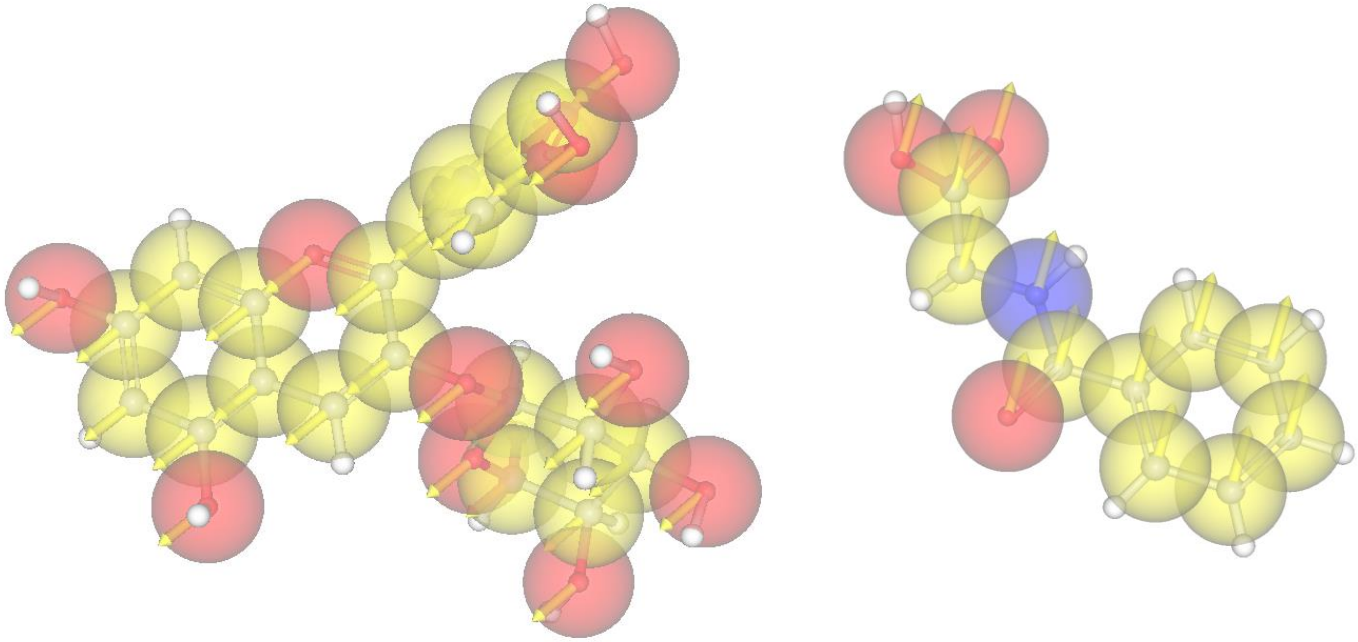
PROGRESS REPORT Grant COST-STSM-FA1403-35655

Virtual screening (docking) of plant bioactives against biomolecular targets involved in their nutrigenomics effects

The aim of the STMS was the virtual screening of a selected list of plant bioactive compounds in order to provide information on molecular aspects to help understanding of their nutrigenomics effects. As virtual screening strategy was used the molecular docking, the process of determining the orientation and energy of binding of a small compound in the active site of a target, usually a protein. To perform a docking there is required to know the 3D structures of the bioactive compounds (the ligands) and of their macromolecular targets.

As ligands were used five anthocyanins (cyanidin-3-O-arabinoside, cyanidin-3-O-galactoside, cyanidin-3-O-glucoside, delphinidin-3-O-glucoside, peonidin-3-O-glucoside) and their metabolites (4-hydroxybenzaldehyde, ferulic acid, hippuric acid, protocatechuic acid, vanillic acid). The selected ligands are from the initial output of data analyzes of nutrigenomic effect of plant food bioactives initiated in the WG2 *COST Action POSITIVe.*

Prior to screening, a pharmacophore search was performed in order to identify the pharmacophoric points and to predict the most probable ligand-protein type of interactions. It was revealed that the phenolic ring system of anthocyanins is planar and have electrons acceptors, meanwhile the planar benzene ring is hydrophobic, but the phenolic hydroxyl group (also present in metabolites) confers polarity and water-solubility and the capacity for establishing hydrogen bonds with proteins (Fig. 1.a). Moreover, the hippuric acid have an electron donor (–NH–) and it may be able to establish supplementary H-bonds with proteins, as donor (Fig. 1.b). Only some AA residues of targets are able to make H-bonds with the ligands: glutamine, asparagine, histidine, serine, threonine, tyrosine, cysteine, methionine, tryptophan. Meanwhile, the extended hydrophobic regions of all ligands make them able to interact preferably with the targets' regions rich in hydrophobic AA residues (alanine, valine, leucine, isoleucine, proline, phenylalanine, tryptophane, cysteine and methionine – glycine do not have a side chain and it may not be able to interact with ligands).



a) cyanidin-3-O-arabinoside b) hippuric acid

Fig. 1. The pharmacophoric points: the bioactives are depicted as stick and balls, yellow spheres indicate the hydrophobic points, the red spheres indicate the electron acceptor points, the blue sphere indicates the electron donor point, meanwhile the yellow arrows indicate the direction of the pharmacophore effects

