Indian Journal of Biotechnology
Vol 15, July 2016, pp 334-342

Heterodimeric interaction of the ADP-glucose pyrophosphorylase (AGPase) enzyme in *Hordeum vulgare*

Bharati Pandey, Chetna Tyagi, Ohika Chakraborty, A K Mishra*, Amrender Kumar and A K Jain

Agricultural Knowledge Management Unit, Indian Agricultural Research Institute, Pusa Campus, New Delhi 110 012, India

Received 30 March 2015; revised 5 June 2015; accepted 17 June 2015

ADP-glucose pyrophosphorylase (E.C. 2.7.7.27; AGPase) is a key regulatory enzyme, constituting two small (SS) and two large (LS) subunits. The crystallographic structure of AGPase enzyme has not been reported yet for barley and its active site residues have also not been identified, which is restraining the complete understanding of structure-function relationships of this enzyme. In the present study, three-dimensional (3D) structures of LS and SS of barley AGPase were built through homology modeling and optimized using MD (molecular dynamics) simulations. Further evaluation resulted in about 74.8% residues of LS and 75.7% residues of SS falling in the favorable regions of Ramachandran Plot, demonstrating the stability and trustworthiness of the models. RMSD (root mean square deviation) of 1.1 Å and 1.2 Å was predicted on superimposition of the deduced LS and SS structure on the template 1YP2, implying the similarity between the structures.

Protein-protein docking was carried out using ZDOCK and GRAMM-X server to obtain the stable heterodimer structure of barley AGPase. Interaction analysis using the Dimplot revealed six hydrogen-bonding interactions between HIS-359, GLN-322, SER-113, LEU-342, SER-324 and ASN-111 residues of LS, and ASN-341, SER-286, SER-8, ASP-261 and VAL-283 residues of SS. The structure-function relationship and substrate binding specificity of AGPase will provide better understanding for the role of specific amino acid accountable for allosteric regulation.

Keywords: ADP-glucose pyrophosphorylase (AGPase), docking, homology modeling, MD simulation, protein-protein interaction

Introduction

Barley (*Hordeum vulgare* L.) is amongst the world's earliest domesticated and most important cereal crops. It is diploid in nature with a large genome of 5.1 gb[1]. Barley grains are the most essential raw material for brewing and its importance has been widely recognized in industry, food and animal livestock. Barley and wheat starch have a bimodal distribution of granular size. It is generally believed that those starch granules increase in size and stay as the large granules at maturity are laid down first, and the small ones are laid down at a later stage in growth and remain as such at maturity[2].

ADP-glucose pyrophosphorylase (AGPase) is an enzyme that plays a leading role in the starch biosynthesis in both photosynthetic and non-photosynthetic plant tissues. In the formation of ADP-glucose using ATP and glucose-1 phosphate as substrates, a regulatory step is catalyzed by AGPase[3]. This enzyme regulates the modulation of photosynthetic efficiency in source tissues and also determines the localization of starch in sink tissues, thereby affecting overall crop yield potential[4]. Plant AGPase is a heterotetramer, composed of two distinct large (LS) and small (SS) subunits, encoded by *Shrunken-2 (Sh2)* and *Brittle-2* (Bt2) genes, respectively[5,6]. Sequence analysis of protein sequence of SSs from various plants showed that SSs are highly conserved (80-90%), as compared to LSs (50-60%)[7].

Although SS and LS share substantial sequence identity, current evidence suggests that these subunits perform distinct roles in AGPase function. The distinct roles of the LS and SS in functioning of AGPase are also supported by both random and site directed mutagenesis[4,8]. One approach to increase the productivity and the nutritional value of crop plant is to enhance the rate of starch biosynthesis[9]. Kleczkowski[10] reported partial reduction of AGPase activity and starch level as a result of missense mutation in the LS gene in *Arabidopsis* leaves. Regulation of AGPase depends on the ratio of 3-phosphoglyceric acid to inorganic phosphate (3PGA/Pi). The enzyme composed of only the SSs required much higher levels of 3-PGA for activation

*Author for correspondence:
Tel: +91-11-25841255
akmishra.usi.iari@gmail.com

as compared to wild type (WT) *Arabidopsis* leaf AGPase [11].

Recombinant SS of AGPase purified from *Escherichia coli* has been crystallized[12]. Its structure has been solved within the resolution range 2.1Å to 2.6Å in the absence of effectors and substrate and also in the absence of ATP and glucose-1 phosphate. Crystal structure of AGPase from potato tuber (SS homotetramer; PDB ID; 1YP2) was obtained with resolution of 2.11Å[12]. Afterwards, Tuncel *et al*[13] modeled the LS of potato AGPase using homology modeling techniques and proposed a model for the heterotetrameric AGPase. AGPase protein model was inspected for key residues that mediate the interactions between LS and SS using both computational and experimental approaches[14]. Two LS dimers were found to be involved in LS-SS subunit interactions based on the molecular mechanics approach.

