

www.sciencevision.in

Colloquium

Sci Vis 11 (3), 125-133 July-September, 2011 ISSN (print) 0975-6175 ISSN (online) 2229-6026

# Protein 3D structure determination using homology modeling and structure analysis

S. Brindha<sup>1</sup>, Sangzuala Sailo<sup>1</sup>, Liansangmawii Chhakchhuak<sup>1</sup>, Pranjal Kalita<sup>3</sup>, G. Gurusubramanian<sup>2</sup> and N. Senthil Kumar<sup>1</sup>

<sup>1</sup>Department of Biotechnology, <sup>2</sup>Department of Zoology, Mizoram University, Aizawl 796 004, India <sup>3</sup>Department of Information Technology, ICFAI University, Aizawl, India

### Abstract

The protein sequence of aspartate aminotransferase of pig was retrieved from the Swiss-Prot database. The appropriate template for homology modeling was determined using Blastp. 3D structures were determined by homology modeling softwares such as Swiss Model and EsyPred3D which passed quality test by ProQ software and set for further analysis. The pockets determined by CASTp server for the predicted structures showed a significant difference in the pocket area and volume, which were due to structural deviation between the residues 30-40 found in the 3d-ss software. Both the structures were analyzed using ProFunc tool which showed different functions as they had different structures and active sites. Thus the structure plays a vital role in determining its function.

Key words: Aspartate aminotransferase; Blastp; CASTp; 3D structure; ProQ.

### INTRODUCTION

The structure of the protein plays a vital role in determining its function.<sup>1</sup> The proteins with same structure and pocket area (active site) have been found to have different functions.<sup>2</sup> The proteins that perform similar function show different structure.<sup>3</sup> The structures are primary factor for any disease-related protein. There are four levels of protein structure: primary, secondary, tertiary and guaternary.

Proceedings of the "National Level Workshop on Random Amplified Polymorphic DNA (RAPD) Markers and It's Applications" organized on 20-21 May 2011 by the Departments of Biotechnology & Zoology, Bioinformatics Infrastructure Facility & State Biotech Hub, Mizoram University, in association with MIPOGRASS, and sponsored by DBT. For correspondence: <u>nskmzu@gmail.com</u> (NSK) <u>gurus64@yahoo.com</u> (GG) Phones. +91-0389-2330859 (NSK)/2331021 (GG)

Primary structure is the linear sequence of amino acids. Secondary structure is local conformation of  $\alpha$ -helixes,  $\beta$ -sheets and random coils. The angle between two adjacent amino acids is called torsion angle, which determines the twists/turns of the sequences resulting in Secondary structure. Tertiary structure (3D) results from the packing of secondary structural elements at a stable conformation. Quaternary structure is combination of more than one subunits or chains.

Today protein sequencing is increasing at exponential rate, till date more than 1 million proteins are sequenced and are available in the databases (SwissProt, Protein Research Foundation and Protein Information Resource). The only protein structure database is PDB (Protein Data Bank). Determining the 3D structure from a protein sequence is a dif-

ficult and challenging problem. Basically Xray crystallography or nuclear magnetic resonance (NMR) techniques are used, which are expensive, time consuming and complex process. Therefore, computational methods/ algorithms including homology modeling, threading, and *ab initio* are developed. Homology modeling is the most accurate. It is based on alignment to a known protein structure which is derived experimentally (template). If the similarity of the sequence is greater than 30% then it can act as template.

Aspartate aminotransferase (AAT) of pig's primary sequence is functionally similar with following functions (Swiss-Prot Annotation): a) plays a vital role in amino acid metabolism, b) significant for metabolic exchange between mitochondria and cytosol, c) aids cellular uptake of long-chain free fatty acids. The present work focuses on predicting the 3D structure of AAT using two homology modeling softwares - ESyPred3D and Swiss model, which runs on Windows systems, operating validating (ProQ), analyzing their active sites (CastP), comparing predicted structures for structural similarities and dissimilarities (3d-ss) and annotating the functions from the structure (ProFunc).

