Essential Steps in Protein-Ligand Preparation With Conserved Waters and Water Networks in Drug Design

Marko Jukić [1,2*], Urban Bren [1,2,4]

[1]*Faculty of Chemistry and Chemical Engineering, University of Maribor (UM FKKT), Smetanova Ulica 17, 2000 Maribor, Slovenia*
[2] *Faculty of Mathematics, Natural Sciences and Information Technologies; University of Primorska (UP FAMNIT), Glagoljaška 8, SI–6000 Koper, Slovenia.*
[3] *Institute of Enviormental Protection and Sensors, Beloruska Ulica 7, SI-2000 Maribor, Slovenia*

correspondence: marko.jukic@um.si

Keywords: in silico drug design, CADD, conserved waters, water networks, protein ligand binding, structure based drug design

Abstract: Drug design is leveraging modern *in silico* techniques, chemoinformatics approaches and molecular modelling to inform early research steps. We have noticed several essential input data preparation steps are being omitted or even not recognized in recent scientific literature and are advocating towards critical data assessment and input preparation, especially involving solvent molecules in the studied systems. The detection of conserved water molecules inside macromolecular structures is a crucial aspect in the fields of medicinal chemistry and structural biochemistry. As eloquently put by Kubinyi et al., conserved waters play a key role in the formation of hydrogen bonding networks between targets and their ligands. Moreover, water has a structural function in macromolecules and their interfaces, acts as a catalyst and facilitates molecular recognition. The significance of investigating conserved water molecules has been previously established, by our prior research on ProBiS H2O where we leverage crystal water data for conserved water identification. Furthermore, we have generalized the tool to include water trajectory data from MD experiment to show similar conservation analyses can be performed. To expand on the method, we incorporated the information on the orientation of hydrogen atoms. The provided information is crucial for evaluating the sustainability of water networks formed upon ligand binding. The recently developed approach is thoroughly described and validated using established results from the scientific literature. Last but not least, we further generalized the method to identify ions and other conserved entities in our newest developed tool for drug design: MADE.

Bibliography:

[1] Jukic, M., Konc, J., Gobec, S., & Janezic, D. (2017). Identification of conserved water sites in protein structures for drug design. Journal of Chemical Information and Modeling, 57(12), 3094-3103.
[2] Jukič, M., Konc, J., Janežič, D., & Bren, U. (2020). ProBiS H2O MD approach for identification of conserved water sites in protein structures for drug design. *ACS medicinal chemistry letters*, *11*(5), 877-882.
[3] Tosovic, J., Fijan, D., Jukic, M., & Bren, U. (2022). Conserved water networks identification for drug design using density clustering approaches on positional and orientational data. *Journal of Chemical Information and Modeling*, *62*(23), 6105-6117.
[4] Ravnik, V., Jukič, M., & Bren, U. (2023). Identifying Metal Binding Sites in Proteins Using Homologous Structures, the MADE Approach. *Journal of Chemical Information and Modeling*, *63*(16), 5204-5219.