

Research Report

Open Access

Comparative structural analysis of cytokinin dehydrogenase enzymes of *O.sativa*, *A. thaliana* and *Zea mays* leading to predict the best enzyme

Pranati Swain[✉], Lambodar Behera

Central rice research institute, cuttack, odisha, India

✉ Corresponding author email: rosylora20@gmail.com

Computational Molecular Biology, 2014, Vol.4, No.14 doi: 10.5376/cmb.2014.04.0014

Received: 03 Dec., 2014

Accepted: 26 Dec., 2014

Published: 30 Dec., 2014

© 2014 Swain and Behera, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Swain and Behera, 2014, Comparative structural analysis of cytokinin dehydrogenase enzymes of *O.sativa*, *A. thaliana* and *Zea mays* leading to predict the best enzyme, Computational Molecular Biology, Vol.4, No.14, 1-9 (doi: [10.5376/cmb.2014.04.0014](https://doi.org/10.5376/cmb.2014.04.0014))

Abstract The cytokinin dehydrogenase enzymes of rice, wheat and maize were selected for this study. The enzymes of these crops show the same function i.e. leading to high grain production. The CKX1, CKX2, CKX3, CKX4, CKX5, CKX6, CKX7 and CKX8 enzymes of these three crops vary from each other on the basis of their physico-chemical analysis. A brief study has been done on all these enzymes showing the similarities, domains, secondary structures, homology models, backbone confirmations, best generated models and functions. From the study we got that Q4ADV8 of *O.sativa* has the best secondary structure with strong helix, turns and sheets along with highest molecular weight. A2XVN3 of *O.sativa* has the longest domain region, Q9FUJ1 of *A.thaliana* has highest numbers of total negatively charged residues (ASP+GLU) and Q9LY71 of *A.thaliana* has highest numbers of total positively charged residues (ARG+LYS). The Q9LTS3 of *A.thaliana* has highest VDW radius from geometric center. Q8LNV6 is the best model depending upon the Z-score calculated by ANOLEA. From the above study it is concluded that the enzymes of *O.sativa* and *A.thaliana* have more strong physico-chemical characteristics than in comparison to the *zea mays*.

Keywords Cytokinin dehydrogenase enzymes; Physico-chemical analysis; Domain region prediction; Predicting the best enzyme

Introduction

Due to the growth of population the demands for more grain production are increasing day by day. The main crops harvested in world are rice, maize and wheat. These are the main source of food. Human beings absorb many proteins from these crops. These crops are blessed with such miraculous proteins which help a lot in high grain production in order to fulfill the demands. Here we have made a brief study about the cytokinin dehydrogenase enzymes of rice, maize and wheat leading to high grain production. Cytokinin dehydrogenase belongs to oxidoreductase family and catalyzes the chemical reaction. Cytokinin dehydrogenase is otherwise known as N⁶-dimethylallylamine: (acceptor) oxidoreductase, 6-N-dimethylallylamine: acceptor oxidoreductase, OsCKX2, CKX, and cytokinin oxidase/dehydrogenase. Cytokinin dehydrogenase is helpful in the degradation of cytokinin isopentenyladenine, zeatin, and their ribosides (Chmülling et al., 2003). Gnl1a gene of rice possesses cytokinin oxidase/dehydrogenase

(OsCKX2), an enzyme which degrades the phytohormone cytokinin (Ashikari et al., 2005). The Gnl1a/OsCKX2 of *Oryza sativa* L, grain number 1a/Cytokinin oxidase 2 gene, encodes a cytokinin oxidase, which acts as a major quantitative trait locus contributing to grain number improvement in rice breeding practice (Lia et al., 2012). In wheat the CKX2 expression is activated by the IKU transcription factor WRKY10 directly and promotes endosperm growth (Li et al., 2013). In wheat the growth control of endosperm by CKX2 integrates genetic and epigenetic regulations. In angiosperms, cytokines are highly active in endosperm, IKU effectors coordinate environmental and physiological factors, resulting in modulation of seed size. Higher activity of the cks3 cks5 in wheat increase in seed yield, highlighting the relevance of sink strength as a yield factor. CKX3 and CKX5 regulate the activity of the reproductive meristems of *Arabidopsis thaliana* (Batrina et al., 2011). In this study we have considered the CKX1 of *zea mays* (Maize), CKX2 of *Oryza Sativa*

(*japonica*), *Arabidopsis thaliana* (*Mouse-ear cress*), CKX3 of *Oryza Sativa* (*japonica*), *Arabidopsis thaliana*(*Mouse-ear cress*), CKX4 of *Oryza Sativa* (*japonica*), CKX5 of *Arabidopsis thaliana*(*Mouse-ear cress*), CKX6 of *Arabidopsis thaliana* (*Mouse-ear cress*) and CKX7 of *Arabidopsis thaliana* (*Mouse-ear cress*) and CKX8 of *Oryza Sativa*(*indica*).The main objective of the study is to find out their physic-chemical properties, similarities and differences among all the enzymes. However the function of all the eznzymes is same, but their properties are explained below.

