Structural Analysis and Mutational Assessment of the Photosystem II Chlorophyll-Binding Protein (CP47) in the Cyanobacterium *Synechocystis* sp. PCC 6803: Modeling and Simulation Study

Mahmood Janlou, Mehr Ali
Department of Biophysics, Faculty of Biological Sciences, Gorgan Branch, Islamic Azad University, Gorgan, I.R. IRAN

Sahebjamee, Hassan*[+]
Department of Biophysics, Faculty of Biological Science, Varamin-Pishva Branch, Islamic Azad University, Varamin, I.R. IRAN

Alaei, Hamid Reza
Department of Physics, Varamin-Pishva Branch, Islamic Azad University, Varamin, I.R. IRAN

Shokravi, Shademan
Department of Plant Biology, Faculty of Biological Sciences, Gorgan Branch, Islamic Azad University, Gorgan, I.R. IRAN

ABSTRACT: *CP47 is one of the essential components of photosystem II (PSII) in green plants, green algae, and cyanobacteria; which is involved in the light reactions of photosynthesis. Various studies have shown that the binding of the extrinsic protein of 33 kDa (PsbO) to the large extrinsic loop of CP47 (E loop) is an essential photoautotrophic activity of the PSII complex. Moreover, the deletion of the amino acids between Gly-351 and Thr-365 within loop E failed to assemble stable PSII centers. In this study, using computational methods, the effect of Phenylalanine (Phe) mutation at position 363 on Synechocystis sp. PCC 6803 CP47 was investigated and then the mutant model was compared with the native one. Because the experimental 3D structure of Synechocystis sp. PCC 6803 CP47 and PsbO proteins are not available in the Protein Data Bank (PDB), the 3D structure of these proteins was modeled by homology modeling. After refining and energy minimization, the quality of protein geometry was assessed by different criteria such as PROCHECK and ProSA. Then, structural analysis of mutant and native models was performed with Molecular Dynamic (MD) simulation and docking method. The analysis of results obtained from MD simulation shows that F363R mutation affects the flexibility of some regions and especially led to an increase in mutation region and changed conformation of CP47. In addition, the results of docking studies indicate that F363R mutation can decrease Buried Surface Area (BSA) at the interface region and decrease the binding energy of CP47 and PsbO. These data reinforce our hypothesis that an increase of flexibility at the position of F363 in the large extrinsic loop of CP47 may be an important factor in reducing interaction between CP47 and PsbO extrinsic protein and then water oxidation.*

KEYWORDS: *Photosystem II, CP47, Mutation, Molecular dynamic simulation, Docking.*

***** *To whom correspondence should be addressed.*
+ *E-mail: sahebjam.has@gmail.com & Sa.Hassan@iau.ac.ir*
1021-9986/2023/1/337-348 12/$/6.02

INTRODUCTION

Photosystem II (PSII) located in the thylakoid membrane of plant chloroplasts and cyanobacteria is the main protein complex for capturing the energy of sunlight and converting it into chemical potential energy. In PSII, chlorophylls and carotenoids are used as antennas to capture photons. The energy of the photons is used for the oxidation of water and reduction of plastoquinone (PLQ), which renders oxygen a waste product [1,2]. *In vivo*, the PSII complex is homodimer. Each monomer is composed of numerous components. The minimum number of protein subunits required to maintain a functional Water-Oxidizing Complex (WOC) is 27 and 20 in plants and cyanobacteria, respectively. However, the essential components of PSII core preparations are the chlorophyll a-binding antenna proteins CP47 and CP43, together with the reaction center proteins, D1 and D2, and the α– and β–subunits of cytochrome b-559, and additional low molecular mass polypeptides. A large number of cofactors, including ions and many glycolipids, may also be important [3]. Structurally, they have six transmembrane helical domains separated by five hydrophilic loop domains A-E, with the N- and C- terminal ends exposed at the cytosol [4]. The loops of A, C, and E are exposed at the lumen [5, 6]. Loop E, containing about 200 amino acids, is located between the fifth and sixth α-helices. In green plants and green algae, three extrinsic proteins with molecular masses of 33, 23, and 17 kDa (PsbO, PsbP, and PslQ) are bound to the luminal side of PSII. In non–green plants and cyanobacteria, PsbP and PsbQ are absent but two different extrinsic proteins, PsbU and cytochrome c-550 (PsbV) are present [7]. Various studies have shown that PsbO protein cross-linked to the E loop and protected it from amino acid modification and digestion with proteases [8, 9]. In addition, a lot of indirect evidence suggests that the large extrinsic loop (loop E) of CP47 and CP43 plays a role in water oxidation [10]. These studies indicate that the binding of PsbO involves the large extrinsic loop of CP47 with the amino acid region 360-440 being specifically important [4]. Also, a number of site-specific mutageneses have been created for amino acids within the large extrinsic loop of CP4. Examination of these substitutions shows that some of them retained a phenotype that was similar to that of the wild type such as P360G, RP357, 358 GG, E364Q, and E364G [11, 12]. Similarly, mutations in thr365, Gly361, Glu353, and Glu355 produced little or no detrimental effect on PSII stability or function [13]. However, biochemical studies by *Shannon et al*. [14] demonstrate that the F363R strain not only had impaired photoautotrophic growth and enhanced sensitivity to photoinactivation but also is required at position 363 for normal PSII function.