Computational methods to predict protein structure and protein-protein interaction have been successfully applied in biochemical research for decades. Three-dimensional (3D) structures of small and large subunits of AGPase from bread wheat was built by homology modelling based on crystal structure of potato tuber AGPase[15]. Recently, structures of the LS and SS of rice AGPase were modeled using comparative modeling and their interaction leading to generation of AGPase heterotetramer in rice were predicted[16]. Since functionality of proteins is important to life itself, so determining the 3D structure is prerequisite for a complete understanding of the biological functions. However, structural properties of large and small granules have not been inspected in detail in barley till date. Thus, it is critical to reveal the barley AGPase structure and examine the crucial residues participating in LS and SS interactions to attain a more comprehensive picture of the enzyme. This strongly limits a complete understanding of the structure-function relationships of the enzyme and also manipulation of the enzyme for increased starch production in grains. Subsequently, amidst homologous proteins, structure is more conserved than primary sequence, thus the 3D structure prediction of AGPase can help us in predicting their biologically significant function.

In the present study, we performed a protein modeling study of AGPase protein using homology modeling to obtain their 3D structure. Predicted models were further energy minimized and rectified by molecular dynamics (MD) simulations to attain a better-quality model. Protein-protein docking was performed using ZDOCK and GRAMMX program both based on Fast Fourier Transformation (FFT) algorithm. The residues mediating the interaction between large and small subunits were analyzed.

Materials and Methods

Secondary Structural Analysis and Disorder Prediction of AGPase Protein

The amino acid sequences of LS and SS of barley AGPase (HvAGPase) were retrieved from the Swiss-Prot database (Acc. id: P30524 & P55238). PROFsec secondary structure prediction tool (https://www.predictprotein.org/) was employed to enumerate the secondary structural features of LS and SS of HvAGPase. The disorder prediction for LS and SS proteins was depicted using a web-based server named IUPred[17], which plots a disorder prediction graph along with the disorder probability for each residue position.

Structure Prediction for HvAGPase LS and SS

3D structures of both protein subunits were built through homology modeling and were obtained using Modeller 9.11 (http://salilab.org/modeller) taking crystal structure of potato tuber ADP-glucose pyrophosphorylase (PDB ID: 1YP2) as the template. The resultant homology model was validated using online protein structure validation suite (PSVS)[18], which comprises a number of software, namely, Verify3D[19], PROCHECK[20] MolProbity[21] and Prosa II[22].

Molecular Dynamics (MD) Simulations of Modeled Protein and Docked Complexes

To get energetically stable conformation of the modeled proteins, MD simulations were done using Desmond Molecular Dynamics module[23-25] of Schrodinger Maestro by applying Optimized Potentials for Liquid Simulations (OPLS) all-atom force field 2005. MD simulations were conducted at 300 K, 1 atm with a time step of 2 femtoseconds (fs). The root mean square deviation (RMSD) values were calculated for the entire simulations trajectory with reference to their respective first frames. The predicted LS and SS were superimposed on the homotetramer crystal structure of potato AGPase (template structure: 1YP2). The superimposition and RSMD calculation was performed by the UCSF Chimera

(http://www.cgl.ucsf.edu/chimera/download.html) to determine the accuracy of the modeled structure.

Protein-Protein Docking

For protein-protein docking, the ZDOCK[26] and GRAMM-X[27] docking servers were used. ZDOCK follows 3D Fast Fourier Transformation (FFT) algorithm to explicitly explore rigid body protein search of the spatial degrees of freedom between two molecules to significantly speed up searching in translational spaces[28]. ZDOCK ranks models in terms of scoring function termed ZScore, that includes shape complementarity, electrostatics and a pairwise atomic statistical potential of transient protein complexes[29]. GRAMM-X also performs rigid body docking using FFT method by applying smooth Lennard-Jones potential, knowledge-based and refinement stage scoring, which gives rise to the best surface match. The RosettaDock server (http://rosettaserver.graylab.jhu.edu/) was used to correct the orientation of docked model obtained from GRAMM-X. The best orientation PDB files with energy scores were retrieved and the minimum energy structure was taken for further reference. The predicted model was energy minimized to remove unfavorable non-bonded contacts and distorted bond angles with force field GROMOS96[30] with the Swiss-PDB Viewer (http://spdbv.vital-it.ch).

Subunit-Subunit Interaction Studies

To study protein-protein interactions, the energy minimized structures were used as input in programs, namely, dimplot (http://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/), to generate plot of interactions for the dimers.

Results and Discussion

Primary Protein Sequence Analysis

Currently, most of the algorithms for the anticipation of secondary structure of a protein are based on machine learning techniques, substantially improving the prediction precision. Among the available methods for secondary structure prediction till date, PROFsec has achieved predictive accuracy, which predicts on average ~76% of all residues correctly[31] The amount of alpha (α) helix, beta (β) sheets, and coils in SS was 31.19, 16.57 and 52.24%, respectively; whereas the corresponding figures in LS was 33.00, 16.90 and 50.10%, respectively. These secondary structure predictions will assist in understanding protein folding, sequence-structure relationships and evolutionary profiles[32].

As the AGPase protein belongs to the ADP-glucose pyrophosphorylase (ADP_Glu_pyroP_CS) family, which catalyzes the synthesis of ADP glucose and inorganic pyrophosphate from glucose-1-phosphate (Glc-1-P) and ATP, and thus generating the nucleotide sugar used by starch synthases to incorporate glucosyl units into starch. In order to carry out such reaction efficiently, proteins are examined for intrinsically unstructured regions based on energy estimates[17]. The disorder prediction outcomes for two protein subunits have been explicitly depicted in Fig. 1. The value of disorder prediction probability value exceeding threshold 0.5, signifies disordered state of a certain amino acid residues. In LS, N-terminal region was predicted to be slightly disordered, whereas both N and C terminal regions were predicted to fairly disorder in SS. However, N-terminal region does not contain any binding site or domain. On the contrary, C-terminal area contains two protein polypeptide binding sites and sulfate binding site, thus might affect the strength of interaction between subunits.

