The 65th ASH Annual Meeting Abstracts

POSTER ABSTRACTS

203.LYMPHOCYTES AND ACQUIRED OR CONGENITAL IMMUNODEFICIENCY DISORDERS

DeepSurf2.0: A Deep Learning Approach for Predicting Interactions of B Cell Receptors with Antigens

Angelos-Michael Papadopoulos [1], *Anastasia Iatrou* [2], *Apostolos Axenopoulos* [1], *Andreas Agathangelidis* [3,2], *Kostas Stamatopoulos* [2], *Petros Daras* [1]

[1] Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece
[2] Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece
[3] School of Biology, National and Kapodistrian University of Athens, Athens, Greece

The B cell receptor immunoglobulin (BcR IG) is a unique molecular identity for each B cell clone, underpinning interactions with foreign and (auto)antigens that eventually affect clonal behavior. BcR signaling is crucial for the homeostasis of B cells, affecting all aspects of their physiology including cell activation, proliferation, differentiation and apoptosis. Moreover, it is highly relevant for pathological conditions implicating B cells, e.g. B cell lymphomas and autoimmune disorders. Structural analysis of the BcR IG and its cognate antigenic epitopes is vital in elucidating the mechanisms of BcR-antigen interactions. While analyzing actual protein crystals would be ideal, the crystallographic procedures are notoriously labor-intensive and challenging. Hence, we pivot to an *in-silico* approach, utilizing 3D analysis of BcR-antigen interactions. Confronted with the inherent variability of BcRs and the arduous nature of experimental analyses, we present a cutting-edge solution: *DeepSurf2.0*. This innovative computational tool leverages deep learning algorithms to predict Protein-Protein Interactions (PPI) and more specifically BcR-antigen interactions, creating a foundation for fast and accurate protein-protein docking. DeepSurf2.0, specifically tailored for the 3D structures of BcR IG and associated antigens, harnesses the power of deep learning to predict PPI: therefore, a carefully curated dataset is of paramount importance. To achieve the latter, we took advantage of SAbDab, a database containing all the antibody structures available in the Protein Data Bank (PDB), annotated and presented in a consistent fashion. We refined the SAbDab dataset by applying the following filtering steps: (i) we retained only complete BcR IG, i.e. those with available heavy and light chains, (ii) we preserved only one biological assembly from multimeric protein complexes, (iii) we excluded BcRs without associated antigens, and (iv) we constructed each BcR-antigen pair to consist of three chains (one each heavy and light for the BcR and one for the antigen). Through these exacting measures, we created a comprehensive collection of 10,543 BcR-antigen pairs. DeepSurf2.0 was evaluated using two metrics: DCA (Distance between Predicted binding site center and nearest antigen Atom) and OVR (Intersection of real and predicted binding sites divided by their union). A binding site prediction was considered as a hit if DCA < 4 Å. For training purposes, we utilized 9,440 BcR-antigen pairs to optimize DeepSurf2.0. The model was then evaluated on a separate test set of 1,103 BcR-antigen pairs. In this evaluation, DeepSurf2.0 achieved a DCA rate of 33%, which means that a hit was detected in 364 out of 1,103 cases. To measure the quality of these predictions, we assessed the OVR metric that resulted in a rate of 22%. To the best of our knowledge, there are no relevant methods that have been tested in a similar dataset. Existing state-of-the-art PPI prediction approaches achieve similar scores in DCA and OVR; however, the utilized datasets consisted of single chains in receptor and ligand. In contrast, our model incorporates a more complex two-chain receptor paradigm, which is a more challenging task but closer to the reality of BcR-antigen interactions. The aforementioned results not only facilitate understanding molecular interactions but also provide valuable insights into potential BcR docking areas for antigens. This ability to predict and locate the most probable interaction sites has immediate practical implications, significantly expediting the docking process by negating the need for time-consuming blind docking. Since our results are not directly comparable with those of the current state-of-the-art methods, our dataset will be provided publicly as a benchmark to evaluate similar methods in two-chain receptor cases. In conclusion, DeepSurf2.0 serves as a foundation for enabling subsequent docking algorithms to target the predicted interaction binding surface rather than the entire protein structure. This advancement underscores the transformative potential of deep learning within the realm of (immuno)hematology, holding the potential to provide novel insights into the pathogenesis and progression of B cell-related disorders.

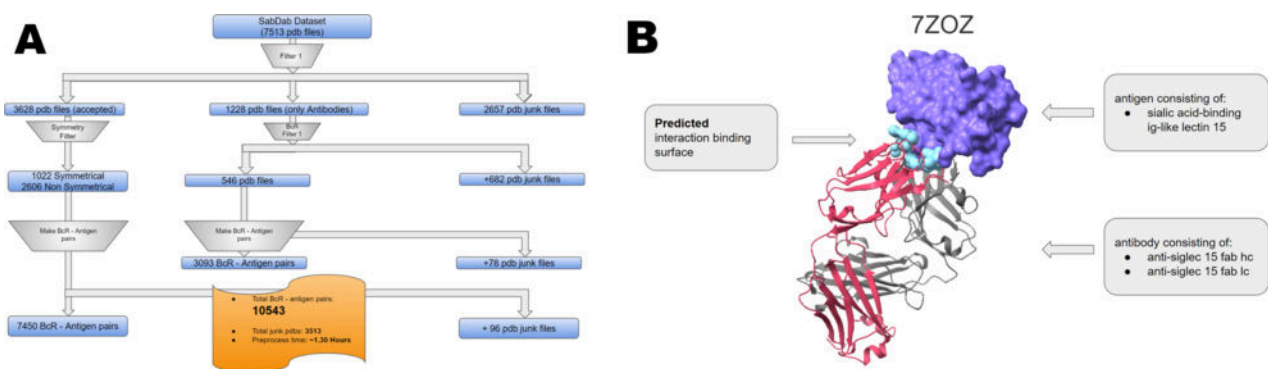**Disclosures** No relevant conflicts of interest to declare.

Figure 1

https://doi.org/10.1182/blood-2023-188537