International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2018; 6(2): 1903-1909

Received: 20-01-2018
Accepted: 23-02-2018

Neha Gupta
a) Department of Chemistry and Biochemistry, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India
b) Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India

Gagan Rani
a) Department of Chemistry and Biochemistry, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India
b) Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India

Neeru Singh Redhu
a) Department of Chemistry and Biochemistry, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India
b) Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India

Sudhir Kumar
a) Department of Chemistry and Biochemistry, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India
b) Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India

Correspondence
Gagan Rani
a) Department of Chemistry and Biochemistry, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India
b) Department of Molecular Biology, Biotechnology and Bioinformatics, Chaudhary Charan Singh Haryana Agricultural University, Hisar, Haryana, India

In silico characterization and modelling of *Drosophila melanogaster* chitin synthase

Neha Gupta, Gagan Rani, Neeru Singh Redhu and Sudhir Kumar

Abstract
Chitin synthase is known to be an important enzyme that catalyzes last step of chitin biosynthetic pathway. It causes linear polymerization of chitin from activated UDP-N-acetylglucosamine monomers in manner joined together by β-(1, 4)-linked. Chitin synthases are thus critical enzymes for synthesis of chitin and as well as for growth and development of insects. In this study, physiochemical properties and modeling of chitin synthase enzyme of *Drosophilla melanogaster* encoded by CS2 gene was analyzed using *in silico* approach. Various isoforms (isoforms C and D) of chitin synthase from CS2 gene were used in this study. Physiochemical properties such as molecular weight, theoretical isoelectric point, extinction coefficient, aliphatic index, instability index, total number of negatively and positively charged residues and grand average of hydrophathicity were computed. Along with these physiochemical properties cellular localilization, no. of transmembrane helices, other proteins with which these isoforms interact were also depicted using various tools.

Keywords: Chitin, chitin synthase 2, *In silico,* CS2

Introduction
After cellulose, chitin is earth's second most abundant organic compound and is synthesized by a broad variety of organisms of different taxonomic groups. Chitin is found not only in arthropods including insects, arachnids and crustaceans but also in lower invertebrates such as sponges, coelenterates, nematodes and molluscs (Merzendorfer, 2011) [3]. Chitin is a linear polymer of N-acetyl- β-D-glucosamine and is a major component of insect cuticle and peritrophic matrices. Chitin is found in the exo- and endocuticle or in the newly secreted, unsclerotized procuticle but not in the epicuticle, the outermost part of the integument. Insect's cuticles form an exoskeleton that exhibits only a limited capacity to keep pace with body growth because it is a more or less rigid structure due to the presence of chitin and sclerotized proteins. To allow growth and development, insects are therefore periodically forced to replace their old cutical with a new one during molting (ecdysis). Chitin functions as light but mechanically strong scaffold material and is always associated with cuticle proteins that mainly determine the mechanical properties of the cuticle (Merzendorfer and Zimoch, 2003) [1]. Chitin is also an integral part of insect peritrophic matrices which functions as a permeability barrier between the food bolus and the midgut epithelium, enhance digestive processes and protect the brush border from mechanical disruption as well as from attack by toxins and pathogens (Tellam, 1996). Thus, insect growth and development is strictly dependent on the capability to remodel chitinous structures. Therefore, insects consistently synthesize and degrade chitin in a highly controlled manner to allow ecdysis and regeneration of the peritrophic matrices. Formation of the different forms of chitin is catalyzed by chitin synthase (CHS) (EC 2.4.1.16), a highly conserved enzyme found in every chitin synthesizing organism. CHS enzymes are classified in CAZy databse as belonging to the GT-2 family. CHS enzymes utilizes UDP-N-acetylglucosamine (UDP-GlcNAc) as the activated sugar donor to form the chitin polymer. Thus chtin synthases represents an attractive target site for combating insects pests as insect growth and development are strictly dependent on precisely tuned chitin biosynthesis and this pathway is absent in humans and other vertebrates.

