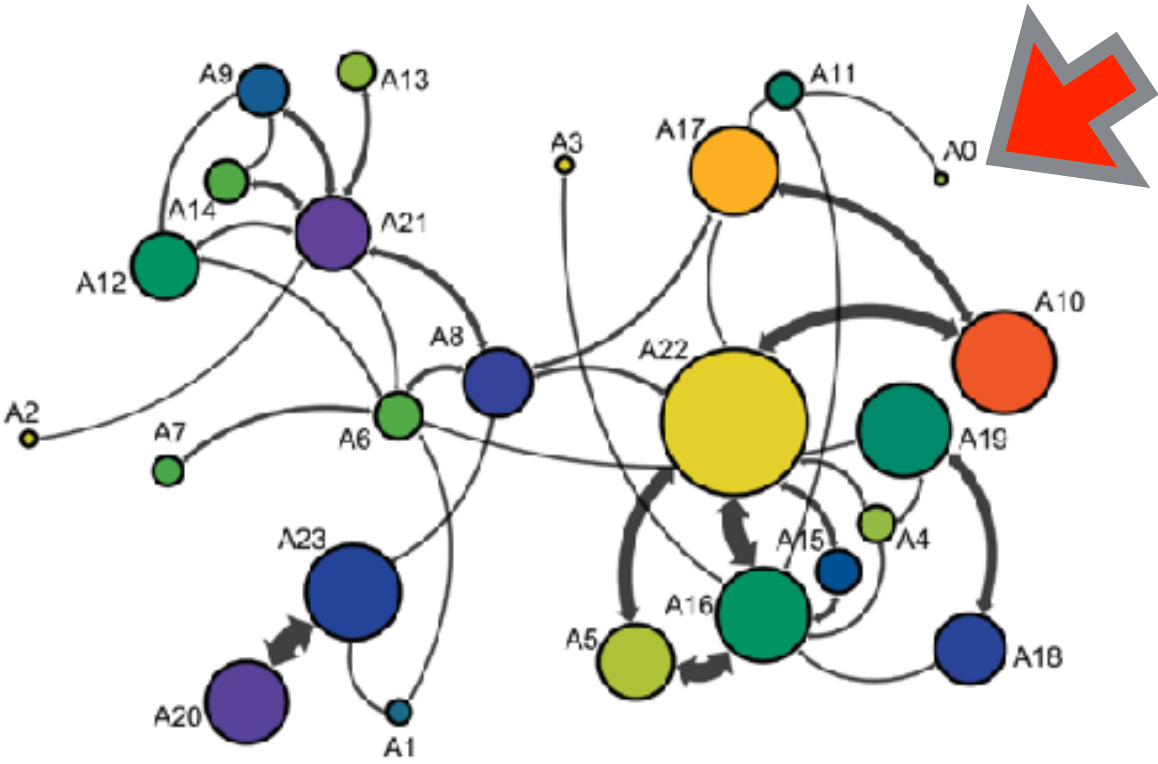
APO SETD8



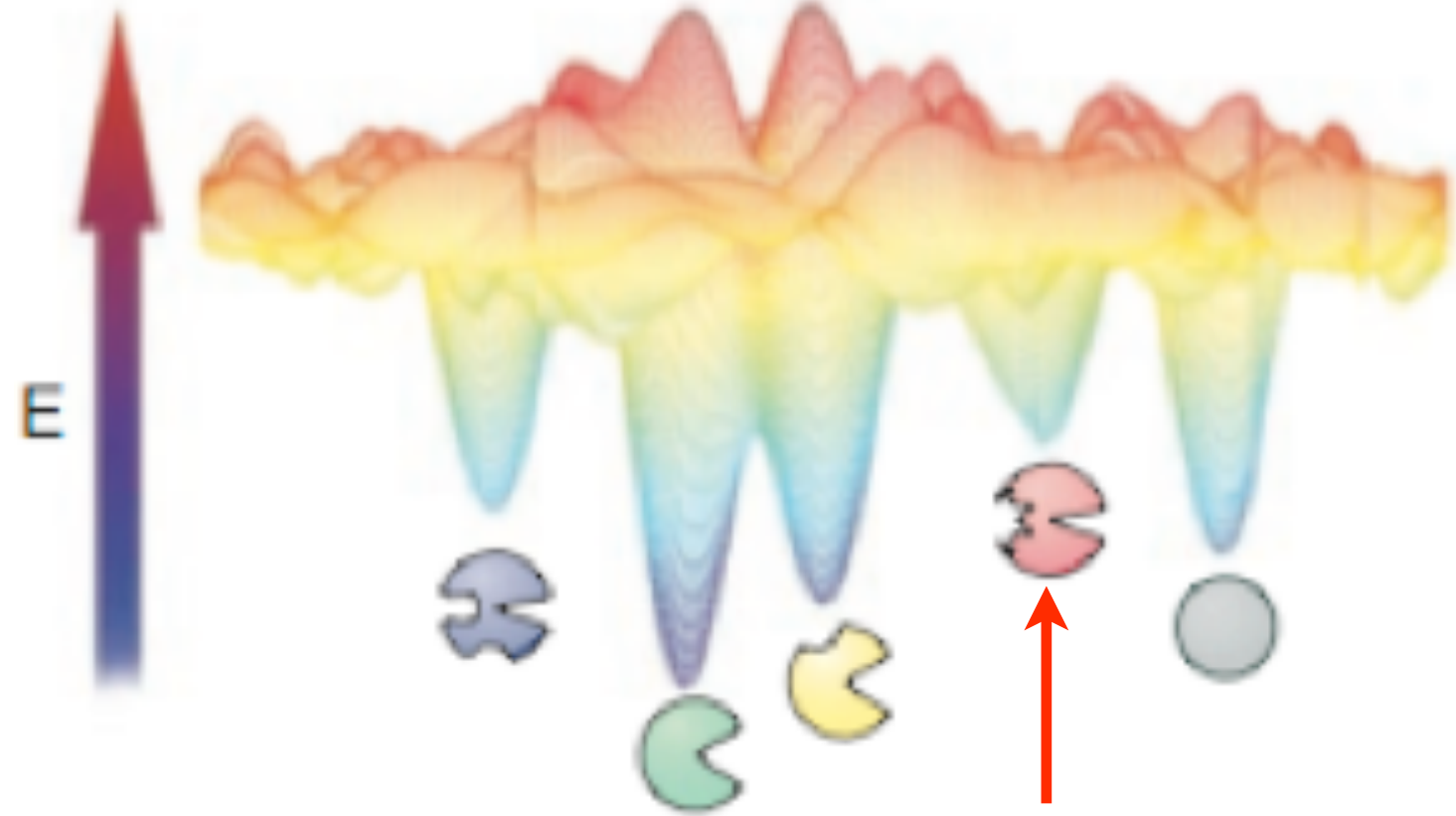
SAM-BOUND AND TERNARY-COMPLEX-LIKE STATES ARE FOUND IN APO SETD8 LANDSCAPE AT LOW POPULATIONS



SAM•SETD8
0.8%



SAM•substrate•SETD8
0.1%

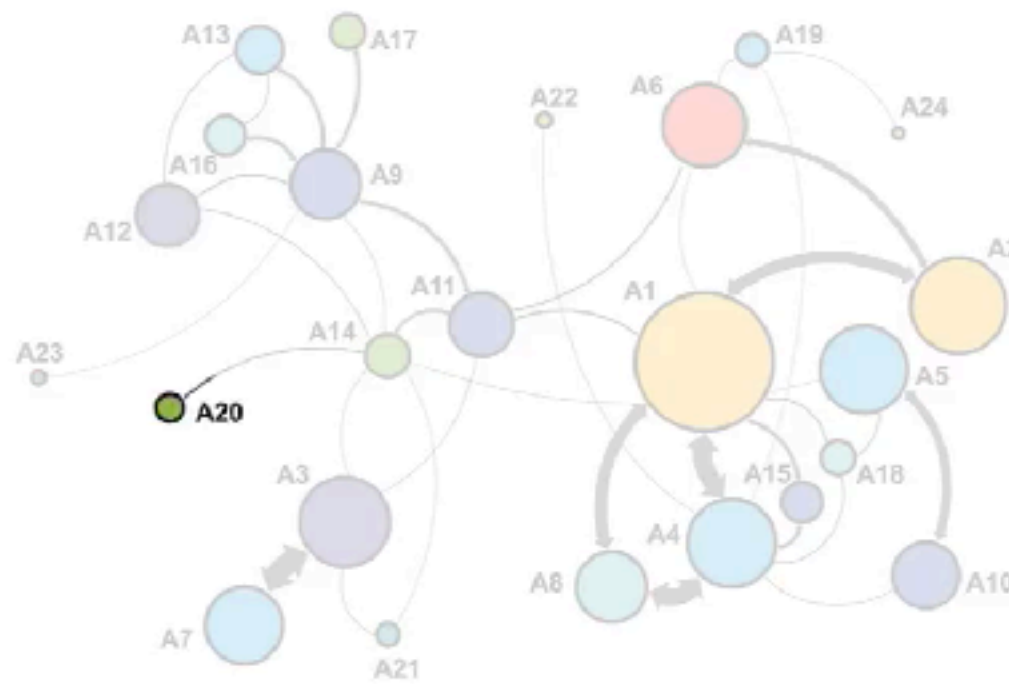
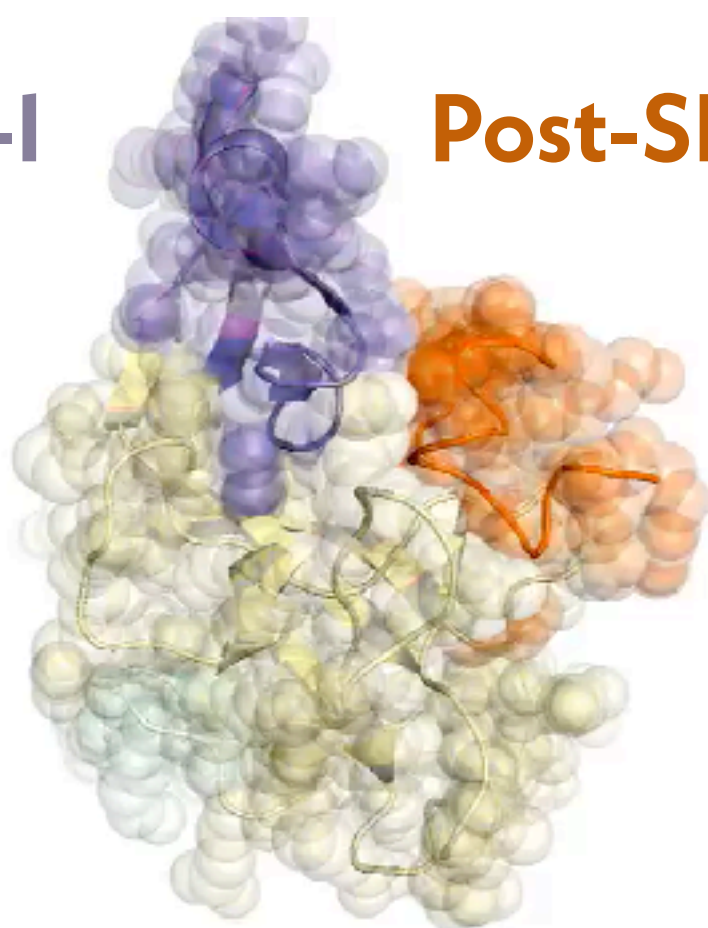


Lee and Craik. Science 324:213, 2009

SAM BINDING SIGNIFICANTLY RESTRICTS CONFORMATIONAL DYNAMICS

APO SETD8

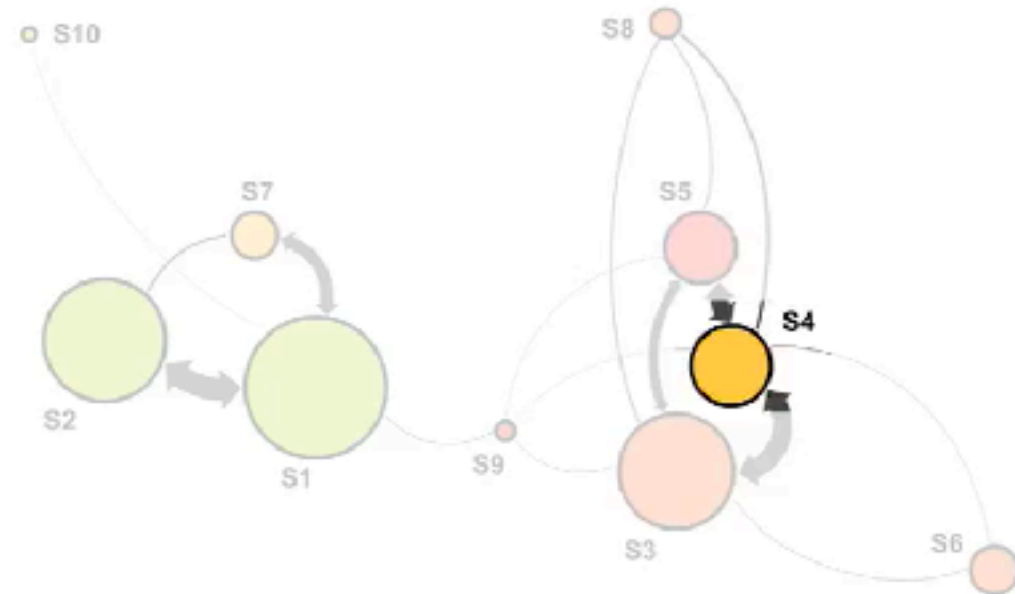
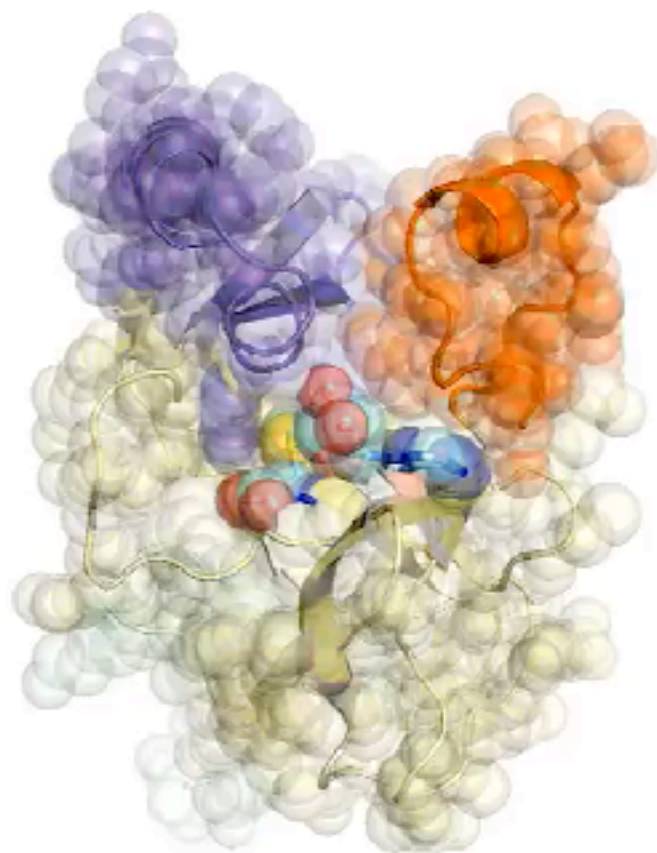
SET-I Post-SET



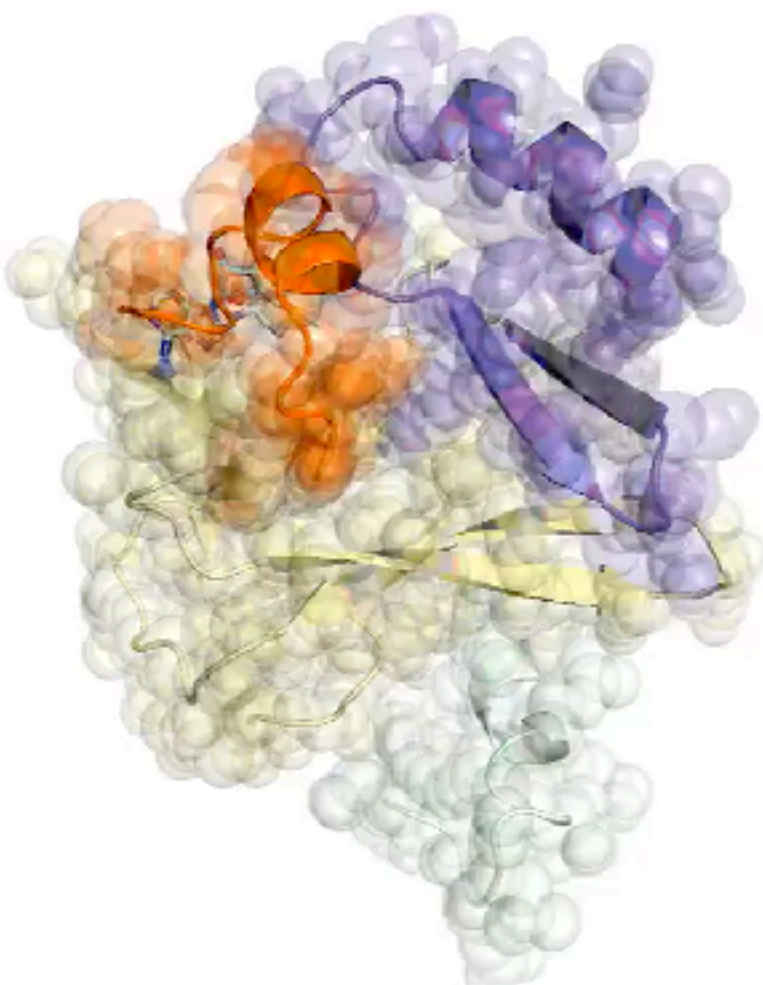
0.000 ms



SAM•SETD8



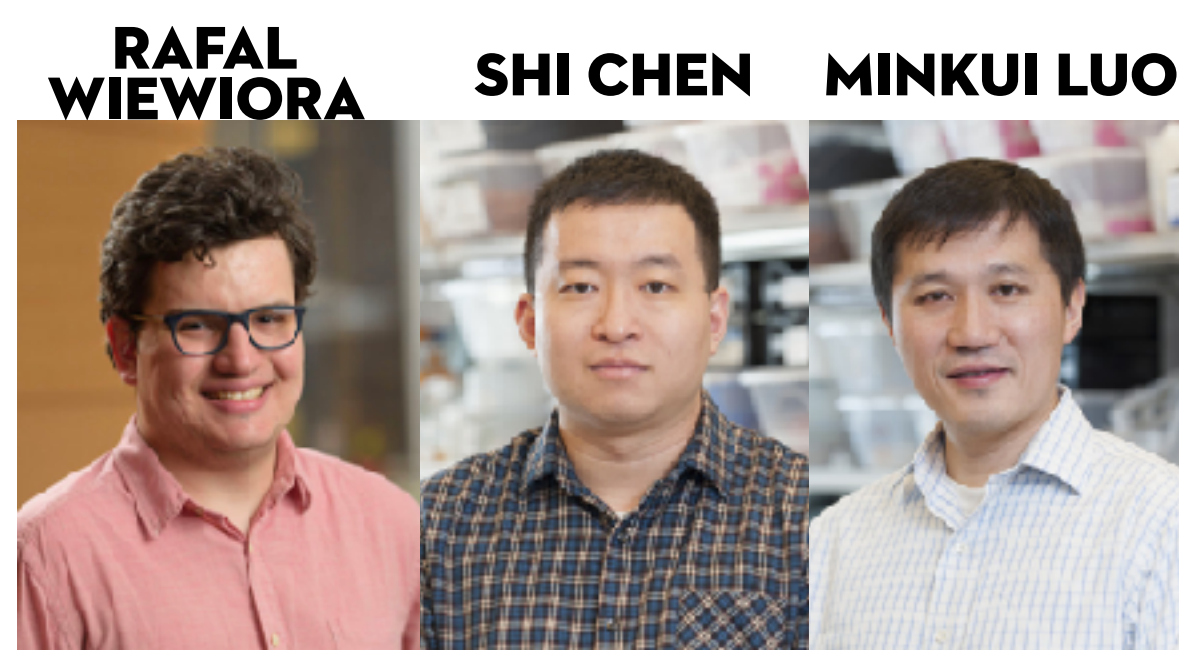
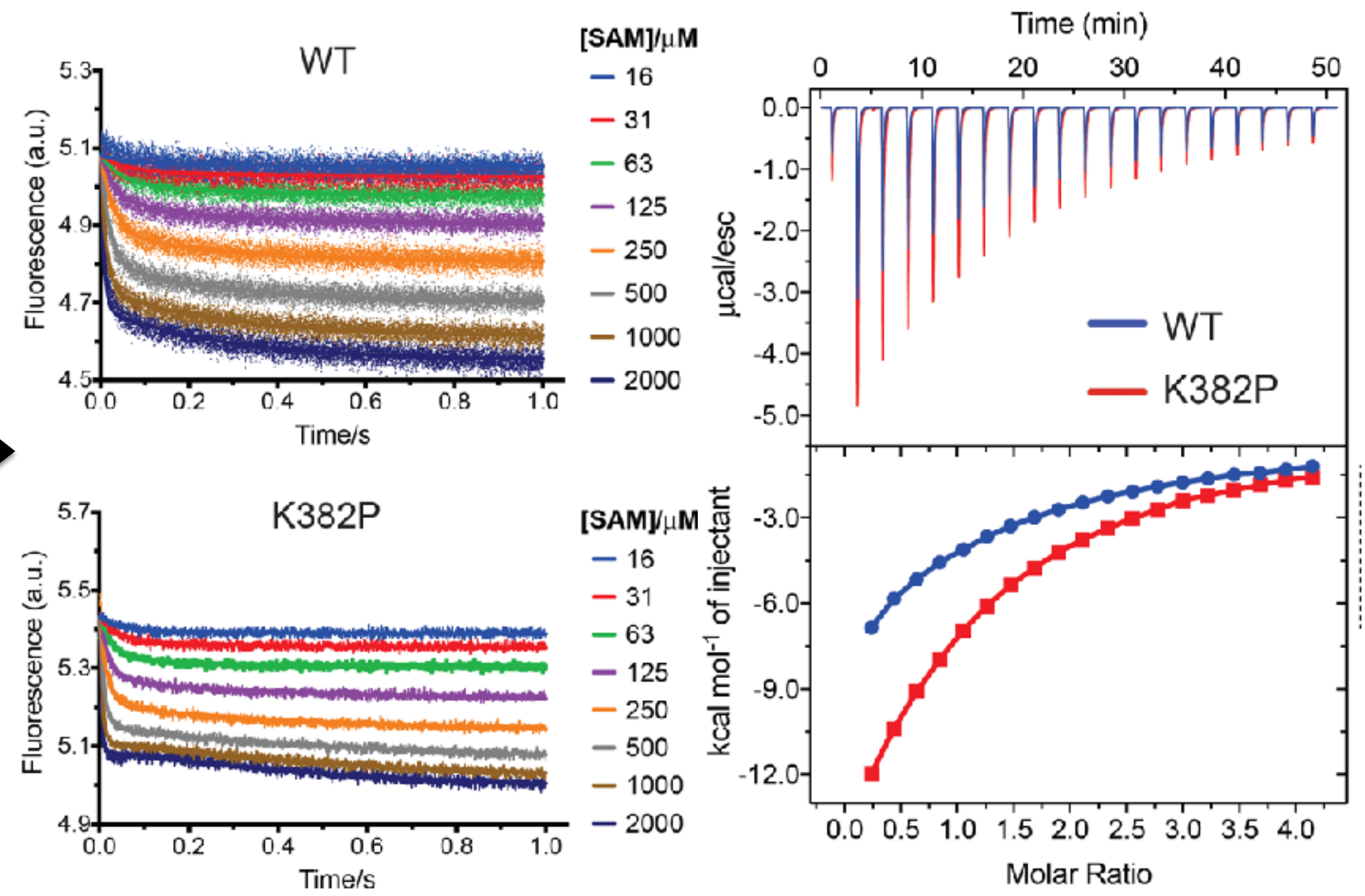
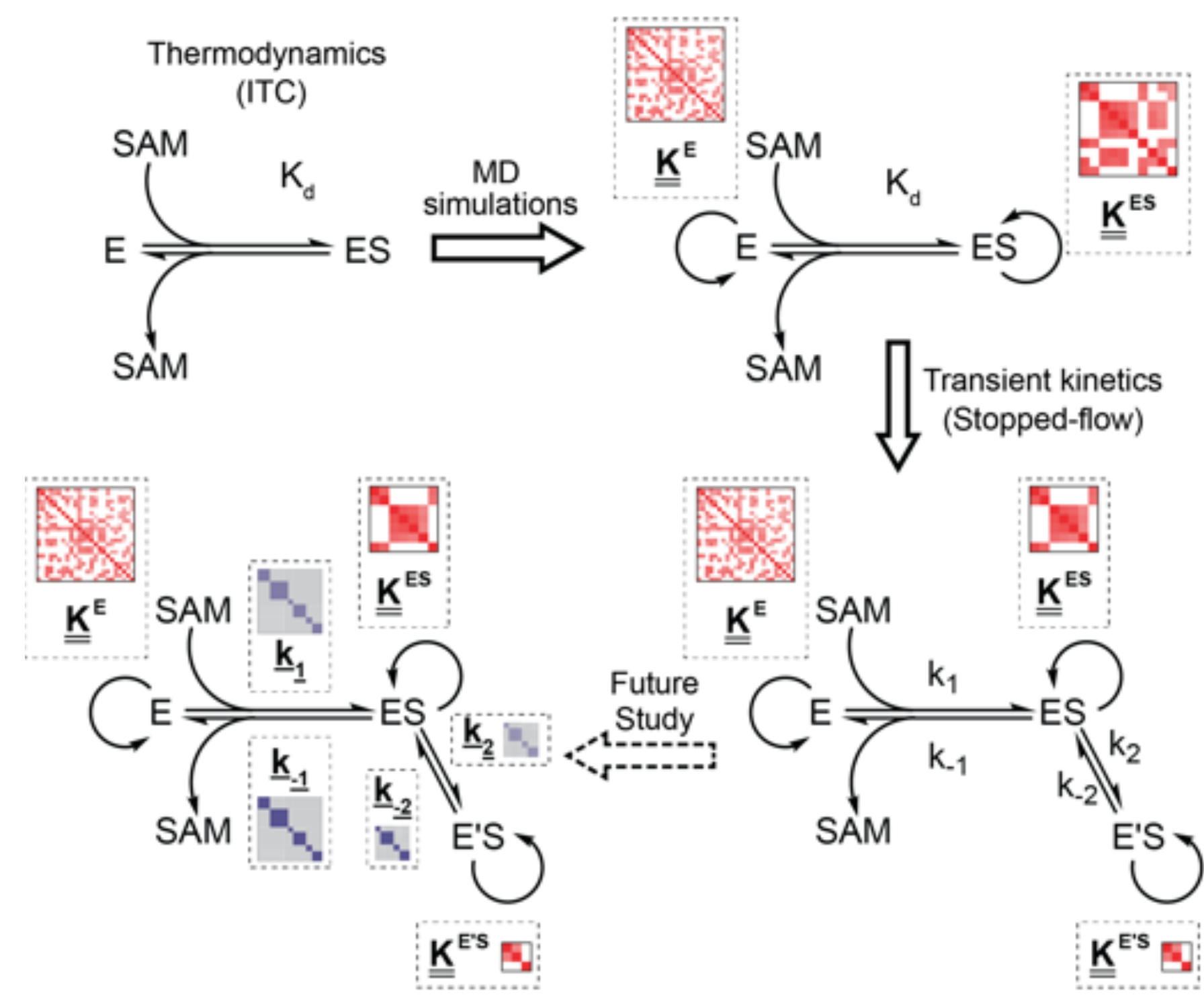
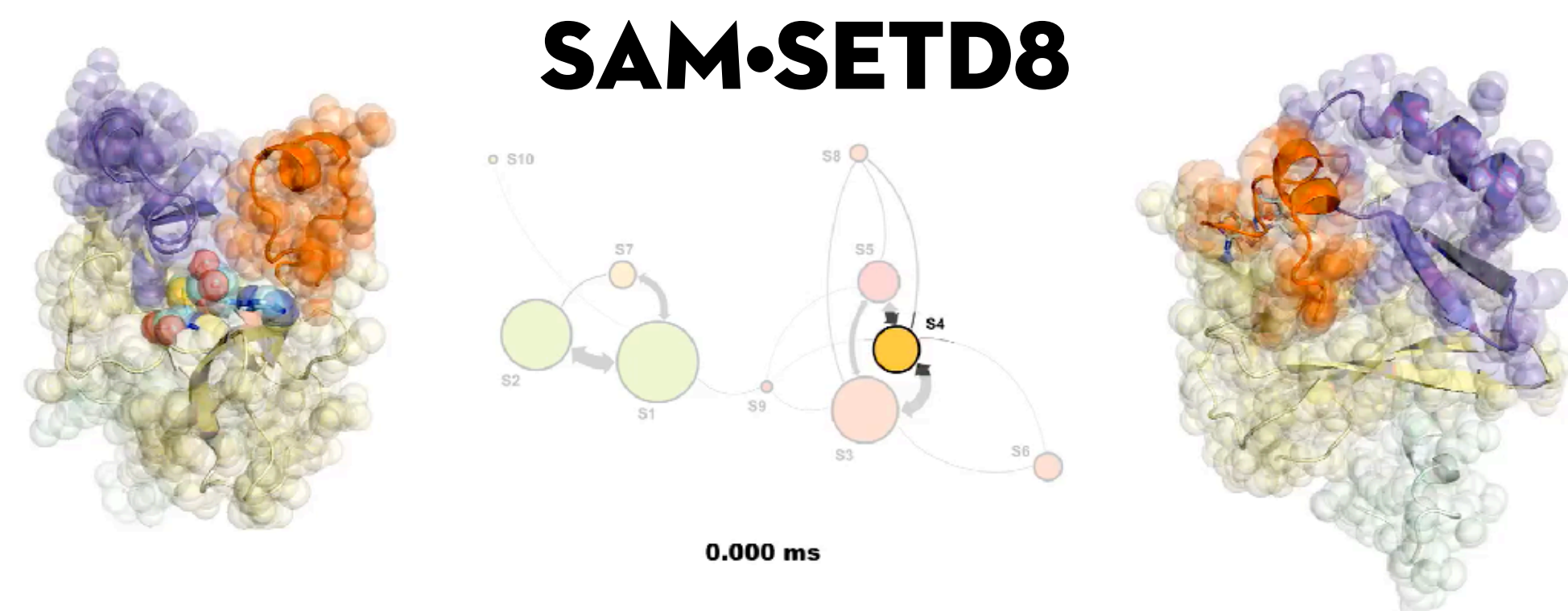
0.000 ms



DYNAMIC MODELS INFORM DESIGN OF GAIN-OF-FUNCTION MUTATIONS

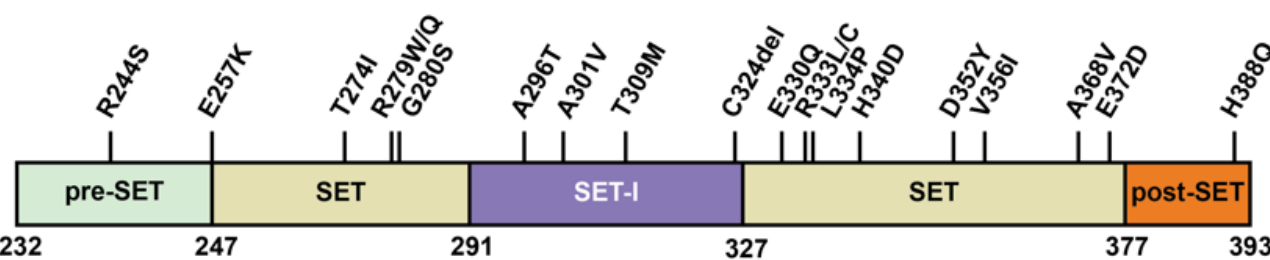
gain-of-function mutations proposed from model

