In silico approaches to predict the impact of leukemic DNMT3a mutations & Identification of leads based on drug decitabine using Complex Based Pharmacophore Mapping and Virtual Screening

Syed Babar Jamal[1*], Matheus Filgueira Bezerra[2*], Sandeep Tiwari[1], Flavia Aburjaile[1,3], Vasco Azevedo[1], Artur Silva[3], Cintia Renata Costa Rocha[2,5], Marcos André Calvancanti Bezerra[4], Antonio Roberto Lucena-Araujo[4], Eduardo Isidoro Carneiro Beltrão[2,5]

*co-authorship of this work

[1]*Laboratory of Cellular and Molecular Genetics, UFMG;* [2]*Keizo Asami Immunopathology Laboratory, UFPE;* [3]*Center of Genomics and System Biology, UFPA;* [4]*Biological Science Center, Genetics Department, UFPE;* [5]*Biological Science Center, Biochemistry Department, UFPE*

DNA methyltransferases are a group of enzymes that catalyzes the addition of a methyl group in cytosines from DNA strand and are considered important regulators of differential gene expression. Recently, the next-generation technologies revealed a high frequency (21-25%) of mutations in the gene coding for DNA methyltransferase 3a (*DNMT3a*) in patients with acute myeloid leukemia (AML). These mutations disrupt the pattern of DNA methylation and conduct hematopoietic cells into malignant transformation. Despite massive efforts to better understand the functional aspects of DNMT3a, experimental studies are limited to only few frequent mutations. The aberrant pattern of DNA methylation in AML led to the use of DNMT3a inhibitors as an alternative treatment. Decitabine is a cytosine analog that binds to the catalytic domain of *DNMT3a* with an inhibitory effect and is used in AML therapy. Although decitabine is a known *DNMT3a* inhibitor, there is a lack of information concerning the interaction of decitabine with mutants *DNMT3a*. Considering the high diversity of mutations described in *DNMT3a*, we aimed to use *in silico* approaches to predict the impact of these mutations on protein function patients outcome and interaction with decitabine. For this, we selected 24 more frequent *DNMT3a* missense mutations described in databases: (COSMIC and TCGA). To predict the impact on protein function, we used the combination of six distinct tools. Based on the best scoring system we found that out of 24 evaluated mutations, 11 were classified as damaging (R882H, R882P, R882S, R803S, D781G, R792H, R736C, R729Q, S714C, G543C, C497Y), 10 as intermediaries (F909C, R882C, R882L, M880V, K841Q, K829R, R736H, R729W, P718L, G646V) and 3 as probably benign (A741V, N501T, K468R). To predict whether the score of the mutations may influence clinical outcome, we used the The Cancer Genome Atlas (TCGA) AML cohort. Of the 190 patients were included, 45 had mutations in *DNMT3a* (24%). We found that AML patients with mutated *DNMT3a* had a minor overall survival (p=0,005). However, when comparing different mutations grouped by the score, there was no significant difference in survival (p=0,214) among them. Although we could not demonstrate any difference in clinical outcome, our analysis suggests a significant biological heterogeneity in *DNMT3a* variants. Additionally, we hypothesize that the change in the enzyme structure caused by mutations may affect the drug-enzyme interaction and consequently, predict the clinical response to decitabine. Based on aforementioned hypothesis, this work reports the complex-based pharmacophore modeling to find the important pharmacophoric features essential for the inhibition of DNMT3a activity by virtual screening, drug-likeness predictions, protein-ligands binding interactions, binding affinity predictions and binding energy calculations.