

# ***Enlighten2: molecular dynamics simulations of protein-ligand systems made accessible***

**Kirill Zinovjev<sup>1</sup> and Marc W. van der Kamp<sup>1,\*</sup>**

<sup>1</sup>School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol BS8 1TD, UK

\*To whom correspondence should be addressed; [marc.vanderkamp@bristol.ac.uk](mailto:marc.vanderkamp@bristol.ac.uk)

## **Abstract**

X-ray crystallography allows detailed insight into protein structure and protein-ligand interactions, which is crucial for many areas of bioscience, including drug design and enzyme engineering. However, crystal structures provide little more than a static picture of protein-ligand interactions, whereas dynamical information is often required for deeper understanding and to assess the effect of mutations. Molecular dynamics (MD) simulations can provide such information but setting up and running these simulations is not straightforward and requires expert knowledge. There is thus a need for a tool that makes protein-ligand simulation easily accessible to non-expert users.

We present Enlighten2: efficient simulation protocols for protein-ligand systems alongside a user-friendly plugin to the popular visualization program PyMOL. With Enlighten2, non-expert users can straightforwardly run and visualize MD simulations on protein-ligand models of interest. There is no need to learn new programs and all underlying tools are free and open source.

The Enlighten2 python package and PyMOL plugin are free to use under the GPL3.0 licence and can be found at [enlighten2.github.io](https://enlighten2.github.io). We also provide a lightweight Docker image via DockerHub that includes Enlighten2 with all the required utilities.

