Computational Molecular Biology 2014, Vol. 4, No. 15, 1-5
http://cmb.biopublisher.ca

Research Report **Open Access**

An in-silico analysis showing the interaction of FAD ligand with cytokinin dehydrogenase enzyme and with its domain part in rice

Pranati Swain, Lambodar Behera
Central rice research institute, India
Corresponding author email: rosylora20@gmail.com
Computational Molecular Biology, 2014, Vol.4, No.15 doi: 10.5376/cmb.2014.04.0015
Received: 10 Dec., 2014
Accepted: 26 Dec., 2014
Published: 30 Dec., 2014

Preferred citation for this article:
Swain and Behera, 2014, An in-silico analysis showing the interaction of FAD ligand with cytokinin dehydrogenase enzyme and with its domain part in rice, Computational Molecular Biology, Vol.4, No.15, 1-5 (doi: 10.5376/cmb.2014.04.0015)

Abstract The cytokinin dehydrogenase enzyme plays an important role in the high grain production of rice. This enzyme interacts with the FAD ligand. This study shows the domain region of the enzyme which was identified by protparam tool, generated homology models using modeller9.12 tool, models were validated using errat, procheck saves server, prosa and anolea server. Then the active sites were predicted by using castP server. Finally the interaction was studied between the FAD and cytokinin dehydrogenase and with its domain part. The stronger interaction was found between the FAD and the domain of the cytokinin degydrogenase with the binding enegy of -10.91. ALA6, ARG1, AR536, TYR10, ASP55, ALA57, CYS38 of FAD is binding to the domain part of cytokinin degydrogenase in rice.
Keywords FAD Ligand; Cytokinin dehydrogenase enzyme; Interaction between ligand and enzyme; Functional domain

Introduction

Cytokinin dehydrogenase of *Oryza sativa (japonica) plays an important role in high grain production. The cytokinin dehydrogenase helps in* the oxidation of cytokinins. This protein belongs to the family of N(6)-substituted adenine derivatives, which is a plant hormone having isopentenyl group. This protein helps in modulating the number of reproductive organs by regulating the cytokinin accumulation in inflorescence meristems. The CKX2 is also known as OsCKX2. The CKX2 is generally located on Os01g0197700. FAD or flavoprotein is the ligand of Cytokinin dehydrogenase 2 proteins in *O.sativa.CKX2 is located on chromosome number 1. CKX2 has 5137 basepairs. Cytokinin positively regulates the growth of shoot apical meristem that helps in more seed production. The Gn1a gene of rice encodes* cytokinin oxidase as a major quantitative trait locus. zinc finger transcription factor directly regulates the OsCKX2 expression in the reproductive meristem (Li et al., 2013) Reduced expression of OsCKX2 causes cytokinin accumulation in inflorescence meristems leading to the increase in number of reproductive organs which results more grain production (Ashikari et al., 2005). CKX helps for high grain production in wheat, maize, barely, sorghum, foxtail millet (Mameaux et al., 2011). Cytokinin dehydrogenase plays an important role in the formation of crown root system in rice, which helps in high grain production (Gao et al., 2014). The main objective of this study is to find out the best interaction among the ligand FAD and enzyme. Here the interaction is studied by considering the whole amino acid sequence and then only the domain part. Domain is the function region of a protein or domain. So generally the binding is more stronger at domain part. This study involves retrieval of amino acid sequence, predicting the domain region of enzyme, finding the common residues among domain part and the whole sequence, generating homologous models, validating and evaluating models, predicting the active site, interaction of FAD with the enzyme.

1 Material and Methods

1.1 Sequence retrieval

The amino acid sequences of cytokinin dehydrogenase was retrieved from uniprot in fasta format.

1.2 Finding the domain part

The sequence of domain in the protein was retrieved

by using protparam tool. However the domain region can be found from CDD (conserved domain database) too.