stopped-flow fluorescence and ITC experiments

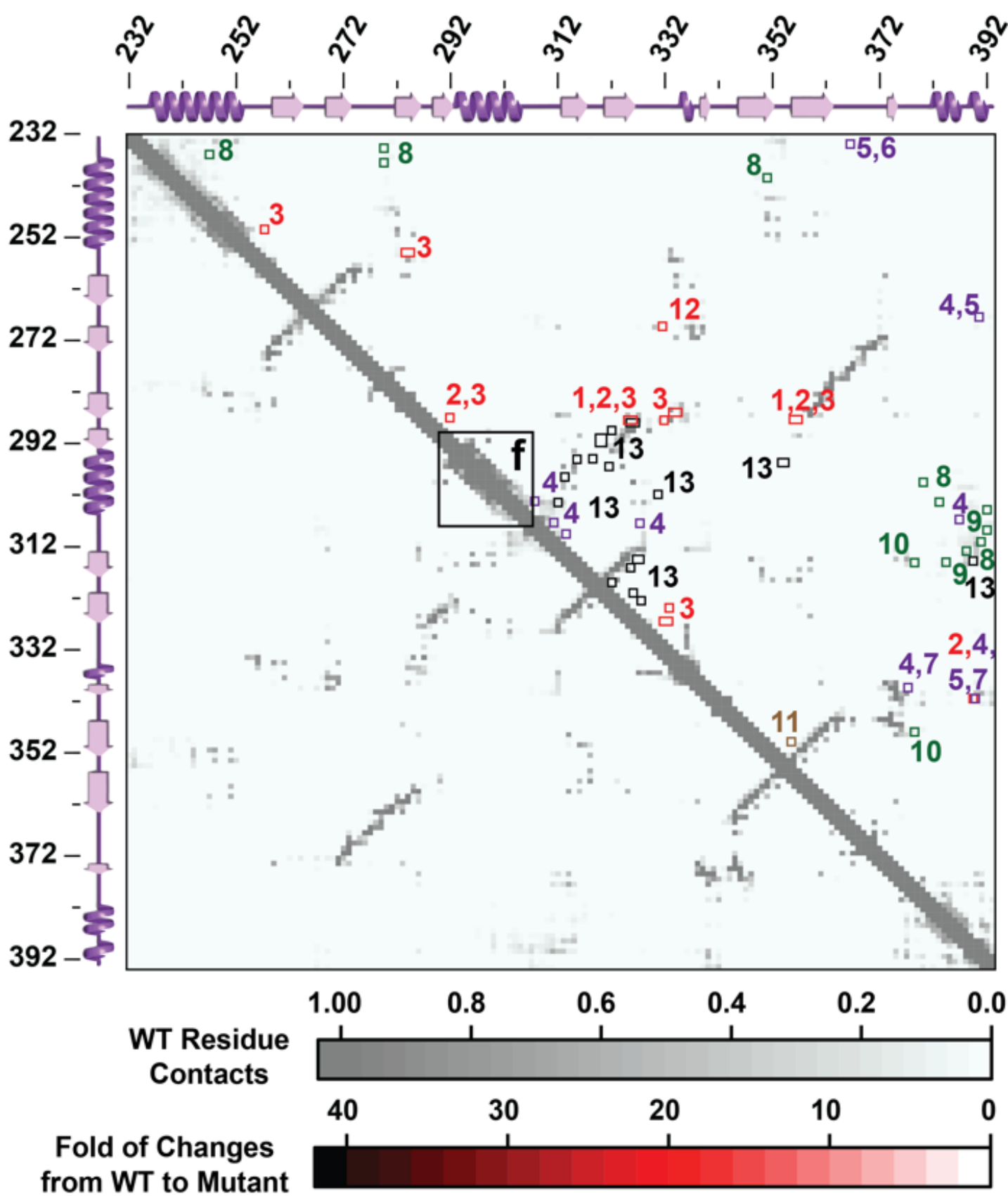
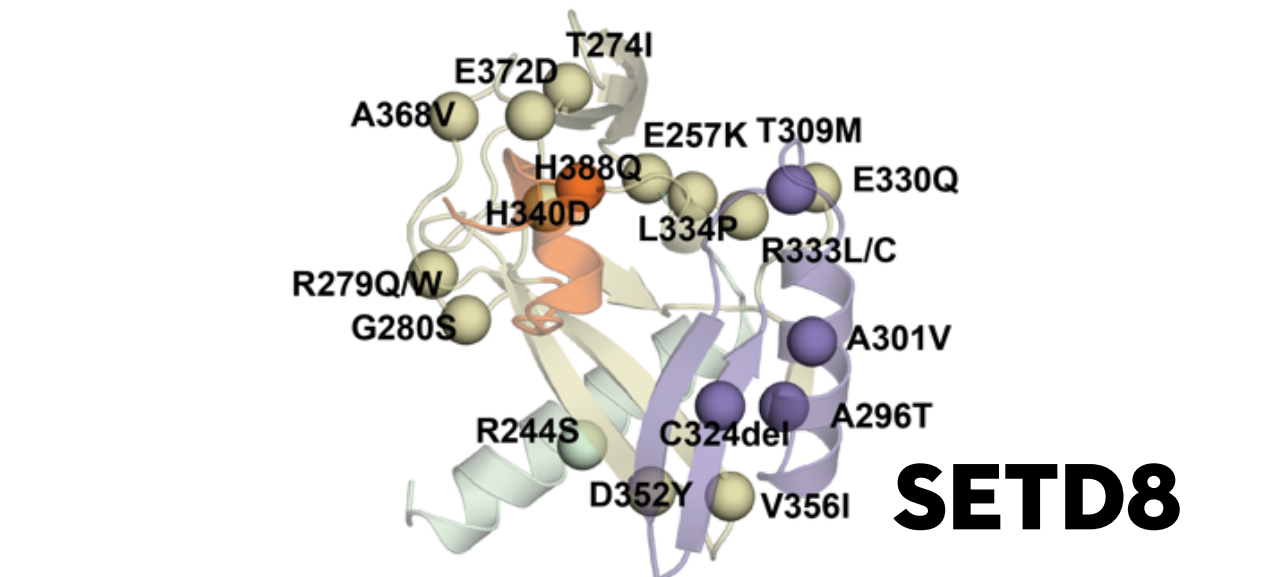
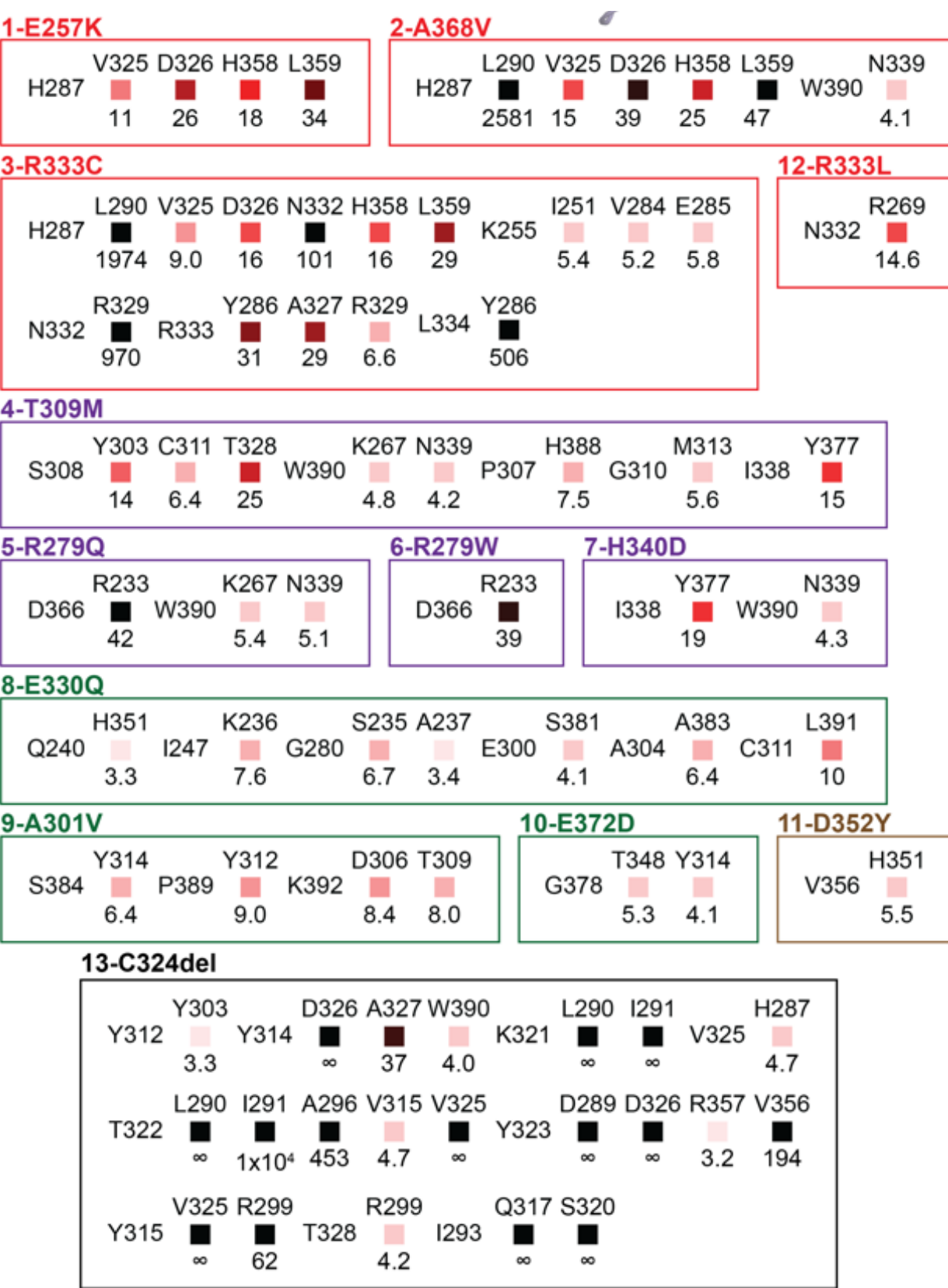


SIMULATIONS OF SETD8 MUTANTS FROM MSK-IMPACT ALLOW US TO ANNOTATE THEIR FUNCTIONAL EFFECTS

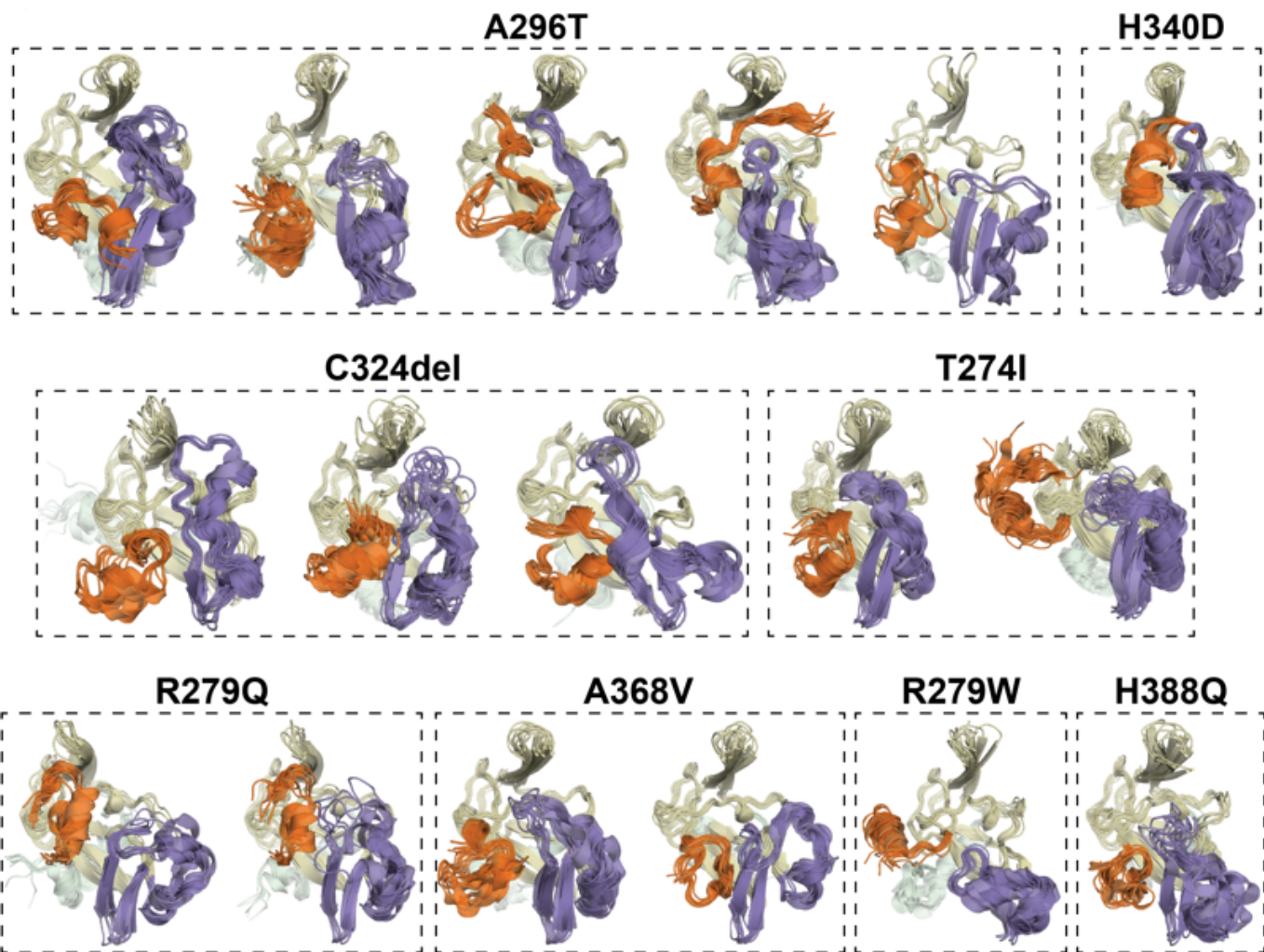
SETD8 clinical cancer mutations from MSK-IMPACT



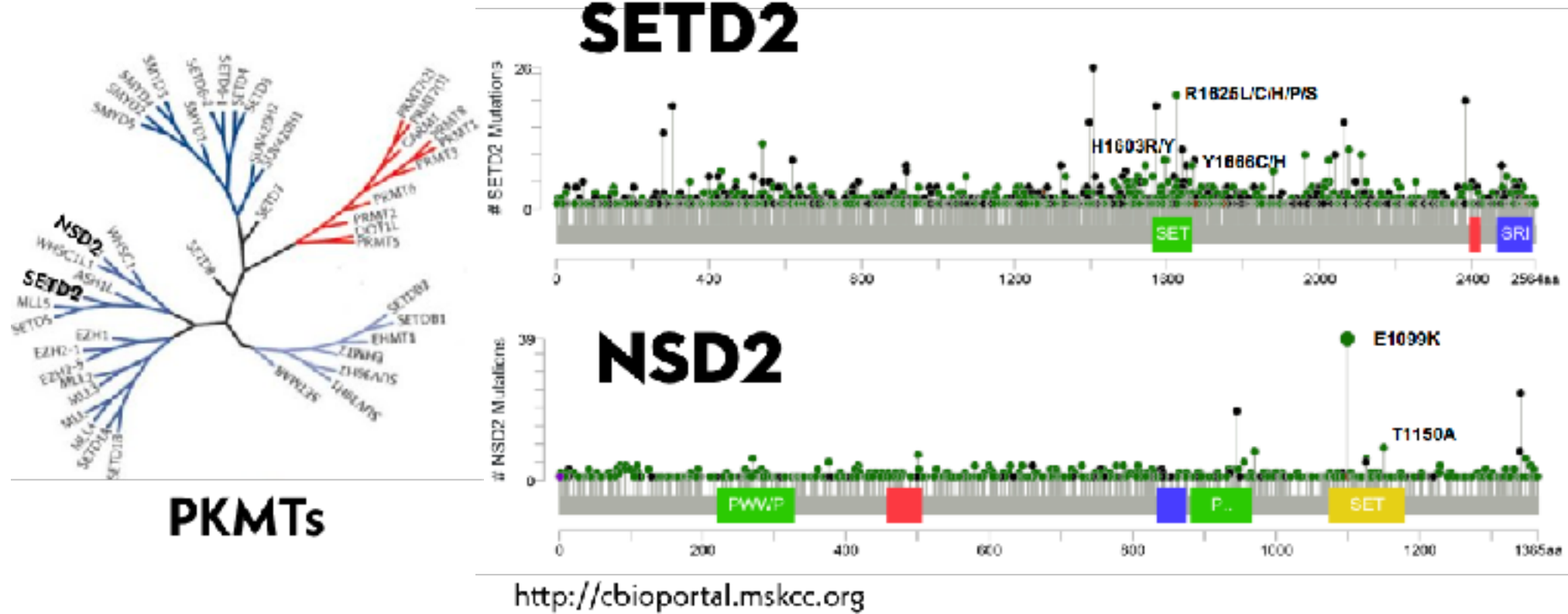
gain/loss of secondary and tertiary structure and population changes



population of neo-conformations

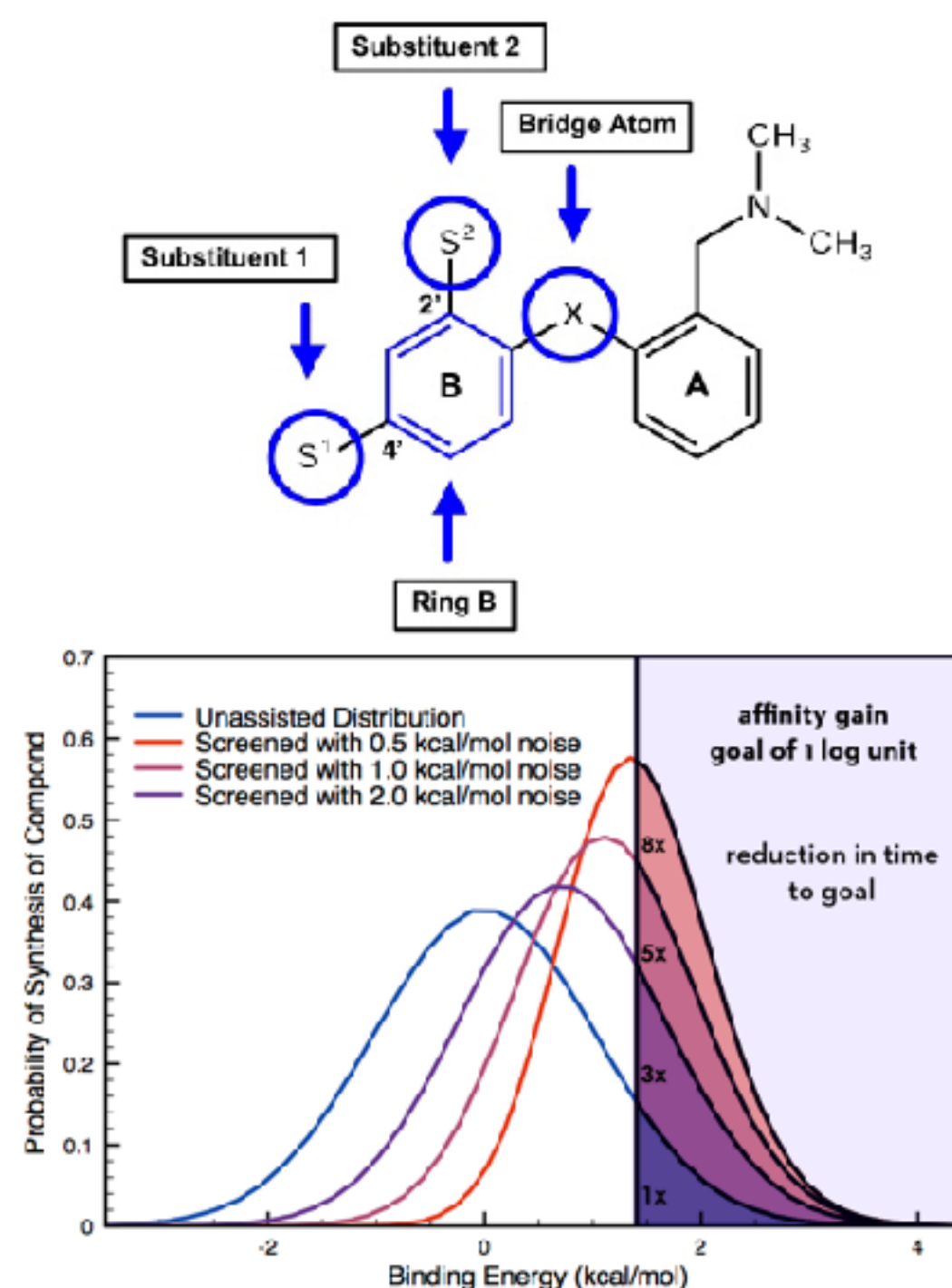


Next steps: What are cancer-associated mutations of SETD2 and NSD2 doing to perturb activity, regulation, and specificity?



HOW CAN BIOPHYSICAL MODELING HELP US IMPROVE SUCCESS RATES IN DRUGGING CANCER TARGETS?

Our lab develops next-generation modeling approaches that fulfill unmet needs in rational drug discovery.



Chodera lab papers:

Cell Chem Biol, in press

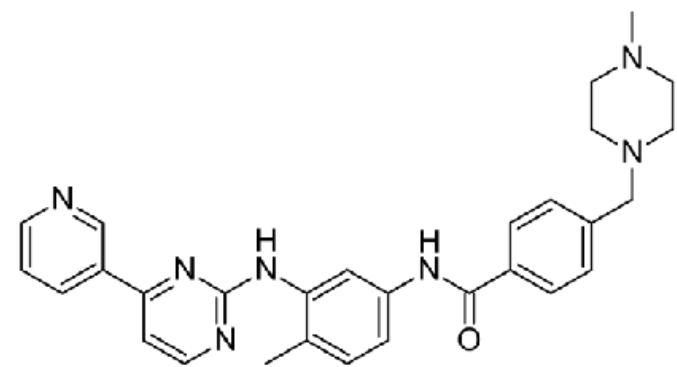
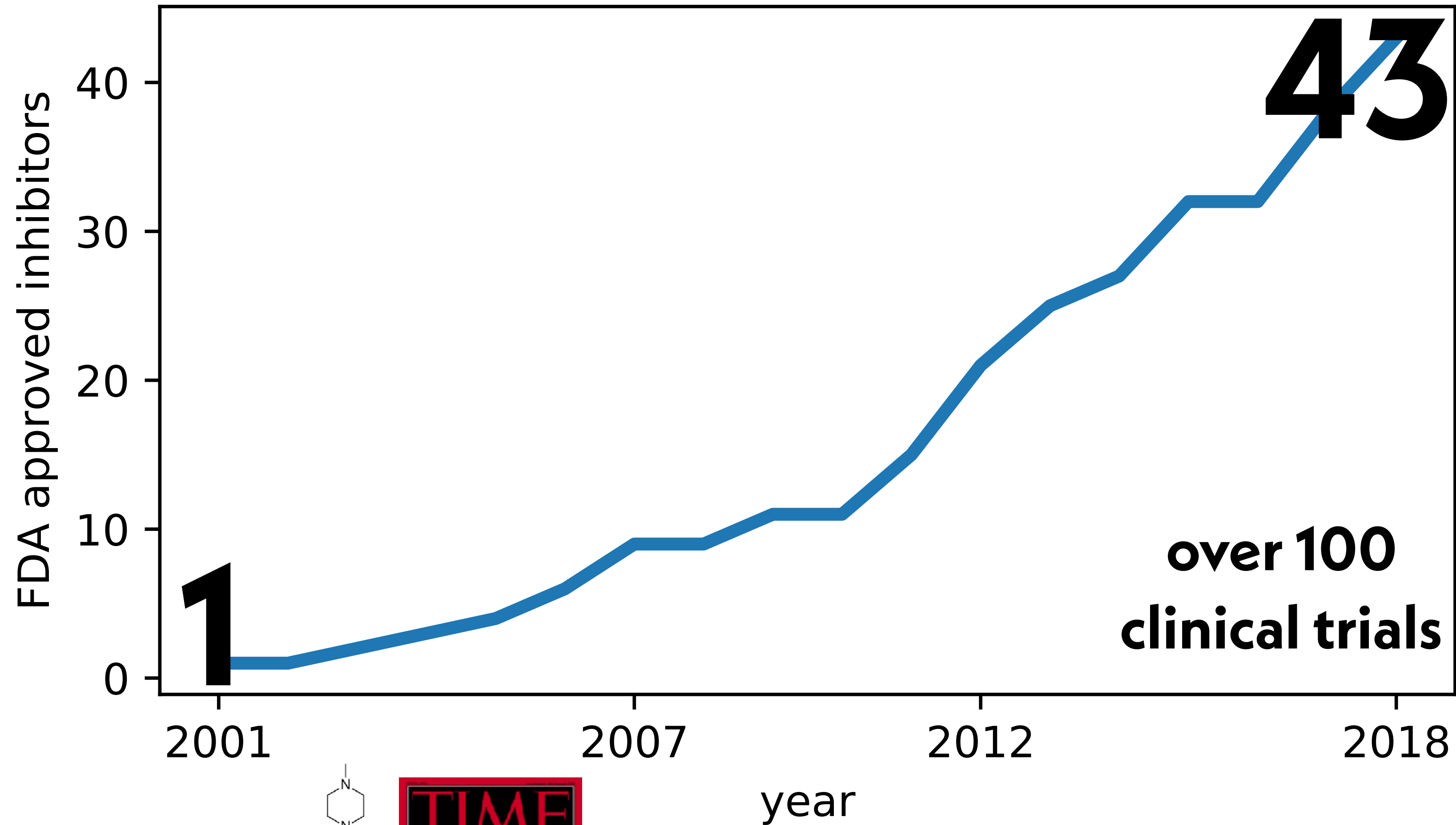
Nature Materials 17:361, 2018

J Chem Eng Data 62:1559, 2017

J Comput Aid Mol Des 30:945, 2016; 29:1073, 2015

J Phys Chem B 122:5466, 2018; 22:5579, 2018; 119:12912, 2015

HUMAN KINASES ARE A MAJOR TARGET IN PRECISION CANCER TREATMENT

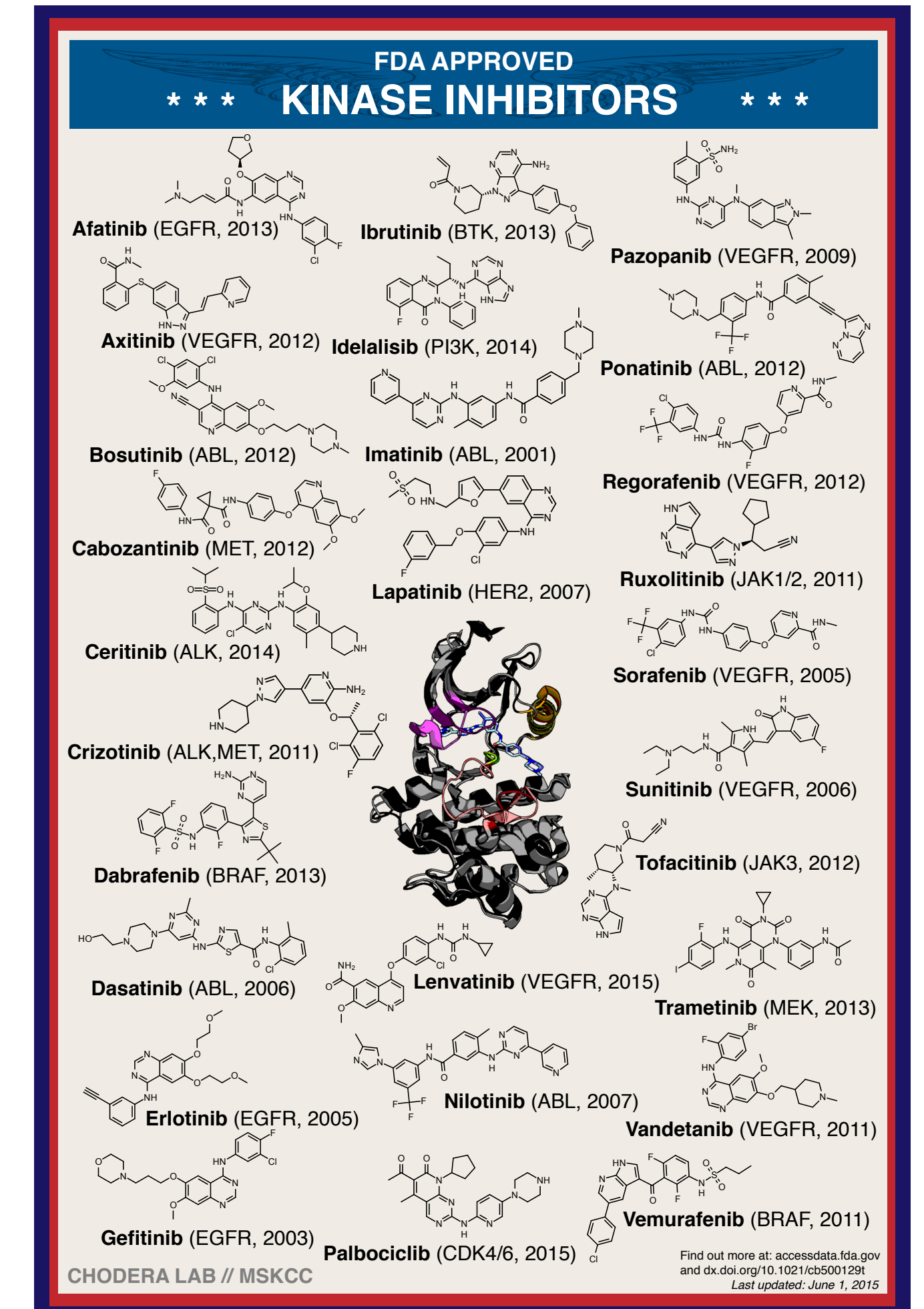


imatinib

approved by USFDA in 2001



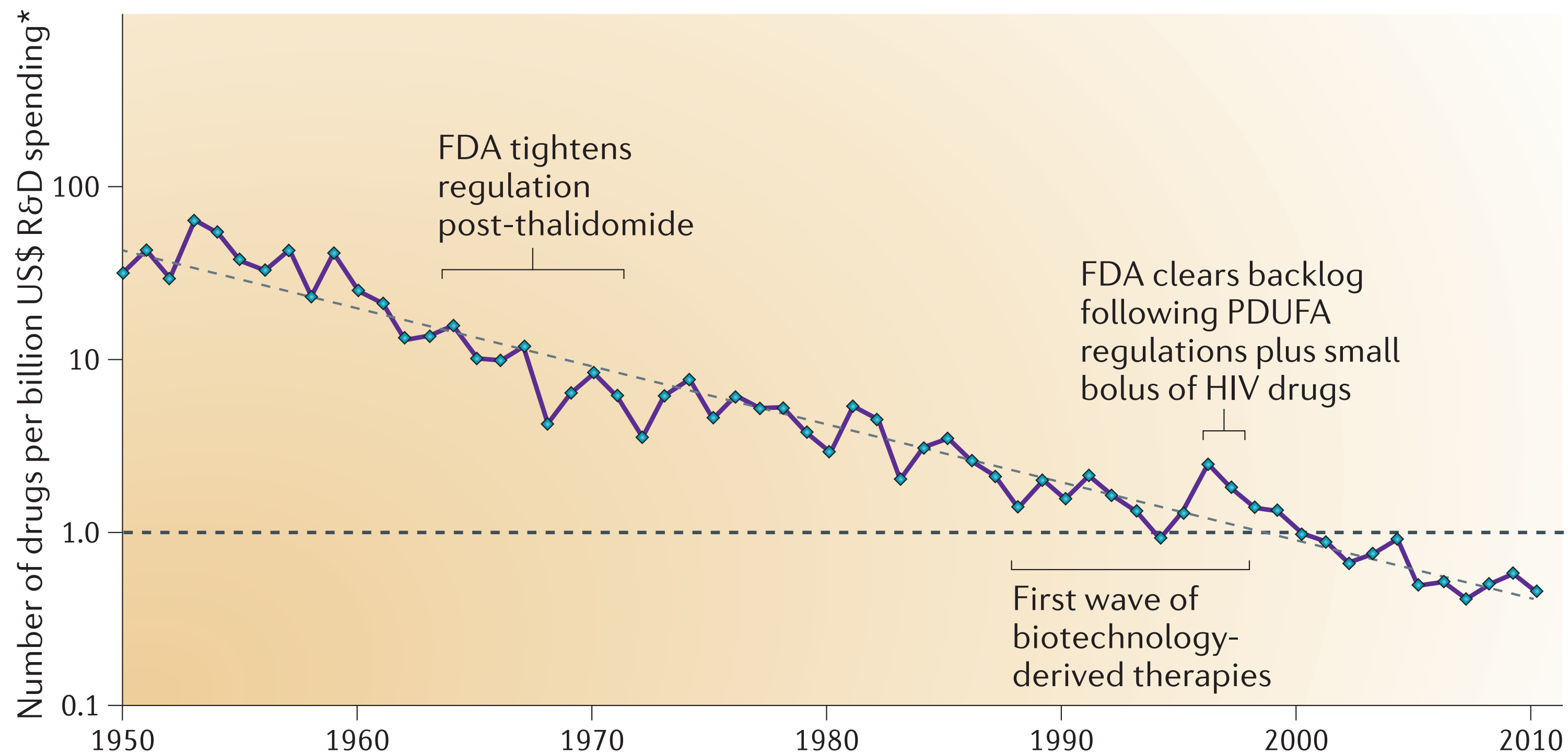
Only a small number of kinases implicated in cancer have been targeted so far



DRUG DISCOVERY USUALLY ENDS IN FAILURE

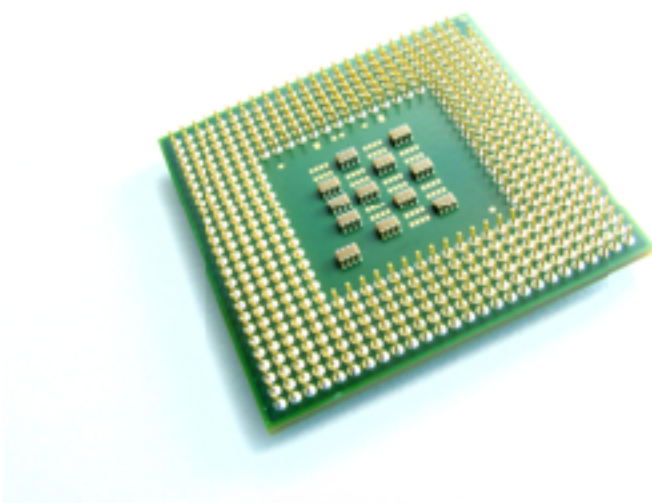
Drugs are getting more expensive to develop due to low success rates

a Overall trend in R&D efficiency (inflation-adjusted)

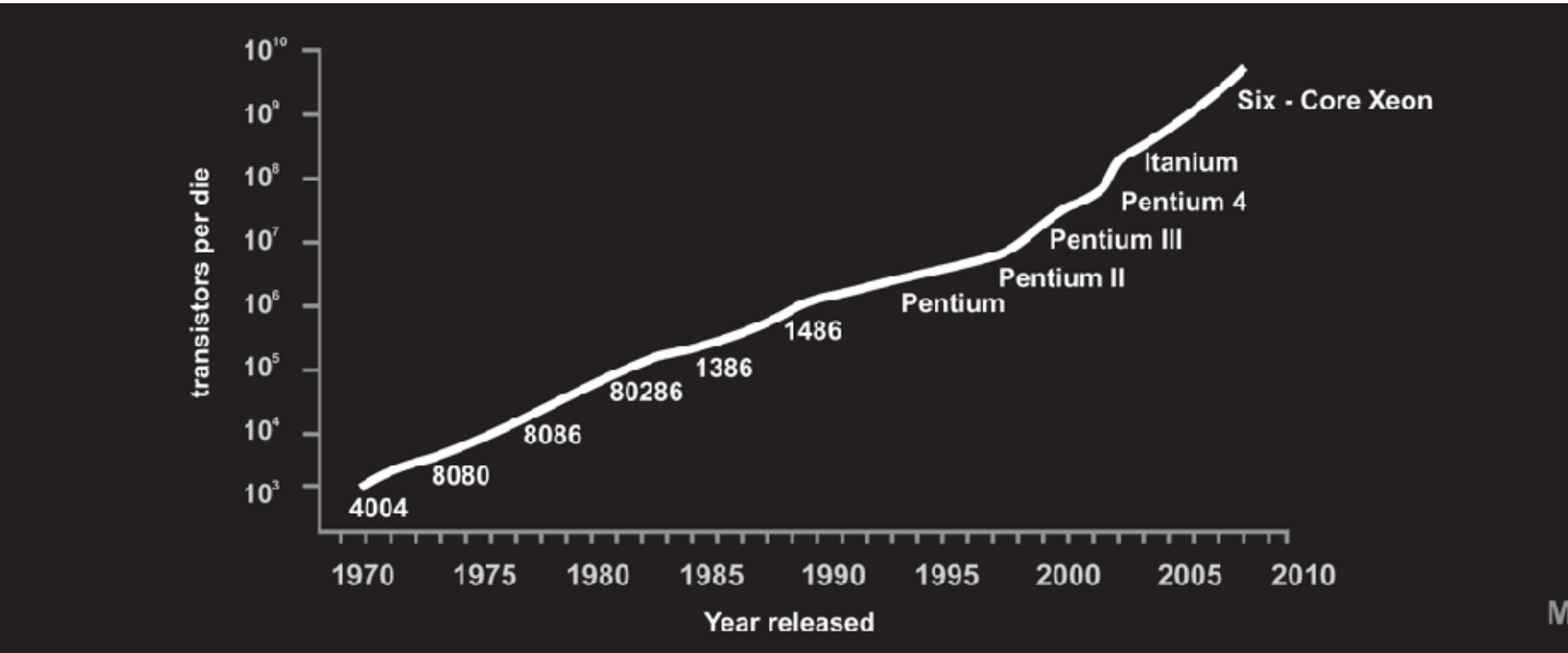
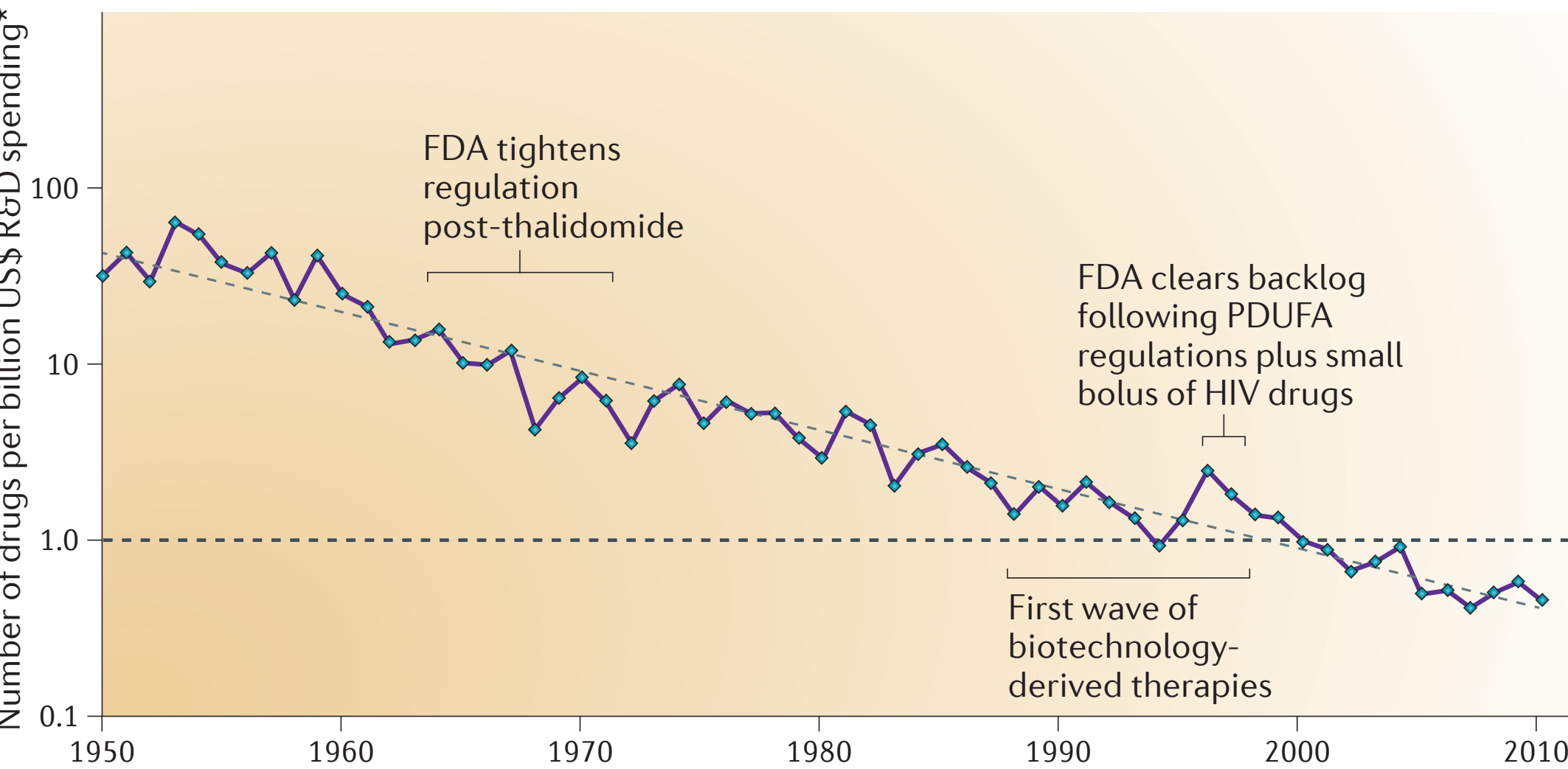


now: \$1.6B/drug

DRUG DISCOVERY USUALLY ENDS IN FAILURE



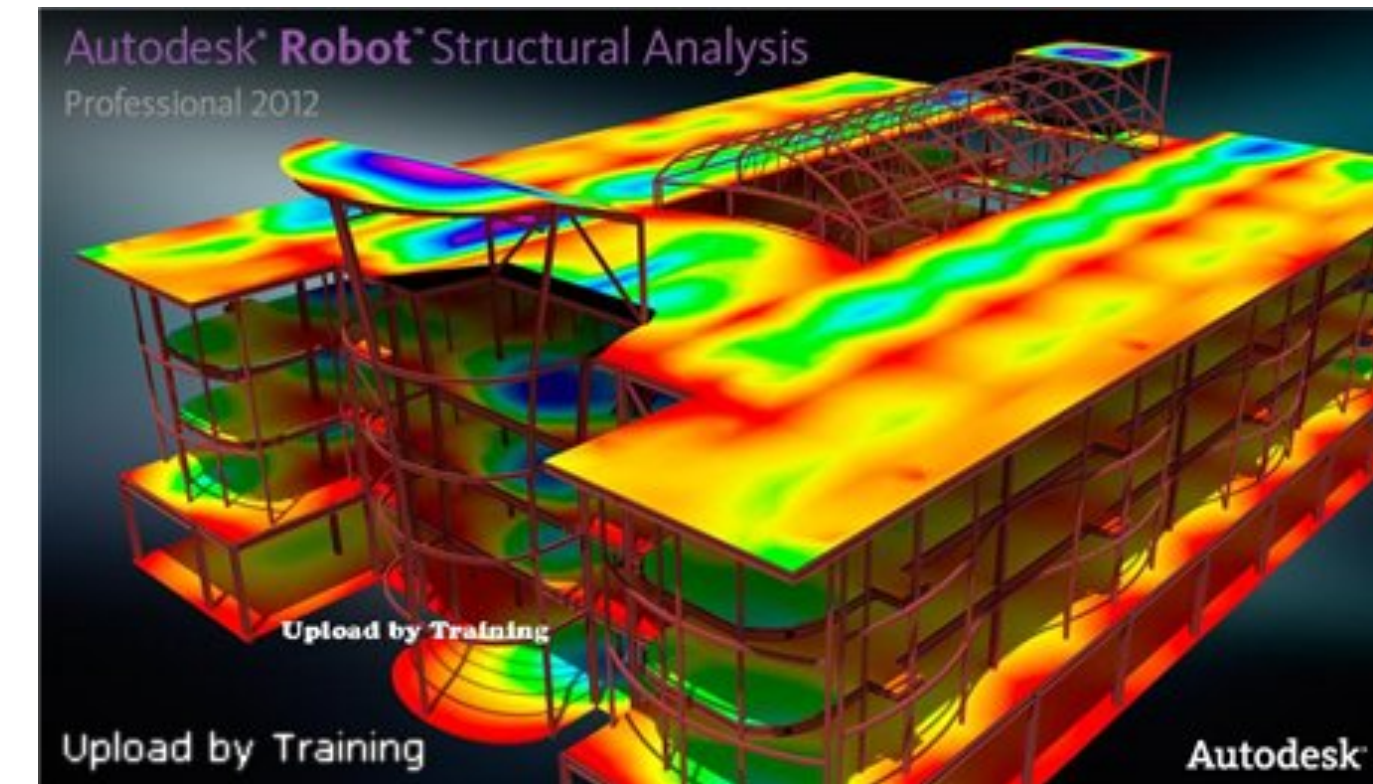
a Overall trend in R&D efficiency (inflation-adjusted)



