WP11: *In silico* prediction of immunogenic and allergenic proteins

1. Main activities and results

Task 11.1*: *In silico prediction of allergenic proteins. The database of allergenic proteins, created in the project SuperCA++ (I phase), was upgraded by novel allergens. A mirror database of non-allergens of the same origin was created. The final version of the database contains 2427 allergens and 2427 non-allergens and is freely accessible at: http://www.ddg-pharmfac.net/AllergenFP/data.html. The structure of proteins was described by amino acid descriptors (E_1, E_2, E_3, E_4 and E_5), reflecting the principal physicochemical properties of amino acids building the proteins, such as lipophilicity, volume, abundance, helix-forming or β-strand forming propensities. Since proteins were of different lengths, they were subjected to ACC (auto-and cross-covariance) transformation and were converted into vectors of equal length. Further, vectors were transformed into binary descriptor fingerprints and were compared in terms of Tanimoto coefficient of similarity. The algorithm was implemented into a server for *in silico* prediction of allergens and is freely accessible via http://www.ddg-pharmfac.net/AllergenFP.

Participants in the task: Ivan Dimitrov and Irini Doytchinova. Results are included in a manuscript submitted for publication [DNDB_11s].

Task 11.2*: *In silico prediction of T-cell epitopes. A database consisted of peptides binding to HLA-DRP1 protein was created and subjected to QSAR analysis. A model for affinity prediction was developed, accounting for the amino acid contribution at each peptide position. The preferred residues at each position are shown at Figure 1.

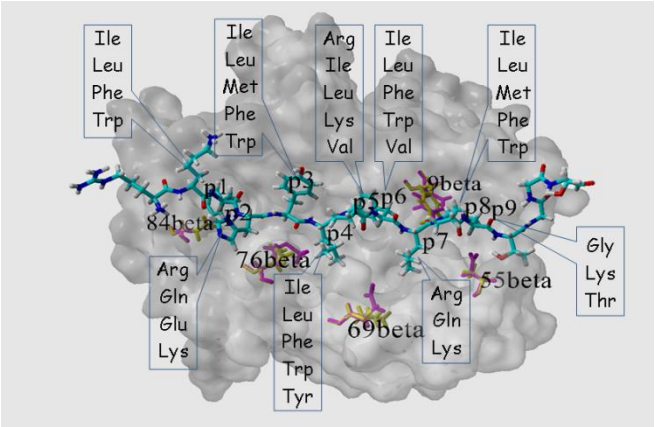

Figure 1. The complex peptide – HLA-DP1 protein modeled by homology from the complex peptide – HLA-DP2 (pdb code: 3LQZ). The protein is given in grey, the peptide – in blue. Preferred amino acids are indicated for each peptide position.

Participants in the task: Ivan Dimitrov, PhD student Stefan Ivanov and Irini Doytchinova. The results were accepted for publication [IDD_11a].

Task 11.3*: *Molecular docking study on peptides binding to MHC class II proteins. The docking-based method for binding affinity prediction to MHC proteins, developed in project SuperCA, was applied to 23 most frequent MHC class II proteins. The method includes homology modeling of proteins using an X-ray template, generation of combinatorial peptide libraries, docking of ligands and compilation of the normalized energies into quantitative matrices (QM) for binding affinity prediction to MHC class II proteins. QMs were implemented into a server freely accessible via http://epidock.ddg-pharmfac.net/.

Participants in the task: Mariyana Atanasova, Ivan Dimitrov, PhD student Atanas Patronov and Irini Doytchinova. Results are published [PDFD_11, PDFD'_11 и PSDFD_11a].

Task 11.4: ***Molecular docking study on acetylcholinesterase (AChE) inhibitors.*** A docking-based method was developed to assess the affinity of ligands binding to enzyme AChE (Figure 2). The method was applied to 2 series of inhibitors. The results were used to design novel more active ligands which will be synthesized and tested.