3D Model Prediction and Validation

HvAGPase LS and SS (57 & 56 kDa) have been cloned and purified[33]. They consist of 523 and 513 amino acid residues, respectively and found in the

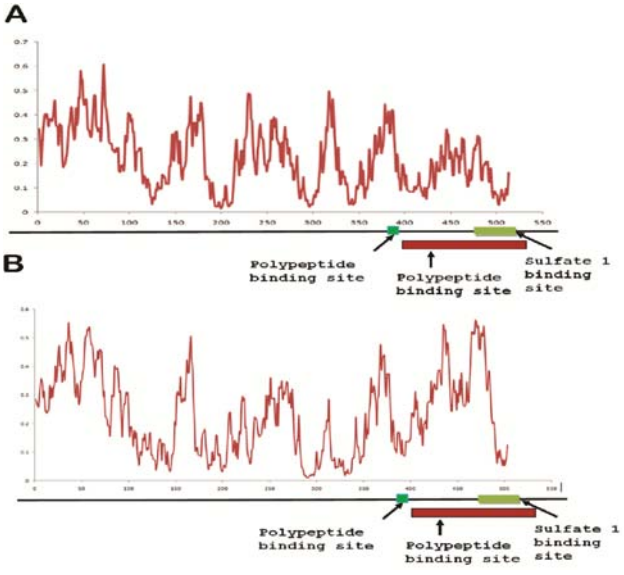

Fig. 1 (A & B)—Disorder prediction graphs of barley AGPase: (A) LS, & (B) SS, with respect to each residue position. A disorder prediction probability value of >0.5 signifies that a particular residue is disordered otherwise not. The different binding sites have been marked along the sequence position (polypeptide binding site in green, polypeptide binding site in red & sulfate binding site in light green).

chloroplast of leaf. With an identity of 92 and 57% and alignment score of 843 and 513, the crystal structure of potato tuber AGPase (PDB ID: 1YP2) has the highest ranked homology in the BLASTP search. Kosloff and Kolodny[34] have demonstrated that 3D structures are similar if the sequence identity between two proteins is higher than 40% and will align to 2.5Å RSMD. Therefore, the 3D structure of HvAGPase was built through homology modeling using chain A of 1YP2 as a template. The crystal structure of potato tuber AGPase was determined at a resolution of 2.21Å[12]. The amino acid sequence of the potato AGPase covers full length of HvAGPase protein sequence, except few residues in the N terminus (Fig. 2). The initial alignment of LS and SS with the template sequence was obtained using ClustalW. The final alignment of LS (57% identity, 76% positivity & 1% gaps) and SS (92% identity, 96% positivity & no gap) was used in the homology modeling as shown in Fig. 2.

Modeller software constructed five different protein models for LS and SS and calculated three scores (DOPE, molpdf & GA341) based on their physiochemical properties, such as, van der Waals interaction, hydrophobicity, hydrophilicity, atomic charges and atomic energy. Model 1 and model 3 were observed as the potential models for further analysis for LS and SS, respectively (Table 1). DOPE (Discrete Optimized Protein Energy) score is a statistical potential for evaluation of homology models. The DOPE score of a protein can be viewed as a conformational energy, which measures the relative stability of a conformation with respect to other conformations of the same protein[35]. GA341 score is a statistical potential for fold assessment. To improve fold assessment, four types of a residue-level statistical potential were optimized, including distance-dependent, contact Ψ/Φ dihedral angle, and accessible surface statistical potentials[36]. According to standard protocol, lower values of DOPE score represent a best probable model.

MD Simulation

The predicted model was further refined by MD simulations. MD was performed on Desmond utility provided in the Schrodinger's Maestro suite to determine stable structure by subjecting to the 20000 picosecond (ps) MD simulation. An average structure was computed for the most stable time frame (0-20 ns) of the simulations trajectory. A trajectory graph has been provided in Fig. 3 for the LS and SS. For the LS, RMSDs gradually increases in the first 4 ns and then become stable in the time frame from 5 to 20 ns, whereas in SS, RMSDs was most stable for 10-20 ns.

The simulated protein structures were again validated by the PSVS suite and the results are summarized in Table 2. Further, Fig. 4 summarizes the extent of variation in the topology of these models occurring after MD simulations were performed. The pre-MD models of LS and SS were superimposed on

Table 1—Summary of predicted protein models of large (LS) and small (SS) subunits using Modeller 9.11

Filename*	molpdf score	DOPE score	GA341 score
HORVUSS.B99990001	4679.3887	-52312.171875	1.00000
HORVUSS.B99990002	5135.7832	-51789.199213	1.00000
HORVUSS.B99990003	4641.4253	-52742.070313	1.00000
HORVUSS.B99990004	4927.4023	-52275.156250	1.00000
HORVUSS.B99990005	5034.0083	-52634.296875	1.00000
HORVULS.B99990001	4641.4253	-53606.062500	1.00000
HORVULS.B99990002	5135.7832	-53493.394531	1.00000
HORVULS.B99990003	4679.3887	-53308.656250	1.00000
HORVULS.B99990004	4927.4023	-52994.644531	1.00000
HORVULS.B99990005	5034.0083	-53151.109375	1.00000