Material and Methods

The cytokinin dehydrogenase enzymes of rice, wheat and maize were selected in this study due to its role in high grain production in crops.

Sequence retrieval

The cytokinin dehydrogenase enzymes of *O.sativa* (*japonica*), *A.thaliana* (*mouse-ear cress*) and *zea mays* (*maize*) were used for this study. Their amino acid sequences were retrived from uniprot (<http://www.uniprot.org/>), which is given in Table 1.

Table 1 The informations of all enzymes which were collected from Uniprot

species	UniprotID	Function	Length of amino acids
<i>o.sativa</i> (<i>indica</i>)	A2XNV3	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	532
<i>o.sativa</i> (<i>japonica</i>)	Q4ADV8	Catalyzes the oxidation of cytokinin, modulates the number of reproductive organs by regulating the cytokinin accumulation in inflorescence meristem, major QTL involved in high grain production	565
<i>o.sativa</i> (<i>japonica</i>)	Q5JLP4	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	529
<i>o.sativa</i> (<i>japonica</i>)	Q8LNV6	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	527
<i>A.thaliana</i> (<i>mouse-ear cress</i>)	Q9FUJ1	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	524
<i>A.thaliana</i> (<i>mouse-ear cress</i>)	Q9FUJ3	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	501
<i>A.thaliana</i> (<i>mouse-ear cress</i>)	Q67YU0	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	540
<i>Zea mays</i>	Q9T0N8	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group. Cleaves zeatin, isopentenyladenine, isopentenyladenosine, zeatinriboside and cis-zeatin, but not dihydrozeatin, kinetin and benzylaminopurine.	534
<i>A.thaliana</i> (<i>mouse-ear cress</i>)	Q9LY71	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenylgrou	533
<i>A.thaliana</i> (<i>mouse-ear cress</i>)	Q9LTS3	Catalyzes the oxidation of cytokinins, a family of N(6)-substituted adenine derivatives that are plant hormones, where the substituent is an isopentenyl group	523

Sequence alignment

The cytokinin dehydrogenase enzymes of all the plants were subjected for multiple sequence alignment using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) analysis (Thompson et al.,

1994). The alignment file was saved in clustal format and conserved residues were identified successfully. A phylogenetic tree was generated. The conserved residues of all the enzymes are shown in Figure 1.