### MATERIALS AND METHOD

### Sequence to 3D structure using homology modeling

The sequence of AAT of pig was obtained from Swiss-Prot (<u>http://expasy.org/sport/</u>). The protein sequence was given as an input



Figure 1. Swiss-Prot webpage with aatm as input.

(Fig. 1). After the result was displayed (Fig. 2), the accession number P00506 was entered to retrieve the sequence. From the fasta link (Fig. 3), fasta sequence was obtained and copied (Fig. 4), it was saved in a notepad.

C and a page of the second sec	* EL A
O CarD to make a second of the	P目 H H (4)(0(0)(0)(0)
angeorge + :	
Na an inche lun an	
STORE AND A CONTRACTOR AND AND A CONTRACTOR AND AND A CONTRACTOR AND AND A CONTRACTOR AND A	FOR 1 HERE & Fails 1 ART
	and Discontinuiting
heat not don terms Dilaters	
for the second	
Prost Vocategeters (1) Profil	a freed a
al results for same it. AND page in its amounts	
A binned by incomery, research part controlsy, include these in pathways (12) Nethers important extentiones in 1979, 1979	A ROLL BREAK
THE REPORT OF THE PROPERTY OF THE REPORT OF THE PROPERTY OF TH	24147
Results (	
Date Shire	
- Facility and "and the superment of the Statement of	
Accession (Pary same Name Promit sprace Collins areas Organia	n tengh
T POINT APA PG T Aquatus anti-diserbrase, edit chouding SDEX See 1005	Pg - 630
T JODOT MATERIA A Appelan annual and san shadowing Gold fast. Falls for	regiona (Fall)
12	Cheld L

Figure 2. Result of the input aatm of pig.

0-0-0) . ULU-1-140		Section Court Descention of
Sec.2	Test" Mp. Rates Di	inaria.
family in	Guery	
Police Roundpress	one B	Tanti (Tant) Advanced Basellin
108606 (AVTN) (FRO)	Hummed, Unithe RB-Barss Prof.	Street hardback
COMMON WAS 1976. NO	N. 104: Meeting 1	UTI CALEND FA (T
Contraction of the local division of the loc	Control provider 10 Manager 12 Segment annuald	
Names and origin		
Polici seriel	Reserve del conse Espective anticescolores, estantecimitat ( 1503-101-1000) 2523-11 Altrasteriores Fally, anticescolores Fally, anticescolores Fally, anticescolores Fally antice	Click here to get sequence in Fasts format

Figure 3. Complete information of accession no. P00506.



Figure 4. Sequence is displayed in the fasta format.

### Selection of template

Blastp (<u>http://blast.ncbi.nlm.nih.gov/</u> <u>Blast.cgi</u>) was used to find template for homology modeling. The sequence similarity greater than 30% with known structure was selected as template. Sequence from the notepad was pasted in input box, Protein Data Bank was selected from the database and BLAST button

was entered (Fig. 5). From the Blastp result (Fig. 6) accession 3PD6\_A had sequence similarity of 93% with known structure marked S in red color, which can be chosen as template for homology modeling.

Farmants				
Today Dunit	Contraction of Factor	17 of the state of	Carlo and an and a second second	Test.
FIGH South	100 C		and the second state of	
freer excession	martinetsi, gibi, or FAST	h sequence(h) (p	Gaugadean	
Di-Son service	1 10-0003 49-1 80-1	contanativiser, actions		1
NALLANDITLAY	VERSION AND DESCRIPTION OF THE OWNER OWNER OF THE OWNER OWN	APPROX PROPERTY AND ADDRESS	the second	- 2
LOUIS VITATI	DETUGLIC CONTRACTOR PORT	THE REPORT OF CASES	inter all and all and all all all all all all all all all al	
Da, agained the	Discourse and	Street. 1	Paste the	e sequence
ing time	Superior Arts Million	fee announderan		
Spins a description	ta No lo plat \$2.027 paint			
T Align term of	And the second second	Choos	e PDB	
Choose See	ACTI DINT	1		
Database	The second second		Click	6) I.
	A Decision of the lower house	of all all all all all all all all all al	Circi	
Congression in the	Finite Internet states in	Contraction of the local sector	Tallada	

Figure 5. BlastP with input and necessary parameters.