*HORVUSS; small subunit; HORVULS; large subunit

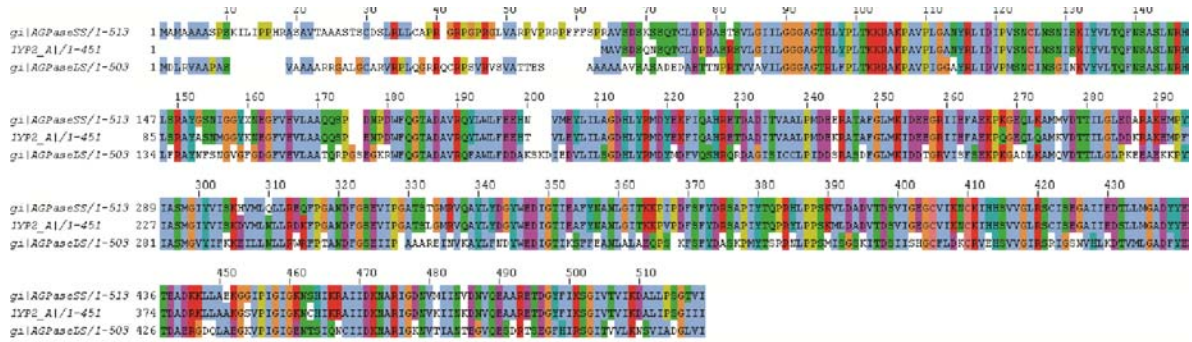

Fig. 2—Sequence alignment between the SS and LS of barley and corresponding sequence of SS of potato (PDB: 1YP2).

Table 2—Ramachandran plot for the validation of large (LS) and small (SS) subunits of AGPase of barley proteins before and after molecular dynamic simulations

	Residues		Most favored (%)	Allowed regions (%)	Generously allowed (%)	Disallowed (%)
Large subunits	Pre-MD	Procheck	85.7	9.6	1.8	2.9
		MolProbity	91.9	3.6	-	4.4
	Post-MD	Procheck	74.8	22.1	2.6	0.5
		MolProbity	85	12.7	-	2.3
Small subunits	Pre-MD	Procheck	85.1	10.8	3.4	0.7
		MolProbity	93.2	3.4	-	3.4
	Post-MD	Procheck	75.7	21.9	2.0	0.5
		MolProbity	84.6	13.2	-	2.1

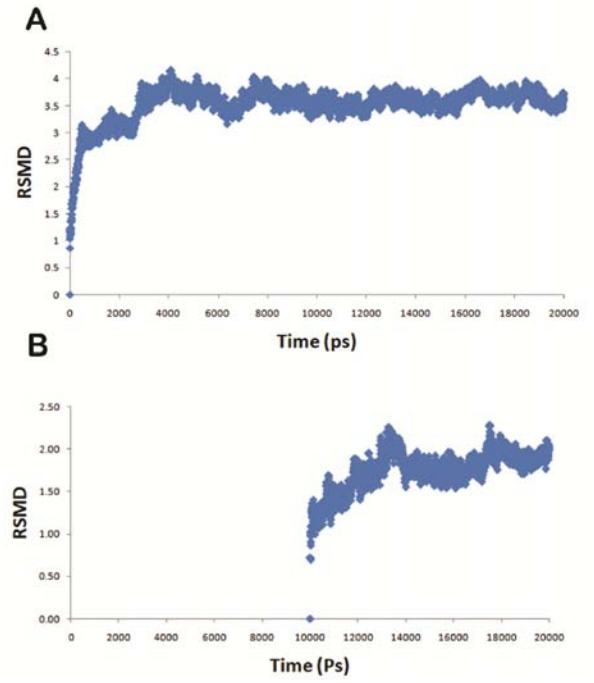

Fig. 3 (A & B)—A RMSD graph showing simulation trajectory acquired by the barley AGPase: (A) LS, & (B) SS.

the post-MD LS and SS structures (Fig. 4). The superimposition showed a RMSD value of 1.17Å and 1.2Å for LS and SS, respectively. The LS (Fig. 4A) had 19 β-strands and 15 α-helices, whereas the SS (Fig. 4B) had 18 β-strands and 15 α-helices. As a result of simulation, the numbers of residues falling under most favored regions and allowed regions in the Ramachandran plot were considerably increased both for LS and SS. This is due to the confiscation of poor contacts from the protein structure through MD simulation, which alters the global quality of protein. Similarly, residues falling in disallowed regions of the Ramachandran plot declined by good amount,

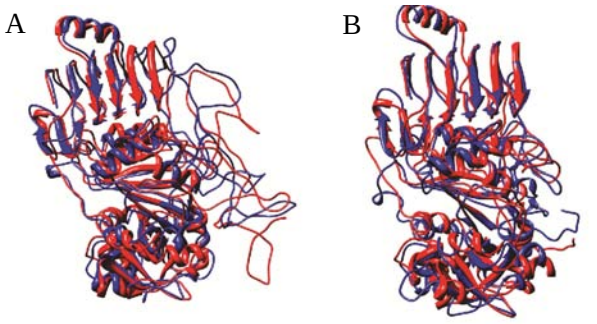

Fig. 4 (A & B)—Comparative topological structure depiction of barley AGPase: (A) LS, & (B) SS, obtained before (shown in red) and after molecular dynamics simulations (shown in blue).

indicating stable and trustworthy models. Pre and Post MD Ramachandran Plot analysis has been provided in Figs 5A and B.

