

Insights of interaction between small and large subunits of ADP-glucose pyrophosphorylase from bread wheat (*Triticum aestivum* L.)

Mohd Danishuddin, Ravish Chatrath, Rajender Singh*

Directorate of Wheat Research, Post Box 158, Karnal - 132001, India; Rajender Singh - Email: rajenderkhokhar@yahoo.com; *Corresponding author

Received July 26, 2010, Accepted January 29, 2011; Published May 07, 2011

Abstract:

Lack of knowledge of three dimensional structures of small and large subunits of ADP- glucose pyrophosphorylase (AGPase) in wheat has hindered efforts to understand the binding specificities of substrate and catalytic mechanism. Thus, to understand the structure activity relationship, 3D structures were built by homology modelling based on crystal structure of potato tuber ADP-glucose pyrophosphorylase. Selected models were refined by energy minimization and further validated by Procheck and Prosa-web analysis. Ramachandran plot showed that overall main chain and side chain parameters are favourable. Moreover, Z-score of the models from Prosa-web analysis gave the conformation that they are in the range of the template. Interaction analysis depicts the involvement of six amino acids in hydrogen bonding (AGP-SThr422-AGP-LMet138, AGP- SArg420-AGP-LGly47, AGP-SSer259-AGP-LSer306, AGP-SGlu241-AGP-LIle311, AGP-SGln113-AGP-LGlu286 and AGP-SGln70-AGP-LLys291). Fifteen amino acids of small subunit were able to make hydrophobic contacts with seventeen amino acids of large subunit. Furthermore, decrease in the solvent accessible surface area in the amino acids involved in interaction were also reported. All the distances were formed in between 2.27 to 3.78Å. The present study focussed on heterodimeric structure of (AGPase). This predicted complex not only enhance our understanding of the interaction mechanism between these subunits (AGP-L and AGP-S) but also enable to further study to obtain better variants of this enzyme for the improvement of the plant yield.

Keywords: AGPase, ADP-glucose pyrophosphorylase, AGP-S, AGP-L, modelling, molecular docking, hydrogen bonding, hydrophobic contacts, accessible surface area.

Background:

Wheat (*Triticum aestivum* L.) is one of the most important staple crops worldwide, with a total production of over 600 million tonnes annually. The seed number and seed weight are critical yield components of wheat. Starch, which accounts 65-75% of wheat seed weight, is a major determinant wheat yield. Starch is known to be an important carbohydrate and the primary energy source for plants, having numerous industrial applications [1, 2]. It is generally accepted that four enzymes may play a key role in starch biosynthesis: ADP-glucose pyrophosphorylase, starch synthase, starch branching enzyme and starch debranching enzyme [3]. ADP-glucose pyrophosphorylase (AGPase) is regarded as rate- limiting enzyme in starch biosynthesis. This enzyme plays a key role in the modulation of photosynthetic efficiency in source tissues and also determines the level of storage starch in sink tissues, thus influencing overall crop yield potential. It catalyzes the first step of starch biosynthesis by generating the sugar nucleotide ADP-glucose and inorganic pyrophosphate (Pi) from glucose-1-phosphate and ATP. ADP-glucose functions as the glucosyl donor for glucan synthesis by starch synthase. AGPase from higher plants is heterotetrameric, consisting of two large subunits (AGP-L) and two small subunits (AGP-S) encoded by two distinct genes [4]. Seed yield and plant biomass increases are conferred by deregulation of endosperm AGPase and thus, AGPase has attracted wide interest for potential crop improvements [5]. In the present study the 3D structures of the subunits (large and small) from wheat were modelled by the *in silico* homology techniques using the crystal

structure of small subunit from potato tuber as a template. Moreover, molecular docking between these subunits was also performed to gain the insight of interaction.

