



Fig.2.The pipeline for finding unbound parts for complexes from SAbDab database

The current results of the pipeline's work on SAbDab database are as follows. The gathered validation set of antibody-antigen complexes has 100 complexes, for which the pipeline has given a definitive answer that it found required unbound parts, and 74 complexes more, for which the pipeline marked all the unbound parts as in need of human's assessment.

IV. CONCLUSION AND FUTURE WORK

We have shown that the use of statistical potential significantly improves the performance of HEDGE, a novel GPU-driven docking algorithm developed by BIOCAD. This makes the development of a new statistical potential that favors complexes of type antibody-antigen reasonable.

We have also presented a plan for developing said statistical potential along with the plan for its comparison to existing solutions, such as DARS and aADARS.

A validation set for comparison of statistical potentials has been gathered using a developed pipeline for finding unbound parts of antibody-antigen complexes.

As for the future work, by the time of the conference, we plan to develop a DARS-like statistical potential for antibody-antigen complexes and conduct the comparison of developed statistical potential with DARS and aADARS.

REFERENCES

- [1] M.F. Lensink, S. Velankar, S.J. Wodak, "Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition.", *Proteins: Structure, Function, and Bioinformatics*, 85.3, 2017, pp. 359-377.
- [2] Katchalski-Katzir, Shariv, Eisenstein, Friesem, Aflalo and Vakser, "Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques.", *Proceedings of the National Academy of Sciences*, 89(6), 1992, pp.2195-2199.
- [3] D. Kozakov, R. Brenke, S.R. Comeau and S. Vajda, "PIPER: an FFT-based protein docking program with pairwise potentials.", *Proteins: Structure, Function, and Bioinformatics*, 65(2), 2006, pp.392-406.
- [4] C. Zhang, G. Vasmatzis, J.L. Cornette and C. DeLisi, "Determination of atomic desolvation energies from the structures of crystallized proteins.", *Journal of molecular biology*, 267(3), 1997, pp.707-726.
- [5] K. Olechnovič and Č. Venclovas, "VoroMQA: Assessment of protein structure quality using interatomic contact areas.", *Proteins: Structure, Function, and Bioinformatics*, 85(6), 2017, pp.1131-1145.
- [6] G.Y. Chuang, D. Kozakov, R. Brenke, S.R. Comeau and S. Vajda, "DARS (Decoys As the Reference State) potentials for protein-protein docking.", *Biophysical journal*, 95(9), 2008, pp.4217-4227.
- [7] H. Lu, and J. Skolnick, "A distance-dependent atomic knowledge-based potential for improved protein structure selection.", *Proteins: Structure, Function, and Bioinformatics*, 44(3), 2001, pp.223-232.
- [8] B. G de la Torre and F. Albericio, "The pharmaceutical industry in 2018. An analysis of FDA drug approvals from the perspective of molecules.", *Molecules*, 24(4), 2019, p.809.
- [9] Brenke, Hall, Chuang, Comeau, Bohnuud, Beglov, Schueler-Furman, Vajda and Kozakov, "Application of asymmetric statistical potentials to antibody-protein docking.", *Bioinformatics*, 28(20), 2012, pp.2608-2614.
- [10] Goodsell, Zardecki, Di Costanzo, Duarte, Hudson, Persikova, Segura, Shao, Voigt, Westbrook and Young, "RCSB Protein Data Bank: Enabling biomedical research and drug discovery.", *Protein Science*, 29(1), 2020, pp.52-65.
- [11] Vreven, Moal, Vangone, Pierce, Kastritis, Torchala, Chaleil, Jiménez-García, Bates, Fernandez-Recio and Bonvin, "Updates to the integrated protein-protein interaction benchmarks: docking benchmark version 5 and affinity benchmark version 2.", *Journal of molecular biology*, 427(19), 2015, pp.3031-3041.
- [12] Raybould, Marks, Lewis, Shi, Bujotzek, Taddese and Deane, "TheraSAbDab: the Therapeutic Structural Antibody Database.", *Nucleic acids research*, 48(D1), 2020, pp.D383-D388.
- [13] T. Ermak, A. Shehovtsov and P. Yakovlev, "HEDGE: Highly accurate GPU-powered protein-protein docking pipeline.", *SI Kabanikhin, Professor, Corresponding Member of the RAS, Institute of Computational Mathematics and Mathematical Geophysics of SB RAS*, 2018, p.22.
- [14] Altschul, Madden, Schäffer, Zhang, Zhang, Miller and Lipman, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.", *Nucleic acids research*, 25(17), 1997, pp.3389-3402.