Modeling of Potato ADP-Glucose Pyrophosphorylase

by

Aytuğ Tunçel

A Thesis Submitted to the
Graduate School of Sciences & Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science
in
Computational Sciences & Engineering

Koç University

August 2007

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Aytuğ Tunçel

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

İ. Halil Kavaklı, Ph. D. (Advisor)

Özlem Keskin, Ph. D. (Co-advisor)

Metin Türkay, Ph.D.

Mehmet Sayar, Ph.D.

Alkan Kabakçıoğlu, Ph.D.

Date: _____

ABSTRACT

In this thesis we modeled the native structure of ADP-glucose pyrophosphorylase (AGPase) and identified important amino acids that take role in subunit-subunit interactions. AGPase is composed of pairs of large (LS) and small (SS) subunits. Although SS was crystallized in a homotetrameric form recently, structure of LS and native enzyme has not been revealed yet. In this study, we first modeled the structure of LS by homology modeling. Then, we proposed three models for AGPase based on the crystal structure of homotetrameric SS. Six dimeric interaction types between the subunits were extracted as the representatives of the proposed models. We performed molecular dynamics simulations of these dimers. Trajectories obtained from the simulations were analyzed by molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method to calculate the binding free energies of subunits. The two most favorable interaction types were used to construct the AGPase structure. Hot-spot residues that are critical for the subunit-subunit interactions were then identified by molecular mechanics generalized born surface area (MM-GBSA) method using free energy decomposition scheme. Importance of residues obtained from computational data was further experimentally verified. The residues were mutated to alanine by site directed mutagenesis and the effects of the mutations were assessed by yeast-two hybrid system. We believe that the AGPase model and the key residues defined here can be used as a guide for further studies that involve engineering a more stable and efficient form of the enzyme.

ÖZET

Bu tezde ADP-glukoz pirofosforilaz (AGPaz) enziminin doğal yapısını ve enzimin yapıtaşları arasındaki etkileşimlerde önemli rolü olan amino asitleri belirledik. AGPaz büyük (BY) ve küçük (KY) ikili yapıtaşlarından oluşur. KY yakınlarda türdeş dörtlü yapı halinde kristalleştirildiği halde BY'nın ve doğal AGPaz enziminin yapısı henüz gösterilemedi. Bu çalışmada önce BY'nın yapısını benzşim modellemesi yöntemini kullanarak modelledik. Daha sonra AGPaz için KY'nın türdeş dörtlü kristal yapısına dayanarak üç model öne sürdük. Yapıtaşları arasındaki altı adet ikili etkileşim çeşidi öne sürülen modellerin temsilleri olarak seçip çıkarıldı. Bu ikili etkileşimlerin moleküler dinamik simulasyonlarını gerçekleştirdik. Simulasyonlardan elde edilen yörüngeler moleküler mekanik Poissson-Boltzmann yüzey alan (MM-PBYA) metodu ile yapıtaşları arasındaki bağlanma enerjisinin hesaplanması ile analiz edildi. En uygun iki etkileşim çeşidi AGPaz' ın yapılandırılmasında kullanıldı. Yapıtaşları arasındaki etkileşimde kritik olan önemli-nokta amino asitleri moleküler mekanik genellenmiş doğuş yüzey alan metodu (MM-GDYA) kullanılarak ve enerji dağılımlandırılması yöntemi uygulanarak belirlendi. Hesaplamalı veriden elde edilen amino asitlerin önemi deneysel olarak doğrulandı. Bu amino asitler konum yönlendirilmiş değitirme yöntemi ile alanine çevrildi ve mutasyonların etkisi maya ikili-hibrid sistemi ile test edildi. AGPaz modelimizin ve burada tanımlanmış önemli-nokta amino asitlerin daha sonraki enzimin kararlı ve verimli formlarının geliştirilmesinde önemli olacağını düşünüyoruz.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisors Assist. Prof. İ. Halil Kavaklı and Assist. Prof. Özlem Keskin for their invaluable support and patience during the completion of this thesis. I am grateful to them for teaching me how to make research and feel fortunate to be able to work with them.

I would also like to thank my thesis committee members, Assist. Prof. Metin Türkay, Assist. Prof. Mehmet Sayar and Assist. Prof. Alkan Kabakçıoğlu for their valuable considerations and contributions.

Special thanks to my colleagues for helping me to develop my skills and providing their kind friendship at anytime. I would especially thank to M. Emre Özdemir, Evrim Besray Ünal, Seda Kılınç, Natali Özber, Hande Asımgil and Dr. Şule Özdaş.

Lastly, I wish to thank my family for everything to whom I dedicate this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

AGP(ase)	ADP-glucose pyrophosphorylase
3-PGA	3-phosphoglycric acid
PPi	pyrophosphate
LS	Large subunit
SS	Small subunit
NVT	constant atom number, volume and temperature
NPT	constant atom number, pressure and temperature
MM-PBSA	molecular mechanics Poisson-Boltzmann surface area
MM-GBSA	molecular mechanics generalized born surface area
$\Delta G_{binding}$	binding free energy
$G_{complex}$, $G_{receptor}$, G_{ligand}	energies of complex, receptor and ligand molecule
E_{MM}	total mechanical energy in gas phase
G_{sol}	solvation free energy
E_{elec}	Columbic energies
E_{vdw}	van der Waals energies
E_{int}	internal strain energies
G_{polar}	polar solvation energy
$G_{non\text{-}polar}$	non-polar solvation energy
G_{gas}	gas phase energy
G_{elec}	sum of E_{elec} and G_{polar}
SASA	solvent accessible surface area
S_{vib}	vibrational entropy
S_{rot}	rotational entropy
S_{trnas}	translational entropy

RMSD	root mean square deviation
RMSF	root mean square fluctuation
WT	wild type
EV	empty vector
$\vec{F}_i$	force applied on an atom *i*
m_i	mass of atom *i*
$\vec{a}_i$	acceleration of atom *i*
$\vec{\nabla}$	gradient
$\langle E_{kin} \rangle$	average kinetic energy
k_B	Boltzmann constant
l	length of arc
z	perpendicular distance

Chapter 1

INTRODUCTION

Starch is an important carbohydrate being the primary energy source for plants. In addition to its role as an indispensable nutritional source it is also used in several industrial applications such as paper or textile manufacturing. Starch biosynthesis occurs by the participation of three main enzymes: ADP-glucose pyrophosphorylase (AGPase), starch synthase, and branching enzymes (1-3). Total product of starch can be modified either qualitatively or quantitatively. While the starch synthase and branching enzymes are the main targets to alter the quality in a desired way, AGPase can be engineered to increase the amount of yield. Such an enhancement in plant productivity by modification of AGPase has many economical and environmental benefits. The cost of production can be decreased by increasing the harvest obtained from a fixed plant population. Engineered crops can also help decrease the agricultural land required to yield the same amount of starch compared to their natural variants.

AGPase is the first enzyme in starch biosynthesis that catalyzes the conversion of Glc-1-P and ATP to ADPglucose and pyrophosphate (PPi) using Mg^{2+}. Then ADP-glucose is utilized by other starch synthase isoforms for the synthesis of polyglucanes. Ample evidence has been obtained showing that the AGPase catalyzed reaction is a pivotal regulatory control point in the biosynthesis pathway in higher plants and its increased activity correlates with increased starch production (4-7). AGPase from higher plants has a heterotetrameric structure ($\alpha_2\beta_2$) composed of pairs of small (SS) and large (LS) subunits

encoded by two different genes (8, 9). Since the enzyme has an oligomeric structure one way to increase its activity, thus the total amount of starch product, would be to enhance the subunit interactions. Validity of such an approach has been proven in maize endosperm AGPase by enhancing the heat stability of enzyme using random mutagenesis in combination with yeast two hybrid methods (10).

Solanum tuberosum potato tuber AGPase has been widely studied as the model enzyme in higher plants because of its relatively efficient recombinant expression and stability. Recently, three dimensional structure of the potato tuber AGPase SS was determined by x-ray crystallography in homotetrameric (α_4) form (11) (Figure1.1).

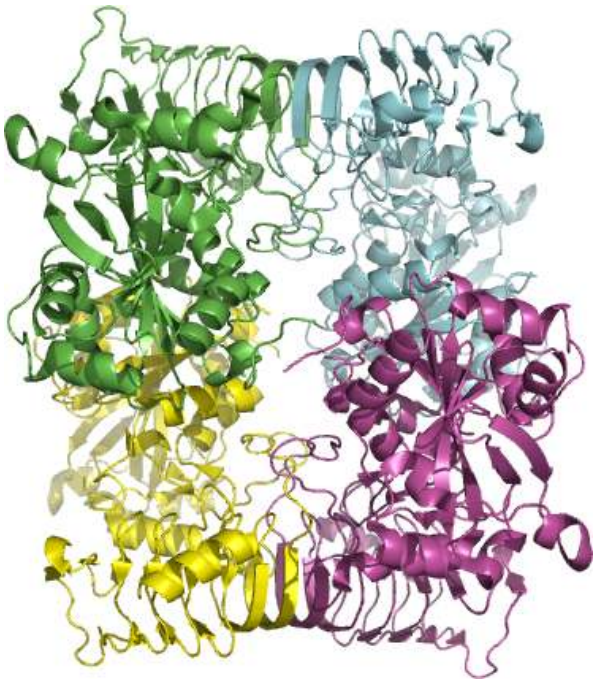


Figure 1.1: Crystal structure of homotetrameric SS composed of four identical chains, chain A (green), chain B (blue), chain C (violet) and chain D (yellow).

This recombinant homotetrameric enzyme elucidates structural information about one type of the subunits and gives structural insights into the heterotetrameric enzyme. However, the LS and heterotetrameric AGPase ($\alpha_2\beta_2$) could not be crystallized yet. This is due to the difficulty of obtaining AGPase in its highly pure and stable form. Nevertheless, it is of crucial importance to elucidate the native AGPase structure and key residues that take role in subunit-subunit interactions for obtaining more detailed information how this enzyme works. Understanding the structure will enable to manipulate the native enzyme and modified AGPase may be utilized to enhance the plant productivity.

In this thesis, first, we have modeled the structure of potato AGPase LS using homology modeling. Then, a model structure was proposed for the native heterotetrameric AGPase (both LS and SS together) based on the binding free energy calculations between subunits by molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method (12, 13). Important amino acids that mediate the subunit-subunit interactions were identified by applying the free energy decomposition scheme with molecular mechanics generalized born surface area (MM-GBSA) approach (12, 13). Finally, these key residues obtained from computational results were mutated to alanine by site directed mutagenesis and it was verified using yeast two hybrid method (14) that they effect the interactions between the subunits. Results show that native AGPase is formed by heterodimeric interactions between the subunits. That is, LS and SS interact with each other both laterally and longitudinally in the proposed model of AGPase. Important amino acids in LS that mediate these interactions were found to be Arg28, Ile322 and Ile323. Also, Arg78 and Gln100 were classified as key residues critical for homodimeric SS association. This study establishes the groundwork for understanding the subunit-subunit interactions and the native structure of AGPase for the first time.

Chapter 2 extends the background information about AGPase and presents a literature review about the computational and experimental methods used in this study.

Chapter 3 describes the materials and methods and their applications for structure prediction of LS, proposition of native AGPase model, identification and verification of important amino acids. Also extended theory about the MM-PB(GB)SA methods is provided in this chapter, since the method plays a central role in this study.

Computational results regarding the structure prediction of LS and model proposition for native AGPase are given in Chapter 4.

Both computational and experimental results about the important amino acids that play critical role in subunit-subunit interactions are discussed in Chapter 5.

The thesis is concluded with a brief summary of the study.

Chapter 2

LITERATURE REVIEW

2.1 ADP-glucose pyrophosphorylase

Molecular weights of AGPases change between 200 and 240 kDa depending on plant species. Particularly, apparent molecular weights of potato SS and LS are shown to be 50 and 51 kDa, respectively (9). Comparison of the primary amino acid sequences of small subunits from different plant species has 85-95% identity. On the other hand, such comparison between the large subunits of different plant species at amino acid level has 50-60% identity. LS and SS amino acid sequences share relatively less but still significant homology. For example, there is a 53% sequence identity between the potato tuber AGPase small and large subunits (15). Such a high homology between different subunits suggests that these two genes might have evolved from a common ancestor, most probably by a gene duplication event (16). Almost all plant AGPases, with a few exceptions, are regulated by 3-phosphoglycric acid to inorganic phosphate (3-PGA/Pi) ratio in cells with 3PGA being the main activator while Pi is the main inhibitor (5, 6, 16, 17). The enzyme is also subject to post-translational redox modification by oxidation/reduction of the Cys12 residues in the small subunits. When oxidized, a disulfide bond forms between the Cys12 residues, which covalently links the small subunits. In the reduced state, the enzyme shows more inter-subunit flexibility, has higher affinity for its substrates and is more sensitive to 3-PGA activation and more resistant to Pi inhibition (11, 18, 19).

Different approaches have been utilized in attempts to decipher the role of the two subunit types in higher plant AGPase function. Genetic, mutagenesis and biochemical studies suggest that, while SS has both catalytic and regulatory activities, LS is mainly responsible for modulating the allosteric regulatory properties of SS in the heterotetrameric enzyme (19-24). The LS is also incapable of forming a catalytically active oligomeric structure in the absence of SS, but the SS can assemble into a homotetramer which possesses catalytic properties (8, 21). However, this so called recombinant α_4 enzyme composed of only SSs requires elevated levels of 3-PGA for activation and is more sensitive to Pi inhibition. This suggests that LS is required for optimal catalysis of the heterotetrameric enzyme (21, 25).

Apart from the above findings, which assign subunit functions, specific regions from both LS and SS were found to be important for subunit association and enzyme stability. Laughlin *et al.* (26) showed that deletion of a 19 amino acid segment at C-terminus of either subunit results in a decrease in enzyme activity due to inability of subunits to assemble into a heterotetrameric enzyme. They also identified a region composed of 28 residues at N-terminus of LS that is essential for stability of the enzyme. In addition, Cross *et al.* (27) using chimeric maize/potato small subunits in *E.coli* system, found a polymorphic motif in SS critical for subunit interaction. They have concluded that a 55-amino acid region between the residues 322-376 directly interacts with LS and significantly contributes to the overall enzyme stability.

2.2 Homology Modeling

Homology modeling is one of the most accurate and reliable protein structure prediction methods (28). In this method, the three-dimensional structure of a given protein

sequence (target) is predicted based on an alignment to one or more known protein structures (templates). This involves mainly four steps (29):

1. Identification of homologs that can be used as template(s)
2. Alignment of the target sequence to template(s) for modeling
3. Building a model for the target based on the information from the alignment(s)
4. Evaluation of the model

The most critical steps of this modeling process that affects the accuracy of the end model are selection of the correct template structure(s) and alignment between the target and template sequences (28). The alignment step is considered to be almost completely correct if the sequence identity between target and template is above 40% (30-32). High levels of model accuracy is achieved when the target and template sequences share at least 50% identity and the quality of these models can approach a low resolution (3 Å) x-ray structure (33). These models can also be used for protein-protein docking (34, 35).

2.3 Molecular Dynamics Simulations

Since the molecular dynamics (MD) simulation of liquid water (36), the method has become a powerful tool for exploring the structural and dynamic behaviors of biological systems. They can give detailed information on the fluctuations and conformational changes of proteins and nucleic acids. Ma *et al.* (37) used the method to investigate the conformational changes in the functional mechanism of GroEL. They were able to find a pathway between the open and closed conformations of GroEL-GroES system which is impossible to determine experimentally. Several nanoseconds simulations of acetylcholinesterase revealed dynamic features relevant to function which were not anticipated from experimental studies (38). The method has also become popular for attacking the protein folding and unfolding problems (39). With the development of faster

computers it is now possible to visualize true biological events. De Groot *et al.* (40) were able to visualize migration of a water molecule through a model of the aquaporin channel with multiple simulations of 10 ns duration.

2.4 MM-PB(GB)SA

In recent years, computational methods are being employed to gain insights in protein-protein or protein-ligand interactions and to guide experimental studies (41). These routines include the estimation of absolute binding free energies between the binding partners or identification of hot spot residues in protein interfaces. Among the methods such as application of empirical free energy functions (42, 43) , Free Energy Perturbation (FEB), Thermodynamic Integration (TI) and MM-PB(GB)SA (12, 13) the latter method provides a preferable alternative in terms of accuracy/computational cost ratio. MM-PB(GB)SA methodologies have been widely used to predict protein structures (44), to estimate the binding energies of protein-ligand (12, 45-48), and protein-protein (49-53) interactions. MM-PB(GB)SA methods couples an explicit solvent molecular dynamics (MD) simulation with an implicit solvent free energy calculation. In this method, binding free energy of a complex macromolecule is calculated by taking snapshots from an MD trajectory and computing the average free energy difference between the complex and the binding partners using these snapshots.

MM-PBSA method also allows the so called “computational alanine scanning” (CAS) scheme to detect the hot spot residues for probing protein-protein interactions (54, 55). The idea behind the definition of a hot spot residue lies in the fact that a single specific amino acid in the interface of a protein complex might have a dominant contribution to the binding process (56, 57). Indeed, they can be defined as those residues which decrease the binding free energy at least 2 kcal/mol when mutated to alanine (56, 58). Systematic search

for the hot-spots by CAS methodology might be time consuming if separate simulations are considered for each of the different mutations. Conversely, application of the method on a single simulation might not be sufficient for conformational sampling.

MM-GBSA method, the closely related and computationally less expensive approach to MM-PBSA, permits an easier means of decomposing the free energies at atomic level compared to MM-PBSA. Free energy decomposition (FED) provides an alternative approach to CAS for identification of hot spots. This is because it does not require separate simulations for each mutation as in the case of CAS and avoids perturbations which might cause various destabilization problems in protein structure as reviewed in DeLano *et al*. (59). Use of FED approach for identifying hot-spots has previously been reported in other studies to yield results that are in good agreement with the experimental alanine scanning data (50, 51, 53, 60).