Methodology
Sequence retrieval
The amino acid sequences of various isoforms of chitin synthase 2 (isoform C and D) were retrieved from NCBI having the accession number NP_01137997.2 and NP_524209.3,

respectively. The characterization of various isoforms was done by using bioinformatics tools *in silico.*

Characterization of target sequence

The physiochemical property of the protein was determined by using Protparam tool such as molecular weight, theoretical pI, total number of negatively and positively charged amino acids, amino acid composition of the protein, extinction coefficient, aliphatic index and GRAVY index. The subcellular localization of the protein was found out using Cello v2.5. The transmembrane helices were predicted using TMHMM tool of Expasy. SignalP 4.1 server was used to check if the protein is a signal peptide or not. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database was used to identify proteins interacting with the isoforms of the chitin synthase 2. Secondary structure prediction was done using SOPMA. *Ab initio* modelling was carried out using phyre 2 server.

Results and Discussion

Protparam

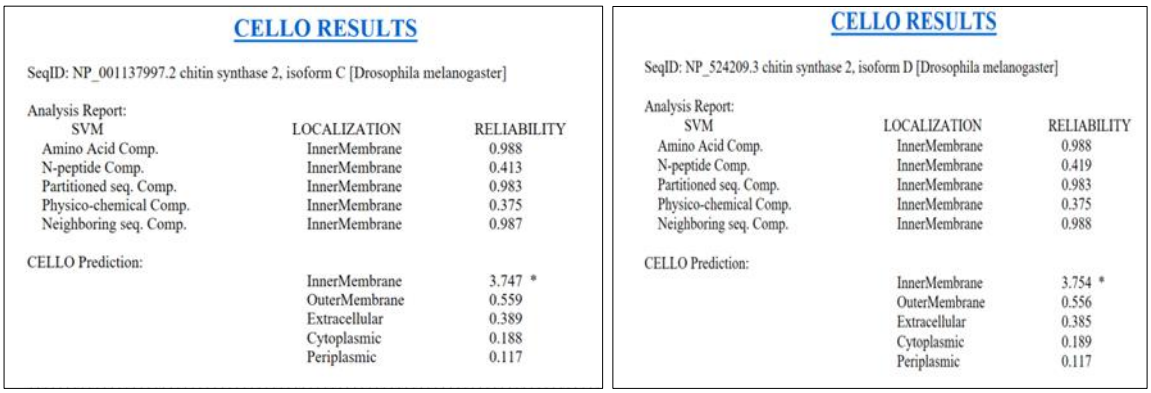
The results of Protparam tool show that the molecular weight of both the isoforms of chitin synthase 2 is different. Theoretical pI of the two isoforms of chitin synthase 2 is same with a value of 7.11. Total no. of negatively charged (Asp+Glu) residues in the two isoforms of chitin synthase 2 is 146 and total number of positively charged (Lys+Arg) residues is 145 leaving them with total charge of -1.Extinction coefficient (in M^{-1} cm^{-1} & at 280 nm measured in water) for the isoforms of chitin synthase 2 protein assuming all pairs of Cys residues form cystines is 256385 and under same conditions extinction coefficient assuming all Cys residues are reduced was found to be 254510 for both the isoforms of chitin synthase 2. Extinction coefficient can be used to separate the protein from the solution. The instability index (II) determines the stability of the protein in a test tube and is computed to be 40.56 for isoforms of chitin synthase 2. Since instability index for both the isoforms of protein is more than 40, it indicates that both the isoforms of the protein may be unstable. A value 98.37 for aliphatic index for the isoforms of chitin synthase 2 indicates the relative volume of a protein that is occupied by aliphatic side chains, which in turn contributes to the increased thermo stability of isoforms of chitin synthase. Positive Grand average of hydropathicity (GRAVY) index of chitin synthase 2 isoforms indicates that the isoforms are in hydrophobic in nature.

Table 1: Various physiochemical properties by PROTPARAM

Protein name		Molecular weight (dalton)	Theortical pi	Total charge	Instability index	Aliphatic index	Gravy index
Chitin synthase 2	Isoform c	158765.11	7.11	-1	40.56	98.37	0.102
	Isoform d	158848.20	7.11	-1	40.56	98.37	0.103

Cello v2.5

The sub-cellular localization of the protein was predicted by using the tool Cello v2.5 which shows that both the isoforms of chitin synthase 2 are located in inner membrane with isoform D of chitin synthase 2 having the highest reliability (3.754).



