WP8: *In silico* prediction of immunogenic and allergenic proteins

1. Main activities and results

Task 8.1*: *In silico prediction of allergenic proteins. A database of allergenic and non-allergenic proteins of food origin was created. The structure of proteins was described by amino acid descriptors (z_1, z_2 and z_3), reflecting the principal physicochemical properties of amino acids building the proteins, such as lipophilicity, volume and electronic properties. Since proteins were of different lengths, they were subjected to ACC (auto-and cross-covariance) transformation to convert them into vectors of equal length. A discriminant analysis based on several methods was performed in order to develop models for comparison between allergens and non-allergens. The methods used were partial least squares (PLS), *k* nearest neighbours (*k*NN), support vector machines (SVM), and logistic regression (LR). The models were validated by external test group and compared to 5 other servers for antigenicity prediction available free online. The *k*NN model showed the highest accuracy of the prediction (83.6%) and had the best performance comparing to others. The model will be implemented into a server for *in silico* prediction of allergens.

Participants in the task: Ivan Dimitrov and Irini Doytchinova.

Task 8.2*: *In silico prediction of T-cell epitopes. The first phase of this task aims to develop a proteochemometric model for affinity prediction of peptides binding to human MHC class II proteins, locus HLA-DR, serotype DRB1. A training set of peptides binding to HLA-DRB1 proteins was compiled and used to obtain several models which were validated by external test set. The model with the best predictive ability was included in a specially designed website (http://www.pharmfac.net/EpiTOP/index.php). All models that will be developed for MHC class II binding prediction will be included in this site. Our model was compared with other similar methods and it showed superior performance. A model for prediction of peptide binding affinity to HLA-DRB3 protein is in development.

Participants in the task: Ivan Dimitrov, Panajot Garnev and Irini Doytchinova

Task 8.3(new)*: *Molecular docking study on peptides binding to HLA-DRB1 proteins. A docking based method for binding affinity prediction to HLA-DRB1 proteins is developing. As input data in the software GOLD is used the X-ray structure of the peptide-HLA-DRB1*0101 complex (pdb code: 2g9h). The peptide position Tyr308 which binds into protein pocket 1 is mutated to the rest 19 naturally occurring amino acids. The binding affinities of the mutated proteins are assessed by a scoring function and compared to experimental data. The method is calibrating in regards to binding radius, peptide and protein flexibility, scoring functions, ranking method and other settings. The settings that give results closest to the experimental data will be used further in the binding assessment of pockets 4, 6 and 9.

Participants in the task: Mariyana Atanasova and Irini Doytchinova

2. Publications

a) published:

[DGFD_09] Dimitrov, I.; Garnev, P.; Flower, D.R.; Doytchinova, I. Peptide binding to the HLA-DRB1 supertype: A proteochemometrics analysis. Eur. J. Med. Chem., 45 (2010) 236-243.

b) accepted:

c) submitted:

[DGFD_09s] Dimitrov, I.; Garnev, P.; Flower, D.R.; Doytchinova, I. MHC class II binding prediction – a little help from a friend. J. Biomed. Biotech.

[GDFD_09s] Garnev, P.; Dimitrov, I.; Flower, D.R.; Doytchinova, I. EpiTOP – a proteochemometric tool for MHC class II binding prediction. Bioinformatics

d) in preparation

[DGFD_09p] Dimitrov, I.; Garnev, P.; Flower, D.R.; Doytchinova, I. Fusion and voting methods for MHC class II binding prediction

3. Presentations

Dimitrov, I., P. Garnev, D. R. Flower, **I. Doytchinova**: T-cell epitope prediction: A proteochemometric approach (oral presentation). *1st National Conference with International Participation on Biomedical and Bioprocess Engineering*, Sofia, 3-4 December 2009.

Dimitrov, I., P. Garnev, D. R. Flower, I. Doytchinova: Comparative study on servers for T-cell epitope prediction (poster presentation). *1st National Conference with International Participation on Biomedical and Bioprocess Engineering*, Sofia, 3-4 December 2009.

Garnev, P., I. Dimitrov, D. R. Flower, I. Doytchinova: EpiTOP – a proteochemometric tool for T-cell epitope prediction (poster presentation). *1st National Conference with International Participation on Biomedical and Bioprocess Engineering*, Sofia, 3-4 December 2009.

4. Other activities

1. A new computer lab was equipped in the Faculty of Pharmacy, Medical University of Sofia. The lab is used by students in the practicals in Physical Chemistry, Pharmacokinetics and Drug Design.

2. The Drug Design course has been renewed and it is based now on the Molecular Conceptor software (Synergix Ltd.).

3. In July 2009 Ivan Dimitrov and Panayot Garnev attended a workshop on Structure-based drug design in Oxford University.

4. A new software product for molecular docking GOLD (CCDC Ltd.) was acquired. The product is used in the current project.