Protein-Protein Docking

Computational approach, such as, protein-protein docking directly maps physical interactions between proteins[37,38]. Owing to greater protein flexibility, high computational complexity restricts the flexible docking algorithms. Rigid body docking algorithm can be used to overcome such problem in order to reduce computational complexity[39]. Therefore, LS as receptor and SS as donor were docked with the GRAMM-X server, generating a heterodimer, which were further energy minimized. According to scoring function, GRAMM-X results top scored ten models. Among them model 1 was selected as the best docked complex for depicting the interaction mechanism (Figs 6A & B). Docking results were also validated by ZDOCK[26], which uses FFT-based protein docking.

Thus, outcome from two approaches increases confidence level and also provides strength in accuracy for analyzing data. Energy score for docked

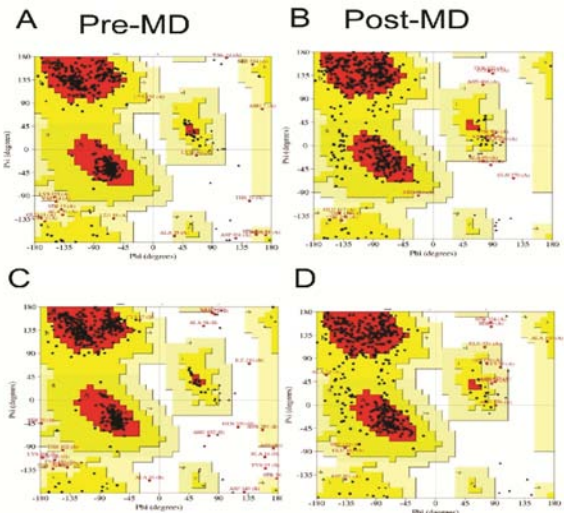

Fig. 5 (A-D)—Comparative diagram depicting Ramachandran Plot analysis of AGPase protein subunits before and after molecular dynamic simulations (A & C) SS (B & D) LS. Ramachandran plots show the phi (φ)-psi (ψ) torsion angles for LS and SS of AGPase amino acid residues in the structure. The plots were generated in PROCHECK

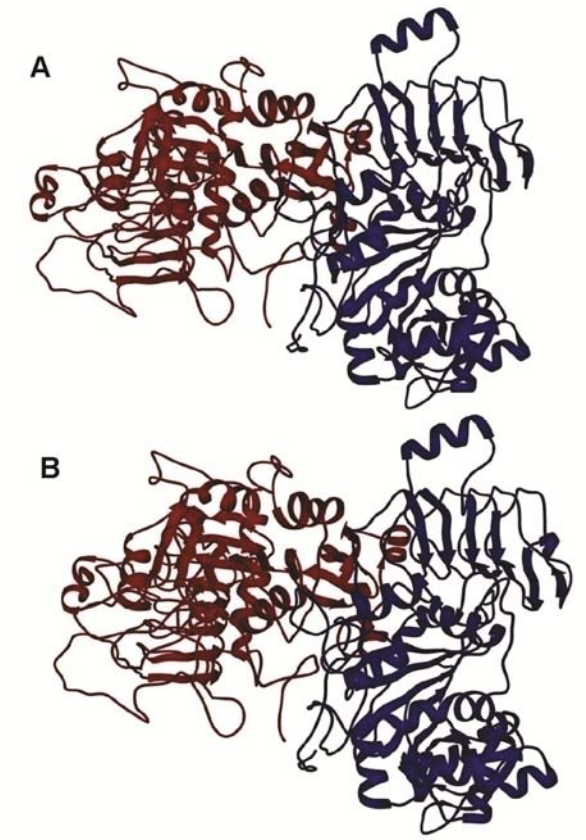

Fig. 6 (A & B)—Docked dimer structure of AGPase from barley obtained from the GRAMM-X server with energy score of 191285054.9 kj/mol. Energy-minimized structure of dimer with energy score of −457,816.6 kj/ mol. [Blue chain, LS; Red chain SS]

complex changed from 191285054.9 kJ/mol (without minimization) to –457,816.6 kJ/mol after energy minimization with force field GROMOS96, suggesting the accuracy of docked model for analysis. Dimplot, module of the Ligplot+ program was used to depict the hydrogen and hydrophobic interaction between subunits. Six amino acids (HIS-359, GLN-322, SER-113, LEU-342, SER-324 & ASN-111) in LS were found to form six hydrogen bonds with amino acids (ASN-341, SER-286, SER-8, ASP-261 & VAL-283) of SS (Table 3; Fig. 7). Similar results and approach were reported for heterodimer complex of AGPase from wheat[15] and rice[16].