2.5 Yeast Two Hybrid Method

The yeast two-hybrid method is a powerful experimental tool for investigating protein-protein interactions. The method was first described by Fields and Song (14) and has been modified and improved (61, 62). In this system an interaction between two proteins is tested by fusing one of the proteins to DNA activation domain of a eukaryotic transcription factor and fusing the other protein to DNA binding domain. The interaction is then detected by expression of reporter genes. The method can either be used in large scale screening of protein-protein interactions or to investigate the effects of specific mutations. This system has many advantages. First of all it is easy to obtain the target genes using a cDNA library. Furthermore, it is highly sensitive to detect low abundant proteins and can also detect weak and transient interactions. However, false negative or false positive results can also be

detected. Indeed, up to fifty percent of the interactions identified using this system might be estimated to be false positives (63, 64).

Yeast-two hybrid method has been used to study protein-protein interactions in many areas such as circadian clock (65), phytohormone signaling (66) and floral development (67). Recently, Uetz *et al.* (68) used the method to screen 192 yeast bait proteins against 6000 predicted proteins and identified 281 interacting protein partners from 87 bait proteins that are involved in protein-protein interactions. Specifically, it has also been used to investigate the subunit interactions in maize endosperm AGPase by Greene *et al.* (10, 69).

Chapter 3

MATERIALS & METHODS

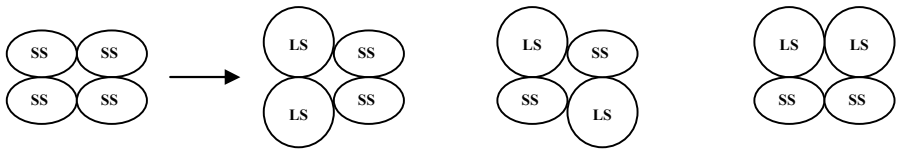
3.1 Homology Modeling of the Large Subunit and Construction of the Heterotetrameric Models

The sequence alignments of the LS and SS using CLUSTALW (70) (1.83) with default parameters reveal that there is a 53% of sequence identity between LS and SS as mentioned in introduction. SWISS–MODEL homology modeling server (first-approach method) (71-73) was used to construct the three dimensional structure of LS. This method was developed for cases in which the target and template sequences share at least 50% sequence identity. Reliability of the first-approach method was also assessed with a set of structurally known 1200 proteins against templates that share sequence identities between 25-95%. In cases where the template and target have 50-59% identity, a deviation less than 3Å from the experimental structures was achieved for 79% of the sequences (32, 74).

When the near-full-length cDNA clones of potato tuber AGPase large and small subunits were compared, LS and SS were found to consist of 470 and 521 residues respectively (15). Both subunits carry amyloplast target sequences at their N-terminus regions. In the crystal structure of homotetrameric SS (pdb id: 1YP2), one chain consists of 442 amino acids, thereby excluding the target sequence of the first 79 residues at the N-terminus. 29 residues at the N-terminus of LS were removed before submission to the SWISS–MODEL server in order to exclude the random coil fragment at this region and to

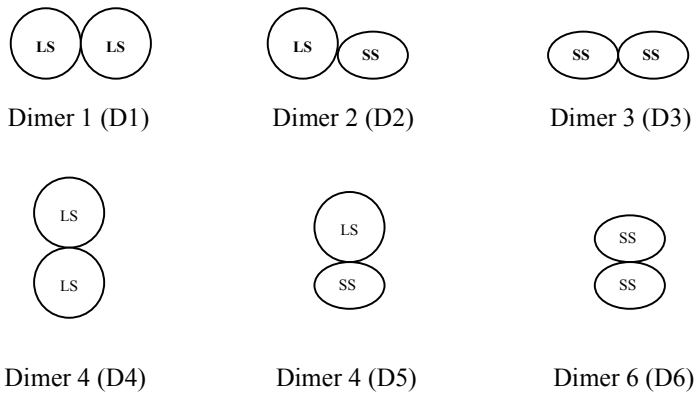
achieve a better global superimposition with SS. It should be noted that the numbering of SS and LS amino acids in this study is different than the literature. In the crystal structure, SS (1YP2_C) starts with Gln at position 10. It also starts with the same residue but at position 1 in this study. mRNA coding sequence for LS is composed of 446 residues (15), however the LS in this study is composed of 441 residues which makes a five residue difference compared to literature. The crystal homotetrameric structure of SS contains structural gaps where C-chain (1YP2_C) is the most complete one. In the crystal structure of this chain, fragments between the residues 27-32 and 91-98 were missing. In order to fill these gaps 1YP2_C was also submitted to SWISS-MODEL and used as the model SS in further calculations. After the three dimensional structure of LS was generated; the homotetrameric structure was used as a template to construct the heterotetrameric AGPase with two large and two small subunits. Three models were proposed: A schematic presentation of the three proposed models (3.1.b, 3.1.c, 3.1.d) is illustrated in Figure 3.1 together with the crystal model of the homotetrameric SS (3.1.a). Each model was built by superimposing the large (or small) subunit with the template homotetrameric structure using backbone atoms. Models were named as Model-1 (3.1.b), Model-2, (3.1.c), Model-3 (3.1.d).

In order to predict the model for the heterotetrameric AGPase structure among the proposed models, different computational approaches were followed (see below). Since the enzyme has a 2-fold symmetry, six different dimeric (either up-down or side-by-side) interactions were possible. Each possible model was subjected to additional analysis (Figure 3.2). First, molecular dynamics (MD) simulations were performed for each of these possible dimers. Trajectories obtained from MD simulations of each dimer were then analyzed by MM-PBSA method in order to identify the most favorable interactions. The most favorable two interaction types, one from set 1 and the other from set 2, were used to propose the native heterotetrameric AGPase structure.



1.a 1.b 1.c 1.d

Figure 3.1: Schematic presentation of the crystal structure of homotetrameric SS and proposed models. For construction of each model each large and small subunits were superimposed with the corresponding chain in the crystal structure and the original SS chains were than deleted. (1.a) Each chain corresponds to a SS taken from the x-ray structure. Homodimers are in right-left (1.b), diagonal (1.c) and up-down (1.d) interactions in Model-1, Model-2 and Model-3 respectively.



Dimer 1 (D1) Dimer 2 (D2) Dimer 3 (D3)

Dimer 4 (D4) Dimer 4 (D5) Dimer 6 (D6)

Figure 3.2: Schematic presentation of dimeric interactions between the subunits that constitute the heterotetrameric models. LS and SS are composed of 441 and 442 residues respectively. Set 1 contains the D1, D2, D3 and set 2 contains the D4, D5, D6.

3.2 Molecular Dynamics Simulations

MD simulations are performed based on the Newton's law of motion, Eq (3.1):

$$\vec{F}_i = m_i \vec{a}_i \tag{3.1}$$

where $\vec{F}_i$ represents the force applied on an atom i, m_i is the mass of atom i and $\vec{a}_i$ is the acceleration of that atom. Simulation procedure follows a deterministic path which means that once the initial positions, velocities and accelerations of each atom in the system are known the trajectory of all the atoms can be calculated at a given time. Initial coordinates of molecules are generally taken from x-ray crystal structures or Nuclear Magnetic Resonance (NMR) images. The acceleration of an atom is given by the gradient of the potential energy, Eq (3.2):

$$\vec{a}_i = -\frac{1}{m_i} \vec{\nabla} \mathrm{E}_i \tag{3.2}$$

where $\vec{a}_i$ is the acceleration of atom i, m_i is the mass of atom i and E_i is the potential energy. At thermal equilibrium, the initial velocity of each atom is assigned according to the kinetic energy of the system as shown in Eq (3.3) assuming that the velocity components have a mean value of zero in a random Gaussian distribution:

$$\left\langle E_{kin} \right\rangle = \frac{1}{2} \sum m_i \upsilon_i^2 = \frac{1}{2} (3\mathrm{N}) \mathrm{k_B T} \tag{3.3}$$

where E_{kin} is the kinetic energy, m_i is the mass of atom *i*, υ_i is the velocity of atom *i*, T is temperature, k_B is Boltzmann constant and N is the number of atoms in the system. Calculation of the above equations at each time step determines the trajectory of the atoms. However, since there is no analytical solution to the equations of motion they are solved numerically by using several algorithms such as Verlet algorithm, Leap-frog algorithm, Velocity Verlet or Beeman's algorithm. Below is the procedure of MD simulations in this study.

All of the six dimers (D1-D6) were solvated in different rectangular boxes including TIP3P water molecules (75). Distances between the edge of the water boxes and the closest atom of solutes were at least 10 Å. Counter ions were added in order to neutralize the systems. All the histidine residues were treated as carrying +1 charge at their Nε atoms. Simulations were performed with the NAMD (76) software using the Kollman *et al* (77) *parm96* force field and periodic boundary conditions (PBC) (78). A direct-space non-bonded cutoff value of 9 Å was used with Particle Mesh Ewald (PME) method (79) to treat the long range electrostatic interactions. SHAKE algorithm (80) was applied to water molecules to treat them as rigid bodies and to hydrogen atoms to constrain their movements. Langevin piston Nose-Hoover method (81, 82) was used to keep the pressure constant. Time step of all simulations was 2 fs. Systems were minimized by conjugate gradient method for 10^4 steps keeping the backbone atoms of solutes fixed followed by an additional 10^4 steps with relaxed backbone atoms. The systems were then gradually heated from 0K to 300K in 150 ps using NVT ensembles in which the Cα atoms of the solutes were restrained by applying 2 kcal mol^{-1} $Å^{-2}$ force constants. Isothermal-isobaric ensembles (NPT) were then applied for 80 ps during which the restraints on Cα atoms were removed gradually with an additional 100 ps of equilibration simulation. Subsequent NPT simulations were performed for 5 ns and the last 3ns of the simulations were analyzed for binding free energy calculations by MM-PBSA method.

3.3 Identification of Interface Residues

Snapshots taken from the last 3 ns of the simulations (150 snapshots with 20 ps intervals) were separated as complex, receptor and ligand structures. Interface residues for every snapshot were identified based on the implementation of Lee and Richards method (83) using NACCESS program (84). The NACCESS program calculates the atomic accessible surface defined by rolling a probe of given size around a van der Waals surface. The Lee and Richards method works by taking thin Z-slices through the molecule and calculating the exposed arc lengths for each atom in each slice, and then summing the arc lengths to the final area ($Å^2$) over all z-values according to the following equations:

$$ASA = R / \sqrt{(r^2 - z^2)} \times \rho \times l \quad (3.4)$$

$$\rho = \Delta z / 2 + \Delta' z \quad (3.5)$$

where l is the length of the arc drawn in a slice, z is the perpendicular distance from the center of the sphere to the section, Δz is the space between the slices and $\Delta' z$ is the smaller of $\Delta z/2$ or $\Delta' z$. A schematic presentation of the method can be seen in Figure 3.3. Probe radius used for calculation of the atomic accessible surface area was taken 1.4 Å together with a z-slice value of 0.05 Å. This z-slice value provides an optimization for accuracy and speed. Hydrogen atoms were not included during the calculations. Set of interface residues were completed by a two step approach. First, residues that show > 1Å decrease in their accessible surface area upon complexation were considered as part of the initial interface set. Second, residues from the initial set that hold the above criteria at least for 80% (120 snapshots) of the last 3 ns part of the simulations were chosen as the actual set of interface residues.

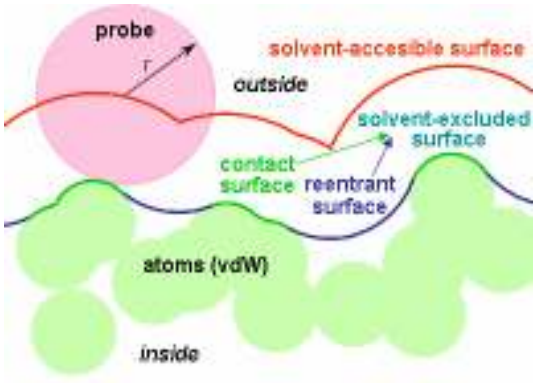


Figure 3.3: Simplified diagram showing the calculation of atomic solvent accessible surface area.

3.4 Binding Energy Calculations

Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) or a related approach of Generalized Born Surface Area (MM-GBSA) methods (12, 13) can be used to calculate the binding free energy of molecules in equilibrium state. In these approaches, binding free energy of a complex is calculated by taking snapshots from a molecular dynamics trajectory and computing the average energy of these snapshots according to the formula in Eq (3.6):

$$\Delta G_{binding} = G_{complex} - G_{receptor} - G_{ligand} \quad (3.6)$$

where $G_{complex}$, $G_{receptor}$, G_{ligand} are the energies of the complex, receptor and ligand respectively. Snapshots of the complex, receptor and ligand can either be taken from separate trajectories or a single trajectory in which the coordinates of receptor and ligand

are extracted from the complex molecule in the latter approach. Energy of a molecule in Eq (3.6) can be represented as shown in the following equation:

$$G_{mol} = E_{MM} + G_{sol} - TS \quad (3.7)$$

where E_{MM} is the total mechanical energy of the molecule in gas phase, G_{sol} is the solvation free energy and TS is the entropic term. Each term in Eq (3.7) can be divided into individual energetic components as shown below:

$$E_{MM} = E_{elec} + E_{vdw} + E_{int} \quad (3.8)$$

In Eq (3.8) E_{MM} represents the bonded and non-bonded interactions as a sum of electrostatic (columbic), van der Waals (Lennard-Jones) and internal strain (bonds, angles and dihedrals) energies. This term is calculated by classical molecular-mechanics methods using standard force fields such as Kollman *et al*. (77) force field.

Solvation free energy of a molecule is calculated as the sum of a polar and a non-polar term:

$$G_{sol} = G_{polar} + G_{non\text{-}polar} \quad (3.9)$$

where electrostatic contribution to the solvation energy (G_{polar}) is computed in a continuum solvent environment by solving either the Poisson-Boltzmann equation (85), or using a generalized born method. Non-polar solvation energy ($G_{non\text{-}polar}$), which is considered to be the sum of a solute-solvent van der Waals interaction and solvent-solvent cavity formation energy, is approximated by using an empirical formula such as $G_{non\text{-}polar} = \alpha \times SASA + \beta$. According to this formula, non-polar solvation energy of a molecule is proportional to the

solvent accessible surface area (SASA) of that molecule in a solvent, where α and β values are the constants (86, 87).

The entropic term in Eq (3.7) is considered as the summation of vibrational, rotational and translational contributions where vibrational term is calculated by normal-mode analysis or quasi-harmonic analysis:

$$- T\Delta S = - T\Delta S_{vib} - T\Delta S_{rot} - T\Delta S_{trnas} \quad (3.10)$$

Since the calculation of entropic contribution is computationally expensive, this term can be omitted if qualitative results, rather than quantitative, are considered to be more important. This is also true for different ligands that show similar binding affinities and modes to a given receptor (48, 51, 88-90).

The last 3 ns parts of the simulations for each dimeric interaction between the LSs and SSs (Figure 3.2) were analyzed by MM-PBSA method as implemented in AMBER8 package (91). The trajectories were post processed in order to strip off the water molecules and counter ions before the calculations. 150 snapshots with 20 ps intervals were extracted for each complex, receptor and ligand structures from single trajectories. In single trajectory approach internal strain energies from complex are always canceled by the strain energies from receptor and ligand. This is due to the fact that coordinates of both the receptor and ligand are taken from the trajectory of complex. This approach is valid provided that the receptor and ligand do not show any significant amount of conformational changes upon complex formation (12, 46, 48, 53-55, 92, 93). In all of the calculations LS was treated as the receptor and the SS as the ligand. Gas phase energies (E_{MM}) of the proteins were calculated by the SANDER module applying no cutoff value for non-bonded interactions. The electrostatic contribution to the desolvation free energy was computed by solving the finite difference Poisson-Boltzmann equation using the PBSA module of

AMBER8 with PARSE parameter set (94). Dielectric constants for the solute and solvent were taken as 1 and 80, respectively; and the solvent probe radius was adjusted to 1.4 Å. Non-polar solvation energies were calculated according to SASA dependent empirical formula by using the LCPO method (95) implemented in AMBER8. The surface tension parameter, α, and the β values were taken as 0.0542 and 0.92, respectively (94). Entropic contributions were not computed since the main interest was calculating relative binding energies between the subunits and the subunits have similar binding modes because of the high structural similarity. The dimers with the most favorable interaction energies (one from set 1 and one from set 2 in Figure 3.2) were used to construct the heterotetrameric AGPase model.

3.5 MM-GBSA Analysis

In this study MM-GBSA (12, 13) method was mainly used to identify the critical amino acids taking role in subunit-subunit interactions. The detailed theory about the method was introduced in section 3.4, and the same parameters were also applied in these calculations. However there are some minor differences about the calculation of ΔG_{polar} term and the choice of parameters. This time polar part of the solvation free energy ($G_{sol} = G_{polar} + G_{non\text{-}polar}$) was calculated using the generalized born scheme in AMBER8 (igb=2) with the modified Bondi radii (mbondi2) (96) which is appropriate for macromolecules such as proteins. Non-polar solvation energy was again approximated by the SASA dependent formula; $G_{non\text{-}polar} = \gamma \times SASA$ and applying a recursive algorithm which uses an icosahedron to estimate the SASA of an atom. The constant of γ was taken 0.005 kcal·Å^{-2} (49, 94, 96). The parameters were chosen according to previously published studies to keep consistency with literature. Residues in interfaces of the subunits that showed at least 3 kcal/mol energy decrease, upon complexation, according to the per-residue free energy

decomposition were considered as hot-spots. Molecular dynamics simulations for MM-GBSA analysis were extended to 8 ns and last 4 ns parts, 200 snapshots with 20 ps intervals, were taken for the corresponding calculations. Interface residues were again identified by the method explained in section 3.3, but this time from 200 snapshots.

3.6 Cloning and Site Directed Mutagenesis of Large and Small Subunits

The LS and SS cDNAs of potato were obtained by PCR using primers, which contain NcoI and BamH I sites. After obtaining PCR products, they were subjected to restriction enzymes and cloned into pGADT7Leu+ and pGBKT7Trp+ vectors. *E. coli* DH5α host strain was used during the manipulation of plasmids.