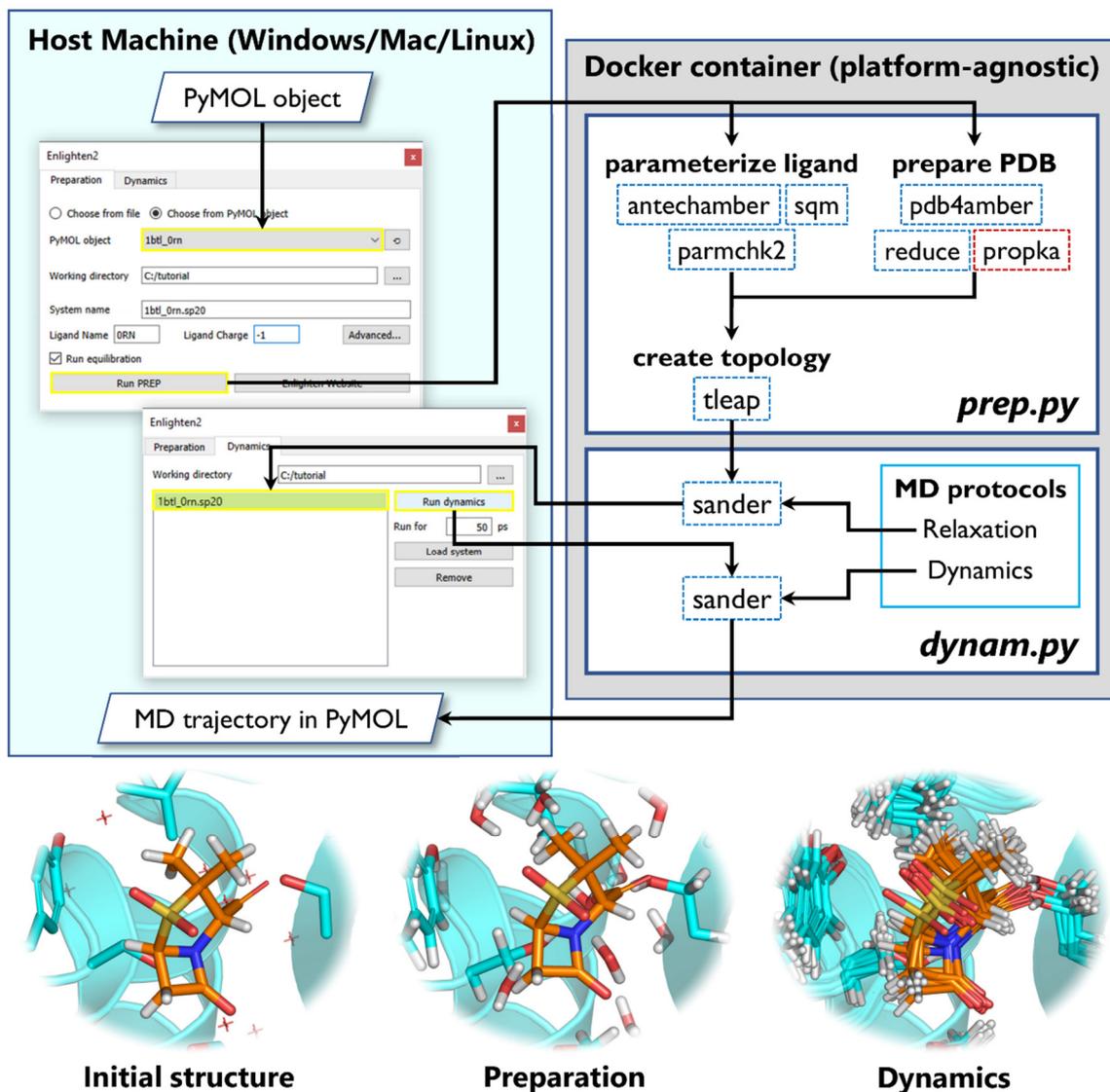
## Introduction

Biomolecular simulation, and especially molecular dynamics (MD) simulation, is increasingly used for understanding and designing biological molecules and their interactions (Huggins, et al., 2019). Atomically detailed, physics-based MD simulation methods are demonstrating growing impact in areas as diverse as biocatalysis, biomaterials, biotechnology, and drug design. Such simulations offer uniquely detailed insights at high time- and space- resolution into mechanisms, dynamics, and processes. Particularly important for the (re)design of enzymes as biocatalysts is assessment of the effect of mutations, for which static structural modelling alone is of limited use.

Although automated MD simulations are available in other visualization programs, such as YASARA Structure (Land and Humble, 2018), the QwikMD plugin to VMD (Ribeiro, et al., 2016) or commercial program packages such as MOE or Discovery Studio, these still require users to obtain and get used to a new, unfamiliar program and interface. Other options, such as the GROMACS Plugin to PyMOL (pymol.org, Schrödinger, LLC, 2018) still require expert knowledge (for ligand parameterization and MD simulation) and the non-trivial installation of MD software. There is thus a need for user-friendly, low-barrier access to MD simulation of protein-ligand complexes through a visualization program familiar to most structural biologists.

As a first attempt to meet this need, we initially developed efficient protocols (implemented as shell-scripts) and a basic PyMOL plugin (for the legacy Tkinter GUI) as Enlighten. These protocols have already been used to analyze enzyme-substrate interactions and dynamics (Rice, et al., 2019; Wang, et al., 2018) as well as for preparation for quantum mechanical / molecular mechanical (QM/MM) simulations of enzyme reactions (Byrne, et al., 2016; Drulyte, et al., 2019).

Now, we introduce Enlighten2, which includes a python package for the protocols, a new user-friendly PyMOL plugin and, notably, a Docker image with all the required software that allows seamless running of simulations on Windows, Mac and Linux operating systems without non-trivial software installation. In addition, both the python package and the plugin give users increased flexibility over the original Enlighten implementation. Enlighten2 thereby enables experimentalists to easily perform basic MD simulations of their protein-ligand system of interest through a familiar and open-source interface, without requiring the knowledge for using or manual installation of an MD software package or utilities.



**Fig. 1. Overview of the Enlighten2 PyMOL plugin and tools.** Top: workflow for setting up and running protein-ligand MD simulation, with PyMOL plugin GUI windows on the right, and utilities run inside the Docker container on the left. *prep.py* and *dynam.py* are the python wrappers. Bottom: example views going from protein-ligand structure (TEM-1  $\beta$ -lactamase [PDB ID: 1BTL] with sulbactam from Enlighten2 tutorial 1) to a system prepared for MD simulation and dynamics output.

