SciForum MOL2NET

Comparative Modeling of the Three-Dimensional Structure of Protein Kinase D from Mycobacterium Tuberculosis.

Alejandro Morales-Bayuelo [1,*].

[1] *Centro de Bioinformática y Simulación Molecular (CBSM), Universidad de Talca, 2 Norte 685, Casilla 721, Talca, Chile.*; E-Mail: almorales@utalca.cl

***** Author to whom correspondence should be addressed; E-Mail: almorales@utalca.cl.

Abstract: Among the main aims of the United Nations is to eradicate communicable diseases such as tuberculosis, which affect millions of people worldwide producing more critical problems in countries with low- and middle-income. *Mycobacterium tuberculosis*, it is the causative agent of tuberculosis; it is one of the most lethal human pathogens further characterized by being strongly resistant. Despite more than 100 years of research developed until today, some two million people die every year around the world due to tuberculosis. For this reason in this study are postulates and analyzed the homology modeled to the protein kinase D from *mycobacterium* tuberculosis. Homology modeling aimed at constructing three-dimensional structure of a protein model using experimentally determined structure of related family members as the template. Protein kinase homology or comparative protein structure modeling was performed with the help of *UniProt* (de universalt protein) and SWISS-MODEL workspace, which is an integrated web-based modeling expert system and by selecting suitable template solved by experimental methods and stored at Protein Data Bank (PDB) database. On the basis of a sequence alignment between the target proteins Pks D and the template structure PKn B, a three-dimensional model for the target proteins were generated. SWISS-MODEL workspace derived the restraints automatically from related known structure (template) present in the database. Three-dimensional structure or model was generated by optimizing the molecular probable density function. The generated model with the highest score was validated by the probable density functions. The validated model was chosen for further studies and refinements. The rough model generated was subjected to energy minimization using the steepest descent technique to eliminate bad contacts between protein atoms. Computations were carried out in vacuo by the GROMOS96 force field set, implemented through Gromacs 4.1, and steriochemical quality of minimized model was performed using Ramachandran Plot, PROCHECK tool, WHAT_CHECK , ERRAT, VERIFY_3D and ProSA-web.

Keywords: mycobacterium tuberculosis, protein Kinase D, homology models, topological analysis, secondary and tertiary structures.

Introduction: Historically, one of the main aims of the United Nations is to eradicate communicable diseases such as tuberculosis, which affect millions of people worldwide producing more critical problems in countries with low- and middle-income. *Mycobacterium tuberculosis*, it is the causative agent of tuberculosis; it is one of the most lethal human pathogens characterized by be strongly resistant. Despite more than 100 years of research developed until today, some two million people die every year around the world due to tuberculosis. According to the World Health Organization (WHO), a third of the world's population carries the infection in an inactive form known as latency **[1]** and a hallmark of the disease is the ability to persist in the host for years and to reactivate under conditions of immune suppression. The situation is worsened by the increasing incidence of multi-drug-resistant strains. In this sense, our current inability to control the spread of this disease can be explained by the lack of an effective vaccine, lack of multidrug resistant **[2-4]** and the great adaptability of the *mycobacterium tuberculosis* in different environments (e.i.; high mutation) **[5]**. It is therefore imperative to identify novel *mycobacterium tuberculosis* antigens/targets for the development of new effective anti-tubercular drugs and vaccines. The Protein Kanases (Pks) plays an important role in controlling proliferation and differentiation in eukaryotic cells in living organisms, are enzymes that catalyze the protein phosphorylation process. One reason to investigate the protein phosphorylation is due to that its rationalization represents an attractive drug target in a variety of diseases such as cancer **[6]**, alzheimer **[7]**, chronic inflammations **[8]**, etc. PKs present in the human body have been widely studied by the interest in their use as therapeutic targets in these diseases; however, not much is known about the PKs involved in tuberculosis for this reason propose new models using the homology modeling can help to get new insights in tuberculosis treatment today.

The determination of the 3D structure also provides valuable information about the function of such proteins whose functions are otherwise unknown. X-ray crystallography is a powerful tool for determining protein 3D structures but it is time consuming and expensive, and not all proteins can be successfully crystallized. This is especially the case with most membrane proteins like kinase that are difficult to crystallize and that do not dissolve in commonly used solvents. Therefore, very few membrane protein structures from *mycobacterium tuberculosis* have been determined by X-ray crystallography so far.

.

Materials and Methods: Homology modeling aimed at constructing three-dimensional structure of a protein model using experimentally determined structure of related family members as the template. Protein kinase homology or comparative protein structure modeling was performed with the help of *UniProt* (de universalt protein) and SWISS-MODEL workspace, which is an integrated web-based modeling expert system and by selecting suitable template solved by experimental methods and stored at Protein Data Bank (PDB) database. On the basis of a sequence alignment between the target protein Pkn D and the template structures Pkn E (homology: 65.7% with PKn D), a three-dimensional model for the target protein was generated. SWISS-MODEL workspace derived the restraints automatically from related known structure (template) present in the database. Three-dimensional structure or model was generated by optimizing the molecular probable density function. The generated model with the highest score was validated by the probable density functions. The validated model was chosen for further studies and refinements. The rough model generated was subjected to energy minimization using the steepest descent technique to eliminate bad contacts between protein atoms.

