

# John D. Chodera



url <http://www.choderalab.org>

email [john.chodera@choderalab.org](mailto:john.chodera@choderalab.org)

github <https://github.com/choderalab>

ORCID iD [0000-0003-0542-119X](#)

twitter [@jchodera](#)

mobile +1.415.867.7384

post 1275 York Ave, ZRC 6-South  
New York, NY 10065

## Education and positions

2019–

Associate Member, Memorial Sloan-Kettering Cancer Center

2019–

BIH Einstein Visiting Professor, Charité, Berlin

2013–

Faculty, Physiology, Biophysics, and Systems Biology Program,  
Weill Cornell Graduate School of Medical Sciences

2012–2019

Assistant Member, Memorial Sloan-Kettering Cancer Center

2008–2012

Independent Distinguished Postdoctoral Fellow, California Institute for Quantitative Biosciences (QB3),  
University of California, Berkeley

2007–2008

Independent research funding, sponsors [Phillip L. Geissler](#) and [Susan Marqusee](#)

Postdoctoral researcher, Department of Chemistry, Stanford University

1999–2006

With [Vijay S. Pande](#) (head of [Folding@Home](#) distributed computing project)

Ph.D. in Biophysics, University of California, San Francisco

1995–1999

Committee: [Ken A. Dill](#), [Matthew P. Jacobson](#), [Vijay S. Pande](#)

B.S. in Biology, California Institute of Technology

Undergraduate research with [Paul H. Patterson](#) (*molecular neurobiology*) and Jerry E. Solomon (*computational chemistry*)

## Fellowships and awards

2022–2024

Einstein Visiting Fellowship

2019–2022

BIH Einstein Visiting Fellowship

2017

Silicon Therapeutics Open Science Fellowship

2013–2016

Louis V. Gerstner Young Investigator Award

2013–2014

Google Exacycle for External Faculty

2008–2012

QB3-Berkeley Distinguished Postdoctoral Fellowship, University of California, Berkeley

2005–2006

IBM Predoctoral Fellowship

2000–2005

Howard Hughes Medical Institute Predoctoral Fellowship

## Research overview

My lab is developing the infrastructure to enable fully autonomous small molecule drug discovery powered by next-generation hybrid physical/machine learning models to predict potency, selectivity, and resistance in a data-efficient manner. We develop advanced predictive computational methodologies for drug discovery in frontier areas where these models are lacking, working with industry collaborators to deploy those solutions in real drug discovery programs. We build open source tools—such as [OpenMM](#), which has been downloaded over 1.2 million times, run on millions of machines, and powers advances such as AlphaFold—that enable these algorithms to have impact across the drug discovery and biomolecular modeling fields. We create open science communities—such as (1) the [Open Force Field Initiative](#), (2) the [COVID Moonshot](#) (which delivered a patent-free COVID antiviral preclinical candidate into an IND-enabling preclinical program funded by the Wellcome Trust), (3) the [Folding@home Consortium](#) which harnesses hundreds of thousands of volunteer computers around the world to accelerate biomolecular simulation and discovery, and (4) the [AI-driven Structure-enabled Antiviral Program](#) that received \$68M in initial funding from the NIH for open science antiviral drug discovery. We apply these tools to cancer, where we develop models for (1) designing small molecule kinase inhibitors with targeted polypharmacology, (2) predicting drug sensitivity or resistance of clinical mutations, and (3) understanding the detailed structural mechanisms underlying oncogenic mutations. And we train the next generation of interdisciplinary scientists equipped to tackle the most challenging problems in health and human disease for industrial drug discovery.

## Major Research Collaborations



### The COVID Moonshot (2020–) · <https://covid.postera.ai/covid>

I co-founded the COVID Moonshot, a global open science collaboration leveraging our free alchemical energy calculations on Folding@home to discovery a patent-free oral antiviral useful against COVID-19. Starting from an X-ray fragment screen, we produced SARS-CoV-2 main viral protease (Mpro) inhibitor preclinical candidates in just 18 months (*Nature Comment*). Together with the Drugs for Neglected Diseases initiative (DNDi)—the world's leading nonprofit developing new therapies for neglected diseases—we were awarded \$11M from the WHO Access to COVID Tools Accelerator (ACT-A) via the Wellcome Trust to carry out an accelerated preclinical program, with clinical trials expected in 2023. Our antiviral is expected to be free of the drug-drug interactions that present difficulties to widespread use of Paxlovid, and has differentiated resistance liabilities. We are working to partner with generics manufacturers to make our antiviral available for globally equitable and affordable access.



### AI-driven Structure-enabled Antiviral Platform (ASAP) (2022–) · <https://asapdiscovery.org>

Together with DNDi and PostEra, I lead the AI-driven Structure-enabled Antiviral Platform, which uses artificial intelligence and computational chemistry to accelerate structure-based open science antiviral drug discovery and deliver oral antivirals for pandemics with the goal of global, equitable, and affordable access. I am the Contact PI on the NIH NIAID Antiviral Drug Discover (AVIDD) Center that currently funds ASAP, currently funded by a \$68M NIH award for the first three-year phase of a five-year \$110M grant.



### Folding@home Consortium (2012–) · <http://foldingathome.org>

I am a founding member of the Folding@home Consortium, a worldwide network of laboratories that run the Folding@home global distributed computing project where hundreds of thousands of volunteers around the world (mostly gamers with GPUs) contribute idle computing cycles to enable us to study molecular mechanisms of disease and search for new potential therapeutics. As part of the COVID Moonshot effort, Folding@home became the world's first exascale distributed computing platform in 2020. Using our alchemical free energy methodologies on Folding@home, we accelerated lead optimization for preclinical candidates for the COVID Moonshot that are expected to enter clinical trials in 2023.



### OpenMM molecular simulation framework (2006–) · <http://openmm.org>

I co-lead the OpenMM Consortium, which develops an open source GPU-accelerated molecular simulation package used both as a stand-alone tool for biomolecular simulation and as a library to enable other tools to use biomolecular simulation, such as Isolde and AlphaFold2. OpenMM is the fastest and most widely deployed GPU-accelerated biomolecular simulation package globally, and has been downloaded over 1.2 million times. OpenMM is currently funded by the NIH to effect a transformation in the biomolecular simulation community to enable the use of fast, accurate machine-learning force fields.



### Open Force Field Initiative and Consortium (2019–) · <http://openforcefield.org>

I co-founded the Open Force Field Initiative (an NIH-funded collaboration) and related Open Force Field Consortium (funded by over a dozen industry partners), a scientific collaboration consisting of dozens of scientists around the world working to develop modern open source infrastructure for building and applying high-quality biomolecular force fields for accelerating drug discovery. These force fields and tools are used by pharma and academic labs around the world, and recently aided the discovery of a new preclinical candidate by the COVID Moonshot.



### Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) (2012–) · <https://samplerchallenges.github.io>

I co-organize the SAMPL Challenges, which are iterative blind predictive modeling challenges for the computational chemistry community aimed at accelerating the generation of solutions to major obstacles limiting the progress of computer-aided drug discovery, focusing the community on solving these challenges, providing critical experimental data to computational researchers that lack access to data generation capabilities, and emphasizing knowledge sharing. The learnings from these challenges have been reported in over 14 special issues of the Journal for Computer-Aided Molecular Design. In some areas, we have managed to move computational methods for wide disagreement among methods to tightly-grouped predictions that can identify when the experimental data is wrong.



### Critical Assessment of Computational Hit-finding Experiments (CACHE) (2022–) · <https://cache-challenge.org/>

I co-founded the CACHE Challenge, which aims to be the "CASP for hit-finding in drug discovery" by providing a new generation of AI- and physics-driven approaches to assess and refine their hit identification approaches against experimental data. Each CACHE Challenge focuses on identifying new chemical matter against an orphan target protein of pharmaceutical relevance, and all experimental data and compounds are disclosed into the public domain to accelerate drug discovery against these targets.



### Open Molecular Software Foundation (OMSF) (2021–) · <http://openforcefield.org>

I work closely with the Open Molecular Software Foundation (OMSF), a 501(c)3 nonprofit organization that aims to be thoughtful stewards for the biomolecular modeling and computer-aided drug discovery open source software communities. Using a funding mechanism I helped pioneer, OMSF can act as a fiscal sponsor that works with open source software communities in the biomolecular sciences to aid them in becoming self-governing organizations that serve the community with fiscally sustainable support from industrial, federal, and philanthropic sources. OMSF currently hosts the Open Force Field Initiative and Consortium, the Open Free Energy Consortium, and OpenFold Consortium.

## Selected Significant Publications

Underline denotes Chodera lab trainee; ‡ denotes co-corresponding authorship; **Chodera JD** highlighted for convenience.

### Selected publications and preprints following promotion to Associate Member



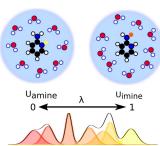
Preprint · [bioRxiv](#)

The COVID Moonshot Consortium, **Chodera JD** ‡, Lee A‡, London N‡, von Delft F‡

#### Open science discovery of oral non-covalent SARS-CoV-2 main protease inhibitors

We report the discovery of a new oral antiviral non-covalent SARS-CoV-2 main protease inhibitor developed by the COVID Moonshot, a global open science collaboration leveraging free energy calculations on Folding@home and ML-accelerated synthesis planning, now in accelerated preclinical studies funded by an \$11M grant from the WHO ACT-A program via the Wellcome Trust. We are currently in discussions with generics manufacturers about partnering with us throughout clinical trials (expected to begin in 2023) to ensure we can scale up production for global equitable and affordable access once approved by regulatory agencies.

Contributions: My lab led computer-aided drug discovery aspects of this project, and performed tens of thousands of free energy calculations to prioritize which compounds were synthesized.

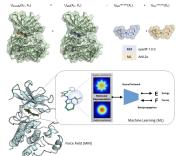


Preprint · [bioRxiv](#)

Wieder M, Fass J, and **Chodera JD**

#### Teaching free energy calculations to learn from experimental data

In this pioneering work, we show, for the first time, how we can use machine learning to systematically improve explicit solvent alchemical free energy calculations using fully machine learning force fields by fitting directly to experimental free energies. This is also the first work to show how to compute alchemical free energies in explicit solvent using fully machine learning force fields, and the first to show how this can also involve breaking and forming covalent bonds.



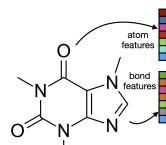
Preprint · [bioRxiv](#)

Rufa DA, Bruce Macdonald, HE, Fass J, Wieder M, Grinaway PB, Roitberg AE, Isayev O, and **Chodera JD**

#### Towards chemical accuracy for alchemical free energy calculations with hybrid physics-based machine learning / molecular mechanics potentials

In this first use of hybrid machine learning / molecular mechanics (ML/MM) potentials for alchemical free energy calculations, we demonstrate how the improved modeling of intramolecular ligand energetics offered by the quantum machine learning potential ANI-2x can significantly improve the accuracy in predicting kinase inhibitor binding free energy by reducing the error from 0.97 kcal/mol to 0.47 kcal/mol, which could drastically reduce the number of compounds that must be synthesized in lead optimization campaigns for minimal additional computational cost.

Contributions: My lab led all work, consulting Adrian Roitberg and Olexandr Isayev on appropriate use of the ANI potentials in alchemical free energy calculations.



Chemical Science, in press · [DOI](#)

Wang Y, Fass J, and **Chodera JD**

#### End-to-end differentiable molecular mechanics force field construction

Molecular mechanics force fields have been a workhorse for computational chemistry and drug discovery. Here, we propose a new approach to force field parameterization in which graph convolutional networks are used to perceive chemical environments and assign molecular mechanics (MM) force field parameters. The entire process of chemical perception and parameter assignment is differentiable end-to-end with respect to model parameters, allowing new force fields to be easily constructed from MM or QM force fields, extended, and applied to arbitrary biomolecules.



Proceedings of the National Academy of Sciences 118:e2111451118 · [DOI](#)

Lyczek A, Tilman-Berger B, Rangwala AM, Paung Y, Tom J, Philipose H, Guo J, Albanese SK, Robers MB, Knapp S,

**Chodera JD**, Seeliger MA

#### Mutation in Abl kinase with altered drug binding kinetics indicates a novel mechanism of imatinib resistance

Here, we characterize the biophysical mechanisms underlying mutants of Abl kinase associated with clinical drug resistance to targeted cancer therapies. We uncover a surprising novel mechanism of mutational resistance to kinase inhibitor therapy in which the off-rate for inhibitor unbinding is increased without affecting inhibitor affinity.

Contributions: My lab curated the clinical cancer mutations, coordinated the production of initial protein constructs, and helped elucidate the molecular mechanism underlying resistance.

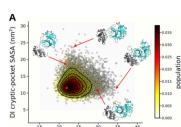
Nature Chemistry 13:651, 2021 · [DOI](#)

Zimmerman MI, Porter JR, Ward MD, Singh S, Vithani N, Meller A, Mallimadugula UL, Kuhn CE, Borowsky JH, Wiewiora RP, Hurley MFD, Harbison AM, Fogarty CA, Coffland JE, Fadda E, Voelz VA, **Chodera JD**, Bowman GR

#### SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome

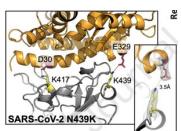
To accelerate a multitude of drug development activities to combat the global threat posed by COVID-19, over a million citizen scientists have banded together through the Folding@home distributed computing project to create the world's first Exascale computer and simulate protein dynamics. An unprecedented 0.1 seconds of simulation of the viral proteome reveal how the spike complex uses conformational masking to evade an immune response, conformational changes implicated in the function of other viral proteins, and cryptic pockets that are absent in experimental structures. These structures and mechanistic insights present new targets for the design of therapeutics.

Contributions: My lab contributed key software infrastructure to power this effort, and led the initial drive for volunteers to join our effort to study COVID.



Cell 184:1171, 2021 · DOI

Thompson EC, Rosen LE, Shepherd JG, Spreafico R, da Silva Filipe A, Wojcechowsky JA, Davis C, Piccoli L, Pascall DJ, Dillen J, Lytras S, Czudnochowski N, Shah R, Meury M, Jesudason N, De Marco A, Li K, Bassi J, O'Toole A, Pinto D, Colquhoun RM, Culap K, Jackson B, Zatta F, Rambaut A, Jaconi S, Sreenu VB, Nix J, Zhang I, Jarrett RF, Glass WG, Beltramello M, Nomikou K, Pizzuto M, Tong L, Cameroni E, Cross TI, Johnson N, Di Julio J, Wickenhagen A, Ceschi A, Harbison AM, Mair D, Ferrari P, Smollett K, Sallusto F, Carmichael S, Garzoni C, Nichols J, Galli M, Hughes J, Riva A, Ho A, Schiuma M, Semple MG, Openshaw PJM, Fadda E, Baillie JK, Chodera JD, Rihm SJ, Lycet SJ, Virgin HW, Telenti A, Corti D, Robertson DL, Snell J, ISARIC4C Investigators



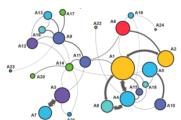
### Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity

New mutations that enhance the affinity of SARS-CoV-2 spike protein for human ACE2—and potentially pose threats to antibody-based therapeutics and vaccines for COVID-19—are already emerging in the wild. We characterize and describe sentinel mutations of SARS-CoV-2 in the wild that herald challenges for combatting COVID-19, and use simulations of the RBD-ACE2 interface on Folding@home to biophysically characterize why these mutations can lead to enhanced affinity.

Contributions: My lab led all the molecular simulation work reported in this paper.

eLife 8:e45403, 2019 · DOI

Rafal P. Wiewiora<sup>†</sup>, Shi Chen<sup>†</sup>, Fanwang Meng, Nicolas Babault, Anqi Ma, Wenyu Yu, Kun Qian, Hao Hu, Hua Zou, Junyi Wang, Shijie Fan, Gil Blum, Fabio Pittella-Silva, Kyle A. Beauchamp, Wolfram Tempel, Hualing Jiang, Kaixian Chen, Robert Skene, Y. George Zheng, Peter J. Brown, Jian Jin, Chodera JD<sup>‡</sup>, and Minkui Luo<sup>‡</sup>



### The dynamic conformational landscapes of the protein methyltransferase SETD8

In this work, we show how targeted X-ray crystallography using covalent inhibitors and depletion of native ligands to reveal structures of low-population hidden conformations can be combined with massively distributed molecular simulation to resolve the functional dynamic landscape of the protein methyltransferase SETD8 in unprecedented atomistic detail. Using an aggregate of six milliseconds of fully atomistic simulation from Folding@home, we use Markov state models to illuminate the conformational dynamics of this important epigenetic protein.