Antima	and the second	1944	1000	Rowers.	1	
-		0.02	10.02	10100A	1.000	
53.4.	Chart A. Chyrlal Structure IV Means Meterlandinal Assarbaile Americk	80	315	125	8,8	10
4,8,8	Outri 6, Coystal Structure Of Mouse Mitochendrial Argustule Amonto-	142	10	- 07%	8.8	
WL+	Chan A, X-Roy Dituction beforement And Companions Of Three-Party	125	191	39%	6.6	
A.444	Outer A, Elyuptural Basis For The Caladytic Activity Of Assarbability	28	:118	8.9%	11	- 10
4.121.4	Dun 6, Dullal Illucture II The Deald From II Dollari Cytocolic	- 14	40.	89%	39.18	
ALM.	Oak A. Orophasiste Induced Centernalismal Olampis In Centerlin	142	- 440	01%	10111	
4.102	Oun 6, Dystal Structure of learner Gutenate Instructure Trans	100	197	896	44(11)	
1,000	Chain &, Rafmannere And Companisate OF The Church Mitschures (0179)	20	294	47%	4410	
4.101	Chair A, Ralissmett And Companies Of The Double Mitacharak (IVW)	28	181	:: 1296.	1010	- 8
MALE -	Than A. Asparlate Annualizate brane Print Bascharomyces Caterinia	100	191	87%	44.00	
111.5	Our A. Overal Structures Of Socherches Cell Aspartate Americana	1.22	-528	47%	0.01	
4.201	Oven 6, Augustate Americansherase Active Site Watant (2017-1026)	700	114	825	1-01	
1.111	Chain &, The Structural State Pur The Radwood Activity Of The 1226	322	110	879.	14-12	
BAT A	Chair A. J. & Wepttone Resolution Crystal Structure Of An Achiev 1	-	182	87%	14-64	
AM. II	20an & Itructure Bass for The Catalytic Activity (I) issueties an	- 22	-111	0.0%	14.41	- 8
DOM: N	(Day A, Stratterid Dasher Bookin The Startbratter OFE, ORG- 44	100	102	10%	10-55	
AAR a	Our 4. The Structure Hass For The different fubstrate (specificity )	100	112	52%	Jan MA	
444.0	Chain A. The Structural Basis for The Radacast Activity Of The \$223	102	100	875	2410	10
41.0	Chart A. Activity and Strachart Of The Soliton line Huberts Station F.		183	44%	14-52	10
111.1	Chan A. Algorithm Animological Areas Front E. Col. (1971) Harakan, 1	100	m	379	14.00	
1.1	Over 4. Assertate interestmetheses indice like Hutger visiteises	100	111	12%	10.00	- B

Figure 6. Results of BlastP.

### Software – ESyPred3D

Software 1: ESyPred3D server predicts 3D structure based on homology modeling.<sup>5</sup>

Inputs: Email id, name of the protein, protein sequence and PDB ID of the template which is chosen from Blastp result (Fig. 7), after submitting, the results were sent to email id provided (Fig. 8), and the attachments of PDB (.pdb) and atomic coordinate (.ali) files were downloaded. The .pdb file was opened in the Rasmol viewer (Fig. 9), which can be downloaded from <u>http://openrasmol.org/</u>.



Figure 7. ESyPred3D webpage with corresponding inputs.

#### ESyPred3D: Aspartate aminotransferase of pig



Figure 8. Predicted 3D structure result in the email.