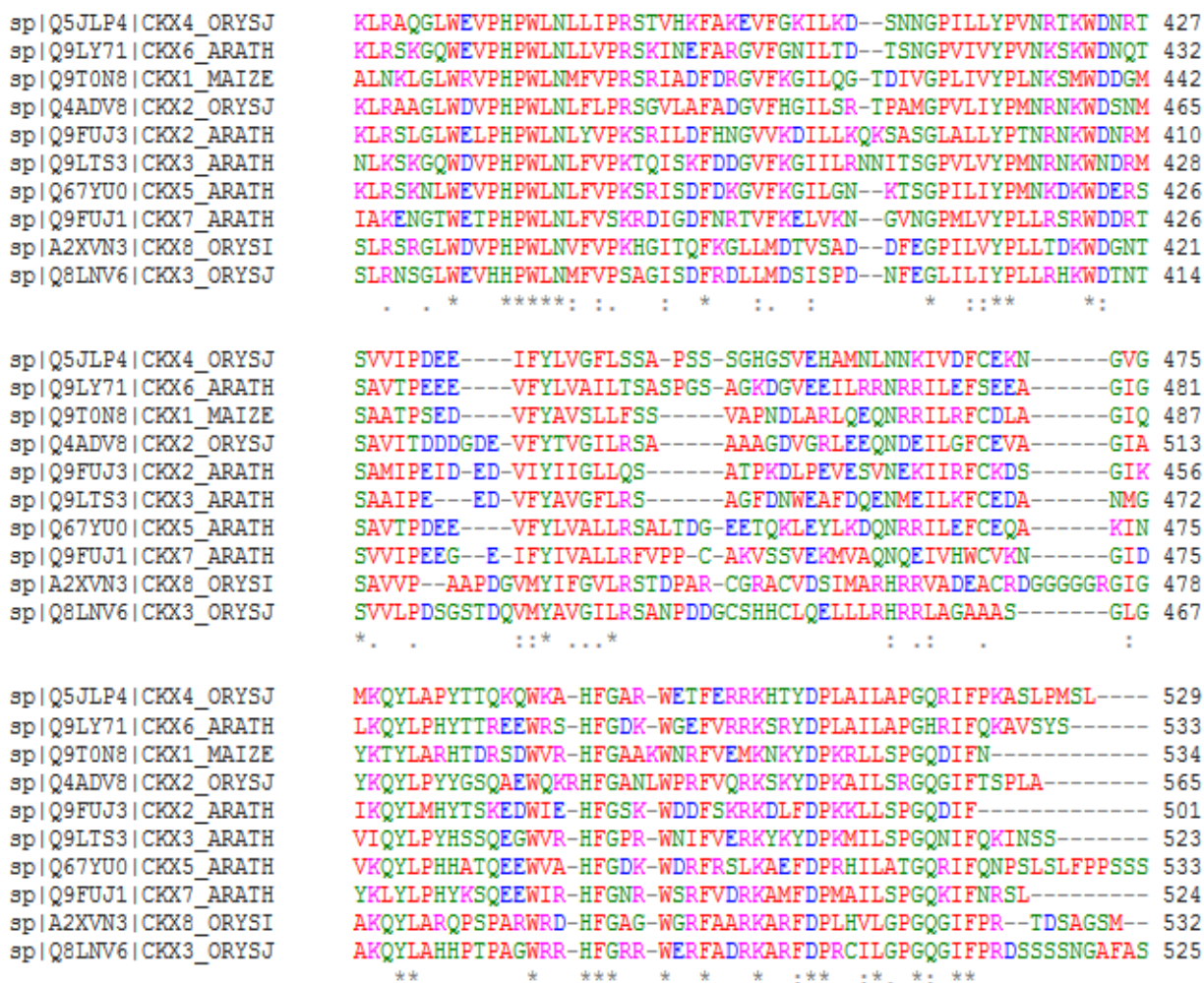


Figure 1 the multiple sequence alignment result of all the enzyme

Physico-Chemical Characterization of all the enzymes

The amino acid sequences of all the enzymes were subjected to ProtParamtool (<http://web.expasy.org/protparam/>) for further analysis, from where we got the physical and chemical properties of all the enzymes including molecular weight, theoretical PI, extinction coefficients, instability index, aliphatic index, grand average of hydrophobicity, total number of positively and negatively charged

residues [9]. Detail information is shown in Table 2, Table 3, Table 4.

Secondary structure prediction of enzymes

The CFFSP server (<http://www.biogem.org/tool/chou-fasman/>) was used for secondary structure analysis where we submitted the amino acid sequences as an input and the result showed the details about helices, sheets and turns respectively (David et al., 1999). The detail information is shown in Table 5 and Figure 2-Figure 11.

Table 2 Physico-chemical characterization of enzymes

Enzymes	Molecular weight	Theoretical PI	Total no. of atoms	Extinction assuming all pair of CYS residues from CYS	Coefficient	Extinction assuming all CYS residues are reduced	Coefficient	Instability index	Aliphatic index
A2XNV3	10056.6	9.05	1917	10095		9970		50.27	109.16
Q4ADV8	60021.1	6.15	8408	77600		77350		37.76	87.98
Q5JLP4	58426.8	7.37	8219	69120		68870		37.30	91.98
Q8LNV6	58428.7	8.19	8194	68215		67840		49.91	93.06
Q9FUJ1	57975.5	5.02	8091	84840		84340		39.85	86.64
Q9FUJ3	55583.0	7.17	7894	68090		67840		39.14	100.54
Q67YU0	60423.7	6.00	8493	81610		81361		35.84	90.02
Q9T0N8	57228.8	5.94	8010	81945		81820		35.44	89.63
Q9LY71	59999.7	8.94	8487	82070		81820		43.83	95.48
Q9LTS3	59422.7	6.33	8324	95800		95800		42.57	86.23

Table 3 Information about the positively and negatively charged residues of all enzymes

Enzymes	GRAVY	Total number of negatively charged residues	Total number of positively charged residues
A2XNV3	0.359	57	49
Q4ADV8	0.026	57	52
Q5JLP4	-0.095	52	52
Q8LNV6	-0.035	52	45
Q9FUJ1	-0.145	69	50
Q9FUJ3	-0.040	53	35
Q67YU0	-0.195	65	55
Q9T0N8	0.059	50	45
Q9LY71	-0.150	51	58
Q9LTS3	-0.228	59	55