Based on accumulated know-how of *COST Action POSITIVe*, were selected as targets: signaling proteins involved in the NF-κB signaling pathway and MAPK pathway as well Focal Adhesion Kinase to start with. For the target identifications process was necessary to cross-reference 3 on-line databases (GeneCards[1], UniProt[2] and PDB[3]) in order to identify the possible targets and their high-resolution complete 3D structure (a resolution better than 2.0 Å being recommended for docking). The selection process revealed 63 possible targets (including here a mutant protein – RHOA, with F25N mutation). From those, only 4 targets were found available in PBD as full length high-resolution 3D structures: 4QTA for MAPK1, 4QTB for MAPK3, 3FMK for MAPK14 and the mutant 5C4M for RHOA. For all the rest of targets, including the normal RHOA, were build homologue models with the help of SWISS-MODEL/ExPASy web server[4]. All docking runs were carried out with PyRx – Python Prescription 0.9.4 using AutoDock Vina as docking software and the search space was set to entirely cover the targets, meanwhile the exhaustiveness was manually increased 10 times from default in order to improve the accuracy of predictions. This screening set-up was aimed to reveal all the possible ligand-protein interactions, disregarding their specific way of interaction, in order to eliminate from equation, the interactions without any biological relevance. This preliminary screening revealed that all ligands are weak binders of CREB1, c-Jun, IκBα and MEKK1 and perhaps their interactions doesn't have biological relevance. Moreover, the metabolites are considerably weaker binder that the anthocyanins, with a binding affinity (BA), mostly between -3.3 and -6.9 kcal/mol, suggesting either a small or even an irrelevant biological signification. The anthocyanins have proved to establish strong interaction with some of the targets (BA between -8.5 and -9.9 kcal/mol): Akt3, ASK1, ASK2, FAK1 (Fig. 2), IKKα, IKKβ, JAK2, JAK3, JNK1, JNK2, JNK3, MEK1, MEK2, MKK4, MSK1, MSK2, mTOR, p38ɣ, ROCK2, RSK2, SMAD2, SMAD3, SMAD4, Smurf2.

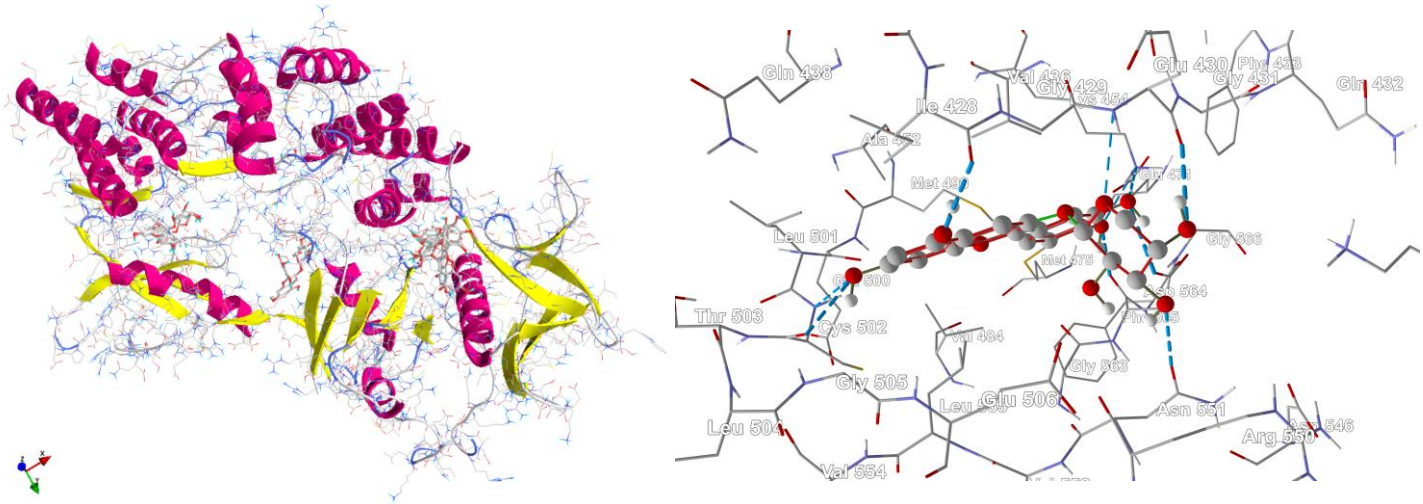


Fig. 2. Docking results for FAK1: left side – general view of FAK1 with binding sites of all ligands; right side – detailed view of binding site of cyanidin-3-O-arabinoside with surrounding AA residues (H-bonds are depicted as dashed blue lines)

This type of strong interactions between anthocyanins and the short list of targets should be investigated to identify more accurate the nature of interactions and the interacting domains/regions of target, together with the possible biological effects.

This STSM allowed to test and identify the protocol for *in silico* docking structure that will be performed in the WG2 in the coming 2 years using nutrigenomic data that will be extracted for different families of plant food bioactives.

Dr. Radu TAMAIAN 08.03.2017

[1] GeneCards® – The Human Gene Database: http://www.genecards.org
[2] The Universal Protein Resource – UniProt: http://www.uniprot.org
[3] RCSB Protein Data Bank – RCSB PDB: http://www.rcsb.org/pdb/home/home.do
[4] SWISS-MODEL is a fully automated protein structure homology-modelling server, accessible via the ExPASy web server: https://swissmodel.expasy.org/