

John D. Chodera



email john.chodera@choderlab.org
url <http://www.choderlab.org>
tel 646.888.3400
fax 510.280.3760
mobile 646.737.3319
post Memorial Sloan Kettering Cancer Center
1275 York Ave, Box 357
New York, NY 10065

Education and positions

- 2013- **Assistant Professor, Physiology, Biophysics, and Systems Biology Program, Weill Cornell Graduate School of Medical Sciences**
- 2012- **Assistant Member, Memorial Sloan-Kettering Cancer Center**
- 2008-2012 **Independent Distinguished Postdoctoral Fellow, California Institute for Quantitative Biosciences (QB3), University of California, Berkeley**
Independent research funding, sponsors [Phillip L. Geissler](#) and [Susan Marqusee](#)
- 2006-2008 **Postdoctoral researcher, Department of Chemistry, Stanford University**
With [Vijay S. Pande](#) (head of [Folding@Home](#) distributed computing project)
- 1999-2006 **PH.D. in Biophysics, University of California, San Francisco**
Committee: [Ken A. Dill](#), [Matthew P. Jacobson](#), [Vijay S. Pande](#)
- 1995-1999 **B.S. in Biology, California Institute of Technology**
Undergraduate research with [Paul H. Patterson](#) (*experimental molecular neurobiology*)
and Jerry E. Solomon (*computational chemistry*)

Fellowships and awards

- 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
- 2005 Frank M. Goyan Award for outstanding work in physical chemistry, University of California, San Francisco
- 2000-2005 Howard Hughes Medical Institute Predoctoral Fellowship

Research interests

- Rational computational design of allosteric small-molecule inhibitors and therapeutics
Kinase inhibitor selectivity and evolution of therapeutic resistance in cancer
Biomolecular dynamics and conformational heterogeneity
Error and uncertainty in biophysical measurements
Computational chemistry, molecular modeling, and forcefield development
Multiscale modeling of the effects of small molecules on biochemical pathways

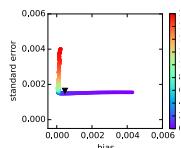
Publications

Google Scholar statistics: <http://goo.gl/qO0JW>

h-index: 29 / i10-index: 42

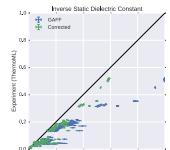
* asterisk denotes equal contributions

PUBLISHED AND IN PRESS



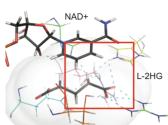
Chodera JD. A simple method for automated equilibration detection in molecular simulations. *J. Chem. Theor. Comput.*, in press. · [bioRxiv preprint](#)

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



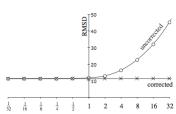
Beauchamp KA, Behr JM, Rustenburg AS, Bayly CI, Kroenlein K, and Chodera JD. Towards automated benchmarking of atomistic forcefields: Neat liquid densities and static dielectric constants from the ThermoML data archive. *J. Phys. Chem. B*, 199:12912, 2015 · [DOI](#)

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.



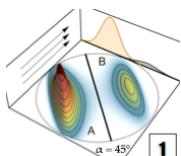
Intlekofer AM, Dematteo RG, Venetti S, Finley LWS, Lu Chao, Judkins AR, Rutenburg AS, Grinaway PB, Chodera JD, Cross JR, and Thompson CB. Hypoxia introduces production of L-2-Hydroxyglutarate. *Cell Metabolism* 22:1–8, 2015 · [DOI](#)

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



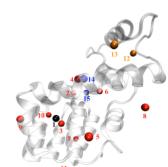
Sivak DA, Chodera JD, and Crooks GE. Time step rescaling recovers continuous-time dynamical properties for discrete-time Langevin integration of nonequilibrium systems. *J. Phys. Chem. B*, 118:6466–6474, 2014. William C. Swope Festschrift · [DOI](#)

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



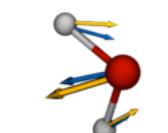
Prinz J-H, Chodera JD, and Noé F. Spectral rate theory for two-state kinetics. *Phys. Rev. X* 4:011020, 2014 · [DOI](#)

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.



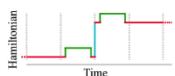
Wang K, Chodera JD, Yang Y, and Shirts MR. Identifying ligand binding sites and poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics. *J. Comput. Aid. Mol. Des.*, 27:989–1007, 2013 · [DOI](#)

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.



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. *J. Phys. Chem. B*, 117:9956–9972, 2013 · [DOI](#)

A new inexpensive polarizable model of liquid water for next-generation forcefields is derived using an automated parameterization engine.



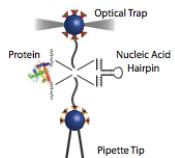
Sivak DA, Chodera JD, and Crooks GE. Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium discrete Langevin dynamics simulations. *Phys. Rev. X.*, 3:011007, 2013 · [DOI](#)

The finite-timestep errors in molecular dynamics simulations can be interpreted as a form of nonequilibrium work. We show how this leads to straightforward schemes for correcting for these errors or assessing their impact.



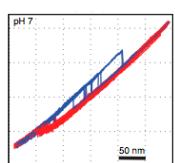
Eastman P, Friedrichs MS, Chodera JD, Radmer RJ, Bruns CM, Ku JP, Beauchamp KA, Lane TJ, Wang L, Shukla D, Tye T, Houston M, Stich T, Klein C, Shirts MR, and Pande VS. OpenMM 4: A reusable, extensible, hardware independent library for high performance molecular simulation. *J. Chem. Theor. Comput.*, 9:461, 2012 · [DOI](#)

We describe the latest version of an open-source, GPU-accelerated library and toolkit for molecular simulation.



Elms PJ, Chodera JD, Bustamante CJ, Marqusee S. The limitations of constant-force-feedback experiments. *Biophys. J.*, 103:1490, 2012 · [DOI](#)

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.



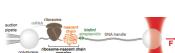
Elms PJ, Chodera JD, Bustamante C, Marqusee S. The molten globule state is unusually deformable under mechanical force. *Proc. Natl. Acad. Sci. USA.*, 109:3796, 2012 · [DOI](#)

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.

$$f_{\text{exp}} = \int d\vec{r} f(\vec{r}) p_1(\vec{r})$$

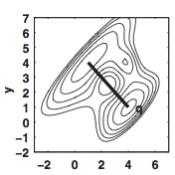
Pitera JW and Chodera JD. On the use of experimental observations to bias simulated ensembles. *J. Chem. Theor. Comput.*, 8:3445, 2012 · [DOI](#)

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.



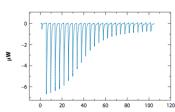
Kaiser CM, Goldman DH, Chodera JD, Tinoco I, Jr., and Bustamante C. The ribosome modulates nascent protein folding. *Science*, 334:1723, 2011 · [DOI](#)

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



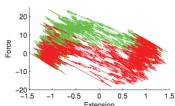
Chodera JD and Pande VS. Splitting probabilities as a test of reaction coordinate choice in single-molecule experiments. *Phys. Rev. Lett.*, 107:098102, 2011 · [DOI](#)

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



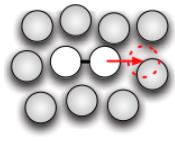
Tellinghuisen JT and Chodera JD. Systematic errors in isothermal titration calorimetry: Concentrations and baselines. *Anal. Biochem.*, 414:297–299, 2011 · [DOI](#)

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



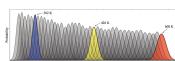
Minh DDL, Chodera JD. Estimating equilibrium ensemble averages using multiple time slices from driven nonequilibrium processes: Theory and application to free energies, moments, and thermodynamic length in single-molecule pulling experiments. *J. Chem. Phys.* 134:024111, 2011 · [DOI](#)

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.



