XIII Argentine Congress of Bioinformatics and Computational Biology
XIII International Conference of the Iberoamerican Society of Bioinformatics
III Annual Meeting of the Ibero-American Artificial Intelligence Network for Big BioData

Advancing Computational Modeling of Protein-Carbohydrate Interactions using Water Sites Biased Docking

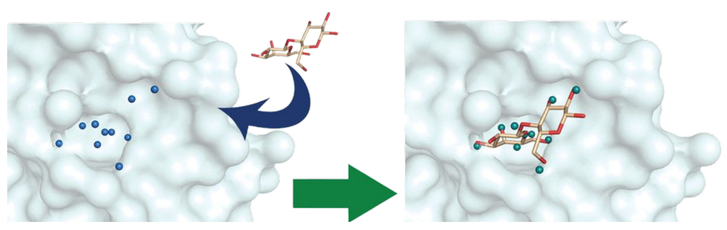
Lannot, J. O. [1,2,#] ; Rey, E. [1,2,#]; Gamarra, M. [1,2] ; Martí, M. A.[1,2]; Modenutti, C. P. [1,2,*]

1: Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. II (CE1428EHA), Buenos Aires, Argentina
2: Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales (IQUIBICEN) - CONICET. Ciudad Universitaria, Pab. II (CE1428EHA), Buenos Aires, Argentina
*: corresponding author. **Contact:** *cmodenutti@qb.fcen.uba.ar* // #: these authors contributed equally to this work

Contact: *jo.lannot@gmail.com*

INTRODUCTION

Molecular docking is a widely used computational method for modeling ligand-receptor interactions, primarily focusing on small rigid drug-like molecules. However, its application to flexible carbohydrate molecules often leads to suboptimal performance due to their hydrophilic nature, low affinity, and ligand conformational flexibility (1, 2). To address this, our study aims to develop a precise method capable of predicting both oligosaccharide conformations and relative binding energies to their receptors. To improve the conventional method (CADM, from *Conventional Autodock Docking Method*), it was proposed to use information derived from the solvent structure at the ligand binding site to modify the scoring function of several docking programs (**WSBDM, for Water Sites Biased Docking Method**) (2, 3, 4, 10). This was possible due to the recently developed ***Autodock Vina Site*** program, which allows the application of user-defined biases to the scoring function.



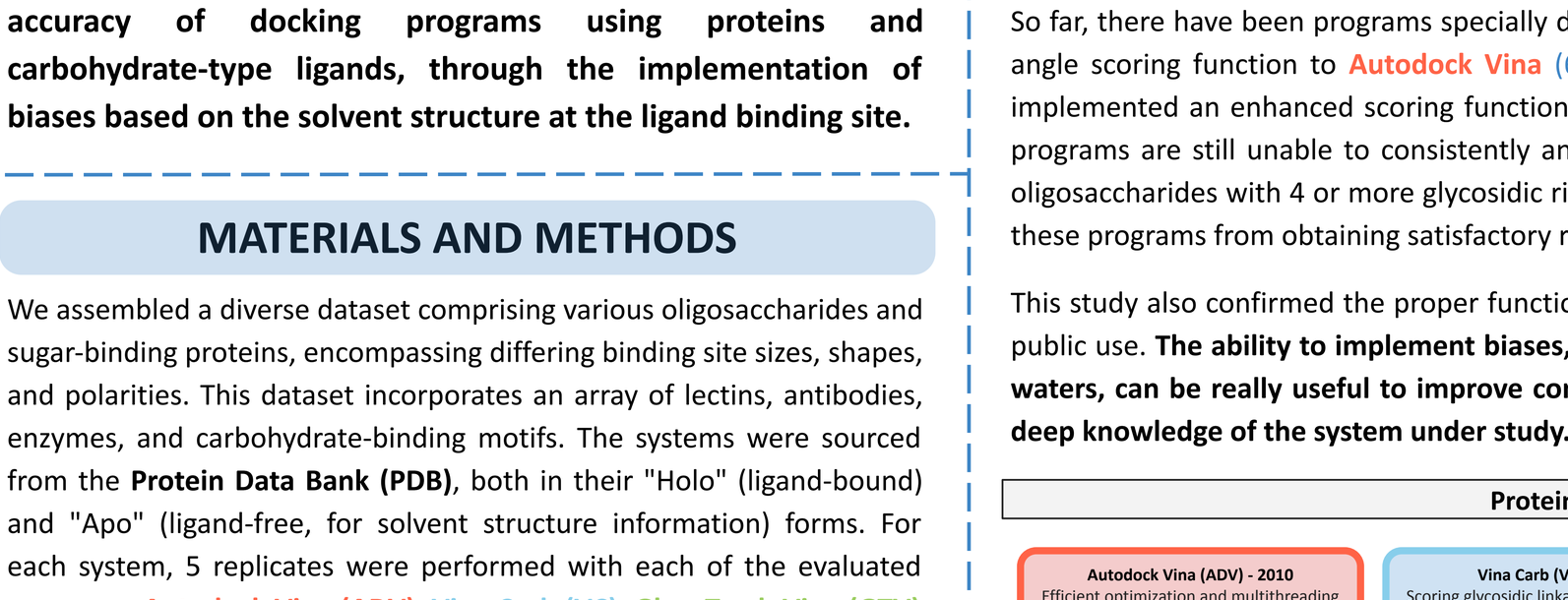
(3)