EROOM'S LAW

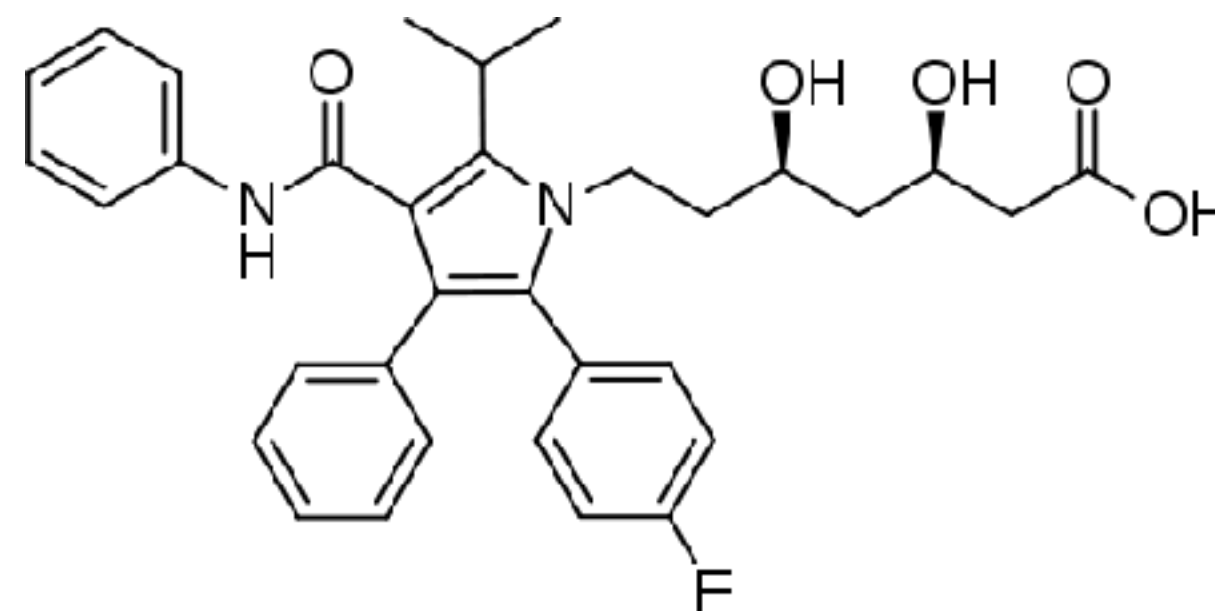
MOORE'S LAW

WE REGULARLY **DESIGN** PLANES, BRIDGES, AND BUILDINGS ON COMPUTERS



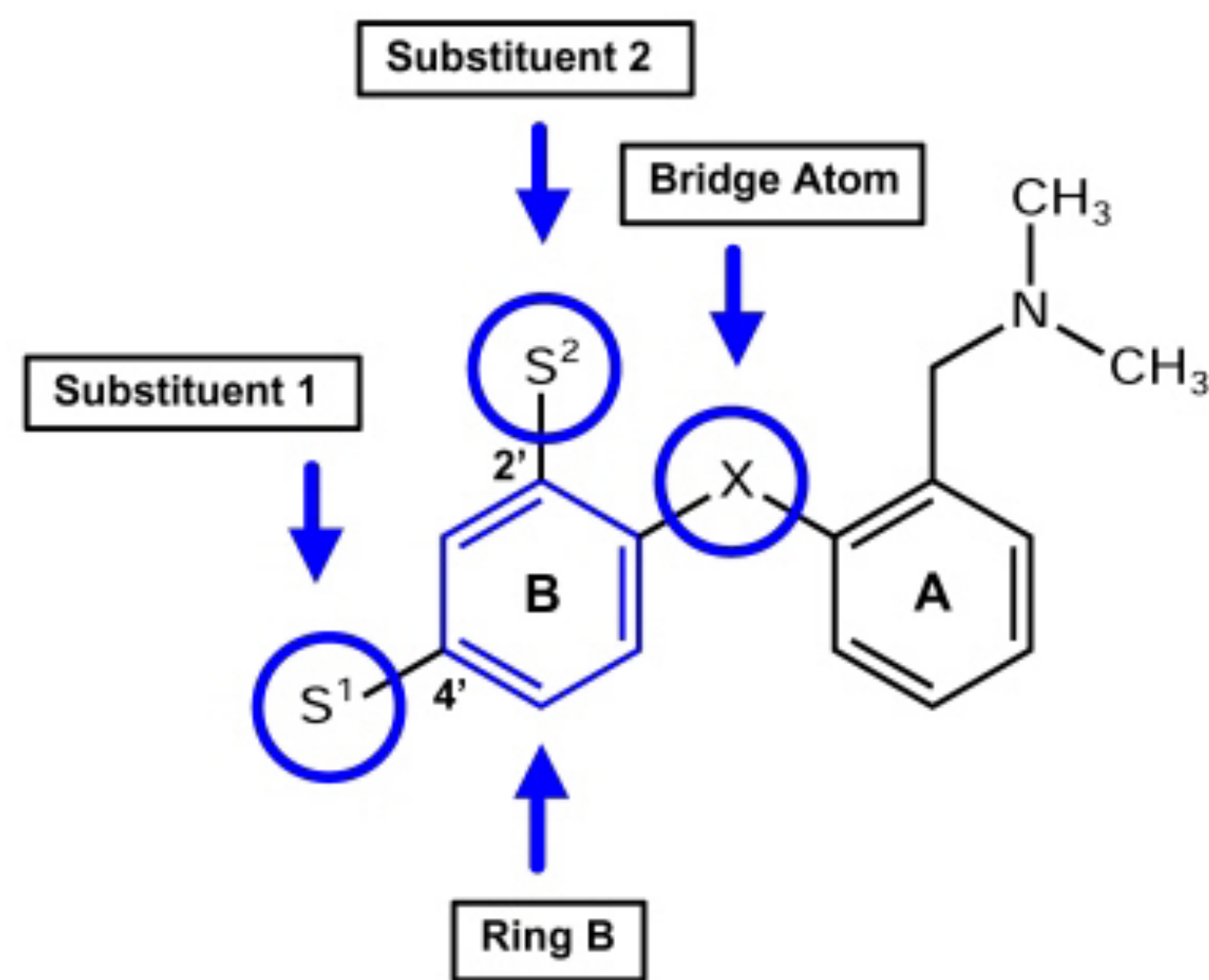
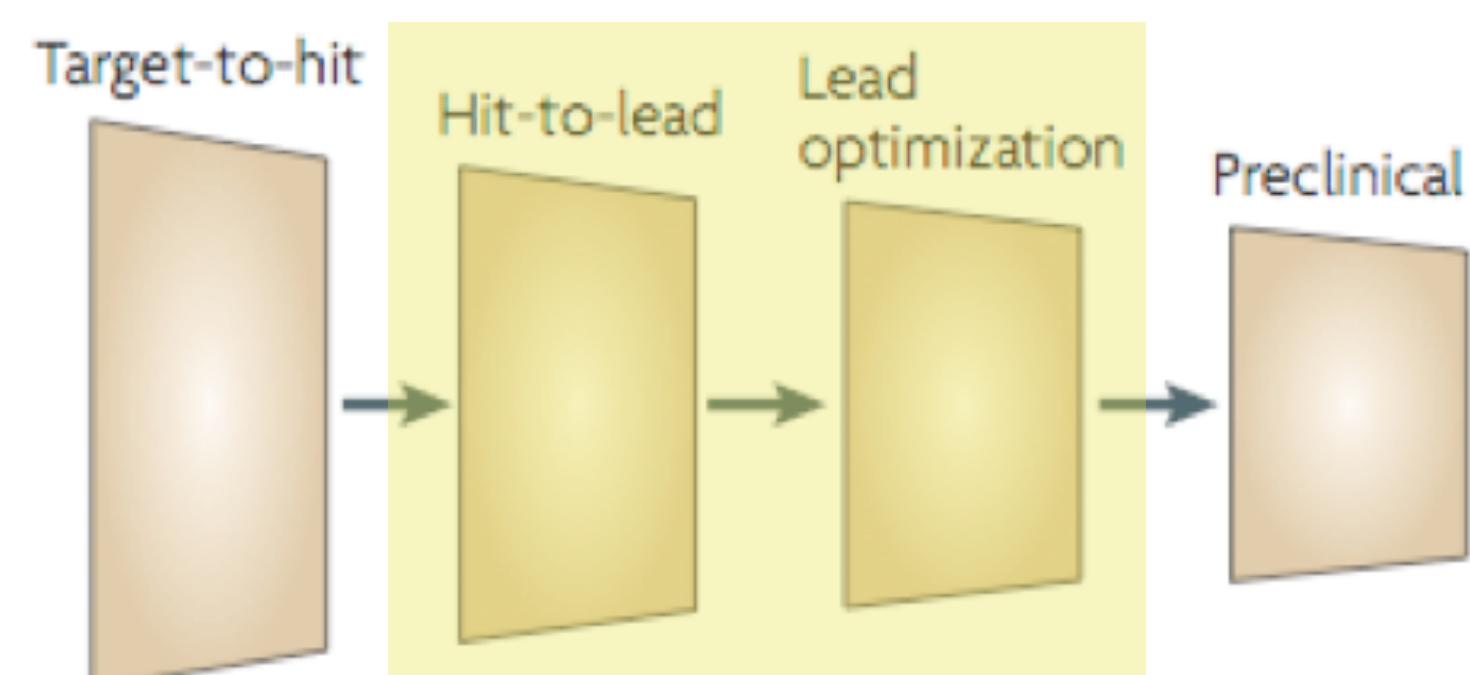
$10^3 - 10^6$ parts

WHY NOT SMALL MOLECULE DRUGS?

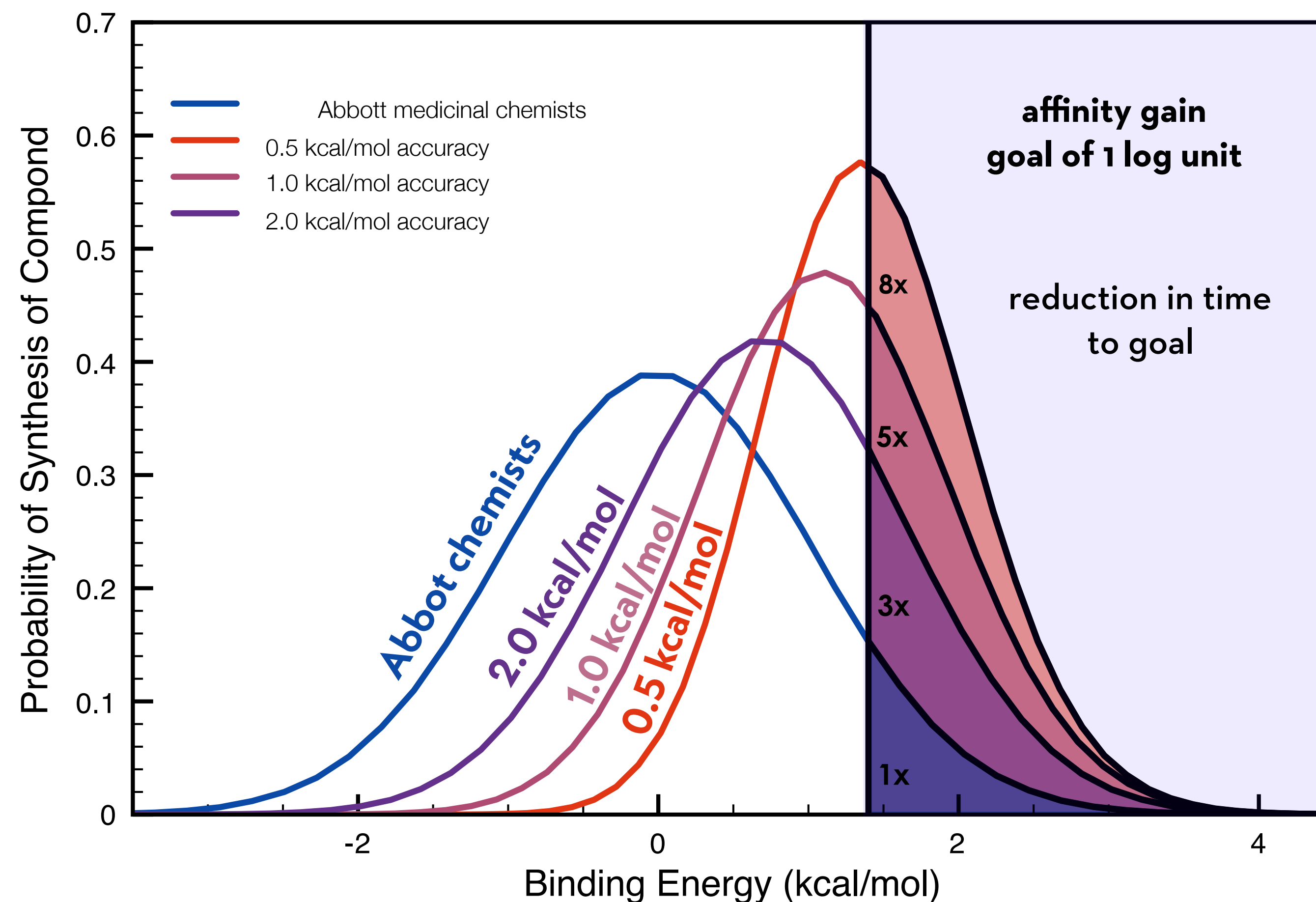


$< 10^2$ atoms

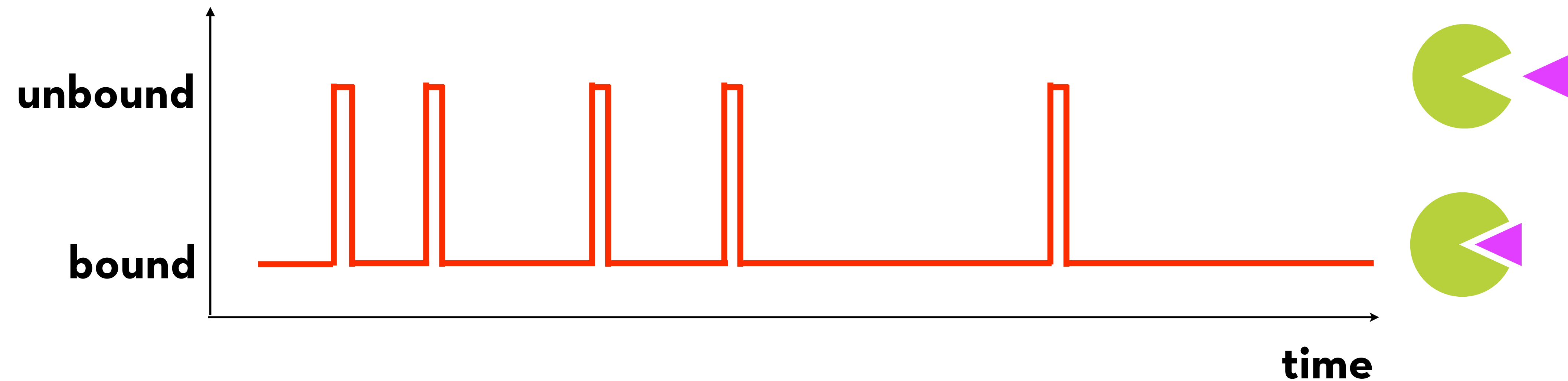
A PREDICTIVE MODEL OF LIGAND AFFINITY WOULD HAVE A HUGE IMPACT ON DRUG DISCOVERY



binding free energy gain in lead optimization synthesis



HOW CAN WE **COMPUTE** BINDING AFFINITIES OF LIGANDS WE HAVEN'T SYNTHESIZED YET?



dissociation
constant

$$K_d \propto \frac{\tau_{\text{unbound}}}{\tau_{\text{bound}}}$$

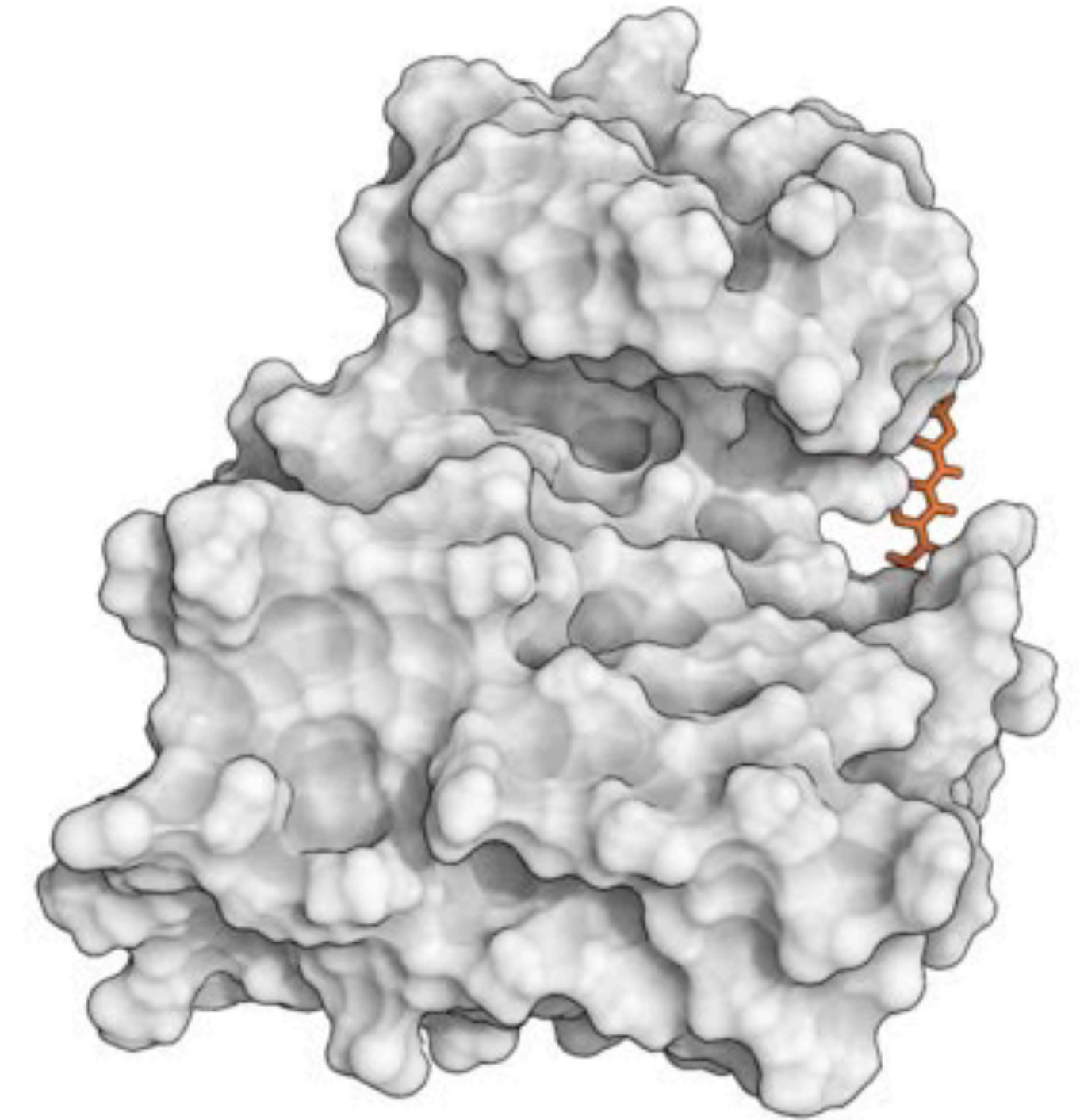
HOW CAN WE **COMPUTE** BINDING AFFINITIES OF LIGANDS WE HAVEN'T SYNTHESIZED YET?

ANTON

\$50M special-purpose supercomputer from D.E. Shaw Research



David E. Shaw

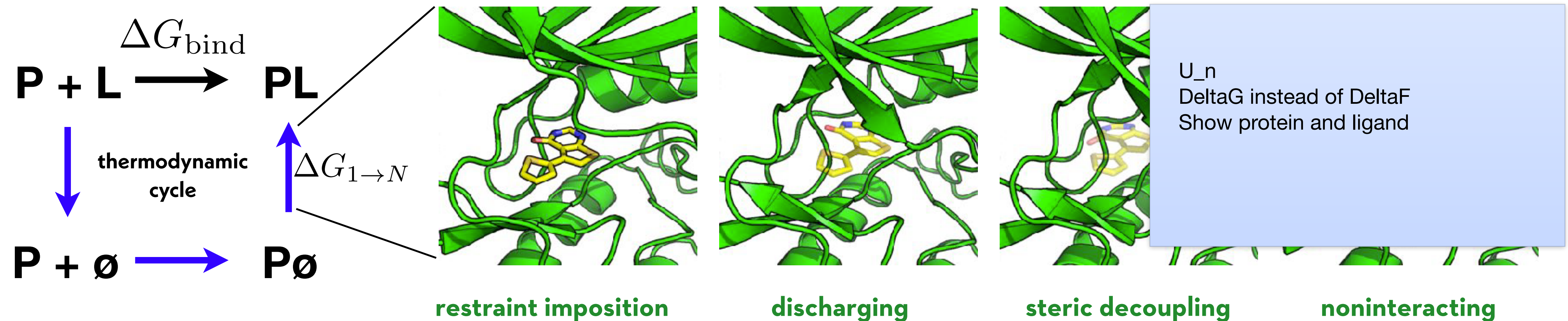


Src:dasatanib
(4 us simulation)

For typical drug off-rates (10^{-4} s^{-1}),
reliable calculation of binding affinities would require hour trajectories,
requiring $\sim 10^6$ years to simulate.

ALCHEMICAL FREE ENERGY CALCULATIONS PROVIDE A RIGOROUS WAY TO EFFICIENTLY COMPUTE BINDING AFFINITIES

multiple simulations of **alchemical intermediates**

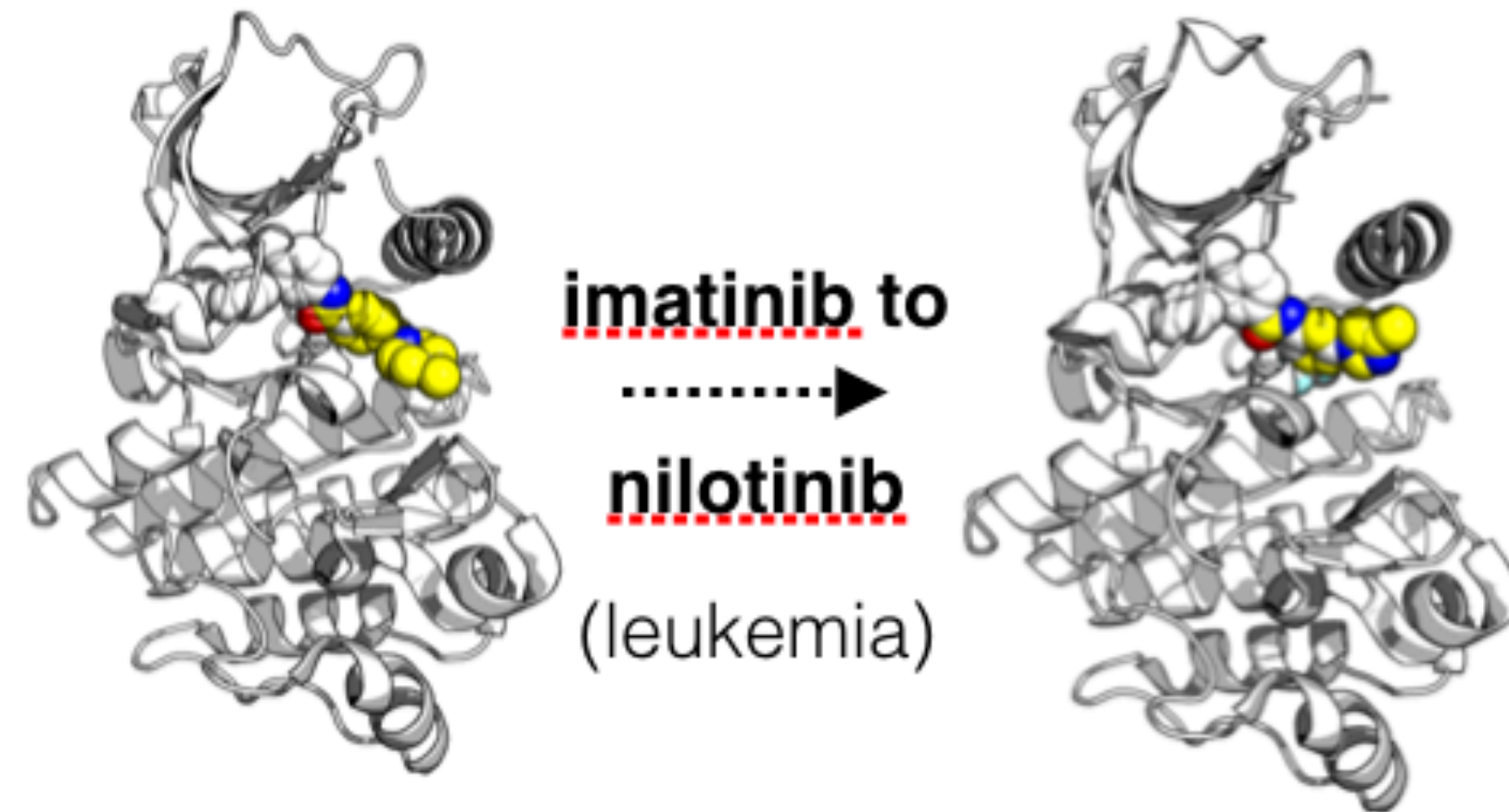


By breaking the problem into statistically easily computable pieces,
calculation can be completed in just **hours**

$$\Delta G_{1 \rightarrow N} = -\beta^{-1} \ln \frac{Z_N}{Z_1} = -\beta^{-1} \ln \frac{Z_2}{Z_1} \cdot \frac{Z_3}{Z_2} \cdots \frac{Z_N}{Z_{N-1}} \quad Z_n = \int dx e^{-\beta U_n(x)} \quad \text{partition function}$$

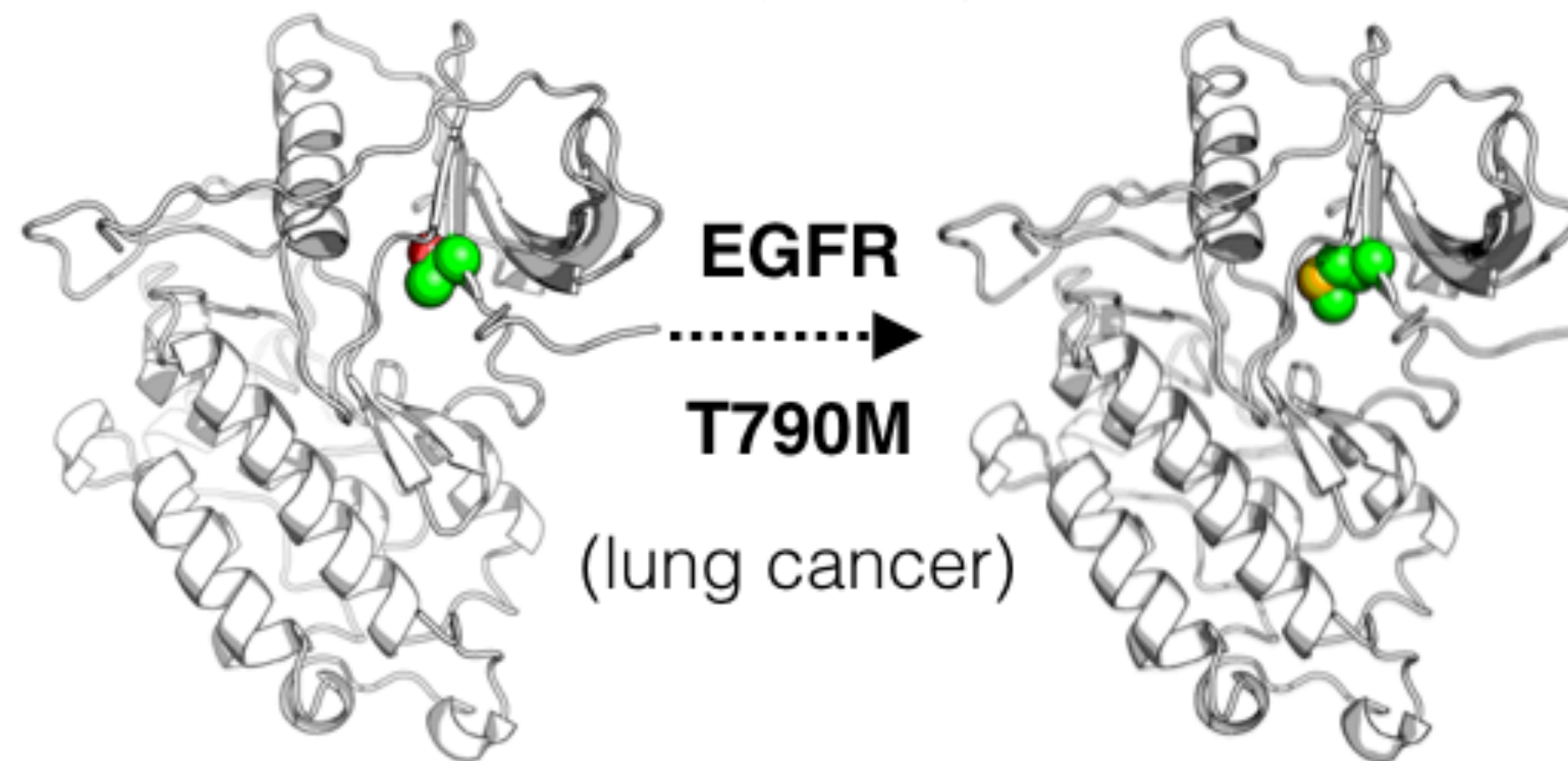
CAN WE USE FREE ENERGY CALCULATIONS TO ADDRESS MAJOR QUESTIONS IN CANCER THERAPY?

CHANGES OF A FEW ATOMS



HOW CAN WE DESIGN
SPECIFICALLY TARGETED
CANCER DRUGS?

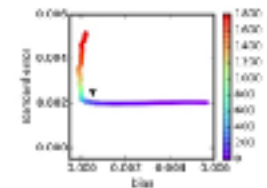
inhibitor modification
for drug design



HOW CAN WE PREDICT
DRUG RESISTANCE
AND SUSCEPTIBILITY?

