COVER LETTER

Submission ID:	17-0790
Title:	Assessing the Accessibility of Protein Tunnels using Motion Planning
Date:	December 9th, 2017

We would like to thank the anonymous reviewers for their helpful comments. We attempted to address all remarks in the revised version of the paper. In this document, we give detailed responses to all reviewer comments and describe the corresponding changes in our manuscript.

As most of the comments were related to the biochemical background and relevance of our work, we decided to add an extra author to the list of authors, who is the protein engineering expert and helped us to prepare the input datasets, to evaluate the results, and to improve the text according to the reviews.

Based on the comment of one reviewer regarding the stability of the simulations, we computed all the experiments again for a different part of the MD simulation. Therefore, the numbers in the tables changed from the original submission, but the conclusions made in the previous paper remain the same.

Summary of all reviews

1. The ligand conformation library is critical to the contribution of this method, as it relates to prior work. However, there are little to no details given as to how this library is generated/selected, what is the diversity of this library, and how is it stored during RRT construction?

   The ligands used in our paper were prepared by the protein engineers from our collaboration group. The models of the ligands were prepared using the R.E.D. Server Dev, converted to the param files [1] containing the topology, rotatable bonds, atom types, and partial charges. The rotamer library was created by the rotation of all rotatable bonds with 30° steps. The energy cut-off for individual conformations was set to 0.075 (-rot_ensemble_ecutoff 0.075) and the Rosetta database 3.2.1 was used during the conformers generation.

   We added the text clarifying the ligand preparation to the paper. The text is in the first paragraph of the section V-B.

   [1] Jens Meiler and David Baker (2006). "ROSETTALIGAND: Protein-Small Molecule Docking with Full Side-Chain Flexibility" Proteins 65, 538-548.

2. The biophysical realism of the RRT solution is not explained. For example, the scaling is not justified biophysically. Perhaps some scaling (as suggested by

Reviewer 2) can be understood, however, it is never proven to be realistic. A justification, that does not just cite prior work, needs to be provided.

The aim of our proposed work is to help the biochemists to decide if a given protein-ligand pair is suitable for subsequent (usually very time-consuming) analysis.

We are interested in two main properties: a) can the ligand reach the active site? (this is answered by our traversability rate), and b) if not, where is the bottleneck that prevents the ligand to pass the tunnel? The identification of the bottlenecks is necessary to design the modification of the protein, e.g., in order to increase its activity towards the ligand.

We used the scaling-down approach, which is necessary in the case of narrow tunnels. Of course, scaling-down the atoms helps the motion planner to find a trajectory, but we agree that too low scaling-down factors may lead to unrealistic results. We are aware of possible problems when too low scales are used and therefore, we used the low-scale values only in our visualization.

Based on the suggestion of one of the reviewers, we extended our experiments to consider only one scale 0.8 (for both the protein and the ligand), as this scale has been used also in other publications [1]. The traversability rate for the DCP ligand was 48 % in the tested sequence and it was proven that DCP can bind the active site (we refer to our response to comment no. 3). This way of scaling can be considered as realistic.

Some software tools, e.g., Rosetta, can also enable closer interactions between the atoms by using soft-potentials. This has been used, e.g., to study protein docking [2], where the authors scaled down the atoms even to 50% of their original vdW radii.

[1] J. Cortés, T. Siméon, V. Ruiz de Angulo, D. Guieysse, M. Remaud, V. Tran. A path planning approach for computing large-amplitude motions of flexible molecules. Bioinformatics, 21(Suppl.1):i116-i125, 2005.

[2] Zhang, Z., Schindler, C. E., Lange, O. F., & Zacharias, M. (2015). Application of enhanced sampling Monte Carlo methods for high-resolution protein-protein docking in Rosetta. PloS one, 10(6), e0125941.

3. Disappointingly, no direct comparison is made. Rather, the paper mentions that the result is in agreement. A quantitative result should be demonstrated, and the parameters of the MD simulation should be provided. This AE also suggests that the MD be quantitatively compared to the different scaling runs, in order to identify a single scaling value that generally holds and allows the method to validate against MD runs. However, this will require MD runs on additional ligand/receptor pairs.

We analyzed the MD simulation of the wild type of the haloalkane dehalogenase DhaAWT [1] (with PDB ID 4E46 without ligand) and we computed the trajectories for several ligands. Another simulation has been already performed (and published), confirming that the DCP ligand (tested also in our work) can bind to the active site [1]. Unfortunately, other ligands tested in our work were not proved experimentally yet.

As ligands and proteins can influence each other, tunnels can sometimes get wider, allowing the ligand to pass through. However, this flexibility is very hard to model for different proteins and ligands. The flexibility of the tunnels is different if they are close to the protein backbone. If so, they cannot open so easily. Therefore, we think that we cannot come with any universal scale that can be used in all possible scenarios. This is the reason why we considered also the other scales.