RasMol Command Line	
RasMol Molecular Renderer	*
Roger Sayle, August 1995	
Copyright (C) Roger Sayle 1992-1993	
Version 2.7.5 June 2009	
Copyright (C) Herbert J. Bernstein 1998-2008	
*** See "help notice" for further notices ***	
Secondary Structure No Assignment	
Experiment Technique MODEL, MODELLER Version	642 20
Number of Groups 401	
Number of Atoms 3144	
Number of Bonds 3218	
Rastol>	
- Figure 9. Rasmol viewer displaying	
arat 1411222421075 ndh	
JIUL_1011333421073.pdb	

Software – Swiss Model Server

Software 2: Swiss model server (http:// swissmodel.expasy.org/) predicts 3D structure based on homology modeling.<sup>6,7</sup> To input the sequence, Automated mode was selected (Fig. 10). Inputs: Email id, name of the protein, and protein sequence. Alternatively, to speed up the process user can enter the template under the advance option and click Submit Modeling Request (Fig. 11). Significance of this model: Blastp will be done by the Swiss model server and from the result it chooses the template on its own. The quality of the model will be displayed along with the 3D structure predicted (Fig. 12). The results were sent to email with direct link the webpage. QMEAN Z-Score was used to evaluate the generated 3D structure.<sup>8</sup> The reliable score was between 0 and 1, i.e. 0.355, which is highly acceptable score. The predicted structure was visualized using Rasmol Software (Fig. 13).

Site Statute of	ZENTRUM	(fe	SW	ISS-MO	DEL
Sortemates 1400	streament I.a	U.	http://s	wissmodel.e	expasy.org/
Nymmaaa Nymmaaa Adamaad Koole Ageneet Koole	SWISS MODEL is a full attuctive transcopy non the ExPIREL web server Deschiles (Sama Pick- terver is to take Picks bochweist and males)	y automptiest prot oning poliver, auto- or from the prop leavers. The purp in Modelling accert for boologens work	elli teolote via tarri tae of the socie to all thatte	WESS MODEL To Tonten Schweite Poran Keler Larveze Benzei Konzanten Avroiz	an Prost Leader SHIDS-MODOL Reportery Selector Development and select support SHIDS-WODEL Violation
Project Mode	Wiefs sew?			References	
Termina Contribution	<ul> <li>New Beta BWS automoted motils and ligands is to East peak team</li> </ul>	S MODEL power elaulong with orgonization of australia for the organization before	e for partent eing	<ul> <li>Amplit public plane de MCCCEL, plane de Amplit K, B T, control 1</li> </ul>	r report veryal carry divides- inter research publications. Iontol 1, Kogo 7, and Botweek Val BANKE JUMP, Internet

Figure 10. Swiss model webpage.



Figure 11. Inputs of the Swiss model webpage.

Gette (Motel Quality E	vimation ( (Rignment) (Model	ing Log (Terpi	we Switchor)	Portfao
Model Details 🥥 [-/.]				
ST.	Model information: Mudeled residue range Eased on temptate Bonark: No search for K Only user specified temp Bestamon identity (NL Evalue	30 to 430 Jan KA (J. 40) anglete was perf fato was used for 54 TR3 E Star	6j formuld r readefing	
-m	Quality information QMEAN 2-Score	0.356		

Figure 12. Re-directed page from the email link showing the predicted 3D structure.

Tanta 1000. LPO INDUCTION PORT IN A CONTRACTOR	Line in the second seco
The Bit Thurk Colour Steine Steine Freet Her	Randhi Milensiza Reminest Auger Tayla, Anguzi 159 Ceptright (D. Sepa Augis 150-1500 Versina T. 1: Jone 200 Versina T. Jone 200 Ve

Figure 13. Rasmol view of aatm\_pig 3D structure modeled by Swiss model server.

### Structure validation

ProQ-Protein Quality Predictor<sup>9</sup> is the software to check the quality of the model (<u>http://www.sbc.su.se/~bjornw/ProQ/</u> <u>ProQ.html</u>) (Fig. 14). If the predicted structure satisfies the validation parameters of ProQ process then the structure was taken for further analysis. The .pdb file predicted by

ESyPred3D uploaded (Fig. 15). The results of the structure file modeled by ESyPred3D is shown in Figure 16. Similarly the file with Swiss-Model was processed and the result is displayed in Figure 17.



Figure 14. Webpage of the ProQ web server.



Figure 15. Upload 3D structure in ProQ webpage.



Figure 16. 3D structure modeled by EsyPred3D.