Site directed mutagenesis for the specified hot spot residues were introduced by PCR. 50 ng of plasmid samples and 30 pmol of each primer were used per reaction mixture together with 2.5 unit Pfu DNA polymerase. Conditions for the14 cycles of amplification reaction were 95 ºC for 30s, 50 ºC for 30 s and 68 ºC for 14 min. Before the first cycle reaction mixtures were kept at 95 ºC for 4 min and at the end of the 14th cycle an additional 68 ºC extension period was applied for 10 min. Samples were then treated by DpnI restriction enzyme to remove the template DNA. Mutagenesis results were verified by DNA sequencing throughout Iontek (Istanbul-Turkey) Company.

3.7 Yeast Manipulations

In yeast-two hybrid system proteins under question are fused to two separate vectors, prey and bait, which also separately code for the distinct domains of eukaryotic transcription factors. While one of the domains is responsible for binding to DNA (DNA binding domain [BD], inserted into bait vector) such as GAL4 or LexA, the other domain

activates the transcription (DNA activation domain [AD], inserted into prey vector) such as GAL4, B42 or VP16. The vectors also separately contain genes that code for the amino acids leucine or tryptophan for growth selection of the cells. Additionally, they possess the *LacZ* gene which is used for the color detection.

In case of the interaction of bait and prey proteins BD and AD also get linked to each other and trigger the expression of reporter genes. These reporter genes either result in color development or provide additional growth selection by the expression of adenine (Ade) or histidine. If the bait and prey proteins do not interact yeast cells can not survive in an Ade^- and His^- medium and no color development is observed. A schematic presentation of the method can be seen in Figure 3.4.

Transformation of constructs into the yeast strain AH109 followed the protocol outlined by Clontech (Palo Alto, CA). AH109 yeast cells containing both plasmids were selected on a synthetic minimal medium containing 6.7 g/L yeast nitrogen base (Difco, Detroit, MI) without amino acids, 2% glucose, amino acid dropout supplement without Leu and Trp (Clontech), and 20 g/L agar (plates only).

The constructs containing wild type (WT) and mutant forms of LS and SS were sequentially transformed into the cells as in the following procedure. First, AH109 cells were separately transformed with pGADT7-SSWT and pGADT7 empty vector (EV) and selected by cultivating for seven days on the appropriate SD medium lacking Leu (agar only). Selected colonies were then retransformed with the counter construct either the WT or mutant sequences. In all of the combinations SS and LS were inserted into pGADT7Leu+ and pGBKT7Trp+ respectively. These cells were then grown for seven days on the SD agar media lacking both Trp and Leu. 40 colonies from each of the agar plates were selected and transferred into the interaction medium (Trp-, Leu-, His-). Number of colonies for mutated and control groups were then compared to investigate the effect of the mutations on subunit-subunit interactions.

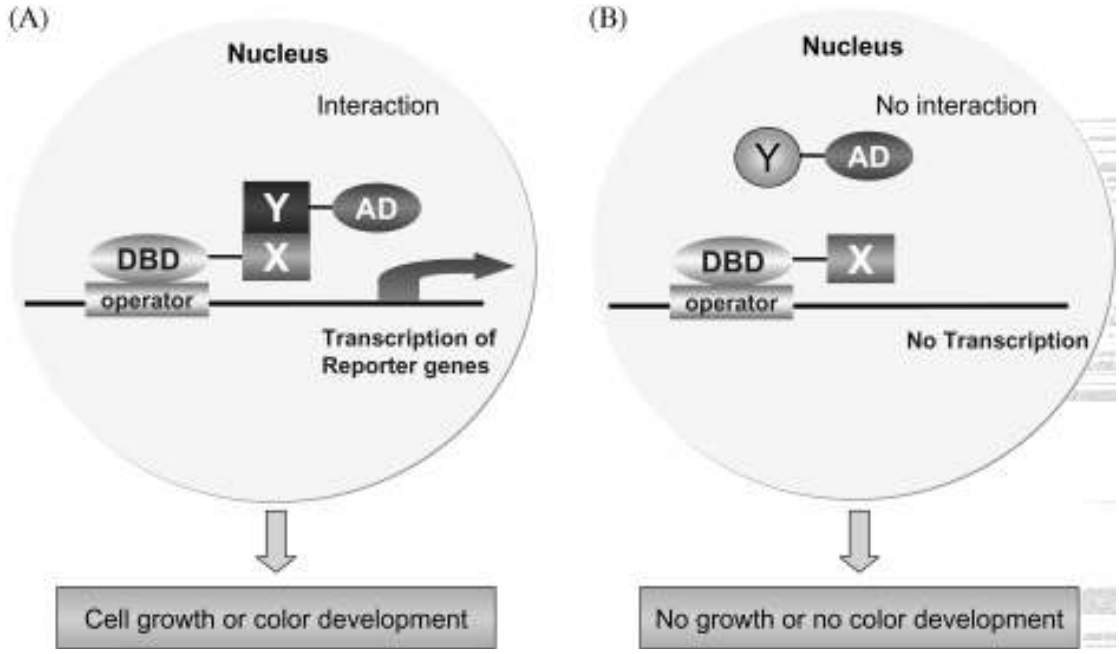


Figure 3.4: Principle of yeast-two hybrid method. (A) X and Y proteins are interacting thereby providing the BD and AD to be linked. This linkage triggers the transcription of reporter genes and color development and cell growth occurs. (B) The proteins are not interacting hence no transcription occurs. Therefore, the cells can not synthesize the amino acids that are absent in the medium and no growth or color development occurs.

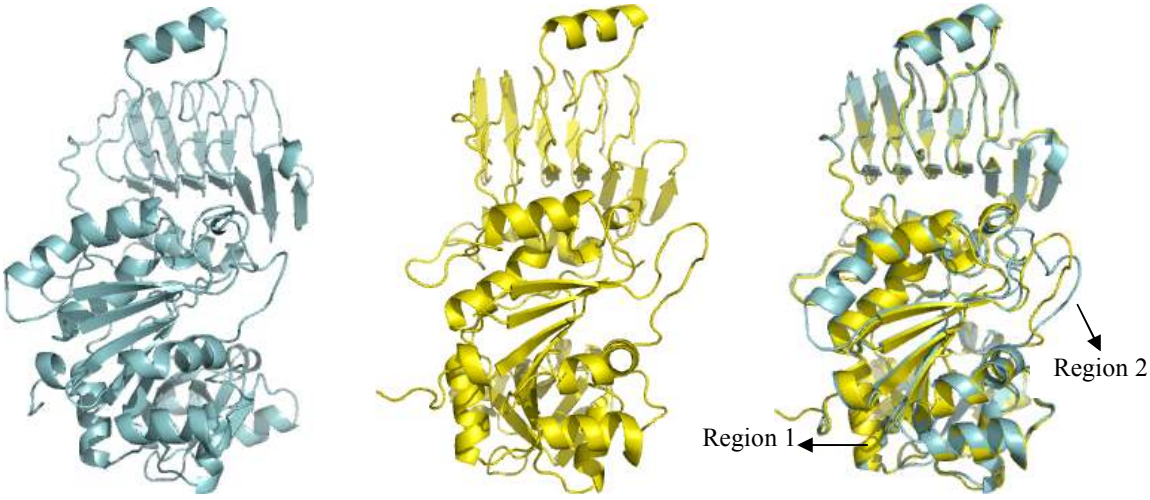
Chapter 4

MODELED STRUCTURE of LS & AGPase COMPLEX

4.1 Modeled Structure of the AGPase Large Subunit

Predicted structure of potato AGPase LS, crystal structure of SS (C chain) and the superimposed images of these subunits can be seen in Figure 4.1. Modeled LS structure shows an rmsd value of 1.3 Å when superimposed with the C-chain of crystal structure of SS (1c) using backbone atoms. This relatively small difference indicates a high structural similarity between the subunits. Visual inspection of the superimposed structures reveals two regions where the subunits differ most. These regions correspond to residues between 95-108 (region 1) and 122-126 (region 2) in LS and 108-119 and 133-134 in SS (1c) and may reflect different functions for the subunits. For example, amino acids 112-117 undergo a conformational change upon binding of glucose-1-p binding in small subunit. Also residues from 106 to 119 are forced to move significantly upon ATP binding (11). While both of the regions constitute loop structures in LS, the first fragment in SS is a loop and the second one is part of a α-helix. When the structurally completed SS was superimposed against the original C-chain of crystal structure (not shown) an rmsd value of 1.09 Å was obtained since the modeling process used the original chain as the template. In addition to the visual differences pointed here several mutagenesis studies revealed differences between LS and SS at specific amino acid positions which are important for substrate binding. Fu *et al.* (19) showed that Lys198 in SS plays a critical role for Glucose-1-

phosphate binding whereas a mutation on the homologous residue in LS do not affect the catalytic properties of the enzyme. Similar to Lys198, Asp143 in SS, but not the homologous Asp160 in LS, was also found to be vital for the catalytic activities of the enzyme (22). Moreover, alanine mutations of Lys 404 and Lys 441 in SS have been discovered to decrease the affinity of AGPase towards 3-PGA much more compared to the equivalent mutations in LS (20).



a b c

Figure 4.1: 3D structures of LS (a), SS (b) and their superimposed images (c). See methods section for details about the missing regions in SS.

4.2 Analysis of Interaction between Small and Large Subunits of Potato AGPase

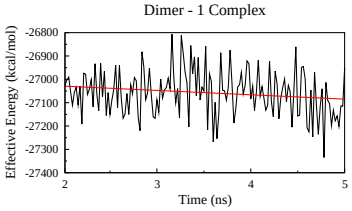
In this study, this method plays a central role for predicting the native heterotetrameric structure of the potato AGPase. As mentioned in material and methods we have investigated possible six different dimeric interactions that can exist between the subunits (Figure 3.2). Since the MM-PB(GB)SA methods require the systems to reach equilibrium, first 2 ns parts of the simulations were considered as the transition phase from the starting

structures into the equilibrium state. This was essential since the starting configurations of the dimers were not taken from experimentally determined crystal structures and the size of the dimers were relatively large which might require significant amount of time to reach equilibrium state.

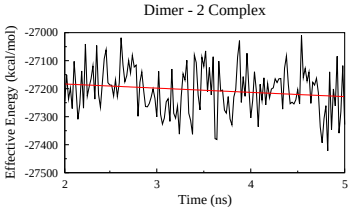
To test the stability of the systems, gas-phase energies and solvation energies were calculated for each snapshot and plotted as shown in Figure 4.2. The graphs show that energy of the systems highly fluctuates except for the Dimer 1 ligand (2g), Dimer 4 ligand (2p) and Dimer 5 receptor (2n) which exhibit relatively smaller oscillations. While Dimer 1 receptor and Dimer 3 ligand display increasing energetic values as can be seen by the regression lines, $E_{MM} + G_{sol}$ for rest of the systems decrease. Four of the systems seem to be converged in terms of $E_{MM} + G_{sol}$ energies when the slopes of the regression lines are considered. These are Dimer 3 ligand, Dimer 4 ligand, Dimer 6 complex and Dimer 6 receptor for which the slopes of the regression lines are 1.2×10^{-3} kcal/(mol•ps), 8.4×10^{-2} kcal/(mol•ps), 5.8×10^{-2} kcal/(mol•ps) and 1.1×10^{-2} kcal/(mol•ps), respectively. A more stringent convergence criterion can be applied by extending the simulation periods, but longer simulations might not guarantee convergence especially when the sizes of the systems are considered. It can be observed that dimers involving SS are energetically more convergent compared to systems possessing LS. A possible explanation to this convergence difference can be that since the structure of SS is an experimentally determined one it might reach the equilibrium state faster than the LS which is a computationally modeled structure.

As explained in materials and methods, the dimers with the most favorable interaction energies (one among D1-D3 and one from D4-D6 in 3.2) were used to construct the heterotetrameric AGPase model. Binding free energies along the trajectories of all dimers are presented in Figure 4.3. The averages of the 150 snapshots are listed in Table 4.1. It can again be observed from Figure 4.3 that ΔG values for dimers comprising of only SSs

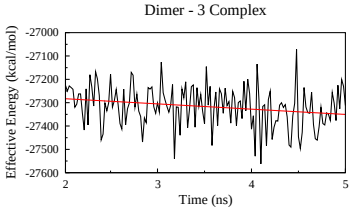
(Dimer 3 and 6) do not fluctuate as much as the other dimers. This fact can also be seen from the standard mean errors for these dimers which are 2.24 and 1.48 kcal/mol.



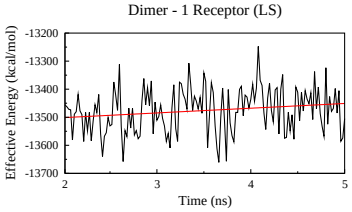
2a



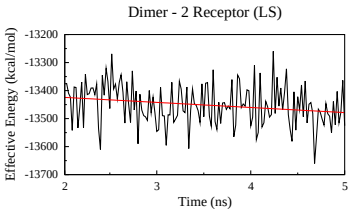
2b



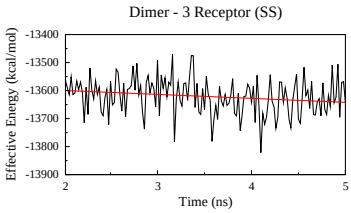
2c



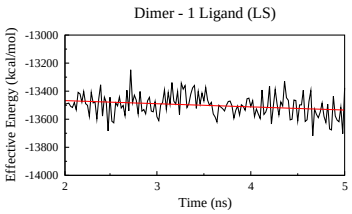
2d



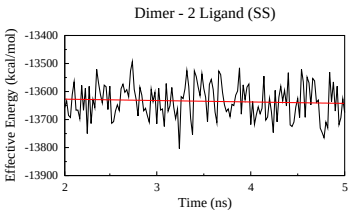
2e



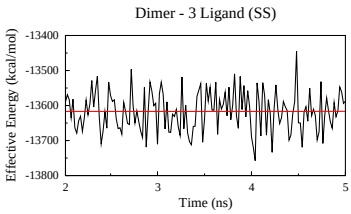
2f



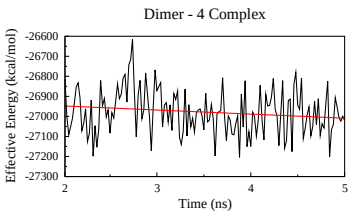
2g



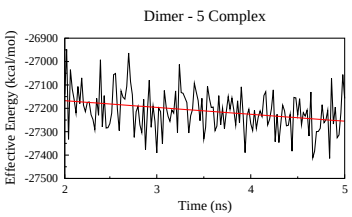
2h



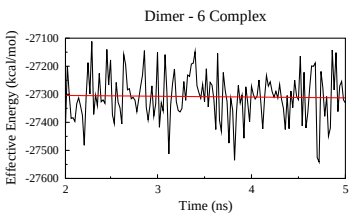
2i



2j



2k



2l

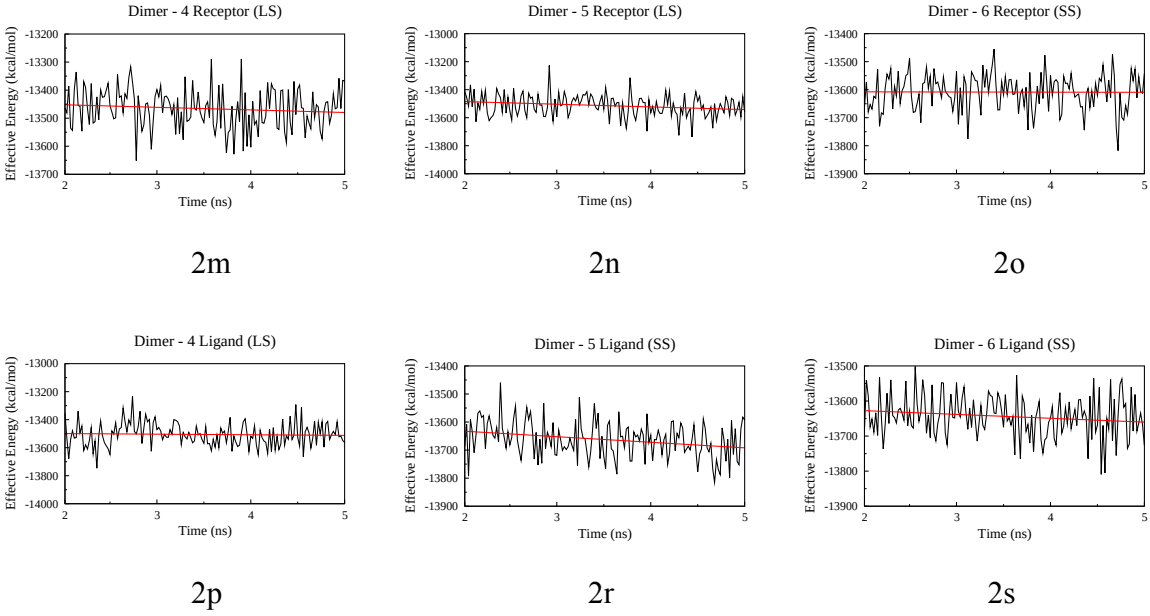


Figure 4.2: Graphs showing the E_{MM} + G_{sol} vs. Time for complex, receptor and ligand in each of the dimeric interactions. Regression lines are plotted in red. In each of the graph complex corresponds to dimer shown in Figure 3.2. For homodimers both the receptor and ligand are the same structures whereas LS was treated as receptor and SS was treated as ligand in heterodimers.

Table 4.1 demonstrates that subunit interactions in set 1 have dominant roles in maintaining the stability of native AGPase structure. In other words, the interaction between the lateral subunits is stronger compared to the longitudinal interactions. Indeed, number of interface residues that take place in binding of these dimers is higher than the set 2 interface residues. In all of the cases, internal energies (ΔE_{int} = Δ(bond + angle + dihedral)) converge to zero which is characteristic of single trajectory approach. This is due to the fact that subunits do not experience significant conformational changes upon binding since the coordinates of binding partners are taken from the complex trajectory.

