Describing drug metabolism and regulatory interactions in the chemoimmune system using systems biology approaches

Background:

Xenobiotics entering our body are neutralized by the cellular chemoimmune (ChI) system formed by metabolic enzymes, membrane transporters, and transcription factors. These proteins recognize chemically different compounds, thus influencing drug absorption, distribution, metabolism, excretion, and toxicity. Our objectives included advancing our knowledge on the function of this system. We aimed to characterize the structure and function of individual members and to describe quantitatively the orchestrated action of the ChI system.

Main results:

We performed molecular dynamics simulations and *in silico* docking to learn functioning of ChI proteins and their interaction with drugs.

First, we studied the small soluble drug/ligand binding domain of **Aryl Hydrocarbone Receptor**, recognizing ligands which are also substrates for the ABCG2 multidrug transporter. We built structural models of the ligand binding PAS-B domain of AhR and performed molecular dynamics simulations using different force fields and methods to characterize the dynamics of the AhR PAS-B domain. The generated conformational ensembles were subjected to *in silico* docking for understanding the structural and ligand binding features of this small domain. Our exhaustive *in silico* binding studies suggest that ligand specificity and selection may be determined not only by the PAS-B domain itself, but also by other parts of AhR and its protein interacting partners. Szollosi *et al.* PLOS ONE 2016, http://ahr.hegelab.org

We have suggested that ligand binding pocket and access channels leading to the pocket play equally important roles in discrimination of xenobiotics from endogenous molecules. Therefore, we continued our studies with the **ligand binding domain of PXR** (Pregnane X Receptor) with well-defined access channels. We generated structural ensembles using various molecular dynamics simulation approaches and employed a diverse set of *in silico* methods to dock molecules to conformational ensembles of PXR. However, our results suggested that none of the methods were capable to accurately describe the binding of hydrophobic drugs and distinguish binders from non-binders.

Our primary target protein became the **ABCG2 (BCRP, Breast Cancer Resistance Protein)** drug transporter, which is a transmembrane ATPase. ABCG2 has unique structural features among human ABC proteins, since functions as a homodimer and the NBD (nucleotide binding domain) is located N-terminally from TMD (transmembrane domain). It also possesses a so called linker region connecting the NBD and TMD. Collaborators' experimental and our *in silico* results suggested that the linker region contains a sequence motif homologous to the canonical ABC signature present in all ABC nucleotide-binding domains and has an important mechanistic role in ATP binding and/or hydrolysis coupled to drug efflux. Macalou *et al.* Cell Mol Life Sci, 2015

Analysis of ABCG2 expression in erythrocyte membranes of gout patients showed decreased expression levels compared to normal red blood cells. Our collaborators identified a relatively frequent, **novel ABCG2 variant** (M71V), which alters functional expression. We proposed a mechanism for the altered dynamics of the nucleotide binding domain containing this M71V mutation based on molecular dynamics simulations. Our combined experimental and computational study

suggested that information regarding this newly discovered ABCG2 variant may be used in the future in the framework of personalized medicine. Zámbó *et al.* Sci Rep, 2018

After the structure of an ABCG2 homolog, the human sterol transporter ABCG5/ABCG8 was published, we were able to generate an *in silico* **ABCG2 structural model**. In order to describe the effect of mutations on structure and dynamics, and characterize substrate recognition, we performed molecular dynamics simulations and *in silico* docking using the full length ABCG2 model embedded in a lipid bilayer. Our results showed that in the presence of the Q141K variant with decreased functional expression, the introduced positive charge diminishes the interaction between the nucleotide binding and transmembrane domain. Interestingly, the R482 position, which plays a role in the substrate specificity of the transporter, is located in one of the substrate binding pockets identified by our *in silico* docking calculations. This pocket exhibits a preference for substrate binding versus non-substrate type molecules and may act as a filter along the translocation pathway. Laszlo *et al.* PLOS ONE, 2016, http://abcg.hegelab.org

At the time of the report preparation, we are investigating the **cholesterol regulation of ABCG2 and the transport process** of a substrate in µs-long MD simulations. Cholesterol does not influence the equilibrium dynamics of this transporter, thus it likely influences a transition step during the transport process. We could observe the binding of a substrate to binding sites we have predicted by *in silico* docking and movements of the cholesterol molecule between the sites. The substrate could not move to the extracellular space in our simulations, which do not allow the hydrolysis of ATP, thus we are performing metadynamics simulations to describe the final step of the export process. Manuscript is under preparation, talks: KeMoMo-QSAR, Szeged, 2019 and Farmakokinetika és Gyógyszermetabolizmus Szimpózium, Galyatető, 2019