Besides detecting how many frames are traversable by the ligand, we aim to identify also the difficult parts of the tunnels (i.e., their bottlenecks). These bottlenecks may temporarily force the ligand to scale down in order to pass through and after that (depending on the protein/ligand combination), the ligand may traverse the the protein with a higher scale. Therefore, we have also presented our visualization which shows the information about the scale. Multiple scales utilized in the paper allow us to visualize these bottlenecks.

[1] Marques, S. M., Dunajova, Z., Prokop, Z., Chaloupkova, R., Brezovsky, J., & Damborsky, J. (2017). Catalytic Cycle of Haloalkane Dehalogenases Toward Unnatural Substrates Explored by Computational Modeling. Journal of Chemical Information and Modeling, 57(8), 1970-1989.

4. The authors have a prior workshop paper on this same topic. A clarification between that work and the submitted work should be provided, even if it is noted that the workshop is non-archival or extensions have been made.

Yes, we presented the following publication at the RoMoCo workshop:

V. Vonasek and B. Kozlikova. Tunnel detection in protein structures using sampling-based motion planning. In 2017 11th International Workshop on Robot Motion and Control (RoMoCo). July 2017, 185-192.

And also one at the MMAR conference:

V. Vonasek and B. Kozlikova. Application of sampling-based path planning for tunnel detection in dynamic protein structures. In 2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR). August 2016, 1010-1015.

The content of these two conference papers is different than the actual RAL submission. In these previous conference papers, we detect static and dynamic tunnels (using sampling-based planning) for spherical ligands (approximated by spherical probes) in a sequence of MD frames. The main goal here was to detect tunnels for the spherical probe.

Contrary, in this RAL paper, we find trajectories for non-spherical ligands, we consider their flexibility, we propose the scaling-down and atom-based techniques. Therefore, we consider tunnel detection (the two conference papers) and the traversability analysis (this RAL paper) as two different topics with different contributions. Namely because of the space limitation of the RAL submission, we decided to omit referencing of our previous workshop papers.

5. Scaling of ligand (up to half its size) is problematic and not adequately
   supported, need better defense that this compensates for keeping the protein
   static, otherwise this is nonsensical from a motion planning perspective.

   This issue is discussed in detail in our comment to reviewer's question no. 18 below.

6. No experimental validation, absolutely necessary for this application domain;
   existing results are not clarified in terms of goals/objectives, observations,
   and inferences.

   This issue is discussed on several places in the detailed answers.

7. Needs support for using a method declared as inadequate because it only uses a
   spherical probe as the basis for this method, in other words, why are tunnels
   defined by spherical probes appropriate as a basis for narrowing the search for a
   flexible ligand?

   There are already dozens of research publications, published in top biochemical and
   biological journals, describing the usage of computational tools detecting tunnels using a
   spherical probe, thus confirming the biochemical relevance of this approach. Of course, the
   tunnels detected by the spherical probe describe the actual void space only very roughly
   and they detect only those tunnels which can accept the bounding sphere of the desired
   ligand. So, the first impression would be that these methods omit many feasible paths which
   can be taken by the actual shape of the ligand. However, in practice, the size of the
   spherical probe is usually much smaller than the actual bounding box of the ligand because
   when the ligand traverses the tunnel, it influences the amino acids in the interaction distance
   and causes their shift. Therefore, the spherical probe based methods provide the
   biochemists with feasible results. However, when studying the tunnel behavior over time,
   these methods still have many limitations and other methods, such as other proposed one,
   should be subsequently used to verify the tunnel shape stability and permeability over time.

   Just as a side note, narrowing the search space helps to significantly reduce the
   computational time and resources, traversing the entire void space inside protein would lead
   to unacceptable costs and in such case, we believe that molecular docking would be already
   a better option for protein void space exploration.

8. Study of which modifications yield which performance gains is needed (atom-based
   distance metric, selecting conformations of various packing properties, region
   restriction, etc).

   To fulfill this requirement, we extended our set of experiments. The proposed method (called
   A1 in the experimental section) was tested in two variants: with the atom-based metric (A1)
   and with the classic Euclidean metric (A1*). The new results are summarized in Table 3. A1
   is able to find the path in more frames than A1* in the difficult cases, i.e., for ligands with
   more atoms and in the case of higher scales (e.g., m040 with 17 atoms at ligand scale 0.8:

A1 detected the path in 6 frames, A1* in 0 frames). On the other hand, both methods A1 and A1* have the same behavior in the case of small ligands and especially when the ligand is scaled to 0.5 or 0.6. We conclude that the proposed atom-based metric should be used when larger scales of the ligands are considered (which is the preferred case, as larger scales lead to more realistic trajectories).