Table 4 Domains of proteins

Enzymes	Start-end positions	Domain names
A2XNV3	51-238	FAD binding PCMH type
Q4ADV8	74-255	FAD binding PCMH type
Q5JLP4	63-244	FAD binding PCMH type
Q8LNV6	52-231	FAD binding PCMH type
Q9FUJ1	58-238	FAD binding PCMH type
Q9FUJ3	53-226	FAD binding PCMH type
Q67YU0	63-241	FAD binding PCMH type
Q9T0N8	65-245	FAD binding PCMH type
Q9LY71	68-248	FAD binding PCMH type
Q9LTS3	66-243	FAD binding PCMH type

Table 5 The secondary structural analysis of all enzymes

Enzymes	Helices	Sheets	Turns
A2XNV3	375= 70.5%	212= 39.8%	62= 11.7%
Q4ADV8	424= 75.0%	181= 32.0%	63= 11.2%
Q5JLP4	382= 72.2%	331= 62.6%	62= 11.7%
Q8LNV6	328= 62.2%	237= 45.0%	59= 11.2%
Q9FUJ1	350= 66.8%	354= 63.7%	64= 12.2%
Q9FUJ3	348= 69.5%	238= 47.5%	72= 14.4%
Q67YU0	394= 73.0%	224= 41.5%	77= 14.3%
Q9T0N8	366= 68.5%	185= 34.5%	61=11.4%
Q9LY71	399= 74.9%	366= 68.7%	63= 11.8%
Q9LTS3	364= 69.6%	252= 48.2%	72= 13.8%

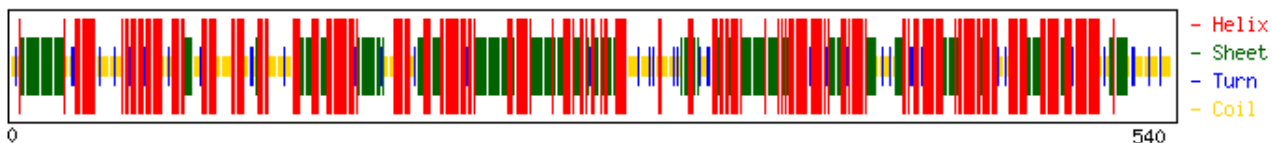


Figure 2 Q67YU0-CKX5 (A.thaliana)

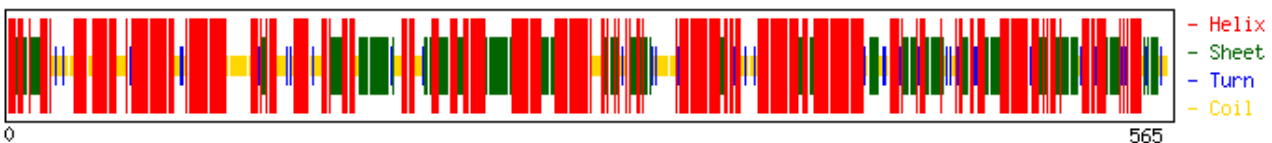


Figure 3 Q4ADV8-CKX2(O.sativa)

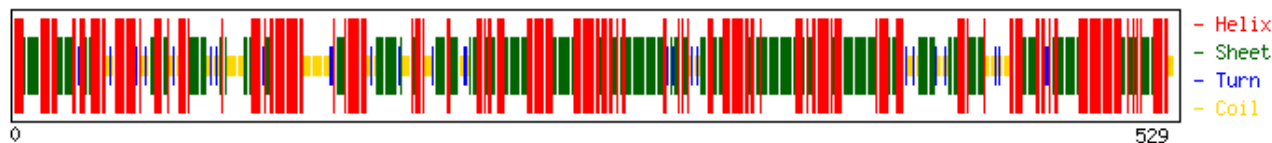


Figure 4 Q5JLP4-CKX4(O.sativa)

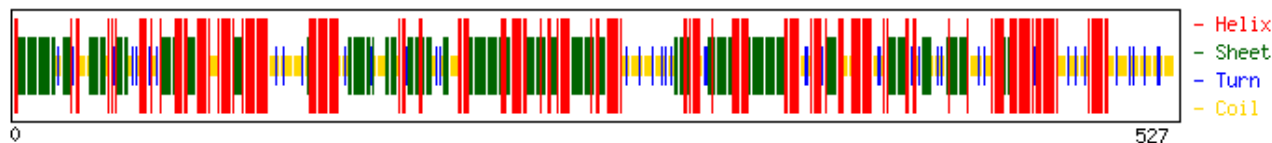


Figure 5 Q8LNV6-CKX3(O.sativa)