1.3 Alignment between whole protein sequence and domain

Multiple sequence alignment is generally done to find out the conserved residues among multiple sequences (Thompson et al., 1994). From the alignment we found the regions of common residues among the domain part and whole sequence depending upon which we could analyse the stronger binding region where ligand (FAD) is binding. The alignment result is shown in Figure 1.

```
sp|Q4ADV8|CKX2_ORYSJ      APDGVVVDMASLGRLQGGGARRLAVSVEGRYVDAGGEQLWVDVLRASMAHGLTPVSWTIDY 180
RLAAAAA                   APDGVVVDMASLGRLQGGGARRLAVSVEGRYVDAGGEQLWVDVLRASMAHGLTPVSWTIDY 60
                          ************************************************************

sp|Q4ADV8|CKX2_ORYSJ      LHLTVGGTLSNAGISGQAFRHGPQISNVLELDVITGVGEMVTCSKEKAPDLFDAVLGGLG 240
RLAAAAA                   LHLTVGGTLSNAGISGQAFRHGPQISNVLELDVITGVGEMVTCSKEKAPDLFDAVLGGLG 120
                          ************************************************************

sp|Q4ADV8|CKX2_ORYSJ      QFGVITRARIPLAPAPARARWVRFVYTTAAAMTADQERLIAVDRAGGAGAVGGLMDYVEG 300
RLAAAAA                   QFGVITRARIPLAPA--------------------------------------------- 135
                          ***************
```

Figure 1 multiple sequence alignment between domain and whole enzyme

1.4 Predicting the structures of domain and whole protein

As the NMR crystallographic structure of enzyme is not available at PDB, so the tertiary structure was generated on the basis of homology modeling concept (Bilal et al., 2013, Sungh et al., 2004, Sali et al., 1993).The models for whole protein and domain parts were generated using modeler 9.12 tool. The align2d.py, model-single.py, evaluate-model.py files were run on the python script by setting the template, target and the number of models to be generated. In this study 10 models were generated for both domain part and for whole enzyme sequence. The best model was selected on the basis of lowest DOPE score. The final generated models were visualized by using visualizing tools i.e., pymol, discovery studio and swiss-pdb viewer. Multiple visualiser tools were used to understand the structures more properly (Table 1, Table 2). The final models were shown in Figure 2 and Figure 3.

Table 1 Information found after visualizing the structures using Pymol tool

Analysis	Protein structure	Domain structure
Atom count	4233	1311
Formal charge sum	-6.0	0.0
Molecular surface area	52371.758A^0	16504.783 A^0
Solvent accessible surface area	23178.145 A^0	10648.261 A^0

Table 2 information found after visualizing the structures using Yasara tool

Analysis	Protein structure	Domain structure
VDW radius	85.337	34.261
Beta factor	109.7	111.3
Mass of object	55813.069g/mol	17316.863g/mol
Stability of the object	971.13kcal/mol	286.78kcal/mol

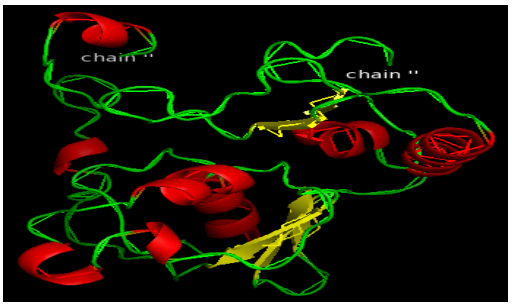

Figure 2 Model generated for whole enzyme

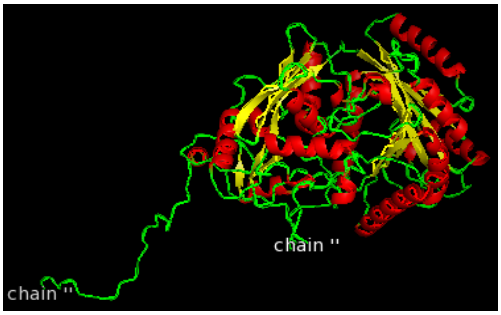

Figure 3 Model generated for domain

1.5 Model validation and optimization

Expasy server provides many online tools with the help of which the generated protein models can be evaluated and validated (Rambabu et al., 2012). The generated models were then validated by using Errat, Verify3d, Procheck saves server. The backbone confirmation of the protein was studied by using procheck saves server from where the residues lying in favoured region, allowed region and outlier regions can be found. When more residues found in favoured region and less residues found in outlier region, then it represents that protein is stable for further study (Table 3). Then the final Z-score of the models were obtained from prosa server, which showed the overall model quality (Table 4b). The models were then submitted to Anolea server which gave the QMEAN score and Z-score (Table 4a).