9. Why is energy never considered? This is often critical in these applications as it is not a purely geometric problem.

We completely agree with the reviewer that energy and forces are a factor which would shift the relevance of the detected tunnels to another level. However, including these into the calculation causes significant increase of computational complexity. Additionally, the purely geometric solutions proved to provide the users with sufficiently precise results when studying the overall protein void space and the detected tunnels mostly correspond to the actual paths taken by the ligands. Therefore, we still believe that our improved geometric solution represents an interesting tradeoff between the widely used Voronoi-based geometric tunnel detection approaches and robust and costly solutions taking into account the energies (such as molecular docking).

10. What happens when the litanies are more elongated (larger deviation in atomic distances)?

Actually, the situation when the ligand is elongated is exactly the case when our method can be advantageous in comparison with the spherical-based approaches to tunnel detection. Of course, the preciseness depends on the size of the rotamer library and the capability of the ligand to easily traverse between these states.

11. Missing related work on biasing sampling to known regions of space, all workspace guided methods do this and there are many available, here are just a few:

D. Hsu, T. Jiang, J. Reif, and Z. Sun. Bridge test for sampling narrow passages with proabilistic roadmap planners. In Proc. IEEE Int. Conf. Robot. Autom. (ICRA), pages 4420–4426, 2003.

This reference was already present in the submitted version (under [14]), but the other authors were not correctly cited. We updated this reference.

van den Berg, J.P., Overmars, M.H.: Using workspace information as a guide to non-uniform sampling in probabilistic roadmap planners. Int. J. Robot. Res. 24(12), 1055–1071 (2005)

Thank you for this reference, we added this to the related work.

An efficient retraction-based RRT planner, L. Zhang, D. Manocha, In Proc. Int. Conf. Rob. and Auto. (ICRA), 2008.

This reference was already in the submitted version (under [37]).

A General Region-Based Framework for Collaborative Planning, Jory Denny, Read Sandstrom, Nancy M. Amato, In Proc. Inter. Symp. On Robotics Research (ISRR), pp. 563-579, Genova, Italy, Sep 2015.

We extended our related work by this publication. In fact, the idea of this paper (the user can identify difficult regions and even the planners suitable for them) is also used in this RAL submission. In our work, the users (protein engineers) decide which tunnel has to be analyzed. This in fact helps to focus the sampling of the configuration space only on a smaller part.

12. Missing related work on other mp techniques for ligand binding and motion:

A.P. Singh, J.C. Latombe, and D.L. Brutlag. A motion planning approach to flexible ligand binding. In 7th Int. Conf. on Intelligent Systems for Molecular Biology (ISMB), pages 252–261, 1999.
Ligand Binding with OBPRM and Haptic User Input, O. Burchan Bayazit, Guang Song, Nancy M. Amato, In Proc. IEEE Int. Conf. Robot. Autom. (ICRA), pp. 954-959, May 2001.

These works utilize the PRM method for the pathway search, and they can also be used to detect the binding sites, so originally we decided not to include them into the original version of our paper. In this revised version, we updated the related work and included these references as well.

13. Subsection headings would help clarify transitions from ligand binding problems to motion planning problems.

We reformulated the beginning of motion planning section in the related work. However, even when we agree that it would increase the readability, we decided to omit the suggested subsection headings, namely because of the page limit. We decided to use the space for adding explanations to more critical comments of the reviewers.

14. Disconnect in flow from introduction to related work.

Corrected.

15. Last 2 paragraphs of results would be better as a separate subsection.

Thank you for this suggestion. We moved the last two paragraphs to the Discussion section.

16. Some parameters need better explanation and support for how they are set: R_init, p_tunnel, d_tunnel.

We updated the last paragraph in Section IV-B: "The distance R_init influences how far the initial configurations are from the start of the tunnel. By placing them too far, the planner would need more iterations to enter the tunnel, which would increase the runtime. Contrary, too low R_init may disable to find any collision-free configurations, especially for large ligands. We suggest to set R_init to the radius of the first sphere in the tunnel and increase it only if no collision-free configuration can be found around."

The parameter p_tunnel determines the probability of generating random samples around the tunnel. Typically, there is not too much space around the tunnel so allowing the tree to sample also outside the tunnel (or even generating the random samples around the protein) does not bring any effect. Therefore, we use the value 0.9.

We added new paragraph to section IV-C to explain how to set up the d_tunnel: "The movements of the ligands inside tunnels are blocked by many atoms forming the tunnel. It is not always possible for the ligand to strictly follow the centerline of the tunnel and slight detours should be allowed, which is controlled by the d_tunnel parameter. This parameter should be set considering the existence of cavities or other tunnels around the tunnel being investigated. In the case of alone tunnels (which is very common), we recommend to use double value of the tunnel average width."