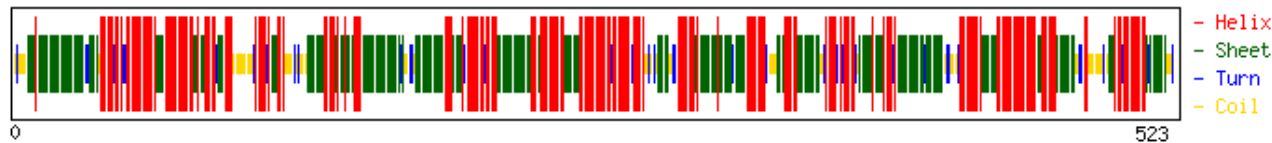


Figure 6 Q9LTS3-CKX3 (A.thaliana)

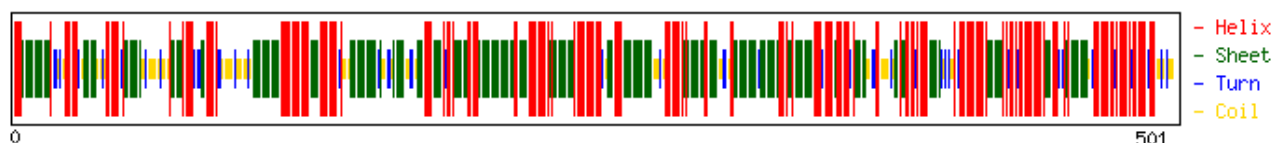


Figure 7 Q9FUJ3-CKX2 (A.thaliana)

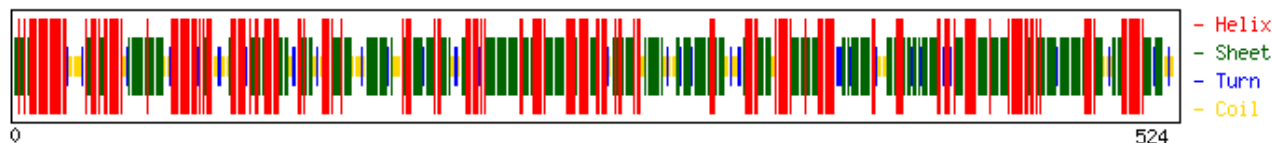


Figure 8 Q9FUJ1-CKX7(A.thaliana)

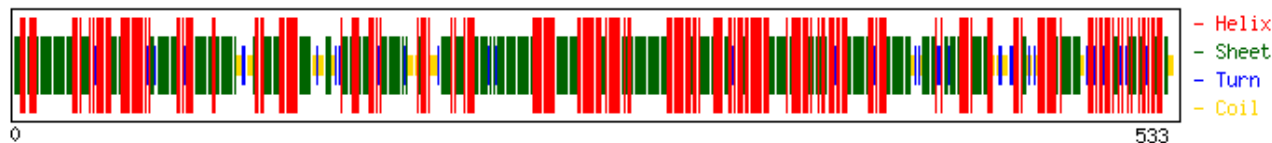


Figure 9 Q9LY71 (A.thaliana)

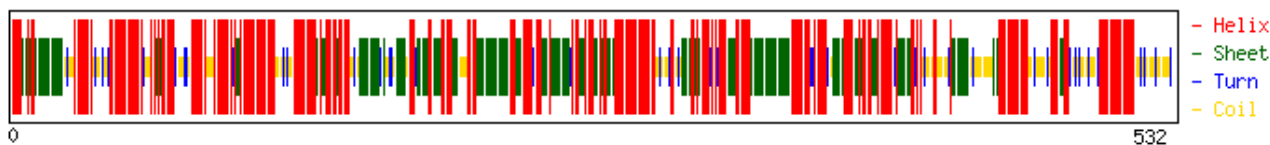


Figure 10 A2XVN3-CKX8(O.sativa)

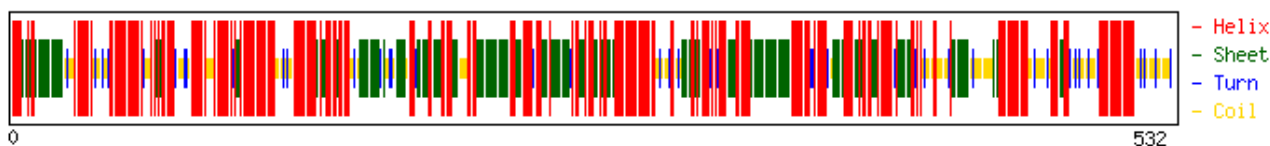


Figure 11 Q9T0N8-CKX1 (Zea mays)