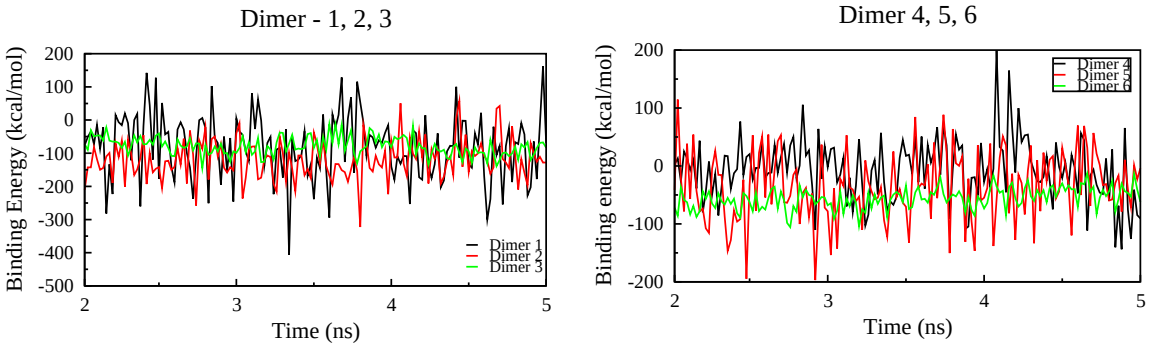


Figure 4.3: ΔG values calculated as a sum of gas phase, polar and non-polar solvation free energies for each of the 150 snapshots. Time interval between the two consecutive frames is 20 ps constituting a total of 3 ns. (a) Binding free energies for set 1 dimers and (b) for set 2 dimers.

Table 4.1: Binding free energy components for each of the dimers averaged over the 150 snapshots. Values in parentheses are standard errors of the means. Explanation for the abbreviations can be found in materials and methods. ΔG_{elec} corresponds to sum of gas-phase electrostatic energy and polar solvation energy. Energetic components for each structure are presented in Appendix A.1.

	Set 1			**Set 2**		
	Dimer 1	**Dimer 2**	**Dimer 3**	**Dimer 4**	**Dimer 5**	**Dimer 6**
ΔE_{elec}	-292.48 (5.49)	-498.12 (3.12)	-167.01 (2.92)	-243.58 (3.61)	-307.96 (3.80)	-366.34 (3.25)
ΔE_{VDW}	-190.73 (0.73)	-190.70 (0.73)	-187.05 (0.53)	-112.23 (0.81)	-104.04 (0.69)	-99.96 (0.62)
ΔE_{int}	0.01 (0.0)	0.01 (0.0)	0.01 (0.0)	0.01 (0.0)	0.01 (0.0)	0.01 (0.0)
ΔG_{gas}	-483.20 (5.76)	-688.81 (3.28)	-354.05 (2.93)	-355.80 (3.92)	-411.99 (3.45)	-466.30 (3.45)
$\Delta G_{non-polar}$	-18.35 (0.06)	-18.97 (0.06)	-17.49 (0.04)	-12.15 (0.07)	-10.83 (0.04)	-10.67 (3.75)
ΔG_{polar}	421.41 (10.38)	588.19 (5.66)	292.47 (3.53)	361.19 (5.42)	387.62 (6.44)	420.99 (3.79)
ΔG_{sol}	403.06 (10.36)	569.22 (5.65)	274.98 (3.52)	349.03 (5.39)	376.79 (6.44)	410.32 (3.75)
ΔG_{elec}	128.92 (7.99)	90.06 (4.77)	125.46 (2.64)	117.60 (4.43)	79.66 (4.70)	54.64 (1.79)
ΔG_{Total}	**-80.15 (7.96)**	**-119.59 (4.85)**	**-79.07 (2.24)**	**-6.76 (4.33)**	**-35.19 (4.83)**	**-55.97 (1.48)**

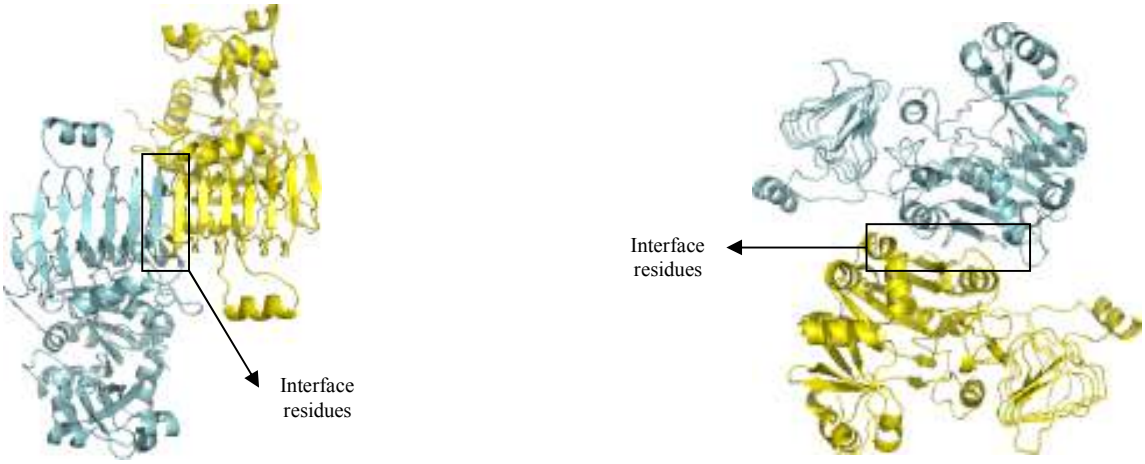
For all dimers in set 1, ΔE_{VDW} values are around 189.49 kcal/mol and are very close to each other. Differences among the set 2 dimers for this energy component are more pronounced, but still very close to each other. In addition to ΔE_{VDW} energies, $\Delta G_{non\text{-}polar}$ energies also show very similar values in each set of dimers. Values in set 1 dimers are higher than the values in set 2. This is an expected result since non-polar solvation energy is directly related to the solvent accessible surface area by a relation of $G_{non\text{-}polar} = \alpha \times SASA + \beta$ and the number of interface residues that are buried upon complex formation in set 1 complexes is higher than the set 2 complexes. Gas-phase electrostatic (E_{elec}) and polar solvation (G_{polar}) energies are observed to be discriminating among the subunits. It is also seen that ΔE_{elec} for Dimer 2 is 1.7 times greater than Dimer 1 and nearly 3.0 times greater than Dimer 3. This indicates a better steric complementation between large and small subunits. For set 2 dimers the most favorable steric complementarity is achieved by the association of two small subunits in Dimer 6. ΔE_{elec} for this dimer is 1.5 and 1.2 times greater than Dimer 4 and 5, respectively. While gas-phase electrostatic energies favor binding, polar solvation energies contribute negatively to the interactions. These two components generally tend to cancel each other. In our study, dimers that have higher E_{elec} values can better compensate the desolvation penalties of complexation (Dimer 2 and 6, Table 4.1), but the overall contribution from ΔG_{elec} disfavors binding of dimers. However, positive contributions from van der Waals and non-polar solvation energies drive the association of dimers, thus yielding overall favorable complexes. These results are consistent with the previously reported studies (49-54, 60).

4.3 Properties of Native Structure of AGPase

Table 4.1 shows that Dimer 2 is a more favorable interaction with a binding free energy of -119.59 kcal/mol compared to Dimer 1 and 3 which have values of -80.15 kcal/mol and

-79.07 kcal/mol, respectively. This means that association of LS and SS is more likely to happen than a homodimer formation. Among the set 2 dimers, although Dimer 6 is the most favorable interaction it is also important to consider the contribution of Dimer 4 while determining the native enzyme structure. Association of two small subunits in an up-down manner makes enforcement to large subunits to interact in the same way (Figure 3.1.b) which results in a total of -62.73 kcal/mol binding free energy (E = -55.97 kcal/mol + (-6.76 kcal/mol)). However, if we consider a heterodimeric interaction between the subunits as in Dimer 5, we obtain a total of 70.38 kcal/mol free energy (E = 2 × -35.19 kcal/mol) which is higher than -62.73 kcal/mol. This indicates that interaction between the LS and SS in the up-down manner would be more preferable for assembly of the enzyme when the overall structure and stability is considered. More detailed pictures of Dimer 2 and Dimer 5 are illustrated in Figure 4.4. According to these results, we report that the native heterotetrameric structure of AGPase should be Model-2 as shown in Figure 3.1.c.

After establishing the native structure of AGPase (Figure 4.5), we have further investigated Dimer 2 and 5, in order to get a more detailed picture. A complete list of the interface residues for Dimer 2 and 5 can be seen in Table 4.2. While the number of interface residues for Dimer 2 is 81 this number decreases to 53 for Dimer 5. Consequently, average buried surface area between the subunits in Dimer 2 is 3330.21 $Å^2$ and 1828.05 $Å^2$ for Dimer 5. Using CONSURF server (97) we found that large and small subunits contain highly conserved residues at their interface (Table 4.2). All of the residues with a conservation score of 9, Thr303 and Pro310 in LS and Thr304, Pro310 and Pro311 in SS, reside in the loop regions taking role in subunit interactions and structural stability as shown in Figure 4.6. In Dimer 2, while Ala93 in LS is part of a β-sheet, Gln117 in SS is part of a α-helix.



a b

Figure 4.4: Snapshots taken from the final structure of the simulations showing (a) Dimer 2 and (b) Dimer 5. In both of the structures LS and SS are colored with cyan and yellow respectively. Interface regions are indicated by arrows. Note that note all of the interface amino acids are visible. For a complete list of the residues see Table 4.2.

The last 3 ns of the simulations were further used to compare the dynamic properties of interface and non-interface residues in large and small subunits. Average root mean square fluctuations for the two proposed dimers are displayed in Figure 4.7. Average Rmsf values for LS in Dimer 2 and 5 are found to be 0.86 Å and 0.92 Å, respectively. These values show a slight increase for SS which are 0.91 Å and 0.96 Å for the corresponding dimers. In general, residues in both of the dimers demonstrate similar modes of fluctuations, but a closer examination of the graphs reveals two regions, comprising a total of 128 amino acids, in Dimer 2 that show relatively smaller fluctuations compared to their equivalents in Dimer 5. These two regions correspond to residues between 299-363 in LS with an average rmsf of 0.6 Å and 300-364 in SS with an average rmsf of 0.59 Å. It should be pointed that 23 of the interface residues in LS and 26 of them in SS lie with in these regions.

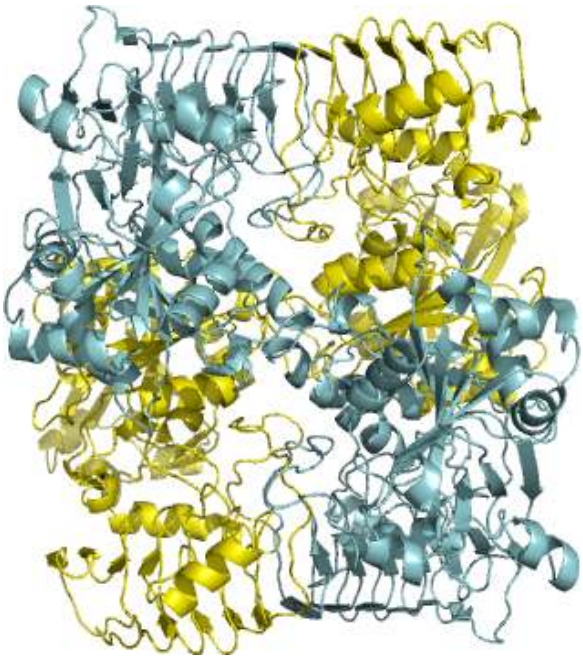


Figure 4.5: Modeled structure of the heterotetrameric potato AGPase. Subunits cyan in color are LS and yellow in color are SS. The model was generated from the final snapshots of the Dimer 2 and Dimer 5 simulations. Schematic presentation of the model can also be seen in Figure 3.2.c.

In LS residues between 299 and 310 are part of a loop region which connects the N-terminal domain with the C-terminal β-helix domain. The corresponding region in SS is between the residues 300 and 341. A possible explanation for the smaller Rmsf values, even they constitute a loop, with in these 9 residue fragments can be that they make interactions with their counterparts in different subunits. These interactions might restrict the movement of amino acids here thus resulting in smaller fluctuations. Indeed, Jin *et al* (11) had also reported that residues between 300 and 320 (291-311 in our case) make several interactions with their equivalent regions in the other subunit.

Table 4.2: Interface residues in Model-2. Residues that show > 1 $Å^2$ change in their solvent accessible surface area upon complexation and hold this condition for at least 80% (120 snapshots) of the last 3 ns of the simulations were considered as part of the interface. Conserved residues obtained from CONSURF (97) were colored according to their conservation score; blue = 6, green =7, yellow = 8 and red = 9 (highest score). Residues with a conservation score of 5 or below were not colored.

	LS	SS
Dimer 2	CYS39 TYR40 THR76 GLY79 ASN80 GLY81 SER83 PHE84 LEU285 THR286 GLU288 PHE289 PRO290 TYR295 PRO297 TYR302 THR303 PRO305 PHE307 LEU308 PRO309 PRO310 THR311 LYS312 ILE313 ASP314 ASN315 CYS316 LYS317 ILE318 LYS319 ASP320 ALA321 ILE322 ILE323 SER324 HIS325 TYR362 THR364	GLY40 ALA41 ASN42 TYR43 SER77 ARG78 SER82 ASN83 ILE285 THR286 LYS287 LYS288 PRO291 TYR296 ARG298 THR304 GLN305 PRO306 ARG307 TYR308 LEU309 PRO310 PRO311 SER312 LYS313 MET314 LEU315 ASP316 ALA317 ASP318 VAL319 THR320 ASP321 SER322 VAL323 ILE324 GLY325 GLU326 LYS331 ARG343 TYR804363 THR365
Dimer 5	ARG28 ASN65 SER66 ALA67 ASN70 ARG71 ILE73 ALA74 PHE78 GLY79 ASN80 SER83 ASP86 PHE88 GLU90 VAL91 LEU92 ALA93 ALA94 GLN96 THR97 PRO98 PHE116 TRP118 ASN125 LYS126	TYR62 ASN68 ALA70 ASN73 ARG74 SER77 ARG78 ALA81 SER82 ASN83 MET84 GLU90 PHE92 VAL93 GLU94 LEU96 ALA98 GLN100 SER101 PRO102 GLN117 TYR118 TRP120 1LEU121 GLU124 HIS125 THR126

Rest of the residues, 311-363 in LS and 312-364 in SS, mostly make up the C-terminal β-helix domains of their corresponding subunits. These results are also in agreement with study of Cross *et al* (27). The region they have identified in SS, which is important for the interaction with LS and enzyme stability, comprises the amino acids between 322 and 376 (289-343 in our case).This 55 residue long fragment highly overlaps with our smaller rmsf region and 29 of them make up an interface with LS. Overall, we can conclude that residues interacting with their counterparts in the other subunit or being a part of the structurally rigid β-helix domain experience smaller rmsf values. In contrast to Dimer 2, we can not observe any obvious stretches in Dimer 5 that represent rigid fragments (Figure 4.7). This might be due to the fact that number of interface residues in this interaction type

is lower and they are more scattered than in Dimer 2, however they still make important interactions. For instance, Glu90, Glu94, Gln100, Trp120, Glu124 residues in SS, which were also reported by Jin *et al* (11) in the crystal structure of SS, were found to be at interface in all of the 150 snapshots taken from the last 3 ns of the simulations. Moreover, residues such as Arg74, Arg78, Glu90 and Glu124 in SS make salt bridges with Glu90, Asp86, Arg75 and Arg71 in LS respectively.

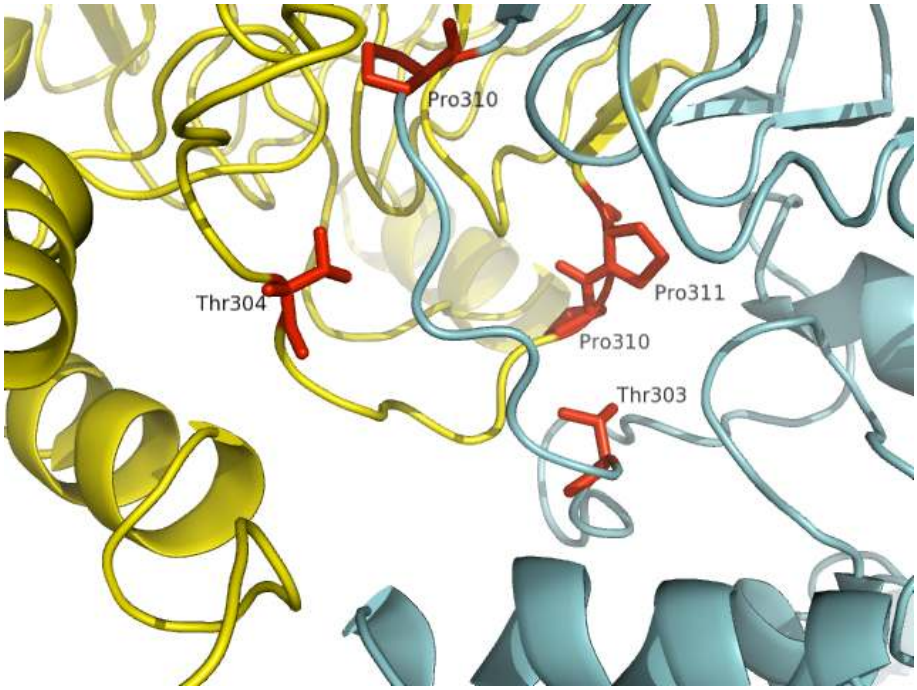


Figure 4.6: Snapshot showing the highly conserved residues (red in color), Thr303, Pro310 in LS and Thr304, Pro310, Pro311 in SS in Dimer 2 interface. LS is cyan and SS is yellow in color.