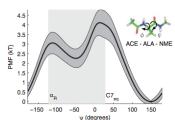
Nilmeier JP, Crooks GE, Minh DDL, and Chodera JD. Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation. *Proc. Natl. Acad. Sci. USA.*, 108:E1009, 2011 · [DOI](#)

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.



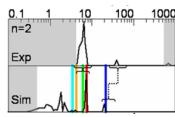
Prinz J-H, Chodera JD, Pande VS, Smith JC, and Noé F. Optimal use of data in parallel tempering simulations for the construction of discrete-state Markov models of biomolecular dynamics. *J. Chem. Phys.* 134:244108, 2011 · [DOI](#)

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



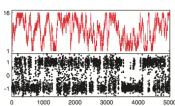
Chodera JD, Swope WC, Noé F, Prinz J-H, Shirts MR, and Pande VS. Dynamical reweighting: Improved estimates for dynamical properties from simulations at multiple temperatures. *J. Chem. Phys.* 134:244107, 2011 · [DOI](#)

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



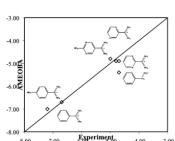
Noé F, Doose S, Daidone I, Löllmann M, Sauer M, Chodera JD, and Smith JC. Dynamical fingerprints: A theoretical framework for understanding biomolecular processes by combination of simulation and kinetic experiments. *Proc. Natl. Acad. Sci. USA*, 108:4822, 2011 · [DOI](#)

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.



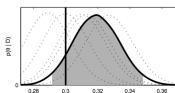
Chodera JD and Shirts MR. Replica exchange and expanded ensemble simulations as Gibbs sampling: Simple improvements for enhanced mixing. *J. Chem. Phys.*, 135:194110, 2011 · [DOI](#)

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.



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 GNI, Johnson ME, and Head-Gordon T. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* 114:2549, 2010 · [DOI](#)

A report on the status of the AMOEBA polarizable force field and its ability to reproduce a diverse set of physical chemical phenomenon to high accuracy.



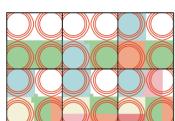
Chodera JD and Noé F. Probability distributions of molecular observables computed from Markov models. II. Uncertainties in observables and their time-evolution. *J. Chem. Phys.* 133:105102, 2010 · [DOI](#)

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



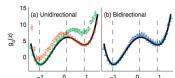
Adelman JL, Chodera JD, Kuo IW, Miller TF, and Barsky D. The mechanical properties of PCNA: Implications for the loading and function of a DNA sliding clamp. *Biophys. J.* 98:3062, 2010 · [DOI](#)

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.



Bacallado S, Chodera JD, and Pande VS. Bayesian comparison of Markov models of molecular dynamics with detailed balance constraint. *J. Chem. Phys.* 131:045106, 2009 · [DOI](#)

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.



Minh DDL, Chodera JD. Optimal estimators and asymptotic variances for nonequilibrium path-ensemble averages. *J. Chem. Phys.* 131:134110, 2009 · [DOI](#)

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.

Shirts MR, Chodera JD. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* 129:124105, 2008 · [DOI](#)

$$\hat{f}_i = -\ln \sum_{j=1}^K \frac{N_i}{\sum_{k=1}^K N_k} \exp[-u_i(x_{jk})]$$

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.



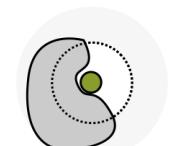
Nicholls A*, Mobley DL*, Guthrie JP, Chodera JD, and Pande VS. Predicting small-molecule solvation free energies: A blind challenge test for computational chemistry. *J. Med. Chem.* 51:769–779, 2008 · [DOI](#)

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.



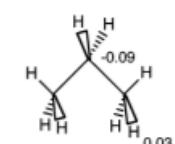
Mobley DL, Dill KA, and Chodera JD. Treating entropy and conformational changes in implicit solvent simulations of small molecules. *J. Phys. Chem. B* 112(3):938–946, 2008 · [DOI](#)

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



Shirts MR*, Mobley DL*, Chodera JD, and Pande VS. Accurate and efficient corrections for missing dispersion interactions in molecular simulations. *J. Phys. Chem. B* 111(45):13052–13063, 2007 · [DOI](#)

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.



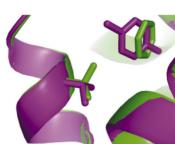
Mobley DL, Dumont E, Chodera JD, Bayly CI, Cooper MD, and Dill KA. Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. *J. Phys. Chem. B* 111:2242–2254, 2007 · [DOI](#)

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



Mobley DL, Graves AP, Chodera JD, McReynolds AC, Shoichet BK, and Dill KA. Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* 371(4):1118–1134, 2007 · [DOI](#)

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.



Mobley DL, Chodera JD, and Dill KA. Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theor. Comput.* 3(4):1231–1235, 2007 · [DOI](#)

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



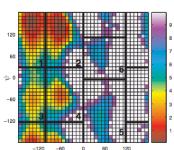
Chodera JD*, Singhal N*, Swope WC, Pitera JW, Pande VS, and Dill KA. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J. Chem. Phys.* 126:155101, 2007 · [DOI](#)

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 the equilibrium populations of these states may too small for standard structural biology techniques to detect.



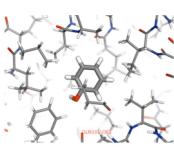
Ozkan SB, Wu GA, Chodera JD, and Dill KA. Protein Folding by Zipping and Assembly. *Proc. Natl. Acad. Sci. USA* 104(29):11987–11992, 2007 · [DOI](#)

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.



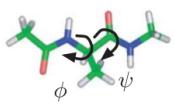
Chodera JD, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill. Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J. Chem. Theor. Comput.* 3(1):26–41, 2007 · [DOI](#)

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.



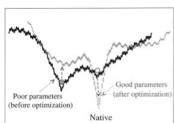
Mobley DL, Chodera JD, and Dill KA. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* 125:084902, 2006 · [DOI](#)

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.



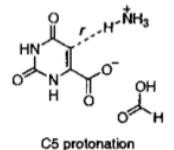
Chodera JD, Swope WC, Pitera JW, and Dill KA. Long-time protein folding dynamics from short-time molecular dynamics simulations. *Multiscale Model. Simul.* 5(4):1214–1226, 2006 · [DOI](#)

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.



Seok C, Rosen JB, Chodera JD, Dill KA. MOPED: Method for optimizing physical energy parameters using decoys. *J. Comput. Chem.* 24(1):89–97, 2003 · [DOI](#)

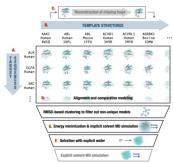
We propose a novel way to optimize parameters for a physical energy function for protein folding studies by making use of 'decoy' structures.



Lee TS*, Chong LT*, Chodera JD, and Kollman PA. An alternative explanation for the catalytic proficiency of orotidine 5'-phosphate decarboxylase. *J. Am. Chem. Soc.* 123(51):12837–12848, 2001 · [DOI](#)

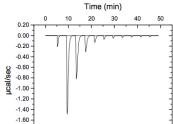
A combined QM and MD analysis of potential plausible mechanisms to explain the enormous catalytic acceleration of one of the most proficient enzymes known.