CELLO RESULTS

SeqID: NP_001137997.2 chitin synthase 2, isoform C [Drosophila melanogaster]

Analysis Report:

SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	InnerMembrane	0.988
N-peptide Comp.	InnerMembrane	0.413
Partitioned seq. Comp.	InnerMembrane	0.983
Physico-chemical Comp.	InnerMembrane	0.375
Neighboring seq. Comp.	InnerMembrane	0.987

CELLO Prediction:

	InnerMembrane	3.747 *
	OuterMembrane	0.559
	Extracellular	0.389
	Cytoplasmic	0.188
	Periplasmic	0.117

CELLO RESULTS

SeqID: NP_524209.3 chitin synthase 2, isoform D [Drosophila melanogaster]

Analysis Report:

SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	InnerMembrane	0.988
N-peptide Comp.	InnerMembrane	0.419
Partitioned seq. Comp.	InnerMembrane	0.983
Physico-chemical Comp.	InnerMembrane	0.375
Neighboring seq. Comp.	InnerMembrane	0.988

CELLO Prediction:

	InnerMembrane	3.754 *
	OuterMembrane	0.556
	Extracellular	0.385
	Cytoplasmic	0.189
	Periplasmic	0.117

Fig 1: Cello v2.5 results showing the cellular localization of the respective isoform (field marked as * indicates the cellular localization with the reliability)

TMHMM

The TMHMM tool predicted the total number of transmembrane helices (TMHs), expected no. of amino acids in TMHs and expected no. of TMHs in first sixty amino acids present in both the isoforms of the protein. All the three values were found to be same in case of isoforms of chitin synthase 2. The total probability that the N-terminal is on the cytoplasmic side of the membrane was found to 0.999 for the two isoforms of chitin synthase 2.

Table 2: TMHMM result showing transmembrane helices and their position

TMHMM result for isoform C
NP_001137997.2 Length: 1392
NP_001137997.2 Number of predicted TMHs: 17
NP_001137997.2 Exp number of AAs in TMHs: 368.49348
NP_001137997.2 Exp number, first 60 AAs: 3.90654
NP_001137997.2 Total prob of N-in: 0.99912
NP_001137997.2 TMHMM2.0 inside 1 57
NP_001137997.2 TMHMM2.0 TMhelix 58 80

NP_001137997.2	TMHMM2.0	outside	81 112
NP_001137997.2	TMHMM2.0	TMhelix	113 135
NP_001137997.2	TMHMM2.0	inside	136 147
NP_001137997.2	TMHMM2.0	TMhelix	148 170
NP_001137997.2	TMHMM2.0	outside	171 179
NP_001137997.2	TMHMM2.0	TMhelix	180 198
NP_001137997.2	TMHMM2.0	inside	199 210
NP_001137997.2	TMHMM2.0	TMhelix	211 233
NP_001137997.2	TMHMM2.0	outside	234 236
NP_001137997.2	TMHMM2.0	TMhelix	237 254
NP_001137997.2	TMHMM2.0	inside	255 278
NP_001137997.2	TMHMM2.0	TMhelix	279 301
NP_001137997.2	TMHMM2.0	outside	302 358
NP_001137997.2	TMHMM2.0	TMhelix	359 381
NP_001137997.2	TMHMM2.0	inside	382 387
NP_001137997.2	TMHMM2.0	TMhelix	388 410
NP_001137997.2	TMHMM2.0	outside	411 776
NP_001137997.2	TMHMM2.0	TMhelix	777 799
NP_001137997.2	TMHMM2.0	inside	800 894
NP_001137997.2	TMHMM2.0	TMhelix	895 917
NP_001137997.2	TMHMM2.0	outside	918 926
NP_001137997.2	TMHMM2.0	TMhelix	927 949
NP_001137997.2	TMHMM2.0	inside	950 955
NP_001137997.2	TMHMM2.0	TMhelix	956 978
NP_001137997.2	TMHMM2.0	outside	979 987
NP_001137997.2	TMHMM2.0	TMhelix	988 1007
NP_001137997.2	TMHMM2.0	inside	1008 1011
NP_001137997.2	TMHMM2.0	TMhelix	1012 1034
NP_001137997.2	TMHMM2.0	outside	1035 1123
NP_001137997.2	TMHMM2.0	TMhelix	1124 1143
NP_001137997.2	TMHMM2.0	inside	1144 1182
NP_001137997.2	TMHMM2.0	TMhelix	1183 1205
NP_001137997.2	TMHMM2.0	outside	1206 1392