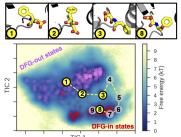
Contributions: My lab led all the molecular simulation work reported in this paper.

### Selected publications prior to promotion to Associate Member

Cell Chemical Biology 26:390, 2019 · DOI

Hanson SM\*, Georghiou G\*, Miller WT, Rest JS, Chodera JD<sup>‡</sup>, and Seeliger MA<sup>‡</sup>

#### What makes a kinase promiscuous for inhibitors?



Using a combination of chemogenomics, structural biology, and molecular simulation approaches, we identify a set of human kinases that are especially promiscuous binders of small molecule kinase inhibitors, and show that a prototypical member of this class, DDR1, achieves this promiscuity by virtue of its more stable Asp-DFG-out conformation. Contributions: My lab led all the molecular simulation work reported in this paper.

Communications Biology 1:70, 2018 · DOI

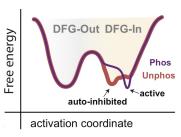
Hauser K, Negron C, Albanese SK, Ray S, Steinbrecher T, Abel R, and Chodera JD, and Wang L

#### Predicting resistance of clinical Abl mutations to targeted kinase inhibitors using alchemical free-energy calculations

The emergence of resistance mutations are a significant problem in precision cancer therapy. With cancer centers now routinely collecting tumor sequencing panels, it is apparent the vast majority of missense mutations in kinases—for which more than 36 small molecule inhibitors are available—are very rare, and no information will be available as to whether these mutations might reduce therapeutic efficacy for some inhibitors by inducing drug resistance. Here, we take the first steps toward using physical modeling to predict individualized therapeutic response, assessing how accurately the impact of clinical Abl mutations on inhibitor binding affinity can be predicted. Our results suggest free energy calculations are a promising tool for aiding therapeutic decisionmaking. Contributions: My lab conceived of this study, design the experiment, helped analyze the data, and helped write the paper.

eLife 7:e32766, 2018 · DOI

Ruff EF, Muretta JM, Thompson A, Lake E, Cyphers S, Albanese SK, Hanson SM, Behr JM, Thomas DT, Chodera JD, and Levinson NM



#### A dynamic mechanism for allosteric activation of Aurora kinase A by activation loop phosphorylation

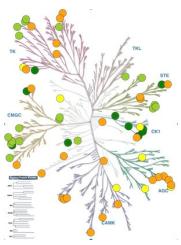
Kinase phosphorylation is canonically believed to induce a conformational change of the DFG loop that borders the ATP binding site from out to in. Surprisingly, it seems that not all kinases behave this way. Through a combination of FRET, IR, and EPR labeling and large-scale molecular dynamics simulations, we show that phosphorylation activates Aurora kinase by a novel mechanism that does not simply correspond to a DFG-out to DFG-in population shift, but rather reorganization of DFG-in subpopulations. This paper follows our prior report in Nature Chemical Biology on the role of Tpx2 in activating a cellular subpopulation of Aurora by another surprising mechanism.

Contributions: My lab led all the molecular simulation work reported in this paper.

Biochemistry 57:4675, 2018 · DOI

Albanese SK\*, Parton DL\*, Isik M†, Rodríguez-Laureano L†, Hanson SM, Gradia S, Jeans C, Levinson NM, Seeliger M, and Chodera JD

#### An open library of human kinase domain constructs for automated bacterial expression



To establish a tractable experimental system for studying the biophysical determinants of selective kinase inhibitor resistance in clinical cancer mutations, we engineer a library of human kinase domains with useful bacterial expression with phosphatase coexpression. Over 350 requests have been made for these plasmids.

Contributions: My lab designed the experiments, directed cloning and expression at the QB3 MacroLab (Gradia and Jeans), and performed the characterization experiments in our laboratory (with guidance from Levinson and Seliger).

## Scientific Advisory Board Memberships

One way I work to translate my research into impact is via participating in Scientific Advisory Boards (SABs) for companies using computational chemistry to accelerate drug discovery. Some of the SABs I have served on include:



### Schrodinger (2013–2018) · <http://schrodinger.com>

I joined the Schrodinger SAB shortly after their first release of the FEP+ alchemical free energy calculation tool, aiming to ensure the first commercial turnkey product version of my field of alchemical free energy calculations had maximum impact for structure-enabled drug discovery. This effort was highly successful, with FEP+ near-universally used by industry to accelerate structure-enabled drug discovery efforts. I left in 2018 to focus my efforts on new, more scalable technologies and ways to use these tools for frontier problems in drug discovery.



### OpenEye Scientific (2018–) · <https://www.eyesopen.com>

I joined the OpenEye SAB excited to help enable an ecosystem of open source computer-aided drug discovery tools and academic research that could rapidly be translated into impactful deployment within both industry and academia using their Orion cloud-based drug discovery platform. OpenEye was recently acquired by Cadence for \$500M.



### Interline Therapeutics (2021–) · <https://www.interlinetx.com>

I am a founding SAB member and scientific cofounder of Interline Therapeutics, which has raised \$92M since its inception in 2021. Interline seeks to harness open source alchemical free energy methodologies that I developed to tackle problems in target and conformational selectivity to enable the design of small molecules that modulate protein communities.



### Ventus Therapeutics (2021–) · <https://www.ventustx.com>

Ventus is a rapidly growing small molecule discovery company that has raised over \$300M in three funding rounds. In addition to featuring extraordinary medicinal chemistry talent, Ventus has developed a novel computational platform, ReSOLVE, to accelerate novel chemotype hit discovery for new targets. I joined the SAB to advise on the integration of structure-enabled open source computer-aided drug discovery workflows to accelerate hit-to-candidate progression.



### Redesign Science (2020–) · <https://www.redesignscience.com>

Redesign Science is a NYC-based startup aiming to commercialize technologies I helped develop as a graduate student—the construction of Markov state models of coarse-grained biomolecular dynamics from distributed biomolecular simulations—to use as a platform to accelerate drug discovery.

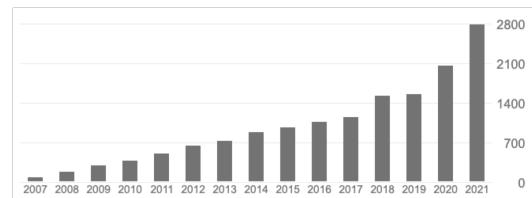
## Open Source Software

All software developed by our lab can be found at <http://github.com/choderalab>

All software is licensed under permissive OSI-approved open source licenses and available on [GitHub](#)

## All Publications

All publications: <http://choderalab.org/publications>  
Google Scholar statistics: <http://goo.gl/qO0JW>  
MyNCBI Bibliography: <http://goo.gl/e3kjgK>  
bioRxiv preprints: <https://bit.ly/2LHGwxJ>  
arXiv preprints: <https://bit.ly/2Rp2Y3T>  
chemRxiv preprints: <https://chemrxiv.org/search?q=chodera>  
Open source software: <http://github.com/choderalab>  
Major datasets: <http://choderalab.org/data>  
*h*-index: 56 / i10-index: 108 / citations: 17464 (31 Oct 2022)  
\* denotes co-first-authors  
† denotes co-second-authors  
‡ denotes co-corresponding authors



annual citation counts via Google Scholar (31 Dec 2021)

## Submitted, Under Review, and Preprints

Underline denotes Chodera lab trainee; ‡ denotes co-corresponding authorship; **Chodera JD** highlighted for convenience.

### Open science antiviral discovery



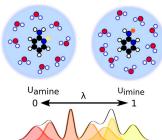
Preprint · [bioRxiv](#)

The COVID Moonshot Consortium, **Chodera JD** ‡, Lee A‡, London N‡, von Delft F‡

#### Open science discovery of oral non-covalent SARS-CoV-2 main protease inhibitors

We report the discovery of a new oral antiviral non-covalent SARS-CoV-2 main protease inhibitor developed by the COVID Moonshot, a global open science collaboration leveraging free energy calculations on Folding@home and ML-accelerated synthesis planning, now in accelerated preclinical studies funded by an \$11M grant from the WHO ACT-A program via the Wellcome Trust. We are currently in discussions with generics manufacturers about partnering with us throughout clinical trials (expected to begin in 2023) to ensure we can scale up production for global equitable and affordable access once approved by regulatory agencies.

### Introducing learnability into biomolecular modeling for drug discovery

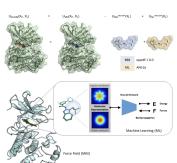


Preprint · [bioRxiv](#)

Wieder M, Fass J, and **Chodera JD**

#### Teaching free energy calculations to learn from experimental data

In this pioneering work, we show, for the first time, how we can use machine learning to systematically improve explicit solvent alchemical free energy calculations using fully machine learning force fields by fitting directly to experimental free energies. This is also the first work to show how to compute alchemical free energies in explicit solvent using fully machine learning force fields, and the first to show how this can also involve breaking and forming covalent bonds..

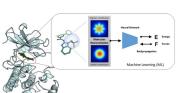


Preprint · [bioRxiv](#)

Rufa DA, Bruce Macdonald HE, Fass J, Wieder M, Grinaway PB, Roitberg AE, Isayev O, and **Chodera JD**

#### Towards chemical accuracy for alchemical free energy calculations with hybrid physics-based machine learning / molecular mechanics potentials

In this first use of hybrid machine learning / molecular mechanics (ML/MM) potentials for alchemical free energy calculations, we demonstrate how the improved modeling of intramolecular ligand energetics offered by the quantum machine learning potential ANI-2x can significantly improve the accuracy in predicting kinase inhibitor binding free energy by reducing the error from 0.97 kcal/mol to 0.47 kcal/mol, which could drastically reduce the number of compounds that must be synthesized in lead optimization campaigns for minimal additional computational cost.



Preprint · [arXiv](#)

Galvelis R, Varela-Rial A, Doerr S, Fino R, Eastman P, Markland TE, **Chodera JD**, and de Fabritiis G

#### NNP/MM: Fast molecular dynamics simulations with machine learning potentials and molecular mechanics

We demonstrate that a new generation of quantum machine learning (QML) potentials based on neural networks—which can achieve quantum chemical accuracy at a fraction of the cost—can be implemented efficiently in the OpenMM molecular dynamics simulation engine as part of hybrid machine learning / molecular mechanics (ML/MM) potentials that promise to deliver superior accuracy for modeling protein-ligand interactions.

## Force field parameterization



Preprint · [chemRxiv](#)

Boothroyd S, Behara PK, Madin OC, Hahn DF, Jang H, Gapsys V, Wagner JR, Horton JT, Dotson DL, Thompson MW, Maat J, Gokey T, Wang L-P, Cole DJ, Gilson MK<sup>‡</sup>, Chodera JD<sup>‡</sup>, Bayly CI, Shirts MR<sup>‡</sup>, Mobley DL<sup>‡</sup>

### Development and benchmarking of Open Force Field 2.0.0—the Sage small molecule force field

We present a new generation of small molecule force field for molecular design from the Open Force Field Initiative fit to both quantum chemical and experimental liquid mixture data.

Preprint · [bioRxiv](#)

Stern CD, Bayly CI, Smith DGA, Fass J, Wang L-P, Mobley DL, and Chodera JD

### Capturing non-local through-bond effects when fragmenting molecules for quantum chemical torsion scans

We show how the Wiberg Bond Order (WBO) can be used to construct small molecule fragmentation schemes that will avoid disrupting the chemical environment around torsions. The resulting fragmentation scheme powers the QCSubmit tool used to fragment and inject small molecule datasets into the QCFractal computation pipeline for deposition into the QCArchive quantum chemistry archive the Open Force Field Initiative uses for constructing force fields, as well as powering bespoke torsion refitting for individual molecules.



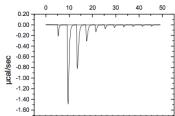
Preprint · [arXiv](#)

Wang Y, Fass J, Stern CD, and Chodera JD

### Graph nets for partial charge prediction

Graph convolutional and message-passing networks can be a powerful tool for predicting physical properties of small molecules when coupled to a simple physical model that encodes the relevant invariances. Here, we show the ability of graph nets to predict partial atomic charges for use in molecular dynamics simulations and physical docking.

## Experimental best practices

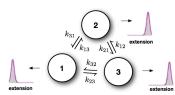


Preprint · [bioRxiv](#)

Boyce SE, Tellinghuisen JT, and Chodera JD

### Avoiding accuracy-limiting pitfalls in the study of protein-ligand interactions with isothermal titration calorimetry

We demonstrate how to avoid accuracy-limiting problems in standard isothermal calorimetry experiments as well as capture the primary sources of uncertainty in thermodynamic parameters.

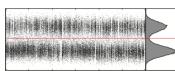


Preprint · [arXiv](#)

Chodera JD, Noé F, Hinrichs NS, Keller B, Elms PJ, Kaiser CM, Ewall-Wice A, Marqusee S, and Bustamante C

### Bayesian hidden Markov model analysis of single-molecule biophysical experiments

We present a Bayesian hidden Markov model analysis scheme that allows biomolecular conformational dynamics—and the corresponding uncertainty due to limited data—to be inferred from single-molecule trajectories. This approach was developed for a single-molecule study examining folding dynamics of nascent proteins exiting the ribosome [Science 334:1723, 2011 · DOI].



Preprint · [arXiv](#)

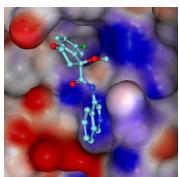
Chodera JD, Elms PJ, Swope WC, Prinz J-H, Marqusee S, Bustamante C, Noé F, and Pande VS

### A robust approach to estimating rates from time-correlation functions

We present a simple, robust approach to estimating two-state rate constants from experimental or simulation data.

## Reviews and Commentaries

### Open science and globally equitable drug discovery



Nature 594:330, 2021 · DOI

von Delft F, Calmiano M, Chodera JD, Griffen E, Lee A, London N, Matviuk T, Perry B, Robinson M, and von Delft A

#### A white-knuckle ride of open COVID drug discovery

The COVID Moonshot is an open science effort to discover a direct-acting SARS-CoV-2 oral antiviral. Here, we share lessons from this effort, including the missed opportunity to develop a phase 2 ready drug more than a decade ago that could have halted the COVID-19 pandemic in its tracks.



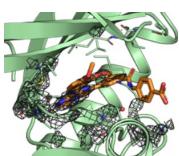
Knowable Magazine 09.27.2021 · web

Lee A, Chodera JD, von Delft F

#### Why we are developing a patent-free COVID antiviral therapy

In this Opinion, we present the case for a new open science driven model for drug discovery for pandemics.

### Best practices for computer-aided drug discovery

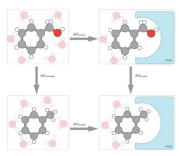


Living Journal of Computational Molecular Science 4:1497, 2022 · DOI

Hahn DF, Bayly CI, Boby ML, Bruce Macdonald HE, Chodera JD, Mey ASJS, Mobley DL, Perez Benito L, Schindler CEM, Tresadern G, Warren GL

#### Best practices for constructing, preparing, and evaluating protein-ligand binding affinity benchmarks

This living best practices paper for the Living Journal of Computational Molecular Sciences describes the current community consensus in how to curate experimental benchmark data for assessing predictive affinity models for drug discovery, how to prepare these systems for affinity calculations, and how to assess the results to compare performance.

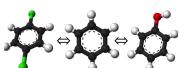


Living Journal of Computational Molecular Science 2:18378, 2020 · DOI

Mey ASJS<sup>‡</sup>, Allen B, Bruce Macdonald HE, Chodera JD<sup>‡</sup>, Kuhn M, Michel J, Mobley DL<sup>‡</sup>, Naden LN, Prasad S, Rizzi A, Scheen J, Shirts MR<sup>‡</sup>, Tresadern G, and Xu H

#### Best practices for alchemical free energy calculations

This living review for the Living Journal of Computational Molecular Sciences (LiveCoMS) covers the essential considerations for running alchemical free energy calculations for rational molecular design for drug discovery.

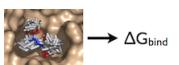


Current Opinion in Structural Biology 21:150, 2011 · DOI

Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson KM, and Pande VS

#### Free energy methods in drug discovery and design: Progress and challenges

A review of current opportunities and challenges for alchemical free energy calculations in drug discovery and design.



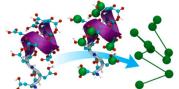
Annual Reports in Computational Chemistry 3:41, 2007 · DOI

Shirts MR, Mobley DL, Chodera JD

#### Alchemical free energy calculations: Ready for prime time?

A review of current alchemical free energy methodologies and their potential for use in drug discovery and ligand design.