Submitted and Under Review



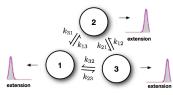
Parton DL, Grinaway PB, Hanson SM, Beauchamp KA, and Chodera JD. Ensembler: Enabling high-throughput molecular simulations at the superfamily scale · [bioRxiv](#)

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.



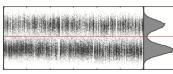
Boyce SE, Tellinghuisen JT, and Chodera JD. Avoiding accuracy-limiting pitfalls in the study of protein-ligand interactions with isothermal titration calorimetry · [bioRxiv](#)

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.



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 · [arXiv](#)

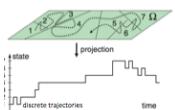
We present a Bayesian hidden Markov model analysis scheme that allows biomolecular conformational dynamics to be inferred from single-molecule trajectories.



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 · [arXiv](#)

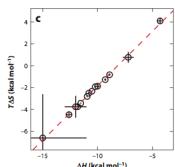
We present a simple, robust approach to estimating two-state rate constants from experimental or simulation data.

Reviews and Commentaries



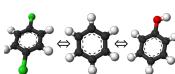
Chodera JD and Noé F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.*, 25:135–144, 2014 · [DOI](#)

A review of the exciting developments in the stochastic modeling of biomolecular dynamics over the last few years.



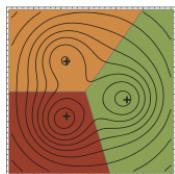
Chodera JD and Mobley DL. Entropy-enthalpy compensation: Role and ramifications for rational ligand design. *Annu. Rev. Biophys.*, 42:121, 2013 · [DOI](#)

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



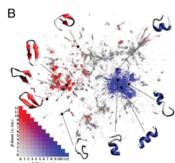
Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson KM, and Pande VS. Free energy methods in drug discovery and design: Progress and challenges. *Curr. Opin. Struct. Biol.*, 21:150–160, 2011 · [DOI](#)

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



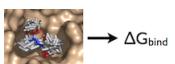
Prinz JH, Wu H, Sarich M, Keller B, Fischbach M, Held M, **Chodera JD**, Schütte, and Noé F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* 134:174105, 2011 · [DOI](#)

A review of current best practices for the generation and validation of Markov state models for describing the stochastic dynamics of biomolecular systems.



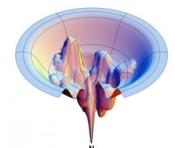
Chodera JD and Pande VS. The Social Network (of protein conformations). *Proc. Natl. Acad. Sci. USA* 108:12969, 2011 · [DOI](#)

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



Shirts MR, Mobley DL, **Chodera JD**. Alchemical free energy calculations: Ready for prime time? *Ann. Rep. Comput. Chem.* 3:41–59, 2007 · [DOI](#)

A review of current alchemical free energy methodologies assessing whether they are ready for practical use in drug discovery and ligand design.



Dill KA, Ozkan SB, Weikl TR, **Chodera JD**, and Voelz VA. The protein folding problem: When will it be solved? *Curr. Opin. Struct. Biol.* 17(3):342–346, 2007 · [DOI](#)

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

Scientific Advisory Boards

2013- Schrödinger, LLC

Conferences organized

BOSTON MAY 2016	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Boston, MA
CAMBRIDGE, MA MAY 2016	Markov State Models in Drug Design Workshop Novartis, Cambridge, MA
BERLIN SEP 2015	World Molecular Kinetics Workshop 2015 Freie Universität Berlin, Germany
BOSTON MAY 2014	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Boston, MA
BERLIN SEP 2013	World Molecular Kinetics Workshop 2013 Freie Universität Berlin, Germany
CAMBRIDGE, MA MAY 2012	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Cambridge, MA
BERLIN SEP 2011	World Molecular Kinetics Workshop 2011 Freie Universität Berlin, Germany
CAMBRIDGE, MA MAY 2010	Free Energy Methods in Drug Design Workshop Vertex Pharmaceuticals, Cambridge, MA
BERLIN MAY 2009	World Molecular Kinetics Workshop 2009 Freie Universität Berlin, Germany

Peer reviewer for scientific journals

Bioinformatics
Biopolymers
Chemical Physics
European Biophysics Journal
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
Journal of the American Chemical Society
Molecular Physics
Nature Physics
Nature Communications
Pacific Symposium in Biocomputing
PLoS Computational Biology
PLoS One
Proceedings of the National Academy of Sciences
Structure

Funding

ACTIVE

Single-Investigator Grants and Contracts

Gerstner Family Foundation

Louis V. Gerstner, Jr. Young Investigator Award

lead PI: John Chodera

direct costs: \$75,000/year

02/01/2013–01/30/2016

SK2015-0252. AstraZeneca 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

lead PI: John Chodera

direct costs: \$100,203 / 15 months

07/30/2015–01/30/2017

Functional Genomics Initiative Rapid Response Grant

Automated biophysical characterization of clinically observed kinase point mutations on inhibitor affinity

lead PI: John Chodera

direct costs: \$25,000/year

05/01/2015–04/30/2016

Functional Genomics Initiative Rapid Response Grant

Probing the biophysics of rare RAS mutations

lead PI: John Chodera

direct costs: \$25,000/year

07/01/2015–06/30/2015

Fellowships

Chaya Stern, TPCB graduate student

NSF GSRA

direct costs: \$32,000/year stipend

08/01/2015–07/30/2018

Collaborative Grants

Starr Center Cancer Consortium

Designing Sinefungin Scaffolds as Specific Protein Methyltransferase Inhibitors

lead PI: Minkui Luo

direct costs to laboratory: \$100,000/year

01/01/2015–12/31/2016

Functional Genomics Initiative Grant

Characterizing cancer-driven mTOR mutations in the MSKCC CMO database for precision therapeutics

lead PI: James Hsieh

direct costs to laboratory: \$59,000/year

05/01/2015–04/30/2017

PENDING

Single-Investigator Grants and Contracts

NIH DP2 New Investigator Award

Predicting patient-specific resistance and sensitivity to targeted kinase inhibitors

lead PI: John Chodera

direct costs: \$300,000/year

09/01/2016–08/31/2021

Targeted Grants in the Mathematical Modeling of Living Systems

Small molecule biomolecular interactions from automated model system mining

lead PI: John Chodera

direct costs: \$200,000/year

09/01/2016–08/31/2019

Fellowships

Ariën S. Rustenburg, PBSB graduate student

HHMI International Student Research Fellowship

direct costs: \$43,000/year stipend

09/01/2016–08/31/2019

Steven Albanese, GSK graduate student

NSF Graduate Research Fellowship Program

direct costs: \$34,000/year stipend

08/01/2016–07/31/2019

Joshua Fass, CBM graduate student

NSF Graduate Research Fellowship Program

direct costs: \$34,000/year stipend

08/01/2016–07/31/2019

Collaborative Grants

NSF PD 09-6881 / CTMC / CHE - Theory, Models, and Computational Methods

A probabilistic framework for automated force field parameterization, refinement, and systematic error quantification

lead PI: John Chodera

direct costs: total \$200,325; \$85,854/year

05/01/2016 – 04/30/2019

Simons Collaborations in Mathematics and the Physical Sciences

Statistical mechanics and statistical inference for systems driven out of equilibrium

lead PI: Aaron Dinner (University of Chicago)

direct costs: \$2,500,000/year (shared among all investigators)