#### ProQ - Results

y structure the prediction will be more reliable)
***********
Result of the Structure modeled from Swiss-Model

Figure 17. 3D structure modeled by Swiss model.

### Structure analysis

CastP server<sup>10</sup> predicts the active sites with their respective volume and area (<u>http://</u><u>sts.bioengr.uic.edu/castp/calculation.php</u>). The active pocket (active site) with more volume and area is always taken for further analysis. The .pdb file was uploaded to get the result (Fig. 18). Active site of the 3D structure predicted by EsyPred3D and Swiss-Model is shown in Figures 19&20, respectively, and the amino acids involved in pocket formation in both the models are displayed in Figure 21.

### Structure comparison using superposition techniques

3d-ss (3-dimensional structure superposition) positions the 3D structure one above



Figure 18. CASTp web server.



Figure 19. Model predicted by ESyPred3D.



Figure 20. Model predicted by ESyPred3D.



Figure 21. Pocket area of the 3D structure modeled by EsyPred3D and Swiss model.

another to find similar and deviated regions (<u>http://cluster.physics.iisc.ernet.in/3dss/</u>) between the structures.<sup>8</sup> Input was given as in Figure 22. There are only two structures for comparison. Hence, the option superimpose only two structures was selected and their respective pdb file were loaded (Fig. 23). All the options were kept in the default mode and proceeded as per Figures 24 to 27.

Here, both the protein structures modeled using two homology modeling techniques (Swiss model and EsyPred3D) were chosen, two ways of superposition were carried out as per the following criteria:

- a. 3D structure by Swiss model was kept as fixed molecule and EsyPred3D as rotated molecule (Fig. 28a).
- b. 3D structure by EsyPred3D model was kept as fixed molecule and Swiss model as rotated molecule (Fig. 28b).

## Predicting function from structure

ProFunc identifies the likely functions of a protein pertaining to its 3D structure (<u>http://</u>www.ebi.ac.uk/thornton-srv/databases/





Figure 25. Step 2 of 3d-SS result.

profunc/index.html) analyzing the protein's sequence and structure by identifying functional motifs.<sup>11</sup> The .pdb file was loaded in the ProFunc online tool as in Figure 29. Preliminary information was entered (Fig. 30). The result was sent to the email provided. Similarly the pdb file modeled using Swiss



### Figure 26. Step 3 of 3d-SS result.

Fired Molecule WISSMODELEPDE		Rotated Molecule Stan		tamp <sup>1</sup> Score	RM	SD (Å)	lignme	ut Long
		ESYPRED	D.PDB	9,564	( 0	313		01
	Ret	ation Matr	ix and Tr	anslat	tion	Vector		
	Rotatis	n Matrix				Frandat	on Vect	ae.
8.99640	1 -0	06713	0.05172	I.		-0.5	1206	
0.06619		99761	0.01961			-2.7	1679	
4.05292	1 -0	01611	0.99547	0		1.3	6554	
umher Fined)	(Fixed)	(Rotated)	(Rotates	1) <b>**</b>	we	w		Value
umher	manie	number	name	A0		Deviati	on Pr	obabilit; Color
rand)	(ennel)	(Rotated)	(Rotates	9				1997
30		70	-	10		77.04	-	0.000
13	W	11	-	10		6.479		6 181
		11	w	1			+	0.102
33	W	11	W	To	AC	4.471	+	0.444
34	A	34	8	10	A	8.799		0.024
35	- 11	-	-	1	-	-	1	
425	-1	425	T	C	A	0.110		0.955
	П	436	- 11	IC	Ast	8.269		0.957
426	Q	427	0	C	A	6.332		8.956
426	1	425	V	C	A	0.218	T	8.958
426 427 425			100000	100	100	6.350	-	0.954
426 427 428 429	T	429 (	1	1.4				

Figure 27. Step 4 of 3d-SS result.

model was performed. The sequence motifs (Fig. 31), matching folds (Fig. 32), nest analysis (Fig. 33) and 3D functional template search results (Fig. 34a&b) were obtained.



Figure 28. Graphical result of 3d-SS showing deviation between residues 30-40.