### Coarse-graining for biomolecular dynamics

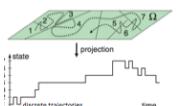


Journal of Chemical Theory and Computation 16:4757, 2020 · DOI

Gkeka P, Stoltz G, Farimani AB, Belkacemi Z, Ceriotti M, Chodera JD, Dinner AR, Ferguson A, Maillet JB, Minoux H, Peter C, Pietrucci F, Silveira A, Tkatchenko A, Trstanova Z, Wiewiora R, Leliévre T

#### Machine learning force fields and coarse-grained variables in molecular dynamics: application to materials and biological systems

We review the state of the art in applying machine learning to coarse grain force fields in space and time to study multiscale dynamics.

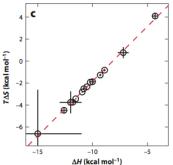


Current Opinion in Structural Biology 25:135, 2014 · DOI

Chodera JD and Noé F

#### Markov state models of biomolecular conformational dynamics

A review of exciting recent developments in the stochastic modeling of biomolecular dynamics using techniques I originally co-developed to study protein folding.

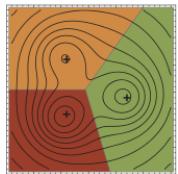


*Annual Reviews in Biophysics* 42:121, 2013 · DOI

Chodera JD and Mobley DL

### Entropy-enthalpy compensation: Role and ramifications for rational ligand design

*Entropy-enthalpy compensation is likely a universal phenomena, but not as severe as widely thought, and irrelevant for drug discovery and ligand design.*

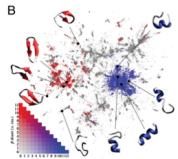


*Journal of Chemical Physics* 134:174105, 2011 · DOI

Prinz JH, Wu H, Sarich M, Keller B, Fischbach M, Held M, Chodera JD, Schütte, and Noé F

### Markov models of molecular kinetics: Generation and validation

*Current best practices for the generation and validation of Markov state models for describing the stochastic dynamics of biomolecular systems.*



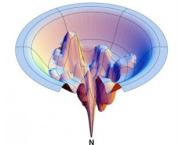
*Proceedings of the National Academy of Sciences* 108:12969, 2011 · DOI

Chodera JD and Pande VS

### The Social Network (of protein conformations)

*A new methodology for mapping protein conformational spaces is reminiscent of how we use two-dimensional maps to navigate a three-dimensional world.*

## Protein folding



*Current Opinion in Structural Biology* 17(3):342, 2007 · DOI

Dill KA, Ozkan SB, Weikl TR, Chodera JD, and Voelz VA

### The protein folding problem: When will it be solved?

*A review of the current state of the protein folding problem.*

## All Published and In Press Publications

### New publications since promotion to Associate Member



Scientific Data, in press · [arXiv](#)

Eastman P, Behara PK, Dotson DL, Galvelis R, Herr JE, Horton JT, Mao Y, Chodera JD, Pritchard BP, Wang Y, De Fabritiis G, and Markland TE

#### SPICE, a dataset of drug-like molecules and peptides for training machine learning potentials

To remedy the lack of large, open quantum chemical datasets for training accurate general machine learning potentials and molecular mechanics force fields for druglike small molecules and biomolecules, we produce the open SPICE dataset, and show how it can be used to build extremely accurate machine learning potentials.



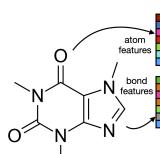
Journal of Chemical Informatics and Modeling 62:22 · [DOI](#)

Horton JT, Boothroyd S, Wagner W, Mitchell JA, Gokey T, Dotson DL, Behara PK, Ramaswamy VK, Mackey M, Chodera JD, Anwar J, Mobley DL, and Cole DJ

#### Open Force Field BespokeFit: Automating bespoke torsion parametrization at scale

In this paper, we describe an automated pipeline for generating tailored force field parameters for small molecules using quantum chemical or quantum machine learning potentials.

Chemical Science, in press · [DOI](#)



Wang Y, Fass J, and Chodera JD

#### End-to-end differentiable molecular mechanics force field construction

Molecular mechanics force fields have been a workhorse for computational chemistry and drug discovery. Here, we propose a new approach to force field parameterization in which graph convolutional networks are used to perceive chemical environments and assign molecular mechanics (MM) force field parameters. The entire process of chemical perception and parameter assignment is differentiable end-to-end with respect to model parameters, allowing new force fields to be easily constructed from MM or QM force fields, extended, and applied to arbitrary biomolecules.



Journal of Chemical Theory and Computation 18:3566, 2022 · [DOI](#)

Boothroyd S, Wang L-P, Mobley DL, Chodera JD, and Shirts MR

#### The Open Force Field Evaluator: An automated, efficient, and scalable framework for the estimation of physical properties from molecular simulation

We describe a new software framework for automated evaluation of physical properties for the benchmarking and optimization of small molecule force fields according to best practices.



Journal of Chemical Theory and Computation 18:3577, 2022 · [DOI](#)

Boothroyd S, Madin OC, Mobley DL, Wang L-P, Chodera JD, and Shirts MR

#### Improving force field accuracy by training against condensed-phase mixture properties

We use a new automated framework for physical property evaluation and fitting to show how molecular mechanics force fields can be systematically improved by fitting to condensed phase properties.

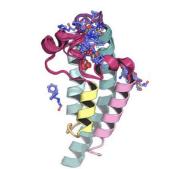


Nature Reviews Chemistry, 6:287, 2022 · [DOI](#)

Ackloo S, Al-awar R, Amaro RE, Arrowsmith CH, Azevedo H, Batey RA, Bengio Y, Betz UAK, Bologa CG, Chodera JD, Cornell WD, Dunham I, Ecker GF, Edfeldt K, Edwards AM, Gilsom MK, Gordijo CR, Hessler G, Hillisch A, Hogner A, Irwin JJ, Jansen JM, Kuhn D, Leach AR, Lee AA, Lessel U, Moult J, Muegge I, Oprea TI, Perry BG, Riley, Singh Saikantendu K, Santhakumar V, Schapira M, Scholten C, Todd MH, Vedadi M, Volkamer A, and Wilson TM

#### CACHE (Critical Assessment of Computational Hit-finding Experiments): A public-private partnership benchmarking initiative to enable the development of computational methods for hit-finding

We describe CACHE: A new public-private partnership that aims to transform computer-aided drug discovery much the way that CASP transformed protein structure prediction into a reproducible, accurate engineering discipline.

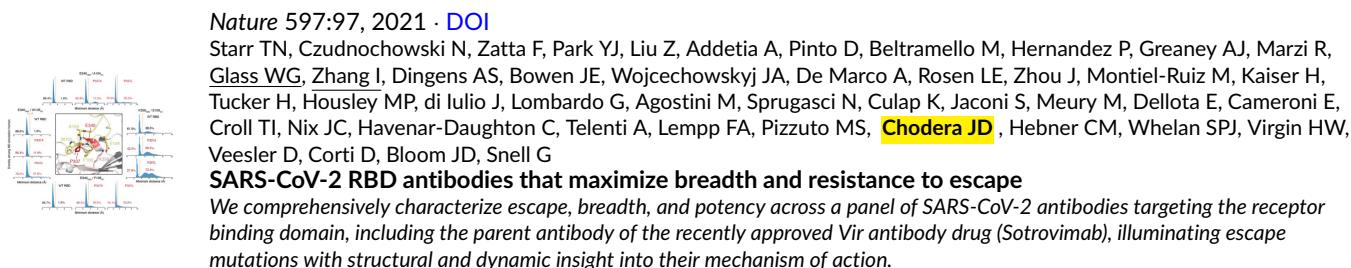
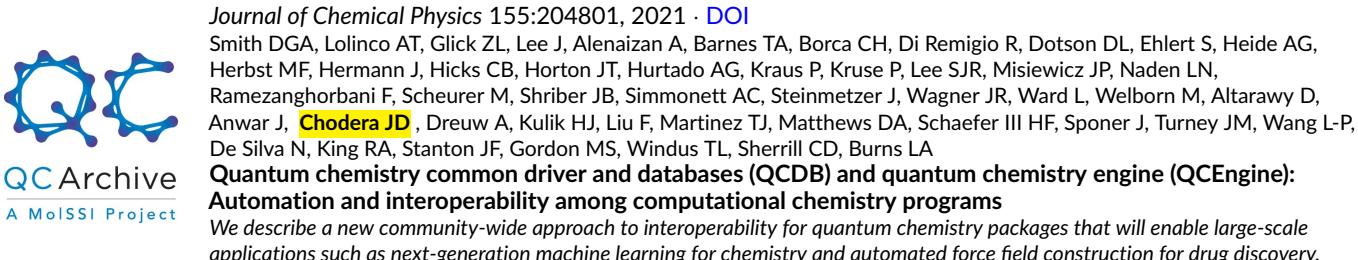
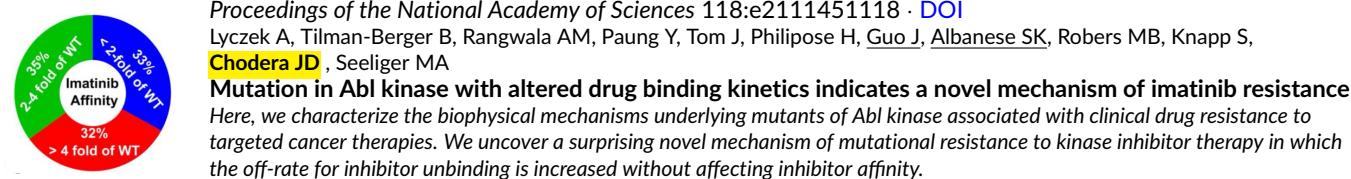
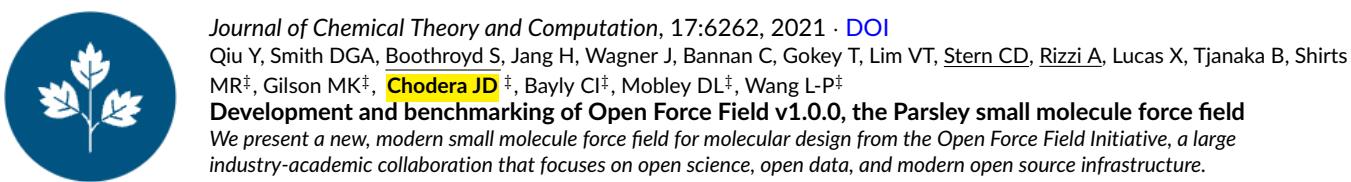
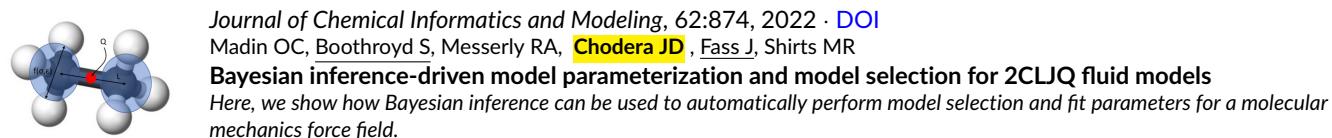
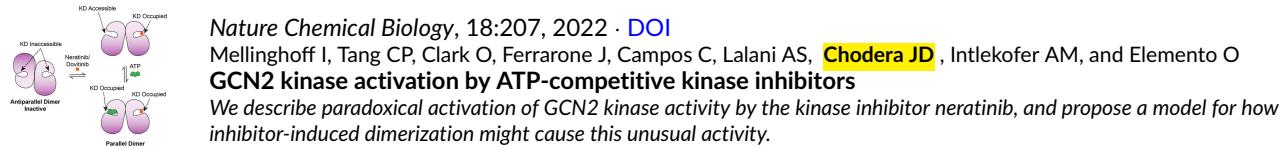
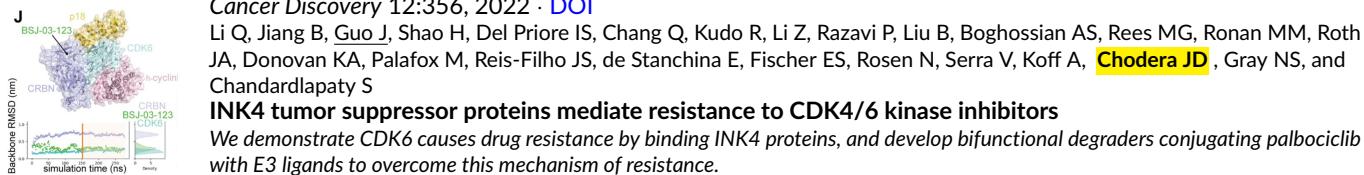


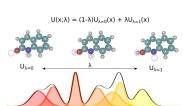
Journal of Computer-Aided Molecular Design 36:291, 2022 · [DOI](#)

Grosjean H, Isik M, Aimon A, Mobley D, Chodera JD, von Delft F, and Biggin PC

#### SAMPL7 protein-ligand challenge: A community-wide evaluation of computational methods against fragment screening and pose-prediction

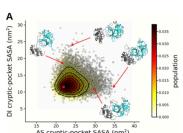
We field a blind community challenge to assess how well state of the art computational chemistry methods can predict the binding modes of small druglike fragments to a protein target for which no chemical matter is known, PHIP2, using fragment screening at the Diamond Light Source.





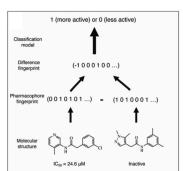
### Fitting quantum machine learning potentials to experimental free energy data: Predicting tautomer ratios in solution

We demonstrate, for the first time, how alchemical free energy calculations can be performed on systems simulated entirely with quantum machine learning potentials and how these potentials can be retrained on experimental free energies to generalize to new molecules from limited training data. We apply this approach to a difficult problem in small molecule drug discovery: Predicting accurate tautomer ratios in solution.



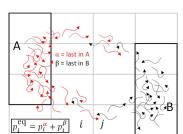
### SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome

To accelerate a multitude of drug development activities to combat the global threat posed by COVID-19, over a million citizen scientists have banded together through the Folding@home distributed computing project to create the world's first Exascale computer and simulate protein dynamics. An unprecedented 0.1 seconds of simulation of the viral proteome reveal how the spike complex uses conformational masking to evade an immune response, conformational changes implicated in the function of other viral proteins, and cryptic pockets that are absent in experimental structures. These structures and mechanistic insights present new targets for the design of therapeutics.



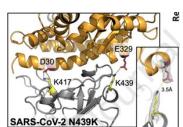
### Discovery of SARS-CoV-2 main protease inhibitors using a synthesis-directed de novo design model

We show how a machine learning models of ligand affinity can be coupled to synthetic enumeration models to rapidly generate potent inhibitors of the SARS-CoV-2 main viral protease.



### What Markov State Models can and cannot do: Correlation versus path-based observables in protein-folding models

Markov state models are now well-established for describing the long-time conformational dynamics of proteins. Here, we take a critical look of what properties can reliably be extracted from these coarse-grained models.



### Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity

New mutations that enhance the affinity of SARS-CoV-2 spike protein for human ACE2—and potentially pose threats to antibody-based therapeutics and vaccines for COVID-19—are already emerging in the wild. We characterize and describe sentinel mutations of SARS-CoV-2 in the wild that herald challenges for combatting COVID-19, and use simulations of the RBD-ACE2 interface on Folding@home to biophysically characterize why these mutations can lead to enhanced affinity.



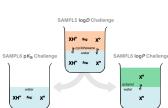
### Crowdsourcing drug discovery for pandemics

The COVID-19 pandemic has left the world scrambling to find effective therapies to stem the tidal wave of death and put an end to the worldwide disruption caused by SARS-CoV-2. In this Correspondence, we argue for the need for a new open, collaborative drug discovery model (exemplified by our COVID Moonshot collaboration) that breaks free of the limitations of industry-led competitive drug discovery efforts that necessarily restrict information flow and hinder rapid progress by prioritizing profits and patent protection over human lives.



### Overview of the SAMPL6 pKa challenge: evaluating small molecule microscopic and macroscopic pKa predictions

The SAMPL6 pKa challenge assessed the ability of the computational chemistry community to predict macroscopic and microscopic pKas for a set of druglike molecules resembling kinase inhibitors. This paper reports on the overall performance and lessons learned, including the surprising finding that many tools predict reasonably accurate macroscopic pKas corresponding to the wrong microscopic protonation sites.