Reviewer 2

17. The MD simulation is not clearly described. What was the force field used? (Which version of AMBER?) What was the MD time step? (maybe 1 fs?) What was the temperature? Were the 100 time frames of MD simulation consecutive time frames, or were they taken at an interval? If the time step was 1 fs and the 100 time frames correspond to 100 fs, are the authors confident that the simulation time was long enough for the protein structure to equilibrate? It is ambiguous whether the authors performed MD for the dynamics of the ligand as it traverses the tunnels or if MD was done only for the protein, in order to understand the tunnel dynamics. It seems from Section V.C that MD was done to test the results for the DCP ligand, but other than stating that MD worked, there is nothing else in the text to substantiate this claim (no results from MD).

Thank you for this valuable comment. We decided to rerun the experiments on a different part of the MD simulation. The used MD simulation did not contain any ligand. The MD simulations were prepared by the protein engineers using the Amber 12 tool with time step 2fs. The total simulation length is 500 ns. We took the 100 frames where the structure was already equilibrated. The details about the MD setup can be found in [1].

Another set of simulations was done separately for the protein and several ligands. These simulations are also described in [1]. For one of the ligands tested in our work (DCP), it was confirmed that it binds to the active site (simulations DhaAwt-DCP in [1]).

[1] Marques, S. M., Dunajova, Z., Prokop, Z., Chaloupkova, R., Brezovsky, J., & Damborsky, J. (2017). Catalytic Cycle of Haloalkane Dehalogenases Toward Unnatural Substrates Explored by Computational Modeling. Journal of Chemical Information and Modeling, 57(8), 1970-1989.

18. While the work has possible applications in the field of biochemistry, where virtual screening of ligands binding to a receptor provides insight on drug discovery, there are still major issues to be resolved before it can be used in biochemistry, as the authors rightly state in the Discussion. The authors justify scaling down ligand atoms' sizes by citing previous works that changed the size of the hard-sphere van der Waals radii of atoms (refs. [6],[8],[12] in the paper, but it is not clear if ref [8] does rescaling explicitly). In refs [6],[12] a reduction of 20% of VDW radii is justified by the collision check: If atomic radii have been reduced to 80% of their size and are still in collision, then this collision truly is unphysical and configurations in collision are prohibited. In other words, if one needs to reduce models below 80% of their size in order to avoid collisions, then the resulting models will be unphysical. So it follows from the same argument given in refs [6],[12], that any reduction below 80% of the original atoms' VDW radii should be forbidden. This is problematic since the central results of this paper focus on atomic radii reductions of 20% to 50% for the ligands. Have the authors considered reducing not only the ligand's but the protein's atoms as well? Since the limit of a 20% reduction is well justified by the previous works, ligand and protein atoms could be reduced by no more than 20% (s_min in [0.8,1]). If the protein atoms are reduced, then the tunnels could be wider and the ligand might not need to shrink too much. In this scenario, possibly a s_min=0.8 might lead to more successful results.

Yes, we have considered scaling both ligand/protein as well. In fact, our algorithm can scale-down both ligand and protein atoms easily (similarly, e.g., as in the MoMa-LigPath tool). We decided to present the results only when the ligand is scaled-down and the protein atoms are used at their original size.

We completely agree with the reviewer that scaling down more than this threshold can result in non-realistic results. As we state also in the answer to comment no. 3, we use the other scaling factors to detect/visualize the bottlenecks (Fig. 4 in the paper). By this we can estimate how difficult the bottleneck is (and how often it appears). By using only one single scale of the ligand, we can only determine the traversable/non-traversable frames.

This is the reason why our algorithm does not use a single scale, but all scales from s_min to 1.0 and it always prefers to expand the tree using larger scales, if possible. The resulting trajectories can be then visualized (e.g., as in Fig. 4) to see the temporal scaling-down required to pass through the bottleneck.

We also considered the scaling of the protein. We extended our experiments by measuring the traversability rate for protein-scale = 0.8 and ligand-scale = 0.8 (similarly to MoMa-LigPath). We expected that the results will be the same as for protein-scale = 1 and ligand-scale = 0.6, which was confirmed. The results are summarized in Table 3, last column (A1p).

Based only on the traversability rates (detected by A1p), also other ligands (e.g., m037t or m038t) seem to be promising, especially if the confirmed ligand DCP (we refer to answer 3) had only 18% of traversability at scale 0.8. However, we cannot confirm this by referencing to any paper.

Scaling down more than 80% seems to be inevitable when computing trajectories on MD simulations that were performed without the ligand, which was our case. The protein tends to fill the void space and the tunnels are very narrow. We propose to utilize the lower scales for the purpose of visualization (to detect possibly difficult regions in the tunnels where the ligand is forced to scale down).

Based on the suggestion of the reviewer, we also removed the reference [8] from the sentence "... the free-space can be dilated by shrinking the atom radii, similarly to … ", as it is really not clear if [8] performs the scaling-down step.