## Methods and implementation

Enlighten2 consists of three main parts (see figure):

1. A *python package* with wrappers for AmberTools programs and predefined protocols for enzyme-ligand system preparation, relaxation and production dynamics. The package consists of a set of wrappers that automate the usage of external programs for system preparation (AmberTools and propka3.1 (Søndergaard, et al., 2011); see Fig. 1). Details can be found on the project github page. It also includes predefined protocols for two types of MD simulations: relaxation (consists of restrained gradual minimization, simulated annealing and final minimization) and production dynamics (consists of heating, production MD and final minimization). Relaxation must be run prior to dynamics. The protocols are time-efficient, because only the dynamics of a (user-defined) sphere around the ligand is included – sufficient for protein-ligand interactions but not for e.g. large-scale conformational change.
2. A *Docker image* (<https://www.docker.com>) with preinstalled enlighten2 package and a reduced version of AmberTools including antechamber, parmchk2 and sqm – for parameterization of the ligand with the General Amber Force Field (Wang, et al., 2004); pdb4amber, reduce(Word, et al., 1999) and tleap – for preparation of Amber topology and coordinate files; sander – for MD simulations. It also includes propka3.1 for assignment of protonation states for titratable residues.
3. A *GUI plugin for PyMOL* consists of two tabs: “Preparation” and “Dynamics” (Fig. 1). The “Preparation” tab allows user to prepare and relax a protein-ligand model system based on a PyMOL object or a PDB file. The advanced options allow for fine tuning of the preparation protocol (setting size of the sphere, pH of the simulation etc.). The “Dynamics” tab is dedicated to running dynamics of specified duration as well as to management of previously run simulations (including loading trajectories to PyMOL). The user is required to provide a protein-ligand model with hydrogens added to the ligand. Two detailed tutorials for novice users are available on the project website. Tutorial 1 features running simulations on a model system constructed from two PDB structures and tutorial 2 shows how to include a common co-factor alongside a ligand.

## Conclusions

Enlighten2 provides non-expert users easy access to efficient, automated MD simulations of protein-ligand complexes, through a familiar interface and without assuming knowledge of setting up and running biomolecular simulations. The usage of Docker guarantees that the code runs and behaves identically on any machine and, apart from PyMOL and Docker, no user installation is required. It is therefore easy to use in classroom teaching or during online workshops. In addition, by using Enlighten2 from the command-line or as a python package, one can systematically run MD simulations for a set of ligands or enzyme variants as part of *in silico* workflows.

## Acknowledgements

M.W.v.d.K. thanks Simon Bennie, Saad Raza, Kara Ranaghan and Adrian Mulholland for their contributions to the initial Enlighten tools, plugin and tutorials.

## Funding

This work has been supported by the by the Biotechnology and Biological Sciences Research Council (BB/L018756/1 and BB/M026280/1) and the UK Catalysis Hub (EPSRC grant EP/M013219/1).

## References

- Byrne, M.J. et al. (2016) The Catalytic Mechanism of a Natural Diels–Alderase Revealed in Molecular Detail. *J. Am. Chem. Soc.*, **138**, 6095-6098.
- Case, D.A. et al. (2018) AMBER 2018, University of California, San Francisco.
- Drulyte, I. et al. (2019) Crystal structure of the putative cyclase IdmH from the indanomycin nonribosomal peptide synthase/polyketide synthase. *IUCrJ*, **6**, 1120-1133.
- Huggins, D.J. et al. (2019) Biomolecular simulations: From dynamics and mechanisms to computational assays of biological activity. *Wires. Comput. Mol. Sci.*, **9**, e1393.
- Land, H. and Humble, M.S. (2018) YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. In: Bornscheuer, U.T. and Höhne, M., editors, *Protein Engineering: Methods and Protocols*, Springer New York, 43-67.
- Ribeiro, J.V. et al. (2016) QwikMD—Integrative Molecular Dynamics Toolkit for Novices and Experts. *Sci. Rep.*, **6**, 26536.
- Rice, K. et al. (2019) The predominance of nucleotidyl activation in bacterial phosphonate biosynthesis. *Nat. Comm.*, **10**, 3698.
- Søndergaard, C.R. et al. (2011) Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theory Comput.*, **7**, 2284-2295.
- Wang, J. et al. (2004) Development and testing of a general amber force field. *J. Comp. Chem.*, **25**, 1157-1174.
- Wang, L. et al. (2018) A Rieske oxygenase/epoxide hydrolase-catalysed reaction cascade creates oxygen heterocycles in mupirocin biosynthesis. *Nat. Catal.*, **1**, 968-976.
- Word, J.M. et al. (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation, *J. Mol. Biol.*, **285**, 1735-1747.