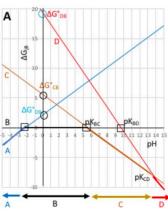


*Journal of Computer Aided Molecular Design* 34:335, 2020 · DOI

Isik M, Bergazin TD, Fox T, Rizzi A, Chodera JD, and Mobley DL

### Assessing the accuracy of octanol-water partition coefficient predictions in the SAMPL6 Part II log P Challenge

We report the performance assessment of the 91 methods that were submitted to the SAMPL6 blind challenge for predicting octanol-water partition coefficient ( $\log P$ ) measurements. The average RMSE of the most accurate five MM-based, QM-based, empirical, and mixed approach methods based on RMSE were  $0.92 \pm 0.13$ ,  $0.48 \pm 0.06$ ,  $0.47 \pm 0.05$ , and  $0.50 \pm 0.06$ , respectively.

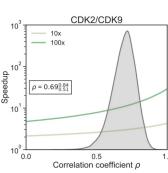


*Journal of Computer Aided Molecular Design* 34:561, 2020 · DOI

Gunner MR, Murakami T, Rustenburg AS, Isik M, Chodera JD

### Standard state free energies, not $pK_a$ s, are ideal for describing small molecule protonation and tautomeric states

Here, we demonstrate how the physical nature of protonation and tautomeric state effects means that the standard state free energies of each microscopic protonation/tautomeric state at a single pH is sufficient to describe the complete pH-dependent microscopic and macroscopic populations. We introduce a new kind of diagram that uses this concept to illustrate a variety of pH-dependent phenomena, and show how it can be used to identify common issues with protonation state prediction algorithms. As a result, we recommend future blind prediction challenges utilize microstate free energies at a single reference pH as the minimal sufficient information for assessing prediction accuracy and utility.

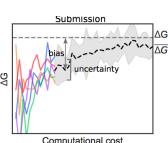


*Journal of Chemical Informatics and Modeling* 60:6211, 2020 · DOI

Albanese SK, Chodera JD, Volkamer A, Keng S, Abel R, and Wang L

### Is structure based drug design ready for selectivity optimization?

We asked whether the similarity of binding sites in related kinases might result in a fortuitous cancellation of errors in using alchemical free energy calculations to predict kinase inhibitor selectivities. Surprisingly, we find that even distantly related kinases have sufficient correlation in their errors that predicting changes in selectivity can be much more accurate than predicting changes in potency due to this effect, and show how this could lead to large reductions in the number of molecules that must be synthesized to achieve a desired selectivity goal.

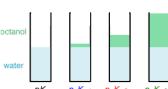


*Journal of Computer Aided Molecular Design* 34:601, 2020 · DOI

Rizzi A, Jensen T, Slochower DR, Aldeghi M, Gapsys V, Ntekoumes D, Bosisio S, Papadourakis M, Henriksen NM, de Groot BL, Cournia Z, Dickson A, Michel J, Gilson MK, Shirts MR, Mobley DL, and Chodera JD

### The SAMPL6 SAMPLing challenge: Assessing the reliability and efficiency of binding free energy calculations

To assess the relative efficiencies of alchemical binding free energy calculations, the SAMPL6 SAMPLing challenge asked participants to submit predictions as a function of computer effort for the same force field and charge model. Surprisingly, we found that most molecular simulation codes cannot agree on the binding free energy was, even for the same force field.

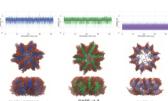


*Journal of Computer Aided Molecular Design* 34:405, 2020 · DOI

Isik M, Levorse D, Mobley DL, Rhodes T, and Chodera JD

### Octanol-water partition coefficient measurements for the SAMPL6 Blind Prediction Challenge

We describe the design and data collection (and associated challenges) for the SAMPL6 part II  $\log P$  octanol-water blind prediction challenge, where the goal was to benchmark the accuracy of force fields for druglike molecules (here, molecules resembling kinase inhibitors).



*Journal of Chemical Theory and Computation* 15:6225, 2019 · DOI

Slochower DR, Hendrikson NM, Wang LP, Chodera JD, Mobley DL, and Gilson MK

### Binding thermodynamics of host-guest systems with SMIRNOFF99Frosst 1.0.5 from the Open Force Field Initiative

We assess the accuracy of the SMIRNOFF99Frosst 1.0.5 force field in reproducing host-guest binding thermodynamics in comparison with the GAFF force field, demonstrating how the SMIRNOFF format for compactly specifying force fields provide comparable accuracy with 20x fewer parameters.



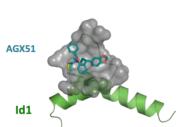
*Journal of Chemical Informatics and Modeling* 59:4093, 2019 · DOI

Abraham MJ, Apostolov R, Barnoud J, Bauer P, Blau C, Bonvin AMMJ, Chavent M, Chodera JD, Condic-Jurkic K, Delemotte L, Grubmüller H, Howard RJ, Jordan J, Lindahl E, Ollila S, Selent J, Smith D, Stansfeld PJ, Tiemann J, Trellet M, Woods C, and Zhumov A

### Sharing data from molecular simulations

There is a dire need to establish standards for sharing data in the molecular sciences. Here, we review the findings of a workshop held in Stockholm in Nov 2018 to discuss this need.

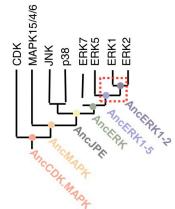
Cell Reports 29:62, 2019 · DOI



Wojnarowicz PM, Lima e Silva R, Ohnaka M, Lee SB, Chin Y, Kulukian A, Chang SH, Desai B, Escolano MG, Shah R, Garcia-Cao M, Xu S, Kadam R, Goldgur Y, Miller MA, Ouerfelli O, Yang G, Arakawa T, Albanese SK, Garland WA, Stoller G, Chaudhary J, Norton L, Soni RK, Philip J, Hendrickson RC, Iavarone A, Dannenberg AJ, Chodera JD, Pavletich N, Lasorella A, Campochiaro PA, Benezra R

### A Small-Molecule Pan-Id Antagonist Inhibits Pathologic Ocular Neovascularization

We report the discovery and characterization of a small molecule, AGX51, with the surprising ability to inhibit the interaction of Ids with E47, which leads to ubiquitin-mediated degradation of Ids.



eLife, 2019;8:e38805 · DOI

Sang D, Pinglay S, Wiewiora RP, Selvan ME, Lou HJ, Chodera JD, Turk B, Gümus Z, and Holt LJ.

### Ancestral reconstruction reveals mechanisms of ERK regulatory evolution

To understand how kinase regulation by phosphorylation emerged, we reconstruct the common ancestor of CDKs and MAPKs, using biochemical experiments and massively parallel molecular simulations to study how a few mutations were sufficient to switch ERK-family kinases from high- to low-autophosphorylation.

eLife 8:e45403, 2019 · DOI

Rafal P. Wiewiora<sup>†</sup>, Shi Chen<sup>†</sup>, Fanwang Meng, Nicolas Babault, Anqi Ma, Wenyu Yu, Kun Qian, Hao Hu, Hua Zou, Junyi Wang, Shijie Fan, Gil Blum, Fabio Pittella-Silva, Kyle A. Beauchamp, Wolfram Tempel, Hualing Jiang, Kaixian Chen, Robert Skene, Y. George Zheng, Peter J. Brown, Jian Jin, Chodera JD<sup>‡</sup>, and Minkui Luo<sup>‡</sup>

### The dynamic conformational landscapes of the protein methyltransferase SETD8

In this work, we show how targeted X-ray crystallography using covalent inhibitors and depletion of native ligands to reveal structures of low-population hidden conformations can be combined with massively distributed molecular simulation to resolve the functional dynamic landscape of the protein methyltransferase SETD8 in unprecedented atomistic detail. Using an aggregate of six milliseconds of fully atomistic simulation from Foldinghome, we use Markov state models to illuminate the conformational dynamics of this important epigenetic protein.

Nature Communications 10:2691, 2019 · DOI

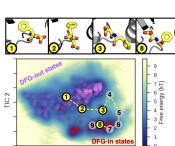
Minuesa G, Albanese SK, Chow A, Schurer A, Park SM, Rotsides CZ, Taggart J, Rizzi A, Naden LN, Chou T, Gourkanti S, Cappel D, Passarelli MC, Fairchild L, Adura C, Glickman FJ, Schulman J, Famulare C, Patel M, Eibl JK, Ross GM, Tan DS, Leslie CS, Beeming T, Golgur Y, Chodera JD, and Kharas MG.

### Small-molecule targeting of MUSASHI RNA-binding activity in acute myeloid leukemia

We use absolute alchemical free energy calculations to identify the likely interaction site for a small hydrophobic ligand that shows activity against MUSASHI in AML.



### Publications prior to promotion to Associate Member

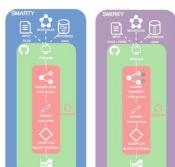


Cell Chemical Biology 26:390, 2019 · DOI

Hanson SM\*, Georgiou G\*, Miller WT, Rest JS, Chodera JD<sup>‡</sup>, and Seeliger MA<sup>‡</sup>

### What makes a kinase promiscuous for inhibitors?

Using a combination of chemogenomics, structural biology, and molecular simulation approaches, we identify a set of human kinases that are especially promiscuous binders of small molecule kinase inhibitors, and show that a prototypical member of this class, DDR1, achieves this promiscuity by virtue of its more stable Asp-DFG-out conformation.

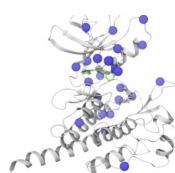


Journal of Chemical Theory and Computation 15:402, 2019 · DOI

Zanette C, Bannan CC, Bayly CI, Fass J, Gilson MK, Shirts MR, Chodera JD, Mobley DL

### Toward learned chemical perception of force field typing rules

We show how machine learning can learn typing rules for molecular mechanics force fields within a Bayesian statistical framework.

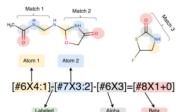


Communications Biology 1:70, 2018 · DOI

Hauser K, Negron C, Albanese SK, Ray S, Steinbrecher T, Abel R, and Chodera JD, and Wang L

### Predicting resistance of clinical Abl mutations to targeted kinase inhibitors using alchemical free-energy calculations

We show how alchemical free energy calculations can be used to predict whether clinical point mutations in human kinase domains confer resistance or susceptibility to targeted kinase inhibitors.



*Journal of Chemical Theory and Computation* 14:6076, 2018 · DOI

Mobley DL‡, Bannan CC, Rizzi A, Bayly CI, Chodera JD, Lim VT, Lim NM, Beauchamp KA, Shirts MR, Gilson MK, and Eastman PK

### Escaping atom types using direct chemical perception with SMIRNOFF v0.1

We describe the philosophy behind a modern approach to molecular mechanics forcefield parameterization, and present initial results for the first SMIRNOFF-encoded forcefield: SMIRNOFF99Frosst.



*Journal of Chemical Theory and Computation* 15:813, 2019 · DOI

Swenson DWH, Prinz JH, Noé F, Chodera JD, and Bolhuis PG

### OpenPathSampling: A Python framework for path sampling simulations. I. Basics

To make powerful path sampling techniques broadly accessible and efficient, we have produced a new Python framework for easily implementing path sampling strategies (such as transition path and interface sampling) in Python. This first publication describes some of the theory and capabilities behind the approach.

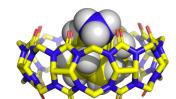


*Journal of Chemical Theory and Computation* 15:837, 2019 · DOI

Swenson DWH, Prinz JH, Noé F, Chodera JD, and Bolhuis PG

### OpenPathSampling: A Python framework for path sampling simulations. II. Building and customizing path ensembles and sample schemes

To make powerful path sampling techniques broadly accessible and efficient, we have produced a new Python framework for easily implementing path sampling strategies (such as transition path and interface sampling) in Python. This second publication describes advanced aspects of the theory and details of how to customize path ensembles.

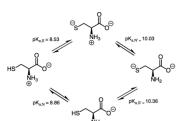


*Journal of Computer Aided Molecular Design* 32:937, 2018 · DOI

Rizzi A, Murkli S, McNeill J, Yao W, Sullivan M, Gilson MK, Chiu MW, Isaacs L, Gibb BC, Mobley DL‡, and Chodera JD ‡

### Overview of the SAMPL6 host-guest binding affinity prediction challenge

We present an overview of the host-guest systems and participant performance for the SAMPL6 host-guest blind affinity prediction challenges, assessing how well various physical modeling approaches were able to predict ligand binding affinities for simple ligand recognition problems where receptor sampling and protonation state effects are eliminated due to the simplicity of supramolecular hosts. We find that progress is now stagnated likely due to force field limitations.

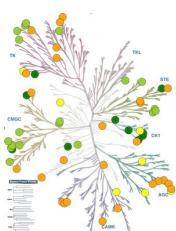


*Journal of Computer Aided Molecular Design* 32:1117, 2018 · DOI

Isik M, Levorse D, Rustenburg AS, Ndukwie IE, Wang H, Reibarkh M, Martin GE, Makarov AA, Mobley DL, Rhodes T‡, and Chodera JD ‡

### pKa measurements for the SAMPL6 prediction challenge for a set of kinase inhibitor-like fragments

The SAMPL5 blind challenge exercises identified neglect of protonation state effects as a major accuracy-limiting factor in physical modeling of biomolecular interactions. In this study, we report the experimental measurements behind a SAMPL6 blind challenges in which we assess the ability of community codes to predict small molecule pKas for small molecule resembling fragments of selective kinase inhibitors.

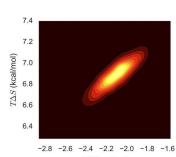


*Biochemistry* 57:4675, 2018 · DOI

Albanese SK\*, Parton DL\*, Isik M†, Rodríguez-Laureano L†, Hanson SM, Gradia S, Jeans C, Levinson NM, Seeliger M, and Chodera JD

### An open library of human kinase domain constructs for automated bacterial expression

To establish a tractable experimental system for studying the biophysical determinants of selective kinase inhibitor resistance in clinical cancer mutations, we engineer a library of human kinase domains with useful bacterial expression with phosphatase coexpression. Over 350 requests have been made for these plasmids.

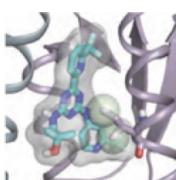


*PLOS ONE* 3(9): e0203224, 2018 · DOI

Nguyen TH, Rustenburg AS, Krimmer SG, Zhang H, Clark JD, Novick PA, Branson K, Pande VS, Chodera JD ‡, MinH DDL‡

### Bayesian analysis of isothermal titration calorimetry for binding thermodynamics

We show how Bayesian inference can produce greatly improved estimates of statistical uncertainty from isothermal titration calorimetry (ITC) experiments, allowing the joint distribution of thermodynamic parameter uncertainties to be inferred.

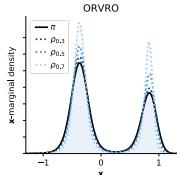


*Nature* 559:125, 2018 · DOI

Intlekofer AM\*, Shih AH\*, Wang B, Nazir A, Rustenburg AS, Albanese SK, Patel M, Famulare C, Correa FM, Arcila ME, Taylor J, Tallman MS, Rosenthal M, Petsko GA, Chodera JD, Thompson CB‡, Levine RL‡, Stein EM‡

### Acquired resistance to IDH inhibition through trans or cis dimer-interface mutations

Clinical double mutations acting in trans in cancer patients receiving IDH2 inhibitors act through a novel biophysical mechanism.

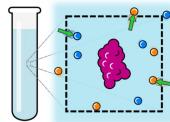


*Entropy* 20:318, 2018 · DOI

Fass J, Sivak DA, Crooks GE, Beauchamp KA, Leimkuhler B, and Chodera JD

### Quantifying configuration-sampling error in Langevin simulations of complex molecular systems

We address a fundamental question regarding why molecular dynamics simulation works despite the fact that the use of finite timesteps leads to error in the sampled probability densities and populations, demonstrating how to measure configuration-space sampling error for an important class of Langevin integrators widely used in biomolecular simulation.



*Journal of Physical Chemistry B* 122:5466, 2018 · DOI

Ross GA, Rustenburg AS, Grinaway PB, Fass J, and Chodera JD

### Biomolecular simulations under realistic salt conditions

We show how NCMC can be used to implement an efficient osmostat in molecular dynamics simulations to model realistic fluctuations in ion environments around biomolecules, and illustrate how the local salt environment around biological macromolecules can differ substantially from bulk.



*Journal of Physical Chemistry* 122:5579, 2018 · DOI

Gill SC, Lim NM, Grinaway PB, Rustenburg AS, Fass J, Ross GA, Chodera JD, and Mobley DL

### Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes Using Nonequilibrium Candidate Monte Carlo

Nonequilibrium candidate Monte Carlo can be used to accelerate the sampling of ligand binding modes by orders of magnitude over instantaneous Monte Carlo.