19. The atom-based metric is interesting, but it isn't clear how it considers ligand shapes. A figure would be helpful. How are rotational constraints imposed? Have the authors considered, instead of taking the shortest Euclidean distance overall between any pair of atoms (atom I in q to any atom j in q'), picking the shortest distance between the same atoms in each ligand (i in q to i' in q')? Could this constrain how much a ligand can rotate from one node to another, avoiding large motions?

For a given 3D position and 3D orientation, this metric depends on the shape of the ligand (that is determined by its conformation), because the metric is computed as the distance between two nearest atoms. We update the corresponding text in Section IV-C: "This metric is sensitive to the shape of the ligand (that depends on its conformation) as it computes the distance using the two nearest atoms of the two configurations. It allows us to expand the tree by a configuration whose nearest atom approaches its nearest atom in q_rand, which supports the retraction of the ligands towards q_rand."

Computing the distance using same pairs of atoms would be faster than our approach (which requires to determine the closest pair of two atoms). The rotation of the ligands is in fact very limited in the narrow tunnels. Moreover, the resolution of our planner is 0.05 Ångströms and the ligand cannot rotate much in such a short distance.

20. Why, when computing the initial configurations, is the initial scale set to s_min instead of 1 (full-size)?

The initial configurations are placed at the beginning of the tunnel. As the tunnels are computed using the CAVER tool, which terminates the tunnel very close to the surface of

the protein, it is not always possible to place the ligands close to beginning of the tunnel at scale 1. In such a case, the ligands need to be placed closer to the beginning of the tunnel and in order to get its path to the protein, the ligand has to be scaled down anyway. To save some computational resources, we decided to place the ligand of the minimal scale (s_min) to the initial configuration.

21. It would be useful to discuss more on ligand conformations, eg. How many are actually chosen by the algorithm in each trajectory?

We added a new table (Tab IV) showing the average number of the conformations used in the trajectories. The number of chosen conformations is less than the number of available (e.g., only ~36 out of 50 available conformations are used for the DCP ligand at scale 0.6).

22. Adding percent signs (%) to the numbers in the tables would be helpful to the reader.

We extended the Table 3 by additional measurements (new columns A3 and the column A1p). To add '%' we would have to decrease the space between the columns which would negatively influence the readability of the table.

23. How many trajectories are shown in Figure 4? Are these all successful trajectories? What was the method used to find these trajectories (A1? A2? MD?). Are the results shown in Figure 4 the same as reported in Table III?

For the demonstration purposes we are not visualizing all the trajectories, otherwise it would lead to visual clutter. Our visualization tool automatically clusters the trajectories according to their similarity and shows only a subset of them (it is also faster than showing all of them). The depicted trajectories were computed by the A1 method and for the DCP ligand.

As only 8% of the results are successful for DCP using A1 (s_min=0.8), what do the successful trajectories have in common? Does the success rate depend on the particular conformations that are added to the tree? Does success depend on initial conformations?

These questions are in fact the reason why we calculate the trajectories for various scaling factors, so we can – by the visual inspection in as the first tool – determine how the 'scaled' ligands perform in the protein. The detail analysis is however beyond the scope of this paper, where we want to focus only on the motion planning part. We already observed, that the RRT tree is built – if it is allowed by the void space inside the protein – using most of the ligand configurations (the number of used conformations in the successful trajectories, which of course relates to the number of conformations used in the whole tree, is shown in Table IV. The success rate depends on the initial configurations, especially in the case of narrow tunnels.

24. Other than the issues stated, the paper would also benefit from an English style and grammar review.

The paper was thoroughly revised by all the co-authors and we believe that we revealed and corrected most of the language issues in the paper.

Reviewer 4

---

25. The paper utilizes a set of ligand configurations (L) but provides no method and no citation as to how these configurations were generated. Sect. IV-D provides no metrics on the variance of the deviations from center. Without any analysis of the diversity contained within the libraries, the sampling strategy employed seems to simplistic. In that ligand configurations with a unique (well separated) deviation from center measurement could be missed and replaced with highly redundant configurations. No cardinality information is given for the protein/ligands listed in table 3.

The ligands are prepared by chemical experts by sampling the rotatable bonds and selecting only conformations with low energy. We updated the text in Section IV-A: "The candidate conformations are prepared by discrete sampling of degrees of freedom of the ligand. Each candidate conformation is ranked according to its energy and only low-energy conformations are used in L."

And in more detail in the Section 5-B: "Models of the ligands in Mol2 formats were converted to param files [26], containing the topology, rotatable bonds, atom types, and partial charges. The pool of conformations L was created by the rotation of all rotatable bonds with 30° step using Rosetta3. The energy cut-off for individual conformations was set to 0.075."