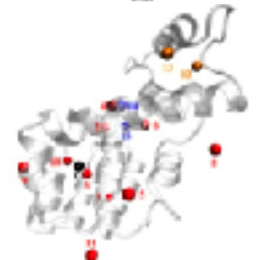
tumor-specific mutation
for therapeutic decisions

AN OPEN-SOURCE PLATFORM FOR INTEGRATING METHODOLOGY ADVANCES INTO GPU-ACCELERATED FREE ENERGY CALCULATIONS



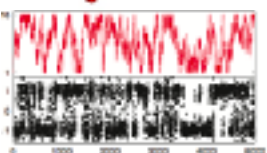
Chodera

J Chem Theor Comput 12:1799, 2016



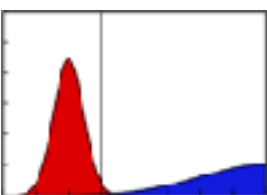
Wang, **Chodera**, Yang, Shirts

J Comput Aid Mol Des 27:989, 2013



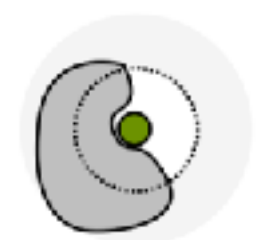
Chodera, Shirts

J Chem Phys 135:194110, 2011



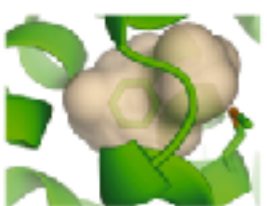
Shirts, **Chodera**

J Chem Phys 129:124105, 2008



Shirts, Mobley, **Chodera**, Pande

J Phys Chem B 111:13052, 2007



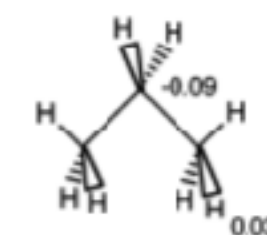
Mobley, Graves, **Chodera**, McReynolds, Shoichet, Dill

J Mol Biol 371:1118, 2007



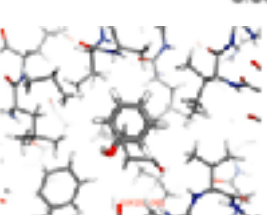
Mobley, **Chodera**, Dill

J Chem Theor Comput 3:1231, 2007



Mobley, Dumont, **Chodera**, Bayly, Cooper, Dill

J Phys Chem B 111:2242, 2007



Mobley, **Chodera**, Dill

J Chem Phys 125:084902, 2006

YANK

A GPU-accelerated Python framework for exploring algorithms for alchemical free energy calculations

Features

- Modular Python framework to facilitate development and testing of new algorithms
- GPU-accelerated via the [OpenMM toolkit](#) (see [benchmarks](#))
- [Alchemical free energy calculations](#) in both explicit (PME, reaction field) and implicit (GBSA) solvent
- Multi-state sampling schemes, including replica exchange with [Gibbs sampling](#) and [self-adjusted mixture sampling \(SAMS\)](#)
- Extensible [Markov chain Monte Carlo](#) framework for exploring enhanced sampling methods
- Built-in [automated equilibration detection](<http://dx.doi.org/10.1021/acs.jctc.5b00784>) and convergence diagnostics
- Support for reading Amber, gromacs, and CHARMM input files
- Support for absolute binding free energy calculations and transfer free energies (such as hydration, partition, or relative solvation free energies)



<http://www.getyank.org>

Powered by OpenMM

<http://openmm.org>

A **free, open-source**, extensible platform
for free energy calculations and ligand design
funded by numerous partners interested in **open science**



LEVI NADEN



SILICON
Therapeutics



ANDREA RIZZI



MSKCC BUILT ONE OF THE FIRST MAJOR ACADEMIC GPU CLUSTERS IN 2013



GTX-Titan
4.5 TFLOP/s (single prec)
\$1000

=

180 TFLOP/s
\$40K



GTX-680
3.0 TFLOP/s (single prec)
\$500

=

240 TFLOP/s
\$40K

420 TFLOP/s for \$80K



HAL @ MSKCC NJDC (2013)

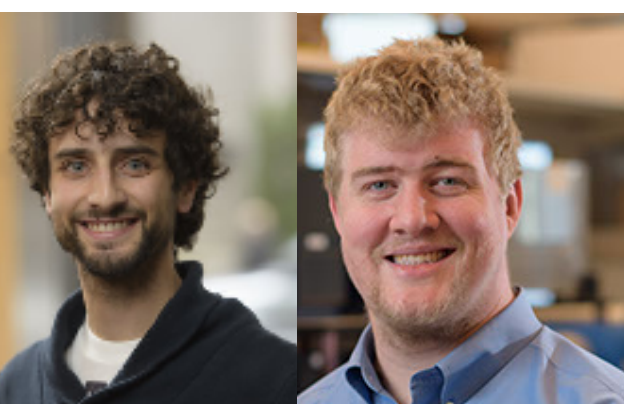
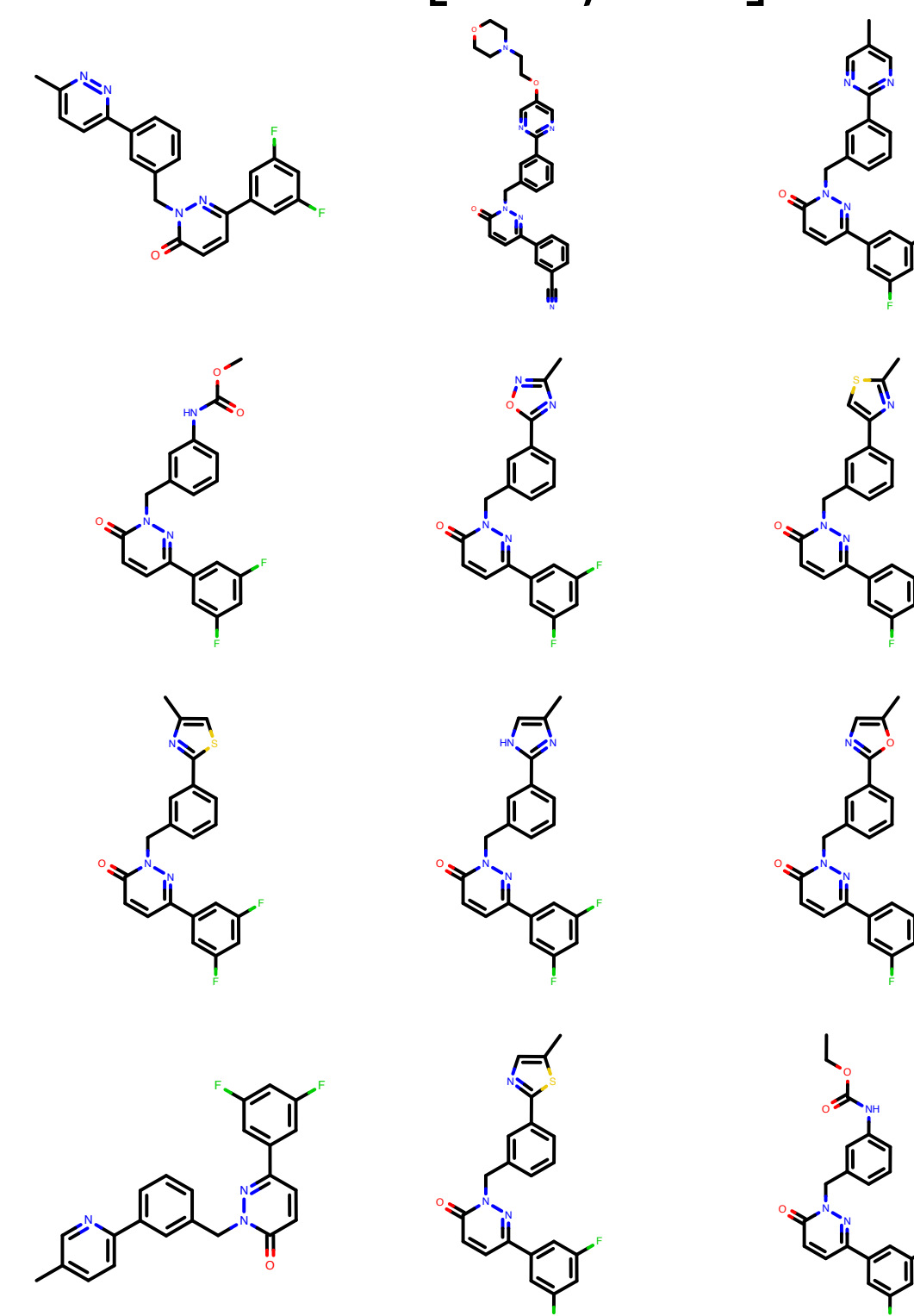
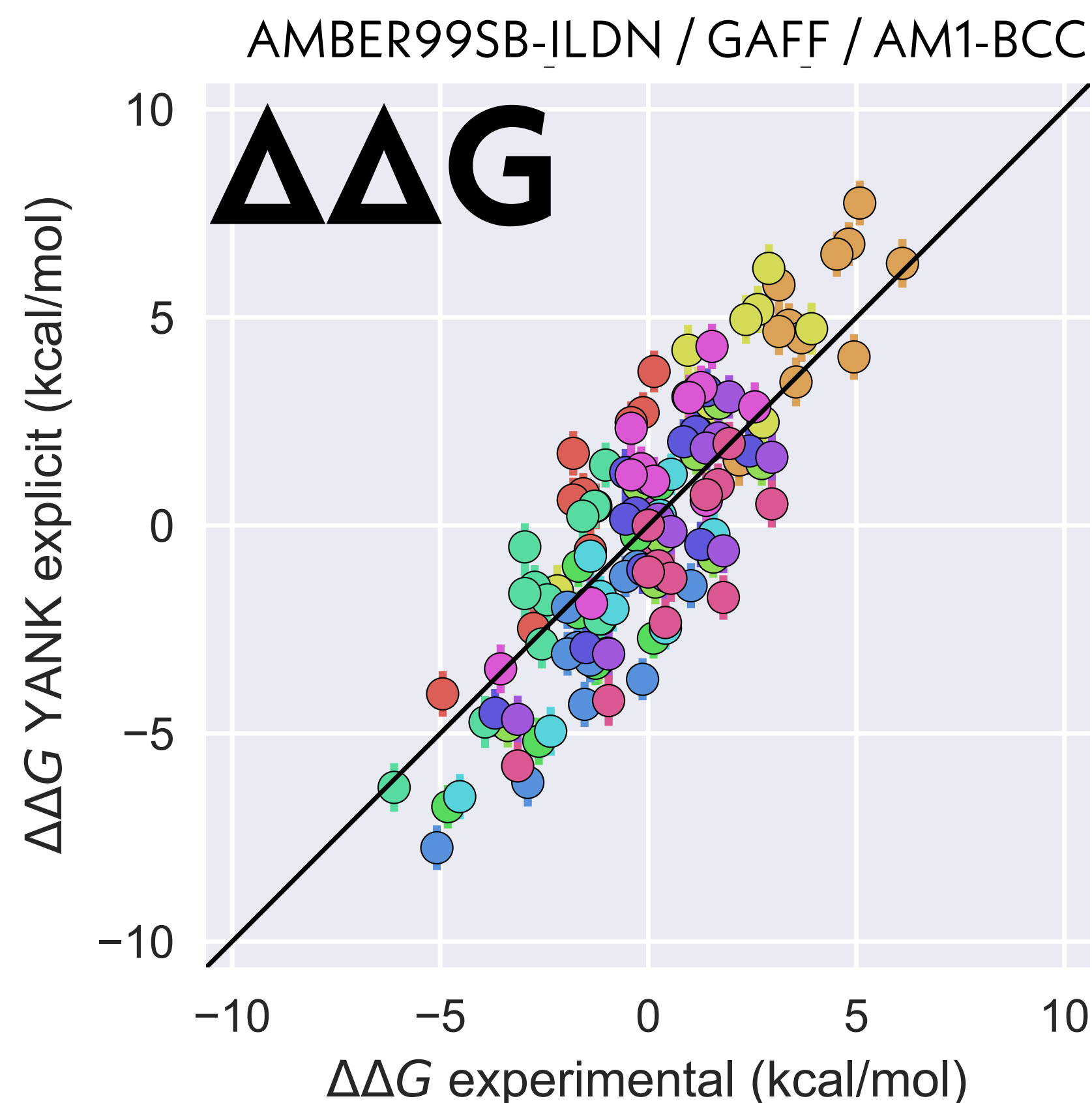
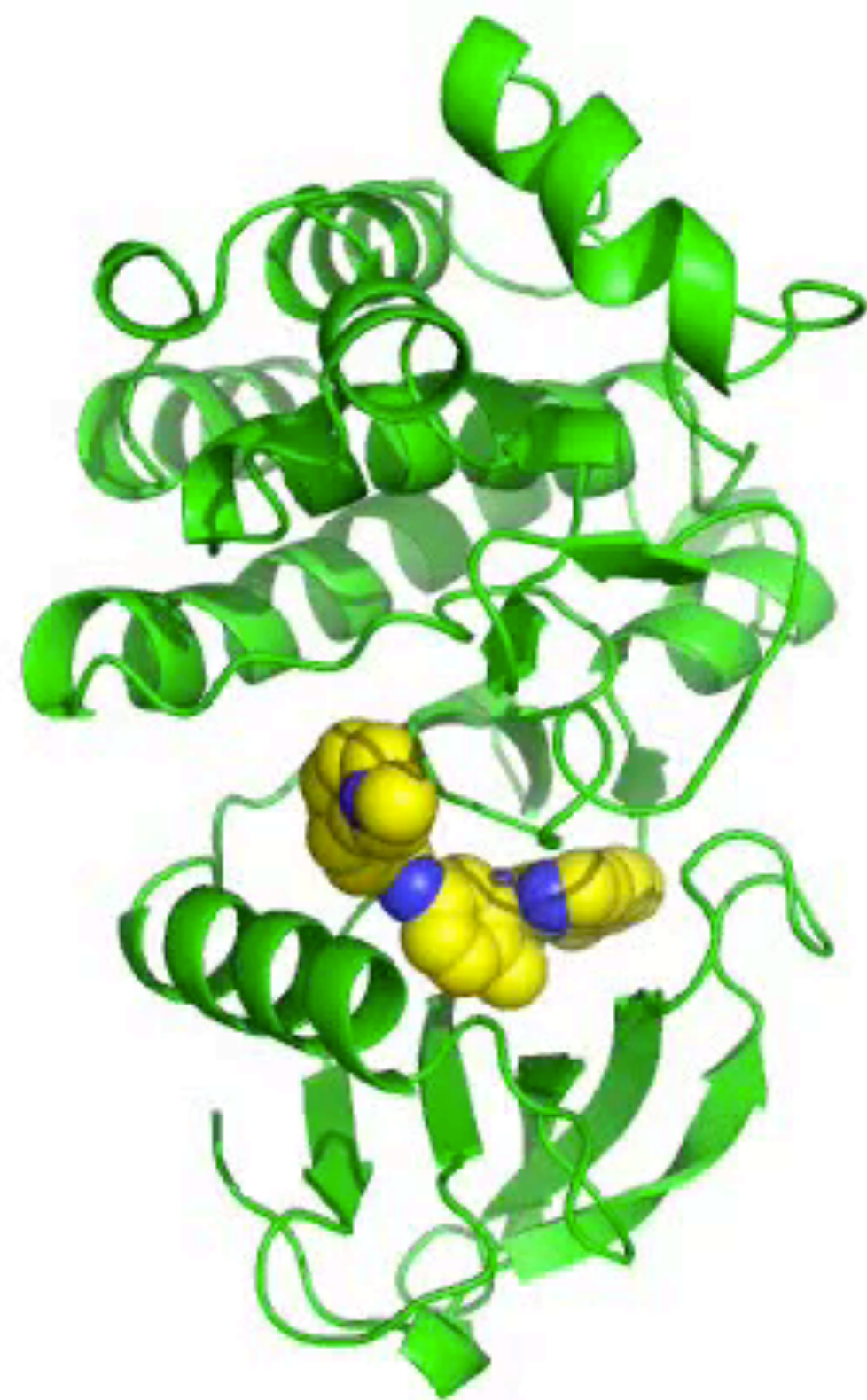


Next-generation GPU cluster **lilac** managed by HPC Director Juan Perin supports **274 users** from the MSKCC community powering machine learning, computational pathology, cryo-EM structural biology, and drug discovery

ALCHEMICAL FREE ENERGY CALCULATIONS CAN HELP PRIORITIZE LIGAND SYNTHESIS

c-Met inhibitors from Bioorg. & Med Chem Lett. 25:1597, 2015
<https://github.com/choderalab/yank-benchmark>

RMSE ~ 1.7 kcal/mol
95% CI: [1.50, 1.83]



LEVI NADEN
ANDREA RIZZI

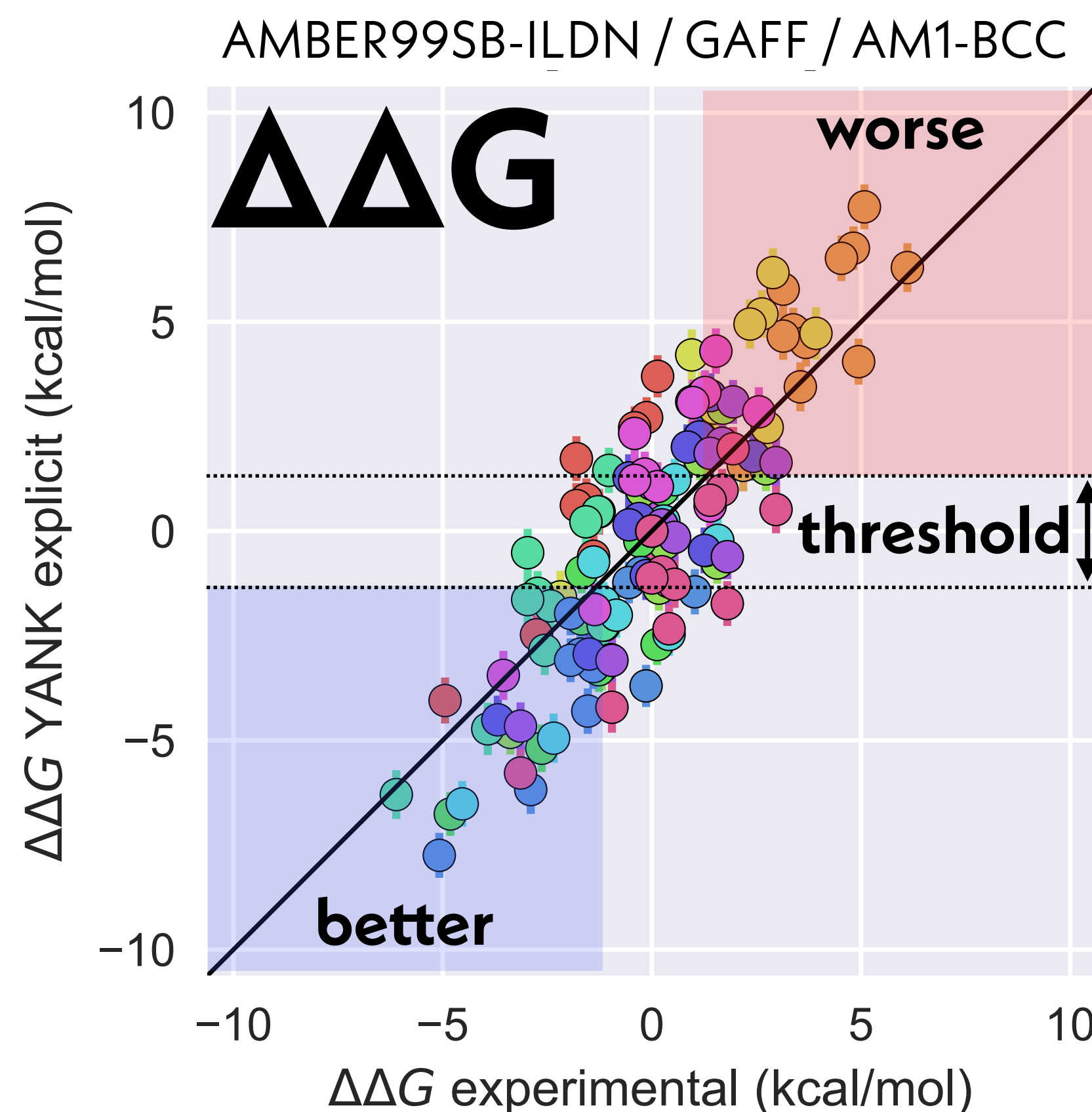
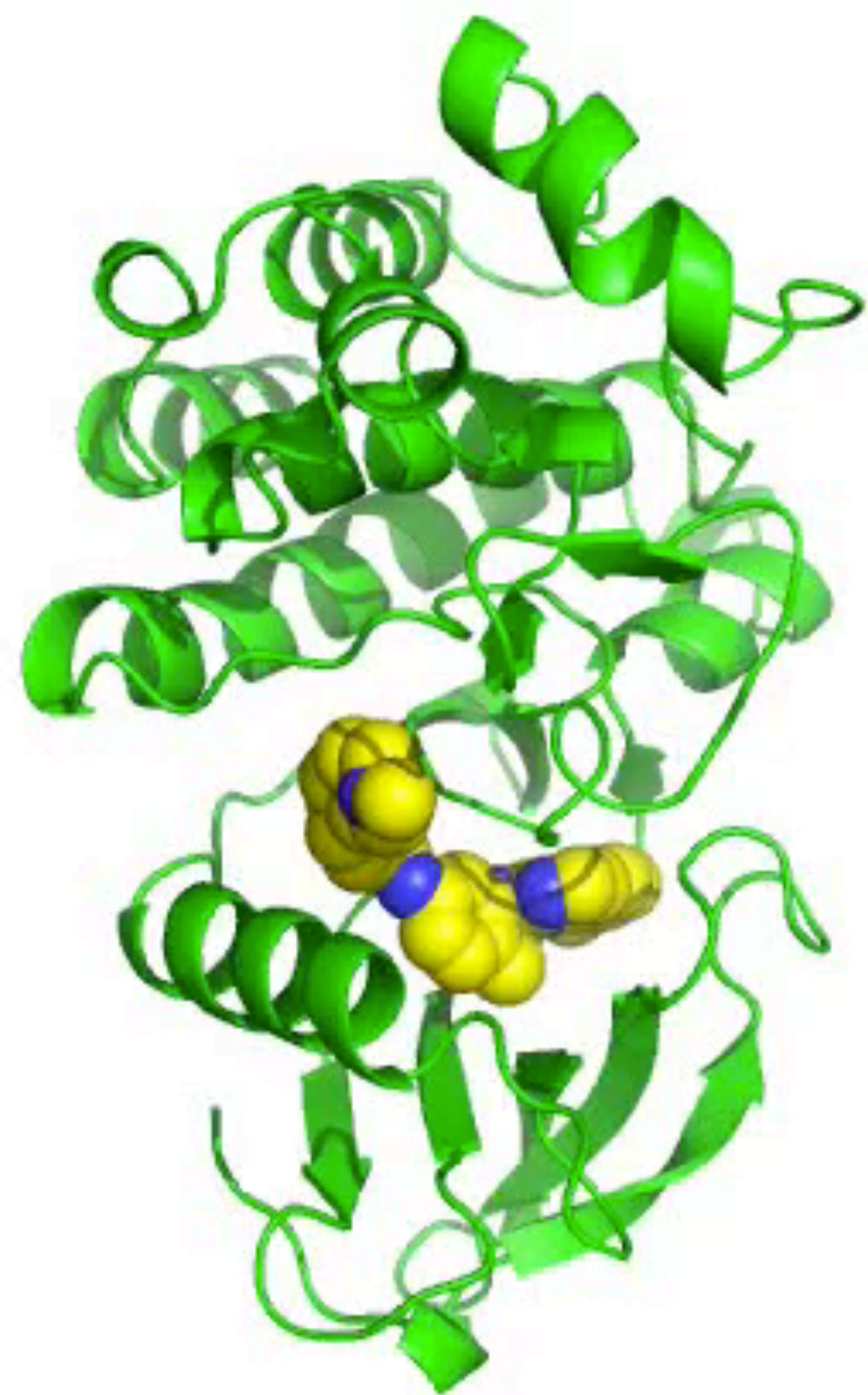


PAUL CZODROWSKI
DANIEL KUHN

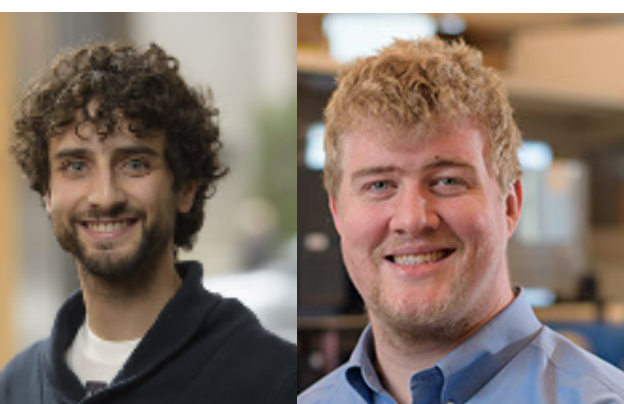


ALCHEMICAL FREE ENERGY CALCULATIONS CAN HELP PRIORITIZE LIGAND SYNTHESIS

c-Met inhibitors from Bioorg. & Med Chem Lett. 25:1597, 2015
<https://github.com/choderalab/yank-benchmark>



How often can this help us make
the **right decision** about which
molecules to synthesize?



LEVI NADEN
ANDREA RIZZI

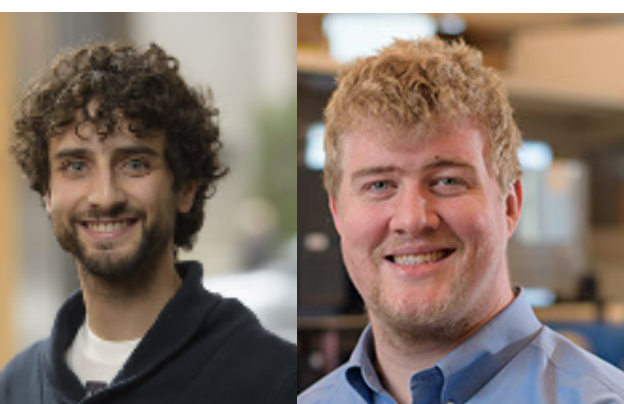
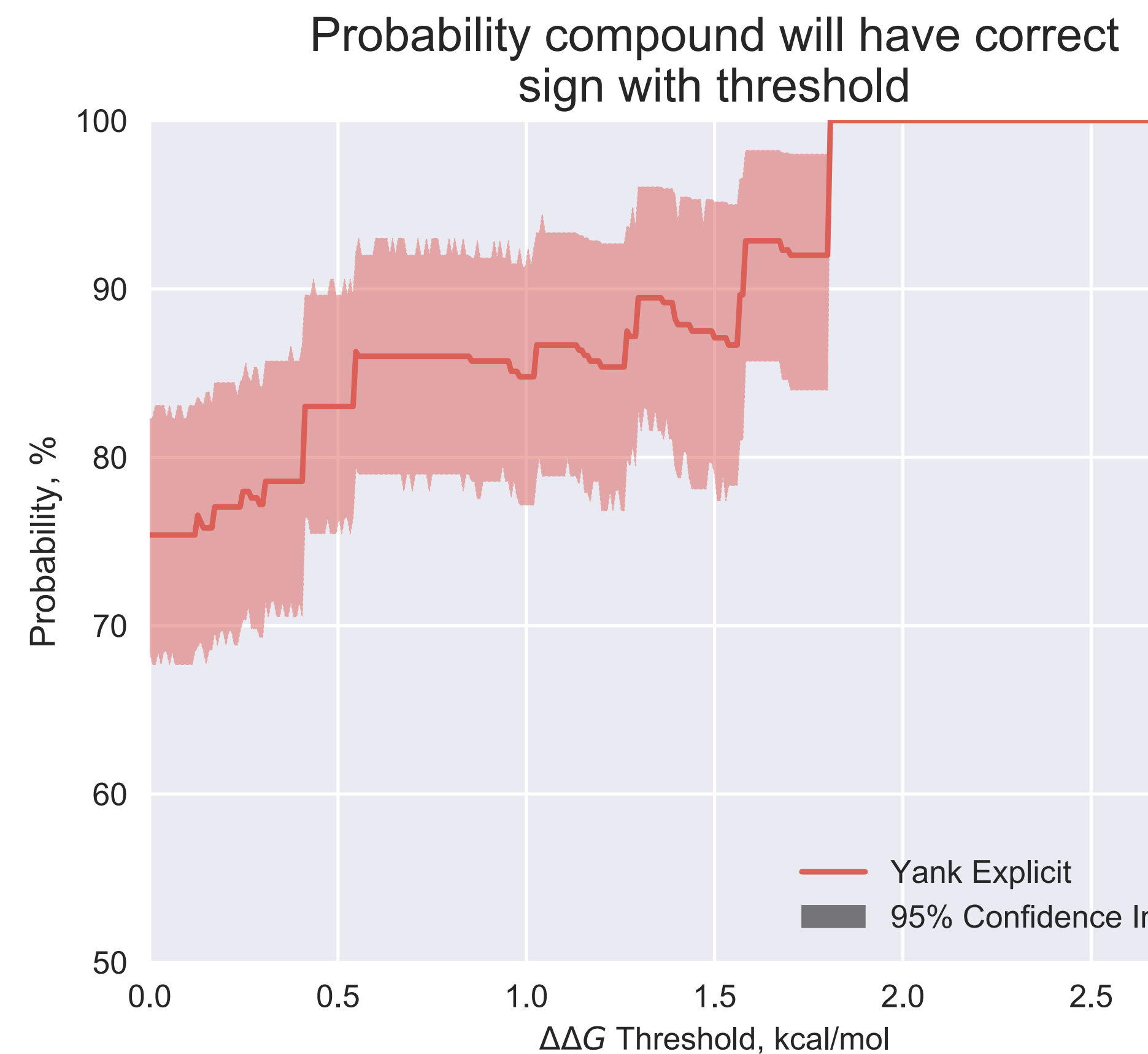
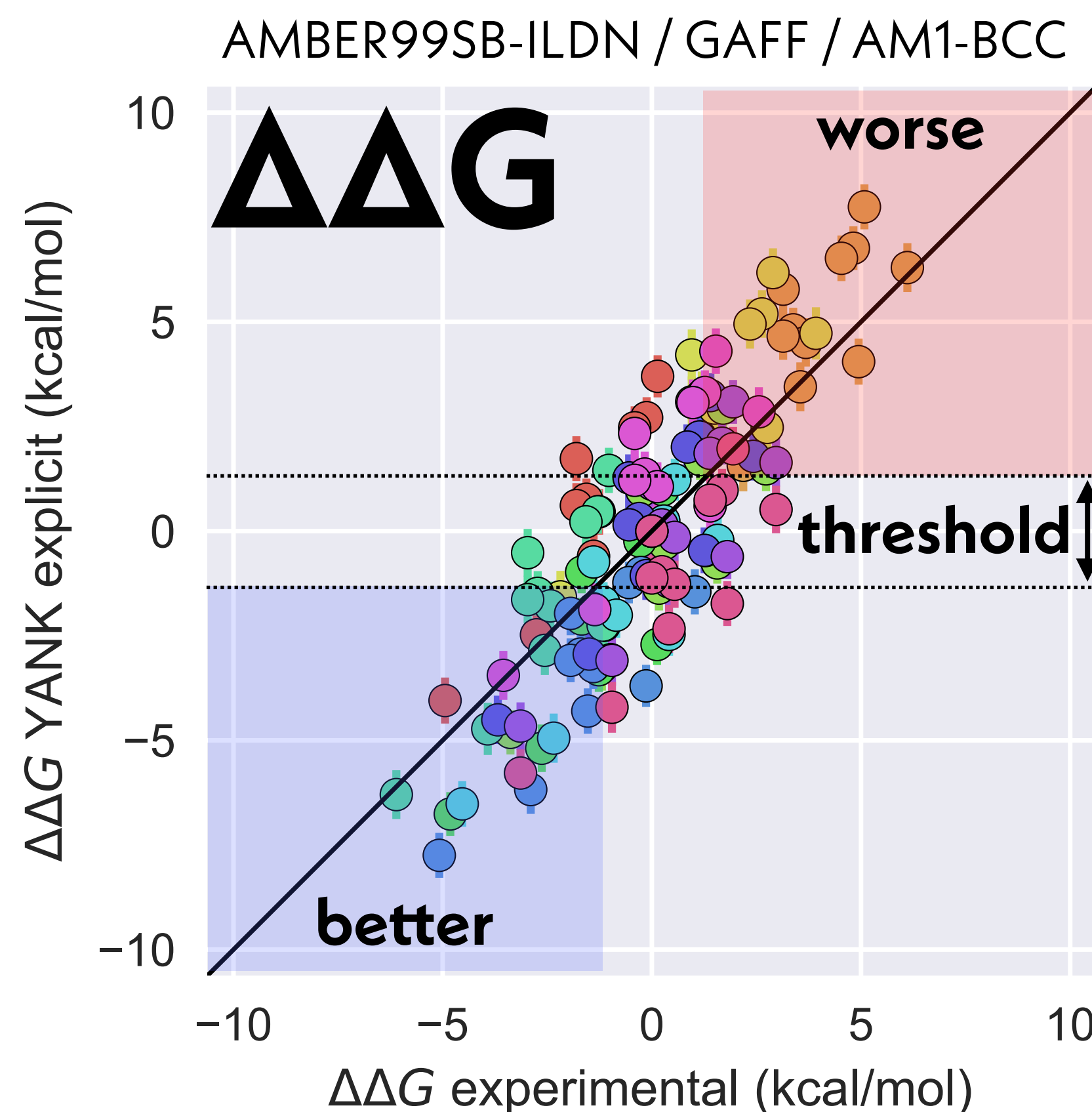
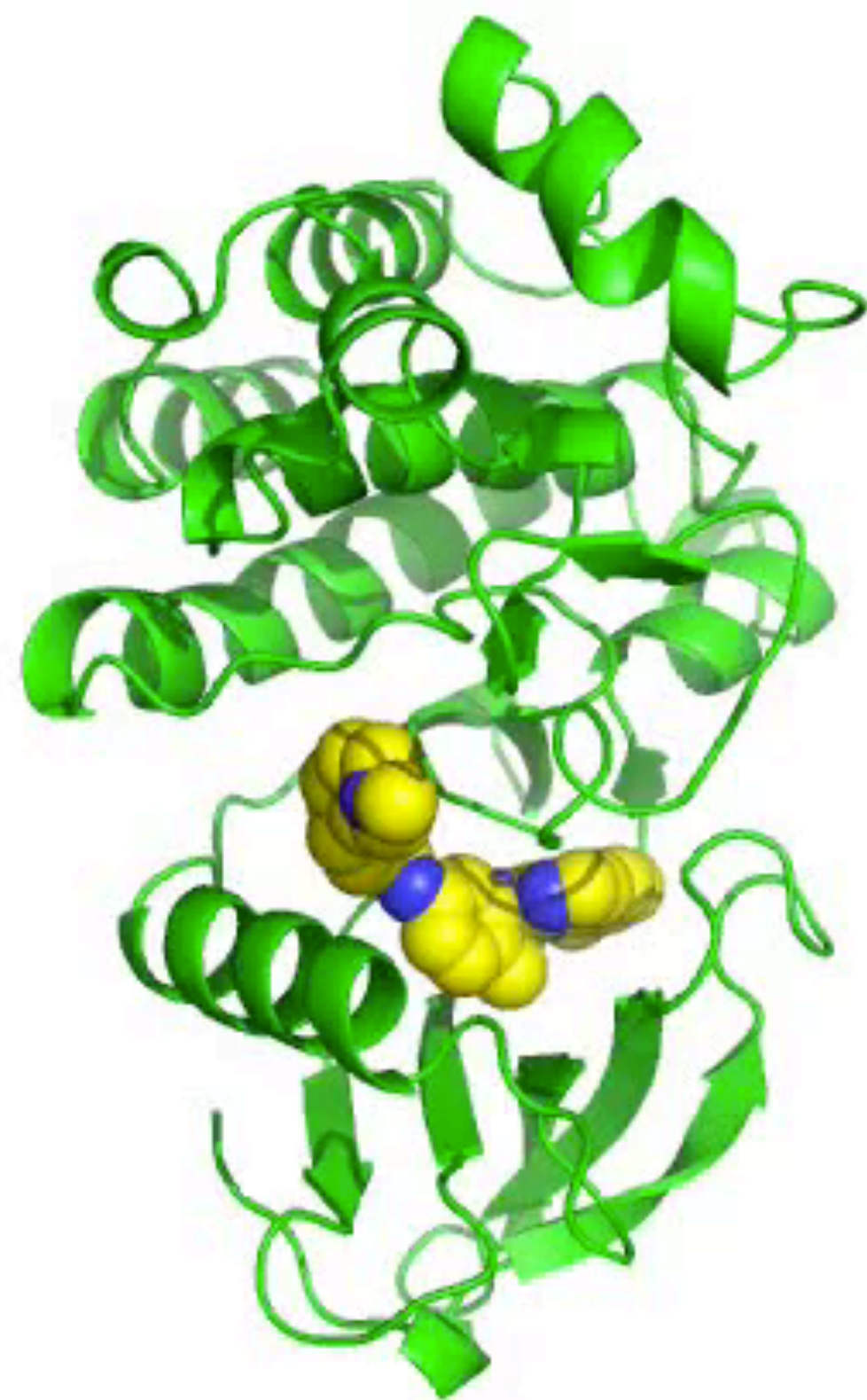


PAUL CZODROWSKI
DANIEL KUHN



ALCHEMICAL FREE ENERGY CALCULATIONS CAN HELP PRIORITIZE LIGAND SYNTHESIS

c-Met inhibitors from Bioorg. & Med Chem Lett. 25:1597, 2015
<https://github.com/choderalab/yank-benchmark>



LEVI NADEN
ANDREA RIZZI



PAUL CZODROWSKI
DANIEL KUHN



ALCHEMICAL FREE ENERGY METHODS HAVE TRANSFORMED DRUG DISCOVERY ALREADY

Gilead Bags Early-Stage NASH Drug in \$1.2B Nimbus Deal

From www.fiercebiotech.com - April 4, 2016 1:54 PM

Gilead is anteing up a \$400 million upfront payment to buy a Phase I NASH drug from Nimbus Therapeutics, a Bill Gates-backed drug discoverer based in Cambridge, MA. And it's backing the deal with another \$800 million in milestones.

