Structural and Dynamics Analysis of Pyruvate Kinase from Erythrocytes: Implications in Pathology

Luis Jordà[#1], Josep Lluís Gelpí[#*2]

[#]*Barcelona Supercomputing Center (BSC)*

[*]*University of Barcelona (UB), Department of Biochemistry and Molecular Biology, 08028-Barcelona, Spain*

[1]luis.jorda@bsc.es, [2]gelpi@ub.edu

Keywords — **Pathogenicity prediction, pathogenic variant, protein dynamics, molecular dynamics, pyruvate kinase.**

EXTENDED ABSTRACT

Our current study revolves around the dynamic characterization of the human erythrocyte pyruvate kinase (R-PYK). The deficiency of this protein is a common cause of nonspherocytic hemolytic anemia, a rare, autosomal recessive disease. We plan to perform a comprehensive set of molecular dynamics simulations of both the wildtype (WT) and mutant variants of R-PYK in different conditions, in order to explore the dynamic behavior of the enzyme, describe the function and allosteric mechanism in terms of its dynamics fingerprint and identify altered dynamic behavior of the known pathogenic variants of the enzyme.

A. Introduction

Amino acid substitutions are implicated in a wide range of human diseases, as they are often the direct cause for a protein to partially or totally lose its function [1]. With the rise of high-throughput sequencing techniques, a very frequent clinical procedure consists in analyzing the genotype of patients who suffer from a disease to confirm or deny the presence of a mutation in the involved genes. Distinguishing such mutations from polymorphisms without significant effect on human health is a necessary step in understanding the etiology of mutation-related diseases [1]. Amino acidic replacements can affect to a different extent the stability, catalytic efficiency and regulatory properties of proteins. However, in the majority of the cases there is no apparent relationship between the nature and location of the replaced amino acid and the type of molecular perturbation [2]. This fact makes predicting the functional effects of missense variants in proteins a challenging task.

This field of study has been addressed in the recent years, and a considerable amount of bioinformatics tools to tackle the problem have been developed. However, while most state-of-the-art pathogenicity predictors rely on sequence-based features of proteins to identify damaging variants [3], structural and dynamic features have been largely neglected, mainly due to the difficulties and computational cost of their obtainment. Although molecular dynamics (MD) simulations have been traditionally employed to study the impact of mutations on the dynamic behavior of proteins, these studies are generally based on qualitative terms, and we still lack techniques to better quantify these dynamic effects. Our research line aims to improve pathogenicity prediction algorithms by establishing a protocol or guidelines that consider the role of dynamics features in the functional behavior of protein variants.

B. Use case

To this end, we are studying the particular case of the human erythrocyte pyruvate kinase (R-PYK). This protein is one of the most widely studied enzymes throughout the history of biochemistry, due to its major role in the regulation of glycolysis. It catalyzes the irreversible conversion of phosphoenolpyruvate (PEP) to pyruvate, generating an ATP molecule in the process. The enzyme needs cofactors (K^+ and Mg^{2+}) for proper catalytic activity, and is allosterically regulated by fructose-1,6-biphosphate (FBP), an activator of its catalytic efficiency. Pyruvate kinase has a highly conserved architecture throughout evolution (Fig. 1) [2].

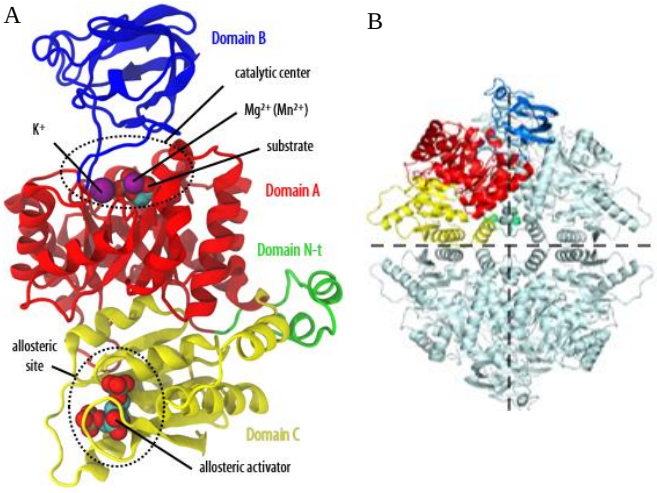

Fig. 1 Tertiary and quaternary structure of R-PYK. The different domains of the protein: N-terminal, A, B and C are colored green, red, blue and yellow respectively. A) View of a monomer of R-PYK and bound cofactors, substrate and allosteric activator. B) View of the tetramer of R-PYK. The two symmetry axes are shown with a dashed line. Only one subunit is colored as indicated above (the rest are colored pale cyan). Images generated with the software VMD.

