Kenichiro Takaba

RESEARCH INTEREST

Drug design, Molecular dynamics simulation, Fragment molecular orbital method, Free energy calculations, De novo drug design, Protein engineering, Machine learning (graph neural networks, reinforcement learning, active learning, transfer learning, etc.)

EDUCATION

- **2016-2020:** Ph.D. in Physics, The University of Tokyo, Japan
  Edge expansion parallel cascade selection molecular dynamics simulation (eePaCS-MD) for investigating protein dynamics

- **2010-2012:** M.S. in Physics, The University of Tokyo, Japan
  Theoretical study of the protonation states of H^+-PPase using molecular dynamics simulations

- **2006-2010:** B.S. in Engineering Science, University of Tsukuba, Japan

WORK EXPERIENCE

2012.04-Present: Computational chemist, Asahi Kasei Pharma Corporation
- Engaged with multiple structure-based and ligand-based drug discovery projects for various drug modalities including small molecules and antibodies
- Infrastructure development of computational tools for medicinal chemists such as automated docking simulations, QSAR/QSPR predictions and *de-novo* compound design
- Development of web services

2022.04-2024.04: Visiting scientist, Memorial Sloan Kettering Cancer Center
- Development of open-source software for computer-aided drug discovery

TECHNICAL EXPERIENCE

- Expertise with biomolecular simulations, cheminformatics, and machine learning for drug discovery
- Expertise in using computational tools such as Maestro and MOE
- Expertise with Python programming language and Unix shell scripts
- Expertise with open-source software packages such as RDKit, OpenMM, and machine learning frameworks such as Scikit-learn and Pytorch
- Fluent with GitHub for code sharing and development
- Experience with flask and SQLAlchemy to develop web services
- Experience with AWS for High Performance Computing (e.g. ParallelCluster)

OTHER EXPERIENCE

- Experience in leading regional and working groups of several industry-academia collaboration consortium

SELECTED PUBLICATIONS

Preprints

- *arXiv:2409.01931*
  On the design space between molecular mechanics and machine learning force fields

  Yuanqing Wang, **Kenichiro Takaba**, Michael S. Chen, Marcus Wieder, Yuzhi Xu, Tong Zhu, John Z. H. Zhang, Arnav Nagle, Kuang Yu, Xinyan Wang, Daniel J. Cole, Joshua A. Rackers, Kyunghyun Cho, Joe G. Greener, Peter Eastman, Stefano Martiniani, Mark E. Tuckerman

  Machine learning force fields (MLFFs) aim to combine the accuracy of quantum mechanics with the speed of molecular mechanics. However, their primary challenge lies in balancing speed with generalizability. This review focuses on the tradeoff between speed and accuracy in force fields, highlighting current efforts to improve both the performance of molecular mechanics (MM) and MLFFs, while considering the future development of more efficient models.

- *bioRxiv:2024.05.28.596296*
  DrugGym: A testbed for the economics of autonomous drug discovery

  Michael Retchin, Yuanqing Wang, **Kenichiro Takaba,** John D. Chodera

  Here, we introduce a frame-work for modeling the stochastic process of drug discovery, representing a realistic testbed for machine learning methods applied to the hit-to-lead phase of small molecules.

Journals

- *Chem Sci.* 2024 (doi: 10.1039/D4SC00690A)
  Machine-learned molecular mechanics force fields from large-scale quantum chemical data

  Kenichiro Takaba, Anika Friedman, Chapin Cavender, Pavan Behara, Iván Pulido, Mike Henry, Hugo MacDermott-Opeskin, Christopher Iacovella, Arnav Nagle, Alexander Payne, Michael Shirts, David L. Mobley, John D. Chodera, and Yaunqing Wang

  Development of self-consistent molecular mechanics force field in an end-to-end differentiable manner directly from quantum chemical calculations using a machine learning framework (Espaloma). This paper extends the original Espaloma method by incorporating energy and force fitting, as well as L2 regularization to torsion and improper torsion force constants, effectively reducing the model variance. The resulting force field, espaloma-0.3.0, can achieve accurate protein-ligand binding free energies when self-consistently parametrizing both the protein and ligand.

- *J. Chem. Phys.* **152**, 225101 (2020)
  Edge Expansion Parallel Cascade Selection Molecular Dynamics Simulation for Investigating Large-amplitude Motion of Proteins

  Kenichiro Takaba, Akio Kitao

  A new adaptive sampling method is proposed to explore the large-amplitude motions of proteins with a focus on domain motions. The proposed method repeats multiple MD simulations from structures that are located at the boundary of a principal component space with the concept of nodes and edges which are designed to expand as the adaptive sampling cycles increase.

INVITED TALKS

- **Machine learned molecular mechanics force fields from large scale quantum chemical data**
  2024 Workshop on Free Energy Methods in Drug Discovery, Leiden, The Netherlands

  A machine-learned molecular mechanics force field is developed using Espaloma, where the rule-base discrete atom-typing scheme is replaced with a continuous atom representation via graph neural networks in an end-to-end differentiable framework. Trained in a single GPU-day to fit a large and diverse quantum chemical dataset of over 1.1M energy and force calculations, we demonstrate how this method and the developed force field are applicable for real drug discovery applications. These applications include self-consistently parametrizing proteins and ligands to produce stable simulations, leading to highly accurate predictions of binding free energies.

- **Application of Molecular Dynamics and Fragment Molecular Orbital Method in Drug Discovery**
  392th CBI seminar and workshop: Application of molecular dynamics in drug discovery
  March 2018, Tokyo, Japan

  Practical application of MD simulations in drug discovery projects such as docking pose validation, predictions of protonation states, and protein-ligand interaction profiling are discussed. Furthermore, application of fragment molecular orbital method (FMO) is introduced as a tool to analyze the protein-ligand interactions based on quantum mechanics providing better understanding of the protein-ligand system. Additionally, an example is shown where systematic averaging of FMO-based interaction energies over multiple conformers generated by MD simulations can be used as a convenient scoring method to predict the protein-ligand binding affinity.

- **Validation of Supervised Molecular Dynamics and its Usefulness in Drug Discovery**
  MOE Forum 2017
  July 2017, Tokyo, Japan

  Supervised molecular dynamics (SuMD) [Sabbadin et al. 2014] is an adaptive sampling method to explore the protein-ligand recognition pathway which utilizes a tabu-like supervision algorithm on the protein–ligand distance. In this study, accelerated MD was combined with the original SuMD to speed up the binding event by lowering the dehydration barrier of the binding site. Furthermore, the original SuMD was extended for protein-ligand dissociation (rSuMD) where the wrong docking poses tend to dissociate faster than the correct binding poses, suggesting that rSuMD may be useful for post processing of structure-based virtual screening to decrease the false positive hits.