09/01/2016–08/31/2020

NIH NCI R01 resubmission

Characterization of cancer-derived mTOR mutations for precision therapeutics

lead PI: James Hsieh

direct costs: \$262,915/year (shared among all investigators)

04/01/2016–03/30/2020

COMPLETED

Single-Investigator Grants and Contracts

None

Fellowships

None

Collaborative Grants

None

Personnel

CURRENT MEMBERS

Postdocs

- Sonya M. Hanson, PhD Oxford (01/01/2015–)
Understanding the role of conformational reorganization energy in targeted kinase inhibitor selectivity
- David Swenson, PhD UC Berkeley (09/21/2015–)
Molecular modeling of kinase activating mutations

Technicians

- Lucelenie Rodriguez (07/01/2015–)
Automated measurements of small molecule kinase inhibitor affinities to mutant kinases

Graduate Students

- Patrick B. Grinaway (07/01/2013–)
Program in Physiology, Biophysics, and Systems Biology (PBSB)
Designing allosteric modulators of Ras; automated lead optimization using synthetic accessibility
- Ariën Sebastian Rustenburg (10/01/2013–)
Program in Physiology, Biophysics, and Systems Biology (PBSB)
Predicting human serum albumin binding; protonation states in selective kinase inhibitors
- Julie M. Behr (01/10/2014–)
Tri-Institutional Program in Computational Biology and Medicine (CBM)
Predicting kinase resistance mutations in response to targeted therapies
- Chaya Stern (04/01/2014–)
Tri-Institutional Training Program in Chemical Biology (TPCB)
Understanding biased agonism and allosteric signal transduction in GPCRs
- Mehtap Isik (06/01/2015–)
Tri-Institutional Training Program in Chemical Biology (TPCB)
Model systems for protein-small molecule recognition
- Rafal Wiewiora (06/01/2015–)
Tri-Institutional Training Program in Chemical Biology (TPCB)
Design of selective inhibitors for protein methyltransferases
- Steven Albanese (06/01/2015–)
Gerstner Graduate Program (GSK)
Biophysical determinants of activating mutants in mTOR
- Andrea Rizzi (09/01/2015–)
Tri-Institutional Program in Computational Biology and Medicine (CBM)
Quantitatively accurate alchemical free energy calculations for kinase inhibitor design
- Joshua Fass (09/01/2015–)
Tri-Institutional Program in Computational Biology and Medicine (CBM)
Machine learning approaches to modeling conformational dynamics and optimizing cell-free expression

PAST TRAINEES

Postdocs

- Sarah E. Boyce (postdoc for six months in 2012 at UC Berkeley; 08/01/2013–10/31/2013 at MSKCC)
Current position: *Senior Scientist, Schrödinger*
- Jan-Hendrik Prinz (postdoc 02/01/2014–01/30/2015)
Current position: *Postdoctoral Fellow, Freie Universität Berlin*
- Kyle A. Beauchamp (postdoc 07/28/2013–06/12/2015)
Current position: *Data Scientist, Counsyl*
- Daniel L. Parton (postdoc 11/01/2012–08/31/2015)
Current position: *Data Scientist, Annalect*

Graduate Students

None

Statement of Research

OVERVIEW

My laboratory at MSKCC uses computation and automated biophysical experiments to develop new algorithms and approaches to the rational design of small molecules that selectively modulate cellular pathways. We develop new GPU-accelerated algorithms for alchemical free energy calculations [1, 2, 3, 4] in order to predict and optimize the affinity and selectivity of small molecules for their intended biomolecular targets. We have designed and built an automated biophysical wetlab to test these predictions, study the evolution of mutational resistance in targeted kinase inhibitors, and refine computational algorithms and forcefields. In addition, we employ the Folding@home worldwide distributed computing platform—where hundreds of thousands of donors around the world contribute their computing power [5, 6]—to study the conformational dynamics of cancer targets (such as kinases, small GTPases like Ras, and protein methyltransferases) in order to understand function, disease, and opportunities for allosteric modulation with small molecules. Major collaborations with both academic laboratories and pharmaceutical companies use computation and biophysical experiment to explore various aspects of biomolecular dynamics, drug pharmacokinetics, design of biological therapeutics, new biophysical measurement and screening technologies, and emerging targets.

BRIEF INVESTIGATOR BACKGROUND

I received my **B.S. in Biology from Caltech**, which included a broad set of elective coursework composed of a mixture of mathematics (both pure and applied), physics (quantum, classical, and statistical mechanics), physical chemistry, computer science, electrical engineering, and machine learning. As a Ph.D. student in the **Biophysics Ph.D. program at the University of California, San Francisco**, I utilized these quantitative tools to tackle various problems in protein folding in biomolecular dynamics. Working with Peter Kollman, we probed the mechanism behind the incredible catalytic proficiency of orotidine decarboxylase [7]. In the laboratory of Ken Dill, I helped pioneer an approach that allows the long-time stochastic conformational dynamics of biological macromolecules to be understood from many molecular dynamics simulations as part of a collaboration with Vijay Pande (Stanford), Frank Noé (Freie Universität Berlin), and William C. Swope (IBM Almaden Research Center) [8, 9, 10, 11, 12, 13]. The resulting *Markov model* formalism allows direct connections between atomistic simulations and experimental measurements in a manner that provides mechanistic insight [8, 9], and is now used by laboratories throughout the world to probe diverse aspects of protein dynamics and biophysics; our key 2007 paper [9] has been cited over 360 times.

Moving to the California Institute for Quantitative Biosciences (QB3) at the University of California, Berkeley as an independent **QB3-Berkeley Distinguished Postdoctoral Fellow** provided autonomy to explore new research interests and extensive opportunities for collaboration. There, I focused on the development of new algorithms, tools, and experimental techniques to enable the quantitative prediction of small molecule binding affinities to biomolecular targets, with the goal of enabling rational computational design of new selective small molecule using computational techniques that directly compute the binding affinity. Building on my earlier work in which I explored challenges to the quantitative prediction of free energies of binding using atomistic forcefields, leading us to realize the importance of the contribution of multiple ligand orientations [14], protein conformational changes [15], electrostatics and polarization [16], solvent far from the binding site [17], and conformational entropy effects in solvation [18], I developed a new GPU-accelerated binding free energy approach [4] and a framework to astronomically increase the efficiency of molecular simulations by making use of new nonequilibrium theorems in a paper that was published in *PNAS* [19].

Working with the thriving single-molecule biophysics community at UC Berkeley, I also developed new approaches for extracting fundamental biomolecular insights from breakthrough optical force spectroscopy instrumentation, leading to a number of joint publications [20, 21, 22, 23], including a study of the role of the ribosome in modulating co-translational folding published in *Science* [24].