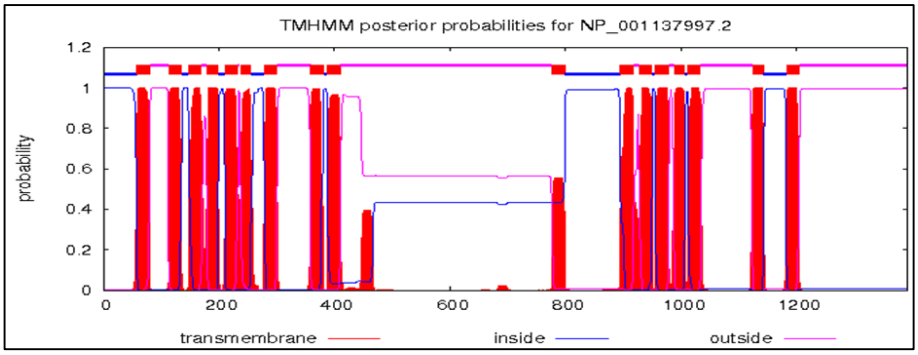


Fig 2: Graphical view of transmembrane helices of isoform C of chitin synthase 2

Table 3: TMHMM result showing transmembrane helices and their position

TMHMM result for isoform D			
# NP_524209.3 Length: 1393			
# NP_524209.3 Number of predicted TMHs: 17			
# NP_524209.3 Exp number of AAs in TMHs: 368.49228			
# NP_524209.3 Exp number, first 60 AAs: 3.90639			
# NP_524209.3 Total prob of N-in: 0.99913			
NP_524209.3	TMHMM2.0	inside	1 57
NP_524209.3	TMHMM2.0	TMhelix	58 80
NP_524209.3	TMHMM2.0	outside	81 113
NP_524209.3	TMHMM2.0	TMhelix	114 136
NP_524209.3	TMHMM2.0	inside	137 148
NP_524209.3	TMHMM2.0	TMhelix	149 171
NP_524209.3	TMHMM2.0	outside	172 180
NP_524209.3	TMHMM2.0	TMhelix	181 199
NP_524209.3	TMHMM2.0	inside	200 211

NP_524209.3	TMHMM2.0	TMhelix	212 234
NP_524209.3	TMHMM2.0	outside	235 237
NP_524209.3	TMHMM2.0	TMhelix	238 255
NP_524209.3	TMHMM2.0	inside	256 279
NP_524209.3	TMHMM2.0	TMhelix	280 302
NP_524209.3	TMHMM2.0	outside	303 359
NP_524209.3	TMHMM2.0	TMhelix	360 382
NP_524209.3	TMHMM2.0	inside	383 388
NP_524209.3	TMHMM2.0	TMhelix	389 411
NP_524209.3	TMHMM2.0	outside	412 777
NP_524209.3	TMHMM2.0	TMhelix	778 800
NP_524209.3	TMHMM2.0	inside	801 895
NP_524209.3	TMHMM2.0	TMhelix	896 918
NP_524209.3	TMHMM2.0	outside	919 927
NP_524209.3	TMHMM2.0	TMhelix	928 950
NP_524209.3	TMHMM2.0	inside	951 956
NP_524209.3	TMHMM2.0	TMhelix	957 979
NP_524209.3	TMHMM2.0	outside	980 988
NP_524209.3	TMHMM2.0	TMhelix	989 1008
NP_524209.3	TMHMM2.0	inside	1009 1012
NP_524209.3	TMHMM2.0	TMhelix	1013 1035
NP_524209.3	TMHMM2.0	outside	1036 1124
NP_524209.3	TMHMM2.0	TMhelix	1125 1144
NP_524209.3	TMHMM2.0	inside	1145 1183
NP_524209.3	TMHMM2.0	TMhelix	1184 1206
NP_524209.3	TMHMM2.0	outside	1207 1393