In order to approach our goal of **structural systems biology** description of the ChI system, we also performed studies on the membrane transporter **ABCB1/MDR1**, which share substrates with and is transcriptionally regulated by **PXR** (pregnane X receptor) (structural ensembles were generated, all atom and coarse grained MD simulations were performed, and *in silico* docking calculations with substrates and non-substrates were initiated). However, as a consequence of the novel **cryo-EM structures** our simulations have to be repeated. Since the number of structural models of membrane ABC proteins increased in the last years because of the cryo-EM revolution, interpreting the structures became more and more challenging. While we tackled the selection of the right structures for MD simulations, we found that standardized, **conformational vectors (conftors)** defined between specific points of ABC proteins (e.g. the extracellular end of a helix and the center of mass of an intracellular coupling helix) can help in quantitative comparison of structures. Csizmadia Comput Struct Biotechnol J. 2018, http://conftors.hegelab.org

The conftors' project required the visual inspection of cryo-EM densities of ABC protein structures. During that process, we realized that these maps contain experimental information on the position of the membrane bilayer surrounding the protein. Since there are sparse experimental data on the topology of membrane proteins and the most used methods are *in silico* algorithms, we developed a computational **pipeline to extract the membrane boundaries from cryo-EM maps**. The pipeline is the first method to determine membrane boundaries based on experiments at a larger scale as compared to inserting epitopes at various sites around a putative transmembrane helix. (1) Farkas *et al.* Bioinformatics, accepted for publication, http://memblob.hegelab.org, (2) Csizmadia, oral presentation at Mechanisms of Membrane Transport Gordon Research Conference, 2019

Although **CFTR (ABCC7)** is not an active transporter and is not a member of the ChI system, it was an important subject of our project. Cystic fibrosis (CF), a lethal monogenic disease, is caused by mutant variants of the CFTR chloride channel. The majority of CF mutations, including the most frequent ΔF508 deletion, affect protein folding and stability and lead to diminished apical anion conductance of epithelial cells. Although extensive efforts have been devoted to CF drug

development, efficacious small molecules capable to correct CFTR folding are still searched that could be partly attributed to our limited knowledge on CFTR structure. Various conformations (+/- ATP and phosphorylation) of this chloride channel have been determined. These structures were available earlier than that of multidrug transporters because of the high medical relevance of CFTR, therefore we also utilized them as setting up and testing simulations with near-atomic resolution, cryo-EM structures.

We studied the structure and dynamics related to CFTR function and the effect of mutations on protein structure and stability. Firstly, we assessed the structural properties of CFTR protein based on experimental and homology models by molecular dynamics simulations. We described in detail the entry sites, identified the intra-protein selectivity filter at the atomic level, and characterized the interaction of the involved amino acids with chloride ions. The principal investigator of this project was awarded a Pilot and Feasibility grant from the Cystic Fibrosis Foundation (108,000 USD/2 years, 2018-2020) to study the folding of wild type and ΔF508 NBD1, using both *in silico* and single molecule force spectroscopy methods. (1) Veit *et al.* Sci Transl Med, 2014 (2) Tordai *et al.* BBRC, 2017 (3) Farkas *et al.* Cell Mol Life Sci, 2016

We applied various systems biology approach to describe the orchestrated action of proteins in the ChI network.

We designed a **reaction kinetic model** based on differential equations describing Phase 0-III participants and regulatory elements and characterized cellular fitness to evaluate toxicity. Based on biological facts about elementary molecular interactions revealed by conventional biochemical experiments and laws of biochemical reaction kinetics, rate equation of every molecular component of the system was formulated. By numerically solving ordinary differential equations of the parametrized model, we performed time course analysis to simulate the temporal behavior of the ChI network. Our results indicated that, the model recapitulates changes associated with acquired drug resistance and allowed toxicity predictions under variable protein expression and xenobiotic exposure conditions in spite of the simplifications applied. Our simulations highlighted the role of multidrug ABC transporters in facilitating the defense function of successive network members (intracellular enzymes) by lowering intracellular drug concentrations. We also extended the model with a novel toxicity framework which opened the possibility of performing *in silico* cytotoxicity assays. Toth *et al.* PLOS ONE, 2015