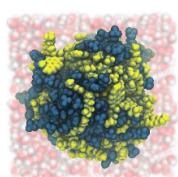


*eLife* 7:e32766, 2018 · DOI

Ruff EF, Muretta JM, Thompson A, Lake E, Cyphers S, Albanese SK, Hanson SM, Behr JM, Thomas DT, Chodera JD, and Levinson NM

### A dynamic mechanism for allosteric activation of Aurora kinase A by activation loop phosphorylation

Through a combination of FRET, IR, and EPR labeling and large-scale molecular dynamics simulations, we show that phosphorylation activates Aurora kinase by a novel mechanism that does not simply correspond to a DFG-out to DFG-in population shift, but rather reorganization of DFG-in subpopulations.

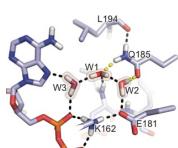


*Nature Materials* 17:361, 2018 · DOI

Shamay Y, Shah J, Tschaharganeh DF, Roxbury D, Budhathoki-Uprety J, Ijsik M, Mizrahi A, Nawaly K, Sugarman JL, Baut E, Neiman MR, Johnson DC, Sridharan R, Chu KL, Rajasekhar VK, Chodera JD, Lowe SW, and Heller DA

### Quantitative self-assembly prediction yields targeted nanoparticles

A decision tree based on predicted physical properties and molecular descriptors is capable of predicting the assembly of drug/dye nanoparticles that can be used in tumor-targeted selective kinase inhibitor therapy to minimize on- and off-pathway toxicity.



*Nature Chemical Biology* 13:402, 2017 · DOI

Cyphers S, Ruff E, Behr JM, Chodera JD, and Levinson NM

### A conserved water-mediated hydrogen bond network governs allosteric activation in Aurora kinase A

Over 50 microseconds of aggregate simulation data on Folding@home reveal a surprisingly stable hydrogen bond network underlies allosteric activation by Tpx2.

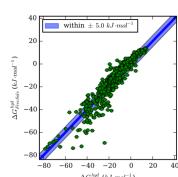


*Nature Chemical Biology* 13:494, 2017 · DOI

Intlekofer A, Wang B, Liu H, Shah H, Carmona-Fontaine C, Rustenburg AS, Salah S, Gunner MR, Chodera JD, Cross JR, and Thompson CB

### Acidification enhances production of L-2-hydroxyglutarate through alternative substrate use by dehydrogenase enzymes

At low pH, metabolic enzymes lactate dehydrogenase and malate dehydrogenase undergo shifts in substrate utilization that have high relevance to cancer metabolism due to surprisingly simple protonation state effects.



*Journal of Chemical & Engineering Data* 62:1559, 2017 · DOI

Matos GDR, Kyu DY, Loeffler HH, Chodera JD, Shirts MR, and Mobley DL

### Approaches for calculating solvation free energies and enthalpies demonstrated with an update of the FreeSolv database

We review alchemical approaches to computing solvation free energies and update FreeSolv—the most popular database of hydration free energies of neutral molecules—with more computed and experimental properties.

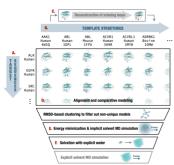


PLoS Computational Biology 13:e1005659, 2017 · DOI

Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang LP, Simmonett AC, Harrigan MP, Brooks BR, and Pande VS

### OpenMM 7: Rapid development of high performance algorithms for molecular dynamics

The latest version of the GPU-accelerated molecular simulation OpenMM features a variety of incredibly flexible but fast tools for rapidly prototyping, evaluating, and deploying new simulation algorithms.



PLoS Computational Biology 12:e1004728, 2016 · DOI

Parton DL, Grinaway PB, Hanson SM, Beauchamp KA, and Chodera JD

### Ensembler: Enabling high-throughput molecular simulations at the superfamily scale

We demonstrate a new tool that enables—for the first time—massively parallel molecular simulation studies of biomolecular dynamics at the superfamily scale, illustrating its application to protein tyrosine kinases, an important class of drug targets in cancer.

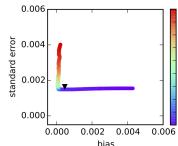


Journal of Clinical Investigation 126:3529, 2016 · DOI

Xu J, Pham CG, Albanese SK, Dong Y, Oyama T, Lee CH, Rodrik-Outmezguine V, Yao Z, Han S, Chen D, Parton DL, Chodera JD, Rosen N, Cheng EH, and Hsieh JJ

### Mechanistically distinct cancer-associated mTOR activation clusters predict sensitivity to rapamycin

We use massively parallel distributed molecular simulations on Folding@home to probe the mechanism activating mutations of the mTOR kinase identified in clinical populations.

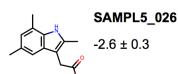


Journal of Chemical Theory and Computation 12:1799, 2016 · DOI

Chodera JD

### A simple method for automated equilibration detection in molecular simulations

We present a simple approach to automatically determining the equilibrated region of a molecular simulation, a longstanding challenge formerly without a good solution.

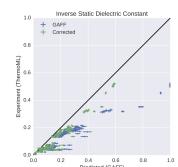


Journal of Computer Aided Molecular Design, 30:945, 2016 · DOI

Rustenburg AS, Dancer J, Lin B, Ortwin D, Mobley DL, and Chodera JD

### Measuring cyclohexane-water distribution coefficients for the SAMPL5 challenge

To test the accuracy of physical modeling techniques in predicting free energies of transfer between aqueous and nonpolar solvents, we worked with Genentech to develop a new protocol to measure cyclohexane-water distribution coefficients for 53 druglike compounds at pH 7.4, fielding a blind community challenge as part of the SAMPL5 exercise. A special issue of JCAMD was published with 16 papers describing various approaches used by participants to predict this data and understand their failures.

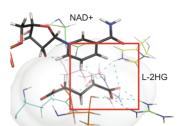


Journal of Physical Chemistry B 199:12912, 2015 · DOI

Beauchamp KA, Behr JM, Rustenburg AS, Bayly CI, Kroenlein K, and Chodera JD

### Towards automated benchmarking of atomistic forcefields: Neat liquid densities and static dielectric constants from the ThermoML data archive

Molecular mechanics forcefields are critical to computer-guide drug design, but the benchmarking and improvement of these forcefields has been hindered by the lack of high-quality machine-readable physical property datasets. We show how the NIST-curated ThermoML Archive, which stores physical property data in an IUPAC-standard XML format, can eliminate these roadblocks and reveal issues with current generation forcefields.

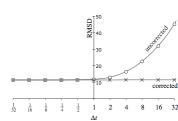


Cell Metabolism 22:1–8, 2015 · DOI

Intlekofer AM, Dematteo RG, Venetti S, Finley LWS, Lu Chao, Judkins AR, Rutenburg AS, Grinaway PB, Chodera JD, Cross JR, and Thompson CB

### Hypoxia introduces production of L-2-Hydroxyglutarate

Molecular docking is used to demonstrate the potential for alternative substrate usage by isocitrate dehydrogenases under hypoxic conditions in cancer.

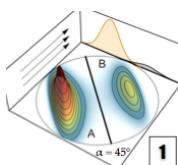


Journal of Physical Chemistry B, 118:6466–6474, 2014. William C. Swope Festschrift · DOI

Sivak DA, Chodera JD, and Crooks GE

### Time step rescaling recovers continuous-time dynamical properties for discrete-time Langevin integration of nonequilibrium systems

We derive a simple, easy-to-implement Langevin integrator that has universally useful properties in molecular simulations.

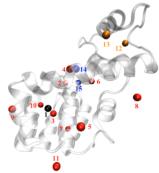


Physical Review X 4:011020, 2014 · DOI

Prinz J-H, Chodera JD, and Noé F

### Spectral rate theory for two-state kinetics

We present a new mathematical framework for unifying various two-state rate theories presented in the physical chemistry literature over many decades, and provide a quantitative way to measure reaction coordinate quality.

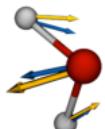


*Journal of Computer Aided Molecular Design* 27:989–1007, 2013 · DOI

Wang K, Chodera JD, Yang Y, and Shirts MR

### Identifying ligand binding sites and poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics

We show how bound ligand poses can be identified even when the location of the binding sites are unknown using the machinery of alchemical modern free energy calculations on graphics processors.

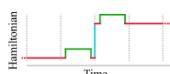


*Journal of Physical Chemistry B* 117:9956–9972, 2013 · DOI

Wang L-P, Head-Gordon TL, Ponder JW, Ren P, Chodera JD, Eastman PK, Martinez TJ, and Pande VS

### Systematic improvement of a classical molecular model of water

Water is the most important molecule in biology, and accurate treatment of its interactions is critical to accurate modeling for drug discovery. While polarizable models of water can achieve very high accuracies, they are both difficult to parameterize and expensive to employ. Here, we show how a high quality inexpensive polarizable model of liquid water can be derived using an automated parameterization engine.



*Physical Review X* 3:011007, 2013 · DOI

Sivak DA, Chodera JD, and Crooks GE

### Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium discrete Langevin dynamics simulations

All molecular dynamics simulations introduce error into the sampled distribution by virtue of the finite timestep used to integrate the equations of motion on a digital computer. While traditional approaches to analyzing this error are extremely complicated, we show how interpreting finite-timestep integrators as a form of nonequilibrium driving leads to simple, straightforward schemes for assessing the impact of these errors, as well as correcting for them.

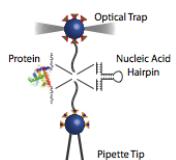


*Journal of Chemical Theory and Computation* 9:461, 2012 · DOI

Eastman P, Friedrichs MS, Chodera JD, Radmer RJ, Bruns CM, Ku JP, Beauchamp KA, Lane TJ, Wang L, Shukla D, Tye T, Houston M, Stich T, Klein C, Shirts MR, and Pande VS

### OpenMM 4: A reusable, extensible, hardware independent library for high performance molecular simulation

Inexpensive consumer GPUs promise a 100-fold increase in simulation power by problems that can effectively exploit their highly specialized structure. Here, we describe the latest advances in an extremely high performance, open-source, extensible GPU-accelerated library and toolkit for molecular simulation.

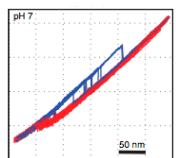


*Biophysical Journal* 103:1490, 2012 · DOI

Elms PJ, Chodera JD, Bustamante CJ, Marqusee S

### The limitations of constant-force-feedback experiments

Popular constant-force-feedback single-molecule experiments can cause severe artifacts in single-molecule force spectroscopy data. We demonstrate a simple alternative that eliminates these artifacts.



*Proceedings of the National Academy of Sciences* 109:3796, 2012 · DOI

Elms PJ, Chodera JD, Bustamante C, Marqusee S

### The molten globule state is unusually deformable under mechanical force

We measure the physical properties of the molten globule state of apo-myoglobin, and show that it is unusually deformable compared to typical protein native states.

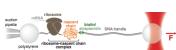
$$f_{\text{exp}} = \int d\vec{r} f(\vec{r}) p_i(\vec{r})$$

*Journal of Chemical Theory and Computation* 8:3445, 2012 · DOI

Pitera JW and Chodera JD

### On the use of experimental observations to bias simulated ensembles

We show how the concept of maximum entropy can be used to recover unbiased conformational distributions from experimental data, and how this concept relates to the popular ‘ensemble refinement’ schemes for NMR data analysis.

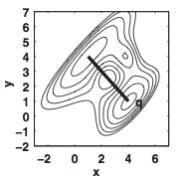


*Science* 334:1723, 2011 · DOI

Kaiser CM, Goldman DH, Chodera JD, Tinoco I, Jr., and Bustamante C

### The ribosome modulates nascent protein folding

Using single-molecule force spectroscopy, we show how the ribosome itself modulates the folding dynamics of nascent protein chains emerging from the exit tunnel.

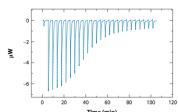


*Physical Review Letters* 107:098102, 2011 · DOI

Chodera JD and Pande VS

### Splitting probabilities as a test of reaction coordinate choice in single-molecule experiments

We demonstrate a simple test for identifying poor reaction coordinates in single-molecule experiments.

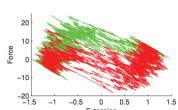


Analytical Biochemistry 414:297, 2011 · DOI

Tellinghuisen JT and Chodera JD

### Systematic errors in isothermal titration calorimetry: Concentrations and baselines

A word of caution about large errors in isothermal titration calorimetry measurements arising from ligand concentration errors.

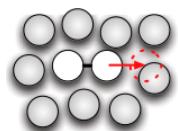


Journal of Chemical Physics 134:024111, 2011 · DOI

Minh DDL, Chodera JD

### Estimating equilibrium ensemble averages using multiple time slices from driven nonequilibrium processes: Theory and application to free energies, moments, and thermodynamic length in single-molecule pulling experiments

We derive a new estimator for estimating equilibrium expectations from nonequilibrium experiments, and show how it can be used to estimate a variety of useful quantities in simulated single-molecule force spectroscopy experiments.

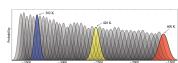


Proceedings of the National Academy of Sciences 108:E1009, 2011 · DOI

Nilmeier JP, Crooks GE, Minh DDL, and Chodera JD

### Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation

We present a significant generalization of Monte Carlo methods that provide an enormously useful tool for enhancing the efficiency of molecular simulations and enabling molecular design.

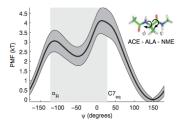


Journal of Chemical Physics 134:244108, 2011 · DOI

Prinz J-H, Chodera JD, Pande VS, Smith JC, and Noé F

### Optimal use of data in parallel tempering simulations for the construction of discrete-state Markov models of biomolecular dynamics

We demonstrate how multitemperature data from parallel tempering simulations can be used to construct fully temperature-dependent models of the dynamics of biomolecular systems.

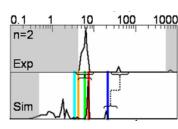


Journal of Chemical Physics 134:244107, 2011 · DOI

Chodera JD, Swope WC, Noé F, Prinz J-H, Shirts MR, and Pande VS

### Dynamical reweighting: Improved estimates for dynamical properties from simulations at multiple temperatures

We describe how reweighing techniques can provide optimal estimates of temperature-dependent dynamical properties from simulations conducted at multiple temperatures.

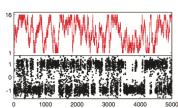


Proceedings of the National Academy of Sciences 108:4822, 2011 · DOI

Noé F, Doose S, Daidone I, Löllmann M, Sauer M, Chodera JD, and Smith JC

### Dynamical fingerprints: A theoretical framework for understanding biomolecular processes by combination of simulation and kinetic experiments

We present a new framework for comparing essential features of the dynamics between experiment and simulation to identify the kinetics processes contributing to individual relaxation timescales in perturbation-response or correlation spectroscopy experiments.

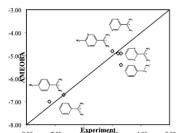


Journal of Chemical Physics 135:194110, 2011 · DOI

Chodera JD and Shirts MR

### Replica exchange and expanded ensemble simulations as Gibbs sampling: Simple improvements for enhanced mixing

We show how a simple change to the way exchanges are handled in the popular replica-exchange simulation methodology can astronomically increase efficiency at no increase in computational cost.

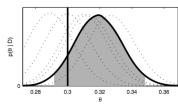


Journal of Physical Chemistry B 114:2549, 2010 · DOI

Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Mobley DL, Schnieders MJ, Haque I, Lambrecht DS, DiStasio RA Jr., Head-Gordon M, Clark GN, Johnson ME, and Head-Gordon T

### Current status of the AMOEBA polarizable force field

The AMOEBA polarizable force field is able to reproduce a diverse set of physical chemical phenomenon to high accuracy.



Journal of Chemical Physics 133:105102, 2010 · DOI

Chodera JD and Noé F

### Probability distributions of molecular observables computed from Markov models.

#### II. Uncertainties in observables and their time-evolution

A simple Bayesian approach for the modeling of statistical uncertainties in kinetic and equilibrium quantities computed from Markov state models of biomolecular dynamics.

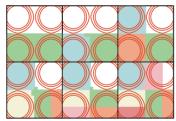


Biophysical Journal 98:3062, 2010 · DOI

Adelman JL, Chodera JD, Kuo IW, Miller TF, and Barsky D

### The mechanical properties of PCNA: Implications for the loading and function of a DNA sliding clamp

Molecular simulations of the PCNA clamp responsible for DNA polymerase processivity show a surprisingly small energetic penalty for the deformation required for clamp loading. Featured on issue cover.

