Journal of
Computer Science & Systems Biology

Pillai and Pardasani, J Comput Sci Syst Biol 2012, 5:1
DOI: 10.4172/jcsb.1000084

Research Article Open Access

Modeling Calcium Dependent Protein Kinase Isoform 1 from *Cicer arietinum* (chick pea)

Lakshmi Pillai* and K. R. Pardasani

Department of Bioinformatics, MANIT, Bhopal, India

Abstract

The homology based 3-Dimensional structure prediction of calcium dependent protein kinase Isoform 1 From *Cicer arietinum* (chick pea) was carried out using bioinformatics tools. The *CPK1* sequence on Protein BLAST analysis for homology search revealed 100 hits and Out which a few had significant E score (E <0.005) and better sequence similarity. Multiple sequence alignment analysis of *CPK1* using MultAlin and HHpred showed above 90% similarities with protein sequences of *Thermotoga petriphila (Hypothetical protein)*, *Geobacillus stereothermophilus* (1w91; 99.5%; E score=1.2 E-11), *Thermoanaerobacterium saccharolyticum* (1uhv; 99.4%; E score=2.9 E-11) and *Bacillus stereothermophilus* (1qw9; 98.5%; E score=4.9 E-6) from the PDB database. The secondary structure of *CPK1* using PSIPRED VIEW revealed many helices, strands and coils in the protein structure. The tertiary structure prediction of *CPK1* by MODELLER 8v2 showed a (ß/á) 8 fold. The program VERIFY 3D assessed the quality of the predicted structure of *CPK1* with acceptable scores.

Keywords: BLAST; Calcium-dependent protein kinase isoforms; Hypothetical protein; Homology; Modeling; Multiple sequence alignment

Introduction

Chickpea (*Cicer arietinum*) is one of the most important grain legume crops worldwide and a major source of protein for millions of families in developing countries.

A protein kinase is a kinase enzyme that modifies other proteins by chemically adding phosphate groups to them (phosphorylation). This class of protein is further separated into subsets such as PKC alpha, PKC beta, and PKC gamma, each with specific functions. Phosphorylation usually results in a functional change of the target protein (substrate) by changing enzyme activity, cellular location, or association with other proteins. Up to 30% of all proteins may be modified by kinase activity, and kinases are known to regulate the majority of cellular pathways, especially those involved in signal transduction, the transmission of signals within the cell. Protein kinases have profound effects on a cell, their activity is highly regulated. Kinases are turned on or off by phosphorylation (sometimes by the kinase itself - *cis*-phosphorylation/autophosphorylation), by binding of activator proteins or inhibitor proteins, or small molecules, or by controlling their location in the cell relative to their substrates.

Calcium regulated protein kinase in plants are CPKs. CPK is a monomeric enzyme that has a wide tissue distribution in *Cicer arietinum* (chick pea) and can phosphorylate a number of substrates. Phosphorylation occur on a loop near the active site.CPK/Cam kinases are autoinhibited in the resting state (low intracellular concentration) by a regulatory segment that follows the catalytic domain.

In one of the GPCR mediated pathway i.e inositol phospholipid pathway activated Phospholipase-C gives product DAG (diacylglycerol) and IP3 (Inositol 1,4,5-trisphosphate) diffuses to the endoplasmic reticulum, binds to a ligand-gated ion channel, and stimulates the efflux of calcium from the lumen into cytosol. Cytosolic calcium may bind, together with diacylglycerol, to another protein kinase and activate them and thus in this cascade Protein Kinase calmodulin independent isoform1 is also activated.

Homology model of CPK1 was built to understand the mechanism of kinase autoinhibition and activation upon calcium binding and subsequent phoshorylation [1].

Materials and Methods

With the development of techniques in molecular biology rapid identification, isolation, and sequencing of genes, is now able to infer the sequences of many proteins. However, it is still a time-consuming task to obtain the three-dimensional structures of these proteins. A major goal of structural biology [2] is to predict the three-dimensional structure from the sequence, a pursuit that has not yet been realized. Thus, alternative strategies are being applied to develop models of protein structure when the constraints from X-ray diffraction or NMR are not yet available. One method that can be applied to generate reasonable models of protein structures is homology modeling [3]. This procedure, also termed comparative modeling or knowledge-based modeling, develops a three-dimensional model from a protein sequence based on the structures of homologous proteins.

Sequence retrieval

Protein sequence corresponding to cpk1 (target protein) was retrieved from non redundant (nr) database in fasta format. To model a protein structure from its sequence is divided in four main steps: finding a template; aligning the target and the template; building models; assessing the models.