This way of preparation the ligand library can still result in too many conformations (hundreds or even thousands). It depends on the resolution of sampling (we use 30 degrees per bond in our case). For fast planning, we further select a subset of conformations from this pool. These conformations should have diverse shapes - from 'elongated' to 'packed'. Therefore, we measure how the atoms of the ligands are spread around the geometric center of the ligands. A simple approach is to compute the deviation of the distance of atoms center from the geometric center of the ligand.

We described this process in the following paragraph (last paragraph of Section IV): "The pool of available conformations L consists all conformations (up to some resolution) with energy lower than a user specified threshold, which can still contain too many conformations (thousands or even more). We select a subset by computing how the atoms of the ligand are spread around its geometric mean. The distance of centers of ligand atoms to the geometric center of the ligand is computed, and the deviation of these distances then characterize the spread. Low deviation is characteristic for the packed conformations, while larger deviations correspond to the longer conformations. To select n conformations from all

|L| available conformations, they are sorted according to the deviation. Each k−th conformation from the sorted list, where k = \floor |L|/n, is used".

26. Although work has been done about ligands exiting proteins with ML-RRT, finding the ligand entrance path is much harder (since the first problem starts with the ligand in a bound state, and thus, you know the ligand and the protein interact). The paper would be strengthen by mentioning that the prior work cannot act as a ligand "filter" (which is one of the goals of this work).

   We thank the reviewer for pointing this out. We added the following sentence to the Introduction section:
   "Another advantage of our method in comparison with the existing approaches is that we enable the users to trace the ligand path leading from the outer solvent towards the protein active site."

27. Continuing from the point above, given the relatively quick run time (under 1 hour according to fig 3), analyzing more than 50 ligand configurations seems prudent. How does increasing this size impact the scaling factor and the number of solutions found. This choice of 50 seems to have been made independent of the size of ligand library.

   We performed an experiment to investigate how the number of conformations influences the success rate. For all tested ligands, using more than 50 conformations leads to the same results (according to the number of traversable frames). Therefore, we decided to use 50 conformations. We also created a new table (Tab. IV), where we show how many conformations were used in the successful trajectories. For example, for the DCP ligand, only 36.6 (at scale 0.6) and 31.4 (at scale 0.8) where used. These numbers are lower than the number of available conformations. The largest number of conformations was used for the m056 (~46) and m037t (~46).

   The minimal scaling factor allowing RRT to find a trajectory may depend on the number of conformations, but only when several conformations are used. For example, if the trajectories are found only using two conformations (one 'elongated' and one 'packed'), then the algorithm may require to scale-down the ligand more than if also other conformations ranging from 'elongated' to 'packed' are used.

   Using more than 50 conformations might have sense in the case of ligands with more degrees of freedom.

28. For a pair of connected nodes in the search tree, these nodes could have different ligand configurations (lets call them L1 and L2). Do the authors assume that the ligand can transition from L1 to L2 within this very small space (epsilon = 0.05). Is this reasoning based on the fact that L1 is collision free at its location and L2 is collision free at its location? It would seem that in some cases (where the ligand goes from an elongated form to a compact one) that

this assumption may not hold. At a minimum, this concern needs to be addressed in the text.

Yes, we assume this. We updated the text accordingly (last sentence in Section IV-A).

29. The end of the introduction indicates the wet lab experimental data was analyzed. Molecular dynamics is not a "wet lab" technique, but rather a computational technique. Thus, this statement is very misleading.

The reviewers are right that this can be very misleading for the reader. Therefore, we corrected this part and changed the sentence to more general statement about the biochemical approach to this problem as there is no space to explain their workflow in detail. We changed the sentence in the following way:
"Our proposed method was tested by the biochemists who compared the results with already known trajectories of ligands, determined by their traditional approaches."