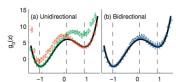


Journal of Chemical Physics 131:045106, 2009 · DOI

Bacallado S, Chodera JD, and Pande VS

### Bayesian comparison of Markov models of molecular dynamics with detailed balance constraint

A Bayesian scheme for comparing state space decompositions for Markov state models of biomolecular dynamics that incorporates the fact that physical systems must obey detailed balance. This paper utilizes recent results from Markov chain theory on edge-reinforced random walks.

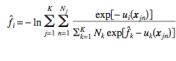


Journal of Chemical Physics 131:134110, 2009 · DOI

Minh DDL, Chodera JD

### Optimal estimators and asymptotic variances for nonequilibrium path-ensemble averages

We derive an optimal estimator and corresponding statistical uncertainties for inferring expectations of bidirectional nonequilibrium processes. These estimators have widespread applicability in single-molecule biophysical force-spectroscopy experiments and nonequilibrium molecular simulations.



Journal of Chemical Physics 129:124105, 2008 · DOI

Shirts MR, Chodera JD

### Statistically optimal analysis of samples from multiple equilibrium states

We present a highly general, statistically optimal approach for producing estimates of free energies and equilibrium expectations from multiple simulations that provably extracts all useful information from the data.



Journal of Medicinal Chemistry 51:769, 2008 · DOI

Nicholls A\*, Mobley DL\*, Guthrie JP, Chodera JD, and Pande VS

### Predicting small-molecule solvation free energies: A blind challenge test for computational chemistry

A blind evaluation of the accuracy of alchemical free energy methods for computing gas-to-water transfer free energies (solvation free energies) of small molecules demonstrates that modern forcefields are likely sufficiently accurate to be useful in drug design.

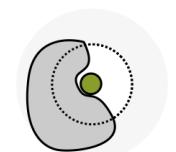


Journal of Physical Chemistry B 112:938, 2008 · DOI

Mobley DL, Dill KA, and Chodera JD

### Treating entropy and conformational changes in implicit solvent simulations of small molecules

An quantitative examination of how much conformational entropy contributes to hydration free energies of small molecules, with implications for ligand binding.

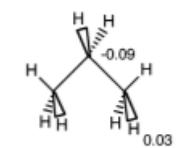


Journal of Physical Chemistry B 111:13052, 2007 · DOI

Shirts MR\*, Mobley DL\*, Chodera JD, and Pande VS

### Accurate and efficient corrections for missing dispersion interactions in molecular simulations

We identify a major source of systematic error in absolute alchemical free energy calculations of ligand binding and show how a simple procedure can inexpensively and accurately eliminate it.

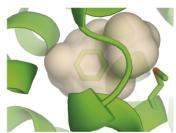


Journal of Physical Chem B 111:2242, 2007 · DOI

Mobley DL, Dumont E, Chodera JD, Bayly CI, Cooper MD, and Dill KA

### Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent

We compare a number of popular methods for deriving charge models for small molecules, deriving lessons about best practices for accurate simulations.

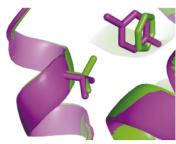


Journal of Molecular Biology 371:1118, 2007 · DOI

Mobley DL, Graves AP, Chodera JD, McReynolds AC, Shoichet BK, and Dill KA

### Predicting absolute ligand binding free energies to a simple model site

We show how alchemical free energy calculations are capable of accurate blind prediction of small-molecule binding affinities to a simple model protein binding site.



Journal of Chemical and Theoretical Computation 3:1231, 2007 · DOI

Mobley DL, Chodera JD, and Dill KA

### Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change

We present a general scheme for obtaining correct ligand binding affinities when protein conformational change is implicated in ligand binding.



Journal of Chemical Physics 126:155101, 2007 · DOI

Chodera JD\*, Singhal N\*, Swope WC, Pitera JW, Pande VS, and Dill KA

### Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics

Proposing one of the first automated algorithms for discovering kinetically metastable states of biomolecules from molecular simulations, this paper shows how many biomolecules can possess numerous distinct long-lived conformational states even though the equilibrium populations of these states may too small for standard structural biology techniques to detect.

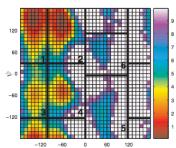


Proceedings of the National Academy of Sciences 104:11987, 2007 · DOI

Ozkan SB, Wu GA, Chodera JD, and Dill KA

### Protein Folding by Zipping and Assembly

A review of the utility of the proposed zipping and assembly mechanism for the concomitant formation of secondary and tertiary structure in protein folding for predicting folding pathways and native structures.

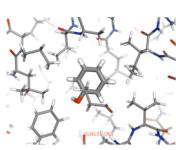


Journal of Chemical Theory and Computation 3:26, 2007 · DOI

Chodera JD, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill

### Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations

The weighted histogram analysis method (WHAM), a mainstay of molecular dynamics simulation analysis, is thoroughly explained and modernized for the analysis of simulated and parallel tempering simulation data.

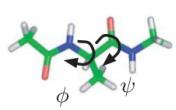


Journal of Chemical Physics 125:084902, 2006 · DOI

Mobley DL, Chodera JD, and Dill KA

### On the use of orientational restraints and symmetry corrections in alchemical free energy calculations

We illustrate how orientational restraints can be used to greatly reduce the computational effort in alchemical calculations of ligand binding free energies, and clarify how symmetry corrections are necessary when molecules contain symmetric or pseudosymmetric substituents.

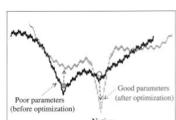


Multiscale Modeling and Simulation 5:1214, 2006 · DOI

Chodera JD, Swope WC, Pitera JW, and Dill KA

### Long-time protein folding dynamics from short-time molecular dynamics simulations

We show how the long-time dynamics of biomolecular systems can be recapitulated from statistics collected from short molecular simulations sampling transitions between kinetically metastable states.

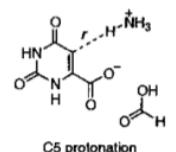


Journal of Computational Chemistry 24:89, 2003 · DOI

Seok C, Rosen JB, Chodera JD, Dill KA

### MOPED: Method for optimizing physical energy parameters using decoys

We propose a new way to optimize parameters for a physical energy function using decoy structures for protein folding studies.



Journal of the American Chemical Society 123:12837, 2001 · DOI

Lee TS\*, Chong LT\*, Chodera JD, and Kollman PA

### An alternative explanation for the catalytic proficiency of orotidine 5'-phosphate decarboxylase

A combined QM and MD analysis of potential plausible mechanisms to explain the enormous catalytic acceleration of one of the most proficient enzymes known.

## Peer reviewer for scientific journals

Bioinformatics, Biopolymers, Chemical Physics, Computation, Drug Discovery Today, European Biophysics Journal, Entropy, International Journal of Molecular Sciences, Journal of the American Chemical Society, Journal of Chemical Theory and Computation, Journal of Computer-Aided Molecular Design, Journal of Computational Chemistry, Journal of Computational Physics, Journal of Physical Chemistry, Journal of Physical Chemistry Letters, Molecular Physics, Multiscale Modeling & Simulation, Nature Chemistry, Nature Communications, Nature Physics, Pacific Symposium in Biocomputing, PLoS Computational Biology, PLoS One, Proceedings of the National Academy of Sciences, Science, Scientific Reports, Structure

## Graduate Programs

- 2013- Program in Physiology, Biophysics, and Systems Biology (PBSB)  
2013- Tri-Institutional PhD Program in Chemical Biology (TPCB)  
2013- Tri-Institutional Program in Computational Biology and Medicine (CBM)  
2015- Gerstner Sloan Kettering Graduate Program (GSK)

## Teaching

- Virginia Tech  
Summer 2017 MolSSI Software Summer School Instructor  
Molecular Software Sciences Institute  
Tri-I  
2015-2018 CBM Journal Club Moderator  
Tri-I Computational Biology and Medicine Program  
WCMC  
2018- BCMB Biochemistry  
WCMC Graduate School of Medical Sciences (statistical mechanics and thermodynamics unit, 3.0 lecture hours)

## Grant Reviews and Study Sections

### **NIH**

Macromolecular Structure and Function B [MSFB] (*early career reviewer*)

### **NSF**

Ad hoc reviewer and virtual panel member for Chemical Theory, Models, and Computational Methods (CTMC)

## Conferences organized

2020-	Pause due to COVID-19 pandemic
London, UK Sep 2019	CompBioMed 2019
Berlin, Germany May 2019	IET London
Göttingen, Germany May 2019	MolKin 2019: Molecular Kinetics: Sampling, Design and Machine Learning Freie Universität Berlin, Germany
Boston, MA May 2018	Alchemical Free Energy Methods Workshop Novartis, Cambridge, MA
San Francisco, CA Sep 2017	Free Energy Methods and Molecular Kinetics in Drug Design Workshop Novartis, Cambridge, MA
Berkeley, CA Jan 2017	MolSSI Workshop on Workflows in Biomolecular Simulation Autodesk
Boston, MA May 2016	SMML//2017: Statistical Mechanics // Machine Learning University of California, Berkeley
Cambridge, MA May 2016	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Boston, MA
Berlin, Germany Sep 2015	Markov State Models in Drug Design Workshop Novartis, Cambridge, MA
Boston, MA May 2014	World Molecular Kinetics Workshop 2015 Freie Universität Berlin, Germany
Berlin, Germany Sep 2013	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Boston, MA
Cambridge, MA May 2012	World Molecular Kinetics Workshop 2013 Freie Universität Berlin, Germany
Berlin, Germany Sep 2011	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Cambridge, MA
Cambridge, MA May 2010	World Molecular Kinetics Workshop 2011 Freie Universität Berlin, Germany
Berlin, Germany May 2009	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Cambridge, MA
	World Molecular Kinetics Workshop 2009 Freie Universität Berlin, Germany

## Recent Invited Talks

### Seminars

Toronto, Canada 6 Jun 2022	The COVID Moonshot: Open science discovery of a novel SARS-CoV-2 antiviral University of Toronto
ZOOM 14 Apr 2022	The future of free energy: Calculations that can learn from experiment OpenEye Scientific
ZOOM 12 Apr 2022	The future of free energy: Calculations that can learn from experiment Amgen
ZOOM 3 Feb 2022	Open science antiviral discovery with the COVID Moonshot and the open source discovery ecosystem NIH BISTI Seminar
ZOOM 16 Dec 2021	The COVID Moonshot: How we got here and where we're going COVID Moonshot Livestream
ZOOM 8 Dec 2021	Redesigning drug design University of Toronto
ZOOM 8 Dec 2021	Frontiers in structure-based drug discovery Rutgers IQB Seminar (celebrating 50 years of the PDB)
ZOOM 4 Oct 2021	Next-generation technologies for computer-driven drug discovery: Teaching free energy calculations to learn, scaling predictions across the kinome, and automating discovery Genentech
ZOOM 14 Jun 2021	Redesigning drug design Abbvie
ZOOM 23 Apr 2021	The COVID Moonshot: An open science collaboration to develop an orally bioavailable inhibitor of SARS-CoV-2 main viral protease Laufer Center, Stony Brook University
ZOOM 12 Apr 2021	Next-generation strategies for designing novel kinase inhibitors and predicting the emergence of cellular resistance Albert Einstein College of Medicine
ZOOM 31 Mar 2021	Redesigning drug design Oxford Centre for Medicines Discovery
ZOOM 16 Dec 2020	Folding@home: Crowdsourcing computing for a cure COVID Moonshot: State of the Moonshot Livestream
ZOOM 8 Apr 2020	The COVID Moonshot: Crowdsourcing a COVID-19 Cure Folding@home 20th Anniversary Livestream
ZOOM 8 Apr 2020	COVID Moonshot: Docking and free energy calculations BioSolvEiT webinar
ZOOM 8 Apr 2020	Drug discovery and how to fix it Google X
San Francisco, CA 3 Mar 2020	Accelerating molecular discovery of small molecule modulators of protein-protein interactions with computational chemistry Foresite Labs
New York, NY 25 Feb 2020	Redesigning drug design Albert Einstein College of Medicine
San Francisco, CA 29 Jan 2020	Building predictive models of antibody:antigen mutations Vir Biotechnology
Berlin, Germany 13 Nov 2019	Redesigning drug design Berlin Pharmaceutical Research Colloquium
Biberach, Germany 7 Nov 2019	Toward adaptive free energy calculations and physically-informed machine learning for predicting affinity, selectivity, and resistance Boehringer Ingelheim
Pittsburgh, PA 6 Sep 2019	Harnessing 100,000 computers around the world to probe kinase activation, inhibition, and resistance University of Pittsburgh
Edinburgh, UK 6 Aug 2019	Alchemical free energy calculations, adaptive sampling, and machine learning University of Edinburgh
Dortmund, Germany 2 Jun 2019	Redesigning drug design Dortmund Universität
Slough, UK 31 May 2019	Predictive models for molecular design UCB Pharma
Wuppertal, Germany 29 May 2019	Predictive models for molecular design

	Bayer
Cambridge, MA 14 May 2019	Predictive models for molecular design XtalPi
Tampa, FL 7 May 2019	Alchemical free energy calculations for biomolecular design Vir Biotechnology
Tampa, FL 29 Apr 2019	Harnessing 100,000 computers around the world to probe kinase activation, inhibition, and resistance University of South Florida
Tempe, AZ 17 Apr 2019	Redesigning drug design Arizona State University
Cambridge, MA 15 Apr 2019	Redesigning drug design: New predictive modeling tools for designing selective small molecules with targeted properties Celgene
Bethesda, MD 15 Nov 2018	Advancing quantitative biophysical predictions: How can we move things forward? National Institute for Advancing Translational Sciences (NIH NCATS)
Philadelphia, PA 4 Oct 2018	What will it take to predict kinase inhibitor selectivity and resistance using computational physical modeling? Fox Chase Cancer Center
Waltham, MA 15 Aug 2018	Predicting porin permeation for antimicrobial drug development Entasis Therapeutics
Göttingen, Germany 18 Jul 2018	Redesigning drug design Max Planck Institute for Biophysical Chemistry
Berlin, Germany 17 Jul 2018	Redesigning drug design: Selectively targeting the kinome and beyond Charité
Darmstadt, Germany 8 Jul 2018	Can free energy really be free? The state of open source absolute and relative free energy calculations Merck KGaA
Cambridge, MA 22 May 2018	What will it take to predict kinase inhibitor selectivity and resistance using computational physical modeling? Blueprint Medicines
Cambridge, MA 21 May 2018	Redesigning drug design: Selectively targeting the kinome and beyond Pfizer
San Francisco, CA 12 Apr 2018	Redesigning drug design: Selectively targeting the kinome and beyond Pharmaceutical Chemistry Departmental Seminar, UCSF
Portland, OR 9 Feb 2018	Selective kinase inhibition and emergence of resistance mutations in cancer: Probing driving forces with physical modeling and automated biophysical experiments Biomedical Engineering Seminar, OHSU
Minneapolis, MN 14 Dec 2017	What will it take to predict kinase inhibitor selectivity and resistance using computational physical modeling? Department of Pharmacology, University of Minnesota
Stanford, CA 27 Oct 2017	How can physical modeling play a major role in drug discovery and biomedicine? Vijay Pande group, Stanford University
New York, NY 13 Oct 2017	Markov chain Monte Carlo and nonequilibrium statistical mechanics in drug discovery Applied Math Seminar, NYU Courant
Darmstadt, Germany 18 Sep 2017	Advancing drug design with alchemical free energy calculations Merck Darmstadt
Mountain View, CA 16 Jun 2017	Redesigning drug design Google Advanced Sciences
Rahway, NJ 9 May 2017	Advancing quantitative biophysical predictions with blind challenges Merck Formulations
San Francisco, CA 10 Apr 2017	Redesigning drug design Diane Barber lab, UCSF
Boston, MA 14 Feb 2017	Is there a use for nonequilibrium statistical mechanics in drug design? Harvard Widely Applied Math Seminar
Collegeville, PA 15 Dec 2016	Redesigning drug design GlaxoSmithKline
Bethesda, MD 8 Dec 2016	Redesigning drug design Bernie Brooks group, NHLBI National Institutes of Health
Stanford, CA 28 Oct 2016	Redesigning drug design Vijay Pande group, Stanford University
Barcelona, Spain 15 Sep 2016	Redesigning drug design Universitat Pompeu Fabra
Slough, UK 13 Sep 2016	Redesigning drug design UCB Pharma
Darmstadt, Germany 6 Sep 2016	Alchemical free energy calculations for drug discovery