### Result and Discussion

Both the 3D structures modeled using ESyPred3D and Swiss model resulted with LG Score of 6.29 and 6.32, respectively, yielding to an extremely good model according to ProQ validation criteria (Fig. 16&17).

CASTp Server shows the pocket (active site) for ESyPred3D model: Area = 1379.9 and volume = 4401.9, respective, amino acids forming the packets are blocked in the sequence as highlighted in Figure 19. The



Figure 29. PDB file uploaded in the ProFunc tool.

FOR STOR		
te uploadest input of	user details	
Here Here	der details Strekten Kossen Base 15.5mm Aufwerto Roundaus (1920) Billette (1930) Billette (1930)	
or PDE Sin has been success iffunc has interpreted the file real.	Character of the a study unstantion. Before commercing the functional analysis (a contents-control)s. Then competito the form before before	n, prease check the defails across to ensure that lisk on Ruri to place your Shukure rite the Profi
or POB fee has been success share has interpreted the fee real. They name	Changing 421 8.8	a, peane checi. The details above to ensure that lets on Run to place your structure role tha there
er POB fan han seen sussen fran he interpreted tie fan rea Fran same Joseffachstartgary	Character Kit A a that you and the converse of the functions analysis or contents controls. Then compare the time issue and co Sector Microsoft	A, panase creat. The details above to ensure that both on their to place your shocking into the Their
or POI for has noon success shure has interpreted the for even the second second second second second break to the second	Champy Kit La chily secand. Before commercing the functional analysis research seconds. Then compare the functional and a Secretar Kurman M2D1 Full-mail second	A, sease check the detail, above to encoure the tot on thur is place, now shackow not the the More than the information
er POH fan han neen hutsen feur han referended fan fan hat Ynge name Jostfacherumgery E maak activeur Protein name	Character Kirl A a thay sustaints. Before commercing the functional analysis of controls controls. Then compute the functional sectors (Long) Sectors (Long) Advices (Long) sectors (Lon	n, presentant for secal above to ensure the late on Rurit place provident sectors region theory Enter the information

Figure 30. ProFunc preliminary information.

ProFund res	uits
Augustalia, Paula	Fig. 34a ESyPred3D Modeled Structure
10000710	Internet Scale Act Sequence Name. Internet Scale Act Sequence Name. Data Scale Act Security Action Scale
ProPunc re	sufts
Seguence met	Fig. 34b Swiss Model modeled Structure
preservo.	HearPre-base for begaviour results. 6 methods in statused in local against PRODUCT, PROFILE, PROFILE, PROFILE, S. and PRODUCT, methods Tope Molf More and Product and Annual Product and Annual Product and Produc

Figure 31. Sequence motifs results of ProFunc.

Matthewn Yold	a senicled by 154 Fig 35e, E5yPred3D modeled Structure
đ	Marching Solds.
	(SATY significant structural matches)
	A market and the second
	KONE SKIE AAK BAR COM MALTANY of ensure minimum that has been ensured as a series to be the series of the ser
	2 1311 131A 14 12.5 Table Cystal structure of wave excision and a spartale amountain complex with mathematics and
	<ol> <li>L. BERD, R. JAA, 24. TEX. These Control and an extension of an extension of the extension of th</li></ol>
	6 6202 0.47A 04 17.2 that 3 yes shorters whereast put comparison of these forms of estimation and estimation an
	5, 8 202, 9 205, 24, 17,2, Tail, 3, by sharfare extraordinal and inequations of these terms of anticolocation associated anomaly available an
	- Aist affert
Mediting to	Fire 35h Saving Meridian Structure
Ø	Matching Kotes.
	\$20% significant structural matches
	store Read 555 acres H10 Warm
	<ol> <li>C200 X.074 - 34 TE.7 Spill Copyloi Ministrari of Images entrophysical apportate and other scorely block his Spreacher and copylogical approximation.</li> </ol>
	3 E.H.I. 5.163 24 18.5 Subjictly structure of movies individuality atparticle analysis with productions and
	<ol> <li>E. B.244, M.24, Dr. (R.R. 2000) 274/04 address of meaning information provides and anti-information dependence of meaning information and provides and anti-information and provides and pr</li></ol>
	4, \$300, \$348, 24, 183, Surt 2 ray structure offeneration of comparison of these forms of administration
	5. E399 8.444, 24 18.5 Test Sing situation of boundaries and comparison of these forms of manufactural
	_ plut attieve

Figure 32. Number of folds matching (predicted structure Vs PDB) - ProFunc software.