The hydrogen bond lengths between LS and SS were found to be in the range of 2.21-3.31Å. Furthermore, fifteen amino acids of LS were engaged in sixty seven hydrophobic contacts with eighteen amino acids of SS, with bond length below 3.0Å distance (Table 4). Hydrophobic interactions provide a very strong driving force in protein folding. Thus the results suggest that the complex is stabilized predominately by hydrogen bonding although hydrophobic interactions remain essentially constant. Out of these, two amino acids were found to be crucial. With the smallest distance in the complex 2.21Å, SER-8 of SS donates the hydrogen bond to ASN-111 of LS, which is a crucial amino acid. VAL-323 from LS made eighteen hydrophobic contacts with TRP-124 and GLU-128 amino acids of SS within the range of 0.58Å to 2.48Å. Furthermore, relative accessible surface area was calculated for LS and SS and also predicts the conformational changes in the docked complex using computationally intensive method, ProtSA[40]. Significant change in solvent accessible surface area of amino acids mediating LS-SS interaction, such as, ASP-261 (SS; 66.638 to 0.770) and SER-324 (LS; 75.839 to 14.188) was examined in docked complex (Table 5). The amino acids ASP-261 and SER-324 were found to be critical as they are involved in both

Table 3—Predicted hydrogen bonds interactions in LS-SS in AGPase of barley

Donor			Acceptor			Distance (Å)
HIS	A	359	ASP	B	261	2.83
ASN	B	341	LEU	A	342	2.91
SER	B	286	SER	A	324	2.7
GLN	A	322	VAL	B	283	2.4
SER	A	113	SER	B	8	3.31
SER	B	8	ASN	A	111	2.21

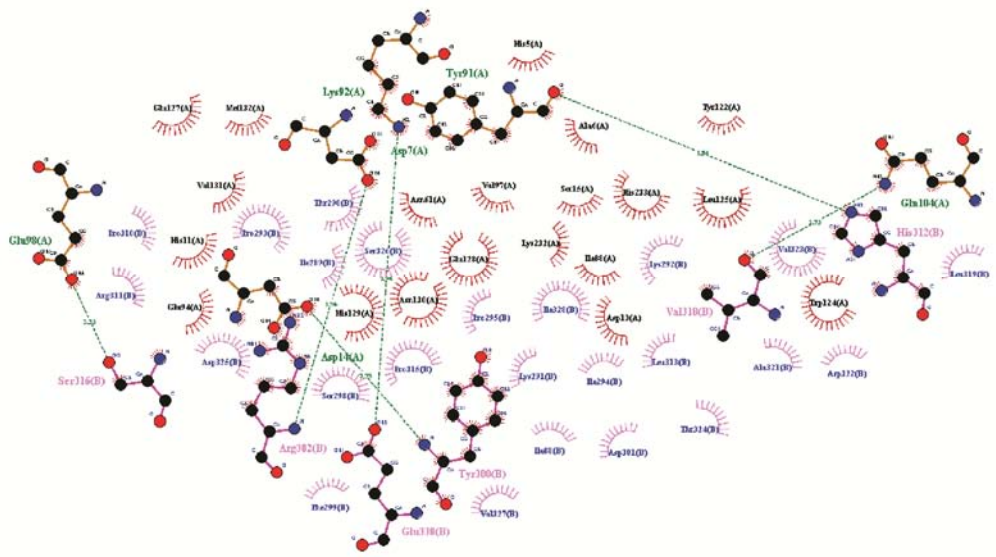

Fig. 7—Hydrogen and hydrophobic interactions of LS (A) & SS (B) as plotted by Dimplot. [Dashed lines hydrogen bonds; Arcs hydrophobic interactions]

Table 4—Predicted hydrophobic interactions in LS-SS in AGPase of barley

Amino acid	Chain	Amino acid	Chain	Distance (Å)	Amino acid	Chain	Amino acid	Chain	Distance (Å)
MET	B	ARG	A	1.44	LEU	B	VAL	A	2.33
MET	B	ARG	A	1.61	LEU	B	SER	A	2.44
MET	B	ARG	A	1.88	LEU	B	GLU	A	2.54
MET	B	ARG	A	1.89	LEU	B	SER	A	2.77
MET	B	GLY	A	2.01	LEU	B	SER	A	2.91
MET	B	ARG	A	2.12	PRO	B	THR	A	2.96
MET	B	ARG	A	2.20	LYS	B	ILE	A	1.04
MET	B	GLY	A	2.20	LYS	B	ILE	A	2.42
MET	B	ARG	A	2.47	LYS	B	ILE	A	2.57
MET	B	ARG	A	2.66	ASP	B	HIS	A	2.57
MET	B	GLY	A	2.84	SER	B	PHE	A	2.98
MET	B	ARG	A	2.92	ARG	B	HIS	A	2.26
MET	B	ARG	A	2.96	ARG	B	HIS	A	2.49
PHE	B	SER	A	2.76	ARG	B	HIS	A	2.75
PHE	B	SER	A	2.80	ASP	B	LYS	A	2.94
VAL	B	PHE	A	1.66	SER	B	SER	A	2.81
VAL	B	PHE	A	1.86	ILE	B	SER	A	1.94
VAL	B	PHE	A	1.90	ILE	B	PRO	A	2.08
VAL	B	PHE	A	1.94	ILE	B	PRO	A	2.10
VAL	B	PHE	A	2.16	ILE	B	SER	A	2.16
VAL	B	PHE	A	2.19	ILE	B	SER	A	2.36
VAL	B	PHE	A	2.70	PRO	B	GLN	A	2.80
VAL	B	PHE	A	2.71	LEU	B	PRO	A	1.76
VAL	B	PHE	A	2.76	LEU	B	PRO	A	1.81
VAL	B	PHE	A	2.78	LEU	B	PRO	A	2.47
VAL	B	VAL	A	2.78	LEU	B	PRO	A	2.80
VAL	B	PHE	A	2.79	LEU	B	PRO	A	2.80
VAL	B	VAL	A	2.82	LEU	B	PRO	A	2.86
VAL	B	PHE	A	2.90	HIS	B	PRO	A	2.13
VAL	B	SER	A	2.92	HIS	B	PRO	A	2.15
VAL	B	ASP	A	3.00	HIS	B	PRO	A	2.76
LEU	B	SER	A	1.33	HIS	B	PRO	A	2.81
LEU	B	SER	A	2.21	HIS	B	PRO	A	2.84
					HIS	B	PRO	A	3.00