PRIMARY RESEARCH PROJECTS

Selective kinase inhibition and resistance

Selective kinase inhibitors—such as the blockbuster drug imatinib (Figure 1, left)—have shown tremendous promise in the treatment of cancers involving kinase dysregulation. Currently, 27 small molecule targeted kinase inhibitors have received FDA approval, representing a substantial fraction of the \$37B U.S. market for oncology drugs. Despite this, major challenges remain in their widespread application in cancer treatment. First, the design of potent inhibitors with desired selectivity profiles is difficult and expensive; existing approaches—based on screening targeted libraries followed by rounds of structure-guided lead optimization—generally require several years and hundreds of millions of dollars to produce a clinical candidate [25]. Second, these molecules are still generally poorly selective (Figure 1), with the risk that off-target effects will limit therapeutic index. Third, the often-rapid emergence of drug resistance can eliminate therapeutic benefit; over 90% of metastatic cancer deaths are believed to result from drug-resistant forms [26]. To meet these challenges, our laboratory develops quantitative physical models of kinase inhibitor efficacy to accelerate the rational design of kinase inhibitors with desired selectivity profiles, an understanding of mutational mechanisms of resistance, and prediction of drug sensitivity and resistance in individual patient tumors.

Design of kinase inhibitors with desired selectivities. Recent evidence suggests that a major effect in the selectivity of imatinib for Abl over the closely-related Src is the larger *conformational reorganization energetic cost* that Src must pay in order to adopt a binding-competent conformation [27, 28, 29]. While failure to account for this contribution will frustrate the ability rationally engineer kinase inhibitors with desired selectivity profiles, it is difficult to measure these reorganization energies experimentally (Figure 2).

We use massively parallel molecular dynamics simulations on the Folding@home worldwide distributed computing platform [5]—the planet’s largest computer for biology—to computationally map the conformations accessible to kinases and their associated energetics. We have developed a tool (*Ensembler* [31]) to automate the modeling of human kinase catalytic domains onto all available kinase structures to broadly capture the variety of active and inactive conformations available to kinases, which are well-known to be highly conformationally heterogeneous. We then employ the Markov state modeling (MSM) approach I originally developed to study protein folding [8, 9, 13] to extract distinct conformational states, energetics, and activation kinetics. We are currently studying a number of kinases of interest to cancer and other diseases—ABL, SRC, MEK, MTOR, AURKA, DDR1, CK2, and SYK—and plan to extend this approach to all human kinases over the next two years.

To exploit these conformational and energetic maps of kinase conformations to understand kinase inhibitor affinity and selectivity, as well as facilitate the engineering of new selective inhibitors, we have developed a new generation of quantitatively accurate physical models based on alchemical free energy calculations [3]. These methods use molecular mechanics forcefields to rigorously compute binding affinities that include all relevant statistical mechanical effects, including entropic and enthalpic contributions. To make these techniques fast enough for practical use, we have built a new open-source GPU-accelerated code YANK [4]. YANK combines numerous algorithmic improvements we have previously developed to achieve quantitative accuracy [14, 15, 17, 9, 32, 13, 19, 33] into a single code, and has an open architecture that allows critically-needed new algorithms to deal with kinase-specific effects to be easily developed, implemented, and validated.

In addition to allowing for accurate inclusion of the reorganization energy for different kinases, our kinase MSMs also reveal opportunities for allosterically targeting kinases by identifying conformations that expose otherwise cryptic binding sites that are not visible in crystallographic structures unless an allosteric inhibitor has already been identified and cocrystallized [34]. As few allosteric kinase inhibitors have been identified to date, this represents an emerging opportunity for engineering new classes of effective therapeutics for cancer.

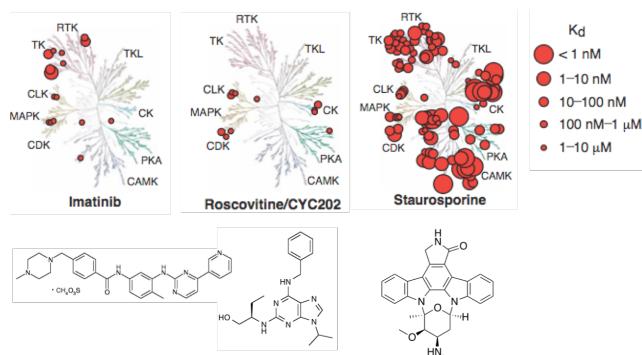


Figure 1: Kinase inhibitor selectivities are highly variable. Even targeted inhibitors like imatinib (left) can inhibit a multitude of kinases; unselective inhibitors (right) are highly toxic. Adapted from [30].

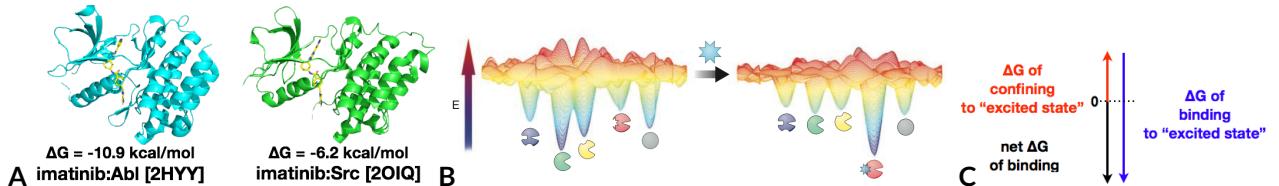


Figure 2: Conformational selection of kinase inactive states as a mechanism for inhibitor selectivity. (A) Crystal structures of imatinib bound to Abl and Src show almost no differences in binding contacts or structure, despite the large 4.7 kcal/mol difference in selectivity. (B) Cartoon depiction of multiple conformational states with differing inhibitor binding competence (from Ref. [35]). The arrow (E) denotes the reorganization free energy required to adopt a particular conformational state. For inhibitor (blue) to bind the the binding-competent conformation (red), an energetic penalty must be paid to confine the kinase to the binding-competent state; favorable binding free energy can then cause this conformation to become the new lowest free energy “ground state”. (C) Illustration of how reorganization energy (red) and binding free energy (blue) give opposing contributions to overall binding free energy (black).

Predicting the emergence of mutational resistance. While drug resistance can emerge via multiple mechanisms, mutations in the therapeutic target drive the emergence of drug resistance in many patients; in commonly-targeted kinases such as Abl, missense mutations are responsible for resistance in as much as 90% of cases [36]. Developing a quantitative and predictive understanding for how these inhibitors elicit resistance through mutational mechanisms would be of immense value, allowing the rapid design of second-generation inhibitors to circumvent resistance, the exploration of therapeutic combinations to minimize the emergence of resistance, and potentially the development of new therapeutics unlikely to elicit resistance.

To predict resistance using physical modeling techniques, we hypothesize that resistance mutations ablate inhibitor binding affinity while maintaining some physically computable surrogate for activity, such as ATP and substrate peptide binding affinities. This allows us to compute a *resistance index* for each possible point mutation, and to construct enhanced sampling schemes that selectively search for point mutations that maximize the resistance index. We are currently building a framework to perform these calculations using our GPU-accelerated alchemical binding free energy code YANK [37], using kinases ABL and AURKA as model systems. Predicted resistance mutations will be compared with clinically-identified mutations in the cBioPortal, and biophysically characterized by engineering these mutations into kinase catalytic domains that can be expressed recombinantly in bacteria. We have developed an automated fluorescence assay to perform affinity measurements against a variety of small molecule kinase inhibitors of clinical interest, and use activity assays to determine whether our hypothesis regarding physical surrogates for activity is valid for individual kinases.

Predicting drug sensitivity and resistance in patient tumors. The decreasing cost of next-generation sequencing has spurred the deep sequencing of patient tumors with the goal of improving patient outcomes by individually tailoring treatment plans to patient-specific genetic alterations—potentially found only in small subpopulations of tumor cells. While this approach will undoubtedly revolutionize cancer therapy, there is a great deal of work to be done over the next decade to connect tumor alterations to optimal therapeutic plans.