Template(s) for the cpk1 were identified in protein data bank using BLAST server. The BLAST search was restricted to PDB only. In order to select suitable templates, following criteria were used: (i) E-value <

***Corresponding author:** Lakshmi Pillai, Department of Bioinformatics, MANIT, Bhopal, India, E-mail: lakshmilster@gmail.com

Received March 16, 2010; **Accepted** January 25, 2012; **Published** February 02, 2012

Citation: Pillai L, Pardasani KR (2012) Modeling Calcium Dependent Protein Kinase Isoform 1 from *Cicer arietinum* (chick pea). J Comput Sci Syst Biol 5: 009-011. doi:10.4172/jcsb.1000084

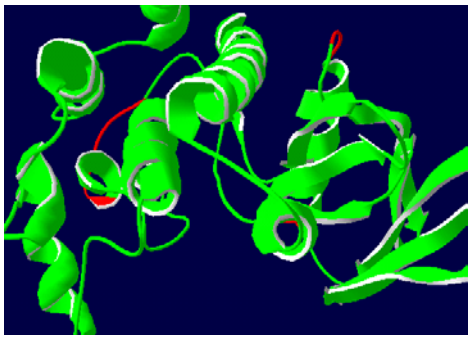

Figure 1: Model for protein kinase isoform1.

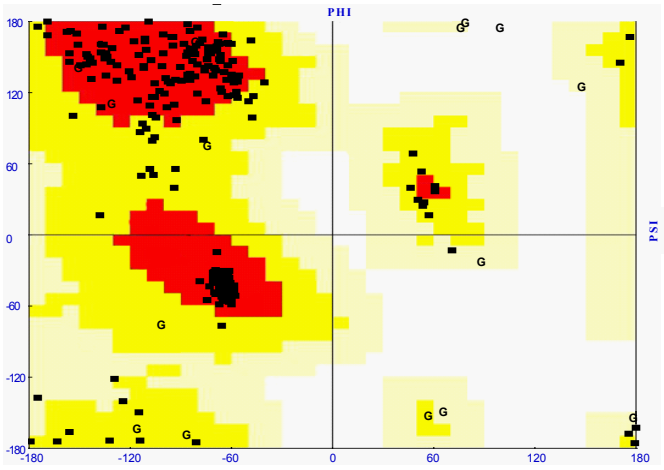

Figure 2: Ramachandran plot.

10-4; (ii) highest resolution and (iii) the lowest R-value from among all template structures. Here the resolution found is 1.8 Armstrong, E-value is 2.73e-54 and the template selected is 2BDW. Then multiple sequence alignment containing the target sequence and all the homologues is done using the tool CLUSTAL W. Multiple alignment provides more information than a single sequence, such as protein domain structure, surface exposure of residues and their involvement in the protein function.

To build the model method [4] used will depend on the quality of the template found and the score of its alignment with the target sequence. Less than 20% of sequence identity, the sequence is divided into domains using secondary structure and possible repeats in the sequence. This is done by PROSITE Scan. Structural domain found is from 91-350, model is built for the target sequence by using tools like SWISS MODEL and Modeller [5] then these models are analyzed either by superimposing two structures or by mathematically calculating the RMSD value i.e. root mean square deviation value which is in 0.2 -1.00 Armstrong range if the two structures are very similar which is applied using Deep view programme. Then R.M.S value is calculated which is observed as 0.37 Armstrong, between the template and protein kinase isoform1 model. Thus the model is build and it is minimized by using Deep view and Molecular orbital environment (MOE).

Model Assessment

The final quality of the model will very much depend on the quality of the initial alignment and the resulting sequence identity [6]. There are different types of evaluation to assess how accurate a model is geometrical checks, empirical energy functions, scoring using statistics. The accuracy of a protein structure model will determine its utility. If a known protein has a high level of identity with the query sequence (over 60%) then it should be feasible to produce a high quality model structure [7], in which the C alpha atom RMSD is less than 1 angstrom. The bioinformatics tool PROCHECK gives the Ramachandran map which tells what conformation is possible for a polypeptide chain and the accuracy of model.

Results and Discussion

The three dimensional structure reveals extensive interaction between the autoinhibitory sequence and the catalytic core which is consistent with the pseudosubstrate model for the activation of kinase. (Figure 1) Here in the procheck (10) results I found that no residues are falling in disallowed region hence I can say that my swiss model and template are identical.

Ramachandran plot gives zero percent of disallowed region, 79.1% core, and 18.7% allowed region. So model is verified as correct. Verify 3D is also used for evaluation in which 17.18% of residues had an average 3D-1D score > 0.2, and more the score better is the model. (Figure 2)

As the model and PSIPRED bioinformatics tool output secondary structures matches the model is verified as correct [8]. (Figure 3a,b)

Conclusions

The Homology model of calcium-dependent calmodulin-independent Protein Kinase isoform1 (cpk1) from chick pea (*C. Arietinum*) was built based on Cam kinase I from *C. eligans*, in order to explain its autoinhibition mechanism and kinase activation.