Table 3 The backbone confirmation of models found from rampage serve

Analysis	Protein structure	Domain structure
Number of residues in favoured region	517 (91.8%)	159 (88.3%)
Number of residues in allowed region	34 (6.0%)	15 (8.3%)
Number of residues in outlier region	12 (2.1%)	6 (3.3%)

Table 4 Model validation result

A Result found from anolea server		
Analysis	protein	Domain
QMEAN score	0.452	0.628
Z-score	-3.206	-1.483
B Result found from prosa server		
Analysis	protein	Domain
Z-score	-8.32	-4.21

1.6 Active site prediction

The active site or binding site is the region where the ligand binds to receptor. In this study the active site was predicted by using castp server. The active sites of both enzyme and domain are shown in Figure 4 and Figure 5. The active sites show the region where the ligand can bind or not. The binding energy dependent upon the type of residue to which it is binding i.e., hydrophobic or hydrophilic or polar or no-polar etc. the residues found at active sites are given in Figure 6 and Figure 7 respectively.

1.7 Protein-ligand interaction

The interaction of ligand (FAD) was checked with whole enzyme and only with the domain parts separately to find out the regions where the lignad binds properly. The interaction was studied by using autodock tool (Jasim et al., 2013, Seeliger et al., 2010). The pdb files of both enzyme, domain, ligand were uploaded one after another by setting the docking