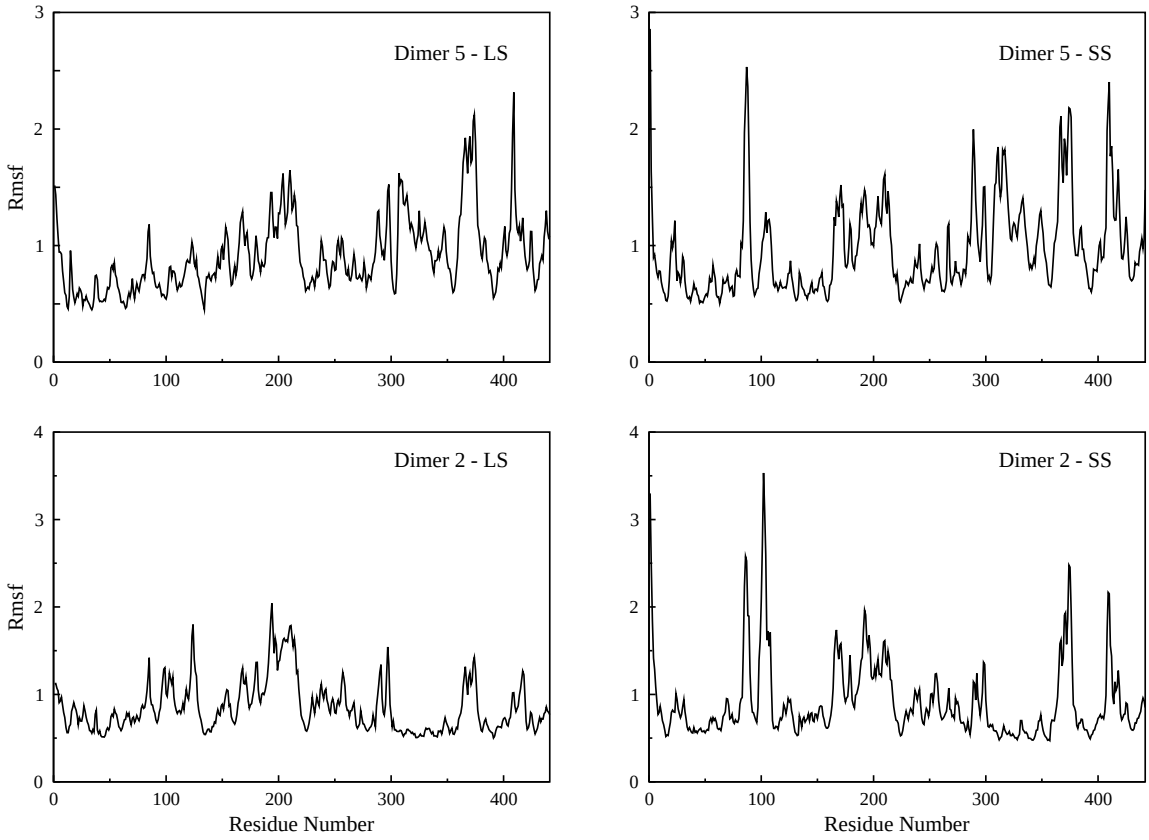


Figure 4.7: Root mean square fluctuations (Rmsf) of backbone atoms (C, CA, N) versus residue number for Dimer 2 and Dimer 5 subunits.

Chapter 5

HOT-SPOTS in POTATO AGPase MODEL & EXPERIMENTAL VERIFICATION

5.1 MM-GBSA Analysis of Dimeric Interactions

The aim of this chapter is to identify the hot-spot residues, both computationally and experimentally, that mediate the interactions between the potato LS and SS in the proposed model (Model 2 in Figure 3.1.c). First, MM-GBSA was used to calculate total binding energy difference between the possible dimers (Figure 3.2) to verify the results of MM-PBSA calculations for our AGPase model and decompose the free energy contributions of amino acids. Then, residues that are identified to be critical were subjected to site directed mutagenesis to change them alanine. Finally, yeast two hybrid systems are used to analyze their role in subunit-subunit interactions. As mentioned in materials and methods simulation times were extended to 8 ns and the first 4 ns parts were excluded from the calculations to achieve an improved environment for the equilibrium state.

The binding free energy differences for all of the dimers using MM-GBSA are shown in Table 5.1. As opposed to the binding free energies calculated by MM-PBSA (Table 4.1), ΔG_{Total} values for both set 1 and set 2 dimers are closer to each other. Another striking difference between the MM-PBSA and MM-GBSA results is the discrepancies among the polar solvation energies which also account for the variations in ΔG_{Total} values. In contrast, gas-phase and non-polar solvation energies calculated by two different methods are parallel to each other since the formulations of E_{elec}, E_{vdw}, E_{int} and $G_{non\text{-}nonpolar}$ are same in both of

the methods. Minor variations in these terms are due to the differences in simulation periods and atomic radius definitions in PARSE (94) and mbondi2 (96) parameter sets used in MM-PBSA and MM-GBSA methods. It is also observed that in all of the dimeric interactions, favorable ΔE_{elec} terms are compensated by unfavorable ΔG_{polar} terms. Hence, total electrostatic interactions ΔG_{elec}, disfavor binding of subunits. However, contributions from van der Waals and non-polar solvation energies favor interactions thus being the major forces that drive the association of subunits.

Table 5.1: Binding free energy components for each of the dimers averaged over the 200 snapshots. Values in parentheses are standard errors of the means. Explanation for the abbreviations can be found in materials and methods. ΔG_{elec} corresponds to sum of gas-phase electrostatic energy and polar solvation energy. Energetic components for each structure are presented in Appendix A.2.

		Set 1			**Set 2**	
	Dimer 1	**Dimer 2**	**Dimer 3**	**Dimer 4**	**Dimer 5**	**Dimer 6**
$\mathbf{\Delta E_{elec}}$	-268.55 (3.25)	-508.73 (2.88)	-135.23 (2.32)	-220.34 (4.40)	-348.80 (3.06)	-391.30 (2.79)
$\mathbf{\Delta E_{VDW}}$	-190.10 (0.60)	-187.61 (0.54)	-174.80 (0.63)	-113.75 (0.48)	-103.68 (0.59)	-101.01 (0.58)
$\mathbf{\Delta E_{int}}$	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
$\mathbf{\Delta G_{gas}}$	-458.64 (3.36)	-696.33 (3.04)	-310.02 (2.58)	-334.07 (4.59)	-452.47 (3.17)	-492.30 (2.66)
$\mathbf{\Delta G_{non\text{-}polar}}$	-18.36 (0.04)	-18.48 (0.05)	-17.25 (0.04)	-13.62 (0.04)	-11.52 (0.04)	-11.83 (0.03)
$\mathbf{\Delta G_{polar}}$	341.53 (3.01)	581.13 (2.81)	209.74 (2.15)	303.54 (4.15)	417.69 (2.75)	456.11 (2.56)
$\mathbf{\Delta G_{sol}}$	323.17 (3.05)	562.65 (2.78)	192.50 (2.13)	289.92 (4.13)	406.17 (2.74)	444.28 (2.56)
$\mathbf{\Delta G_{elec}}$	72.98 (0.57)	72.40 (0.54)	74.52 (0.56)	83.20 (4.13)	68.88 (0.66)	64.81 (0.64)
$\mathbf{\Delta G_{Total}}$	**-135.47 (0.65)**	**-133.67 (0.50)**	**-117.52 (0.69)**	**-44.15 (0.67)**	**-46.30 (0.63)**	**-48.01 (0.50)**

Shown in Table 5.2 are the dimeric interactions between the subunits that constitute the corresponding models and their contributions to the overall energy of the model enzymes calculated both by MM-PBSA and MM-GBSA methods. As previously mentioned MM-PBSA results strongly favor Model-2 with a total energy of -309.56 kcal/mol. However, energy results obtained from MM-GBSA method slightly favors Model-2 with a total energy of -359.94 kcal/mol. It is also probable for Model-1 to represent the heterotetrameric AGPase structure if MM-GBSA results are considered. Although the energies for Model-1 (-359.50 kcal/mol) and Model-2 (-359.94 kcal/mol) are very close to each other according to MM-GBSA results, extension of the simulations and exclusion of longer periods of initial phases may better pronounce the differences. It should also be noted that a disulfide bond between the Cys12 residues in SSs can only be formed in Model-2. Since only dimeric interactions were investigated in this study such a high favorable energy contribution coming from this covalent bond was not considered. Addition of this energy term can enhance the interaction between the subunits and further stabilize the Model-2.

Table 5.2: Dimer and model enzyme energies calculated by MM-PBSA and MM-GBSA methods. Values are in kcal/mol.

	Lateral Interactions	Energy		Longitudinal Interactions	Energy		Total Energy	
		PBSA	GBSA		PBSA	GBSA	PBSA	GBSA
Model-1	D2+D2	-239.18	-267.34	D4+D6	-62.73	-92.16	-301.91	-359.50
Model-2	D2+D2	-239.18	-267.34	D5+D5	-70.38	-92.60	-309.56	-359.94
Model-3	D1+D3	-159.22	-252.99	D5+D5	-70.38	-92.60	-229.6	-345.59

5.2 Free Energy Decomposition for Dimer 2 and Dimer 5

In this study the definition for hot spots is as follows: If a residue shows 3.0 kcal/mol energy drop in dimer structure compared to its subunit form ($|\Delta G_{binding}| > 3.0$ kcal/mol), then it is called a hot spot. Since the AGPase model (Model-2, Figure 3.1.c) consists of pairs of D2 and D5 free energy decomposition was applied only to these dimers. Again, entropic contributions were not considered in the calculations since the main interest is to identify the hot-spots qualitatively. However, addition of entropic effects may give quantitatively more realistic results.

Table 5.3[¶]: Free energy decomposition of hot spot residues in Dimer 2.

Residue	ΔE_{ele}	ΔE_{vdw}	ΔG_{polar}	$\Delta G_{non\text{-}polar}$	$\Delta G_{binding}$ Backbone	$\Delta G_{binding}$ Side Chain	$\Delta G_{binding}$ Total
LS							
Asn80	-9.99±4.37	-4.00±1.20	11.06±4.61	-0.44±0.13	-0.57±0.44	-2.79±1.25	**-3.36**±1.12
Pro310	-0.50±0.40	-5.28±0.50	1.32±0.32	-0.57±0.03	-1.80±0.31	-3.24±0.47	**-5.03**±0.53
Thr311	-4.21±0.94	-2.23±0.58	3.35±0.61	-0.19±0.03	-2.26±0.57	-1.03±0.22	**-3.29**±0.57
Ile313	-2.99±1.18	-4.32±0.61	2.82±0.72	-0.38±0.03	-1.91±0.54	-2.97±0.35	**-4.88**±0.61
Ile318	-0.68±1.15	-4.13±0.74	0.29±0.67	-0.40±0.03	-1.73±0.42	-3.19±0.41	**-4.92**±0.56
Ile322	-0.43±0.47	-3.09±0.44	0.10±0.39	-0.24±0.03	-1.54±0.26	-2.12±0.42	**-3.66**±0.49
Ile323	-3.73±1.27	-3.27±0.54	4.01±0.82	-0.29±0.08	-1.54±0.46	-1.74±0.64	**-3.28**±0.96
His325	-8.47±1.34	-2.34±0.76	7.72±0.86	-0.38±0.04	-0.21±0.09	-3.27±0.60	**-3.48**±0.62
SS							
Met84	-0.57±1.19	-3.67±0.99	1.50±1.00	-0.52±0.13	-0.22±0.30	-3.03±1.08	**-3.25**±1.13
Lys288	-38.56±7.60	-0.47±0.70	36.09±6.67	-0.21±0.05	-0.44±0.17	-2.71±1.10	**-3.15**±1.13
Tyr308	-5.42±2.11	-6.63±0.99	6.12±1.32	-0.81±0.06	-1.99±0.88	-4.75±0.94	**-6.75**±1.16
Pro310	-3.24±0.71	-3.57±0.46	3.94±0.64	-0.33±0.04	-0.78±0.24	-2.41±0.42	**-3.19**±0.48
Pro311	1.59±0.82	-4.47±0.56	-0.50±0.60	-0.57±0.03	-0.97±0.40	-2.97±0.50	**-3.94**±0.68
Lys313	-102.96±5.58	-3.56±1.01	102.40±4.61	-0.65±0.04	-1.24±0.28	-3.54±1.57	**-4.77**±1.59
Met314	-3.70±1.23	-4.23±0.58	3.93±1.13	-0.40±0.04	-2.23±0.37	-2.18±0.52	**-4.41**±0.66
Val319	-3.17±1.00	-3.57±0.57	2.60±0.58	-0.32±0.02	-2.36±0.53	-2.11±0.33	**-4.47**±0.58
Thr320	-7.80±2.56	-2.21±0.44	6.49±1.75	-0.15±0.03	-2.19±0.35	-1.49±0.83	**-3.67**±0.96
Val323	-0.39±0.38	-2.78±0.38	0.07±0.30	-0.19±0.04	-1.55±0.19	-1.74±0.38	**-3.29**±0.43
Ile324	-5.92±0.89	-3.09±0.61	5.12±0.45	-0.29±0.03	-2.29±0.49	-1.89±0.31	**-4.18**±0.57

Table 5.4[¶]: Free energy decomposition of hot spot residues in Dimer 5.

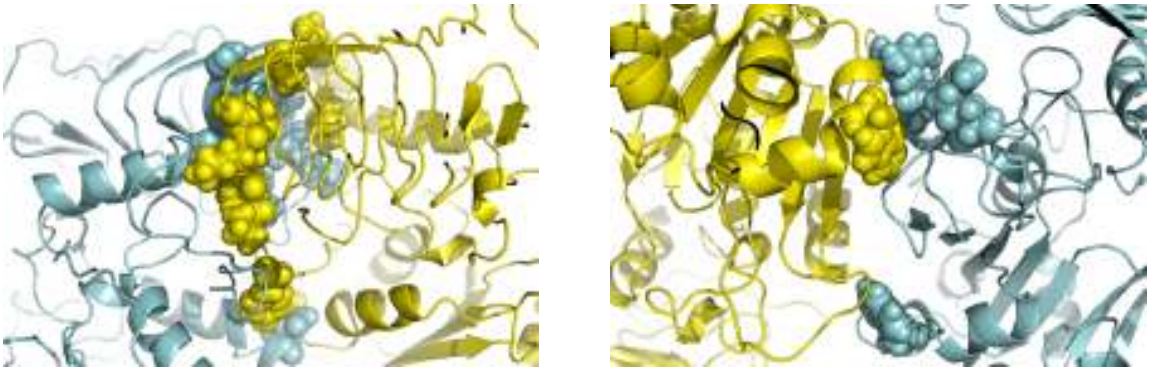
Residue	ΔE_{ele}	ΔE_{vdw}	ΔG_{polar}	$\Delta G_{non\text{-}polar}$	$\Delta G_{binding}$		
					Backbone	Side Chain	Total
LS							
Arg28	-89.00±14.39	0.08±0.78	83.81±11.20	-0.16±0.08	0.07±0.04	-5.35±3.14	**-5.28**±3.15
Arg71	-89.46±11.40	-4.15±0.81	90.33±8.50	-0.55±0.05	-0.18±0.08	-3.64±2.85	**-3.82**±2.86
Arg75	-88.88±3.83	-0.53±0.90	83.25±3.17	-0.21±0.04	-0.03±0.07	-6.34±1.03	**-6.37**±1.01
Trp118	-5.63±1.25	-4.44±0.63	6.20±0.82	-0.52±0.06	-0.11±0.05	-4.28±0.86	**-4.39**±0.88
SS							
Trp120	-5.46±1.09	-6.04±0.59	6.17±0.88	-0.62±0.03	-0.06±0.07	-5.88±0.76	**-5.95**±0.77

[¶] Values are in kcal/mol.

Hot-spot residues for Dimer 2 and Dimer 4 and their binding free energy components are shown in Table 5.3 and Table 5.4, respectively. It should be noted that, this time for a residue to be considered as interface its absolute SASA must decrease at least 1Å^2 upon complexation and it must satisfy this condition for at least 160 of the snapshots. Based on these requirements, total of 79 (38 in LS and 41 in SS) residues in D2 and 53 residues (27 in LS and 26 in SS) in D5 were classified to be part of interfaces.

We observe that out of 79 interface residues in D2, 19 of them (8 in LS and 11 in SS) are hot-spots. The hot-spot residues in LS for D2 are non-polar in general. Only LS_{Asn80}, LS_{Thr311} and LS_{His325} are polar or charged. It is seen that this interface is highly populated by Ile. Seven of the hot-spots in SS for D2 are also non-polar too. Residues SS_{Lys288}, SS_{Tyr308}, SS_{Lys313} and SS_{Thr320} make up the polar region in this interface. According to these ratios we can say that key interactions for D2 are hydrophobic in nature. An illustrated picture of the hot-spots in D2 can be seen in Figure 5.1.a. Number of hot-spots in D5 is relatively smaller than the residues in D2. This is an expected result since the buried surface upon dimer formation for this interaction is smaller too. In contrast to D2 hot-spots, which are generally non-polar, there are three basic amino acids (Arg28, Arg71, Arg75 in

LS) in this interface. The remaining residues (Trp118 in LS and Trp120 in SS) are non-polar. Figure 5.1.b shows the hot-spot residues in D5.



A B

Figure 5.1: Snapshots taken from the final structures of MD simulations showing the hot-spot residues (A) in D2 and (B) in D5. LS is cyan and SS is yellow in color. Hot-spots are shown in spheres.

5.3 Computational Analysis of Hot-Spot Interactions in D2 and D5

In the following sections we will mainly focus on the hot-spot residues in LS, but will also discuss about some of the key residues in SS. Then a detailed picture for the energetic contributions of residues in the interfaces of homotetrameric SS will be presented.

5.3.1 Dimer 2

As can be seen from Table 5.3 Tyr308 in SS shows the highest free energy difference with a $|\Delta G_{binding}|$ value of 6.75 kcal/mol upon complexation. We see that favorable

contributions to $\Delta G_{binding}$ for this residue are dominated by E_{ele} (–5.42 kcal/mol) and E_{vdw} (-6.63 kcal/mol) terms for which Thr76, Asn80 and Thr303 in LS are responsible for the former term. Indeed, several H-bonds are formed by Tyr308 and these polar residues. However, the negative contribution of polar solvation energy (-6.12 kcal/mol) compensates the electrostatic term. Tyr749 is also in close contact with residues Pro302 and Pro305 which account for the favorable van der Waals interactions.

Pro310 in LS has the second highest $|\Delta G_{binding}|$ energy difference with a value of -5 kcal/mol. It should be noted that this residue is highly conserved. It makes van der Waals contacts with Gly481, Ala482, Ile726, Ile765 and the aromatic ring of Tyr484. These interactions explain the hydrophobic nature of Pro310 $|\Delta G_{binding}|$ difference.