30. No analysis of the distance traveled from the Qinit configurations to the active site are provided. At a minimum, the experiments outlined in table 2/3 should include the number of spheres considered in each problem. This goes directly towards validating the premise of the planner modifications. This makes the thresholds of 0.5 A for each planner and 2.0 A for each traversability analysis (last sentence of Section V) very difficult to interpret. Further, for the larger ligands, is the final configuration compact of elongated. Is it the center of mass that is measured for proximity to the active site, or just the closest atom in the ligand?
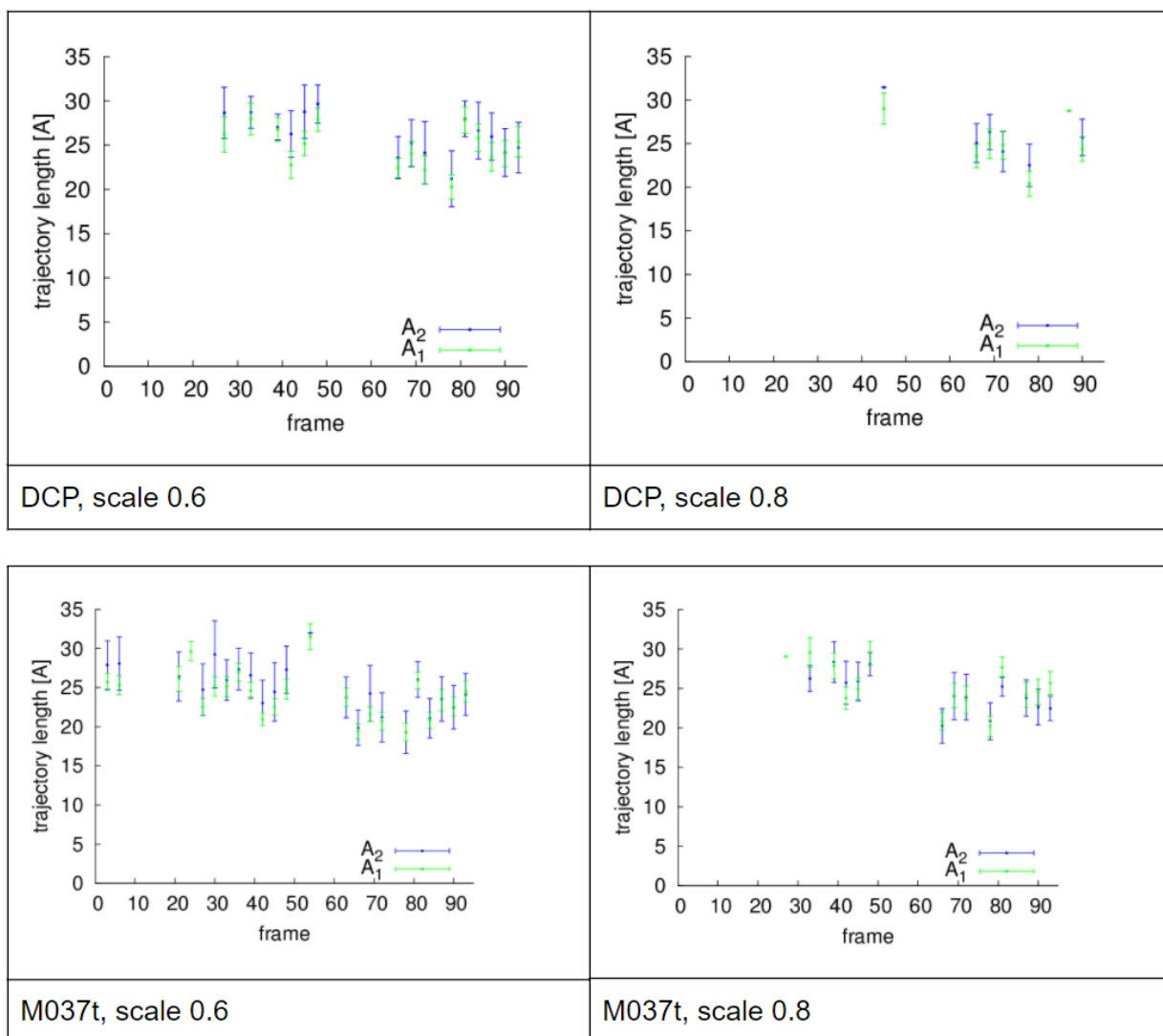
The number of protein atoms was really missing. We added it to the text to the beginning of Section 5 ("The proposed method was used to compute the ligand traversability towards the active site using two main tunnels in the Haloalkane dehalogenase protein (PDB ID 4E46) which consists of 4650 atoms.") and we also repeated this number in Tables 2/3. The number of spheres (atoms) of the ligands are included in the ligands' names (e.g., m037t/8 - this ligand has 8 atoms).
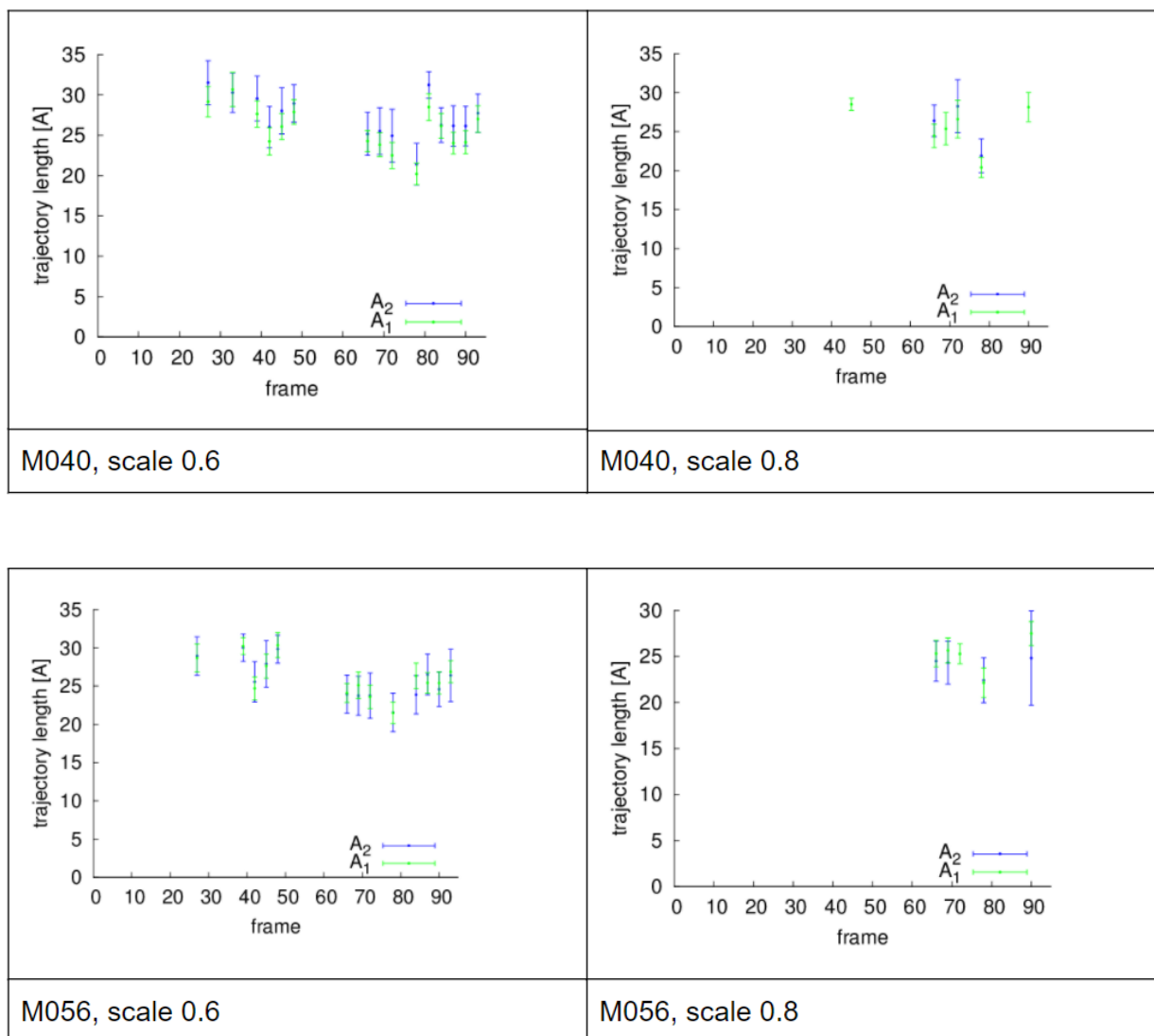
Thank you for pointing out the problem with two thresholds 0.5 and 2 A. We performed all experiments again with only one threshold – 2 A (which was already used in the original paper to compute the traversability rate). We updated the information about terminating the planners ("The planning was terminated after I_max iterations or if the tree approached the active site to the distance less than 2 A (i.e., the geometric center of the ligand distance less than 2 A from the active site").