Methodology:

Amino acids sequences of large subunit (P12299) and small subunit (P30523) were retrieved from Swiss-Prot database [6]. Template was searched by BLAST-P analysis against PDB database. The crystal structure of potato AGPase (1YP2) available at PDB was used as template for modelling. Sequence alignment between the model sequence and template was done with ClustalW [7]. Homology models of large and small subunits were built by Modeller [8] version 9v7. One model from each subunit was selected using PROCHECK [9] and Prosa-web [10]. Further, models were subjected to energy minimization using GROMOS96 implemented via Swiss-pdb viewer [11]. RMSD of the model structures of AGP-L and AGP-S were evaluated from its template using the SUPERPOSE [12]. Docking of model structure was performed by GRAMM-X [13]. Model structure was further evaluated by Patchdock [14]. Best docked structure based on RMSD and Patchdock score was chosen for further analysis. Hydrogen and hydrophobic interactions between these subunits were analyzed by Ligplot [15]. Figures representation was generated with Discovery studio visualizer programme [16]. Accessible surface area was calculated by Mark Gerstein's calc-surface programme implemented in Chimera [17].

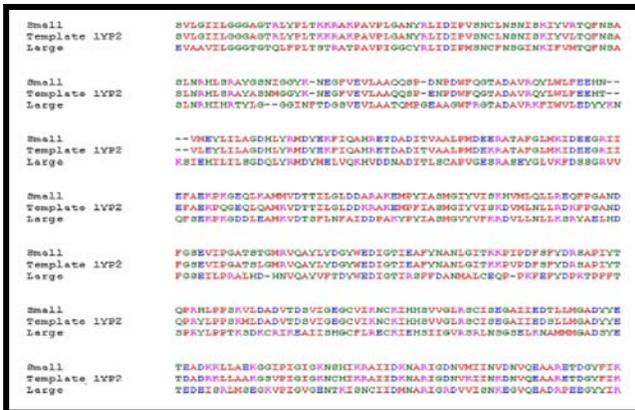


Figure 1: Sequence alignment between the small subunit, large subunit of wheat and corresponding sequence of small subunit of potato (PDB: 1YP2)

Results and Discussion:

Model building and quality assessment:

BLAST analysis of gene sequences showed highest homology with the small subunit of ADP-glucose pyrophosphorylase (AGPase) from potato. The alignment between the sequence of small subunit of ADP-glucose pyrophosphorylase (AGP-S) from wheat and potato tuber revealed at least 90% identity (**Figure 1**), which allowed for a predictable homology modelling approach. However, the alignment between sequences of small and large subunits revealed 52% identity (Data not shown). Template (1YP2) from potato AGPase was used for modelling the small and large subunits of wheat with Modeller 9v7. The final models were selected based on DOPE score. Evaluation of the selected models was done by Procheck and Prosa-web server. The analysis of the Ramachandran plot and Prosa-web of the template (1YP2) were used to compare the overall stereochemical quality of AGPase subunits (AGP-L, AGP-S). Ramachandran analysis showed 89% and 90% amino acids of AGP-L and AGP-S, respectively, in most favourable region (**Figure 2 & 3**). The main chain conformation within the favoured or allowed region of Ramachandran plot and G-factor indicated the accuracy of generated models. Analysis of AGP-L and AGP-S models revealed the Z-score value of -9.56 and -8.57, respectively, which is in the range of nature conformation of the template (**Table 1 see Supplementary material**). Furthermore root mean square deviation of AGP-S and AGP-L models with respect to the equilibrated structure of AGPS (1YP2) from potato and was found to be 1.08 and 1.30, respectively. Thus overall quality assessment of the selected models assisted us to use these models for further analysis.

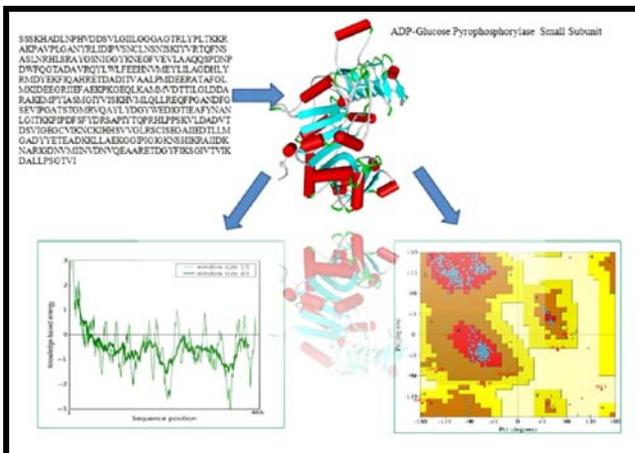


Figure 2: Model structure of AGP-S and quality assessment evaluated by Procheck and Prosa-web

In silico docking:

We used GRAMM-X programme for docking purpose. This programme perform a rigid-body docking using Fast Fourier Transformation methods (FFT) by applying smoothed Lennard-Jones potential, knowledge-based and refinement stage scoring, which give rise to best surface match. Simulation of structural flexibility is a computational expensive process for protein-protein

docking. Thus, computationally docking is difficult for putting two proteins in a complementary manner. High computational complexity restricts the flexible docking algorithms and is rarely applicable to practical protein docking at present. This problem can be overcome by using Rigid body docking algorithm. Rigid body algorithm assumes two proteins as rigid bodies. The conformation change is tolerated by allowing certain degree of penetration between proteins. This assumption will limit the problem to a six-dimensional (three for translation and three for rotation) search space [18].

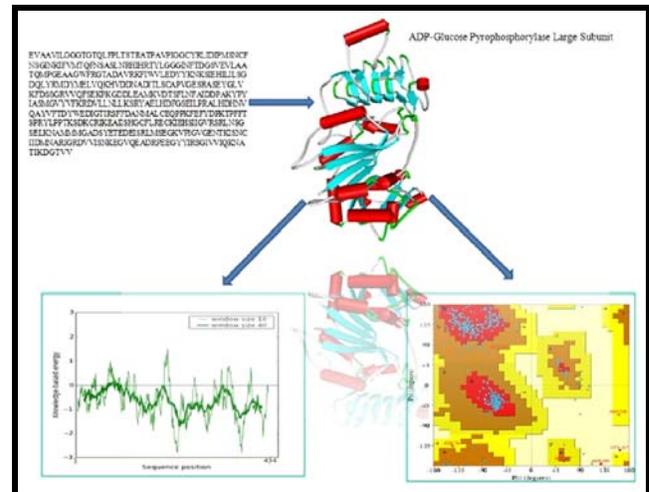


Figure 3: Model structure of AGP-L and quality assessment evaluated by Procheck and Prosa-web.

GRAMM-X generated a PDB file containing the structures of ten models ranked according to the scoring function. We selected best docked complex for analyzing the interaction mechanism (**Figure 2 & 3**). For accuracy of our result we checked our model by Patchdock [14] which uses molecular docking algorithm based on shape complimentary principle. Complementary patches are matched in order to generate candidate transformation which further evaluated by scoring function, considering both geometric fit and atomic desolvation energy [19]. For each complex top 15814 score from the Patchdock was retained for further consideration, while other were rejected due to lower score. Thus, best result from two algorithm not only increase our confidence level but also provide strength in accuracy for our output data. Ligplot analysis depicted the involvement of six amino acids in hydrogen bonding (**Figure 4 & Table 2 see Supplementary material**). Amino acids (Gln70, Gln113, Glu241, Ser259, Arg420, Thr422) of AGP-S were found to form hydrogen bonds with amino acids (Gln47, Met138, Glu286, Lys291, Ser306, Ile311) of AGP-L (**Figure 4**). Six hydrogen bonds were also formed in complex of small and large subunits of AGPase from potato [20]. The hydrogen bond lengths between AGP-L and AGP-S were found to be shorter than 3.3Å. Moreover, seventeen amino acids of AGP-L were involved in seventy five hydrophobic contact with fifteen amino acids of AGP-S (**Table 3 see Supplementary material**). This shows that complex is stabilized by both hydrogen bonding and hydrophobic interactions. Out of these amino acids two amino acids were found to be crucial. Amino acid Gln113 of small subunit donate the hydrogen bond to Glu286 of large subunit with 1.96Å distance which was the smallest in the complex. This amino acid also made five hydrophobic contacts with Glu286 in the distance of 2.3Å to 3.7Å. Amino acids Ile311 of AGP-L made a hydrogen bond with Glu241 of AGP-S with a distance of 3.1Å. Eight hydrophobic contacts were formed by this amino acid (Ile311-AGP-L) with Pro244 of AGP-S within the range of 2.5Å to 3.5Å. Moreover, decrease in accessible surface area in the docking complex has been observed. Change of accessible surface area of Gln113 AGP-S was found to be significant and it changed from 109.75 to 17.85 in complex (**Table 4 see Supplementary material**). As mention previously this amino acids was also found to be critical for hydrogen bonding and hydrophobic interaction (**Table 3 see Supplementary material**). Similarly accessible area of other amino acids of AGP-L; Lys291 (167.82 to 36.53), Thr300 (203.14 to 104.67) and Pro303 (102.86 to 20.94) were also found critical in the complex (**Table 5 see Supplementary material**). Thus these outcomes suggest that the interaction complex is feasible and useful for understanding the binding mechanism between AGP-L and AGP-S subunits of ADP-glucose pyrophosphorylase in wheat.