The ultimate goal is to **significantly increase the precision and accuracy of docking programs using proteins and carbohydrate-type ligands, through the implementation of biases based on the solvent structure at the ligand binding site.**

MATERIALS AND METHODS

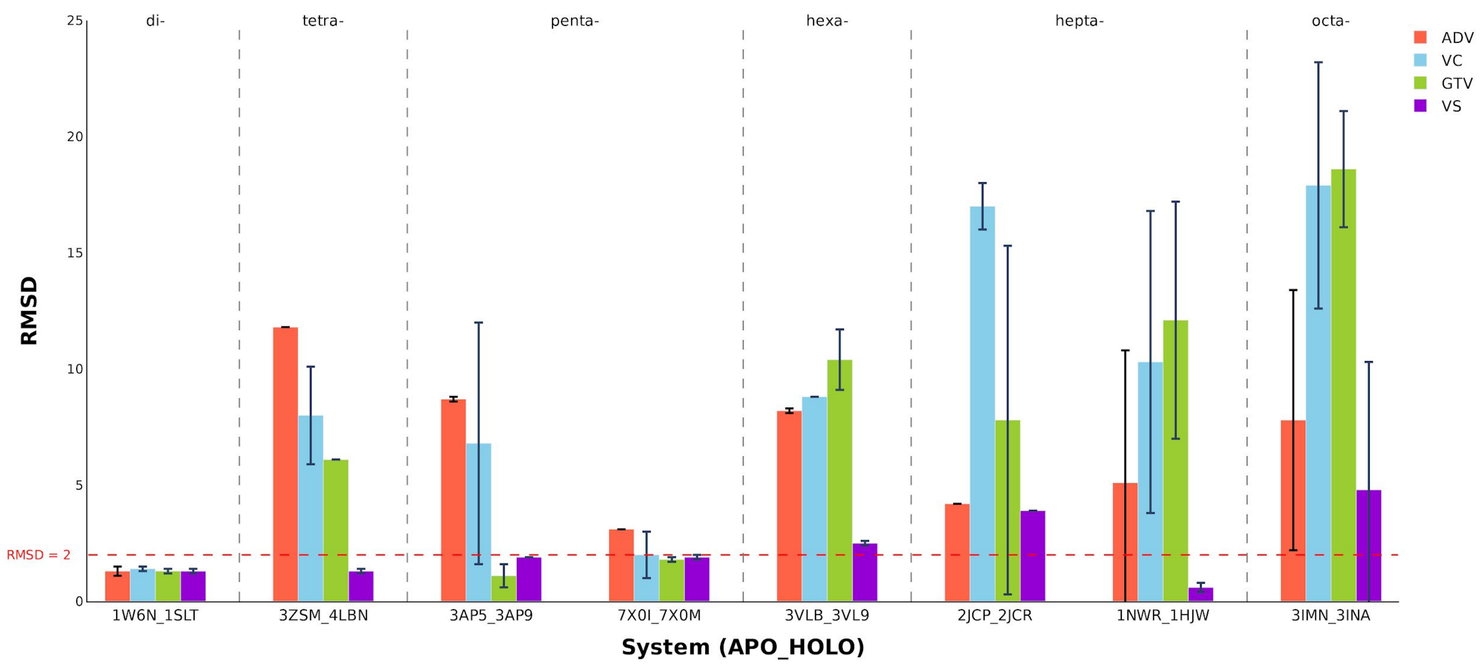
We assembled a diverse dataset comprising various oligosaccharides and sugar-binding proteins, encompassing differing binding site sizes, shapes, and polarities. This dataset incorporates an array of lectins, antibodies, enzymes, and carbohydrate-binding motifs. The systems were sourced from the **Protein Data Bank (PDB)**, both in their "Holo" (ligand-bound) and "Apo" (ligand-free, for solvent structure information) forms. For each system, 5 replicates were performed with each of the evaluated programs: **Autodock Vina (ADV)**, **Vina Carb (VC)**, **GlycoTorch Vina (GTV)** and **Vina Site (VS)**. The protocol parameters were set according to the bibliography for each program. The performance of the programs was evaluated using *Root Mean Square Deviation* (RMSD).