The structure of cpk1 reveals extensive interaction between autoinhibitory sequence and the kinase catalytic core, which partially blocks the active site in the inactive state of kinase. Ca^{2+} binding at the Cterm Cam-like domain release autoinhibitory part which opens up the active site. This process activates kinase.

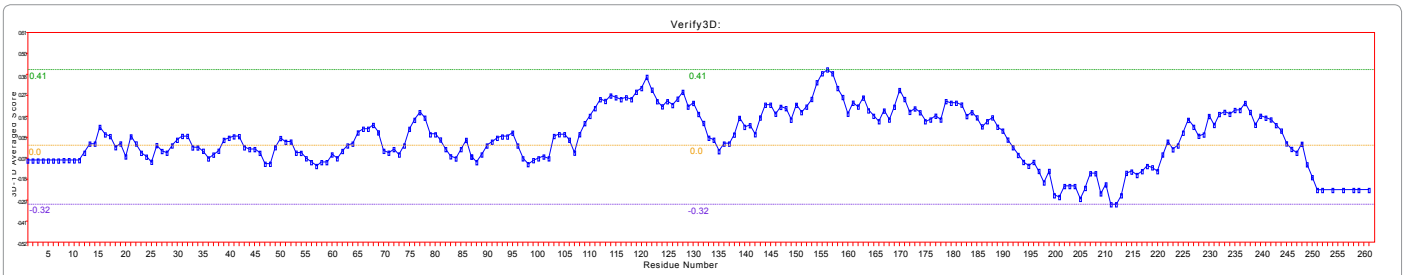

Figure 3a: Verify 3D output.

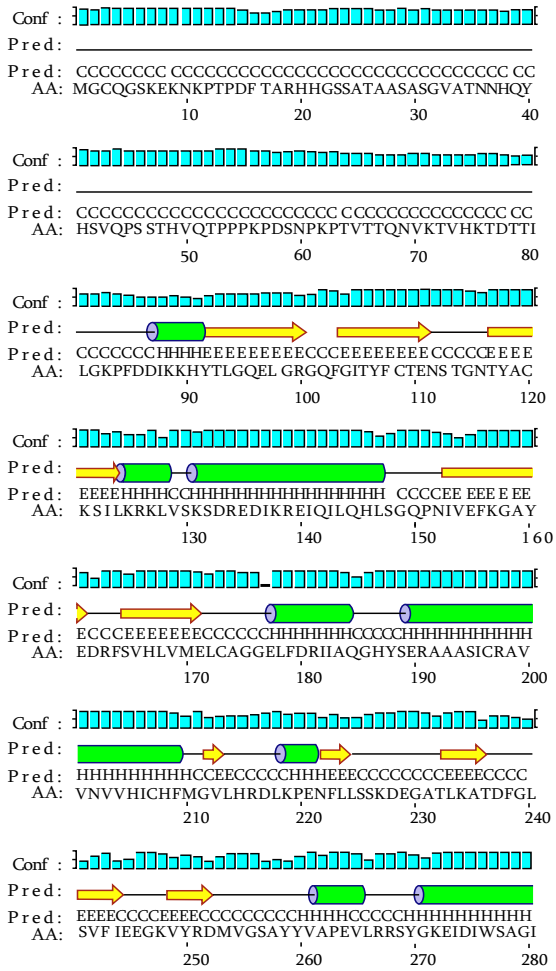

Figure 3b: PSIPRED tool output of secondary structure prediction.

Here a correlation between structure and function is observed as if protein kinase is not activated then transcription factor is not able to bind the active site and this may result in degradation of plant growth and development.

The homology model was built with a purpose to understand the mechanism of autoinhibition and kinase activation as well as to identify unstructured part in the model. The model reveals N terminal residues (1-90) are unstructured. This unstructured part can be biochemically cleaved. Then the truncated protein which is mostly structured could be crystallized and that will be useful for further crystallographic studies.

Acknowledgment

The authors are highly thankful to the Department of Biotechnology, Delhi, India and M.P. Council of Science and Technology M.P., Bhopal, India for providing support in the form of Bioinfomatics infrastructure facility to carry out this work.

References

1. ExPASy Proteomics Server.
2. Vitkup D, Melamud E, Moult J, Sander C (2001) Completeness in structural genomics. Nat Struct Biol 8: 559-566.
3. Homology modeling principles and application by Andrew R.Leach.
4. Homology modeling by Erik Lindahl, Olle Edholm , Arne Elofsson.
5. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2006) Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics 5: 5.6.
6. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, et al. (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 29: 291-325.
7. Chothia C, Lesk AM (1986) The relation between the divergence of sequence and structure in proteins. EMBO J 5: 823-826.
8. Sali A, Blundell TL (2004) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234: 779-815.
9. Baker D, Sali A (2001) Protein structure prediction and structural genomics. Science 294: 93-96.
10. Rathore RS (2005) Conformational investigations on analogs of inflammation response inducing chemotactic tripeptide fMLP. Biopolymers 80: 651-664.