Shown in Figure 5.2 are the three important isoleucine residues in LS, Ile313, Ile318 and Ile323 that constitute a hydrophobic core at the inner layer of β-helix domain. The bulky side-chain groups of these residues make strong hydrophobic interactions with each other as well as their counterparts in SS. In fact, favorable $\Delta G_{binding}$ binding energies for these amino acids are mainly driven by the van der Waals forces (see Table 5.3). E_{ele} terms on the other hand are canceled by the desolvation penalties during dimer formation. Also, noteworthy about Ile313, 318 and 323 is that they form a total of six highly conserved H-bonds with Ser322, Ala317 and Ser312 in SS respectively (Figure 5.2). Even though the effects of these H-bonds are inversely balanced with the polar solvation terms, they help the LS and SS β-helix domains to maintain their correct orientation relative to each other by providing structural constrains in binding process. Modeled structure of LS reveals that side-chain of Ile322 is excluded from the hydrophobic core of β-helix domain. Instead this bulky group faces Ile338, another isoleucine whose side-chain is also excluded from the hydrophobic core, Thr286 of the catalytic domain and Lys313 in SS. Again E_{vdw} term plays a dominant role for the low energy state of this residue.

Lys313 in SS has a striking feature in terms of electrostatic and polar solvation energies. Upon complex formation this residue is surrounded by many non-polar amino

acids such as Leu285, Ala321, Ile322 in LS and Leu315, Val329 in SS which are responsible for the high ΔG_{polar} term. However, it also contacts with LS Gln287, Glu288 and Thr286 polar groups and take part in several H-bonds with these residues. These electrostatic interactions strongly favor the $\Delta G_{binding}$ for Lys754.

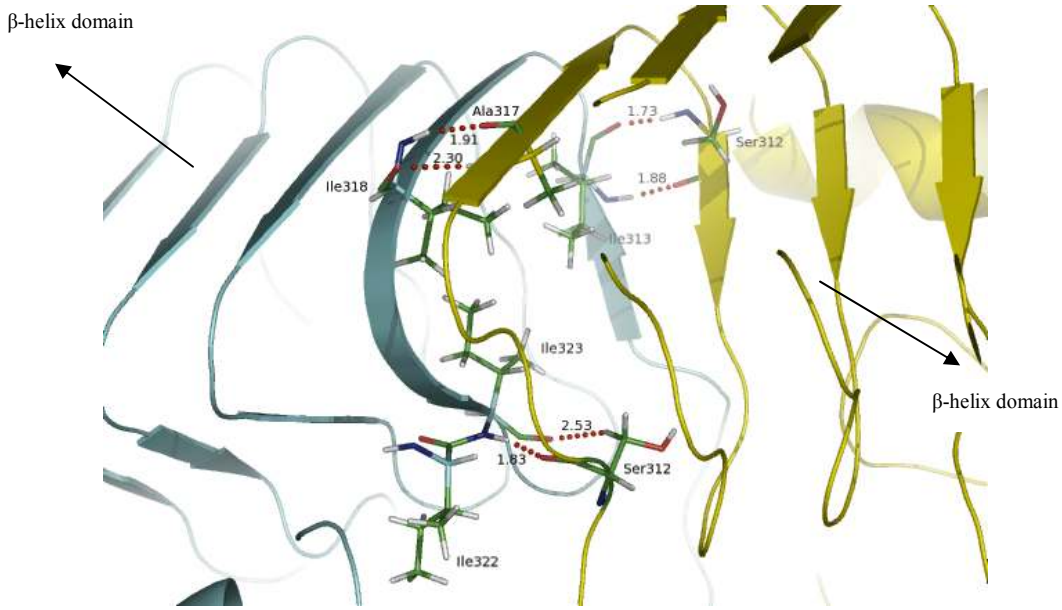


Figure 5.2: Snapshot showing the six H-bonds (red dashed lines) between Ile313-Ser312, Ile318-Ala317 and Ile323-Ser312 and their corresponding distances. These H-bonds are broken and reformed throughout the simulation. Ile322 is also illustrated in the picture. LS is shown cyan and SS is shown yellow in color.

5.3.2 Dimer 5

As shown in Table 5.4 80% of the hot-spots in Dimer 5 belong to LS. It is also worth to mention that contributions of side-chain atoms in dimer stabilization are much higher than

the backbone atoms in this group. This might suggest that longitudinal interaction is not as optimized as the lateral interactions and further might mean that even single alanine mutations on these residues can have deleterious effects in subunit-subunit interactions (LS_{R28A} is such an example, see later). In addition, these residues are responsible for 56% of the total binding free energy ((-5.38 kcal/mol + (-3.82 kcal/mol) + (-6.37 kcal/mol) + (-4.39 kcal/mol) + (-5.95 kcal/mol)) / (-46.30 kcal/mol = 0.56) for D5 complexation.

We see that E_{vdw} term has no contribution for Arg28 stabilization in LS during dimerization. Consequently, nearly all the contributions come from electrostatic interactions. This residue makes H-bonds with Ser66 and Glu431 in LS and Glu124 in SS. Although Arg28 is not caged with hydrophobic amino acids upon complexation it suffers from desolvation effects. One possible explanation for the high desolvation free energy might be that residues found in the close proximity of Arg28 can not sufficiently mimic the solvent environment in complex form.

Trp118 in LS is enclosed by both polar and non-polar groups. Residues in the first group are Asn125 in LS, Asn68, Gln100 and Ser101 in SS which are the constituents of E_{ele}. Second group of amino acids include Val119 in LS, Ala70 and Pro102 in SS. Contributions of Val119 and Ala70 to the E_{vdw} term might be smaller than Pro102 since the aromatic group of Trp118 can get involved in strong hydrophobic interactions with the side-chain of this residue.

5.4 Comparison of Important Interactions Between Dimers and Crystal Structure of Homotetrameric SS

Several important residues were reported in the crystal structure of homotetrameric SS by Jin *et al* (11). To compare these amino acids with the corresponding residues in our AGPase Model, free energy decomposition scheme was also applied to D3 and D6. Shown in Table 5.5 and Table 5.6 are the important interface amino acids and their $\Delta G_{binding}$ values in our AGPase model and in the crystal structure of homotetrameric SS, respectively.

Table 5.5: $\Delta G_{binding}$ values of important residues in SS reported by Jin *et al* (11) in our AGPase model. Values are in kcal/mol. Note that interface residues in A and C chains are not listed since these chains are occupied with LSs in our AGPase model. Results were obtained from the free energy decomposition of LS-SS interaction (D2 and D5 in Figure 3.2). Standard deviations are given in parenthesis.

	Residues in B Chain									Residues in D Chain				
	Y308	L309	P310	K313	V319	D321	S322	I324	Y363	E90	E94	Q100	W120	E124
$\Delta G_{binding}$	-6.75 (1.16)	-1.67 (0.54)	-3.19 (0.48)	-4.77 (1.59)	-4.47 (0.58)	-1.29 (0.59)	-1.97 (0.58)	-4.18 (0.57)	-0.61 (0.34)	-1.09 (1.80)	2.17 (0.60)	-1.97 (0.84)	-5.95 (0.77)	-1.07 (1.53)

Table 5.6: $\Delta G_{binding}$ values of important residues in single chain SS of homotetrameric SS reported by Jin *et al* (11). Values are in kcal/mol. Results were obtained from the free energy decomposition of SS-SS interaction (D3 and D6 in Figure 3.2). Standard deviations are given in parenthesis.

	Residues in B Chain									Residues in D Chain				
	Y308	L309	P310	K313	V319	D321	S322	I324	Y363	E90	E94	Q100	W120	E124
$\Delta G_{binding}$	-2.48	-1.71	-2.94	-1.67	-4.20	-2.69	-3.07	-3.58	-0.98	-0.74	3.30	-2.79	-5.16	-0.32

All of the residues listed in Table 5.5 were found to be part of interfaces according to our analysis. Four of the residues in the B chain (Y308, P310, V319 and I324) and W120 in D chain were also classified as hot-spots in our AGPase model. All the other residues, except for the E94, also have negative $\Delta G_{binding}$ values which mean that they are stabilized

upon complex formation. However, they were not considered as hot-spots since their change in $\Delta G_{binding}$ values according to free energy decomposition are higher than our cutoff value (-3.0 kcal/mol).

We see that while the important amino acids reported by Jin *et al.* (11) have a total of -36.81 kcal/mol $\Delta G_{binding}$ energy in our model, they are less stabilized in the homotetrameric SS with a $\Delta G_{binding}$ value of -29.03 kcal/mol (Table 5.6). In other words, those residues are more stabilized if they interact with residues from LS instead of residues from SS. This is especially true for Y308, K313, I324 and E124. In addition, E94 has smaller positive $\Delta G_{binding}$ energy in our model.

5.5 Analysis of Hot-Spots by Yeast Two Hybrid Method

Yeast-two hybrid method was used in order to experimentally verify the importance of hot-spot residues obtained from per-residue free energy decomposition scheme. First, we have tested our yeast-two hybrid system by mutating three of the residues in SS reported by Jin *et al* (11). Arg78, Gln100 and Thr286 of SS were separately mutated to alanine by site directed mutagenesis. It should be noted that selection of these residues was random. $pGBKT7_{Trp}^{+}$ constructs carrying these mutants were then sequentially transformed into the yeast cells together with the $pGADT7_{Leu}^{+}$ vector carrying the wild type (WT) SS sequence. At this point we did not use WT_{LS} for testing purposes as mentioned earlier. In LS, Arg28, Arg75, Trp118, Thr311+Ile313 (double mutant) and Ile322+Ile323 (double mutant) were mutated to alanine. $pGBKT7_{Trp}^{+}$ vectors carrying these LS mutants were sequentially transformed with the $pGADT7_{Leu}^{+}$ construct carrying the WT_{SS} sequence. Finally, we have obtained total of 13 separate transformants including the positive and negative controls.

Shown in Table 5.7 are the colony formation results of the transformants in the interaction medium lacking Leu, Trp and His amino acids. Cells carrying wild type forms

of the subunits (SS_{WT}/SS_{WT} and SS_{WT}/LS_{WT}) showed colony formation in the interaction medium. This is an expected result since these cells were used as positive controls.

Table 5.7: Results showing the effects of mutations on the subunit-subunit interactions, thus colony formation of yeast cells. Residues R78, Q100 in SS and R28, R75 and W118 in LS are involved in the longitudinal interactions. T286 in SS and T311, I313, I322 and I323 in LS are part of the lateral interactions. SS_{WT}/SS_{WT}, SS_{WT}/LS_{WT} and SS_{WT}/EV, LS_{WT}/EV are treated as positive and negative controls respectively.

Construct Type	Colony Formation
SS_{WT}/SS_{WT}	Growth
SS_{WT}/SS_{R78A}	No Growth
SS_{WT}/SS_{Q100A}	No Growth
SS_{WT}/SS_{T286A}	Growth
SS_{WT}/LS_{WT}	Growth
SS_{WT}/LS_{R28A}	No Growth
SS_{WT}/LS_{R75A}	Growth
SS_{WT}/LS_{W118A}	Growth
$SS_{WT}/LS_{T311A,I313A}$	Growth
$SS_{WT}/LS_{I322A,I323A}$	No Growth
SS_{WT}/EV	No Growth
LS_{WT}/EV	No Growth

We also see that no growth occurred in the negative control groups (SS_{WT}/empty vector(EV) and LS_{WT}/EV). It can be concluded that out yeast-two hybrid system is working since no growth is observed in SS_{WT}/SS_{78} and SS_{WT}/SS_{100} groups meaning that these

mutations disturb the interactions between two SSs as in D3 (Figure 3.2). Indeed, per-residue free energy decomposition for D6 reveals that Arg78 shows -2.66 kcal/mol $\Delta G_{binding}$ energy change in one of the subunits and -4.08 kcal/mol in the other subunit upon complex formation. Average of these values gives -3.37 kcal/mol binding free energy changes, a value which qualifies this residue as a hot-spot. Gln100 also experiences -1.15 kcal/mol free energy change in one of the subunits and -2.79 kcal/mol in the other subunit, thus yielding a -1.47 kcal/mol $\Delta G_{binding}$ value on the average. Even though this value is not enough for a residue to qualify as a hot-spot, it may still have a significant contribution to the dimer formation. Cells carrying the double mutant ($SS_{WT}/SS_{R78A+Q100A}$) constructs also showed no growth. This result is obviously expected since even the single mutations constituting this double mutant are effective in disturbing the interactions. However, cells carrying the mutant SS_{T286A} were able to grow in the interaction medium. This indicates that this mutation has no effect on the interaction of subunits in D3. Although this might be an unexpected result it is reasonable when we also consider the free energy decomposition results for this residue. Tyr286 shows -0.01 kcal/mol binding free energy change in both of the subunits. This computational result also supports the fact that this residue is unlikely to disturb the subunit interactions in D3.

We see that mutation of Arg28 in LS to alanine disturbs the interactions between LS and SS in D5. This is an expected result since the side-chain of this residue has a dominant contribution (-5.35 kcal/mol) to the $\Delta G_{binding}$ for this residue. As a consequence replacement of the bulky side-chain with a methyl group is likely to effect the dimer formation. Indeed, an alanine, which carries a non-polar side chain, is unlikely to make the electrostatic interactions formed by an arginine (-89.0 kcal/mol, see Table 5.3) at this position.

It is observed that mutation of arginine to alanine at position 75 in LS do not effect the interactions between the subunits in D5. This was not expected since this residue has the highest $|\Delta G_{binding}|$ value (6.37 kcal/mol) in this group. A possible explanation to this

unexpected result may be related to position of this residue. Arg75 is part of an α-helix. Mutation of Arg75 to alanine, which has a higher propensity for α-helix, might possibly better stabilize the overall α-helix structure. This stabilization thus allows the neighbor residues to better interact with the SS.

Mutation of Trp118 is also unlikely to disturb the interactions in D5. This is again an unexpected result according to the computational results since this residue has a -4.39 kcal/mol $\Delta G_{binding}$ value. When we look at the energetic terms in Table 5.3 we see that E_{vdw} (-4.44 kcal/mol) has a dominant role for the interactions of this residue. Most probably replacement of tryptophan side-chain with a methyl group as a consequence of alanine mutation at this position can still be able to keep these van der Waals interactions to a certain point, thus the interactions between the subunits.

Thr311 and Ile313 double mutation to alanine seems to have no effect on D2 formation according to yeast-two hybrid results although these residues account for a total of -8.17 kcal/mol $\Delta G_{binding}$ energy. We see in Table 5.2 that side-chain of Thr311 has a minor contribution (-1.03 kcal/mol) to $\Delta G_{binding}$ compared to its backbone (-2.26 kcal/mol). This decreases the chance of an alanine mutation to disturb the interactions between the subunits. In addition, although an alanine mutation at this position might result in a decrease in the electrostatic interactions, which is -4.21 kcal/mol in the case of threonine, it might increase the favorable contributions of van der Waals interactions, which is -2.23 kcal/mol in the case of threonine. Thus, this balancing changes decrease the overall effect of the alanine mutation. It is obvious that an alanine mutation at Ile313 position will certainly decrease the van der Waals effects (-4.21 kcal/mol) of this residue, but still keep the hydrophobicity of the β-helix domain. As previously mentioned Ile313 makes two H-bonds with Ala317 in SS which accounts for the -2.99 kcal/mol electrostatic interactions. However, an alanine mutation at this position can still able to keep these H-bonds with Ala317 in SS which as a result might decrease the affect of mutation. Overall, we can say

that Thr311 and Ile313 double mutation is not strong enough to perturb the interactions between the subunits.

We have previously mentioned that Ile322 side-chain is excluded from the inner part of the β-helix domain and van der Waals energies (-3.09 kcal/mol) have a dominant contribution to the favorable $\Delta G_{binding}$ of this residue which is mainly due to the interactions with Ile338. Mutation of this residue to alanine will not only decrease the E_{vdw} term, but may also force the side-chain of alanine to be included in the inner layer of β-helix domain. This is highly possible since Ile323, whose side-chain is involved in the inner layer, is mutated to alanine at the same time and the residue at position 321 is also an alanine. Such an inclusion will certainly result in steric clashes between the side-chains and disturb the interface structure of β-helix domain where it makes important interactions with SS. In other words, this mutation can effect the interactions both energetically and structurally.

According to these results, we can accept that computational and experimental data support each other.

Chapter 6

CONCLUSION

AGPase is a key regulatory enzyme catalyzing the first and rate limiting reaction in starch biosynthesis in higher plants. The enzyme has the capacity to affect the total yield of starch production. However, LS and native heterotetrameric structure of AGPase has not been revealed yet. Elucidation of native structure and important amino acids that play critical roles in subunit-subunit interaction is of crucial importance since they can provide valuable information about working principles of the enzyme. This knowledge can then be used to engineer a more stable and effective form of the enzyme which can be used to increase the total amount of starch yield. Increasing the yield thus brings both economical and environmental benefits.

In this study we combined both computational and experimental techniques to investigate the *Solanum tuberosum* potato tuber AGPase structure. First, we predicted the structure of LS by homology modeling. This was an essential preliminary step for the next procedures. Then, we have proposed three models for the native structure of AGPase based on the crystal structure of homotetrameric SS (11). In order to find the correct structure we extracted representative dimers from the proposed models and calculated the binding free energies between the subunits by MM-PBSA (12, 13) method. Most favorable interactions between the subunits were used to model the heterotetrameric AGPase structure. Results showed that AGPase is composed of pairs of heterodimers that interact both literally and longitudinally.

After establishing the native enzyme structure we further investigated the interactions between LS and SS. Hot-spot residues that are critical for the communication of LS and SS were identified by the free energy decomposition scheme using the MM-GBSA method (12, 13). Results obtained from this approach revealed many residues that have favorable contributions to the stability of dimers. Finally, we used this computational data as a guide to our experimental study. We have mutated some of these key residues to alanine by site directed mutagenesis. Using yeast-two hybrid method (14) effects of these mutations on the subunit-subunit interactions were evaluated. Combined results from computational and experimental data show that Arg28 and Ile322+Ile323 in LS are the hot-spots that are important for the longitudinal and lateral interactions of LS and SS, respectively. Also the SS residues, Arg78 and Glu100, have deleterious effects in homodimeric interactions of SSs. With an increased computational power interactions between the subunits can be studied at heterotetrameric level rather than heterodimeric level. This can provide more vigorous computational and experimental results.