Re-docking using crystallographic structures (*model: human Galectin-3*)



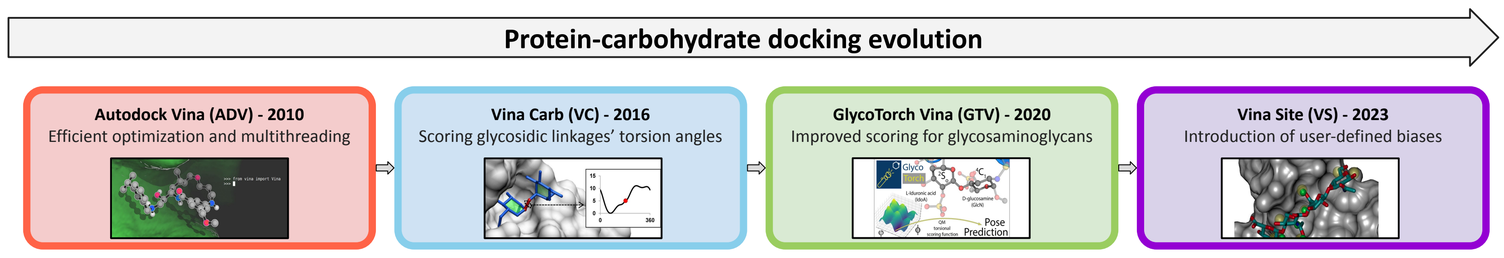
RESULTS

The Water Sites Biased Docking Method (WSBDM) exhibited significant enhancement in precision and accuracy across the analyzed programs for the studied systems. By incorporating solvent structure data from the receptor's ligand binding site, our method yielded predictions closer to experimental observations. Most notably, substantial improvements were evident in systems involving large ligands, where other programs tend to underperform.



So far, there have been programs specially designed to work with carbohydrates. Vina Carb introduced a dihedral angle scoring function to Autodock Vina (6, 7, 8), significantly improving results. Years later, GlycoTorch Vina implemented an enhanced scoring function specific to glycosaminoglycans (9). Although results improved, these programs are still unable to consistently and reliably produce results for systems with large ligands, specifically oligosaccharides with 4 or more glycosidic rings. The vast number of possibilities for these types of ligands hinders these programs from obtaining satisfactory results (1, 7, 9).

This study also confirmed the proper functionality of the new program, Vina Site, which will soon be available for public use. **The ability to implement biases, although in this case was used with the criterion of crystallographic waters, can be really useful to improve computational predictions based on different criteria coming from the deep knowledge of the system under study.**



CONCLUSIONS

Proteins and carbohydrates are two of the main types of biomolecules in living organisms, and their interaction plays a fundamental role in a wide variety of biological processes. The interaction between these molecules can regulate protein activity and function, influence membrane structure, and participate in cellular signaling, among other vital processes (5). **WSBDM, using the new program Vina Site, not only enhances the comprehension of protein-carbohydrate interactions but also improves the reliability, quality, and reproducibility of computational predictions. Its utilization promises a more nuanced understanding of molecular-level protein-ligand interactions, significantly impacting the field of glycobiology**. Furthermore, it holds potential for improving carbohydrate docking predictions in homology-based models when crystal structures are unavailable, which nowadays is essential for understanding key biological processes and for the development of glycomimetic drugs.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to the ***Universidad de Buenos Aires*** and the ***Consejo Nacional de Investigaciones Científicas y Técnicas*** for their invaluable support and resources throughout this research. Their contributions were crucial in the successful completion of this study.

REFERENCES

1: Modenutti et al. (2015). **Using crystallographic water properties for the analysis and prediction of lectin-carbohydrate complex structures.** Glycobiology, 25(2), 181-196. doi: 10.1093/glycob/cwu102
2: Blanco Capurro et al. (2019). **An efficient use of X-ray information, homology modeling, molecular dynamics and knowledge-based docking techniques to predict protein-monosaccharide complexes.** Glycobiology. doi: 10.1093/glycob/cwy102
3: Saraboji et al. (2012). **The carbohydrate-binding site in galectin-3 is preorganized to recognize a sugarlike framework of oxygens: ultra-high-resolution structures and water dynamics.** Biochemistry, 51(1), 296–306. doi: 10.1021/bi201459p
4: Arcon et al. (2019). **AutoDock Bias: improving binding mode prediction and virtual screening using known protein-ligand interactions.** Bioinformatics, btz152. doi: 10.1093/bioinformatics/btz152
5: Gabius et al. (2011). **From lectin structure to functional glycomics: principles of the sugar code.** Trends in Biochemical Sciences, 36(6), 298-313. doi: 10.1016/j.tibs.2011.01.005.
6: Trott et al. (2010). **AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.** Journal of Computational Chemistry, 31(2), 455-461. doi: 10.1002/jcc.21334
7: Nivedha et al. (2016). **Vina-Carb: Improving Glycosidic Angles during Carbohydrate Docking.** Journal of Chemical Theory and Computation, 12(2), 892-901. doi: 10.1021/acs.jctc.5b00834
8: Nivedha et al. (2014). **Importance of ligand conformational energies in carbohydrate docking: Sorting the wheat from the chaff.** Journal of Computational Chemistry, 35(7), 526-539. doi: 10.1002/jcc.23517
9: Boittier et al. (2020). **GlycoTorch Vina: Docking Designed and Tested for Glycosaminoglycans.** Journal of Chemical Information and Modeling, 60(12), 6328-6343. doi: 10.1021/acs.jcim.0c00373
10: Samsonov et al. (2011). **Docking glycosaminoglycans to proteins: Analysis of solvent inclusion.** Journal of Computational Aided Molecular Design, 25(5), 477-489. doi: 10.1007/s10822-011-9433-1