Computational Tools for Allergenicity Prediction in SDAP

Werner Braun
Sealy Center for Structural Biology and Department of Human Biological Chemistry & Genetics
University of Texas Medical Branch, Galveston, Texas 77555-1157

Abstract

Similarities in the sequences and 3D structures of allergens may explain clinically observed cross-reactivities between different biological triggers. Towards the goal to establish quantitative descriptors for IgE epitopes based on sequence and 3D structure information, we have established a Structural Database of Allergenic Proteins (SDAP) as a web server (http://fermi.utmb.edu/SDAP/) for the research community [1-3]. Recent improvements of SDAP include the implementation of bioinformatics search rules for allergens as proposed by the FAO/WHO committee, and a classification of allergens based on PFAM domains.

The user of SDAP can automatically perform FASTA alignments for all 80 aa windows along the user sequence. The user can change the sequence identity cutoff from the value proposed by the FAO/WHO allergenicity rules (35%). The sub-sequences of 80 amino acids windows are compared with all allergens in SDAP, and the alignments that have a sequence identity higher than a threshold are presented. The user can also search in SDAP with a full FASTA alignment between the user sequence and all allergens from the SDAP database. The output contains a summary table containing all high scoring allergens, several statistical indices that measure the sequence homology of the FASTA alignments, and the individual pairwise alignments and full sequence identities.

The PFAM classification of allergens in SDAP is useful for a precise and convenient procedure to find family relations of newly discovered allergens. The results collected for the SDAP allergens include: SDAP allergen accession number, TrEMBL or SwissProt accession number, Pfam Uniprot name, Pfam family domain name, GenBank number, Pfam family type (A or B), Pfam accession number, the start and end of the sequence domain. Most of the allergens in SDAP have assigned PFAM domains. To allow a convenient overview of Pfam A domains found in allergens, the user has access to lists with all allergens that contain a certain Pfam A domain, or to the domain description from Pfam (http://www.sanger.ac.uk/Software/Pfam/).


Acknowledgements. This work was supported by grants from the John Sealy Memorial Endowment Fund (#2535-01), the FDA (FD-U-002249-1), and Texas Higher Education Coordinating Board ATP (004952-0036-2003).