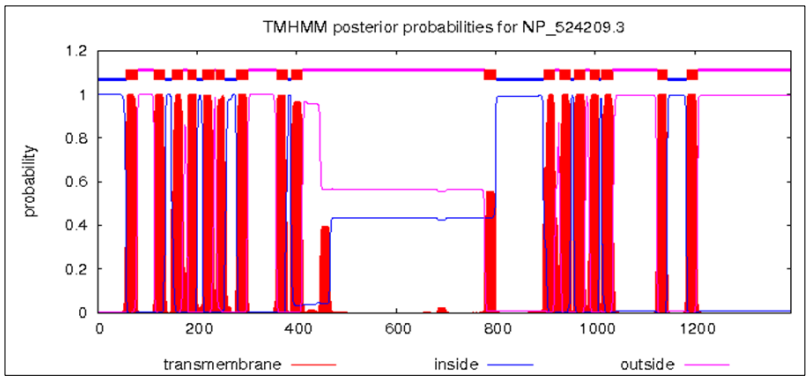
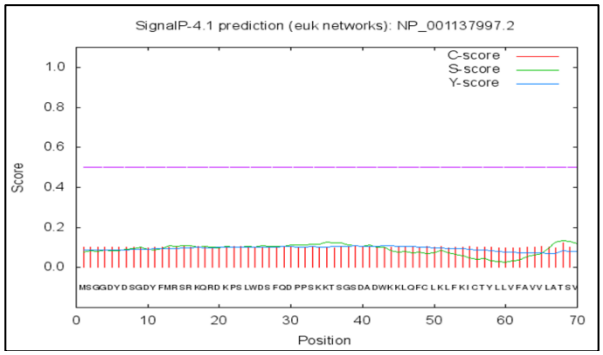


Fig 3: Graphical view of transmembrane helices of isoform D of chitin synthase 2

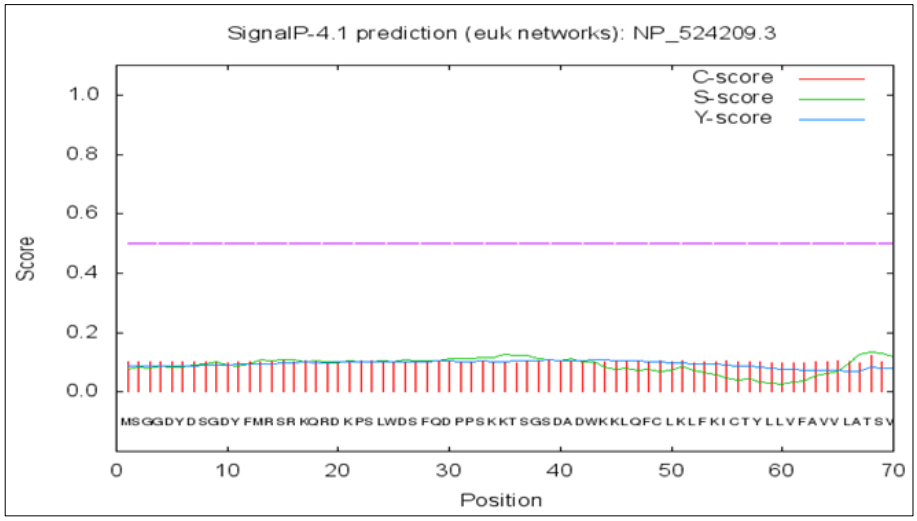
SignalP

The SignalP tool predicts whether the protein is a signal peptide or not. Fig. 3a to 3e shows the results of signalP for the isoforms of chitin synthase. Results clearly shows that none of the isoform is a signal peptide since the D-score (discrimination score) is lesser than the cutoff in each case. So both the isoforms of the protein are non-secretory.



Measure	Position	Value	Cutoff	signal peptide?
max.	C	68	0.123	
max.	Y	39	0.109	
max.	S	68	0.133	
mean	S	1-38	0.103	
D	1-38	0.106	0.500	NO

Fig 4a: SignalP graphical output showing C-score (raw cleavage site score), S-score (signal peptide score), Y-score (combined cleavage site score) for the isoform C of chtin synthase 2.