Evaluating whether missense mutations in therapeutic targets represent resistance, neutral, or sensitizing mutations could significantly shape the use of targeted kinase inhibitors for cancer therapy. Though a number of frequently-occurring mutations have been noted (and, less frequently, characterized biochemically and in cell lines)—such as the T315I gatekeeper mutation in Abl, which causes resistance to first-generation targeted inhibitors—the long tail of infrequently observed mutations makes it impossible to build a catalog of mutations presented in more than a minor fraction of cancers.

We take a two-pronged approach to predicting the therapeutic resistance or sensitivity of patient-derived tumor mutations: First, we utilize the computational model built to predict the emergence of mutational resistance to identify potential resistance mutations (mutations that ablate inhibitor affinity while maintaining some surrogate for activity), and test these by engineering these mutations into bacterially-expressing kinases in our laboratory. Success of this computational model in identifying mutations that cause resistance or susceptibility could eventually aid the development of individually-tailored treatment plans that maximize therapeutic success, avoiding drugs for which intrinsic resistance (even at the subclonal level) already exists, or employing combinations of therapeutics capable of overcoming inherent resistance.

Second, we are developing an automated high-speed cell-free expression pipeline that would allow mutant kinases to be engineered, expressed, and assayed against all available clinical kinase inhibitors within a day. New cost-effective cell-free expression technologies based on S12 extracts and fed by glucose [38, 39] allow

for inexpensive and rapid scalable production of many proteins. Aided by an FGI Rapid Response grant, our laboratory is currently exploring the automated optimization of these techniques (with GFP fluorescence or mass spectrometry as a readout) using machine-learning algorithms based on Gaussian processes to improve yields for kinase domains represented in the IMPACT assay to allow the production of kinases containing engineered patient mutations in a matter of hours. Kinase production would be immediately followed by the fluorescence assay of kinase inhibitor affinity against all FDA-approved kinase inhibitors to aid the determination of whether inhibitors clinicians would like to use would be effective in light of tumor-specific mutations. **We anticipate this would reduce the entire engineering and assay pipeline to turn around sensitivity/resistance results within a single day—fast enough to have substantial impact on therapeutic decisions.**

Rational design of anticancer therapeutics

In addition to targeting kinases, our laboratory is further developing the technology of alchemical free energy calculations [3] and Markov state models (MSMs) [8, 9, 13] to enable the rational design of small molecules to a variety of other targets implicated in human cancers.

Allosteric modulation of Ras. Ras family proteins, important in the control of cell growth via signaling, are commonly mutated in human cancer [40]. Activating mutations in Ras are a leading cause of resistance to modern targeted therapy, and patients who harbor Ras mutations have considerably poorer prognoses than those with wild type Ras [41]. Targeting Ras has proven difficult because oncogenic mutations activate Ras primarily by ablating enzymatic activity, leaving classical enzyme inhibition strategies unworkable. The high affinity of Ras for GTP—which locks Ras in an active conformation—combined with high intracellular GTP concentrations makes outcompeting the bound nucleotide extremely difficult [42]. A new covalent inhibitor for K-Ras G12C from the Shokat lab shows the potential for inactivation by allosteric modulators, but the G12C mutation is present in only small fraction of cancers and these inhibitors, which covalently bind mutant Cys12, present a ready path to inhibitor resistance by mutating position 12 to another activating mutation [42].

Our laboratory is pursuing a combined computational-experimental strategy to develop noncovalent allosteric modulators of Ras family proteins that circumvents these limitations. Initially focusing on K-Ras, we have carried out massively parallel simulations on Folding@home to construct Markov state models of the conformational states accessible with a small energetic penalty (up to $10k_B T$). Using a site-mapping algorithm, we identify conformations that may be functionally inactive (such as those similar to the GDP-bound state) and use virtual screening of soluble small molecule fragment libraries to identify potential scaffolds that could be derivatized for enhanced potency. Using a panel of K-Ras mutants in which tryptophan residues have been introduced at various sites to function as fluorescence reporters of ligand-stabilized allosteric conformational change, we plan to experimentally screen these computationally-identified fragments for weak binding using our automated biophysical laboratory. We will then utilize an automated ligand design strategy (discussed in the next section) to identify synthetically-feasible derivatives that will be synthesized to increase potency.

Automated ligand design guided by synthetic feasibility. The lead optimization stage of drug discovery—in which many derivatives of a low-affinity scaffold are synthesized and tested in an attempt to increase affinity and selectivity—is a crucial step in all small molecule drug discovery projects. Current approaches are driven by weak structure-activity relationships and (if structural biology support is available) rapid rounds of structural biology and molecular graphics, allowing slow and costly progress toward the goal of increased potency and minimal off-target toxicity. As a result, this process often takes many months and consumes a great deal of cost, both in terms of synthetic chemist salaries and reagent costs.

Our laboratory has developed an enhanced sampling approach that allows us to explore combinatorially large spaces of inhibitor designs in a way that automatically biases the simulation toward inhibitors with higher affinity for one or more targets. This approach—based on expanded ensemble simulations [43] and made possible by a new nonequilibrium Monte Carlo algorithm we developed [19]—promises to provide a time- and cost-effective solution to the problem of optimizing small molecules for affinity. By restricting the space of compounds to those accessible from a given set of commercially-available starting materials and a library of common synthetic transformations [44], we can propose a set of compounds that have a high likelihood of increased potency and are likely to be readily synthesizable.

MAJOR COLLABORATIVE PROJECTS

Selective inhibition of protein methyltransferases (Minkui Luo, MSKCC)

Protein lysine methyltransferases (PKMTs) regulate numerous epigenetic processes, and there is increasing evidence for their dysregulation playing a role in cancer through downregulation of tumor suppressors or upregulation of oncogenes that promote cancer progression and proliferation [45]. There is tremendous interest in utilizing a variety of selective inhibitors to illuminate the role of PKMTs in disease and develop potential therapeutics, but few such inhibitors exist. Like kinases, PKMTs are conformationally heterogeneous, providing an opportunity to exploit binding to distinct conformations to achieve selectivity. Funded by a STARR Cancer Consortium grant with Minkui Luo, we are working to apply the same methodology used for designing selective kinase inhibitors to aid the Luo lab in developing useful chemical probes for elucidating PKMT function.

GPCR dynamics and activation (Scott Blanchard and Harel Weinstein, WCMC)

G-protein coupled receptors (GPCRs) represent a significant fraction of all drug targets, and are gradually being recognized as playing important roles in tumor progression and metastasis [46]. In collaboration with Scott Blanchard and Harel Weinstein (WCMC), we are working to illuminate the detailed conformational changes involved in biased agonism and signal transduction using massively parallel simulations on Folding@home that will bridge from static crystal structures to dynamic single-molecule FRET data collected in the Blanchard lab. The resulting detailed model will aid the subsequent design of new GPCR biased agonists and inhibitors.

Activating mutants of MTOR (James Hsieh, MSKCC)

The MSK-IMPACT panel has identified a large diversity of tumor-associated mutations in the MTOR gene, an atypical serine/threonine kinase that acts as a master growth regulator. By engineering these mutations into cell lines, the Hsieh lab has demonstrated that many of these mutations are activating, and identified several distinct mechanistic classes of mutations. Funded by an FGI grant, we are working with the Hsieh lab to elucidate the structural and energetic ramifications of these mutations with the goal of identifying opportunities for therapeutic intervention.

