Conserved domain and structure analysis of a putative polyphosphate kinase from Buruli ulcer causing bacterium

Zarrin Basharat*

Department of Environmental Sciences, Fatima Jinnah Women University, Rawalpindi 46000, Pakistan.

*Corresponding author

Email: zarrin.iiui@gmail.com

Abstract

With increasing sophistication of instruments and techniques, in addition to the increment in intricacies, girth and complexities of the problems being addressed, simple methods (especially computational biology techniques) are being overlooked, replaced or phased out. One such technique on the twilight of survival is simple computational analysis of protein sequence i.e. property determination, homology modeling etc. Manuscripts reporting solely such type of analysis face upfront rejection, although some exceptions might exist. Only some predatory or beginner journals might accept such publications. This continues despite the fact that simple, cost effective, quick computational analysis of protein sequence has its merits and paves way for further research. This report is basically an attempt to keep the dying venture of protein structure modeling alive.

Keywords: Computational sequence analysis, Homology modeling, Dynamics simulation, Mycobacterium ulcerans.
Buruli ulcer is a neglected tropical disease caused by *Mycobacterium ulcerans* (Asiedu *et al*., 2000). The causative bacterium is a slow growing microaerophilic bacterium. The disease is spread over various regions and cases have been reported in almost 32 countries (Nakanaga *et al*., 2013). It usually affects poor people in remote places with very limited access to healthcare. The disease may impact people of all age groups but children under the age of 15 years (range 2–14 years) are more prone to this type of ulcer.

Availability of a wide variety of efficient tools and web servers has benefited computational biologists and they can perform reliable analysis of protein sequence and structure in a very short time span (Basharat, 2015). The sequence analysis of the kinase using conserved domain database of NCBI, revealed two conserved domains, a Catalytic C-terminal domain matching repeat of *Pseudomonas aeruginosa* polyphosphate kinase (residue 53-132) and another domain (residue 1-45) responsible for the synthesis of Poly P. Inorganic polyphosphate plays an important role in bacterial stress responses and stationary-phase survival. Both domains can catalyze the reversible conversion of the terminal-phosphate of ATP to Polyphosphate. These domains also catalyze the synthesis of polyphosphate from GTP or ATP, with a preference for manganese over magnesium ions, with chain lengths of up to a thousand or more orthophosphate molecules. In this study, structure of the polyphosphate kinase from this bacterium has been determined using I-TASSER (Zhang, 2008; Roy *et al*., 2010) with Chain A of a crystal structure of polyphosphate kinase from *Escherichia coli* and *Porphyromonas Gingivalis* as a template. The structure was energy minimized using Molecular Operating Environment software and then visualized (Figure 1).

Homology modeling and fold recognition approaches were unsuccessful in structure prediction due to lack of similarity. When there are no comparable protein sequences to the target protein sequence in the structure databases, *ab initio* method is used because it tries to predict the protein tertiary structure using physical principles to fold the protein from an arbitrary conformation (Skolnick and Kolinski, 2001; Basharat and Yasmin, 2016).

Protein structure was predicted and energy minimized as lowest energy conformation means stable structure. Classically, *ab initio* protein tertiary structure prediction carries out a conformational search with the guidance of an energy function and the goal is to search the protein conformational search space to find the lowest free energy conformation. Firstly, protein conformation was represented based on the treated degree of freedom, ranging from all atoms representation to simplified or reduced representation. Next, an energy function well-matched with the protein conformation representation was used to compute the conformation energy; and then, a conformational search algorithm was utilized to search the conformation search space to find the lowest free energy conformation. The protein conformational search space consisted of all possible conformations of the protein. This study is the first to report the structure of a putative polyphosphate kinase from *Buruli ulcer* causing bacterium.
Figure 1. Modeled structure of putative polyphosphate kinase from *Mycobacterium ulcerans*

The structure was simulated for a short time span using Molecular Operating Environment (according to Basharat *et al.*, 2016). Potential energy fluctuations were observed throughout the simulation span. After 120 picoseconds, the structure became stable in conformation (Figure 2).

Figure 2. Correlation plot of time versus energy of the studied protein. Data for first 200 picoseconds is shown.
It has been previously reported that kinases can serve as drug targets (Schindler et al., 2007; Szekely et al., 2008; Basharat and Yasmin, 2015) as they block host-pathogen signalling mechanism necessary for pathogenesis. This structure can, therefore, be useful for docking studies with different drugs for further analysis as a drug target.

References