Structural alignment

As the NMR crystallographic structures of these enzymes are not available at PDB, so the tertiary structures of all the enzymes of plants were designed by using phyre2 server which generate reliable protein models when other widely used methods such as PSI-BLAST cannot. Phyre2 server (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) generate the model using the principles and techniques of homology modeling. The target can be modelled with reasonable accuracy on a very distantly related sequence of template. The phyre2 server uses a profile-profile alignment algorithm based on each proteins position-specific scoring matrix. Phyre2 server include protein structure prediction, function prediction, domain prediction, domain boundary prediction, evolutionary classification of proteins, guiding site-directed mutagenesis and solving protein crystal structures by molecular replacement (Christie et al., 2012, Bilal et al., 2013, Singh et al., 2009). The homology models of enzymes are shown in Figure 12-Figure 21 and the detail information after visualization is shown in Table 6.

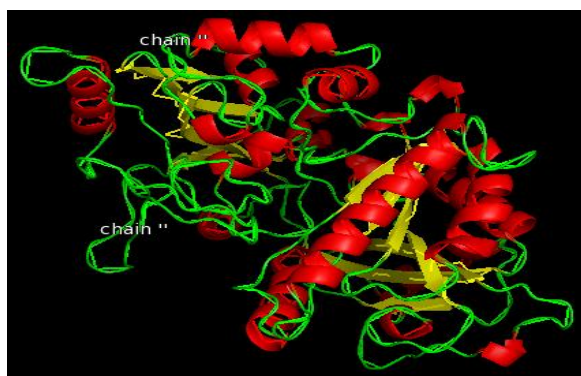


Figure 12 A2XVN3-CKX8(O.sativa)

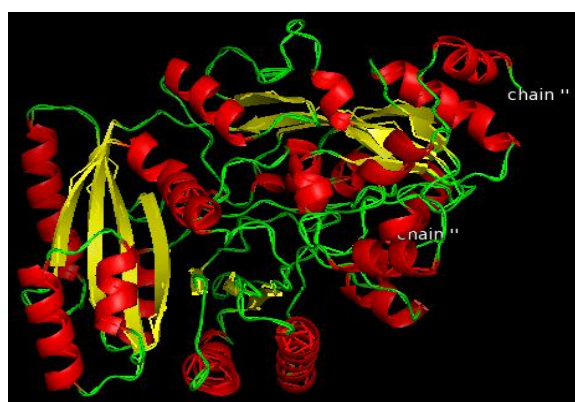


Figure 13 Q4ADV8-CKX2(O.sativa)



Figure 14 Q5JLP4-CKX4(O.sativa)

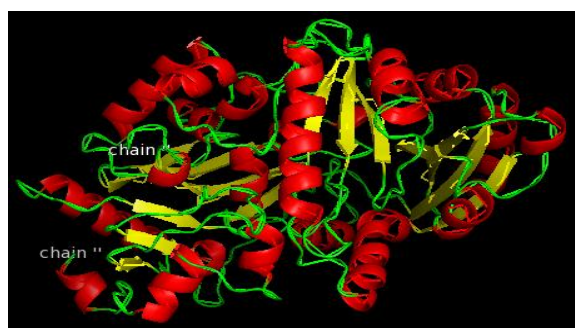


Figure 15 Q8LNV6-CKX3(O.sativa)

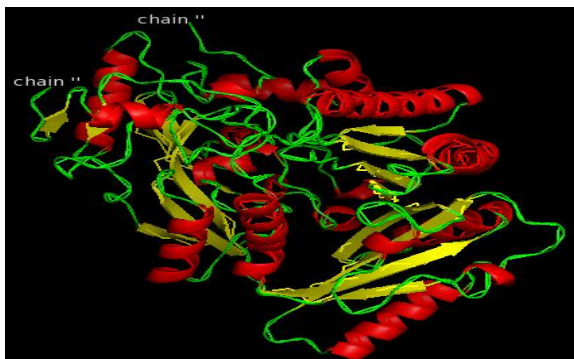


Figure 16 Q9FUJ1-CKX7(A.thaliana)



Figure 19 Q9LY71 (A.thaliana)