Figure 34. 3D functional template searches – ProFunc.

pocket for Swiss model: Area = 534.9 and volume = 1015.5, respective amino acids forming the packets are blocked in the sequence region (Fig. 20). The pocket size of 3D structure by Swiss model is aligning well with 3D structure by EsyPred3D (Fig. 21) because the pocket size of the 3D structure by Swiss model is smaller. The residue deviation from 30-40 (> 0.5 Armstrong) is high (Fig. 28), which plays a vital role in the pocket difference observed in the CASTp server (Fig. 21). Amino acids between 30-40 residues are involved in the pocket formation in the model determined by EsyPred3D which is not seen in structure predicted by Swiss model. But the residues from 43 to 430 remain less deviated (< 0.5 Å), which is negligible.

The results of sequence motifs of the ESyPred3D and Swiss model 3D structures reveal the presence of transaminase region in ESyPred3D model and presence of G3DSA motif in Swiss model structure (Fig. 31). There are more matching folds and nests in ESyPred3D model compared to Swiss model structure (Fig. 32&33). In reverse template comparison versus structures in PDB one hit/function is observed only in ESyPred3D model (Fig. 34). The differences in the functions from protein structure are due to deviated site regions between two models which are observed in Figure 28.

3D structures of AAT of pig showed high deviation between residues 30-40, which affects the active site areas in the structure. The functions of the 3D structures are also in motifs, nests and reverse template comparison versus structures. Further study is required to prove the functions in the wet lab.

#### ACKNOWLEDGEMENT

The authors thank DBT, New Delhi, for supporting research related to bioinformatics by providing Bioinformatics Infrastructure Facility (BIF) through BTISNet DBT/CTEP/01/201100105 dt. 23 March 2011 and State Biotech Hub (BT/04/NE/2009 dtd. 22nd Dec. 2010).

#### References

- Whisstock JC & Lesk AM (2003). Prediction of protein function from sequence and structure. *Q Rev Biophys*, 36, 307-340.
- 2. Wise E, Yew WS, Babbitt PC, Gerlt JA & Rayment I (2002). Homologous ( $\beta/\alpha$ )8-barrel enzymes that catalyze unrelated reactions: orotidine 5'-monophosphate decarboxylase and 3-keto-L-gulonate 6-phosphate decarboxylase. *Biochemistry*, 41, 3861-3869.
- 3. Chan HS & Dill K (1993). The protein folding problem. *Physics Today*, 2, 24-32.
- Sivakumar K (2006). Proteomics: from sequence to structure. Advanced BioTech, 4(11), 18-23.
- Lambert C, Leonard N, De Bolle X & Depiereux E (2002). ESyPred3D: Prediction of proteins structures. *Bioinformatics*, 18, 1250-1256.
- Arnold K, Bordoli L, Kopp J & Schwede T (2006). The SWISS-MODEL Workspace: A web- based environment for protein structure homology modelling. *Bioinformatics*, 22, 195-201.
- Kiefer F, Arnold K, Künzli M, Bordoli L & Schwede T (2009). The SWISS-MODEL repository and associated resources. *Nucleic Acids Res*, 37, D387-D392.
- 8. Bjorn W & Arne E (2003). Can correct protein models be identified? *Protein Sci*, 12, 1073-1086.
- Joe D, Zheng O, Jeffery T, Andrew B, Yaron T & Jie L (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res*, 34, 116-118.
- Sumathi K, Ananthalakshmi P, Md. Roshan MNA & Sekar K (2006). 3dSS: 3-dimensional structural superposition. *Nucleic Acids Res*, 34, 128-134.
- Roman AL, James DW & Janet MT (2005). ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res*, 33, 89-93.