	Merck KGaA
San Francisco, CA 29 Oct 2015	Redesigning drug design Autodesk
Rahway, NJ 8 Oct 2015	Redesigning drug design Merck
Cambridge, MA 24 Jul 2015	Redesigning drug design Novartis
Tokyo, Japan 3 Jun 2015	Redesigning drug design University of Tokyo / RCAST
Waltham, MA 15 May 2015	Redesigning drug design AstraZeneca
New Haven, CT 13 May 2015	Redesigning drug design Yale Theoretical Chemistry Seminar
Chicago, IL 15 Apr 2015	Redesigning drug design Illinois Institute of Technology
New Brunswick, NJ 8 Apr 2015	Making the most of limited data in biophysics: Challenges in single-molecule biophysics, nonequilibrium statistical mechanics, and drug discovery Rutgers – Department of Statistics and Biostatistics Seminar

## Meetings and Conferences

Toronto, Canada 31 Aug 2022	Enabling autonomous decisionmaking in drug discovery ACCELERATE
Barcelona, Spain 15 Jul 2022	Teaching free energy calculations to learn MLSMM
Boston, MA 2 Jun 2022	Next-generation technologies for structure-enabled drug design MPM Capital Scientific Advisors Meeting
San Diego, CA 23 May 2022	The COVID Moonshot: Open science discovery of a novel SARS-CoV-2 antiviral RECOMB
ZOOM 29 Apr 2022	Teaching free energy calculations to learn ICML MLDD
ZOOM 24 Mar 2022	New frontiers in machine learning for drug discovery Acceleration Consortium
ZOOM 20 Mar 2022	The future of free energy: Calculations that can learn from experiment ACS San Diego: Kate Holloway Award Symposium
Santa Fe, NM 10 Mar 2022	A Moonshot on theory and practice OpenEye CUP XXI
ZOOM 21 Apr 2021	The COVID Moonshot: An open science collaboration to develop an orally bioavailable inhibitor of SARS-CoV-2 main viral protease OpenEye Spring MiniCUP 2021
ZOOM 5 Apr 2021	The COVID Moonshot: An open science collaboration to develop an orally bioavailable inhibitor of SARS-CoV-2 main viral protease ACS Spring 2021 Meeting
ZOOM 16 Mar 2021	The COVID Moonshot: Closing in on an orally-bioavailable non-peptidomimetic small molecule inhibitor of SARS-CoV-2 Mpro with an open science collaboration Winter RosettaCon
ZOOM 26 Feb 2021	The COVID Moonshot: Closing in on an orally-bioavailable non-peptidomimetic small molecule inhibitor of SARS-CoV-2 Mpro with an open science collaboration Biophysical Society Annual Meeting
ZOOM 29 Jan 2021	OpenMM: Integrating ML to transform drug discovery NVIDIA Round Table on AI/ML Workloads for MD
ZOOM 26 Jan 2021	The COVID Moonshot: Closing in on an orally-bioavailable non-peptidomimetic small molecule inhibitor of SARS-CoV-2 Mpro with an open science collaboration CSHL COVID/SARS-CoV-2 Rapid Research Reports 5
ZOOM 8 Dec 2020	OpenMM: Key infrastructure for biomolecular modeling and simulation CZI Essential Open Source Software for Science
ZOOM 25 Nov 2020	Learning biomolecular potentials for drug discovery NeurIPS

ZOOM 5 May 2020	Future directions in force field science Open Force Field Initiative Virtual Meeting
Santa Fe, NM 14 Mar 2020	Analysis, Interpretation, and Presentation of Free Energy Results OpenEye CUP XX
Santa Fe, NM 13 Mar 2020	Update on the Open Force Field Initiative OpenEye CUP XX
Cambridge, MA 21 Nov 2019	Setting free energy calculations free: Opening up the free energy ecosystem OpenEye MiniCUP Cambridge MA
College Park, MD 17 Nov 2019	Powering the next generation of research in machine learning in chemistry with OpenMM, the Open Force Field Initiative, and Blind Predictive Challenges MoSSI Workshop in Machine learning in Chemistry
Mainz, Germany 5 Nov 2019	Looking to the future of drug discovery: Integrating free energy calculations, machine learning, and autonomous discovery to drive molecular design German Conference on Cheminformatics
Mainz, Germany 5 Nov 2019	Looking to the future of drug discovery: Integrating free energy calculations, machine learning, and autonomous discovery to drive molecular design German Conference on Cheminformatics
St Louis, MO 21 Oct 2019	Folding@home Consortium: Perspectives from the Chodera lab Folding@home Consortium Meeting, Washington University, St Louis
La Jolla, CA 23 Aug 2019	Looking forward to the future of assessment of predictive modeling and blinded prediction challenges SAMPL/D3R Workshop
Mount Snow, VT 15 Jul 2019	Alchemical free energy calculations: Differentiating approaches, measuring impact, and surveying the open source ecosystem Computer Aided Drug Discovery GRC
Telluride, CO 27 Jun 2019	Protonation state effects in small molecule recognition and design TSRC Protein Electrostatics
Berlin, Germany 19 Jun 2019	Molecular kinetics, free energy calculations, and machine learning for drug discovery MolKin 2019
Mainz, Germany 7 Jun 2019	Molecular kinetics, free energy calculations, and machine learning for drug discovery Mainz Materials Simulation Days 2019
Göttingen, Germany 27 May 2019	What's next for alchemical free energy calculations? Alchemical Free Energy Calculations in Drug Discovery 2019
New York, NY 26 Apr 2019	Redesigning drug design (together) Mid-Atlantic Computational Chemistry / NY Area Group in Informatics and Modeling
Santa Fe, NM 8 Mar 2019	Free energy calculations on Orion: Research and Development for 2019 OpenEye CUP XIX Free Energy Summit
Baltimore, MD 5 Mar 2019	Closing the loop in automated design and measurement: Scalable inference for biophysical experiments Biophysical Society Meeting: Symposium on Bayesian methods in biophysics
Stockholm 27 Nov 2018	Streamlining molecular simulation data BioExcel Workshop on Interoperability and Reproducibility in Molecular Simulations
Berlin 11 Nov 2018	Redesigning drug design Future Medicine Berlin
Tempe, AZ 7 Oct 2018	What can protein folding teach us about drug discovery? Statistical Physics in Biology
Vienna, Austria 2 Oct 2018	Rare events and large-scale conformational changes in drug discovery ECAM Workshop: Large scale activated event simulations
Biddeford, ME 12 Jun 2018	What can physical modeling tell us about mutational mechanisms of kinase inhibitor resistance? Human Genetic Variation GRC
Santa Fe, NM 8 Mar 2018	Testing the Friesner Conjecture: Are relative free energy calculations always more efficient than absolute? OpenEye CUP XVIII
San Diego, CA 22 Feb 2018	The SAMPL6 Challenges: Advancing molecular modeling accuracy with targeted blind challenges Drug Design Data Resource (D3R) Workshop
New York, MA 13 Nov 2017	Bridging scales in cancer Simons MPS Conference
Cambridge, MA 9 Nov 2017	Building next-generation forcefields and moving free energy calculations to the cloud OpenEye MiniCUP
Berlin, Germany 20 Sep 2017	Next-generation small molecule forcefields

	<b>RDKit User Group Meeting</b>
Gaithersberg, MD 24 Aug 2017	Challenges and opportunities for reproducibility and reliability in molecular simulations MolSSI Best Practices Workshop, NIST
Leiden, Netherlands 15 Aug 2017	Developing GPU-accelerated molecular simulation tools with OpenMM and the Omnia ecosystem ECAM Software Workshop on Path Sampling
Telluride, CO 14 Jul 2017	Toward realizing the dream of free-energy based small molecule design Telluride Free Energy Workshop
New York, NY 2 Jun 2017	What can physical modeling tell us about mutational mechanisms of kinase inhibitor resistance? NYU Genomics Symposium
Toronto, Canada 29 May 2017	The statistical mechanics of drug discovery Canadian Society of Chemistry 100th Anniversary National Meeting
Waltham, MA 3 May 2017	What can physical modeling tell us about mutational mechanisms of kinase inhibitor resistance? EMBL-EBI Informatics and -omics for Oncology Drug Resistance
Santa Fe, NM 8 Mar 2017	Can free energies really be free? Alchemical free energy calculations in the cloud OpenEye CUP XVII
Houston, TX 8 Oct 2016	Advancing Quantitative Biophysical Predictions Molecular Software Sciences Institute — Biomolecular Simulation
London, UK 14 Sep 2016	Redesigning drug design CCPBioSim — Kings College London
Philadelphia, PA 23 Aug 2016	What is required for alchemical free energy calculations to be useful for predicting drug polypharmacology? ACS Fall National Meeting
Philadelphia, PA 22 Aug 2016	Advancing quantitative biophysical predictions: What can be gained from industry-academic data sharing? ACS Fall National Meeting
Boston, MA 17 May 2016	How can we combine computation and experiment to move the field forward? Free Energy Calculations in Drug Discovery, Vertex
San Diego, CA 16 May 2016	Redesigning drug design ACS San Diego
Edinburgh, UK 4 Mar 2016	Redesigning drug design International Centre for Mathematical Sciences (ICMS)
La Jolla, CA 1 Mar 2016	What will it take to design kinase inhibitors with desired selectivity and resistance profiles? GTC 11th Protein Kinases in Drug Discovery Meeting
Edinburgh, UK 19 Oct 2015	Biomolecular software interoperability CECAM Software Interoperability Workshop
Vienna, Austria 18 Sep 2015	Reaction coordinates in drug discovery ESI Insight from Molecular Simulations
Berlin, Germany 8 Sep 2015	Redesigning drug design World Molecular Kinetics Meeting
Tokyo, Japan 5 Jun 2015	Redesigning drug design OpenEye JCUP VI
New York, NY 21 Jan 2015	Redesigning drug design New York Structural Biology Discussion Group

## Funding

All funding sources are publicly available online at <http://choderalab.org/funding>

### Active

#### Single-Investigator Grants

Vir Biosciences Sponsored Research Agreement SK2019-0582

##### **Development of alchemical free energy methods for predicting the impact of point mutations**

The goal of this project is to develop alchemical free energy methods for predicting the impact of point mutations on antibody:antigen affinities.

8/1/2019 – 7/31/2023

\$55,556 annual direct costs

Relay Therapeutics Collaboration SK2018-0162

##### **Development of efficient open source cloud-enabled free energy based lead optimization algorithms and integrative Bayesian model of experimental biophysical and molecular simulation data**

The goal of this project is to develop open source scalable cloud workflows for lead optimization using relative alchemical free energy calculations, as well as Bayesian integrative modeling techniques for experimental biophysical data and molecular simulations.

4/5/2018 – 4/4/2023

\$124,828 annual direct costs

Interline Therapeutics Sponsored Research Agreement SK2020-0898

##### **Identification of key biophysical interactions in ligase-substrate recognition**

The goal of this project is to identify key biophysical interactions necessary for ligase-substrate recognition and developing quantitative accurate open source tools for predicting changes in affinity associated with point mutations in either ubiquitin ligases or their substrate peptides.

1/15/2021 – 1/14/2023

\$187,998 direct costs

#### Collaborative Grants

NIH NIAID U19 1U19AI171399-01 (MPI: Chodera [Contact PI]; Alpha Lee, Potera; Ben Perry, DNDI)

**AI-driven Structure-enabled Antiviral Platform (ASAP)** co-Is: Karla Kirkegaard (Stanford), Matt Bogyo (Stanford), Jesse Bloom (Fred Hutch), Frank von Delft (Diamond Light Source), Martin Walsh (Diamond Light Source), Alpha Lee (PostEra), Nir London (Weizmann), Ed Griffen (MedChemica), Laurent Fraisse (DNDI), Haim Barr (Weizmann), Daren Fearon (Diamond Light Source), Kris White (Mount Sinai), Adolfo García-Sastre (Mount Sinai), Randy Albrecht (Mount Sinai)

05/16/2022 – 04/30/2025

*The goal of this antiviral drug discovery (AViDD) Center is to use an AI-accelerated structure-enabled discovery platform to build a pipeline of novel oral antivirals against coronaviruses, flaviviruses, and picornaviruses using an open science, IP-free discovery paradigm to act as a nexus for global antiviral discovery.*

Total federal funds awarded: \$67,662,387

Total direct costs: \$49,000,085

Total F&A for MSKCC: \$18,662,302 Eligible for non-competitive renewal for two more years for an additional \$42M.

NSF CHI 1904822 (PI: David Minh, IIT)

##### **Collaborative Research: CDS&E: Elucidating Binding using Bayesian Inference to Integrate Multiple Data Sources**

The goal of this project is to develop a scalable approach to Bayesian inference from biophysical experiments

9/1/2019 – 8/31/2023

\$47,327 annual direct costs to the Chodera lab

NIH R01 1R01GM132386 (MPI: John Chodera; Michael Shirts, University of Colorado [Contact PI])

##### **Open data-driven infrastructure for building biomolecular force fields for predictive biophysics and drug design**

co-Is: David Mobley (UCI), Lee-Ping Wang (UC Davis), Michael Gilson (UCSD)

The goal of this project is to develop a modern open infrastructure for building fully consistent biopolymer and small molecule force fields for drug design and predictive biophysical modeling.

6/1/2019 – 5/30/2023

\$20,684 annual direct costs to the Chodera lab

*Note:* Chodera was a PI of this multi-PI/PD grant for the first year, but moved to co-I status in subsequent years due to NIGMS limit two active GMS PI grants.

NIH R01 GM124270 (PI: David Mobley, UCI)

##### **Advancing predictive physical modeling through focused development of model systems**

### **to drive new modeling innovations**

The goal of this project is to enable blind community challenges that drive progress toward quantitative accuracy in the field of computational physical modeling of drug-receptor interactions.

9/10/2018 – 8/31/2023

\$42,711 annual direct costs to the Chodera lab

NIH 1R01GM140090 (PI: Tom Markland, Stanford University)

### **OpenMM: Scalable biomolecular modeling, simulation, and machine learning**

co-I: Gianni de Fabritiis (Universitat de Pompeu Fabra, Barcelona)

The goal of this project is to support continued development of OpenMM, transition to a community governance and distributed development model, and extension to integrate physical and machine learning models to enable genomic-scale biomolecular modeling, simulation, and prediction.

7/1/2021 – 3/31/2025

\$36,685 annual direct costs to the Chodera lab

BIH Einstein Visiting Professorship at the Charité Berlin

### **Computational polypharmacology: A new paradigm for selectively promiscuous kinase inhibitors**

The goal of this project is to develop structure-informed machine learning approaches to predicting kinase inhibitor polypharmacology for computer-guided small molecule design.

1/1/2019 – 12/31/2024

EUR150,000 annual direct costs to Chodera (via Charité) for Berlin working group (with six-month no-cost extension)

NIH 3R01GM132386-02S1 (PI: Michael Shirts, University of Colorado Boulder)

### **Supplement to: Open data-driven infrastructure for building biomolecular force fields for predictive biophysics and drug design**

co-Is: David Mobley (UCI), Lee-Ping Wang (UC Davis), Michael Gilson (UCSD)

The goal of this project is to support continued development of OpenMM, transition to a community governance and distributed development model, and extension to integrate physical and machine learning models to enable genomic-scale biomolecular modeling, simulation, and prediction.

3/1/2021 – 2/28/2022

\$132,694 annual direct costs for this supplement

*JDC wrote this proposal to secure additional funding for a critical community resource (QCArchive/QCFractal) he will help direct.*

### **Fellowships**

Sukrit Singh // Postdoctoral Fellow

### **Damon Runyon Quantitative Biology Fellowship**

5/1/2022 – 4/30/2025

\$240,000 stipend support

## Pending

### Single-Investigator Grants and Contracts

I plan to consolidate my NIH NIGMS funding ending 2022–2023 into an R35 (MIRA) for the Jan 2023 deadline.

### Collaborative Grants

NIH NCI R01 GM145739-01 (MPI: Chiosis/Gewirth)

4/1/2022 – 3/31/2027

#### Mechanistic investigations into Glyc62GRP94 epichaperones

*The role of the Chodera lab in this project is to use MD simulations to uncover the effect of the N-glycan on GRP94 structure and ligand binding.*

\$50,000 annual direct costs to the Chodera lab

NIH 1P01 CA262136-01A1 (PI: Li)

7/1/2022 -- 6/30/2027

#### GRP94 in Cancer Biology and Immuno-oncology (Project 1; Project leader: Chiosis)

*The role of the Chodera lab in this project is to use MD simulations to discover improved inhibitors that target this cancer-specific GRP94 variant.*

\$50,000 annual direct costs to the Chodera lab

### Fellowships

None

## Completed

### Single-Investigator Grants and Contracts

NIH R01 GM121505

#### **The role of reorganization energy in achieving selective kinase inhibition**

The goal of this project is to probe the role of protein reorganization energy in achieving selectivity in targeted kinase inhibition with small molecules.

