

**Workshop on Integrative
modelling with HADDOCK in
conjunction with Workshop
on Novel Computational
Methods for Structural
Biology**

Report of Contributions

Contribution ID: 1

Type: **not specified**

General introduction to docking, integrative modelling and HADDOCK

Monday, March 20, 2023 9:00 AM (1h 30m)

Presenter: BONVIN, Alexandre M.J.J. (Utrecht University)

Contribution ID: 2

Type: **not specified**

Computer practical: Antibody-antigen docking with the HADDOCK server

Monday, March 20, 2023 11:00 AM (1h 30m)

HADDOCK2.4 antibody-antigen docking tutorial: This tutorial demonstrates the use of HADDOCK2.4 for predicting the structure of an antibody-antigen complex using information about the hypervariable loops of the antibody and either the entire surface of the antigen or a loose definition of the epitope. This tutorial does not require any Linux expertise and only makes use of our web servers and PyMol for visualisation/analysis.

<https://www.bonvinlab.org/education/HADDOCK24/HADDOCK24-antibody-antigen/>

Presenter: BONVIN, Alexandre M.J.J. (Utrecht University)

Contribution ID: 3

Type: **not specified**

Lecture & practical: Introduction to HADDOCK3 and practical

Monday, March 20, 2023 2:00 PM (1h 30m)

HADDOCK3 antibody-antigen docking: This tutorial demonstrates the use of HADDOCK3 for predicting the structure of an antibody-antigen complex using information about the hypervariable loops of the antibody and either the entire surface of the antigen or a loose definition of the epitope. It illustrates the modularity of HADDOCK3 by introducing a new workflow not possible under the current HADDOCK2.X versions. As HADDOCK3 only exists as a command line version, this tutorial does require some basic Linux expertise.

<https://www.bonvinlab.org/education/HADDOCK3/HADDOCK3-antibody-antigen/>

Presenter: BONVIN, Alexandre M.J.J. (Utrecht University)

Contribution ID: 4

Type: **not specified**

Different Facets of Distance Geometry

Monday, March 20, 2023 4:00 PM (30 minutes)

The Distance Geometry Problem (DGP) asks whether a simple weighted undirected graph $G=(V,E,d)$ can be realized in the K -dimensional Euclidean space so that the distance constraints implied by the weights on the graph edges are satisfied. This problem was proven to be NP-hard in the context of graph embeddability, and has several applications. In this talk, we will focus on various currently ongoing works in this very rich research context: (1) We will talk about a particular class of DGP instances where, under particular assumptions, it is possible to represent the search space as a binary tree, and where in ideal situations, vertex positions can be assigned to each of its nodes. In real-life applications, however, the distances are generally provided with low precision, and they are actually likely to carry measurement errors, so that local continuous search spaces are actually assigned to each tree node. This gives rise to special combinatorial problems which are locally continuous. (2) We will review the main applications of the DGP, ranging from structural biology, passing through sensor network localization and adaptive maps, until the dynamical component is included in DGP instances for computer graphics applications. (3) Finally, we present an alternative computing approach for the solution of the DGPs in dimension 1, where an analog optical processor is employed for the computations, which is based on some properties of laser light beams.

Presenter: Prof. MUCHERINO, Antonio (Institut de Recherche en Informatique et Systèmes Aléatoires, University of Rennes 1, France)

Session Classification: Workshop on Novel Computational Methods for Structural Biology

Contribution ID: 5

Type: **not specified**

A Curvilinear-Path Umbrella Sampling Approach to Characterizing Thermodynamics and Mechanisms of Biomolecular Interactions

Monday, March 20, 2023 4:30 PM (30 minutes)