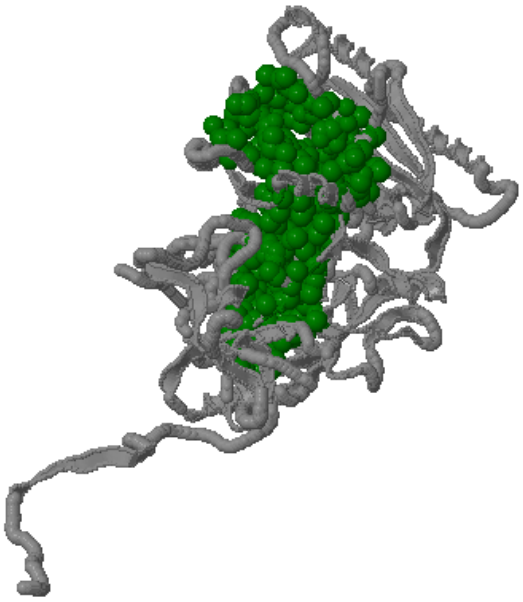

Figure 4 Binding site of whole enzyme

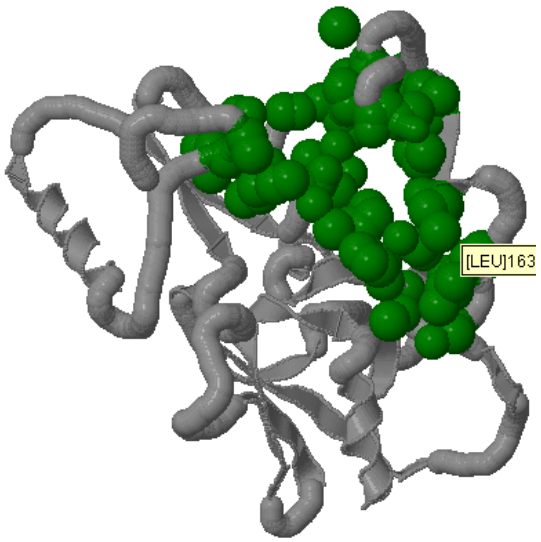

Figure 5 Binding site of domain

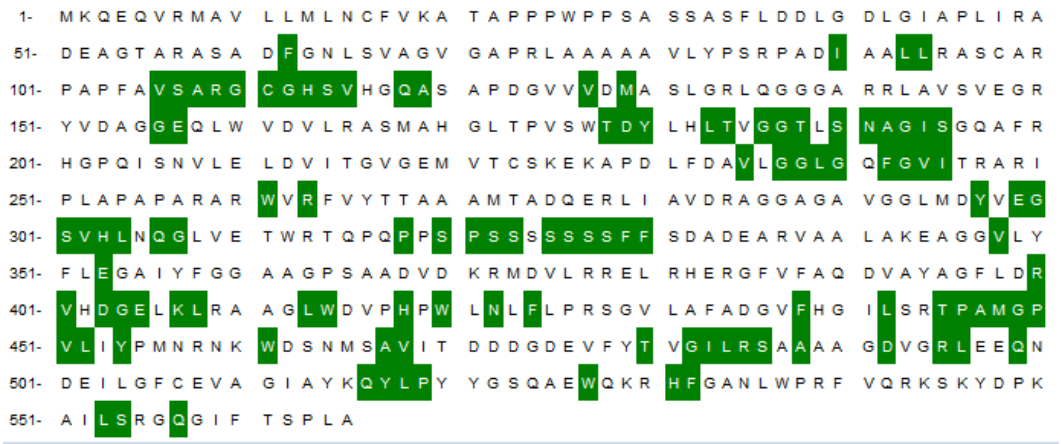

Figure 6 residues present at the active site of whole enzyme

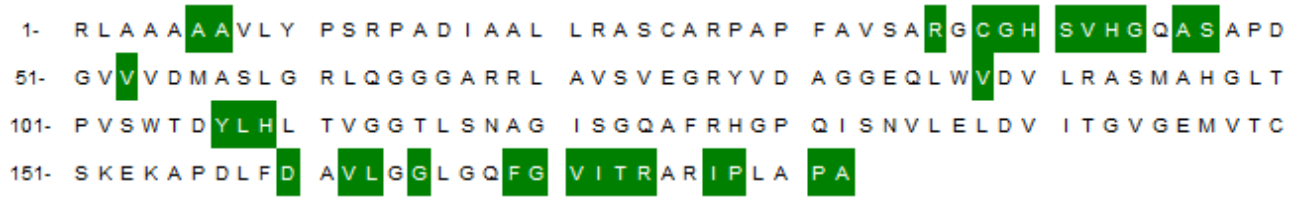

Figure 7 residues present at the active site of domain of the enzyme

parameters. The autodock and autogrid files were generally run on python script after which the interactions found by analyzing the dock.dlg files. Less the binding energy stronger is the interaction. However the binding energy is dependent on the length of amino acids. The interaction result is shown in Figure 8 and Figure 9 respectively.

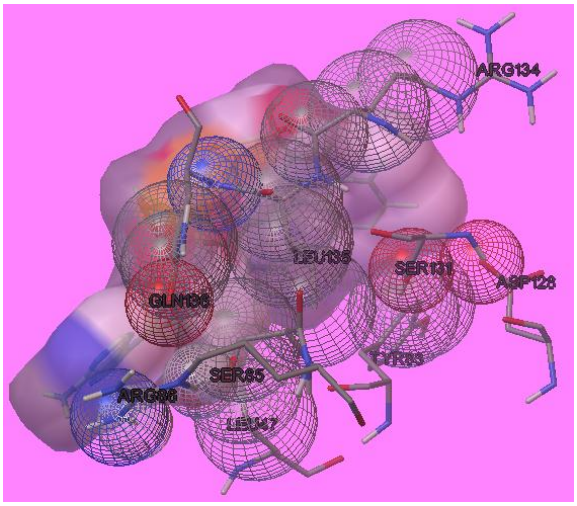

Figure 8 FAD binding to whole enzyme

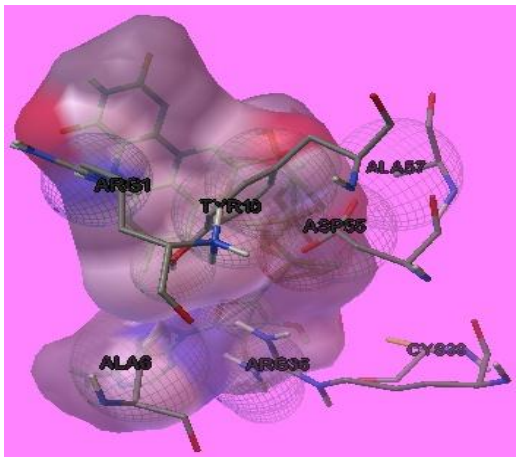

Figure 9 FAD binding to domain

2 Results

2.1 Retrieved sequence and finding the domain region

The amino acid sequence of enzyme was retrieved from uniprot with the uniprotID of Q4ADV8. The protein contains 565 numbers of amino acids. The domain region is found from 74-225. The information about the domain region was found from both uniprot and from protparam output.

2.2 Alignment between domain region and whole enzyme

The domain region was extracted from the whole protein sequence and subjected to clustalw2 for multiple sequence alignment. From this study the conserved residues among domain region and whole enzyme sequence were obtained. The match regions were denoted with *, mis-match regions with . and the gaps were denoted with -. Here the Figure 1 shows the common residues.