Hundreds of missense mutations in various sites of R-PYK have been clinically associated to a disease called hereditary nonspherocytic hemolytic anemia (5 cases per 10,000 individuals in Europe) [4]. The main aim is to go deeper into the coupling between structure and functions of this protein, as well as to provide information about its dynamic properties for both the WT form and some of its mutant variants of known pathological consequences.

C. Preliminary results

So far, we have started to characterize the main motions occurring in the WT enzyme, as well as some potential alterations of these patterns in the mutated forms of the protein. These preliminary descriptions are the result of the analysis of 100 ns long MD simulations run on both the WT (5 replicas) and 62 relevant pathogenic and neutral variants extracted from literature and clinical sources. The amino acidic substitutions affect various regions of the protein, such as the active and allosteric sites, subunit interfaces and hydrophobic cores.

A thorough examination of the generated trajectories allowed us to understand the main concerted motions that govern the dynamics of the enzyme, in accordance with the

previously described features of its biological function and mechanism.

D. Future development

In the current stage of the project we are expanding the amount of MD simulations to cover a more comprehensive representation of the biologically meaningful conditions of the enzyme. Until now, the simulations consisted of the apoenzyme (only protein residues), therefore, we will establish several additional conditions of the enzyme bound to its cofactors (K^+ and Mg^{2+}), substrates (PEP and ADP) and the allosteric activator (FBP), in different combinations, and both for the WT and the mutant variants. To achieve this, we will use quantum mechanics (QM) approaches to model the bound ligands, prior to the MD simulations. The intended production MD simulation length of each variant will be of 200 ns. Through the analysis of all the generated data, we aim to establish a model to better explain the potential dynamic and/or structural alterations of each amino acid replacement in contrast with the WT behavior.

Later stages of the project will include, at least, generating a set of metrics (structural and derived from essential dynamics analyses) with discriminative power, training an R-PYK specific machine learning model to predict pathogenicity, validating it with more R-PYK variants and testing it with more proteins of similar features.

ACKNOWLEDGEMENTS

We thankfully acknowledge the computer resources at MareNostrum and the technical support provided by Barcelona Supercomputing Center - Centro Nacional de Supercomputación (BSC-CNS) and Red Española de Supercomputación (RES) to cope with the computational requirements of our project (RES-BCV-2019-1-0004).

References

[1] C. Feinauer and M. Weigt, "Context-Aware Prediction of Pathogenicity of Missense Mutations Involved in Human Disease", 1–21, 2017.

[2] G. Valentini *et al.*, "Structure and function of human erythrocyte pyruvate kinase: Molecular basis of nonspherocytic hemolytic anemia", J. Biol. Chem., 277, 23807–23814, 2002.

[3] S.E. Flanagan *et al.*, "Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations", Genet. Test. Mol. Biomarkers, 14, 533–537, 2010.

[4] L. Montllor *et al.*, "Red cell pyruvate kinase deficiency in Spain: A study of 15 cases", Med. Clínica (English Ed.), 148, 23-27, 2017.

Author biography

Luis Jordà was born in Barcelona, Spain, in 1993. He received the BSc in Biochemistry from the University of Barcelona (UB), in 2016, and later the MSc in Bioinformatics for Health Sciences, from the University Pompeu Fabra (UPF), in 2018, in Barcelona, Spain.

He is currently pursuing his PhD in Biomedicine (UB) and performing his research at the facilities of the Barcelona Supercomputing Center (BSC), under the supervision of Prof. Josep Lluís Gelpí. He is also an active collaborator of the Molecular Modeling and Bioinformatics (MMB) group from the Institute for Research in Biomedicine (IRB) of Barcelona. His current research interests include structural bioinformatics, protein dynamics and modeling of biomolecules.