The news comes just weeks ahead of Phase I data for the drug, which is aimed at throwing up a hurdle in the cascade of biologic events that leads to fatty liver disease, a condition that has attracted some widespread interest from big biotechs as well as big pharma. Gilead's pipeline lists four separate programs for NASH in early and mid-stage development.

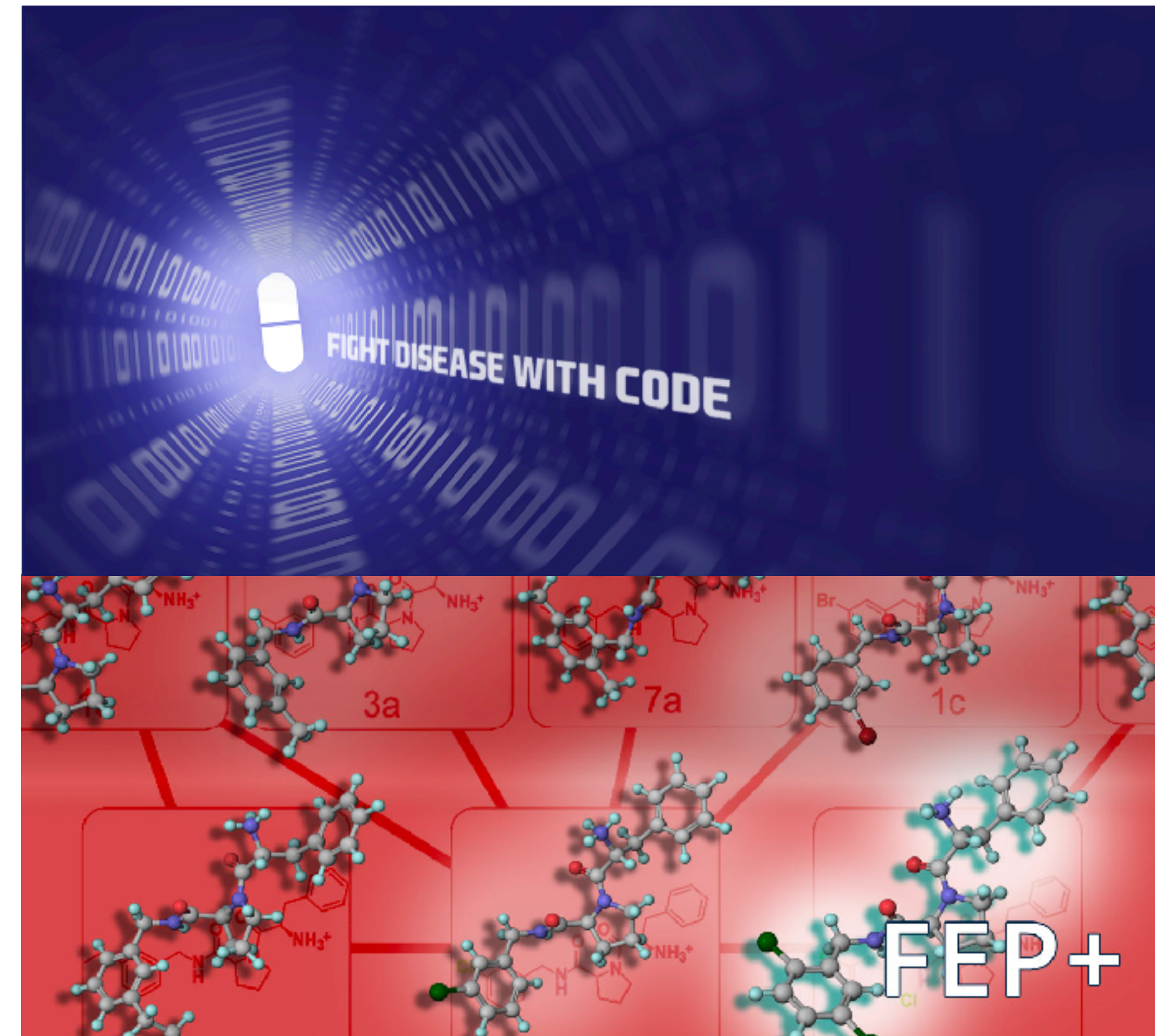


IntellaTurn's insight:

The Nimbus Apollo program includes the lead candidate NDI-010976, an ACC inhibitor, and other preclinical ACC inhibitors for the treatment of non-alcoholic steatohepatitis (NASH), and for the potential treatment of hepatocellular carcinoma (HCC) and other diseases.

<https://www.fiercebiotech.com/biotech/updated-gilead-bags-early-stage-nash-drug-1-2b-nimbus-deal>

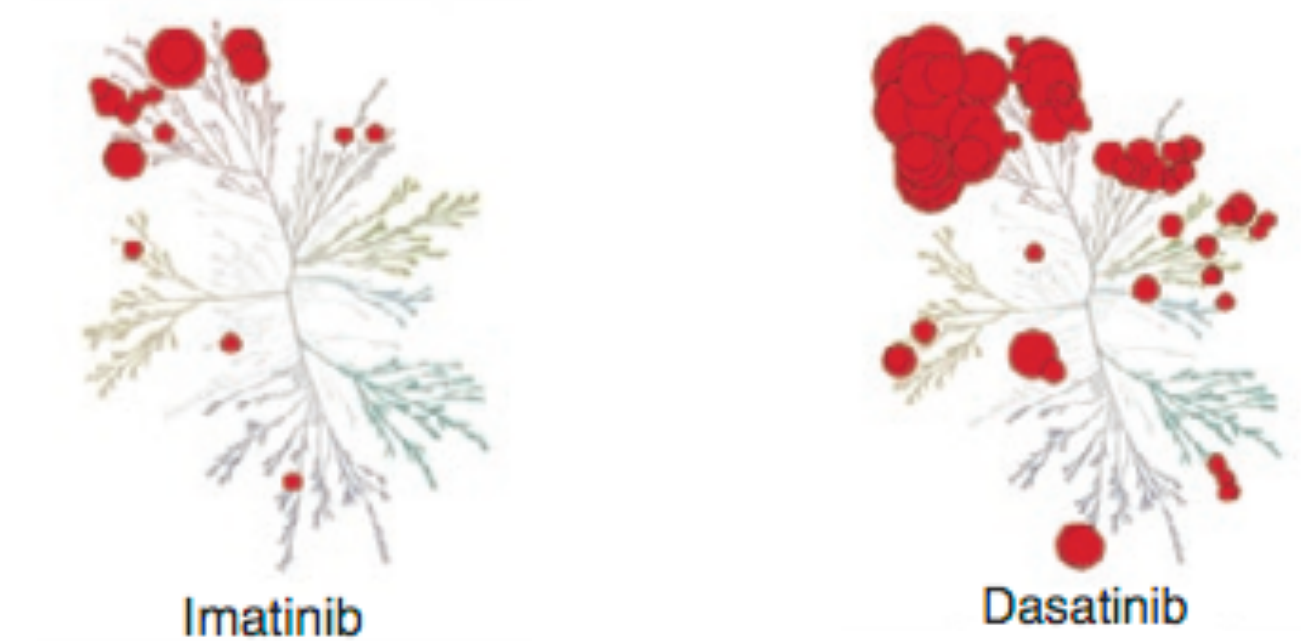
SCHRÖDINGER®



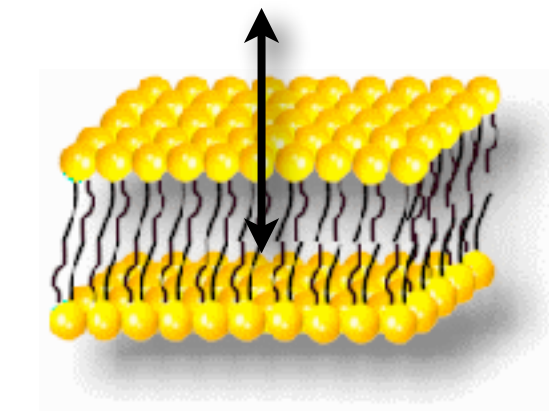
**Impact has currently been limited to affinity optimization.
What about selectivity?**

ALCHEMICAL METHODS CAN ALSO COMPUTE MANY OTHER USEFUL PROPERTIES

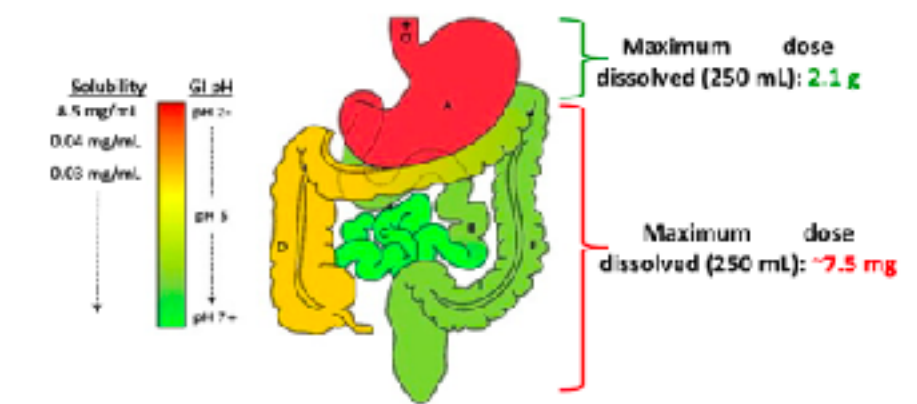
selectivity for targets/off-target effects
(of high relevance to selective kinase inhibition)



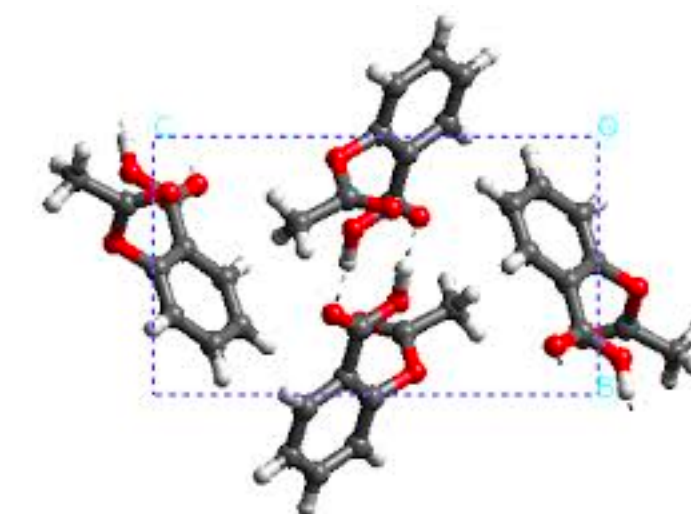
partition coefficients ($\log P$, $\log D$) and permeabilities



environment-dependent solubilities

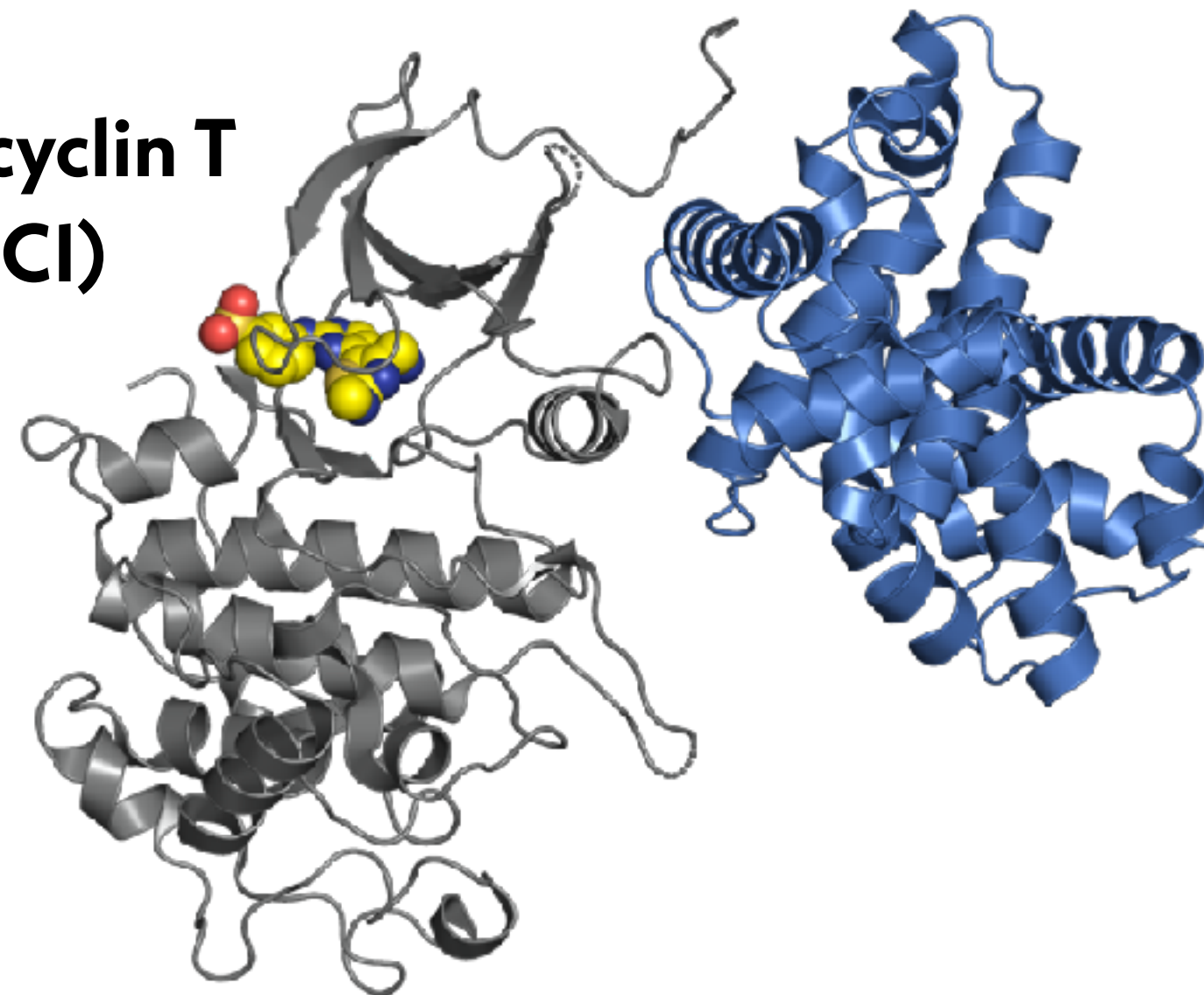


crystal polymorphs



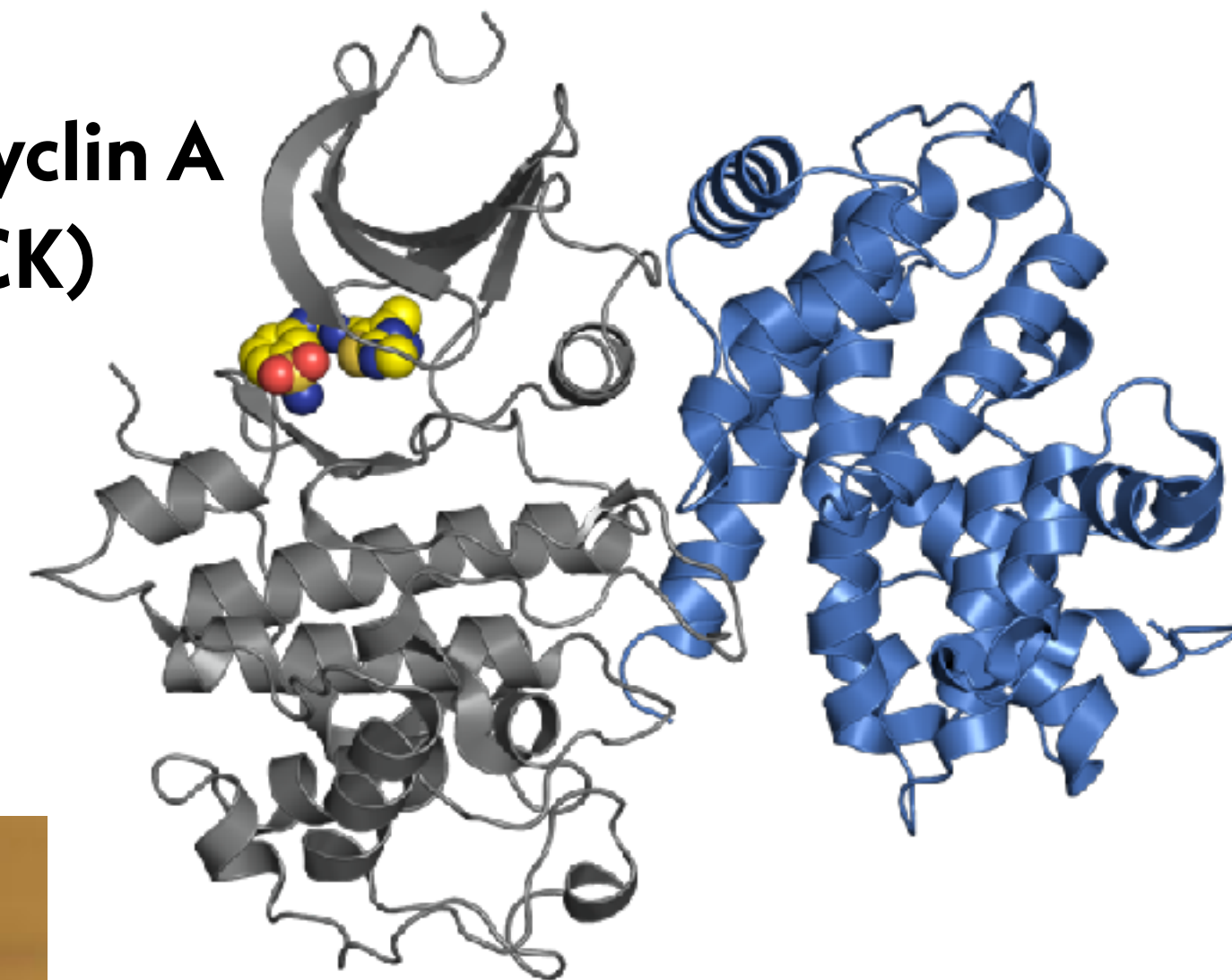
HOW WELL CAN WE PREDICT **SELECTIVITY**?

**CDK9/cyclin T
(4BCI)**

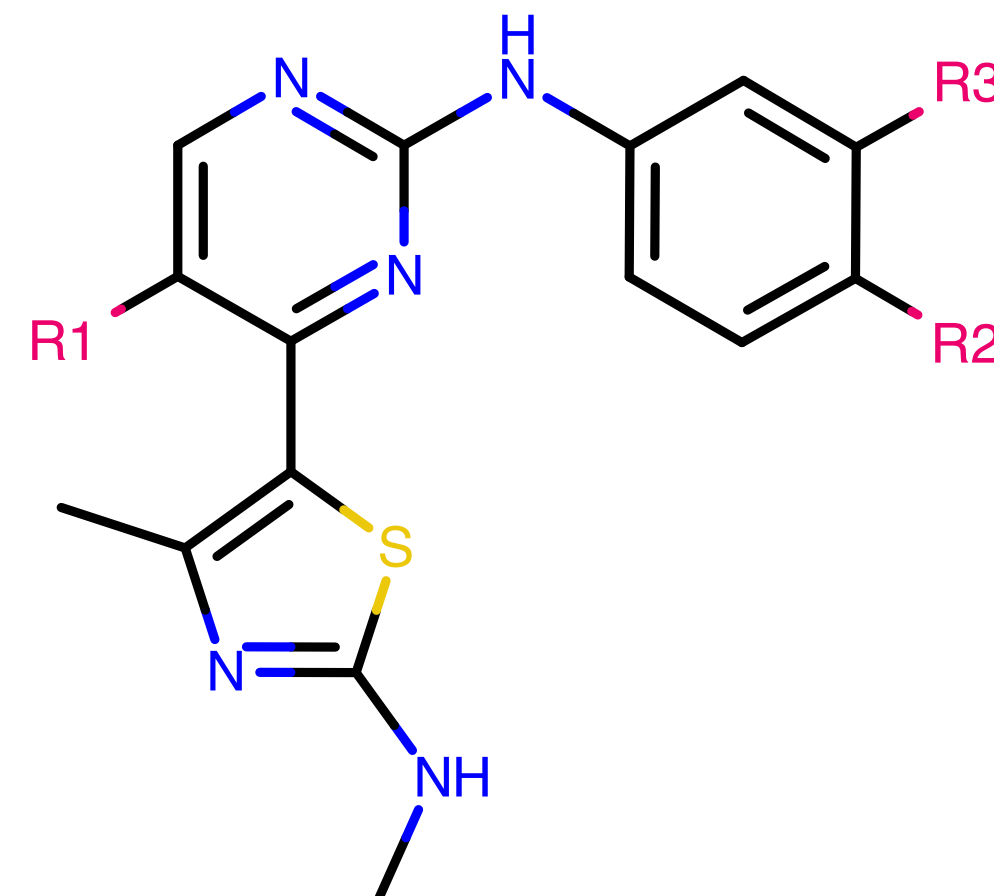


inhibition reinstates apoptosis in cancer cells

**CDK2/cyclin A
(4BCK)**



essential for S-phase progression



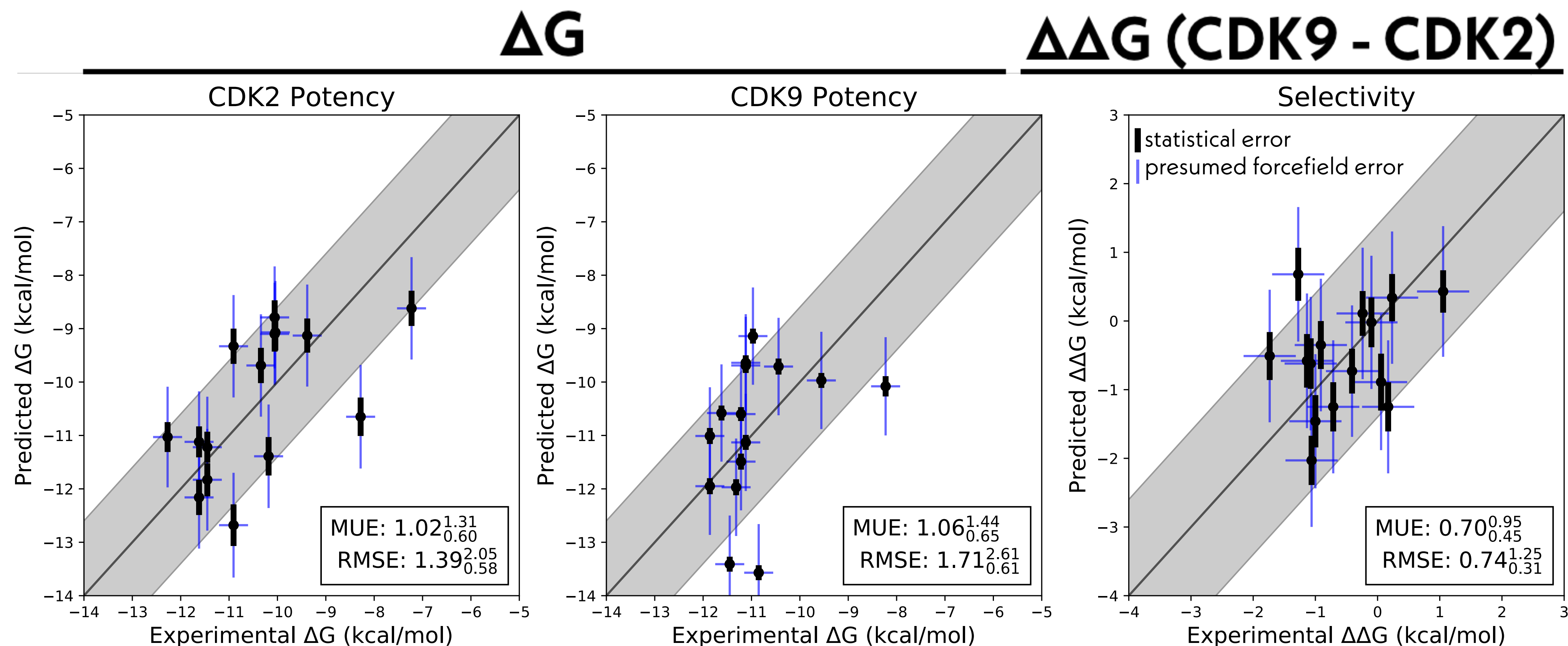
Ligand	R1	R2	R3	ΔG CDK2 (kcal/mol)	ΔG CDK9 (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
12a	CN	H		-12.27	-11.21	-1.64
12b	OH	H		-7.23	-8.22	-1.57
12c	CN	H		-11.45	-11.21	-1.57
12e	F	H		-11.62	-11.45	-1.57
12f	Cl	H		-10.91	-10.85	-2.36
12g	Methyl	H		-10.18	-11.32	-1.97
12h	Ethyl	H		-8.28	-9.56	-2.37
12j	CN	H		-10.04	-11.12	-1.56
12l	CN		H	-10.34	-10.44	-1.34
12n	CN	H		-10.06	-10.97	-2.47
12o	F	H		-10.06	-11.12	-0.75
12q	F	H		-10.91	-11.62	-2.31
12t	CN	H		-9.38	-11.12	-1.91
1a	H	H		-11.62	-11.86	-2.77
1b	H	H		-11.45	-11.86	-1.77

Shao et al., J Med Chem 56(3), 640–659



STEVEN ALBANESE

ALCHEMICAL METHODS CAN ACCURATELY PREDICT BINDING AFFINITIES TO INDIVIDUAL CDKS

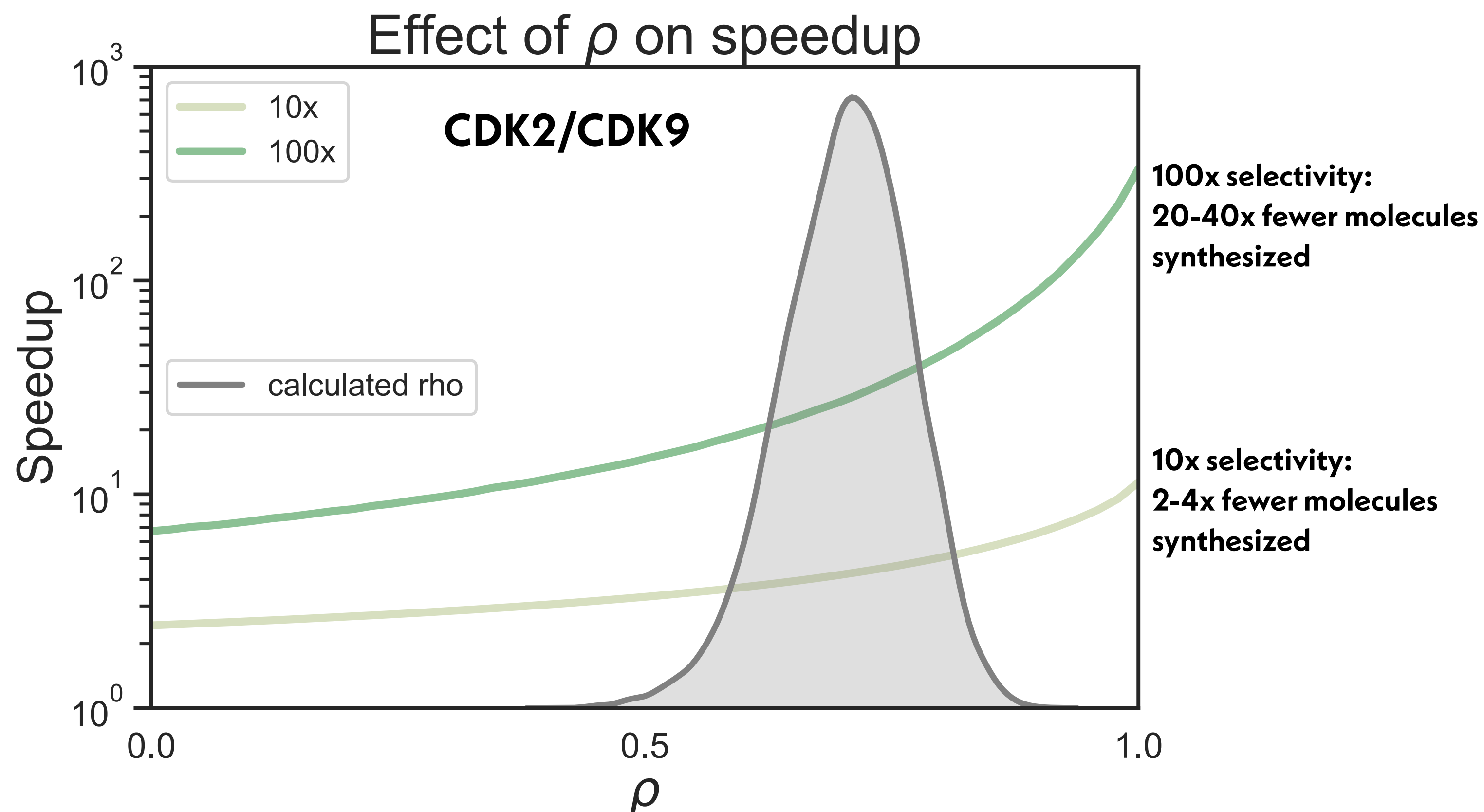
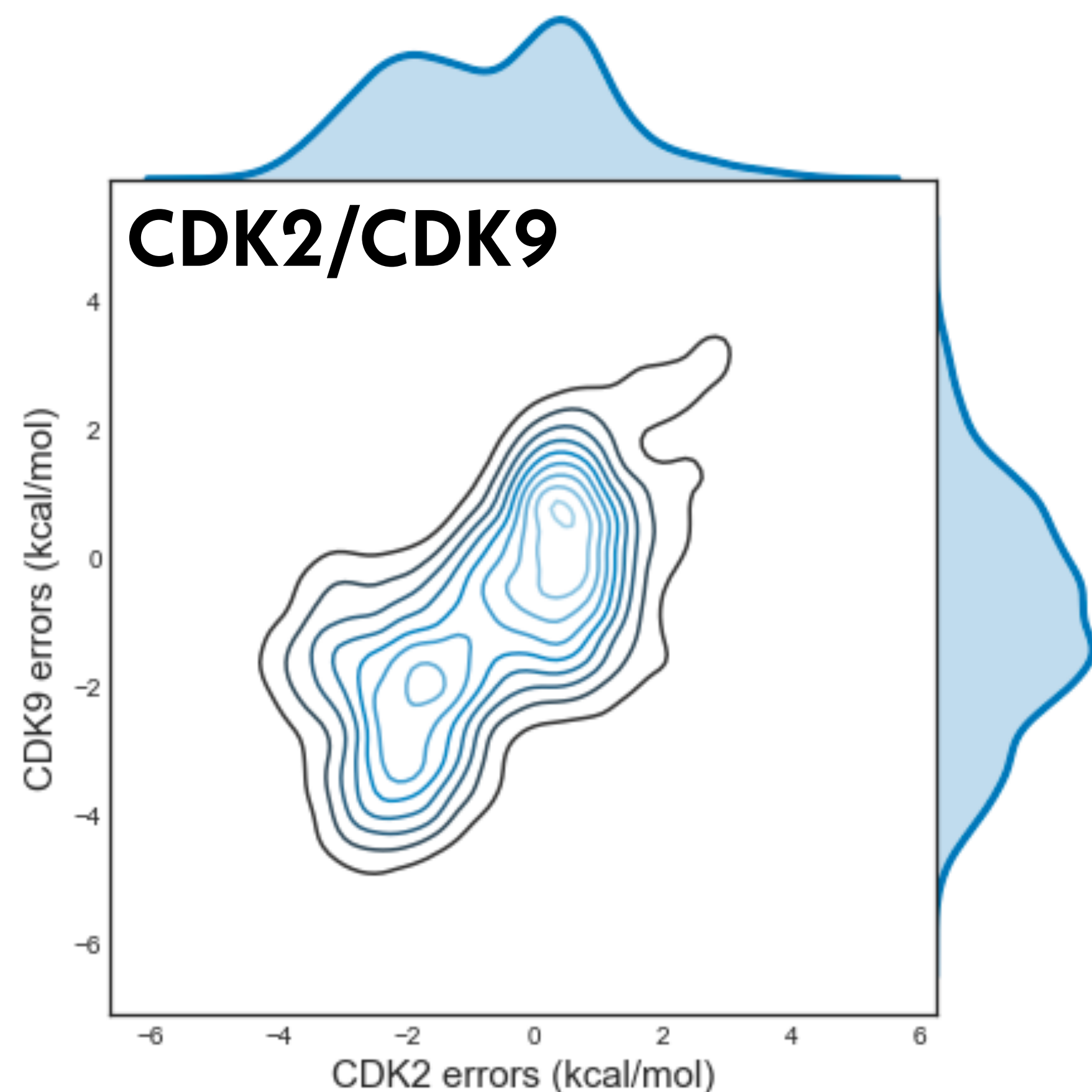


Individual affinities predicted confidently,
but what does this mean for selectivity?

STEVEN ALBANESE

FEP+/OPLS3
LINGLE WANG
SCHRÖDINGER

HOW MUCH CAN FREE ENERGY CALCULATIONS ACCELERATE SELECTIVE INHIBITOR DISCOVERY?



Achieving 100x selectivity is difficult,
but predictive modeling can have substantial impact.

STEVEN ALBANESE

FEP+/OPLS3
LINGLE WANG
SCHRÖDINGER

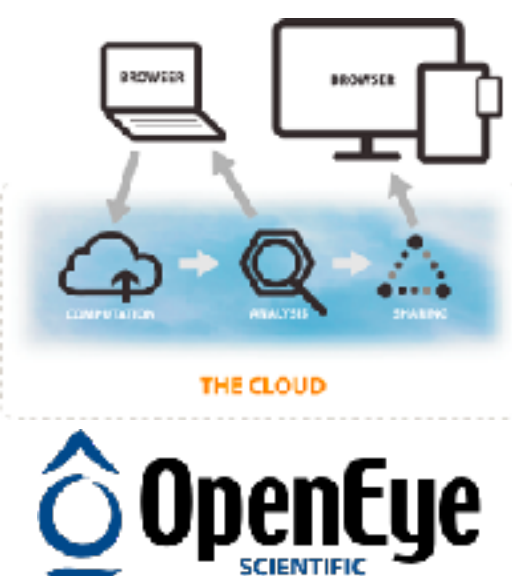
SETTING FREE ENERGY CALCULATIONS FREE



Free for academics, through the browser

The Orion software interface is a web-based platform for molecular modeling. It features a central 3D viewer showing a protein-ligand complex. To the left is a sidebar with navigation icons and a list of visible objects, including various chemical structures identified by IDs like CHEMBL3402755_4200. To the right of the 3D viewer are several tool panels: "LIGAND TOOLS" with options for "Generate Conformers", "Overlay", "Target", and "Align"; "BINDING SITE TOOLS" with "Binding Site View" and "Docking" options; and a "Data Handling" panel. At the bottom, a "Layout" panel displays a grid of chemical structures, each with a button to "Add columns to load in column list". The interface is clean and professional, with a blue and white color scheme.