Protein-protein and protein-ligand interactions are central in biological mechanisms. These interactions can be classified into thermodynamics and mechanistic pathways. Estimating accurate and reliable interaction energetics along the thermodynamic pathway is one of the ongoing challenges in computational biophysics. Umbrella sampling simulation-based potential of mean force calculations is one of the methods to estimate the interaction energetics. Previously this method was implemented by first choosing a predefined path of dissociation, which is often chosen as a straight-line/vectorial path. However, there are several unresolved issues such as choices of predefined direction, corrections of potential of mean force to standard free energy of binding, etc. To unleash these limitations, we developed a curvilinear-path umbrella sampling molecular dynamics (MD) simulation approach to address some of the issues. We have applied the new method for evaluating the standard free energy of binding for the barnase-barstar protein-protein system and then on a protein-ligand system, where the interaction energetics of FKBP12-rapamycin protein-ligand system is estimated. The computed energetics for both systems are in good agreement with the experimental values. The revealed mechanistic insight for the protein-protein complex matches very-well with the computationally expensive adaptive biasing MD based brute-force methods. Further, we also conducted the simulations of dissociation reactions of ternary complex FKBP12-rapalog-FRB, which indeed demonstrated a tug-of-war between FRBP12 and FRB to bind with the rapamycin, and revealed that the rapamycin prefers to bind with FKBP12 more than FRB. Thus, the glue-like molecule rapamycin and other rapalogs seem to follow a step-wise path of forming FKBP12-rapalog complex first and then the ternary complex with FRB. Thus, the developed curvilinear-path approach offers accurate and reliable binding energetic, is sensitive enough to distinguish the change in interaction energetics upon mutations, and can reliably reveal mechanistic details towards the fulfillment of the characterization.

Presenter: Dr JOSHI, Dhananjay (Research Center for Applied Science, Academia Sinica)

Session Classification: Workshop on Novel Computational Methods for Structural Biology

Contribution ID: 6

Type: **not specified**

Quantum Machine Learning for Structure-Based Virtual Screening of the Entire Medicinal Chemical Space

Monday, March 20, 2023 5:00 PM (30 minutes)

It has been estimated based on the graph theory that there are at least 1060 organic molecules that are relevant for small-molecule drug discovery. Using machine learning to estimate the binding free energies for screening of large chemical libraries to search for the tightly binding inhibitors would take a considerable amount of computational resources, yet it is not possible to explore the entire biologically relevant chemical space. Quantum computing provides a unique opportunity to accomplish such a computational task in the near future. Here, we demonstrate how to use 512 occupancies to describe the structures of protein-ligand complexes, how to convert the classical occupancies to the quantum states using nine qubits, and to estimate the binding free energies (ΔG_{bind}) of the complexes using quantum machine learning. We showed that it is possible to use only 450 parameters to prepare the quantum states for describing the structure of one protein-ligand complex. In this work the entire 2020 PDBbind dataset was adopted as the training set, and we used 45 parameters as the first attempt to construct the model for predicting the binding free energies (ΔG_{bind}). The Pearson correlation coefficient (PCC) between the estimated binding free energies and the corresponding experimental values are 0.49. By slightly increasing to 1,440 parameters for constructing the neural network model for the prediction of the ΔG_{bind} , the PCC is improved to be 0.78, which is even slightly better than to the results achieved by recent classical convolutional neural network models using more than millions of parameters. In this work, for the first time, we demonstrated the feasibility of using quantum computers to explore the entire medicinal chemical space with a concrete, implementable approach.

Presenter: Prof. LIN, It has been estimated based on the graph theory that there are at least 1060 organic molecules that are relevant for small-molecule drug discovery. Using machine learning to estimate the binding free energies for screening of large chemical libraries to search for the tightly binding inhibitors would take a considerable amount of computational resources, yet it is not possible to explore the entire biologically relevant chemical space. Quantum computing provides a unique opportunity to accomplish such a computational task in the near future. Here, we demonstrate how to use 512 occupancies to describe the structures of protein-ligand complexes, how to convert the classical occupancies to the quantum states using nine qubits, and to estimate the binding free energies (ΔG_{bind}) of the complexes using quantum machine learning. We showed that it is possible to use only 450 parameters to prepare the quantum states for describing the structure of one protein-ligand complex. In this work the entire 2020 PDBbind dataset was adopted as the training set, and we used 45 parameters as the first attempt to construct the model for predicting the binding free energies (ΔG_{bind}). The Pearson correlation coefficient (PCC) between the estimated binding free energies and the corresponding experimental values are 0.49. By slightly increasing to 1,440 parameters for constructing the neural network model for the prediction of the ΔG_{bind} , the PCC is improved to be 0.78, which is even slightly better than to the results achieved by recent classical convolutional neural network models using more than millions of parameters. In this work, for the first time, we demonstrated the feasibility of using quantum computers to explore the entire medicinal chemical space with a concrete, implementable approach. (Research

Center for Applied Science, Academia Sinica)

Session Classification: Workshop on Novel Computational Methods for Structural Biology