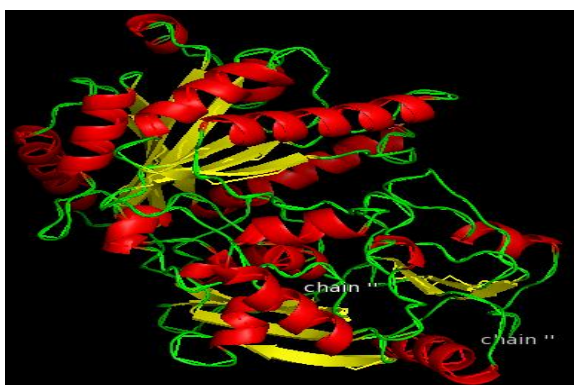


Figure 17 Q9FUJ3-CKX2 (A.thaliana)

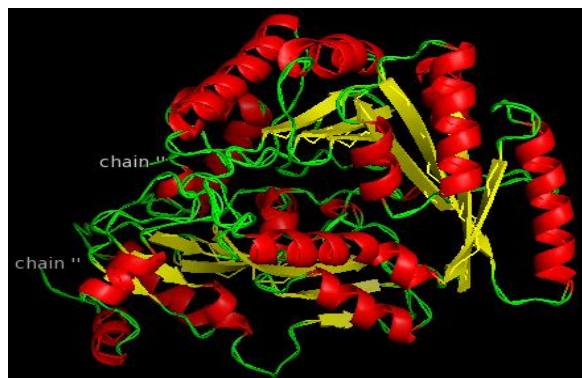


Figure 20 Q9T0N8-CKX1 (Zea mays)

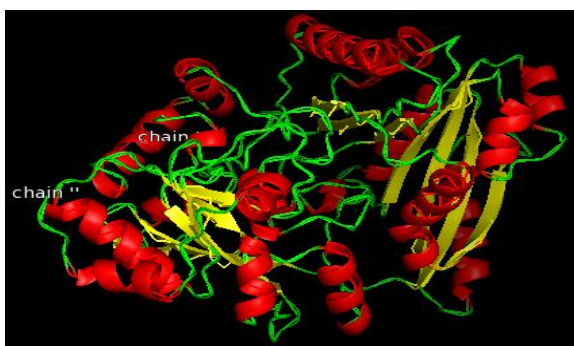


Figure 18 Q9LTS3-CKX3 (A.thaliana)

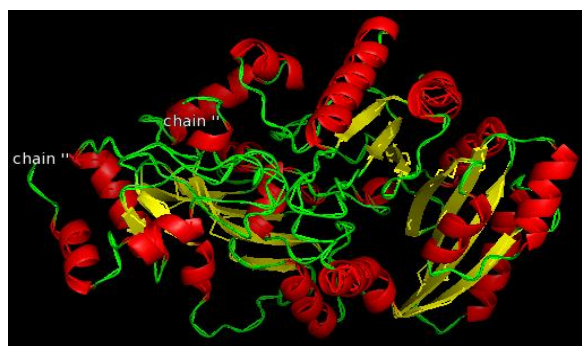


Figure 21 Q67YU0-CKX5 (A.thaliana)

Table 6 properties of all enzymes which were obtained by using pymol visualiser

Models	Atom count	Formal charge sum	Molecular surface area	Solvent accessible surface area
A2XNV3	3722	-5.0	464221.184 A ⁰	18886.068 A ⁰
Q4ADV8	3798	-4.0	47784.027 A ⁰	18796.070 A ⁰
Q5JLP4	3791	1.0	47014.855 A ⁰	19352.930 A ⁰
Q8LNV6	3904	-7.0	48019.445 A ⁰	18558.955 A ⁰
Q9FUJ1	3853	-15.0	49141.406 A ⁰	17943.436 A ⁰
Q9FUJ3	3680	-3.0	45956.980 A ⁰	18055.246 A ⁰
Q67YU0	3922	-11.0	49096.207 A ⁰	19347.986 A ⁰
Q9T0N8	3826	-5.0	48934.703 A ⁰	17650.164 A ⁰
Q9LY71	3869	1.0	48135.730 A ⁰	19421.123 A ⁰
Q9LTS3	3902	-7.0	48695.211 A ⁰	18992.123 A ⁰

Model validation and optimization

The final tertiary structures of the enzymes were visualized using PyMol visualization tool. Then by using PyMolvisualiser we got the number of atoms present in all the enzymes, formal charge sum, and partial charge sum, molecular and solvent accessible surface area. The backbone confirmation of all the enzymes were checked

using Rampage server (mordred.bioc.cam.ac.uk/~rapper/rampage.php) from where we got the residues lying in allowed, favoured and outlier regions [13]. Then the models were subjected to ANOLEA-swiss model from where we got the Z-score, QMEAN score. The model with least Z-score is generally selected as the best model. The detail information is shown in Table 7.