This screenshot shows the "Binding Affinity Replica Exchange" workflow page in the Orion interface. The page is divided into several sections. At the top, there's a "Browse Workflows" section with a search bar. Below that is a detailed description of the Absolute Binding Affinity Free Energy (ABFE) protocol, explaining that it uses YANK Replica Exchange to calculate binding affinities. The "Required Input Parameters" section lists "Ligands: Dataset of the prepared ligands" and "Proteins: Dataset of the prepared protein". The "Outputs" section indicates that the output is a "Dataset of the solvated systems with the calculated binding free energies and free reports". The "Job Properties" section shows the job name "Binding Affinity Replica Exchange" and the date "Oct 14, 12:13 AM". The "Inputs" section has two required fields: "Protein Reader - Protein Input File" and "Ligand Reader - Ligand Input File", both with "Choose input..." buttons. The "Outputs" section has two required fields: "Out - Output Dataset" and "Failures - Output Dataset", both with empty text boxes and "This parameter is required" labels.



THE OPEN FORCE FIELD CONSORTIUM: BETTER BIOMOLECULAR FORCE FIELDS THROUGH OPEN SOURCE, OPEN DATA, AND OPEN SCIENCE

INDUSTRY



COORDINATING INTERMEDIARY



MOLECULAR SOFTWARE
SCIENCES INSTITUTE

coordination of funding

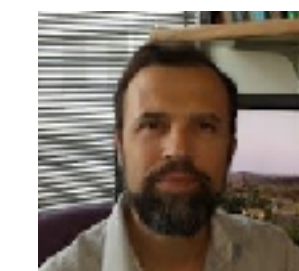
ACADEMIC



CHRISTOPHER BAYLY
OPENEYE SCIENTIFIC



JOHN CHODERA
SLOAN KETTERING INSTITUTE



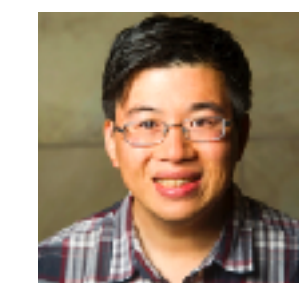
MICHAEL GILSON
UNIVERSITY OF CALIFORNIA, SAN DIEGO



DAVID MOBLEY
UNIVERSITY OF CALIFORNIA, IRVINE



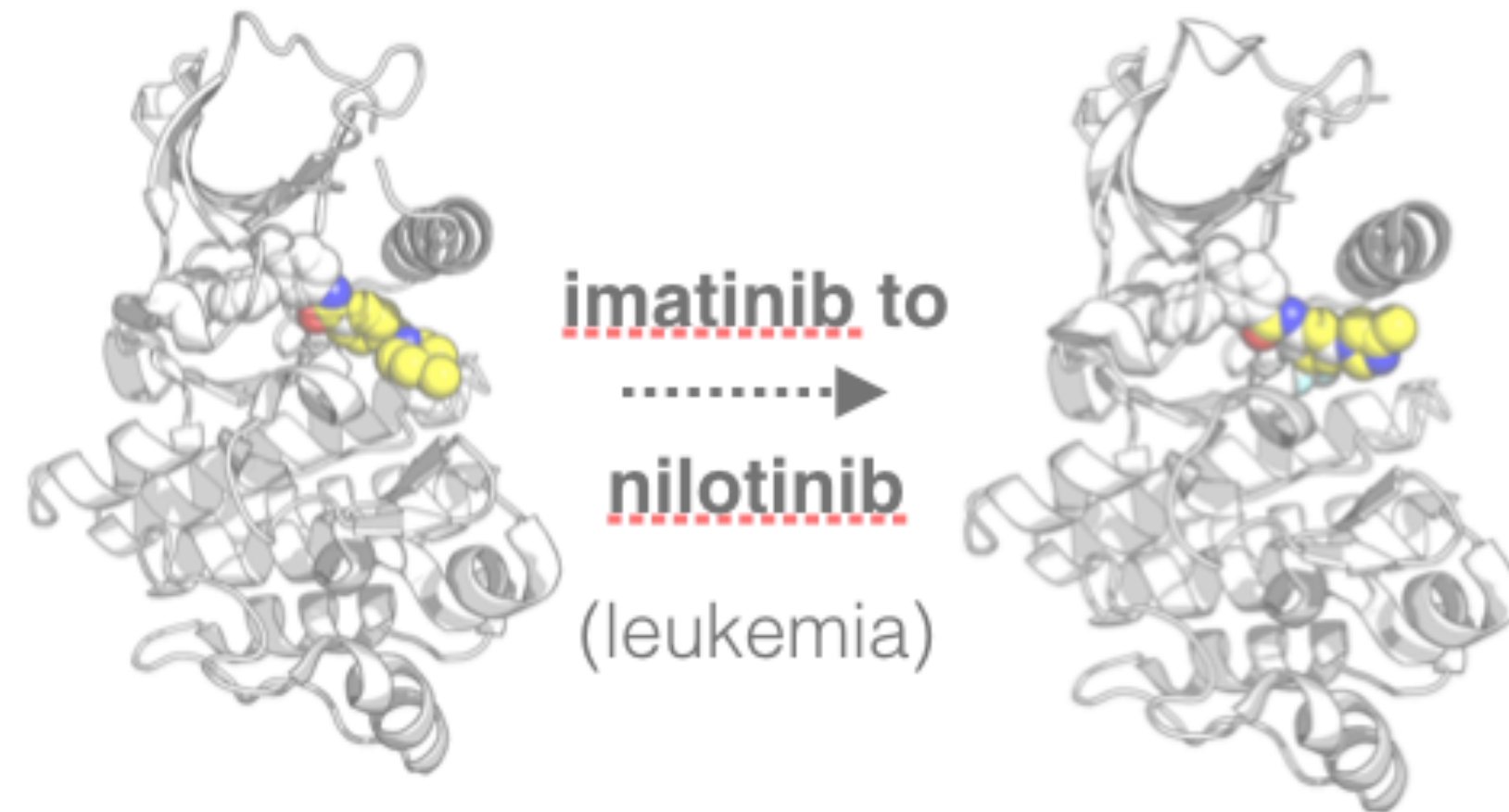
MICHAEL SHIRTS
UNIVERSITY OF COLORADO, BOULDER



LEE-PING WANG
UNIVERSITY OF CALIFORNIA, DAVIS

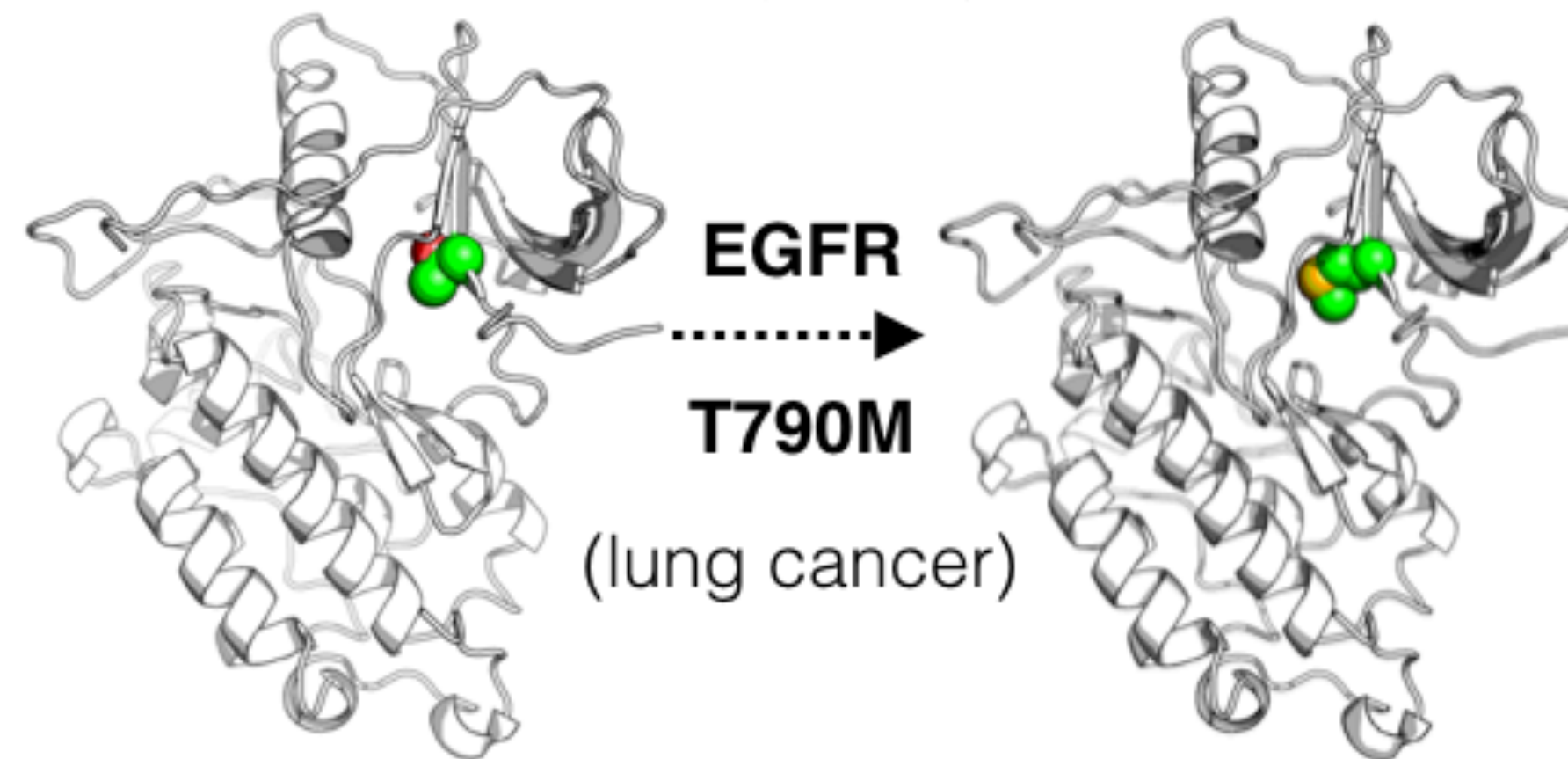
CAN WE USE FREE ENERGY CALCULATIONS TO ADDRESS MAJOR QUESTIONS IN CANCER THERAPY?

CHANGES OF A FEW ATOMS



HOW CAN WE DESIGN
SPECIFICALLY TARGETED
CANCER DRUGS?

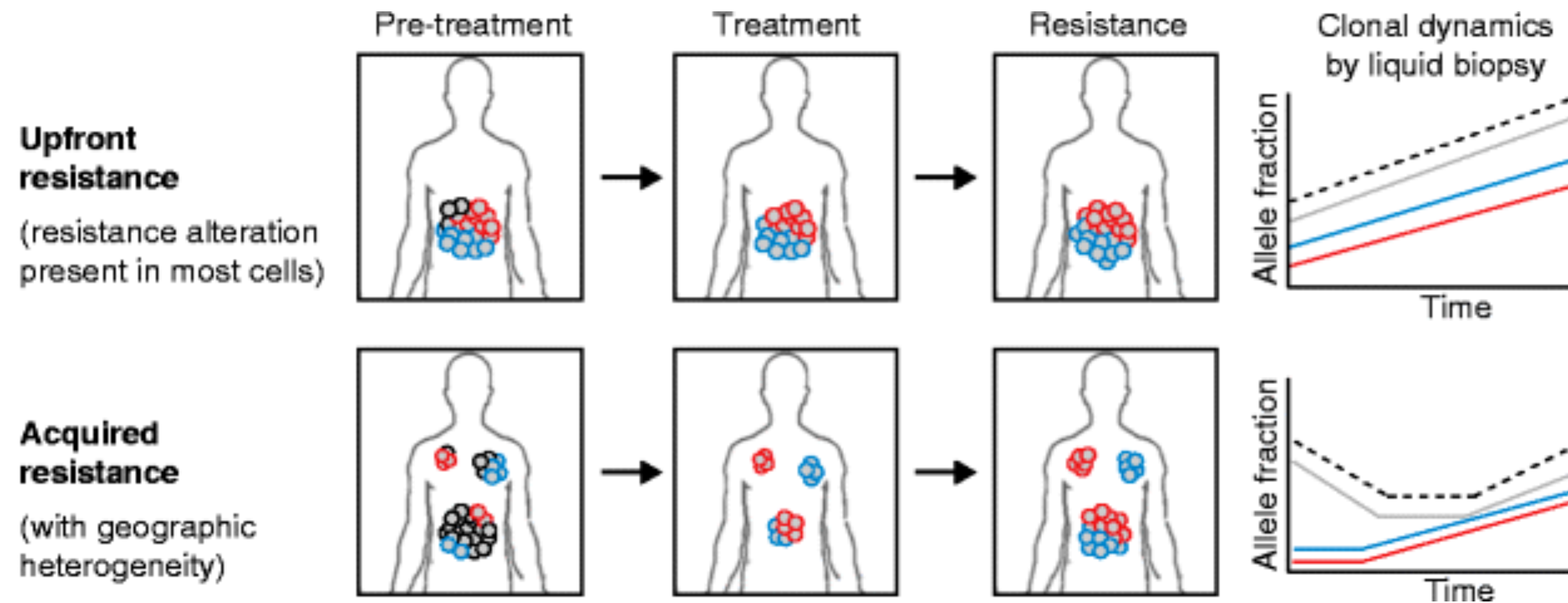
inhibitor modification
for drug design



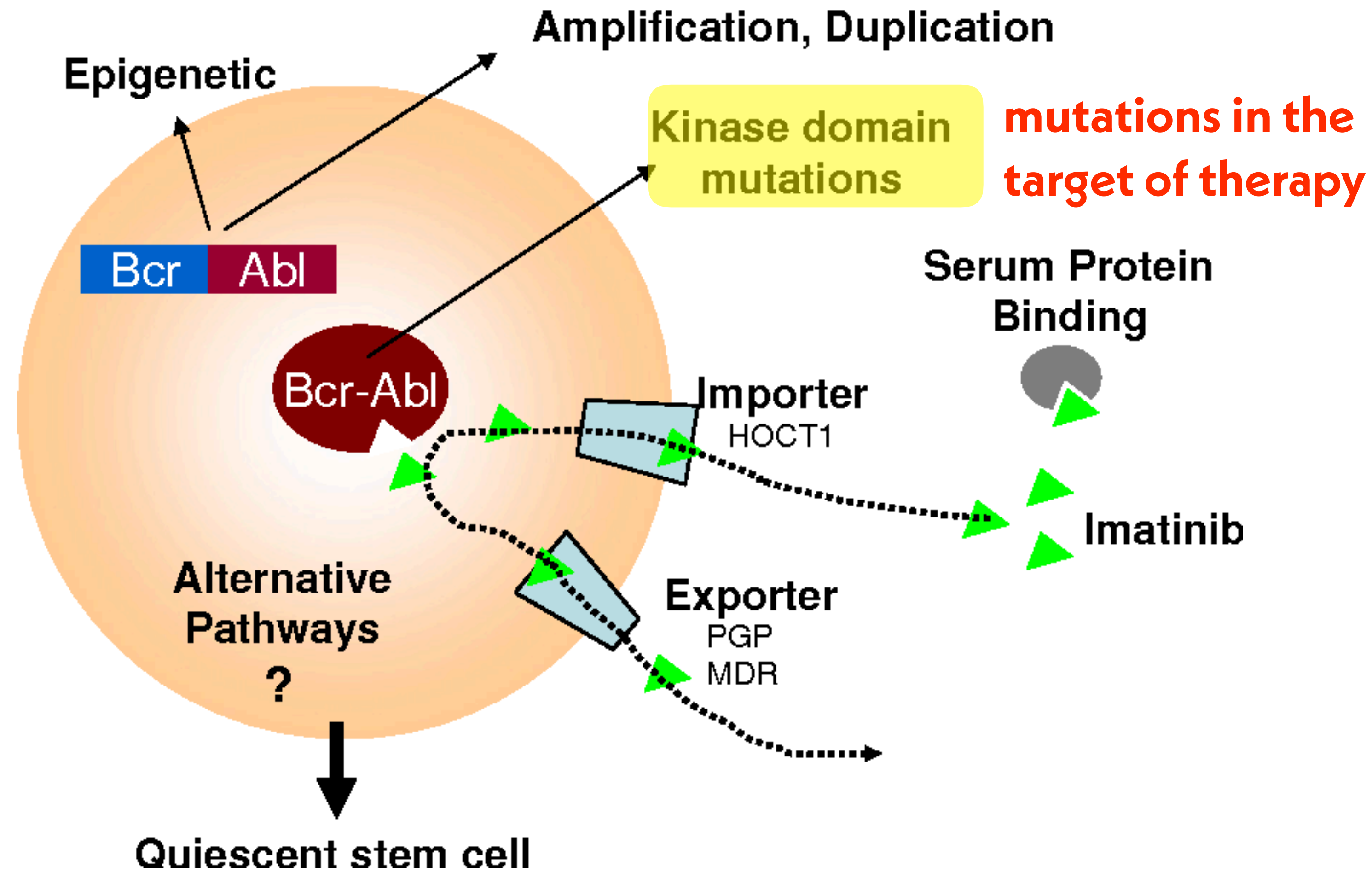
HOW CAN WE PREDICT
DRUG RESISTANCE
AND SUSCEPTIBILITY?

tumor-specific mutation
for therapeutic decisions

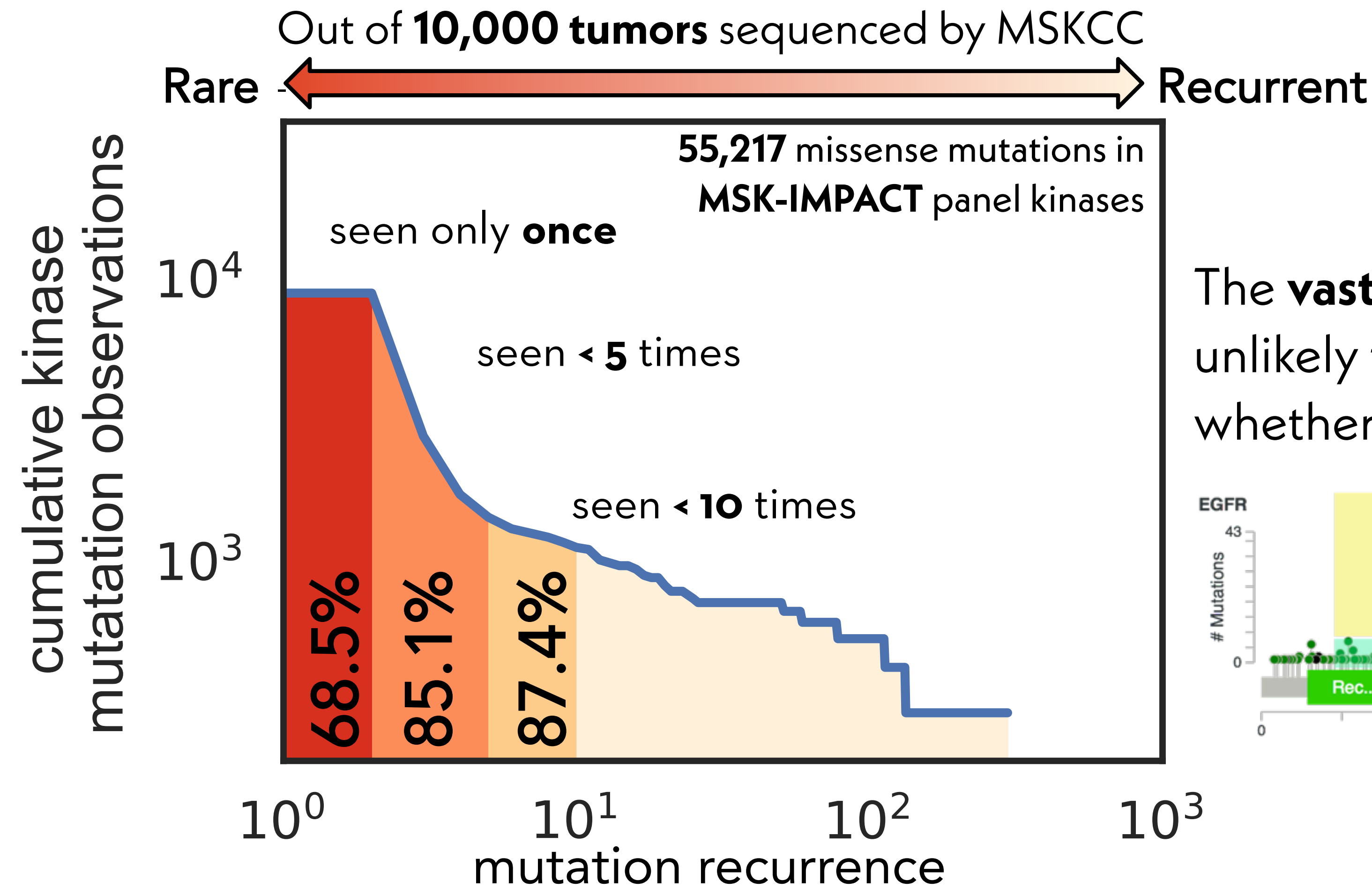
DRUG RESISTANCE IS A MAJOR CHALLENGE FOR TARGETED KINASE INHIBITOR THERAPY



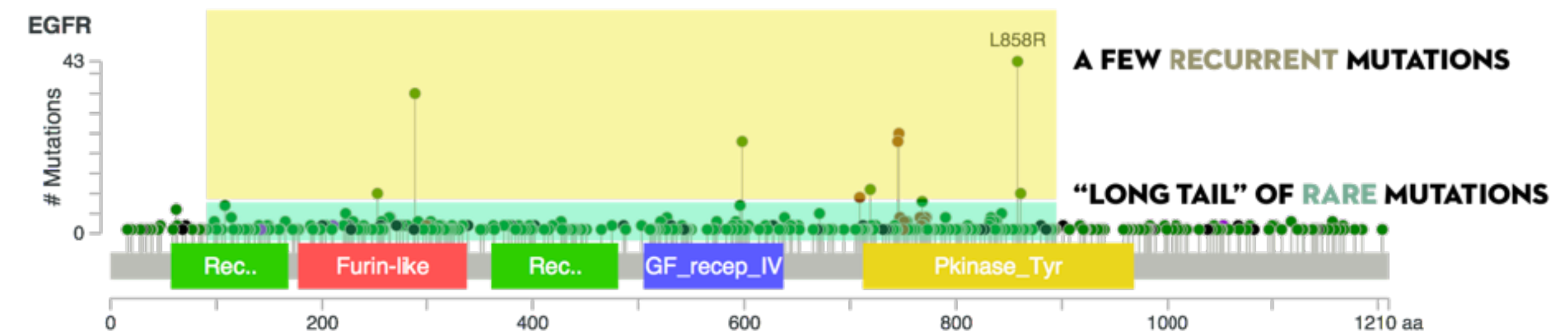
MULTIPLE MECHANISMS CAN MEDIATE DRUG RESISTANCE IN CANCER



THE LONG TAIL OF CANCER MUTATIONS FRUSTRATES THE PREDICTION OF RESISTANCE

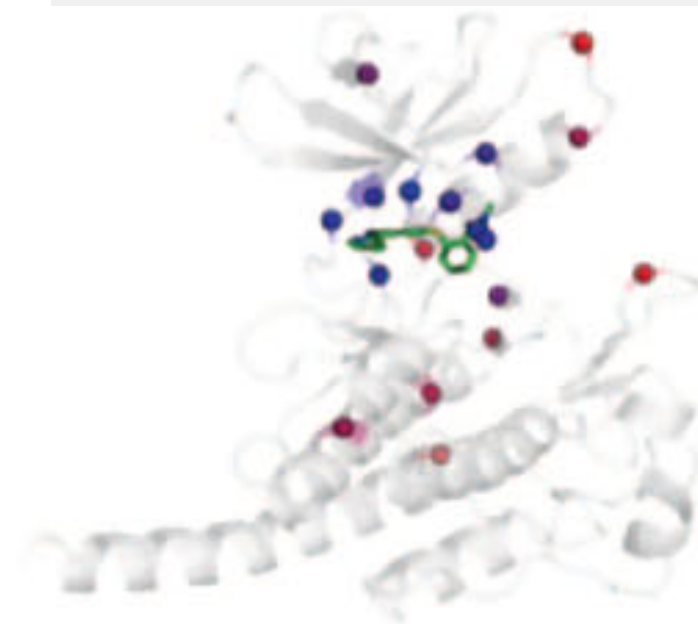


The **vast majority of mutations** are so rare there is unlikely to be clinical or biochemical evidence of whether they confer **drug resistance** or **susceptibility**

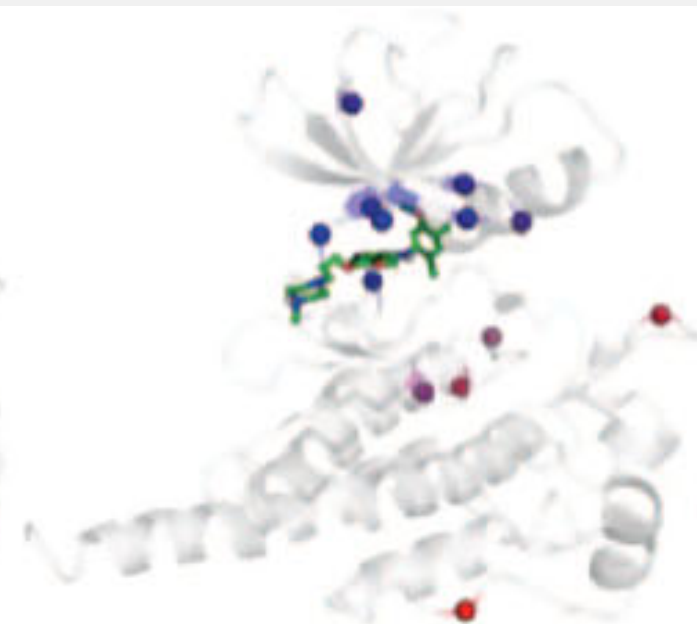


ABL KINASE DOMAIN POINT MUTANTS ARE WIDELY DISTRIBUTED

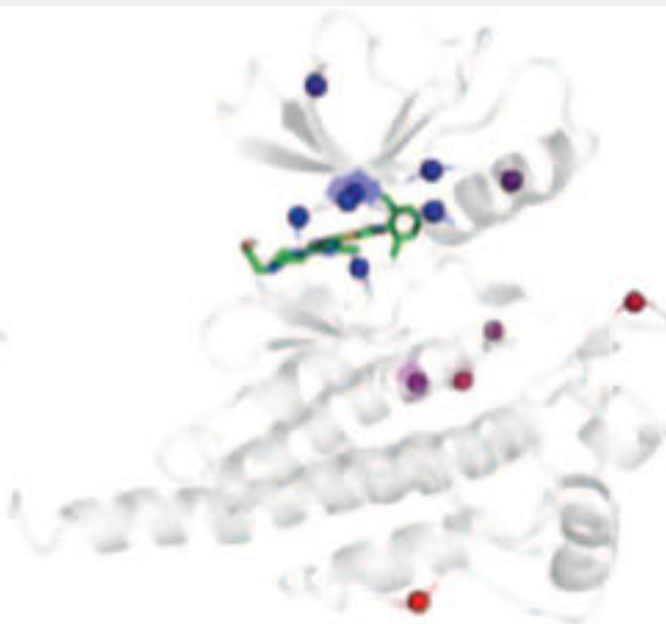
Abl:TKI Co-crystal structures



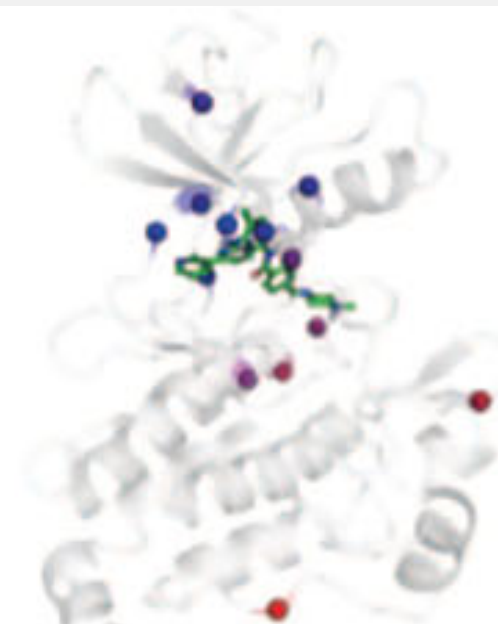
Abl:axitinib
PDB: 4WA9



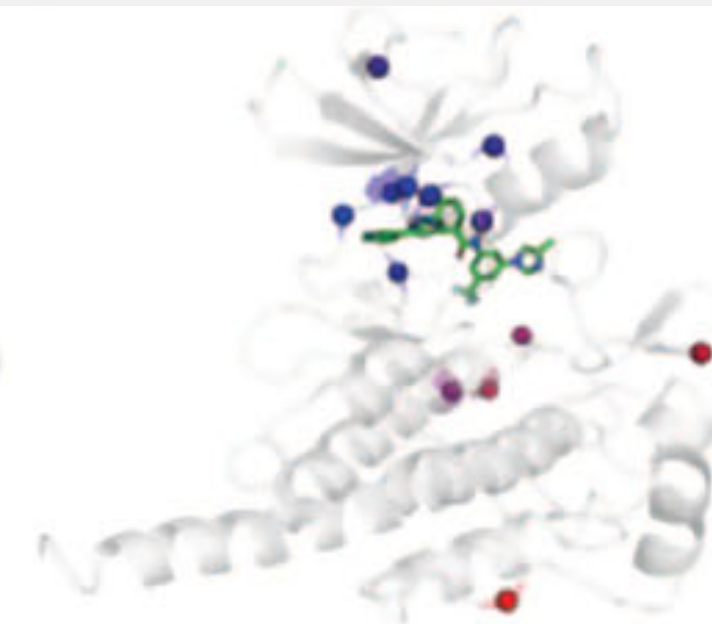
Abl:bosutinib
PDB: 3UE4



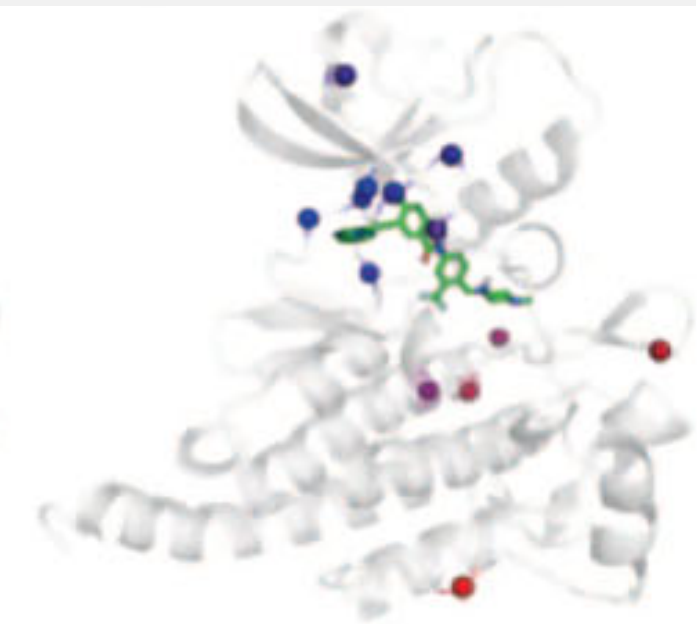
Abl:dasatinib
PDB: 4XEY



Abl:imatinib
PDB: 1OPJ



Abl:nilotinib
PDB: 3CS9

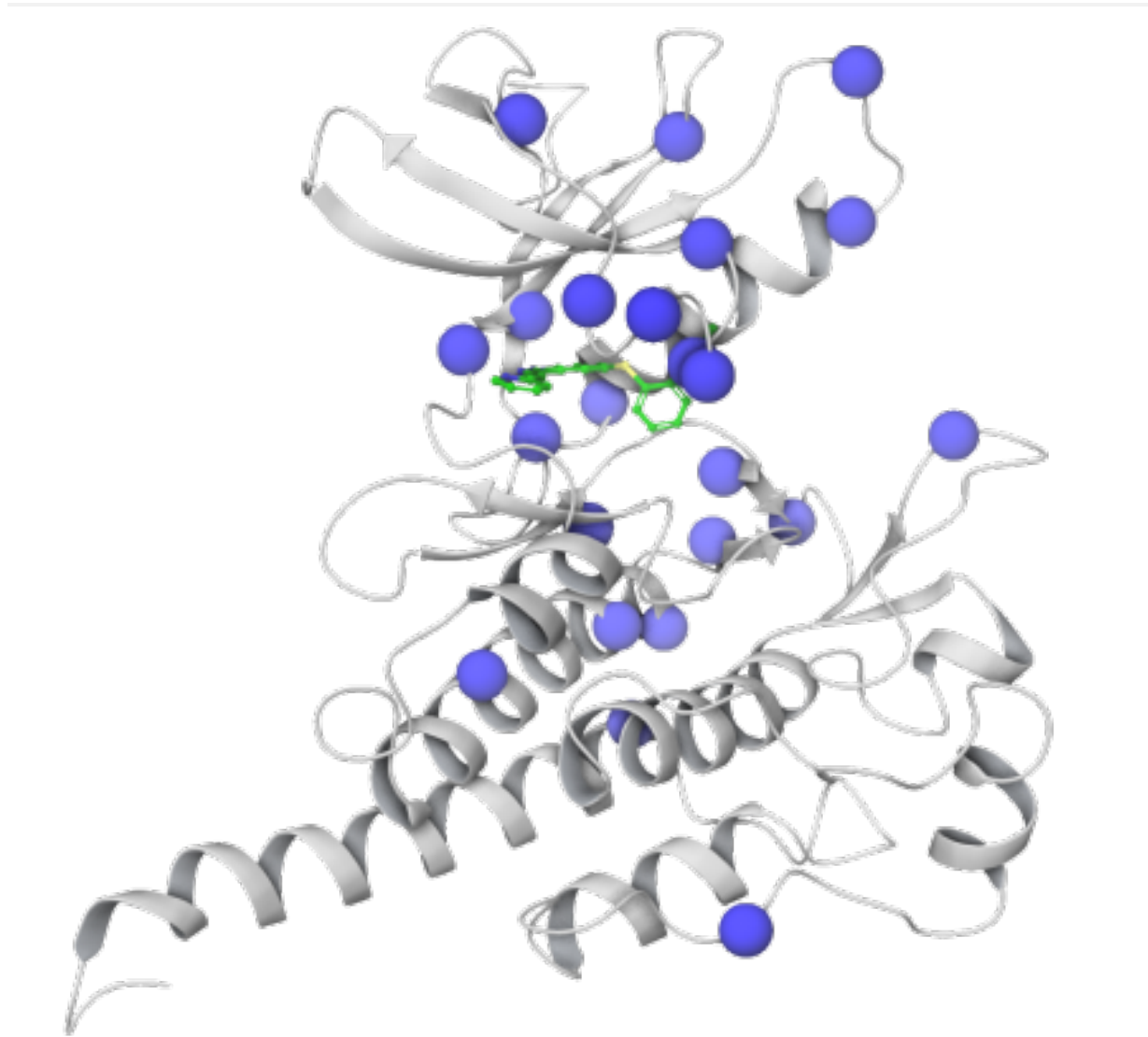


Abl:ponatinib
PDB: 3OXZ

KEVIN HAUSER
SCHRÖDINGER

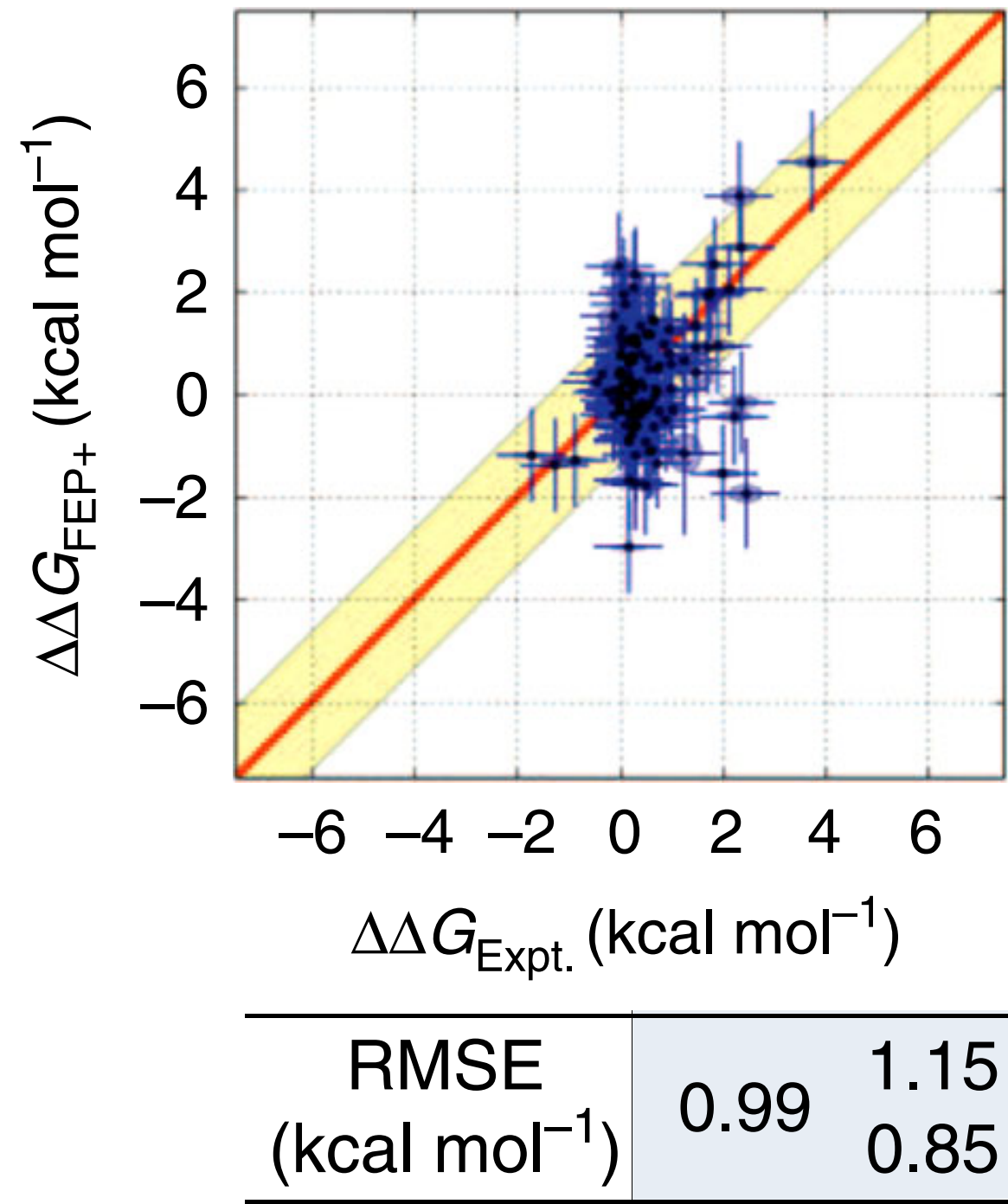


ALCHEMICAL FREE ENERGY CALCULATIONS SHOW PROMISE IN PREDICTING WHETHER CLINICAL MUTANTS ARE SUSCEPTIBLE OR RESISTANT



TKI	<i>N</i> _{mut}	<i>R</i>	<i>S</i>
Axitinib	26	0	26
Bosutinib	21	4	17
Dasatinib	21	5	16
Imatinib	21	5	16
Nilotinib	21	4	17
Ponatinib	21	0	21
Subtotal	131	18	113
Erlotinib	7	1	6
Gefitinib	6	0	6
Total	144	19	125