We hypothesized that the level of cooperation between ChI proteins could be characterized by interactions at the transcriptional level. Therefore, our aims included the investigation of **gene expression pattern triggered by drug treatment**. We selected samples exposed to various drugs from data deposited in GEO (Gene Expression Omnibus, NIH). The observed patterns turned out to be tissue or cell line specific, indicating that in this heterogeneous dataset it is impossible to distinguish whether a specific expression profile is characteristic for the drug treatment or for the applied cell line/tissue. We started to analyze Connectivity Map (Broad Institute) in a very similar way, since this database contains more homogenous data: the effects of over 1,000 drugs on mRNA levels in three cell lines. Unfortunately, the majority of changes in ChI genes expression did not exceed 5-10% compared to control. Our results obviated shortcomings in the experimental data sets or our methods. To test our methods on another dataset, we collaborated with Laszlo Cervenak on gene expression changes connected to MASP2 signaling. The results were published and justified our approaches. This indicated that experimental data sets were insufficient to study transcriptional level interactions in the ChI system. Schwaner *et al.* Sci Rep, 2017

In the framework of this systems biology part of the project, we aimed to develop public databases enhancing data quality and research of the field. We updated our **ABC mutation database**

with additional useful information regarding the effects of variants affecting non-coding regions. We also launched a web application containing data on an unexpectedly large number of ABC transporters and other membrane proteins present in red blood cell (RBC), previously thought to be specific for other tissues and/or related to major human diseases. Since the expression levels of these proteins provide a method to indicate pathological alterations, our RBCC (Red Blood Cell Collection) database may facilitate the development of membrane biomarker platforms and may provide a unique resource to aid related further research and diagnostics.
(1) Hegedűs *et al.* Database, 2015, http://rbcc.hegelab.org, (2) Tordai *et al.* Database, 2017, http://abcm2.hegelab.org

Changes in the budget and employment. Their effect on the project:

In order to create an alternative kinetic model, we initiated a collaboration with Kristóf Szalay and Péter Csermely. One of my young colleagues, Kinga András was expected to perform this work and started her PhD in systems biology at CEU to gain more competence. Unfortunately, she has decided to leave CEU because of the political situation.

Two of my colleagues, Attila Tóth and Kristóf Jakab, who became experts in systems biology tools, left for programming to a telecommunication company (one of them has four children).

The most talented student, Dániel Szöllősi, an owner of the Pro Scientia Gold Medal, moved back to Sopron because of family reasons and is a postdoc at Medical University of Vienna.

These unexpected changes in human resources slowed down the project and prompted us to focus more on the structural biology parts. These developments collectively caused budget changes and also the extension of the timeline of the project with half a year. On the other hand, this situation allowed to start connected experimental work in the last year, after securing complementary funds.

Publications:

https://m2.mtmt.hu/api/publication?cond=authors;in;10015687

peer-reviewed papers: 42; independent citations: 1 999; H-index: 20; ΣIF: 195.1

Most important, invited oral presentations:

Georgina Csizmadia: MemBlob database and server for identifying transmembrane regions using cryo-EM maps, Mechanisms of Membrane Transport *Gordon Research Conference*: Folding and Evolution in the Membrane Environment: New Ways of Understanding Transport Mechanisms, New London, NH, USA, 2019

Tamás Hegedűs: Cisztás fibrózis: gyógyszertervezés membránfehérje-szerkezet alapján? (Cystic Fibrosis: drug design based on membrane protein structure?), Kisfaludy Lajos Foundation's Conference, *Gedeon Richter Co*, 2018

Tamás Hegedűs: Effect of cystic fibrosis mutations on the CFTR protein structure and dynamics, *Chemistry towards Biology (CTB9)*, Budapest, 2018

Tamás Hegedűs: In silico modeling of ABCG2/BCRP sheds light on xenobiotics recognition and transport, *XIX Linz Winter Workshop on single molecule force spectroscopy*, Linz, Austria, 2017

Tamás Hegedűs: Ligand access may be as important as binding for xenobiotic selection of AhR, *IWBBIO 2016: International work-conference on bioinformatics and biomedical engineering*, Granada, Spain, 2016