We believe that this study establishes the groundwork for the native heterotetrameric structure of AGPase for the first time and reveals many important interactions between the subunits. Our AGPase model and the key residues defined here can be used as a guideline for further studies that involve engineering a more stable and efficient form of the enzyme.

APPENDIX

A.1. MM-PBSA Calculations

Table A.1.1[¥]: Free energy components of Dimer 1 calculated by MM-PBSA method.

	Dimer 1	LS	LS	ΔG
ΔE_{elec}	-25217.40±239.53	-12482.28±134.53	-12442.63±115.52	-292.48±5.49
ΔE_{VDW}	-4024.81±44.95	-1914.21±32.97	-1919.87±31.17	-190.73±0.73
ΔE_{int}	14774.79±84.72	7398.45±58.25	7376.34± 63.53	0.01±0.0
ΔG_{gas}	-14467.41±237.35	-6998.05±135.13	-6986.16±120.71	-483.20±5.76
$\Delta G_{non\text{-}polar}$	212.64±1.93	115.68±1.60	115.30±1.35	-18.35±0.06
ΔG_{polar}	-12802.24±252.07	-6593.68±123.69	-6629.97±130.87	421.41±10.38
ΔG_{sol}	-12589.60±251.45	-6477.99±123.45	-6514.66±130.46	403.06±10.36
ΔG_{elec}	-38019.63±79.87	-19075.96±61.58	-19072.60±57.57	128.92±7.99
ΔG_{Total}	-27057.01±97.71	-13476.04±78.62	-13500.83±81.39	**-80.15±7.96**

Table A.1.2[¥]: Free energy components of Dimer 2 calculated by MM-PBSA method.

	Dimer 2	LS	SS	ΔG
ΔE_{elec}	-25140.22±248.23	-12529.52±118.06	-12112.58±173.37	-498.12±3.12
ΔE_{VDW}	-4036.94±43.40	-1927.27±30.86	-1918.97±32.01	-190.70±0.73
ΔE_{int}	14748.58±80.92	7418.21±60.21	7330.35±62.62	0.01±0.0
ΔG_{gas}	-14428.58±258.26	-7038.58±124.87	-6701.19±177.43	-688.81±3.28
$\Delta G_{non\text{-}polar}$	212.44±2.71	115.68±1.52	115.72±2.16	-18.97±0.06
ΔG_{polar}	-12989.67±254.74	-6528.57±127.64	-7049.29±175.20	588.19±5.66
ΔG_{sol}	-12777.24±252.78	-6412.88±127.12	-6933.57±173.58	569.22±5.65
ΔG_{elec}	-38129.90±66.07	-19058.09±63.43	-19161.87±41.48	90.06±4.77
ΔG_{Total}	-27205.82±84.07	-13451.46±73.55	-13634.76±62.58	**-119.59±4.85**

Table A.1.3[¥]: Free energy components of Dimer 3 calculated by MM-PBSA method.

	Dimer 3	SS	SS	ΔG
ΔE_{elec}	-24618.61±151.23	-12271.20±103.64	-12180.40±119.09	-167.01±2.92
ΔE_{VDW}	-4044.03±44.67	-1945.91±32.60	-1911.06±32.86	-187.05±0.53
ΔE_{int}	14698.07±84.29	7352.64±62.66	7345.41±55.43	0.01±0.0
ΔG_{gas}	-13964.57±173.03	-6864.47±119.02	-6746.04±129.91	-354.05±2.93
$\Delta G_{non-polar}$	211.27±2.63	112.70±1.74	116.05±1.57	-17.49±0.04
ΔG_{polar}	-13563.40±140.50	-6869.09±97.10	-6986.78±117.70	292.47±3.53
ΔG_{sol}	-13352.13±139.41	-6756.39±96.25	-6870.73±117.11	274.98±3.52
ΔG_{elec}	-38182.01±55.46	-19140.29±38.62	-19167.17±36.16	125.46±2.64
ΔG_{Total}	-27316.70±87.90	-13620.86±64.83	-13616.77±56.50	**-79.07±2.24**

Table A.1.4[¥]: Free energy components of Dimer 4 calculated by MM-PBSA method.

	Dimer 4	LS	LS	ΔG
ΔE_{elec}	-25596.72±167.52	-12776.17±135.50	-12576.97±103.40	-243.58±3.61
ΔE_{VDW}	-3972.87±45.96	-1951.91±29.33	-1908.73±31.37	-112.23±0.81
ΔE_{int}	14796.37±84.01	7409.08±54.72	7387.28±62.27	0.01±0.0
ΔG_{gas}	-14773.22±168.84	-7319.00±134.70	-7098.42±108.51	-355.80±3.92
$\Delta G_{non-polar}$	213.53±2.49	110.46±1.70	115.22±1.72	-12.15±0.07
ΔG_{polar}	-12418.67±169.17	-6257.37±126.16	-6522.48±107.65	361.19±5.42
ΔG_{sol}	-12205.13±168.50	-6146.91±125.09	-6407.26±107.33	349.03±5.39
ΔG_{elec}	-38015.39±97.21	-19033.54±61.56	-19099.45±66.88	117.60±4.43
ΔG_{Total}	-26978.36±110.94	-13465.91±71.96	-13505.68±83.95	**-6.76±4.33**

Table A.1.5[¥]: Free energy components of Dimer 5 calculated by MM-PBSA method.

	Dimer 5	LS	SS	ΔG
ΔE_{elec}	-25072.17±125.37	-12588.21±102.90	-12176.00±98.90	-307.96±3.80
ΔE_{VDW}	-3991.28±47.07	-1951.62±33.46	-1935.63±31.92	-104.04±0.69
ΔE_{int}	14722.83±78.73	7392.99±61.37	7329.83±59.02	0.01±0.0
ΔG_{gas}	-14340.62±149.08	-7146.83±122.20	-6781.80±119.68	-411.99±3.45
$\Delta G_{non-polar}$	216.01±3.42	111.18±1.82	115.66±1.99	-10.83±0.04
ΔG_{polar}	-13086.55±139.45	-6477.95±126.51	-6996.23±97.07	387.62±6.44
ΔG_{sol}	-12870.54±137.36	-6366.77±125.41	-6880.57±96.16	376.79±6.44
ΔG_{elec}	-38158.72±62.62	-19066.15±63.03	-19172.23±34.61	79.66±4.70
ΔG_{Total}	-27211.16±87.46	-13513.60±78.51	-13662.37±64.33	**-35.19±4.83**

Table A.1.6[¥]: Free energy components of Dimer 6 calculated by MM-PBSA method.

	Dimer 6	SS	SS	ΔG
ΔE_{elec}	-25096.78±152.06	-12363.92±109.76	-12366.52±119.69	-366.34±3.25
ΔE_{VDW}	-3981.24±46.98	-1942.64±31.95	-1938.64±34.74	-99.96±0.62
ΔE_{int}	14691.19±86.51	7356.43±65.52	7334.75±62.29	0.01±0.0
ΔG_{gas}	-14386.83±162.05	-6950.12±116.79	-6970.41±123.19	-466.30±3.45
$\Delta G_{non\text{-}polar}$	213.69±2.02	112.35±1.58	112.00±1.49	-10.67±3.75
ΔG_{polar}	-13135.27±142.74	-6770.81±104.30	-6785.44±107.09	420.99±3.79
ΔG_{sol}	-12921.58±141.88	-6658.46±103.52	-6673.44±106.41	410.32±3.75
ΔG_{elec}	-38232.04±45.54	-19134.73±34.16	-19151.96±41.86	54.64±1.79
ΔG_{Total}	-27308.40±85.01	-13608.58±63.84	-13643.85±59.54	**-55.97±1.48**

¥: Values are in kcal/mol.

A.2. MM-GBSA Calculations

Table A.2.1[*]: Free energy components of Dimer 1 calculated by MM-GBSA method.

	Dimer 1	LS	LS	ΔG
ΔE_{elec}	-25005.23±154.00	-12414.07±100.78	-12322.61±129.60	-268.55±45.91
ΔE_{VDW}	-4005.63±48.99	-1910.78±34.40	-1904.75± 32.10	-190.10±8.44
ΔE_{int}	14729.44±86.84	7376.88±62.14	7352.54±60.20	0.01±0.02
ΔG_{gas}	-14281.43±155.37	-6947.98±114.55	-6874.81±128.14	-458.64±47.51
$\Delta G_{non\text{-}polar}$	194.34±1.51	106.34±1.17	106.35±0.84	-18.36±0.51
ΔG_{polar}	-10939.80±136.19	-5605.09±93.38	-5676.24±108.80	341.53±43.20
ΔG_{sol}	-10745.46±135.64	-5498.74±93.01	-5569.89±108.62	323.17±43.05
ΔG_{elec}	-35945.03±49.83	-18019.16±33.32	-17998.86±36.83	72.98±8.00
ΔG_{Total}	-25026.89±86.77	-12446.72±60.64	-12444.71±61.27	**-135.47±9.14**

Table A.2.2[*]: Free energy components of Dimer 2 calculated by MM-GBSA method.

	Dimer 2	LS	SS	ΔG
ΔE_{elec}	-24939.61±164.08	-12293.21±138.69	-12137.67±184.14	-508.73±40.66
ΔE_{VDW}	-4013.02±45.11	-1920.18±34.56	-1905.22±31.06	-187.61±7.63
ΔE_{int}	14731.21±86.51	7400.37±64.95	7330.83±62.74	0.01±0.02
ΔG_{gas}	-14221.42±171.37	-6813.02±137.43	-6712.07±188.23	-696.33±43.04
$\Delta G_{non\text{-}polar}$	195.41±1.42	106.86±1.61	107.03±1.08	-18.48±0.67
ΔG_{polar}	-11215.40±150.14	-5732.95±124.28	-6063.58±173.60	581.13±39.68
ΔG_{sol}	-11019.99±150.01	-5626.09±123.00	-5956.55±172.90	562.65±39.37
ΔG_{elec}	-36155.02±51.38	-18026.16±34.69	-18201.26±33.55	72.40±7.62
ΔG_{Total}	-25241.41±80.20	-12439.12±58.48	-12668.62±62.48	**-133.67±7.07**

Table A.2.3*: Free energy components of Dimer 3 calculated by MM-GBSA method.

	Dimer 3	SS	SS	ΔG
ΔE_{elec}	-24737.03±116.17	-12316.87±96.25	-12284.94±93.64	-135.23±32.87
ΔE_{VDW}	-4009.70±44.48	-1927.97±34.75	-1906.92±32.56	-174.80±8.88
ΔE_{int}	14666.59±83.34	7326.70±64.22	7339.88±53.92	0.01±0.02
ΔG_{gas}	-14080.14±125.12	-6918.15±101.30	-6851.98±105.64	-310.02±36.47
$\Delta G_{non-polar}$	193.02±1.65	104.40±1.00	105.87±1.13	-17.25±0.52
ΔG_{polar}	-11612.84±97.94	-5902.00±82.39	-5920.58±82.00	209.74±30.36
ΔG_{sol}	-11419.82±97.90	-5797.60±82.13	-5814.71±81.91	192.50±30.06
ΔG_{elec}	-36349.88±44.79	-18218.87±31.68	-18205.52±32.17	74.52±7.94
ΔG_{Total}	-25499.96±82.45	-12715.75±60.05	-12666.69±57.38	**-117.52±9.69**

Table A.2.4*: Free energy components of Dimer 4 calculated by MM-GBSA method.

	Dimer 4	LS	LS	ΔG
ΔE_{elec}	-25505.22±177.31	-12795.36±122.10	-12489.52±116.57	-220.34±65.24
ΔE_{VDW}	-3995.59±47.80	-1965.58±35.66	-1916.27±28.69	-113.75±6.81
ΔE_{int}	14770.88±80.50	7399.23±57.03	7371.64±54.97	0.01±0.02
ΔG_{gas}	-14729.93±167.46	-7361.72±119.85	-7034.14±116.13	-334.07±64.86
$\Delta G_{non-polar}$	193.93±1.38	101.36±0.94	106.18±0.85	-13.62±0.59
ΔG_{polar}	-10480.07±140.73	-5238.81±102.12	-5544.80±96.41	303.54±58.75
ΔG_{sol}	-10286.14±140.02	-5137.44±101.73	-5438.62±96.17	289.92±58.35
ΔG_{elec}	-35985.29±64.37	-18034.17±38.57	-18034.32±39.91	83.20±12.11
ΔG_{Total}	-25016.07±80.49	-12499.16±57.68	-12472.76±58.66	**-44.15±9.53**

Table A.2.5*: Free energy components of Dimer 5 calculated by MM-GBSA method.

	Dimer 5	LS	SS	ΔG
ΔE_{elec}	-25016.15±146.39	-12503.59±99.61	-12163.75±92.15	-348.80±43.28
ΔE_{VDW}	-3964.72±45.31	-1931.69±30.83	-1929.35±31.08	-103.68±8.28
ΔE_{int}	14708.54±77.27	7381.68±61.54	7326.85±56.38	0.01±0.01
ΔG_{gas}	-14272.33±161.90	-7053.61±108.58	-6766.25±105.70	-452.47±44.88
$\Delta G_{non-polar}$	198.97±1.57	103.84±0.93	106.65±1.01	-11.52±0.55
ΔG_{polar}	-11160.25±135.61	-5536.94±83.61	-6040.99±82.29	417.69±38.88
ΔG_{sol}	-10961.27±135.01	-5433.10±83.39	-5934.34±82.04	406.17±38.70
ΔG_{elec}	-36176.39±45.12	-18040.54±35.49	-18204.74±32.39	68.88±9.36
ΔG_{Total}	-25233.60±76.27	-12486.72±58.20	-12700.59±57.75	**-46.30±8.86**

Table A.2.6*: Free energy components of Dimer 6 calculated by MM-GBSA method.

	Dimer 6	SS	SS	ΔG
ΔE_{elec}	-25161.78±137.40	-12362.05±87.37	-12408.43±108.25	-391.30±39.44
ΔE_{VDW}	-3985.81±47.54	-1943.29±32.18	-1941.51±34.35	-101.01±8.19
ΔE_{int}	14711.34±84.35	7360.11±55.77	7351.21±64.76	0.01±0.02
ΔG_{gas}	-14436.25±148.57	-6945.22±92.77	-6998.73±117.19	-492.30±37.65
$\Delta G_{non\text{-}polar}$	193.59±1.15	102.68±0.95	102.74±0.80	-11.83±0.42
ΔG_{polar}	-11226.43±117.19	-5841.49±72.94	-5841.04±96.79	456.11±36.18
ΔG_{sol}	-11032.83±116.94	-5738.81±72.80	-5738.31±96.44	444.28±36.16
ΔG_{elec}	-36388.20±47.42	-18203.54±33.58	-18249.47±31.04	64.81±9.08
ΔG_{Total}	-25469.08±86.05	-12684.03±58.69	-12737.04±62.58	**-48.01±7.06**

*: Values are in kcal/mol.

B.1. Sample Configuration File for Molecular Dynamics Simulations

```
amber                   on
parmfile                ../common/Dimer2.prmtop
ambercoor               ../common/Dimer2.inpcrd
outputname              Dimer2_minimization
restartname             Dimer2_res_minimization

binaryoutput            yes
rigidbonds              all
rigidTolerance          0.00001
outputEnergies          1000
outputPressure          1000
xstfreq                 1000
restartfreq             1000
DCDfreq                 2000
timestep                2.
temperature             0
cutoff                  9.
switching               off
stepspercycle           10
nonbondedFreq           1
fullElectFrequency      2
readexclusions          no

PME                     on
cellBasisVector1        106.   0.   0.
cellBasisVector2        0.  117.   0.
cellBasisVector3        0.    0.  89.
cellOrigin              32.   10.  50.

PMEGridSizeX            108
PMEGridSizeY            120
PMEGridSizeZ            90
wrapAll                 on
margin                  5

exclude                 scaled1-4
1-4scaling              0.833333
scnb                    2

fixedAtoms              on
fixedAtomsForces        on
fixedAtomsFile          ../common/fix_backbone.pdb
fixedAtomsCol           B

constraints             on
```

```
consRef                 ../common/restrain_ca.pdb
consKFile               ../common/restrain_ca.pdb
consKCol                B

langevin                on
langevinDamping         5
langevinTemp            0
langevinHydrogen        off

useGroupPressure        yes
useFlexibleCell         no
useConstantArea         no

langevinPiston           on
langevinPistonTarget    1.01325
langevinPistonPeriod    100.
langevinPistonDecay     50.
langevinPistonTemp      300

minimize                0

langevinPiston          off

minimize                10000
output                  min_fix

fixedAtoms              off
minimize                10000
output                  min_all

for { set TEMP 10 } { $TEMP < 300 } { incr TEMP 10 } {
        langevinTemp            $TEMP
        run 5000
}
langevinTemp            300
output                  heat

langevinPiston          on
run                     10000
output                  equil_ca_1

constraintScaling 0.75
run                     10000
output                  equil_ca_075

constraintScaling        0.50
run                     10000
```

```
output                  equil_ca_050

constraintScaling       0.25
run                     10000
output                  equil_ca_025

constraintScaling       0
run                     50000
output                  equil_ca_0
```

B.2. Sample Run File for AMBER8 MM-PBSA Binding Free Energy Calculations

```
@GENERAL

PREFIX          Dimer2
PATH            ../snapshots

COMPLEX         1
RECEPTOR        1
LIGAND          1

COMPT           ../../common/Dimer2.prmtop
RECPT           ../../common/LS.prmtop
LIGPT           ../../common/SS.prmtop

GC              0
AS              0
DC              0

MM              1
GB              1
PB              1
MS              0
NM              0

@PB

PROC            2
REFE            0
INDI            1.0
EXDI            80.0
SCALE           2.0
LINIT           1000
PRBRAD          1.4
```

```
RADIOPT          0

SURFTEN  0.00542
SURFOFF      0.92

@MM

DIELC            1.0

@GB

IGB              1
GBSA             1
SALTCON      0.00
EXTDIEL      80.0
INTDIEL          1.0

SURFTEN    0.00542
SURFOFF      0.92

@MS

PROBE            0.0
```

B.3. Sample Run File for AMBER8 MM-GBSA Binding Free Energy Calculations

```
PREFIX           Dimer2
PATH             ../snapshots

COMPLEX          1
RECEPTOR         1
LIGAND           1

COMPT            ../common/lsmbondi2.prmtop
RECPT            ../common/chainAmbondi2.prmtop
LIGPT            ../common/chainBmbondi2.prmtop

GC               0
AS               0
DC               0

MM               1
GB               1
PB               0
MS               0
```

```
NM                  0

@PB

PROC                2
REFE                0
INDI              1.0
EXDI             80.0
SCALE             2.0
LINIT            1000
PRBRAD            1.6

RADIOPT             1

SURFTEN         0.005
SURFOFF           0.0

@MM

DIELC             1.0

@GB

IGB                 2
GBSA                2
SALTCON          0.00
EXTDIEL          80.0
INTDIEL           1.0

SURFTEN         0.005
SURFOFF          0.00

@MS

PROBE             0.0
```

B.4. Sample Run File for AMBER8 MM-GBSA Free Energy Decomposition

```
PREFIX              Dimer2
PATH                ../snapshots

COMPLEX             1
RECEPTOR            1
LIGAND              1
```

```
COMPT                ../common/Dimer2.prmtop
RECPT                ../common/LS.prmtop
LIGPT                ../common/SS.prmtop

GC                    0
AS                    0
DC                    1

MM                    1
GB                    1
PB                    0
MS                    0
NM                    0

@DECOMP

DCTYPE                2

COMREC              1-441
COMLIG            442-883
COMPRI   1-441 442-883
RECRES              1-441
RECPRI              1-441
RECMAP              1-441
LIGRES              1-442
LIGPRI              1-442
LIGMAP            442-883


@MM

DIELC                 1.0

@GB

IGB                   2
GBSA                  2
SALTCON            0.00
EXTDIEL            80.0
INTDIEL             1.0

SURFTEN           0.005
SURFOFF            0.00
```

BIBLIOGRAPHY

1. Martin C, Smith AM. Starch biosynthesis. Plant Cell. 1995 Jul;7(7):971-85.

2. Preiss J. Biology and molecular biology of starch synthesis and its regulation. Oxf Surv Plant Mol Cell Biol. 1991;7:59-114.