Measure	Position	Value	Cutoff	signal peptide?
max.	C	68	0.123	
max.	Y	39	0.109	
max.	S	68	0.133	
mean	S	1-38	0.103	
D	1-38	0.106	0.500	NO

Fig 4b: SignalP graphical output showing C-score (raw cleavage site score), S-score (signal peptide score), Y-score (combined cleavage site score) for the isoform D.

String

String predicted that both the isoform of chitin synthase 2 interacts with same proteins such as mummy containing 520 amino acids, kkv, rudimentary protein, topoisomerase 2, myo2881, chitinase 3, furin etc.

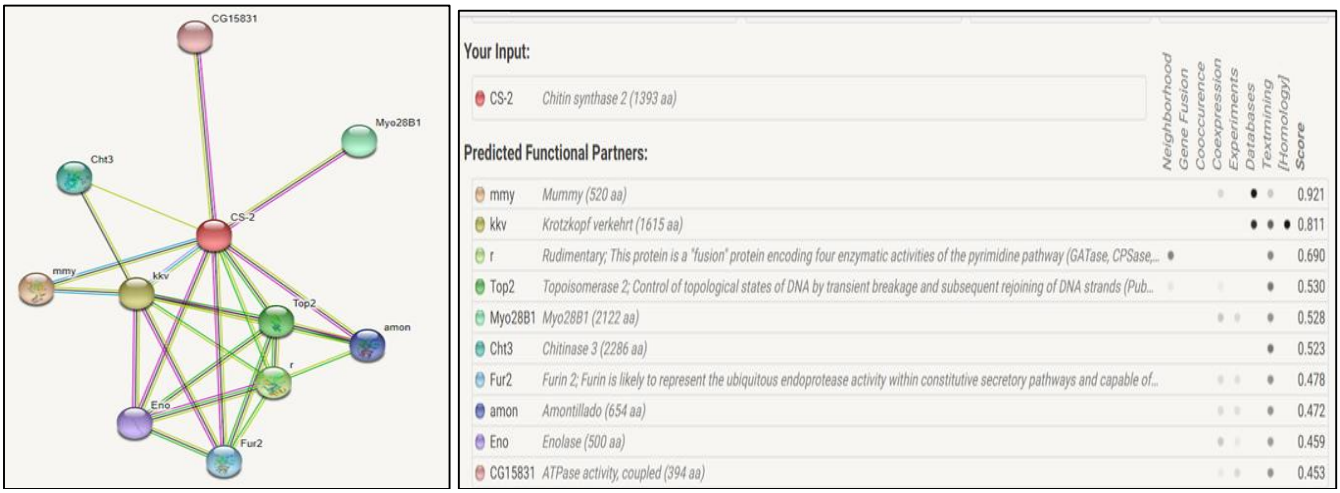


Fig 5: STRING results showing the proteins with which isoforms of chitin synthase 2 interact.

Secondary structure prediction

Secondary structure prediction was done using SOPMA. The results of secondary structure prediction are shown in fig. 5 and 6. Different secondary structures are colour coded with different colours in the sequence, alpha helix(h) with blue, extended strand(e) with red, beta turn (b) with green and random coil(c) with yellow colour in SOPMA. Predicted secondary structure in SOPMA shows that in case of isoforms of chitin synthase 2 percentage of alpha helix is exactly same but there is a slight difference in the percentage of extended strand, beta turn and random coil as shown in fig 5b and 6b. Higher number of helices makes the protein more flexible for folding that might increase interactions. There are no 3_{10} helix, Pi helix, Beta Bridge, bend region or ambiguous states in either of the isoform.

Fig 6a: SOPMA result showing secondary structure for isoform C of chitin synthase2.(Helix represented by ‘h’ in blue, random coil by ‘c’ in yellow, extended strand by ‘e’ in red and beta turn by

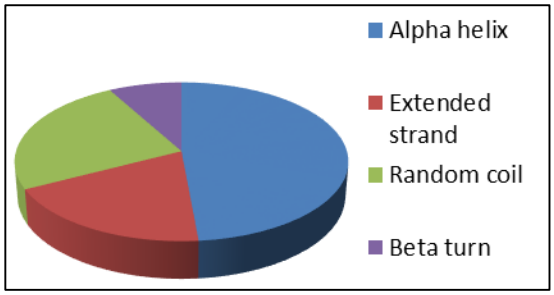


Fig 6b: Pie chart showing the percentage of alpha helix, extended strand, beta turn and random coil for isoform C of chitin synthase 2.