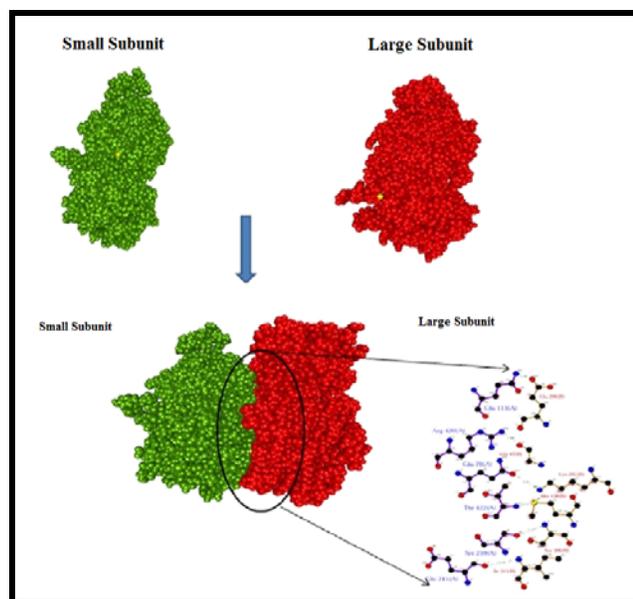


Figure 4: Model of AGP-S and AGP-L complex interaction and hydrogen bonding pattern

Conclusion:

ADP-glucose pyrophosphorylase (AGPase) is a major enzyme controlling starch synthesis, and has been demonstrated in many different plant species. The crystal structure of AGPase (small and large subunits) from wheat has not yet been solved. Thus, the rationale in building the AGPase subunits models and performing *in silico* docking study was to gain the details of interaction between two subunits (AGP-L and AGP-S). This docking study revealed the important residues involved in formation of the complex. Six amino acids of AGP-S and AGP-L were involved in hydrogen bonding whereas fifteen amino acids of AGP-S were involved in hydrophobic interaction with seventeen

amino acids of AGP-L. Importantly, all the amino acids that were involved in hydrogen bonding were found to be crucial as they were also caught up by hydrophobic interactions. Thus, this study hypothesises the model which can be used for further study to elucidate the role of AGPase gene in starch biosynthesis to increase the starch content in wheat and thus, grain yield.

Acknowledgments:

The financial support for Agri-Bioinformatics Promotion Program provided by Bioinformatics Initiative Division, Department of Information Technology, Ministry of Communications & Information Technology, Government of India, New Delhi is gratefully acknowledged. The authors are thankful to Project Director, DWR, Karnal for providing facilities for this work.

References:

- [1] Slattey CJ *et al. Trends Plant Sci.* 2000 **5**: 291 [PMID: 10871901]
- [2] Tuncel A *et al. Biophys J.* 2008 **95**: 3628 [PMID: 18641076]
- [3] Martin C & Smith AM. *Plant Cell* 1995 **7**: 971 [PMID: 7640529]
- [4] Salamone PR *et al. FEBS Lett.* 2000 **482**: 113 [PMID: 11018533]
- [5] Cross JM *et al. Plant Physiol.* 2004 **135**: 137 [PMID: 15122037]
- [6] Boutet E *et al. Methods Mol Biol.* 2007 **406**: 89 [PMID: 18287689]
- [7] Thompson JD *et al. Nucleic Acids Res.* 1994 **22**: 4673 [PMID: 7984417]
- [8] Sali A & Blundell TL. *J Mol Biol.* 1993 **234**: 779 [PMID: 8254673]
- [9] Laskowski RA *et al. J Appl Cryst.* 1993 **26**: 283
- [10] Wiederstein M & Sippl MJ. *Nucleic Acids Res.* 2007 **35**: 407 [PMID: 17517781]
- [11] Guex N & Peitsch MC. *Electrophoresis* 1997 **18**: 2714 [PMID: 9504803]
- [12] Maiti R *et al. Nucleic Acids Res.* 2004 **32**: W590 [PMID: 15215457]
- [13] Tovchigrechko A & Vakser IA. *Nucleic Acids Res.* 2006 **34**: W310 [PMID: 16845016]
- [14] Schneidman-Duhovny D *et al. Nucleic Acids Res.* 2005 **33**: W363 [PMID: 15980490]
- [15] Wallace AC *et al. Protein Eng.* 1995 **8**: 127 [PMID: 7630882]
- [16] <http://accelrys.com/products/discovery-studio/visualization>
- [17] <http://www.cgl.ucsf.edu/chimera/>
- [18] <http://www.comp.nus.edu.sg/~leowwk/hyp/huangwenfan.pdf>
- [19] Zhang C *et al. J Mol Biol.* 1997 **267**: 707 [PMID: 9126848]
- [20] Baris I *et al. Plos Comput Biol.* 2009 **5**: e1000546 [PMID: 19876371]