Allosteric and slow off-rate inhibition of CK2 and SYK (Melissa Vasbinder, AstraZeneca)

In a closed system, the binding affinity of an inhibitor for its target fully determines target inhibition, but in an open system such as a cell where drugs can be exported or metabolized and cleared from the blood, the off-rate of an inhibitor can more directly determine the efficacy of inhibition [47]. Through a funded collaboration with Astra-Zeneca, we are investigating the mechanistic origins of new slow off-rate inhibitors of CK2, and examining opportunities for engineering allosteric or slow-off-rate inhibitors for SYK.

Protonation state effects in kinase inhibitor recognition (Marilyn Gunner, CCNY)

Recently, the complicated role that protonation states play in the binding of imatinib to Abl has suggested that both *equilibrium mixtures* of protonation states and *changes* in protonation state upon binding are relevant to accurately describing inhibitor binding [48, 27, 29], posing an enormous challenge to computational drug design paradigms, which do not consider protonation state effects explicitly. We believe protonation states may play a much larger role in both on- and off-target kinase inhibition than previously appreciated, since over half of all FDA-approved small molecule kinase inhibitors have two or more protonation states predicted to be accessible during binding. Working with Marilyn Gunner (CCNY), who has extensive experience in protonation state modeling in proteins, we are embarking on a large-scale effort to elucidate the role of protonation state effects in kinase inhibitor binding and explicitly allow protonation state effects to be considered in inhibitor design and profiling of off-target effects. We plan to apply for NIH funding for this project in early 2016.

MHC-peptide-TCR interactions (David Scheinberg, MSKCC)

Recently, the development of immunotherapies that promote the recognition of mutated tumor antigen has spurred the desire to develop quantitative, predictive models for the association of major histocompatibility complex (MHC) with T cell receptors (TCR) mediated by antigen peptide binding interactions (Figure 3), in order to develop novel biologics that may be useful in cancer treatment. These biologics may be recombinant variants of TCRs, TCR-mimicking antibodies (TCRm), or even synthesized antigen peptides for use as potential vaccines. In all of these cases, a detailed, accurate model for the affinity of the MHC-peptide-TCR/TCRm complex (including the details of an individual's MHC variants) for both cancer and self antigen peptides is essential.

While simple experiment-derived additive models derived from experimental data have been proposed [49], these models are insufficiently accurate for the purpose of therapeutic engineering. We are working with Ron Gejman and David Scheinberg at MSKCC to utilize atomistically detailed free energy calculations for this purpose. First, we are exploring what biophysical detail is necessary to recapitulate known affinities for the MHC-peptide-TCR ternary complex. Next, we will utilize advanced techniques for sampling over sequence space to perform the computational equivalent of evolution selection experiments, selecting for peptide sequences that bind tightly to specific MHC-TCR complexes. In addition to aiding the development of novel therapeutics, accurate computational models of MHC-peptide-TCR complex affinity will also be useful in predicting the degree of therapeutic response from individual tumor sequences.

Human serum albumin binding and interference (Ron Blasberg, MSKCC)

Binding to the abundant human serum protein albumin (HSA, Figure 4) is well-known to modulate the pharmacokinetics of drugs, including common drugs such as aspirin and ibuprofen [50]. While this may sometimes be advantageous, in many cases, HSA binding prevents accumulation of bioactive conformations at target sites of interest. The Blasberg lab encountered just such a case with a radiolabeled dye molecule intended for the visualization of CXCR-expressing cells in brain cancer associated with poor prognosis: While the molecule binds CXCR effectively and crosses the blood-brain barrier, the molecule binds HSA to a degree that useful quantities cannot accumulate in the brain.

To enable these radioligands to be useful in imaging, we are working with the Blasberg lab to identify safe, well-tolerated drugs that could be coadministered with the label compound to saturate the HSA binding site and ensure sufficient compound can reach target receptors in the brain. While complicated by the fact that HSA has at least seven characterized binding sites (Figure 4), this could function as a novel approach to modulating the pharmacokinetics of HSA-binding drugs or imaging agents. Our approach first aims to identify the binding sites to which the radioimaging compound has high affinity using a novel GPU-accelerated binding free energy computation technique we developed for characterizing multiple weak binding events. We then utilize automated isothermal titration calorimetry and fluorescence competition assays to experimentally verify both binding sites (using well-characterized reporter compounds) and compound affinity. Once candidate well-tolerated competitive drugs are identified, the Blasberg lab will carry out subsequent experiments in rodents to demonstrate the practicality of this approach.

We are currently developing this concept into an NIH grant application with James Fraser (UCSF), who will use room-temperature X-ray crystallography to provide complementary experimental validation of our findings.

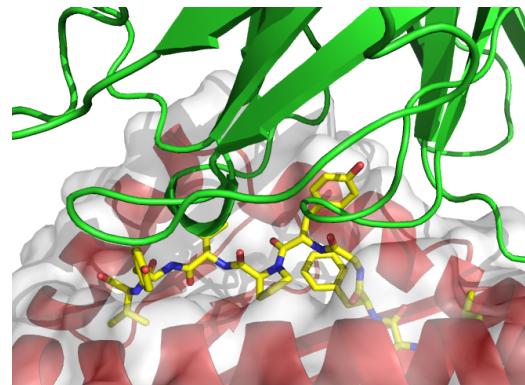


Figure 3: MHC-peptide-TCR complex. Human major histocompatibility complex (MHC; red) HLA-A2 and T cell receptor (TCR; green) is shown in complex with the viral TAX peptide (LLF-GYPVYV; yellow) [rcsbid:1AO7].

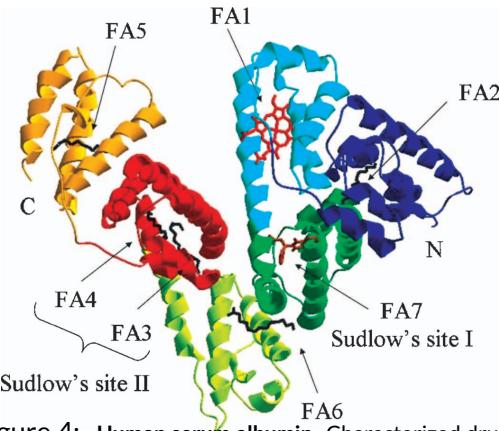


Figure 4: Human serum albumin. Characterized drug and fatty acid binding sites are shown. Adapted from [50].

