

Title: *Computational Detection of Pathways for Flexible Ligands in Protein Structures*

Author: *Bc. Tom Jankovec*

Supervisor: *Vojtěch Vonásek, Ph.D.*

Reviewer: *Adam Jurčák, Ph.D.*

This master thesis presents a novel algorithm for the detection of paths that could be used by a ligand when binding or unbinding to/from a protein receptor. The path detection algorithm employs a robotics approach. Specifically, it is based on the RRT-Path motion-planning algorithm which uses precomputed CAVER tunnels as auxiliary paths. Furthermore, the algorithm employs ligand flexibility and scaling of van der Waals (vdW) radii of both a ligand and a receptor in order to accommodate for the dynamic and ligand-induced behaviour of molecules. The author also compares the proposed algorithm to the state of the art MoMA-LigPath software.

Overall, the thesis is very well written. First, the author describes (Chapter 2) how the motion planning methods relate to the ligand path problem in general. Then, the algorithm description follows (Chapter 3), and it is easy to comprehend. An important part of the thesis is also the presentation of the experiments (Chapter 4). The proposed algorithm is compared to MoMA-LigPath, and the algorithm's ability to consider potential energies is demonstrated. Finally the results from the experiments are discussed and two main conclusions are drawn (Chapters 5, 6):

1. The proposed algorithm surpasses MoMA-LigPath in terms of the path detection success rate, and also time performance.
2. The proposed algorithm considers potential energies which was not enabled by any similar algorithm before.

The thesis is well structured, its language is clear, and the included illustrations and algorithm listings are supportive.

At the first sight, this thesis could be rated as excellent. However, I have concerns about the design of the experiments w.r.t. MoMA-LigPath (MMLP) and therefore also about the reliability of the conclusion nr. 1. First, it is not obvious why significantly different settings in terms of the allowed vdW radii scales were used for MMLP – 0.75 (guessed from web defaults) – and the proposed algorithm – [0.5-0.8]. There might be a rationale behind such settings, but it is missing in the thesis, and the default settings of MMLP are missing as well. Looking at the results, several questions that might be interrelated arise:

- What were the vdW scales of ligand-receptor complexes that enabled unbinding from active sites? How these relate to the settings used for MMLP, i.e., could these values be disallowed in terms of used MMLP settings?

- How were the ligand-receptor complexes prepared?
- Why the settings of MMLP were not assessed as the settings of the proposed method?

Unless these questions are answered, I cannot accept the author's interpretation of the results as fully reliable. In other words, the comparison results might still be (and probably are) fair to MMLP, but the argumentation is insufficient.

Besides the main concern the thesis exhibits few small issues which I point out as a feedback for the author:

- Not all proteins possess an active site, see collagen, as you state it in Section 1.2.1.
- In Figure 1.7, the molecule's sticks model seems wrong compared to the vdW model as there are less hydrogens depicted. Further, the depiction of the DOFs seems misleading as the dihedral angles would share the axis of rotation according to it. Additionally, I would suggest to use different colors for carbons and chlorines.
- Probably, by the term *tunnel surface* the author means the surface of the union of CAVER tunnel spheres, but I would suggest to describe it.
- In Figure 4.1, the molecule *2-Chloropropane* is in fact *2-Chlorobutane* since there are four carbons in the hydrocarbon chain.

I have two additional questions related to the thesis:

- In Section 3.4.5, a new ligand conformation generation is described. Do you treat possible similarity of the newly generated conformations? How?
- In the intramolecular potential energy plots, there is an upper threshold of $0.5 \text{ kcal} \cdot \text{mol}^{-1}$. Were the shown data directly produced by AutoDock Vina or were the data somehow modified?

Regarding all the arguments stated above I suggest to accept this thesis as a master thesis and rate it as very good (B).

Ostrava June 4, 2018

Adam Jurčík, Ph.D.