Computations were carried out in vacuo by the GROMOS96 force field set, implemented through Gromacs 4.1, and steriochemical quality of minimized model was performed using Ramachandran Plot, PROCHECK tool (checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry), WHAT_CHECK (derived from a subset of protein verification tools from the WHATIF program, this does extensive checking of many sterochemical parameters of the residues in the model), ERRAT (analyzes the statistics of non-bonded interactions between different atom types and plots the value of the error function versus position of a 9-residue sliding window, calculated by a comparison with statistics from highly refined structures), VERIFY_3D (determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures), PROVE (calculates the volumes of atoms in macromolecules using an algorithm which treats the atoms like hard spheres and calculates a statistical Z-score deviation for the model from highly resolved (2.0 Å or better) and refined (R-factor of 0.2 or better) PDB-deposited structures). Finally, was used ProSA-web, this program calculates an overall quality score. If this score is outside a range characteristic for native proteins the structure probably contains errors. A plot of local quality scores points to problematic parts of the model which are also highlighted.

Results and Discussion: In the model predicted to Pkn D the Ramachandran Plot analysis showed that amino acids of model (**Figure 1B**) and template in the most favorable region were 76.4% and in the additional allowed region were 22.7% amino acids, respectively. This stipulates that protein backbone dihedral angles φ-ψ occupied reasonably accurate positions in the 3D model.

The template was selected based on amino acid sequence similarity and crystal resolution, 100 models were generated and the model showing the least RMSD with respect to trace (C α atoms) of the crystal structure of the template was saved for further modification and validation (**Figure 1**). The modification was performed to obtain the best conformation of modeled Pkn D. Among the available potential templates, crystal structure of Pkn E with 2.8 Å resolution was selected as the model quality assessment template structure to build molecular model of the Pkn D.

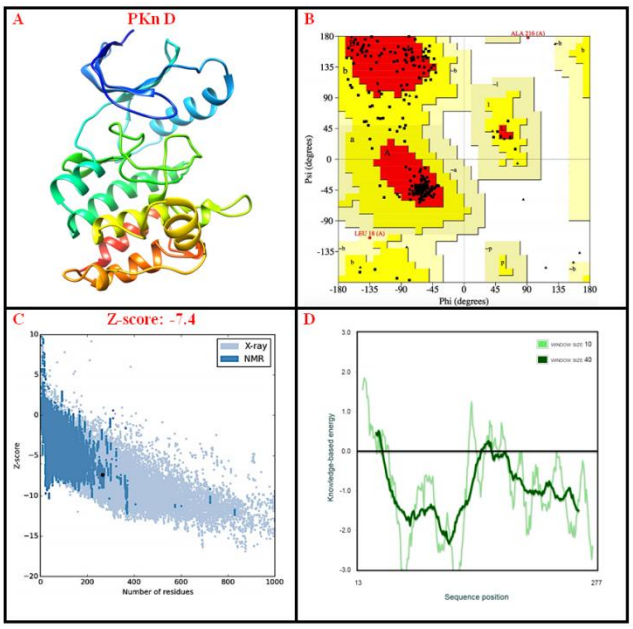

Figure 1. **(A)** three-dimensional model predicted to Pkn D. **(B)** Ramachandran plot for the predicted model of Pkn D. Residues in most favoured regions [A,B,L] 172 (76.4%), residues in additional allowed regions [a,b,l,p] 51 (22.7%), residues in generously allowed regions [~a, ~b, ~l, ~p] 1 (0.4%),

residues in disallowed regions 1 (0.4%), number of non-glycine and non-proline residues 225 (100%), number of end-residues (excl. Gly and Pro) 1, number of glycine residues (shown as triangles) 21, number of proline residues 18, total number of residues of 265. **(C)** the *z*-score indicates overall model quality. Its value is displayed in a plot that contains the *z*-scores of all experimentally determined protein chains in current PDB. In this plot, groups of structures from different sources (X-ray, NMR) are distinguished by different colors. **(D)** this plot shows local model quality by plotting energies as a function of amino acid sequence position *i*. In general, positive values correspond to problematic or erroneous parts of the input structure.

The ProSA-web study, the interaction energy per residue was calculated by PROSA 2003 program. The PROSA Z-score indicates overall model quality (acceptable values are below 0.5). The overall model quality showed the Z-score of −7.4 in **Figure 1C**. The score achieved by our model is within the range of scores found in native-protein structures of similar size. The overall quality parameters (stereochemical, geometrical as well as energetic) achieved by our Pkn D model suggested that it can be trusted. Another test for quality assessment is to examine the ProSA profile which describes the energy of residues as a function of sequence position; in which positive values indicate a problematic or erroneous part of the structure and vice versa. The energy analysis for our model (**Figure 1D**) demonstrated favorable residues energy as indicated by the negative values achieved along the amino acids sequence. In **Figures 2A** and **2B** is shows the ERRAT and VERYFY_3D plots to this Pkn D.