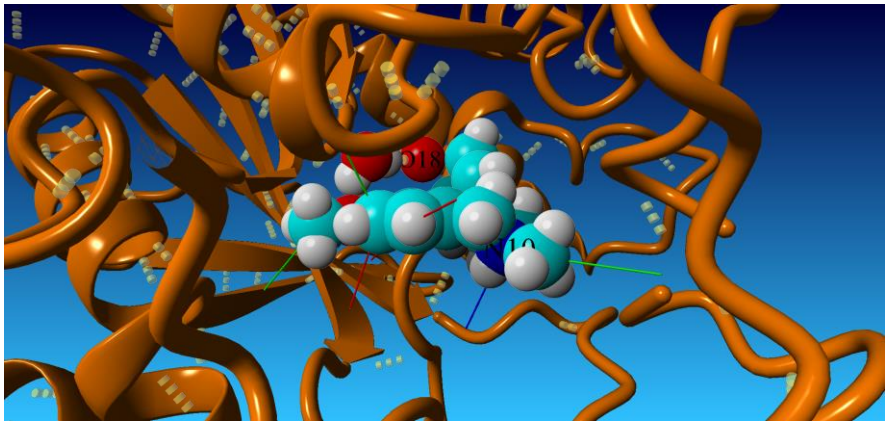

Figure 2. The complex galantamine – AChE (pdb code: 1QTI). The protein is given in orange, galantamine – in blue. The interactions between the molecules are shown: hydrogen bonds (in yellow), hydrophobic interaction (in green), cation – π interaction (in dark blue), π- π stacking (in red).

Participants in the task: Mariyana Atanasova, Ivan Dimitrov, PhD student Nikola Yordanov and Irini Doytchinova. Results are prepared for publication [AYDBTD_11p].

2. Publications

a) published:

[DFD_11] Dimitrov, I.; Flower, D.R.; Doytchinova, I. Improving MHC class II Peptide binding prediction by union and intersection of single predictors. World J. of Vaccines, 1 (2011) 15-22.

[PDFD_11] Patronov, A.; Dimitrov, I.; Flower, D. R.; Doytchinova, I. Peptide binding prediction for the human class II MHC allele HLA-DP2: a molecular docking approach. BMC Struct. Biol., 11 (2011) 32. IF(2011) = 2,476

[DPDAF_11] Doytchinova, I.; Petkov, P.; Dimitrov, I.; Atanasova, M.; Flower, D. R. HLA-DP2 binding prediction by molecular dynamics simulations. Protein Sci.., 20 (2011) 1918-1928. IF(2011) = 2,798

[PDFD'_11] Patronov, A.; Dimitrov, I.; Flower, D. R.; Doytchinova, I. Peptide binding to HLA-DP2 proteins at pH 5.0 and pH 7.0: a quantitative molecular docking study. BMC Struct. Biol., 12, (2012) 20. IF(2012) = 2,099

[PD_11] Patronov, A.; I.A. Doytchinova. T-cell epitope vaccine design by immunoinformatics. Open Biology, 3 (2013) 120139.

[DFD_11] Dimitrov I.; Flower D. R.; Doytchinova I. AllerTOP – a server for *in silico* prediction of allergens. BMC Bioinformatics, 14(Suppl. 6) (2013) S4. IF(2012) = 3,024

[BNMDD_11] Bangov, I., Naneva, L., Moskovkina, M., Dimitrov, I., Doytchinova, I. Application of Descriptor Fingerprints in QSAR Studies. Annual of Konstantin Preslavsky University of Shumen, 22B1 (2013) 28-36.

[FDD_11] Flower, D.R.; Davies, M.N.; Doytchinova, I.A. Identification of Candidate Vaccine Antigens In Silico. In: Immunomic Discovery of Adjuvants and Candidate Subunit Vaccines (Eds. D. R. Flower, Yvonne Perrie), Springer New York Heidelberg Dordrecht London, 39-73, 2013.

b) accepted:

[PSDFD_11a] Patronov, A.; Salamanova, E.; Dimitrov, I.; Flower, D. R.; Doytchinova, I. Histidine hydrogen bonding in MHC at pH 5 and pH 7 modeled by molecular docking and molecular dynamics simulations. Curr. Comp.-Aid. Drug Des., in press, 2013. IF(2012) = 1,540

[APDFD_11a] Atanasova, M.; Patronov, A.; Dimitrov, I.; Flower, D. R.; Doytchinova, I. EpiDOCK – a molecular docking-based tool for MHC class II binding prediction. Protein Eng. Des. Sel., in press, 2013. IF(2012) = 2,588

[NDBD_11a] Naneva, L., Dimitrov, I., Bangov, I., Doytchinova, I. Allergenicity prediction by partial least squares-based discriminant analysis. Bulg. Chem. Commun., in press, 2013. IF(2012) = 0,320