9/1/2017 – 8/31/2022

\$196,250 annual direct costs to the Chodera lab

Bayer Sponsored Research Agreement SK2019-1289

#### **Development of a combined machine learning / alchemical free energy approach to predict the impact of mutations on small molecule kinase inhibitor binding**

The goal of this project is to develop open source scalable free energy and machine learning workflows in collaboration with the Volkamer lab (Charité Berlin) to predict kinase mutations small molecule therapeutics may be effective against.

3/1/2020 – 4/30/2022

\$117,105 annual direct costs to the Chodera lab

NSF RAPID CHE2033426

#### **RAPID: Identifying Biophysical Determinants of Binding to the SARS-CoV-2 Main Viral Protease**

The goal of this proposal is to assess the accuracy of large-scale alchemical free energy calculations in identifying potent inhibitors of the SARS-CoV-2 main viral protease (Mpro) with the COVID Moonshot.

7/1/2020 – 6/30/2021

\$112,994 direct costs to Chodera lab

Entasis Therapeutics Sponsored Research Agreement SK2018-0163

#### **Development of efficient equilibrium and nonequilibrium algorithms for predicting small molecule porin permeation with potential of mean force methods**

The goal of this project is to develop algorithms and open source tools for efficient equilibrium and nonequilibrium potential of mean force calculations for bacterial porins.

4/5/2018 – 4/5/2019

\$74,074 annual direct costs to the Chodera lab

Parker Institute for Cancer Immunotherapy Pilot Grant

#### **Physics-based computational prediction of cancer-associated mutant MHC class II epitopes**

The goal of this project is to develop algorithms and open source tools for efficient computation of MHC class II epitope binding affinities.

4/2/2018 – 4/1/2019

\$75,000 annual direct costs to the Chodera lab

NSF D3SC EAGER CTMC

#### **A probabilistic framework for automated forcefield parameterization from experimental datasets**

The goal of this project is to develop a scalable Bayesian inference infrastructure to parameterize molecular mechanics forcefields using experimental datasets.

9/1/2017 – 8/31/2019

\$52,030 annual direct costs to the Chodera lab

Silicon Therapeutics Sponsored Research Agreement

#### **Development of efficient open source free energy based lead optimization algorithms**

The goal of this project is to develop open source scalable parallel workflows for lead optimization using relative alchemical free energy calculations.

5/17/2017 – 11/17/2018

\$48,655 annual direct costs to the Chodera lab

Merck KGaA Sponsored Research Agreement

#### **Developing automated workflows for absolute alchemical free energy calculations**

The goal of this project is to develop open source automated workflows using absolute alchemical free energy calculations for use in prioritizing compounds for synthesis in drug discovery applications.

7/26/2016 – 1/25/2019

\$90,273 annual direct costs to the Chodera lab

Astra-Zeneca iMed Sponsored Research Agreement

#### **Evaluating the potential for Markov state models of conformational dynamics to advance quantitative prediction of thermodynamics and kinetics of selective kinase inhibitors**

The goal of this project is to evaluate the potential for Markov state models of conformational dynamics to quantitatively

predict the thermodynamics and kinetics of selective kinase inhibitors to CK2 and SYK.

7/30/2015 – 1/30/2017

\$117,505 annual direct costs to the Chodera lab

Gerstner Family Foundation

**Louis V. Gerstner, Jr. Young Investigator Award**

Dates: 2/1/2013 – 1/30/2016

\$75,000 annual direct costs

MSKCC Functional Genomics Initiative (FGI) Rapid Response Grant

**Biophysical characterization of clinically-identified kinase mutations**

The goal of this project is to quantify the impact of clinically-identified kinase domain mutations on FDA-approved selective kinase inhibitor binding.

4/1/2016 – 3/31/2016

\$25,000 annual direct costs

MSKCC Functional Genomics Initiative (FGI) Rapid Response Grant

**Biophysical characterization of clinically-identified K-Ras mutants**

The goal of this project is to biophysically characterize clinically-identified K-Ras mutants.

7/1/2016 – 6/30/2016

\$25,000 annual direct costs

## Fellowships

Dominic Rufa // TPCB Graduate Student

**MolSSI Seed Software Fellowship**

1/1/2022 – 6/30/2022

\$20,000 stipend support

Yuanqing Wang // PBSB Graduate Student

**Anagenex Open Science Fellowship**

9/1/2021 – 2/28/2022

\$21,750 stipend support

Karmen Čondić-Jurkić, Postdoctoral Research Fellow

**Open Force Field Consortium Postdoctoral Fellowship**

4/1/2020 – 3/31/2021

\$88,864.10 stipend, supplies, and travel support

Ivy Zhang CBM Graduate Student

**MolSSI COVID-19 Seed Software Fellowship**

7/1/2020 – 12/31/2020

\$20,000/year stipend support

Hannah Bruce Macdonald, Postdoctoral Fellow

**MolSSI Investment Software Fellowship**

7/1/2019 – 12/31/2020

\$50,000/year stipend support

Marcus Wieder, Postdoctoral Research Fellow

**Austrian Schrödinger Stipendium**

10/1/2018 – 9/30/2019

Fellowship recipient direct paid

Simon Boothroyd, Postdoctoral Research Fellow

**XtalPi / Open Force Field Consortium Distinguished Postdoctoral Fellowship**

10/1/2018 – 9/30/2019

Fellowship recipient direct paid

Andrea Rizzi, CBM Graduate Student

**MolSSI Open Force Field Predoctoral Fellowship**

1/1/2019 – 6/30/2019

\$20,000/six months stipend support

Rafal P. Wiewiora, TPCB Graduate Student

**DOD Peer-Reviewed Cancer Research Program Award**

8/15/2017 – 8/14/2019

\$75,000/year for stipend and supplies

Karmen Čondić-Jurkić, Postdoctoral Research Fellow  
Open Force Field Consortium Postdoctoral Fellowship

1/1/2019 – 6/30/2019

\$60,628 stipend and travel support

Chaya Stern, TPCB Graduate Student

**MolSSI Phase II Fellowship**

7/1/2018 – 12/31/2018

\$40,000/year stipend support

Chaya Stern, TPCB Graduate Student

**MolSSI Phase I Fellowship**

7/1/2018 – 12/31/2018

\$40,000/year stipend

Chaya Stern, TPCB Graduate Student

**NSF Graduate Research Fellowship**

8/1/2015 – 7/30/2018

\$32,000/year stipend support

Mehtap Isik, TPCB Graduate Student

**Doris J. Hutchison Fellowship**

7/1/2017 – 6/30/2018

\$38,668/year stipend support

Gregory A. Ross, Postdoctoral Fellow

**Molecular Software Sciences Institute (MolSSI) Graduate Research Fellowship**

\$50,000/year stipend support

Steven Albanese, GSK Graduate Student

**Summer Internship, Schrödinger**

Summer 2017

Ariën S. Rustenburg, PBSB Graduate Student

**Summer Internship, Genentech**

Summer 2015

Andrea Rizzi, CBM Graduate Student

**Tri-I CBM Student Stipend Assistance**

\$13,500/year stipend support

## Collaborative Grants

Chan Zuckerberg Initiative – Essential Open Source Software for Science (PI: Tom Markland, Stanford University)

co-I: Gianni de Fabritiis (Universitat de Pompeu Fabra, Barcelona)

**OpenMM: Key infrastructure for biomolecular modeling and simulation**

The goal of this project is to support continued development of OpenMM, transition to a community governance and distributed development model, and extension to integrate physical and machine learning models to enable genomic-scale biomolecular modeling, simulation, and prediction.

9/1/2020 – 4/1/2021

*JDC wrote this proposal to secure funding for a key community infrastructure project that his lab depends upon.*

\$180,000 total direct costs to support this project

Functional Genomics Initiative (FGI)

**Integrated Approaches Annotate Functions of Cancer-associated H3K36 Methyltransferases**

The goal of this project is to develop approaches for using large-scale physical modeling and biochemical experiments to annotate the functions of clinical cancer mutations.

1/1/2019 – 12/31/2020

co-PI: Minkui Luo

\$125,000 annual direct costs to the Chodera lab

2014 Functional Genomics Initiative (PI: James Hsieh, MSKCC)

**Characterization of Cancer-derived mTOR Mutations for Precision Therapeutics**

The goal of this project is to understand the mechanism underlying clinically-identified mTOR activating mutants and evaluate the potential for rapalog therapy to aid the 2% of cancer patients harboring mTOR mutations.

5/1/2015 – 4/30/2017

\$118,000

2014 STARR Cancer Consortium (PI: Minkui Luo, MSKCC)

**Designing sinefungin scaffolds as specific protein methyltransferase inhibitors**

The goal of this project is to use computational techniques that explicitly incorporate protein flexibility to design selective inhibitors for protein lysine methyltransferases.

1/1/2015 – 6/30/2017

\$125,000

## Personnel

### **Current members**

#### **Postdocs**

- Sukrit Singh, PhD Washington University St. Louis (6/1/2021–)  
Damon Runyon Quantitative Biology Fellow  
*Selective inhibition and drug resistance in clinical kinase inhibitors*

#### **Graduate Students**

- Ivy Zhang (8/1/2018–)  
Program in Physiology, Biophysics, and Systems Biology (PBSB)  
*Scalable free energy calculations for predicting the impact of mutations on protein-protein interactions*
- Yuanqing Wang (7/1/2018–)  
Program in Physiology, Biophysics, and Systems Biology (PBSB)  
*Machine learning of physics for drug discovery*
- Dominic Rufa (4/1/2019–)  
Tri-Institutional Program in Computational Biology and Medicine (CBM)  
*Machine learning for alchemical free energy calculations*
- Talia Kimber (4/1/2019–) [co-supervised with Andrea Volkamer, Charité Berlin]  
Charité Universitätmedizin (supervisor: Andrea Volkamer)  
Freie Universität Berlin, Computer Science PhD program (supervisor: Frank Noé)  
*Machine learning for optimizing affinity/selectivity and predicting resistance*
- Alexander Payne (8/1/2020–) [co-supervised with Rich Hite, MSKCC Structural Biology Program]  
Program in Physiology, Biophysics, and Systems Biology (PBSB)  
*Combining cryo-EM with Markov state models, virtual screening, and free energy calculations for drug discovery*
- Viktor Belay (8/1/2020–) [co-supervised with Rich Hite, MSKCC Structural Biology Program]  
Program in Physiology, Biophysics, and Systems Biology (PBSB)  
*Combining cryo-EM with Markov state models, virtual screening, and free energy calculations for drug discovery*
- Michael Retchin (7/1/2021–)  
Tri-Institutional Program in Computational Biology and Medicine (CBM)  
*New technologies for autonomous drug discovery*
- Benjamin Kaminow (1/1/2022–)  
Tri-Institutional Program in Computational Biology and Medicine (CBM)  
*Structure-enabled machine learning for drug discovery*

#### **Software Scientists**

- Iván Pulido (7/1/2021–)  
*Alchemical free energy calculation software infrastructure; software infrastructure support for ASAP*

#### **Research Associates**

- Erica Goldberger (7/1/2019–)  
*Lab manager; Automated measurements of small molecule binding affinities to model drug targets for improving quantitative predictions for drug discovery*

## Past trainees

My lab trains interdisciplinary scientists to tackle the most challenging problems in drug discovery.

### Postdocs

- David Schaller, PhD Freie Universität Berlin (7/1/2020–7/31/2022) (**Berlin group**)  
Current position: Associate Scientist, [Nuvisan](#)
- Ana Silveira, PhD Federal University of Rio de Janeiro (8/1/2018–7/31/2019)  
Current position: Senior Investigator, [Rovant Sciences](#)
- Jiaye Guo, PhD Stony Brook University (10/1/2018–5/31/2021)  
Current position: Senior Scientist I, [Schrodinger](#)
- Simon Boothroyd, PhD Lancaster University (11/1/2018–10/31/2019)  
Current position: Senior Researcher, Scientific Programmer, [Rovant Sciences](#)
- William Glass, PhD Southampton (5/1/2020–7/31/2021)  
Current position: Molecular Dynamics Research Scientist, [Exscientia](#)
- Jaime Rodríguez-Guerra Pedregal, PhD Autonomous University of Barcelona (4/1/2019–5/31/2021) (**Berlin group**)  
Current position: Software Engineer, [Quansight](#)
- Karmen Čondić-Jurkić, PhD Friedrich Alexander University, Erlangen (4/6/2019–)  
Current position: Executive Director and Co-Founder, [Open Molecular Software Foundation](#)
- Hannah Bruce Macdonald, PhD University of Southampton (10/1/2018–9/30/2020)  
Current position: Senior Scientist I, [Merck Sharp & Dohme \(MSD\)](#)
- Marcus Wieder, PhD University of Vienna (9/1/2018–4/16/2020)  
Current position: Faculty, [University of Vienna](#)
- Levi N. Naden (until 8/10/2018)  
Current position: Software Scientist at the NSF [Molecular Sciences Software Institute \(MolSSI\)](#)
- Sonya M. Hanson (until 6/30/2017)  
Next position: Postdoctoral Researcher with Joachim Frank, [Columbia University](#)  
Current position: Group Leader and Research Scientist, Center for Computational Biology, [Flatiron Institute](#), Simons Foundation, NYC
- Gregory Ross (until 6/30/2017)  
Current position: Principal Scientist, [Schrödinger](#)
- David W. H. Swenson (until 12/30/2015)  
Current position: Software Scientist, [Open Free Energy Consortium](#)
- Daniel L. Parton (until 8/31/2015)  
Current position: Lead Data Scientist, [Bardess Group](#)
- Kyle A. Beauchamp (until 6/12/2015)  
Current position: Senior Director of Data Science, [Tempus](#)
- Jan-Hendrik Prinz (until 1/30/2015)  
Current position: Digital Solutions Architect, [Keylight GmbH](#)
- Sarah E. Boyce (until 10/31/2013)  
Current position: Director, [Schrödinger](#)

### Graduate Students

- Joshua Fass (9/1/2015–11/9/2020)  
Tri-Institutional Program in Computational Biology and Medicine (CBM)  
Current position: Scientist, [Relay Therapeutics](#)
- Mehtap Isik (6/1/2015–9/11/2020 anticipated)  
Tri-Institutional Program in Chemical Biology (TPCB)  
Current position: Scientist, Computational Chemistry, [Moderna](#)
- Rafal Wiewiora (6/1/2015–8/31/2020)  
Tri-Institutional Training Program in Chemical Biology (TPCB)  
Current position: Scientist, [Rovant Sciences](#)
- Ariën Sebastian Rustenburg (10/1/2013–10/19/2018)  
Program in Physiology, Biophysics, and Systems Biology (PBSB)  
Current position: DevOps Engineer, Software for Chemistry & Materials, [SCM](#)

- Andrea Rizzi (9/1/2015–12/31/2019)  
Tri-Institutional Program in Computational Biology and Medicine (CBM)  
Current position: *Postdoctoral Fellow, Michele Parrinello Group, ETH Zürich*
- Chaya Stern (4/01/2014–6/30/2020)  
Tri-Institutional Training Program in Chemical Biology (TPCB)  
Current position: *Senior Principal Scientist, Odyssey Therapeutics*
- Steven Albanese (6/1/2015–3/31/2019)  
Gerstner Graduate Program (GSK)  
Current position: *Principal Scientist II, Schrödinger*
- Patrick B. Grinaway (7/1/2013–12/31/2018)  
Program in Physiology, Biophysics, and Systems Biology (PPSB)  
Current position: *Scientist, Onai*
- Julie M. Behr (1/10/2014–12/31/2016)  
Tri-Institutional Program in Computational Biology and Medicine (CBM)  
Current position: *Transferred to Imielinski lab after publishing three papers with the Chodera lab*

### **Technicians**

- Liza Casella (1/19/2021–11/6/2021)  
Current position: *Research Associate, Kingdom Supercultures*
- Binisha Karki (8/1/2019–8/28/2020)  
Current position: *Research Associate, BioNTech*
- Lucelenie Rodriguez (7/1/2015–12/31/2016; 12/4/2017–7/20/2018)  
Current position: *Medical student, NYU School of Medicine*
- Erin Grundy (6/1/18–7/31/2019)  
Current position: *Graduate student, George Washington University*