References

- [1] B. L. Tembe and J. A. McCammon, *Computers & Chemistry* **8**, 281 (1984).
- [2] D. L. M. Michael R. Shirts and J. D. Chodera, *Annu. Rep. Comput. Chem.* **3**, 41 (2007).
- [3] J. D. Chodera, D. L. Mobley, M. R. Shirts, R. W. Dixon, K. M. Branson, and V. S. Pande, *Curr. Opin. Struct. Biol.* **21**, 150 (2011).
- [4] K. Wang, J. D. Chodera, Y. Yang, and M. R. Shirts, *J. Comput. Aid. Mol. Des.* **27**, 989 (2013).
- [5] M. Shirts and V. S. Pande, *Science* **290**, 1903 (2000).
- [6] <http://fah-web.stanford.edu/cgi-bin/main.py?qtype=osstats>.
- [7] T.-S. Lee, L. T. Chong, J. D. Chodera, and P. A. Kollman, *J. Am. Chem. Soc.* **123**, 12837 (2001).
- [8] J. D. Chodera, W. C. Swope, J. W. Pitera, and K. A. Dill, *Multiscale Model. Simul.* **5**, 240 (2006).
- [9] J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, *J. Chem. Phys.* **126**, 155101 (2007).
- [10] S. Bacallado, J. D. Chodera, and V. S. Pande, *J. Chem. Phys.* **131**, 045106 (2009).
- [11] J.-H. Prinz, J. D. Chodera, V. S. Pande, J. C. Smith, and F. Noé, *J. Chem. Phys.* in press (2011).
- [12] F. Noé, S. Doose, I. Daidone, M. Löllmann, M. Sauer, J. D. Chodera, and J. C. Smith, *Proc. Natl. Acad. Sci. USA* **108**, 4822 (2011).
- [13] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Fischbach, M. Held, J. D. Chodera, C. Schütte, and F. Noé, *J. Chem. Phys.* **134**, 174105 (2011).
- [14] D. L. Mobley, J. D. Chodera, and K. A. Dill, *J. Chem. Phys.* **125**, 084902 (2006).
- [15] D. L. Mobley, J. D. Chodera, and K. A. Dill, *J. Chem. Theor. Comput.* **3**, 1231 (2007).
- [16] D. L. Mobley, Élise. Dumont, J. D. Chodera, and K. A. Dill, *J. Phys. Chem. B* **111**, 2242 (2007).
- [17] M. R. Shirts, D. L. Mobley, J. D. Chodera, and V. S. Pande, *J. Phys. Chem. B* **111**, 13052 (2007).
- [18] D. L. Mobley, J. D. Chodera, and K. A. Dill, *J. Phys. Chem. B* **112**, 938 (2008).
- [19] J. P. Nilmeier, G. E. Crooks, D. D. L. Minh, and J. D. Chodera, *Proc. Natl. Acad. Sci. USA* **108**, E1009 (2011).
- [20] P. J. Elms, J. D. Chodera, C. J. Bustamante, and S. Marqusee, *Biophys. J.* **103**, 1490 (2012).
- [21] P. J. Elms, J. D. Chodera, C. J. Bustamante, and S. Marqusee, *Proc. Natl. Acad. Sci. USA* **109**, 3796 (2012).
- [22] J. D. Chodera, P. Elms, W. C. Swope, F. Noé, and V. S. Pande, submitted .
- [23] J. D. Chodera, P. Elms, F. Noé, B. Keller, C. M. Kaiser, A. Ewall-Wice, S. Marqusee, C. J. Bustamante, and N. S. Hinrichs, submitted .
- [24] C. M. Kaiser, D. H. Goldman, J. D. Chodera, I. T. Jr., and C. Bustamante, *Science* **334**, 1723 (2011).
- [25] S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, and A. L. Schacht, *Nature Rev. Drug Discovery* **9**, 203 (2010).
- [26] D. B. Longley and P. G. Johnston, *J. Pathol.* **205**, 275 (2005).
- [27] M. A. Seeliger, B. Nagar, F. Frank, X. Cao, M. N. Henderson, and J. Kuriyan, *Structure* **15**, 299 (2007).
- [28] A. Aleksandrov and T. Simonson, *J. Biol. Chem.* **285**, 13807 (2010).
- [29] Y.-L. Lin, Y. Meng, W. Jiang, and B. Roux, *Proc. Natl. Acad. Sci. USA* **110**, 1664 (2013).
- [30] M. A. Fabian, W. H. Biggs, D. K. Treiber, C. E. Atteridge, M. D. Azimioara, M. G. Benedetti, T. A. Carter, P. Ciciri, P. T. Edeen, M. Floyd, J. M. Ford, M. Galvin, J. L. Gerlach, R. M. Grotzfeld, S. Herrgard, D. E. Insko, M. A. Insko, A. G. Lai, J.-M. Lélias, S. A. Mehta, Z. V. Milanov, A. M. Velasco, L. M. Wodicka, H. K. Patel, P. P. Zarrinkar, and D. J. Lockhart, *Nature Biotech.* **23**, 329 (2005).
- [31] D. L. Parton, P. B. Grinaway, K. A. Beauchamp, S. M. Hanson, and J. D. Chodera, submitted .
- [32] M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 124105 (2008).
- [33] J. D. Chodera and M. R. Shirts, *J. Chem. Phys.* **135**, 194110 (2011).
- [34] G. R. Bowman and P. L. Geissler, *Proc. Natl. Acad. Sci.* **109**, 11681 (2012).
- [35] G. M. Lee and C. S. Craik, *Science* **324**, 213 (2009).
- [36] N. P. Shah, J. M. Nicoll, B. Nagar, M. E. Gorre, R. L. Paquette, J. Kuriyan, and C. L. Sawyers, *Cancer Cell* **2**, 117 (2002).
- [37] J. D. Chodera, K. Branson, M. R. Shirts, D. L. Mobley, and P. Vijay S, YANK, an open platform for GPU-accelerated binding free energy calculations, 2010, <http://simtk.org/home/yank>.
- [38] T.-W. Kim, J.-W. Keum, I.-S. Oh, C.-Y. Choi, C.-G. Park, and D.-M. Kim, *Journal of Biotechnology* **126**, 554 (2006).
- [39] K. A. Calhoun and J. R. Swartz, *Biotechnology Progress* **21**, 1146 (2008).
- [40] Y. Pylayeva-Gupta, E. Grabocka, and D. Bar-Sagi, *Nat. Rev. Cancer* **11**, 761 (2011).
- [41] R. H. Mak, G. Hermann, J. H. Lewis, H. J. Aerts, E. H. Baldini, A. B. Chen, Y. L. Colson, F. H. Hacker, D. Kozono, J. O. Wee, Y.-H. Chen, P. J. Catalano, K.-K. Wong, and D. J. Sher, *Clin. Lung Cancer* **16**, 24 (2015).
- [42] J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, and K. M. Shokat, *Nature* **503**, 548 (2013).
- [43] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *J. Chem. Phys.* **96**, 1776 (1992).
- [44] S. D. Roughley and A. M. Jordan, *J. Med. Chem.* **54**, 3451 (2011).
- [45] R. A. Copeland, M. E. Solomon, and V. M. Richon, *Nat. Rev. Drug Discov.* **8**, 724 (2009).
- [46] R. T. Dorsam and J. S. Gutkind, *Nat. Rev. Cancer* **7**, 79 (2007).
- [47] R. A. Copeland, D. L. Pompliano, and T. D. Meek, *Nat. Rev. Drug Disc.* **5**, 730 (2007).
- [48] Z. Szakács, S. Béni, Z. Varga, L. Örfi, G. Kéri, and B. Noszál, *J. Med. Chem.* **48**, 249 (2005).
- [49] M. E. Birnbaum, J. L. Mendoza, D. K. Sethi, S. Dong, J. Glanville, J. Dobbins, E. Özkan, M. M. Davis, K. W. Wucherpfennig, and K. C. Garcia, *Cell* **157**, 1073 (2014).
- [50] M. Fasano, S. Curry, E. Terreno, M. Galliano, G. Fanali, P. Narciso, S. Notari, and P. Ascenzi, *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)* **57**, 787 (2005).