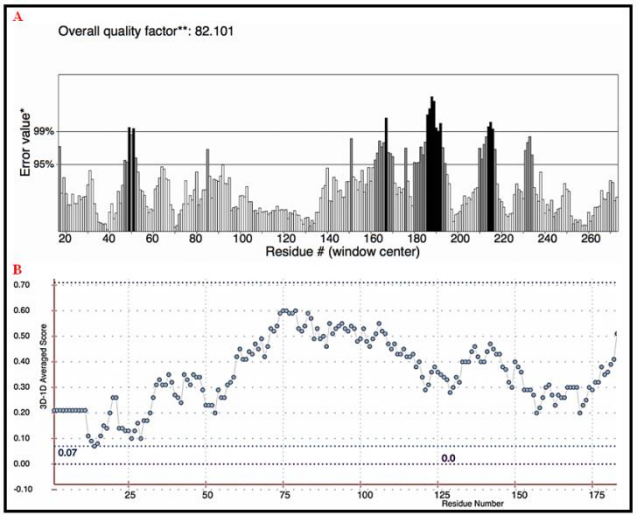

Figure 2. ERRAT and **VERIFY_3D** plots to the Pkn D model. **(A) ERRAT** plot analyzes the statistics of non-bonded interactions between different atom types and plots the value of the error function versus position of a 9-residue sliding window, calculated by a comparison with statistics from highly refined structures. **(B) VERIFY_3D** determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures.

In **Figure 2**. ERRAT for non-bonded atomic interactions and higher scores means better quality. The normal received range for a high quality model is >50. The ERRAT score of Pkn D model was shown in **Figure 1A**. The above validation suggests that the backbone conformation and non-bonded interactions of Pkn D homology model was all reasonable within a normal range. The prediction of the modeled Pkn D structure was checked by VERIFY_3D, **Figure 2B**. The VERIFY 3D analysis indicated a reasonably good sequence-to-structure agreement because none of the amino acids had a negative score and it is 0.61 Pkn D. It is to be noted that compatibility scores above zero correspond to acceptable side chain environment. Therefore, the validation used shown that our homology model can be an alternative to the Pkn D from tuberculosis unknown today.

Conclusions: the homology modeling has been used to propose the first 3D structures for Protein kinases D from *mycobacterium tuberculosis*. With the assistance of the well-defined features associated to protein kinases involved the "gatekeeper door", hinge zone, C Helix, Asp-Phe-Gly (DFG), C-terminal and N-terminal, we can predict functional and binding sites, which can help in understanding what biological role it fulfills and desing new inibithors to the tuberculosis treatment. The validation used shown that our homology model can be an alternative to the protein kinase B from tuberculosis unknown today.

Conflicts of Interest: *The author declare no conflict of interest*

Acknowledgements: Thanks to the Universidad de Talca (CBSM)) for the continuous support to this investigation and finally to the postdoctoral project N^0 3150035 (FONDECYT, CHILE).

References:

1. World Health Organization (WHO), www.who.org, Health topics (Tuberculosis), reviewed September 2016.
2. M. Cully. *Nat. Rev. Drug Discov, 2014, 13, 256.*
3. R. E., Lee. *Nature Medicine, 2014, 20, 152.*
4. A. Zumla; P. Nahid; S. T. Cole. *Nat. Rev. Drug Discov. 2013, 12, 388.*
5. G. B. Migliori; C. Lange; R. Centis; G. Sotgiu; R. Muetterlein. H. Hoffmann; K. Kliiman; G. De laco; F. N. Lauria; M. D. Richardson; A. Spanevello; D. M. Cirillo; T. S. Grp. *Eur. Resp. J. 2008, 31, 1155.*
6. K. R. Hallows; R. Alzamora; H. Li. *Am. J. Physiology - Cell Physiology.* 2009, 296, C672.
7. A. Salminen; K. Kaarniranta; A. Haapasalo; H. Soininen; M. Hiltunen. J. Neurochem. 2011, 118, 460.
8. A. Kotlyarov; A. Neininger; C. Schubert; R. Eckert; C. Birchmeier; H. D. Volk; *M. Gaestel. Nat. Cell. Biol. 1999, 1, 94.*

Notes: *MDPI do not released one specific template for Sciforum conference. Consequently, this is not official template released by MDPI. In principle, the papers may be presented without specific format. However, the chairperson and the secretariat of the conference decided to create this template to give homogeneity to Mol2Net works. As is, this template should be used only for Mol2Net conference. Please, do not use this template for MDPI journals. Please, delete these notes before saving your final version. Type of the Paper (Proceeding, Letter, Expert Opinion, Communication, etc.). MDPI generates doi upon author request.*

YouTube link: *this option is only for those authors with welcome videos and/or oral presentations, plenary conferences uploaded to Mol2Net YouTube site.*