Table 5—Solvent accessible surface area for amino acid mediating interaction between LS and SS in AGPase of barley

Amino acid	Position	Chain	Area ($Å^2$)	
			Subunit	Docked complex
HIS	359	A	107.016	24.693
GLN	322	A	147.029	34.800
SER	113	A	85.238	44.883
LEU	342	A	23.835	8.749
SER	324	A	75.839	14.188
ASN	111	A	37.154	17.883
ASP	261	B	66.638	0.770
ASN	341	B	39.046	16.070
SER	286	B	39.338	10.058
SER	8	B	97.997	17.429
VAL	283	B	16.096	11.689

hydrogen bonding and hydrophobic interaction (Tables 3 & 4).

Conclusion

AGPase enzyme plays a central role in the modulation of photosynthetic efficiency in source tissues and also regulates the level of starch storage in them. Thus, it influences the global crop yield potential. In view of these, computational approaches are viable options to obtain 3D structure and to understand further the structure-functional relationships. The analysis of the sequences revealed that LS and SS proteins are mixture of α, β and coil but containing more than 50% of random coils, which make highly flexible conformation. N-terminal region of LS and N and C-terminal of SS were found to be slightly disordered.

The 3D structure of heterodimer of barley AGPase was predicted using homology modeling and further model optimization was done using MD simulations. The results illustrated that the structure obtained in barley was similar to the crystal structure of potato AGPase (template) as evident by the RMSD value of their superimposition. Subunit interaction studies by Dimplot listed the key residues, which form hydrogen bonding between LS and SS. Extrapolation of these key amino acid residues between LS and SS will pave way for engineering of the enzyme to increase starch yield in plants. Thus, combined data support for binding interaction of two protein subunits further enhance the understanding of AGPase regulation in barley.

Acknowledgement

Authors are thankful to the Department of Biotechnology, Government of India, New Delhi for the financial support to the present study. The authors are also grateful to the School of Biotechnology, Jawaharlal Nehru University, New Delhi for providing the necessary facilities to carry out molecular dynamics simulations.

References

1 International Barley Genome Sequencing Consortium, Mayer K F, Waugh R, Brown J W, Close T J *et al*, A physical, genetic and functional sequence assembly of the barley genome, *Nature (Lond)*, 491 (2012) 711-716.

2 Singh S, Choi S B, Modi M K & Okita T W, Isolation and characterization of cDNA clones encoding ADP-glucose pyrophosphorylase (AGPase) large and small subunits from chickpea (*Cicer arietinum* L.), *Phytochemistry*, 59 (2002) 261-268.

3 Geigenberger P, Regulation of starch biosynthesis in response to a fluctuating environment, *Plant Physiol*, 155 (2011) 1566-1577.

4 Kavakli I H, Park J S, Slattery C J, Salamone P R, Frohlick J *et al*, Analysis of allosteric effector binding sites of potato ADP-glucose pyrophosphorylase through reverse genetics, *J Biol Chem*, 276 (2001) 40834-40840.

5 Bae J M, Giroux M J & Hannah L C, Cloning and characterization of the *brittle-2* gene of maize, *Maydica*, 35 (1990) 317-322.

6 Bhave M R, Lawrence S, Barton C & Hannah L C, Identification and molecular characterization of *shrunken-2* cDNA clones of maize, *Plant Cell*, 2 (1990) 581-588.

7 Smith-White B J & Preiss J, Comparison of proteins of ADPglucose pyrophosphorylase from diverse sources, *J Mol Evol*, 34 (1992) 449-464.

8 Fu Y, Ballicora M A & Preiss J, Mutagenesis of the glucose-1-phosphate-binding site of potato tuber ADP-glucose pyrophosphorylase, *Plant Physiol*, 17 (1998) 989-996.

9 Okita T W, Nakata P A, Kathryn B, Smith-White B J & Preiss, J, Enhancement of plant productivity by manipulation of ADP-glucose pyrophosphorylase, in *Gene conservation and exploitation*, edited by J P Gustafson, R Appels & P Raven, (Springer, USA) 1993, 161-191.

10 Kleczkowski L A, Is leaf ADP-glucose pyrophosphorylase an allosteric enzyme?, *Biochim Biophys Acta*, 1476 (2000) 103-108.

11 Lin T P, Caspar T, Somerville C & Preiss J, Isolation and characterization of a starchless mutant of *Arabidopsis thaliana* (L.) Heynh lacking ADPglucose pyrophosphorylase activity, *Plant Physiol*, 86 (1988) 1131-1135.

12 Jin X, Ballicora M A, Preiss J & Geiger J H, Crystal structure of potato tuber ADP-glucose pyrophosphorylase, *EMBO J*, 24 (2005) 694-704.