The distance to the active site is measured as the 3D distance between the geometric center of the ligand and the 3D position of the active site. We updated the information about computing the traversability rate: "This distance is measured as the 3D distance between the geometric center of the ligand and the position of the active site."

The distance traveled by the ligands is around 27 A – it varies through the MD frames, as the main tunnel also moves and its length changes. The traveled distance is similar for all tested ligands. Unfortunately, we cannot include all images to the paper (due to the page limit), but we included them to this cover letter.

We show the average (+std.dev) of the trajectory length for several frames. In the following graphs, each third-frame is shown to maintain the graphs readable. The trajectories are shown only for frames, where the ligand visited the active site to the distance less than 2 A (measured by the above discussed metric). The traversability rate decreases with the large scale of the ligands, therefore, the following graphs with scale = 0.8 have less number of data than the graphs with ligand scale 0.6.

DCP, scale 0.6

DCP, scale 0.8

M037t, scale 0.6

M037t, scale 0.8

M040, scale 0.6

M040, scale 0.8

M056, scale 0.6

M056, scale 0.8

We analyzed the conformations at the active site for the ligand m040 which has 17 atoms. We show how often the conformations reached the active site by the following histograms for scale 0.6 (left) and 0.7 (right). The conformation No. 0 is the most packed one, the conformation No. 50 is the most elongated one. The histograms are computed aver all 100 MD frames. We do not observe that either packed or elongated conformations dominate when the ligand reaches the active site. This is probably caused by the protein dynamics, as it causes also the change of the void space in the active site, which allows various conformations to reach it.