Edited by N Pattabiraman

Citation: Danishuddin *et al.* Bioinformatics 6(4): 144-148 (2011)
 and reproduction in any medium, for non-commercial purposes,
 provided the original author and source are credited.

Supplementary material:

Table 1: Ramchandran plot and Z-score of the models

Model	Ramachandran plot	Prosa2003
AGPase Large subunit	88.4	-9.54
AGPase Small subunit	89.3	-8.17
Template (1YP2)	87.5	-8.76

Table 2: Hydrogen bonding pattern in complex

Donor	Acceptor	Distance(Å)
THR A* 422	MET B* 138	2.11
ARG A* 420	GLY B* 47	1.98
SER B* 306	SER A* 259	3.17
ILE B* 311	GLU A* 241	3.14
GLN A* 113	GLU B* 286	1.96
LYS B* 291	GLN A* 70	3.31

*A- Small subunit; *B- Large subunit

Table 3: Hydrophobic contacts in the complex

Total contacts	*SS amino acids	*LS amino acids	Distance(Å)
10	Tyr30	Lys 291	2.72-3.80
2	Trp111	Asp289	3.71-3.87
5	Gln113	Glu286	2.34-3.77
4	Glu192	Pro302, Pro303	3.40-3.82
6	Lys193	Tyr300	3.35-3.76
4	Phe243	Ser306, Cys309	3.13-3.63
11	Pro244	Ile311, Lys312	2.83-3.72
3	Gly245	Ala314, Ile315	2.84-3.83
3	Ser251	Pro303	3.40-3.77
4	Ser259	Lys305	2.59-3.59
2	Ala413	Thr78	3.35-3.89
6	Val415	Thr78	2.09-3.85
5	Ala419	Asn45, Gly47	3.38-3.57
2	Glu421	Met138	3.25-3.65
8	Thr422	Met138	2.89-3.87

*SS- Small subunit; *LS- Large subunit

Table 4: Accessible surface area values for AGP-S amino acids in unbound and bound stage

SS* Amino acids	Uncomplex	Complex
Tyr30	75.88	28.04
Trp111	89.44	33.41
Gln113	109.75	17.85
Glu192	89.97	33.02
Lys193	102.56	51.7
Phe243	50.89	6.16
Pro244	70.32	17.79
Gly245	63.04	15.07
Ser251	40.8	14.78
Ser259	96.06	13.41
Ala413	102.1	65.39
Val415	65.94	28.25
Ala419	62.7	33.05
Glu421	41.9	12.1
Thr422	103.87	69.52

*SS- Small subunit

Table 5: Accessible surface area values for AGP-L amino acids in unbound and bound stage

LS* Amino acids	Uncomplex	Complex
Asn45	59.41	21.49
Gly42	5.2	4.63
Gln70	19.17	17.05
Thr78	109.51	30.17
Met138	95.99	20.27
Glu286	69.69	10.06
Asp289	26.29	4.64
Lys291	167.82	36.53
Thr300	203.15	104.67
Pro302	70.93	41.6
Pro303	120.86	20.94
Lys305	118	85.79
Ser306	50.81	12.61
Cys309	33.28	6.89
Ile311	61.78	3.34
Lys312	141.95	134.54
Ala314	26.08	2.6

*LS- Large subunit