2.3 Homology modeling analysis

The homologous models for both enzyme and domains were generated. The helices are denoted red colours, helices with yellow colours and the loops are denoted with green colours respectively. The VDW radius, beta factor, mass of structures, stability of the objects were found from yasara tool. Then total number of atoms, formal charge sum, molecular surface area and solvent accessible surface area were found from pymol tool.

2.4 Model validation and optimization

The backbone confirmations of the structures were generated from rampage server. 91.8% of the residues lie in most favored region whereas 3.3% of the residues lie in outlier region. The QMEAN score and Z-scores were found from both anolea server and from prosa server.

2.5 Active site analysis result

The active sites were found from castp server. The balls with green color show the binding regions of whole protein and the domain region. Many pockets were found, but out of them the first volume was selected. The pockets of domain were mainly ALA, ARG, CYS, HIS, GLY, VAL, SER, TYU, LEU, PHE, ILE, THR, PRO and similarly the pockets of whole enzyme were PHE, ILE, LEU, VAL, SER, HIS, GLY, CYS, ALA, MET, GLU, THR, ASP, TYR, ASN, GLN, PRO, TRP, GLN respectively.

2.6 Protein-ligand interaction

The pdb files of protein, domain and the ligand were submitted and by adding polar hydrogens to protein and domain the interaction was studied. The binding energy which was found between the interactions of whole protein-ligand (FAD) was -11.23, with -14.81 intermolecular energy, 3.58 torsion energy and 13.66 internal energy. Similarly the interaction which was performed between the domain region and ligand (FAD) showed -10.91 binding energy, -14.49 intermolecular energy, 3.58 torsion energy and 14.04 internal energy respectively.

From the interaction it is found that the FAD is binding to ALA6, ARG1, AR536, TYR10, ASP55, ALA57, CYS38 at the domain part and in whole protein the domain is binding to ARG134, ASP128, SER131, TYR83, SER85, LEU135, ARG86, GLN136, LEU135, LEU47.

3 Discussions

From the above study it is found that the model of whole enzyme and the domain are suitable for future analysis as most of the residues lie in favored region. The main objective of this study was to find the stronger interaction of ligand (FAD) with the domain part of enzyme and with the whole sequence, which showed that the binding interaction is stronger between the domain part and FAD than in comparison to the interaction between FAD and whole amino acid sequence

References

Ashikari et al., 2005, Cytokinin Oxidase Regulates Rice Grain Production, science journal, vol.309, No.5735,741-745

Bilal et al., 2013, Generation of a 3D model for human cereblon using comparative modeling, journal of bioinformatics and sequence analysis, vol. 5, No.1, 10-15

Gao et al., 2014, CYTOKININ OXIDASE/DEHYDROGENASE4 Integrates Cytokinin and Auxin Signaling to Control Rice Crown Root Formation, plant physiology journal, vol.165, 1035-1046

Jasim et al., 2013, Protein-Ligand interaction studies on 2, 4, 6-trisubstituted triazine derivatives as anti-malarial DHFR agents using AutoDock, IOSR Journal of Applied Physics, vol.4, No.5, 2278-4861

Li et al., 2013, Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression, plant biology, vol. 10, No.8, 3167-72.

Mameaux et al., 2011, Molecular, phylogenetic and comparative genomic analysis of the *cytokinin oxidase/dehydrogenase* gene family in the Poaceae, plant biotechnology journal, vol.10, No.1, 67-82

Rambabu et al., 2012, Computational analysis and function prediction of a hypothetical protein 1RW0, International Journal of Computational Bioinformatics and In Silico Modeling, vol.1, No.5, 58-62

Sali et al., 1993, Comparative Protein Modelling by Satisfaction of Spatial Restraints, Journal of Molecular Biology, vol.234, No.3, 779-815
http://dx.doi.org/10.1006/jmbi.1993.1626

Seeliger et al., 2010, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, J Comput Aided Mol Des, vol.4, 417-422
http://dx.doi.org/10.1007/s10822-010-9352-6

Sungh et al., 2004, Comparative modeling and analysis of 3-D structure of Hsp 70, in Cancer irroratus, Biological Forum — An International Journal, vol. 1, No.2, 1-4

Thompson et al., 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Research journal, vol. 22, 4673-468