

## New strategy for optimizing knowledge-based docking parameters: application to ssRNA-RRM docking

<sup>1</sup>\*Kravchenko A., <sup>1</sup>Smail-Tabbone M., <sup>1</sup>Chauvot de Beauchene I., <sup>2</sup>De Vries S.J.

<sup>1</sup>LORIA, campus scientifique, 54000, Nancy, France.

<sup>2</sup>RPBS, 35 rue Hélène Brion, 75013, Paris, France

Corresponding Author: [anna.kravchenko@loria.fr](mailto:anna.kravchenko@loria.fr)

### Topic: 1 - RBP-RNA binding

Docking is a computational prediction of a 3D structure of a molecular complex. It is a great tool to get an idea of the 3D structure of a complex relatively fast and much cheaper, compared to the other types of *in silico*, and even more so, *in vitro* experiments. We are interested in fragment-based docking for the single-stranded RNA-RRM (RNA binding motif) complexes. Generally speaking, flexibility is a scourge of docking. Why? Because it increases docking complexity and decreases its reliability. High flexibility leads to a very large number of molecular conformations, and thereof of putative docking models, processing of which is too expensive computationally. Hence, ssRNAs are a challenging target. Previously we had developed a method to model ssRNA bound to protein. It splits RNA into fragments of manageable conformational diversity, docks them individually from random positions using gradient descent, and then assembles the compatible docked poses back into the ligand. The current parameters of its energy function for RNA-protein atom-atom interaction can be optimized for the ssRNA-RRM case.

Here we present a novel histogram-based optimization approach: convert the energy function into the relative probabilities of each atom-atom distance in RRM-RNA interfaces, adjust these probabilities by their log odds ratio in correct/incorrect docking poses, convert them into new parameters of the energy function, and dock again (iteratively). Upon convergence, this procedure should generate almost equal distributions of atom-atom distances in correct/incorrect poses, thus maximising the proportion of correct poses. Our first results on a small data set consisting of a few complexes of known structures are very promising. Obtained poses could be further identified on other criteria (e.g. angles between atoms) with the help of machine learning.

With updated parameters, ssRNA-RRM docking can become a powerful tool for solving ssRNA-RRM complexes. Combined with AlphaFold for the determination of RRM structure, ssRNA-RRM docking can be performed by a researcher with little to no knowledge of structural biology, making it accessible to the wide scientific community.

### **Acknowledgements**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813239.