Fig 7a: SOPMA result showing secondary structure for isoform D of chitin synthase2. (Helix represented by ‘h’ in blue, random coil by ‘c’ in yellow, extended strand by ‘e’ in red and beta turn by ‘t’ in green color.)

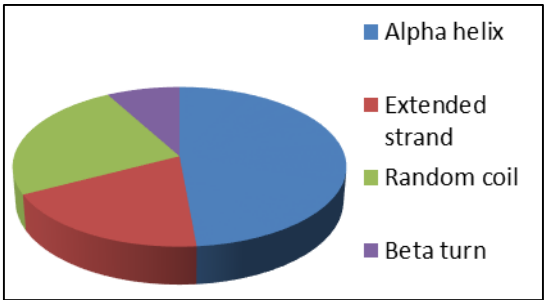


Fig 7b: Pie chart showing the percentage of alpha helix, extended strand, beta turn and random coil for isoform D of chitin synthase 2.

Ab initio modeling

Ab initio modeling of isoform C & D was carried out using Phyre2 server. For the isoform C which contains 1392 amino acids 402 residues(29%) were modelled at >90% accuracy whereas for isoform D which consists of 1393 residues 403(29%) residues were modeled at >90% accuracy.

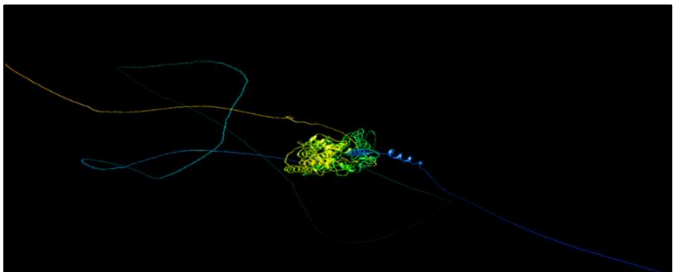


Fig 8a: *ab initio* model for isoform C

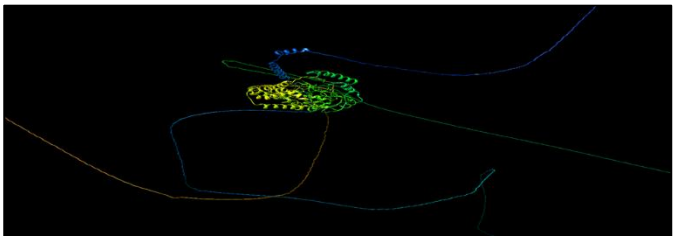


Fig 8a: *ab initio* model for isoform D

Conclusion

The current study is performing ab initio modeling of the isoforms of chitin synthase 2 from *Drosophila melanogaster* using *in silico* methods. The physiochemical properties were investigatedusing various in silico tools. Both the isoforms were classified asinner membrane proteins with no signal peptide and their functional partners were also found to be same. Alpha helices and random coil were computed to be dominating in secondary structure of both the isoforms followed by extended strand.

Acknowledgement

Authors thank Department of Science and Technology (DST), Government of India, for supporting the research by providing INSPIRE fellowship.

References

1. Merzendorfer H, Zimoch L. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases, Journal of Experimental Biology. 2003; 206:4393-4412.
2. Xin Zhang, Jianzhen Zhang, Yoonscong Park, Kun Yan Zhu. Identification and characterization of two chtin synthase genes in African malaria mosquito, Anopheles gambiae, Insect Biochemistry and Molecular Biology, 2012; 42:674-682.
3. Merzendorfer H. The cellular basis of chitin synthesis in fungi and insects: Common principles and differences, European journal of Cell Biology, 2011; 90:759-769.
4. Bhagath KP, Nagendra PS, Kasi VK, Sampath KR, Dinakara RA. UDP-N-acetylglucosamine pyrophosphorylase as novel target for controlling Aedes aegypti - molecular modeling, docking and simulation studies. International Journal of Mosquito Research 2014; 1(4):17-24.
5. Peterson TN, Brunak S, Von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods, 2011; 8:785-786.
6. Kelley LA *et al*. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols 2015; 10:845-858.