Table 7 Backbone confirmation of all enzymes

Models	Residues in favoured region	Residues in allowed region	Residues in outlier region	QMEAN score	Z-score
A2XNV3	92.5%	4.6%	2.9%	0.628	-1.673
Q4ADV8	94.9%	3.0%	2.2%	0.718	-0.57
Q5JLP4	92.1%	5.8%	2.1%	0.668	-1.188
Q8LNV6	91.1%	5.8%	3.0%	0.564	-2.445
Q9FUJ1	94.3%	5.1%	0.6%	0.747	-0.256
Q9FUJ3	94.0%	3.8%	2.1%	0.66	-1.284
Q67YU0	93.3%	4.9%	1.8%	0.704	-0.774
Q9T0N8	97.0%	2.4%	0.6%	0.784	0.216
Q9LY71	93.8%	4.6%	1.7%	0.69	-0.927
Q9LTS3	93.2%	4.6%	2.3%	0.676	-1.096

Result and Discussion

Sequence retrieval analysis

The amino acid sequences were successfully retrieved from uniprot, which showed the following informations.

Multiple sequence alignment result

From the MSA analysis we got that there are many conserved residues among all the enzymes i.e, D, F, G, V, P, L, H, S, Q, W, Y, N, C, T, R, A, E, V, P, K, I. the similarities are denoted with *, dissimilarities with . and gaps with -.

Physico-chemical analysis result

From the physico-chemical analysis it is found that the A2XVN3 enzyme of *O.sativa* has longest domain region whereas Q9FUJ3 of *A.thaliana* has smallest domain region. The Q67YU0 of *A.thaliana* has highest molecular weight, A2XVN3 enzyme of *O.sativa* has highest theoretical PI and aliphatic index, Q4ADV8 enzyme of *O.sativa* has highest total number of atoms.

Secondary structure prediction result of all enzymes

The secondary structures of all the enzymes were predicted using CFFSP server from where we could

got the percentage of alpha, helix, turns. The helix is denoted with red colours, sheet is denoted with green colours, turn is denoted with blue colours. The secondary structures of all the enzymes are given from Figure 2-Figure 11.

Homology modeling results

The models which were generated by using phyre2 server were visualized by using PyMol visualization tool, which showed the following result.

Backbone confirmation result of enzymes

The final generated models of all the enzymes were submitted to RAMPAGE server from where we got the residues lying in allowed, favoured and outlier regions. The detail information is given below.

Discussions

From the above study we got the functions of all the enzymes, the most common residues among them, from the physico-chemical analysis we got that A2XVN3 of *O.sativa* has the most longest domain region. The best model which was selected on the basis of Z-score was Q8LNV6 of *O.sativa*. All the enzymes has most common domain i.e, FAD binding PCMH type. Depending upon the back-bome confirmation of enzymes the best model was Q9T0N8

of *zea mays*. The function of all the enzymes are same leading to high grain production, however they are having many structural and chemical differences which is represented in this study.

References

- Ashikari et al., 2005, Cytokinin Oxidase Regulates Rice Grain Production. *Science journal*, vol.309, No.5735, 741-745
- Batrina et al., 2011, Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*, *The Plant Cell journal*, vol.23, No.1, 69-80
- Bilal et al., 2013, Generation of a 3D model for human cereblon using comparative modeling. *Journal of Bioinformatics and Sequence Analysis* Vol. 5(1), pp. 10-15
- Chmilling et al., 2003, Structure and function of cytokinin oxidase/dehydrogenase genes of maize, rice, *Arabidopsis* and other species. *Journal of Plant Research*. Vol.116, No.3,241-52
- Christie et al., 2012, phylogenetic analysis and Homology modelling of laccase gene of *Fusarium Oxysporum* Icc2. *International Journal of Pharma Sciences and Research (IJPSR)*, Vol 3, No 12
- David et al., 1999, Protein Secondary Structure Prediction Based on Position-specific Scoring Matrices, *Journal of molecular biology*, vol.292, 195-202
- Lia et al., 2012, Rice zinc finger protein DST enhances grain production through controlling *Gn1a/OsCKX2* expression
www.pnas.org/lookup/suppl/doi:10.1073/pnas.1300359110/DCSupplemental
- Li et al., 2013, Integration of epigenetic and genetic controls of seed size by cytokinin in *Arabidopsis*. *Plant biology journal*, vol.110, No.38, 15479-15484
- Singh et al., 2009, Comparative modeling and analysis of 3-D structure of Hsp 70, in *Cancer irroratus* *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*, Vol. 22, No. 22, 4673-4680
<http://dx.doi.org/10.1093/nar/22.22.4673>