[IDD_11a] Ivanov, S.; Dimitrov, I.; Doytchinova, I. Quantitative prediction of peptide binding to HLA-DP1 protein. IEEE-ACM Trans. Comput. Biol. Bioinform., in press, 2013. IF(2012) = 1,616

c) submitted:

[DNDB_11s] Dimitrov, I.; Naneva, L.; Doytchinova, I.; Bangov, I. AllergenFP: Allergenicity prediction by descriptor fingerprints.

d) in preparation

[AYDBTD_11p] Atanasova, M.; Yordanov, N.; Dimitrov, I.; Berkov, S.; Tsachev, H.; Doytchinova, I. Molecular docking study on galantamine derivatives as cholinesterase inhibitors.

3. Presentations

[1] Dimitrov, I. Allertop – a Bioinformatics Tool for Allergenicity Prediction (oral presentation). VIII EWDD Eighth European Workshop in Drug Design, Certosa di Pontignano, Siena (Italy), 22 – 28 May 2011.

[2] Dimitrov, I. Application of machine learning techniques for allergenicity prediction (oral presentation). 2nd Regional Conference "Supercomputing Applications in Science and Industry" Rodopi Hotel, Sunny Beach, Bulgaria, 20-21 September, 2011.

[3] Atanasova, M. Prediction of peptide binding to swine leukocite antigen (SLA-1) proteins by molecular docking (oral presentation). VIII EWDD Eighth European Workshop in Drug Design, Certosa di Pontignano, Siena (Italy), 22 – 28 May 2011.

[4] Atanasova, M. Structural Immunoinformatics – two case studies (oral presentation). 2nd Regional Conference "Supercomputing Applications in Science and Industry" Rodopi Hotel, Sunny Beach, Bulgaria, 20-21 September, 2011.

[5] Atanasova, M. Modelling of anticholinesterase activity by molecular docking (poster presentation). Innovative Approaches to Computational Drug Discovery, Lausanne, Switzerland, 1 – 4 October, 2013.

[6] Doytchinova, I. In silico prediction of immunogenic and allergenic proteins (oral presentation). Supercomputer Applications SuperCA++ 2012, Hotel St. Ivan Rilski, Bansko, 22 – 24 April 2012.

[7] Doytchinova, I. Drug design – the art of drug creation (oral presentation). 9th EPSA Autumn Assembly, Hotel Princess Dedeman, Sofia, 29 October – 3 November 2012.

[8] Doytchinova, I. In silico prediction of immunogenic and allergenic proteins (oral presentation). Supercomputer Applications SuperCA++ 2012, Hotel Kalina Pallace, Tryavna, 31 March – 2 April 2013.

4. Other activities

1. During the last three years Prof. Doytchinova's group participates in the course "Supercomputer application in natural sciences" with lectures and practicals in module: Drug design and the use of supercomputers in the drug design. The course is visited by young scientists, PhD students, postgraduates and undergraduates working or studying in the field of natural sciences.

2. A server EpiDOCK for *in silico* prediction of peptides binding to MHC class II proteins was constructed. It is freely available at: http://epidock.ddg-pharmfac.net/.

3. A server AllergenFP for *in silico* prediction of allergens was constructed. It is freely available at: http://www.ddg-pharmfac.net/AllergenFP.

4. Administration of the Drug Design Group web page at http://www.ddg-pharmfac.net.

5. Administration of all servers available at ddg web page: AllergenFP. AllerTOP, EpiDOCK, EpiTOP, EpiJen, MHCPred and VaxiJen.

6. In 2012/2013 two PhD students (Atanas Patronov and Panaiot Garnev) graduated with theses on tasks from the project SuperCA++. A third PhD student continues his study in the University of Manchester, UK. A forth PhD student, Nikola Yordanov, currently prepares his thesis on a task from the project.

7. The group collaborates with Prof.Bangov's research group from the University of Shumen. Results are included in 3 manuscripts: 1 published, 1 accepted for publication and 1 submitted for publication.

8. Mariyana Atanasova participated in a workshop titled "Innovative Approaches to Computational Drug Discovery", in Lausanne, Switzerland, from 1 to 4 October, 2013. There she presented her results on *task 11.4. Molecular docking study on acetylcholinesterase (AChE) inhibitors.*