*N*_{mut} Total number of mutants for which Δ*p**I*₅₀ data was available
Number of **R**esistant, **S**usceptible mutants using 10-fold affinity change threshold



		<u>Prediction</u>	
		S	r
<u>Experiment</u>	S	105	8
	r	9	9

Accuracy	0.89	0.92 0.86
Specificity	0.91	0.94 0.89
Sensitivity	0.69	1.00 0.46

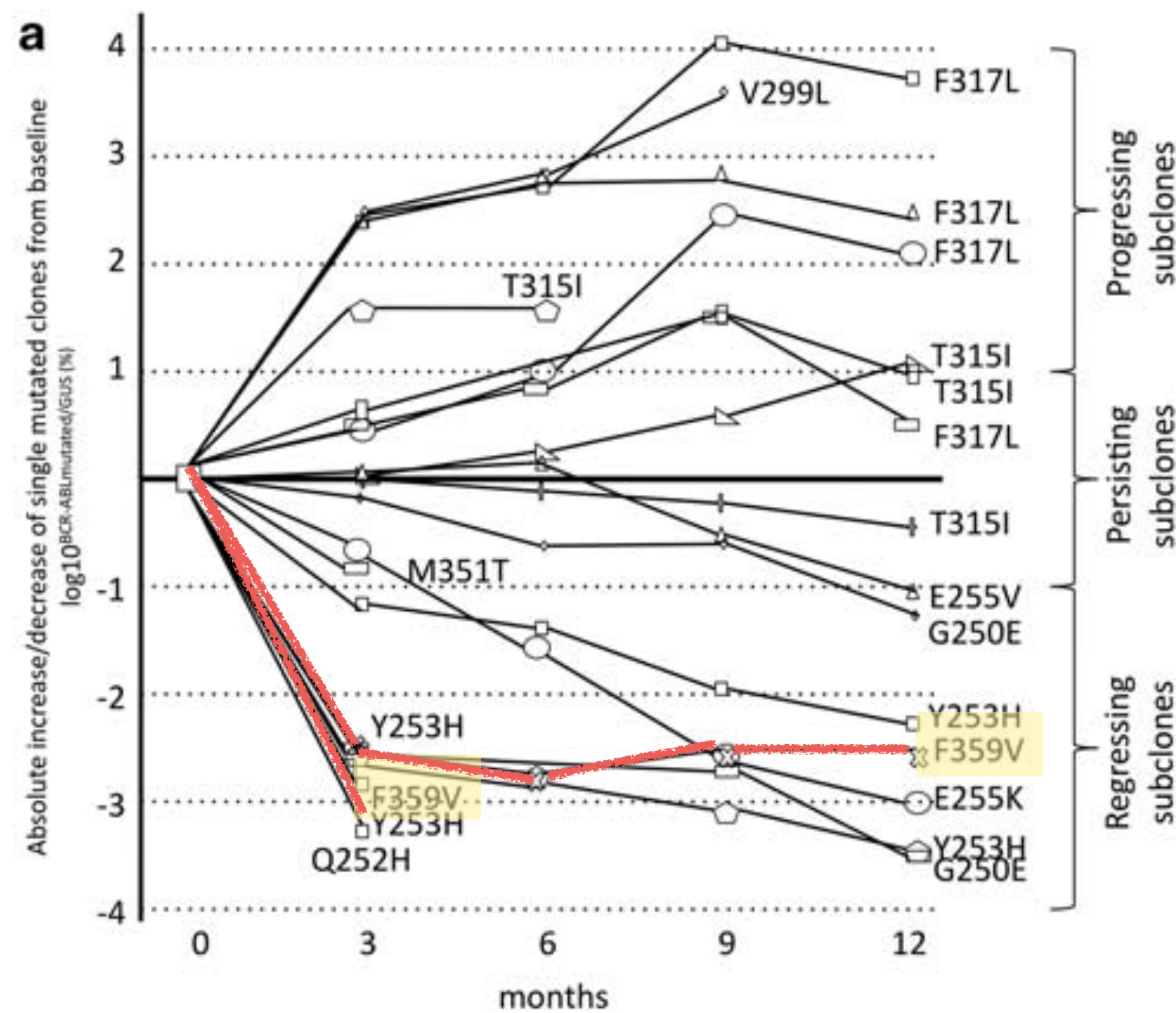
KEVIN HAUSER
SCHRÖDINGER



DIFFERENT DRUGS APPEAR TO EXERT DISTINCT SELECTIVE EVOLUTIONARY PRESSURES

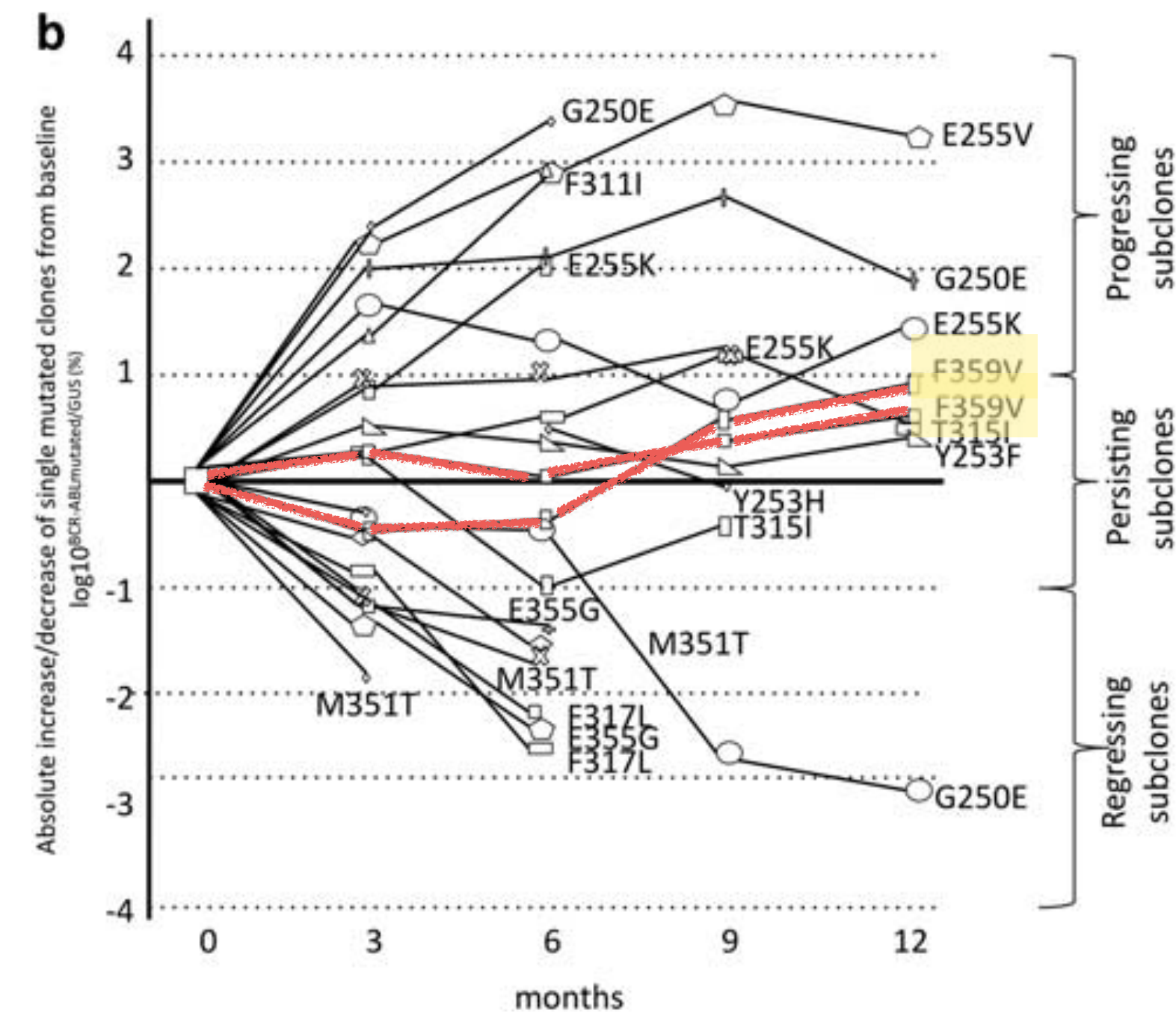
CML patients failing out of **imatinib** therapy often different kinds of resistance depending on the choice of second-line therapy:

dasatinib treatment

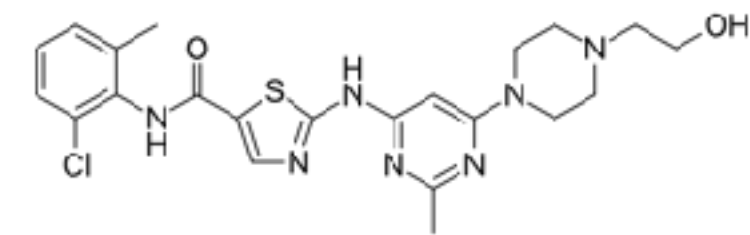


F359V DEPLETED

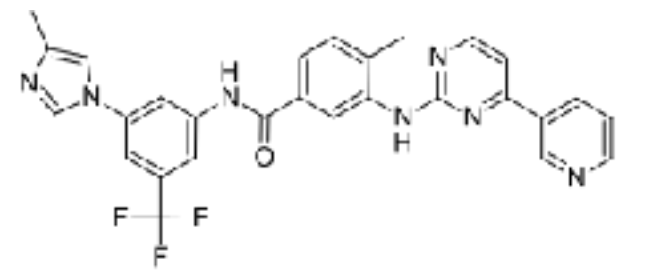
nilotinib treatment



F359V ENRICHED



dasatinib



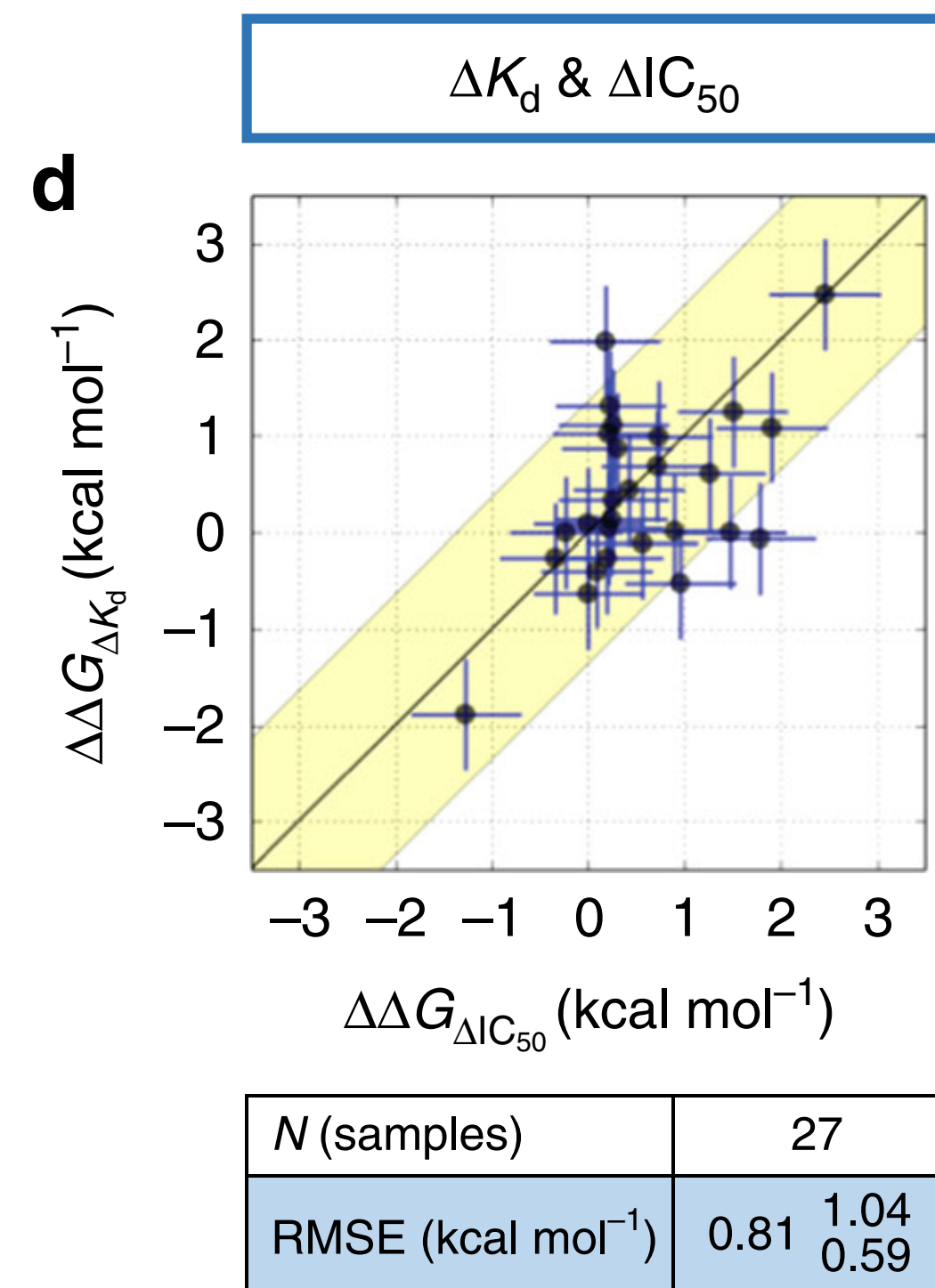
nilotinib

Can we model evolution of resistance using predicted affinities as a surrogate for fitness?

PHYSICAL MODELING OF DRUG RESISTANCE CAN OPEN UP NEW AVENUES OF RESEARCH

- Could we predict whether mutation confers **resistance** or **sensitivity** to any kinase inhibitor approved by the FDA or currently in clinical trials?
Can we **anticipate** resistance mutations before they appear in the clinic?
Get a head start on second- and third-generation inhibitors
- Could we **quantify** potential to elicit mutational resistance?
Prioritize molecules to promote to the clinic that are less likely to elicit resistance
- Can we evaluate selective pressures from **combinations**?
- Could we **engineer** inhibitors unlikely to elicit mutational resistance?

WE NEED MORE GOOD BIOPHYSICAL AFFINITY DATA FOR CLINICAL CANCER MUTATIONS



IC50 data **does not agree well** with affinity data
IC50 data is **highly limited**, while cbiportal has 10^4 - 10^5 mutations

AUTOMATED BACTERIAL EXPRESSION YIELDS A DIVERSE PANEL OF KINASE DOMAINS INTO WHICH WE CAN INTRODUCE CLINICAL MUTANTS



STEVEN ALBANESE



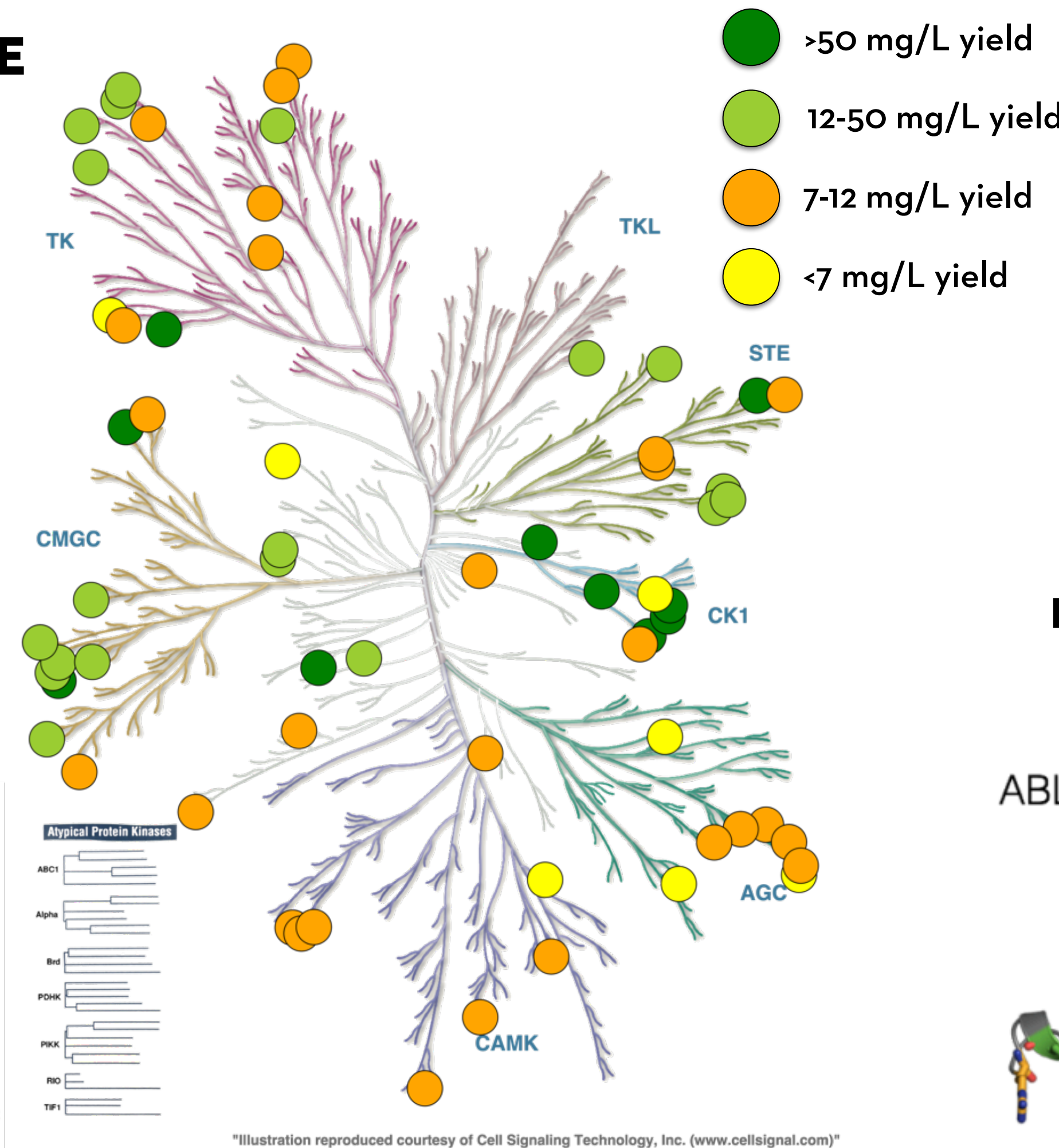
DANIEL PARTON



MEHTAP ISIK



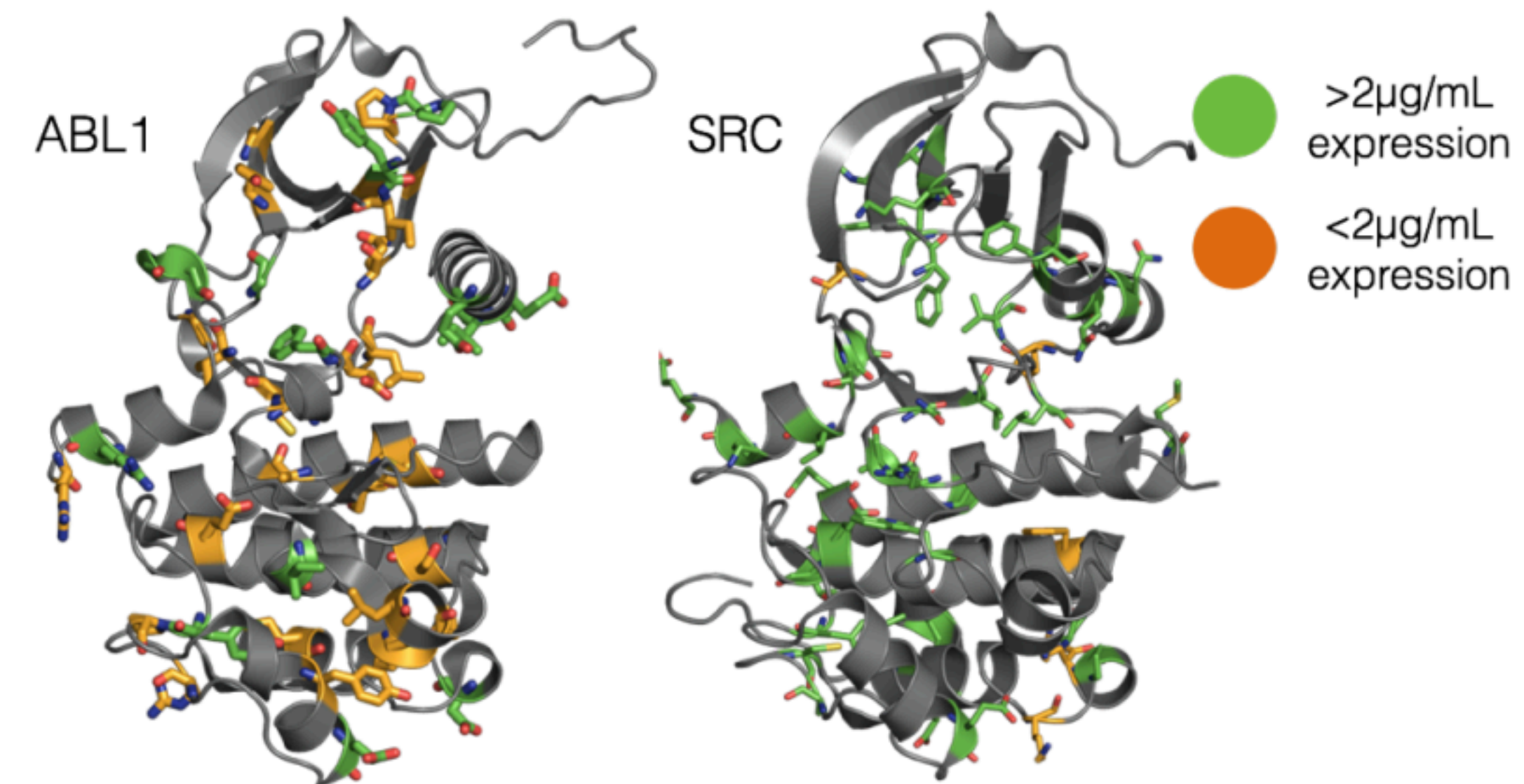
**LUCELENIE
RODRIGUEZ-
LAUREANO**



co-transform kinases with phosphatase



Expression of MSK-IMPACT mutants for Abl and Src

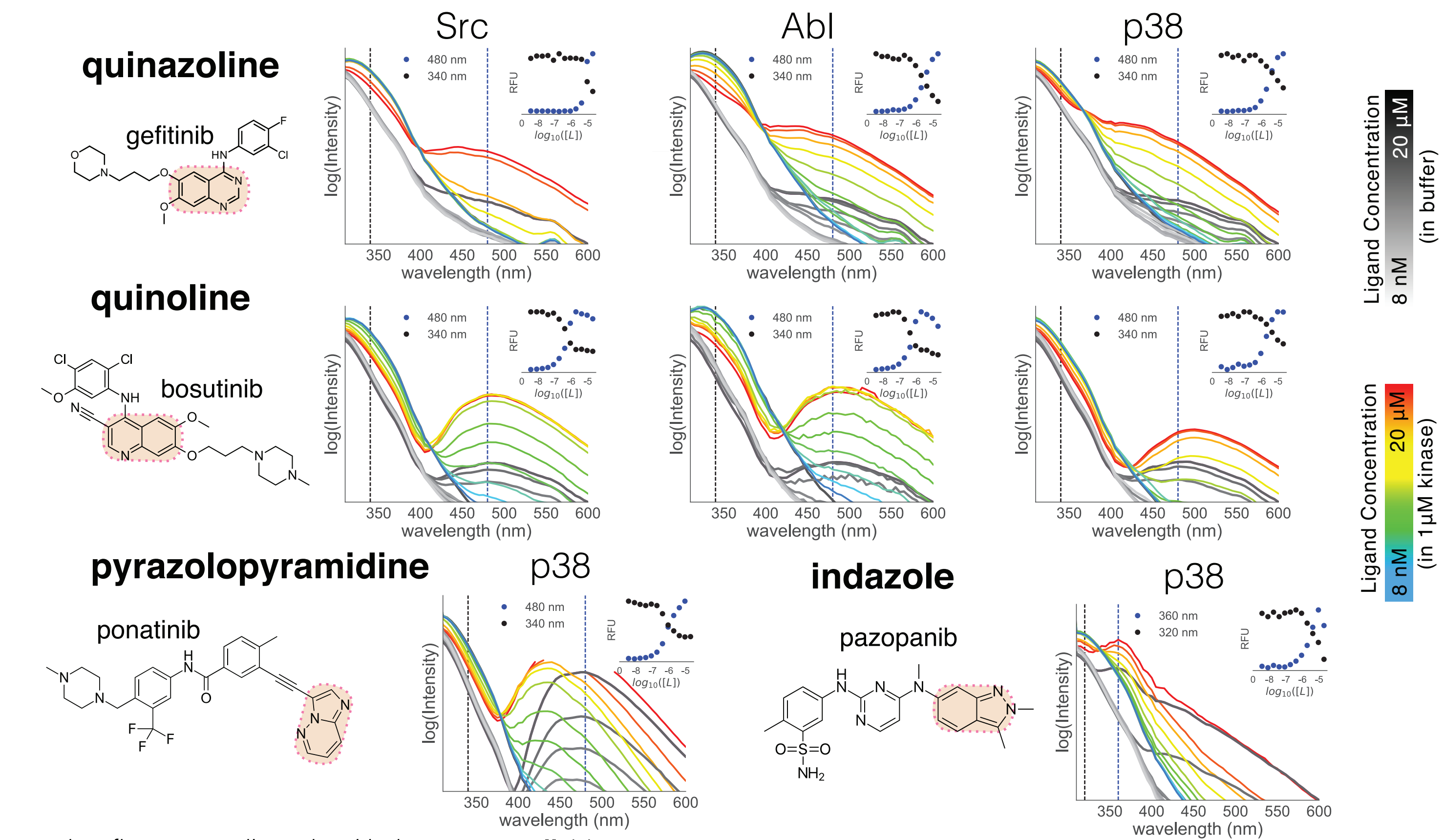


Biochemistry 57:4675, 2018

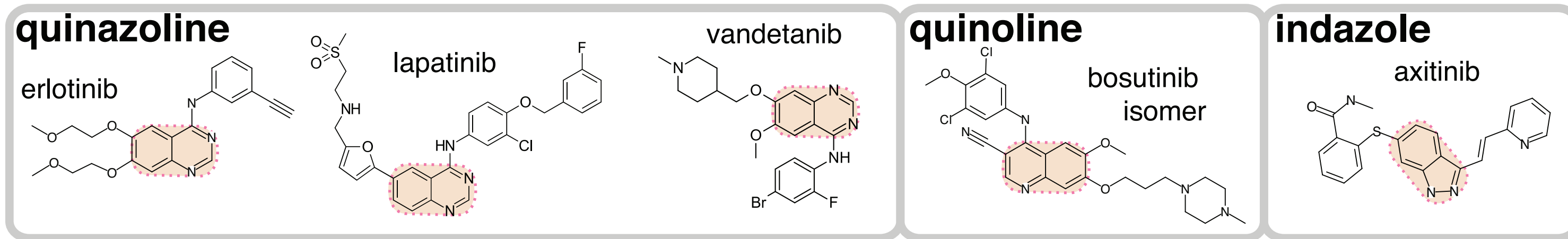
data: <http://choderalab.org/kinome-expression>

plasmids: <https://www.addgene.org/kits/chodera-kinase-domains>

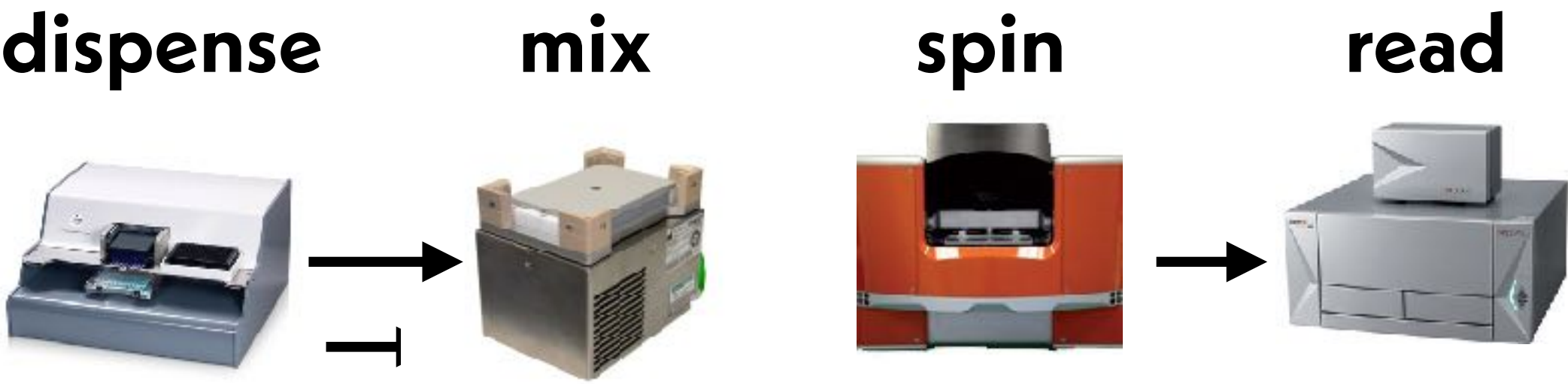
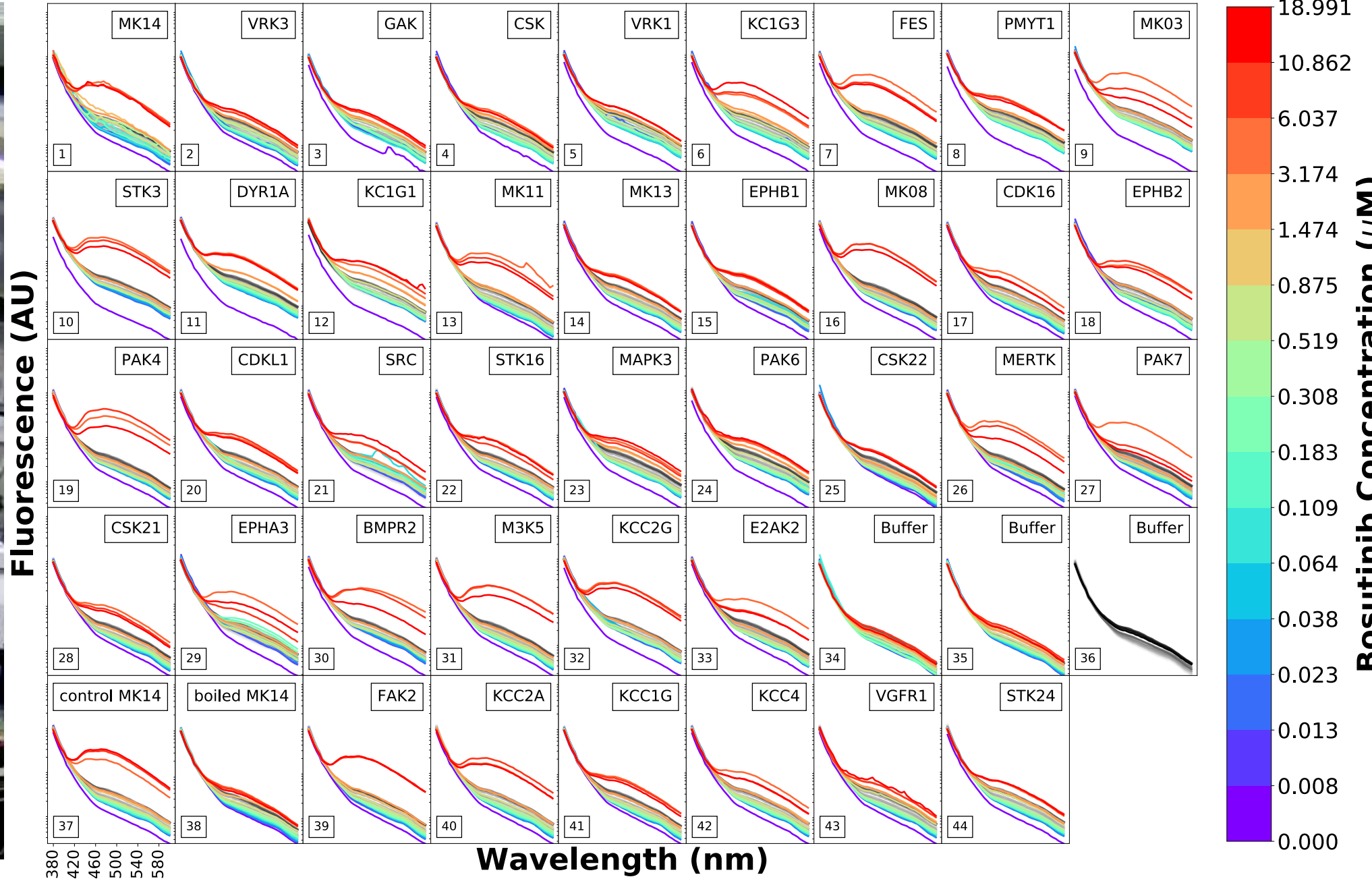
THE INTRINSIC FLUORESCENCE OF MANY FDA-APPROVED INHIBITORS ALLOWS US TO DIRECTLY MEASURE BINDING AFFINITIES TO WT AND MUTANT KINASES



other fluorescent ligands with the same scaffold



ASSAY AUTOMATION ALLOWS US TO SCALE THIS UP TO MSK-IMPACT SCALE WHILE CONTROLLING SOURCES OF ERROR



SONYA HANSON LUCELENIE RODRIGUEZ MEHTAP ISIK ERIN GRUNDY STEVEN ALBANESE



We're scaling up automated measurement of kinase inhibitor binding affinities to clinical cancer mutations in human kinase domains from **cbioportal**