13 Tuncel A, Kavakli I H & Keskin O, Insights into subunit interactions in the heterotetrameric structure of potato ADPglucose pyrophosphorylase, *Biophys J*, 95 (2008) 3628-3639.

14 Baris I, Tuncel A, Ozber N, Keskin O & Kavakli I H, Investigation of the interaction between the large and small subunits of potato ADP-glucose pyrophosphorylase, *PLoS Comput Biol*, 5 (2009) e1000546.

15 Danishuddin M, Chatrath R & Singh R, Insights of interaction between small and large subunits of ADP-glucose pyrophosphorylase from bread wheat (*Triticum aestivum* L.), *Bioinformation*, 6 (2011) 144-148.

16 Dawar C, Jain S & Kumar S, Insight into the 3D structure of ADP-glucose pyrophosphorylase from rice (*Oryza sativa* L.), *J Mol Model*, 19 (2013) 3351-3367.

17 Dosztanyi Z, Csizmok V, Tompa P & Simon I, IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics*, 21 (2005) 3433-3434.

18 Bhattacharya A, Tejero R & Montelione G T, Evaluating protein structures determined by structural genomics consortia, *Proteins*, 66 (2007) 778-795.

19 Luthy R, Bowie J U & Eisenberg D, Assessment of protein models with three-dimensional profiles, *Nature (Lond)*, 356 (1992) 83-85.

20 Laskowski R A, MacArthur M W & Thornton J M, PROCHECK: Validation of protein structure coordinates, in *International tables of crystallography, (Volume F) Crystallography of biological macromolecules*, edited by M G Rossmann & E D Arnold (Kluwer Academic Publishers, The Netherlands) 2001, 722-25.

21 Lovell S C, Davis I W, Arendall W B 3rd, de Bakker P I, Word J M *et al*, Structure validation by Calpha geometry: phi, psi and Cbeta deviation, *Proteins*, 50 (2003) 437-50.

22 Sippl M J, Recognition of errors in three-dimensional structures of proteins, *Proteins*, 17 (1993) 355-62.

23 Beckstein O, Fourrier A & Iorga B I, Prediction of hydration free energies for the SAMPL4 diverse set of compounds using molecular dynamics simulations with the OPLS-AA force field, *J Comput Aided Mol Des*, 28 (2014) 265-76.

24 Guo Z, Mohanty U, Noehre J, Sawyer T K, Sherman W *et al*, Probing the alpha-helical structural stability of stapled p53 peptides: Molecular dynamics simulations and analysis, *Chem Biol Drug Des*, 75 (2010) 348-359.

25 Bowers K J, Chow E, Xu H, Dror R O, Eastwood M P *et al*, Scalable algorithms for molecular dynamics simulations on commodity clusters, *Proc ACM/IEEE* 2006 *Conf on Supercomputing*, held at Tampa Florida USA, 11-17 November 2006, pp 43-44.

26 Pierce B G, Wiehe K, Hwang H, Kim B H, Vreven T *et al*, ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers, *Bioinformatics*, 30 (2014) 1771-1773.

27 Tovchigrechko A & Vakser I A, GRAMM-X public web server for protein-protein docking, *Nucleic Acids Res*, 34 (2006) W310-W314.

28 Chen R & Weng Z, Docking unbound proteins using shape complementarity, desolvation, and electrostatics, *Proteins*, 47 (2002) 281-294.

29 Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R *et al*, Integrating statistical pair potentials into protein complex prediction, *Proteins*, 69 (2007) 511-520

30 van der Spoel D, Lindahl E, Hess B, van Buuren A R, Apol E *et al*, GROMACS User manual version 3.2, 2004. [www.gromacs.org/]

31 Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T *et al*, PredictProtein—An open resource for online prediction of protein structural and functional features, *Nucleic Acids Res*, 42 (Web Server issue) (2014) W337-W343.

32 Rost B & EyrichV, EVA: Large-scale analysis of secondary structure prediction, *Proteins*, 45 (2001) 192-199.

33 Villand P, Olsen O A, Kilian A & Kleczkowski L A, ADP-glucose pyrophosphorylase large subunit cDNA from barley endosperm, *Plant Physiol*, 100 (1992) 1617-1618.

34 Kosloff M & Kolodny R, Sequence-similar, structure-dissimilar protein pairs in the PDB, *Proteins*, 71 (2008) 891-902.

35 Colubri A, Jha A K, Shen M Y, Sali A, Berry R S *et al*, Minimalist representations and the importance of nearest neighbor effects in protein folding simulations, *J Mol Biol*, 363 (2006) 835-857.

36 Melo F, Sanchez R & Sali A, Statistical potentials for fold assessment, *Protein Sci*, 11 (2002) 430-448.

37 Sternberg M J, Gabb H A & Jackson R M, Predictive docking of protein-protein and protein-DNA complexes, *Curr Opin Struct Biol*, 8 (1998) 250-256.

38 Vajda S, Sippl M, & Novotny J, Empirical potentials and functions for protein folding and binding, *Curr Opin Struct Biol*, 7 (1997) 222-228.

39 Wenfan H, Rigid body protein docking by fast fourier transform. Honor year project report, Department of Computer Science, School of Computing, National University of Singapore, 2005

40 Estrada J, Bernado P, Blackledge M & Sancho J, ProtSA: A web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble, *BMC Bioinformatics*, 10 (2009) 104.