3. Preiss J, Sivak MN. Starch and glycogen biosynthesis. In Comprehensive Natural Products Chemistry. Pinto BM ed. Oxford, UK: Pergamon Press; 1998. p. 441-95.

4. Okita TW. Is There an Alternative Pathway for Starch Synthesis? Plant Physiol. 1992 Oct;100(2):560-4.

5. Okita TW, Nakata P, Smith-White B, Preiss J. Enhancement of plant productivity by manipulation of ADPglucose pyrophosphorylase. In Gene Conservation and Exploitation. New York: Plenum Press; 1993. p. 161-91.

6. Preiss J, Sivak MN. Starch Synthesis in sinks and sources. In Photoassimilate distribution in plants an crops: source-sink relationships In: Zamski E, Schaffer AA, editors. New York:Dekker; 1995. p. 139-68.

7. Stark DM, Timmerman KP, Barry GF, Preiss J, Kishore GM. Regulation of the Amount of Starch in Plant Tissues by ADP Glucose Pyrophosphorylase. Science. 1992;258:287-92.

8. Iglesias AA, Barry GF, Meyer C, Bloksberg L, Nakata PA, Greene T, et al. Expression of the potato tuber ADP-glucose pyrophosphorylase in Escherichia coli. J Biol Chem. 1993 Jan 15;268(2):1081-6.

9. Okita TW, Nakata PA, Anderson JM, Sowokinos J, Morell M, Preiss J. The Subunit Structure of Potato Tuber ADPglucose Pyrophosphorylase. Plant Physiol. 1990 Jun;93(2):785-90.

10. Greene TW, Hannah LC. Enhanced stability of maize endosperm ADP-glucose pyrophosphorylase is gained through mutants that alter subunit interactions. Proc Natl Acad Sci U S A. 1998 Oct 27;95(22):13342-7.

11. Jin X, Ballicora MA, Preiss J, Geiger JH. Crystal structure of potato tuber ADP-glucose pyrophosphorylase. EMBO J. 2005 Feb 23;24(4):694-704.

12. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res. 2000 Dec;33(12):889-97.

13. Srinivasan J, Cheatham TE, 3rd, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. J Am Chem Soc. 1998 Sep;120(37):9401-09.

14. Fields S, Song O. A novel genetic system to detect protein-protein interactions. Nature. 1989 Jul 20;340(6230):245-6.

15. Nakata PA, Greene TW, Anderson JM, Smith-White BJ, Okita TW, Preiss J. Comparison of the primary sequences of two potato tuber ADP-glucose pyrophosphorylase subunits. Plant Mol Biol. 1991 Nov;17(5):1089-93.

16. Smith-White BJ, Preiss J. Comparison of proteins of ADP-glucose pyrophosphorylase from diverse sources. J Mol Evol. 1992 May;34(5):449-64.

17. Sowokinos JR. Pyrophosphorylases in Solanum tuberosum: II. Catalytic properties and regulation of ADP-Glucose and UDP-Glucose pyrophosphorylase activities in potatoes. Plant Physiol. 1981 Oct;68(4):924-9.

18. Ballicora MA, Frueauf JB, Fu Y, Schurmann P, Preiss J. Activation of the potato tuber ADP-glucose pyrophosphorylase by thioredoxin. J Biol Chem. 2000 Jan 14;275(2):1315-20.

19. Fu Y, Ballicora MA, Leykam JF, Preiss J. Mechanism of reductive activation of potato tuber ADP-glucose pyrophosphorylase. J Biol Chem. 1998 Sep 25;273(39):25045-52.

20. Ballicora MA, Fu Y, Nesbitt NM, Preiss J. ADP-Glucose pyrophosphorylase from potato tubers. Site-directed mutagenesis studies of the regulatory sites. Plant Physiol. 1998 Sep;118(1):265-74.

21. Ballicora MA, Laughlin MJ, Fu Y, Okita TW, Barry GF, Preiss J. Adenosine 5'-diphosphate-glucose pyrophosphorylase from potato tuber. Significance of the N terminus of the small subunit for catalytic properties and heat stability. Plant Physiol. 1995 Sep;109(1):245-51.

22. Frueauf JB, Ballicora MA, Preiss J. ADP-glucose pyrophosphorylase from potato tuber: site-directed mutagenesis of homologous aspartic acid residues in the small and large subunits. Plant J. 2003 Feb;33(3):503-11.

23. Greene TW, Chantler SE, Kahn ML, Barry GF, Preiss J, Okita TW. Mutagenesis of the potato ADPglucose pyrophosphorylase and characterization of an allosteric mutant defective in 3-phosphoglycerate activation. Proc Natl Acad Sci U S A. 1996 Feb 20;93(4):1509-13.

24. Kavakli IH, Park JS, Slattery CJ, Salamone PR, Frohlick J, Okita TW. Analysis of allosteric effector binding sites of potato ADP-glucose pyrophosphorylase through reverse genetics. J Biol Chem. 2001 Nov 2;276(44):40834-40.

25. Salamone PR, Greene TW, Kavakli IH, Okita TW. Isolation and characterization of a higher plant ADP-glucose pyrophosphorylase small subunit homotetramer. FEBS Lett. 2000 Sep 29;482(1-2):113-8.

26. Laughlin MJ, Chantler SE, Okita TW. N- and C-terminal peptide sequences are essential for enzyme assembly, allosteric, and/or catalytic properties of ADP-glucose pyrophosphorylase. Plant J. 1998 Apr;14(2):159-68.

27. Cross JM, Clancy M, Shaw JR, Boehlein SK, Greene TW, Schmidt RR, et al. A polymorphic motif in the small subunit of ADP-glucose pyrophosphorylase modulates interactions between the small and large subunits. Plant J. 2005 Feb;41(4):501-11.

28. Tramontano A, Morea V. Assessment of homology-based predictions in CASP5. Proteins. 2003;53 Suppl 6:352-68.

29. Sanchez R, Sali A. Advances in comparative protein-structure modelling. Curr Opin Struct Biol. 1997 Apr;7(2):206-14.

30. Rost B. Twilight zone of protein sequence alignments. Protein Eng. 1999 Feb;12(2):85-94.

31. Saqi MA, Russell RB, Sternberg MJ. Misleading local sequence alignments: implications for comparative protein modelling. Protein Eng. 1998 Aug;11(8):627-30.

32. Schwede T, Diemand A, Guex N, Peitsch MC. Protein structure computing in the genomic era. Res Microbiol. 2000 Mar;151(2):107-12.

33. Sanchez R, Sali A. Evaluation of comparative protein structure modeling by MODELLER-3. Proteins. 1997;Suppl 1:50-8.

34. Totrov M, Abagyan R. Detailed ab initio prediction of lysozyme-antibody complex with 1.6 A accuracy. Nat Struct Biol. 1994 Apr;1(4):259-63.

35. Vakser IA. Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex. Proteins. 1997;Suppl 1:226-30.

36. Stillinger FH, Rahman A. Improved simulation of liquid water by molecular dynamics. J Chem Phys. 1974;60:1545-57.

37. Ma JP, Sigler PB, Xu ZH, Karplus M. A dynamic model for the allosteric mechanism of GroEL. J Mol Biol. 2000;302:303-13.

38. Tai K, Shen T, Borjesson U, Philippopoulos M, McCammon JA. Analysis of a 10-ns molecular dynamics simulation of mouse acetylcholinesterase. Biophys J. 2001 Aug;81(2):715-24.

39. Cavalli A, Ferrara P, Caflisch A. Weak temperature dependence of the free energy surface and folding pathways of structured peptides. Proteins. 2002 May 15;47(3):305-14.

40. de Groot BL, Grubmuller H. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. Science. 2001 Dec 14;294(5550):2353-7.

41. Selzer T, Albeck S, Schreiber G. Rational design of faster associating and tighter binding protein complexes. Nat Struct Biol. 2000 Jul;7(7):537-41.

42. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. Proc Natl Acad Sci U S A. 2002 Oct 29;99(22):14116-21.

43. Novotny J, Bruccoleri RE, Saul FA. On the attribution of binding energy in antigen-antibody complexes McPC 603, D1.3, and HyHEL-5. Biochemistry. 1989 May 30;28(11):4735-49.

44. Liu Y, Beveridge DL. Exploratory studies of ab initio protein structure prediction: multiple copy simulated annealing, AMBER energy functions, and a generalized born/solvent accessibility solvation model. Proteins. 2002 Jan 1;46(1):128-46.

45. Gouda H, Kuntz ID, Case DA, Kollman PA. Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and thermodynamic integration methods. Biopolymers. 2003 Jan;68(1):16-34.

46. Huo S, Wang J, Cieplak P, Kollman PA, Kuntz ID. Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design. J Med Chem. 2002 Mar 28;45(7):1412-9.

47. Masukawa KM, Kollman PA, Kuntz ID. Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. J Med Chem. 2003 Dec 18;46(26):5628-37.

48. Wang J, Morin P, Wang W, Kollman PA. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to

HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc. 2001 Jun 6;123(22):5221-30.

49. Gohlke H, Case DA. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem. 2004 Jan 30;25(2):238-50.

50. Gohlke H, Kiel C, Case DA. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J Mol Biol. 2003 Jul 18;330(4):891-913.

51. Strockbine B, Rizzo RC. Binding of antifusion peptides with HIVgp41 from molecular dynamics simulations: quantitative correlation with experiment. Proteins. 2007 May 15;67(3):630-42.

52. Wang W, Kollman PA. Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. J Mol Biol. 2000 Nov 3;303(4):567-82.

53. Zoete V, Meuwly M, Karplus M. Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. Proteins. 2005 Oct 1;61(1):79-93.

54. Huo S, Massova I, Kollman PA. Computational alanine scanning of the 1:1 human growth hormone-receptor complex. J Comput Chem. 2002 Jan 15;23(1):15-27.

55. Massova I, Kollman PA. Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. J Am Chem Soc. 1999;121(36):8133-43.

56. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol. 1998 Jul 3;280(1):1-9.

57. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. Science. 1995 Jan 20;267(5196):383-6.

58. Thorn KS, Bogan AA. ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. Bioinformatics. 2001 Mar;17(3):284-5.

59. DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol. 2002 Feb;12(1):14-20.

60. Lafont V, Schaefer M, Stote RH, Altschuh D, Dejaegere A. Protein-protein recognition and interaction hot spots in an antigen-antibody complex: free energy decomposition identifies "efficient amino acids". Proteins. 2007 May 1;67(2):418-34.

61. Chien CT, Bartel PL, Sternglanz R, Fields S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc Natl Acad Sci U S A. 1991 Nov 1;88(21):9578-82.

62. Gyuris J, Golemis E, Chertkov H, Brent R. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. Cell. 1993 Nov 19;75(4):791-803.

63. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc Natl Acad Sci U S A. 2001 Apr 10;98(8):4569-74.

64. Mrowka R, Patzak A, Herzel H. Is there a bias in proteome research? Genome Res. 2001 Dec;11(12):1971-3.

65. Jarillo JA, Capel J, Tang RH, Yang HQ, Alonso JM, Ecker JR, et al. An Arabidopsis circadian clock component interacts with both CRY1 and phyB. Nature. 2001 Mar 22;410(6827):487-90.

66. Ouellet F, Overvoorde PJ, Theologis A. IAA17/AXR3: biochemical insight into an auxin mutant phenotype. Plant Cell. 2001 Apr;13(4):829-41.

67. Davies B, Egea-Cortines M, de Andrade Silva E, Saedler H, Sommer H. Multiple interactions amongst floral homeotic MADS box proteins. EMBO J. 1996 Aug 15;15(16):4330-43.

68. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, et al. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature. 2000 Feb 10;403(6770):623-7.

69. Greene TW, Hannah LC. Maize endosperm ADP-glucose pyrophosphorylase SHRUNKEN2 and BRITTLE2 subunit interactions. Plant Cell. 1998 Aug;10(8):1295-306.

70. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, et al. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 2003 Jul 1;31(13):3497-500.

71. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997 Dec;18(15):2714-23.

72. Kopp J, Schwede T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. Nucleic Acids Res. 2004 Jan 1;32(Database issue):D230-4.

73. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. 2003 Jul 1;31(13):3381-5.

74. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006 Jan 15;22(2):195-201.

75. Jorgensen WL, Chandrasekhar J, Madura J, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983;79:926-35.

76. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005 Dec;26(16):1781-802.

77. Kollman PA, Dixon R, Cornell W, Fox T, Chipot C, Pohorille A. The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of ab initio calculations and experimental data. In Computer

Simulation of Biomolecular Systems, Vol 3, A Wilkinson, P Weiner & WF van Gunsteren, Ed Elsevier, pp 83-96; 1997.

78. Brooks CLI, Karplus M, Pettitt BM. A theoretical perspective of dynamics structure and thermodynamics. In: Prigogine I, Rice SA, editors Advances in chemical physics Vol 71 New York: Wiley-Interscience; 1988.

79. Darden T, York D, Pedersen L. Particle Mesh-Ewald - an N Log(N) method for Ewald sums in large systems. J Chem Phys. 1993;98(12):10089-92.

80. Ryckaert JP, C, G. , Berendsen HJC. Numerical-integration of Cartesian equations of motion of a system with constraints molecular dynamics of n-alkanes. J Comput Phys. 1977;23:327-41.

81. Feller SE, Zhang Y, W. PR, Brooks BR. Constant pressure molecular dynamics simulation: The Langevin Piston Method. J Chem Phys. 1995;103(11):4613-21.

82. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. J Chem Phys. 1994;101(5):4177-89.

83. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol. 1971 Feb 14;55(3):379-400.

84. Hubbard SJ, Thornton JM. 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London. 1993.

85. Gilson MK, Honig BH. Energetics of charge-charge interactions in proteins. Proteins. 1988;3(1):32-52.

86. Amidon GL, Yalkowsky SH, Anik ST, Valvani SC. Solubility of non-electrolytes şn polar solvents. V. Estimation of the solubilty of aliphatic monofunctional compounds in water using a molecular surface area approach. J Phys Chem. 1975;79:2239-46.

87. Hermann RB. Theory of hydrophobic bonding. II. Correlation of hydrocarbon solubility in water with solvent cavity surface area. J Phys Chem. 1972;76(19):2754-59.

88. Massova I, Kollman PA. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict lignad binding. Persp Drug Disc Des. 2000;18:113-35.

89. Wang W, Kollman PA. Computational study of protein specificity: the molecular basis of HIV-1 protease drug resistance. Proc Natl Acad Sci U S A. 2001 Dec 18;98(26):14937-42.

90. Zhang N, Jiang Y, Zou J, Zhuang S, Jin H, Yu Q. Insights into unbinding mechanisms upon two mutations investigated by molecular dynamics study of GSK3beta-axin complex: role of packing hydrophobic residues. Proteins. 2007 Jun 1;67(4):941-9.

91. Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, Merz KM, Jr., et al. The Amber biomolecular simulation programs. J Comput Chem. 2005 Dec;26(16):1668-88.

92. Kuhn B, Kollman PA. Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. J Med Chem. 2000 Oct 5;43(20):3786-91.

93. Noskov SY, Lim C. Free energy decomposition of protein-protein interactions. Biophys J. 2001 Aug;81(2):737-50.

94. Sitkoff D, Sharp KA, Honig BH. Accurate calculation of hydration free-energies using macroscopic solvent models. J Phys Chem. 1994;98:1978-88.

95. Weiser J, Shenkin PS, Still WC. Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities. Biopolymers. 1999 Oct 5;50(4):373-80.

96. Onufriev A, Bashford D, Case DA. Modification of the Generalized Born Model Suitable for Macromolecules. J Phys Chem B. 2000;104:3712-20.

97. Armon A, Graur D, Ben-Tal N. ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. J Mol Biol. 2001 Mar 16;307(1):447-63.

VITA

Aytuğ Tunçel was born in İstanbul, 1982. After graduating from Middle East Technical University, Department of Molecular Biology and Genetics in 2005 he attended Computational Sciences and Engineering MS program at Koç University. He studied on the thesis project titled as “Modeling of Potato ADP-Glucose Pyrophosphorylase”. After graduating from Koç University he is going to attend Washington State University, Molecular Plant Sciences program, to do his Ph.D. study.