INSPIRE: INTEGRATED AND SCALABLE PREDICTION OF RESISTANCE



SHANTENU JHA, RUTGERS

Massively scalable robust computational infrastructures



PETER COVENEY, UNIVERSITY COLLEGE LONDON

Parallel binding free energy calculations for drug discovery



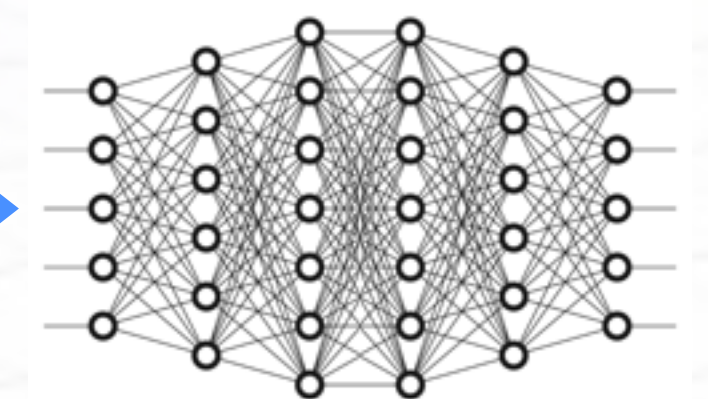
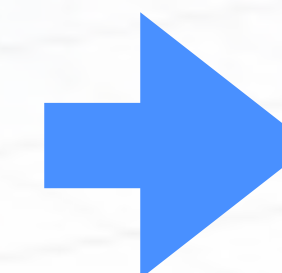
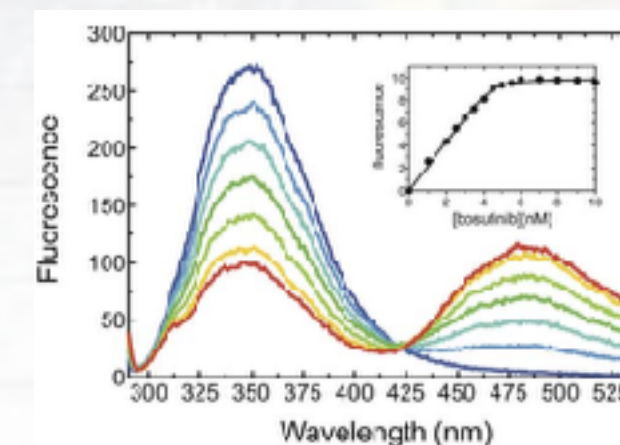
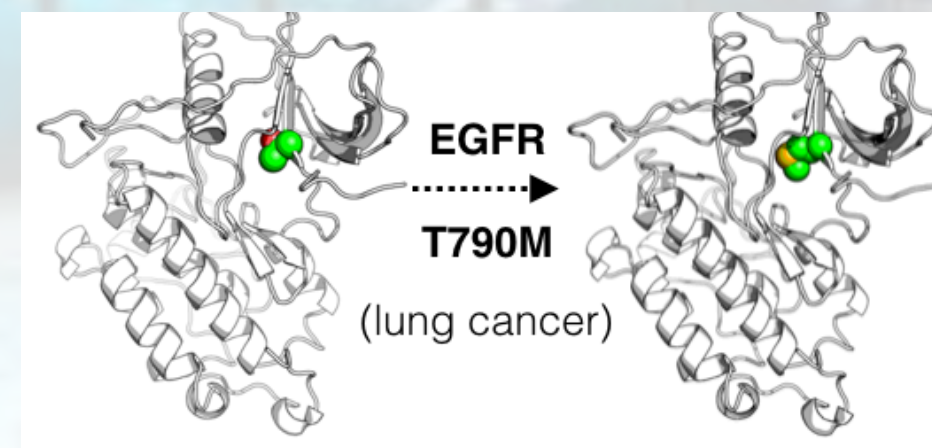
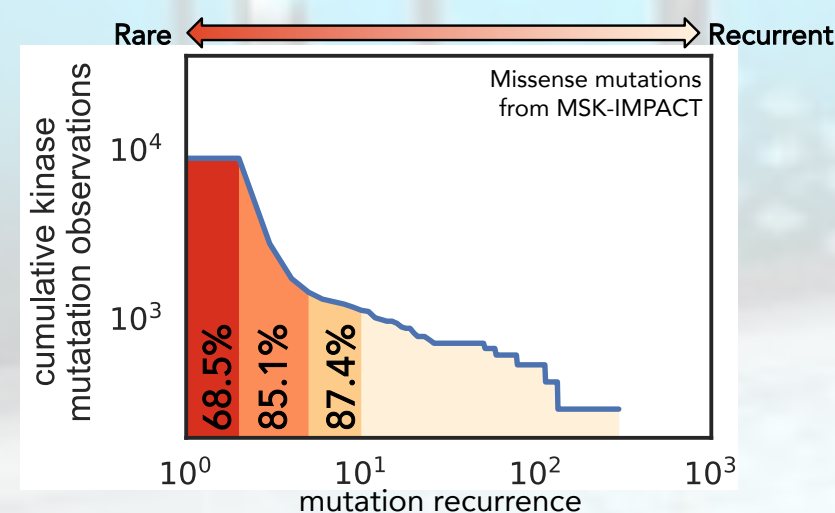
RICK STEVENS, ARGONNE NATIONAL LABORATORY

Scalable machine learning for cancer / CANDLE



JOHN CHODERA, SLOAN KETTERING INSTITUTE

Alchemical free energy calculations for resistance, automated biophysics

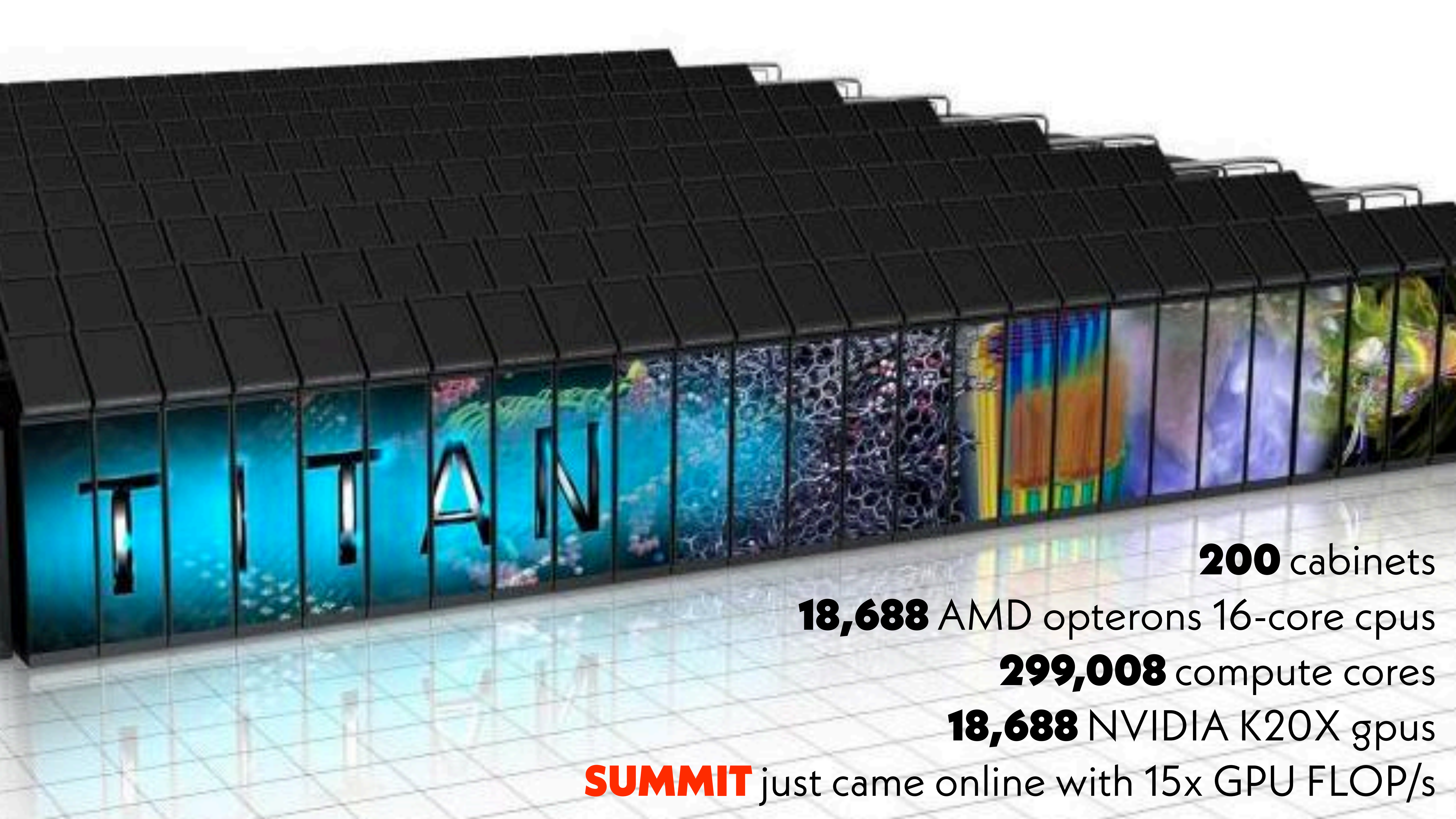


**55,000 clinical
cancer mutations**

**predict kinase inhibitor
resistance/susceptibility**

**automated biophysical
experiments**

**train machine
learning models**



200 cabinets

18,688 AMD opterons 16-core cpus

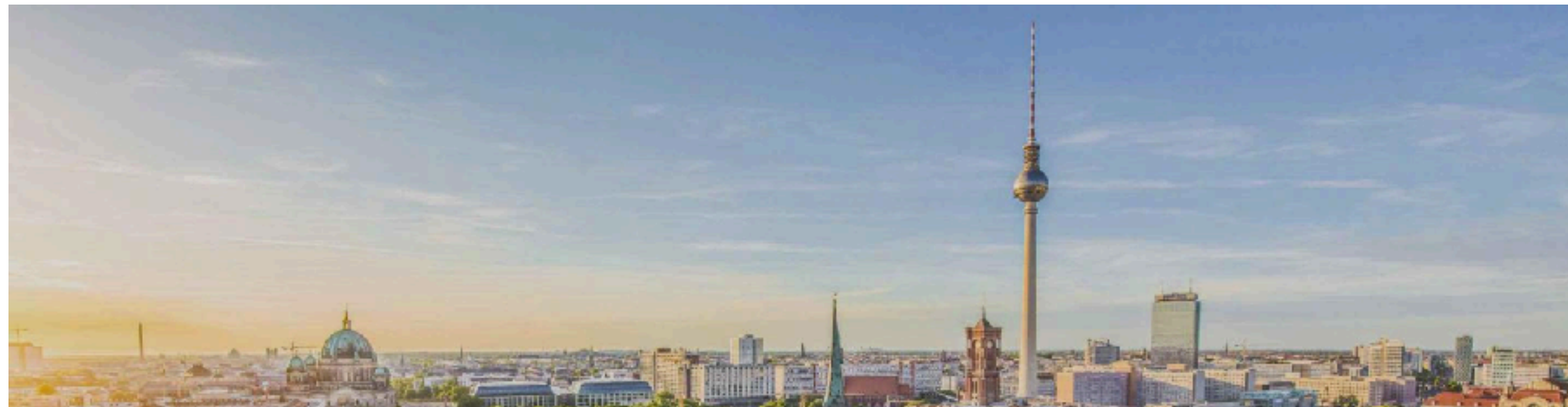
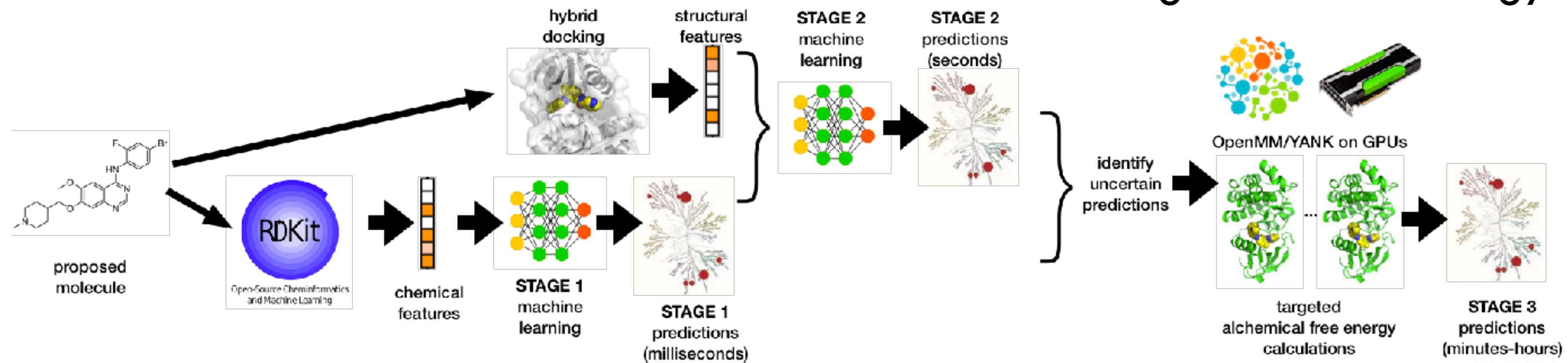
299,008 compute cores

18,688 NVIDIA K20X gpus

SUMMIT just came online with 15x GPU FLOP/s

THE FUTURE IS EXCITING

Exploring the interface of structure-informed machine learning and free energy calculations



**PROF. DR. ANDREA
VOLKAMER**

STIFTUNG CHARITÉ

PRESSEMITTEILUNG

Berlin, 24. September 2018 / AL
PM 2018-09-01

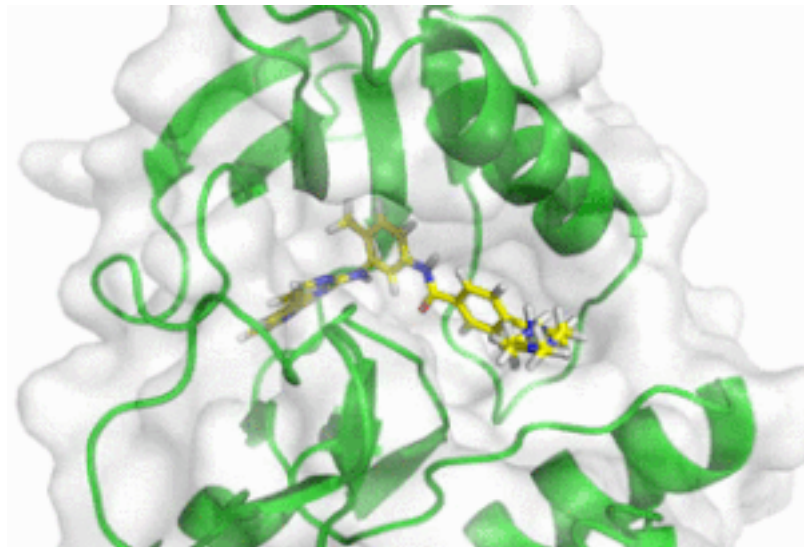
FÖRDERINITIATIVE VON JOHANNA QUANDT UMFASST NUNMEHR ÜBER 300 PERSONEN AUS DEN BERLINER LEBENSWISSENSCHAFTEN

Stiftung Charité zeichnet sieben weitere Einstein BIH Visiting Fellows aus
und richtet neues Innovationsprogramm ein

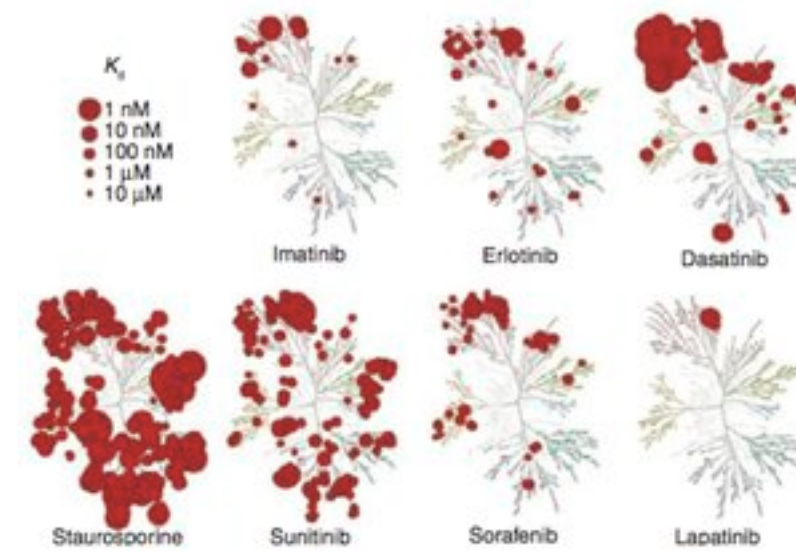
Prof. Dr. John Chodera, Memorial Sloan Kettering Cancer Center, New York (USA)
Vorhaben: Computational polypharmacology: A new paradigm for selectively promiscuous kinase inhibitors
Gastgeber/in: Prof. Dr. Andrea Volkamer & Prof. Dr. Wolfgang Kübler, Charité – Universitätsmedizin Berlin

CHODERA LAB

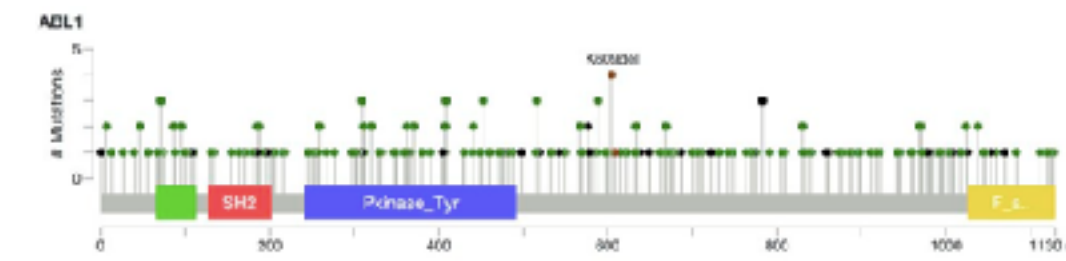
HOW CAN COMPUTATIONAL BIOPHYSICS PLAY A MAJOR ROLE IN THE ERA OF CANCER GENOMICS?



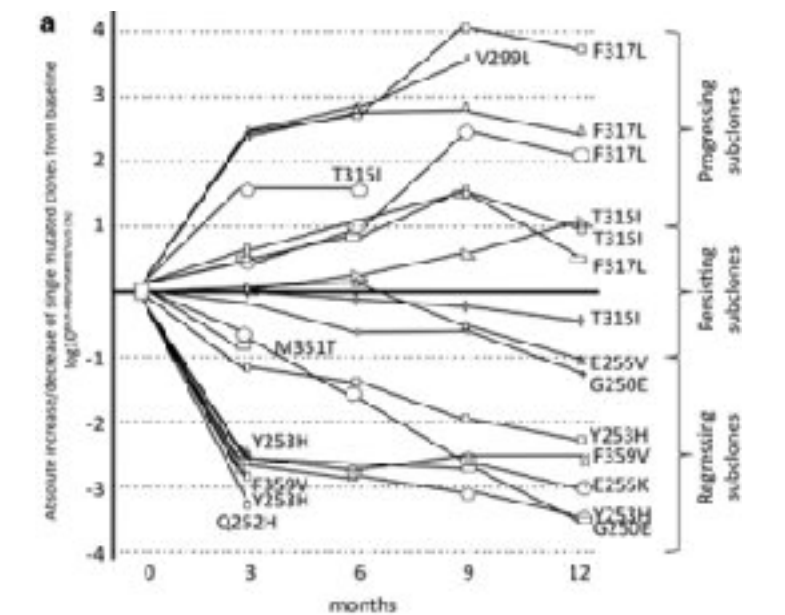
**SELECTIVE INHIBITOR DESIGN:
TARGETS/ANTITARGETS**



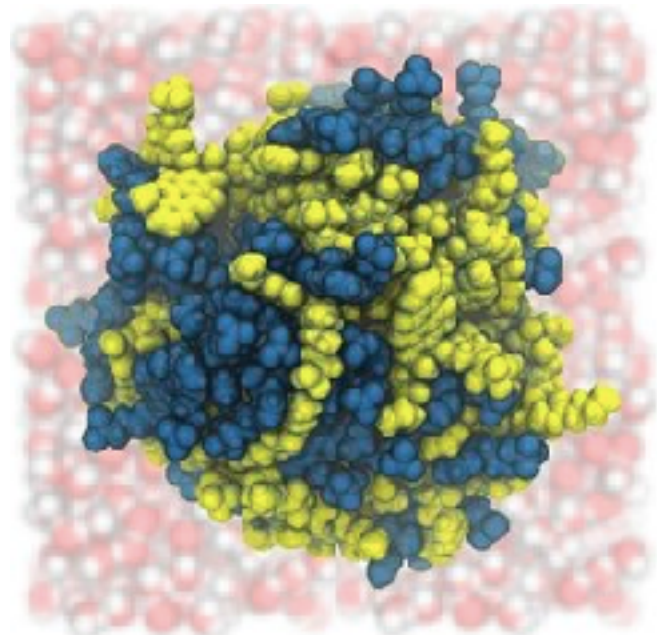
**KINASE INHIBITOR
SELECTIVITY**



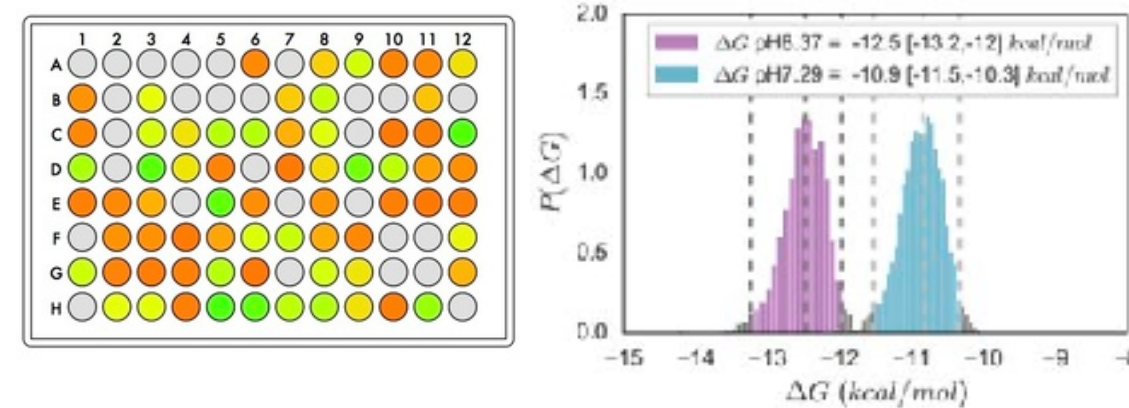
**PREDICTING DRUG
SENSITIVITY/RESISTANCE**



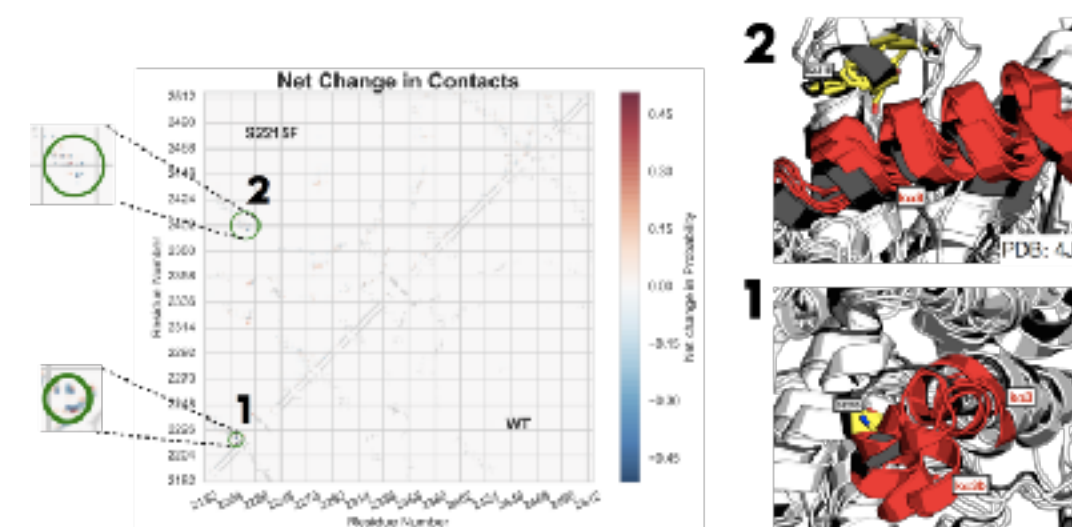
**ANTICIPATING
DRUG RESISTANCE**



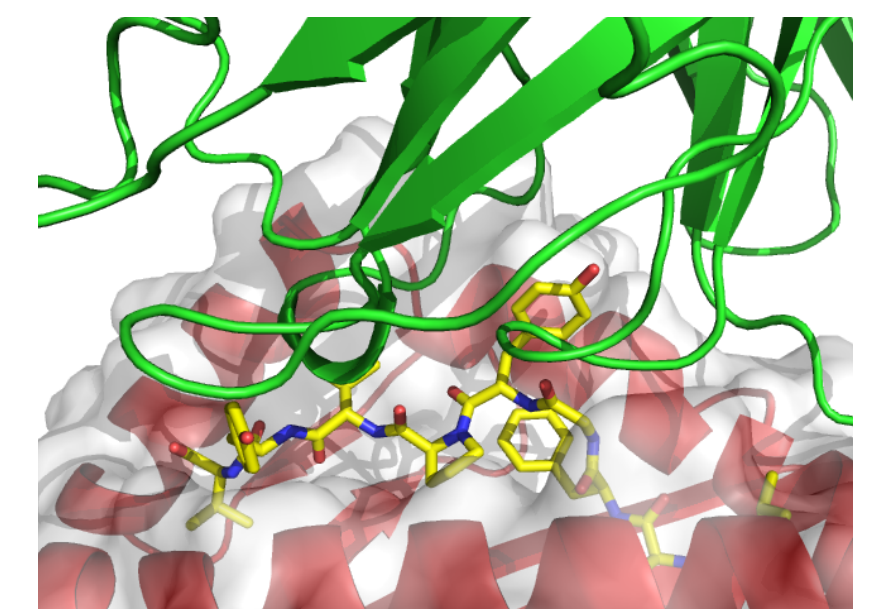
**NOVEL DRUG DELIVERY
MODALITIES**



**AUTOMATED BIOPHYSICAL
ASSAYS AND INFERENCE**



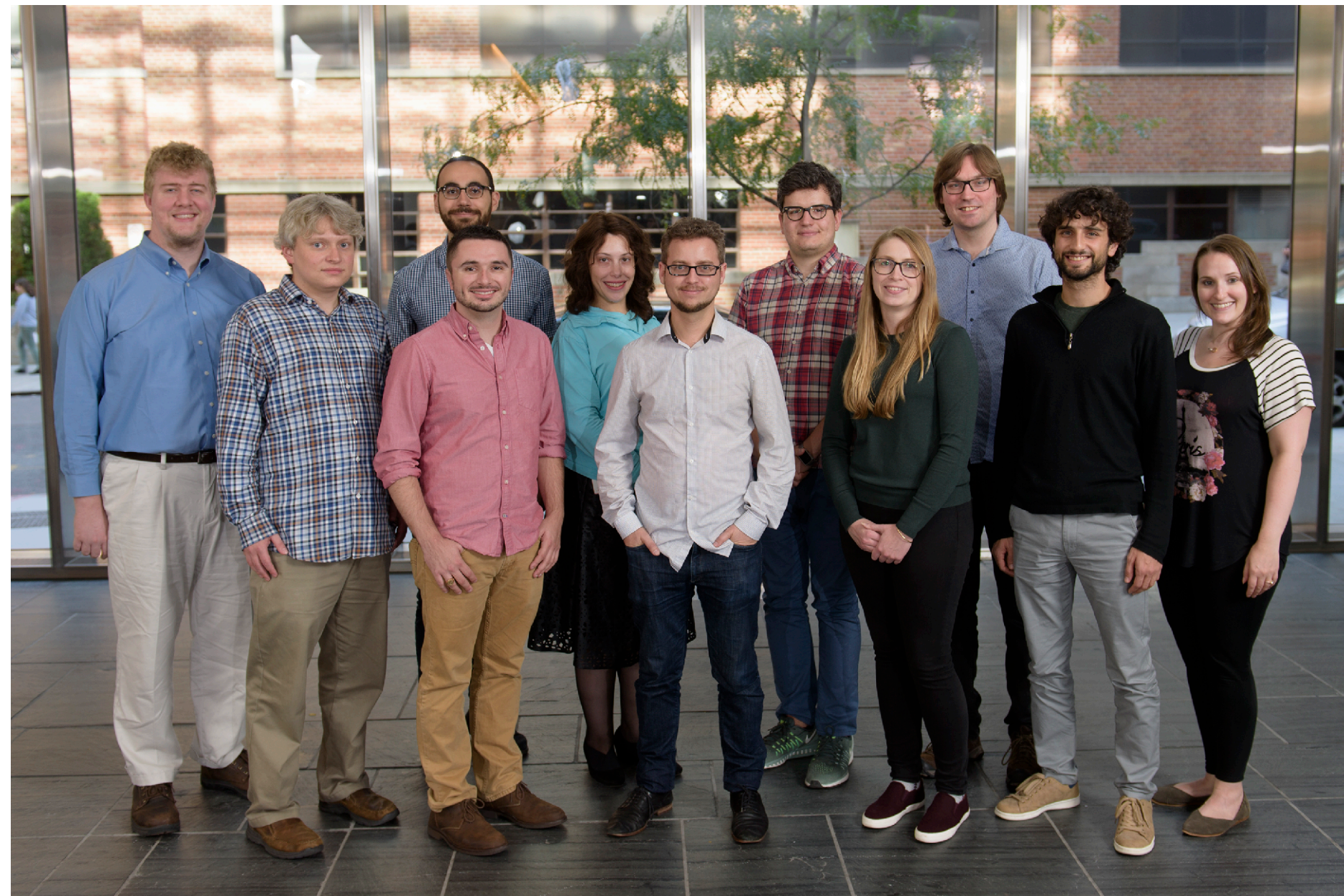
**MECHANISMS OF
ONCOGENIC ACTIVATION**



**CANCER
IMMUNOTHERAPY**

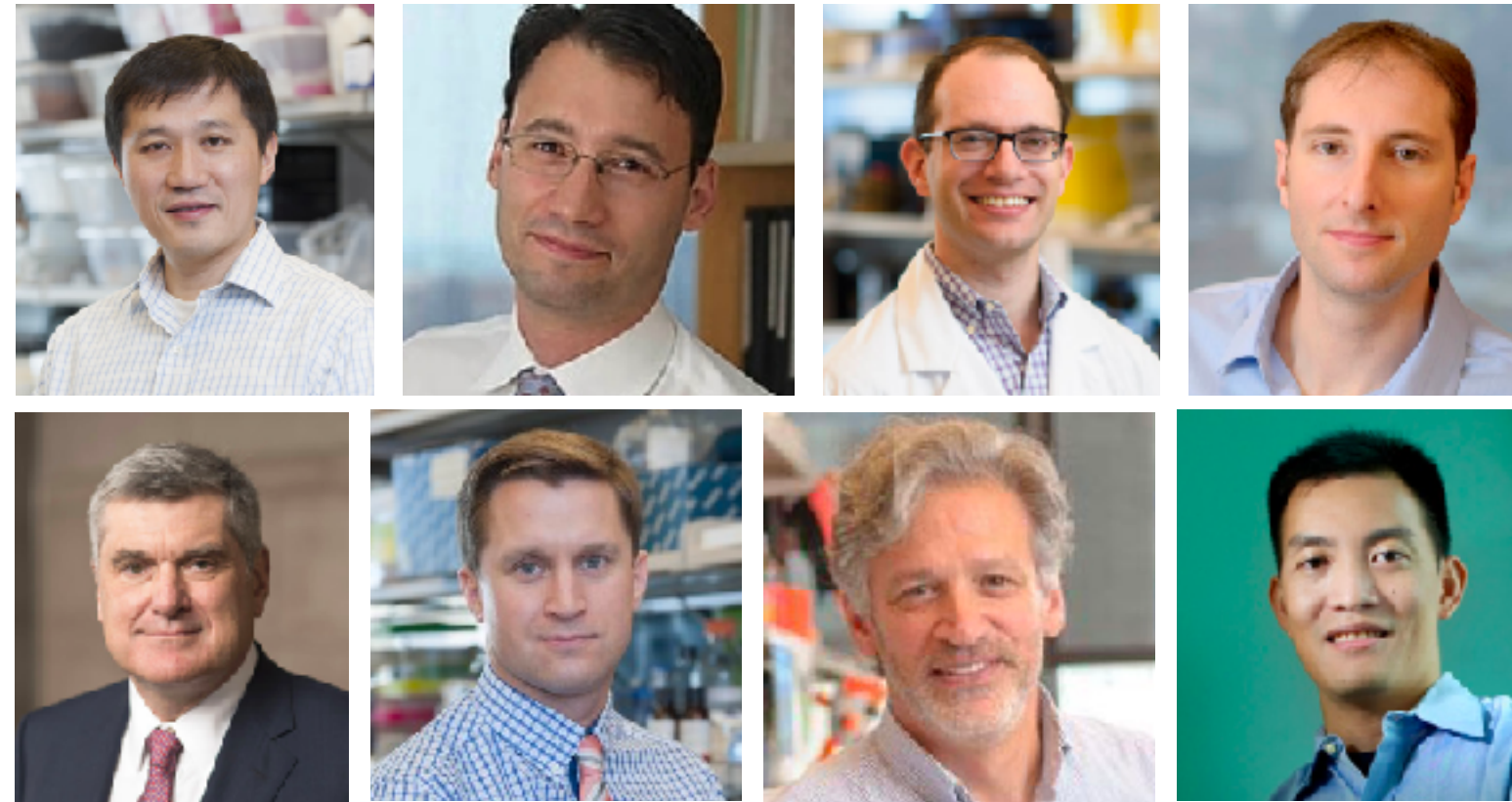
THANKS FOR AN EXCITING SIX YEARS!

CHODERA LAB



Daniel Parton	Mehtap Isik
Kadeem Ho Sang	Ashleigh Fowlkes
Patrick Grinaway	Andrea Rizzi
Sarah Boyce	Josh Fass
Sonya Hanson	Levi Naden
Jan-Hendrik Prinz	Steven Albanese
Kyle Beauchamp	Lucelenie Rodriguez
Bas Rustenburg	Ana Silveira
Chaya Stern	Jiaye Guo
Rafal Wiewiora	Hannah Bruce Macdonald
Erin Grundy	Marcus Wieder
Julie Behr	David W. H. Swenson

MSKCC COLLABORATORS



Minkui Luo
Daniel Heller
Michael Kharas
Alex Kentsis
Andy Intlekofer

Craig Thompson
Robert Benezra
Derek Tan
Richard Kolesnick
Ouathek Ouerfelli

FUNDING



FUNDING

NIH R01 GM121505	Parker Institute for Cancer Immunotherapy	
NIH R01 GM124270	Starr Foundation I8-A8-058	
NSF CHE-1738979	AstraZeneca iMED	Entasis Therapeutics
Cycle for Survival	Louis V Gerstner Young Investigator	XtalPi
Merck KGaA	BIH Einstein Visiting Fellowship	Schrödinger
Silicon Therapeutics	Open Force Field Consortium	OpenEye