Although we know that F363R mutation may lead to enhanced sensitivity to photoinactivation, however, the molecular processes associated with this mutation have not yet been clearly determined. Therefore, structural studies at the atomic level would offer a better understanding of the impact of F363R mutation on CP47 structure, PsbO-CP47 interaction, and the mechanism of impaired photoautotrophic growth. Hence, in this investigation, by means of Molecular Dynamics (MD) simulation and docking study, the effect of F363R mutation on the molecular structure of CP47 and interaction with PsbO protein were considered. The former is a physics-based modeling method providing detailed information on the fluctuations and conformational changes of atoms and molecules in a variety of biological systems [15, 16]. The latter is performed by a HADDOCK server for determining protein-protein interaction analysis [17].

COMPUTATIONAL SECTION

Retrieval of proteins

The native protein sequences of *Synechocystis* sp. PCC 6803 CP47 and PsbO were retrieved from UniProt (*https://www.uniprot.org/*) (IDS: P05429 and Pl0549 respectively), which is a protein data repository maintained by the Swiss Institute of Bioinformatics. These sequences were further used for modeling protein 3D structure development.

Developing 3D structures

Since the 3D structures for wild-type CP47 and PsbO proteins of *Synechocystis* sp.PCC 6803 is not available in the Protein Data Bank (PDB) (*http://www.pdb.org/*), their structures were determined by SWISS-MODEL (*https://swissmodel.expasy.org/*), based on homology or comparative modeling [18, 19]. The program automatically provides an atom model using alignment between the query sequence and known homologous structure as a template. The template protein was searched through the Blastp program against the PDB database. The X-ray crystal structure of the *Thermosynechococcus*

vulcanus (PDB ID:5V2C) variant, a complex protein with over 20 chains, and a resolution of 1/92Å have been chosen. Chains B and O, which showed 82.61% and 56.25% identity with the target proteins of CP47 and PsbO respectively, were used as a template.

To improve the quality of predicted models of CP47 and PsbO, energy minimization was performed by CHIMERA (*https://www.cgl.ucsf.edu/chimera/*), a highly extensible program for analysis of molecular structures, at the steepest descent of 1000 steps followed by conjugated gradients of 500 steps.

Structural Validation

The constructed CP47 and PsbO structures were validated by the inspection of Φ/Ψ distributions of the Ramachandran plot obtained through PROCHECK analysis [20] (*http://services.mbi.ucla.edu/PROCHECK/*). The significance of consistency between the template and modeled CP47 and PsbO were evaluated using the ProSA server [21]

(*http://prosa.services.came.sbg.ac.at/prosa.php*).
In addition, the root means square deviation (RMSD) was analyzed by the 3D-ss server [22] on the superimposition of templates 5V2C_B, and O with predicted structures of CP47 and PsbO as models respectively to check the reliability of the model.

Induce mutagenesis in modeled native CP47

The program CHIMERA is also a graphical tool for computational-aided protein engineering. It implements the methodology of computational site-directed mutagenesis to design new protein mutants with the required properties. Therefore, using this program F363R mutant of CP47 was constructed and the energy minimization was done according to step 2.2 of materials and methods.

Embedding in the Thylakoid Membrane

The six transmembrane helical domains of CP47 including cofactors and ligands were embedded in a realistic thylakoid membrane of *T. vulcanus* [23]. Building and assembly of CP47/thylakoid membrane systems were achieved by the step-by-step protocol developed by Emilia and co-workers [24-25, 26] based on the procedure in CHARMM-GUI Membrane Builder [27-28, 29]. The composition of the membrane was taken from *van Eerden et al.* [30], which was experimentally determined by *Sakurai et al.* [23]. The modeled thylakoid membrane comprises a mixture of 10% 18:1(9)-16:0 phosphatidylglycerol (PG), 25% 18:1(9)-16:0 glycolipids digalactosyldiacylglycerol (DGDG), 40% 18:1(9)-16:0 monogalactosyldiacylglycerol (MGDG), 15% 18:1(9)-16:0 sulfoquinovosyldiacylglycerol (SQDG), and 10% of the lipids di16:0 SQDG.

Molecular dynamics simulations

The structure of native and mutant CP47 was used as starting point for molecular dynamics simulation. All MD simulations were carried out by Gromacs5.1.4 and membrane systems were equilibrated for 30 ns using CHARMM-GUI [31] with the C36 force field for lipids [32] and the TIP3P water model [33]. We used a 2-fs time step together with the Linear constraint solver (LINCS) algorithm [34]. NVT (constant particle number, volume, and temperature) and NPT (constant particle number, pressure, and temperature) CHARMM equilibrations were followed by a 30 ns NPT production run for all systems using Gromacs5.1.4 with the input scripts generated by CHARMM-GUI [35]. Periodic boundary conditions were applied to x = 12.58492 nm, y = 12.58492 nm, z = 14.20217 nm directions under the NPT ensemble. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm [36] allowing for an integration step of 2 fs. The smooth Particle Mesh Ewald (PME) methods [37] were used to calculate long-range electrostatic interactions. The trajectories were saved every 4 ps for analysis. We then computed the comparative analysis of structural deviations in native and mutant structures. RMSD, RMSF, SAS, and DSSP analysis were carried out by using g_rmsd, g_rmsf, g_sas, and do_ dssps tool respectively.

Docking

HADDOCK (*http://www.bonvinlab.org/software/haddock2.2*), is a flexible docking program where experimental or bioinformatics information is directly used to predict the binding affinity of protein-protein interactions [17]. It is user-friendly software, which uses biochemical or biophysical information to enhance sampling and scoring [17]. We have implemented some restrictions to run this software, the active and passive residues were our restrictions. In the CP47, the region

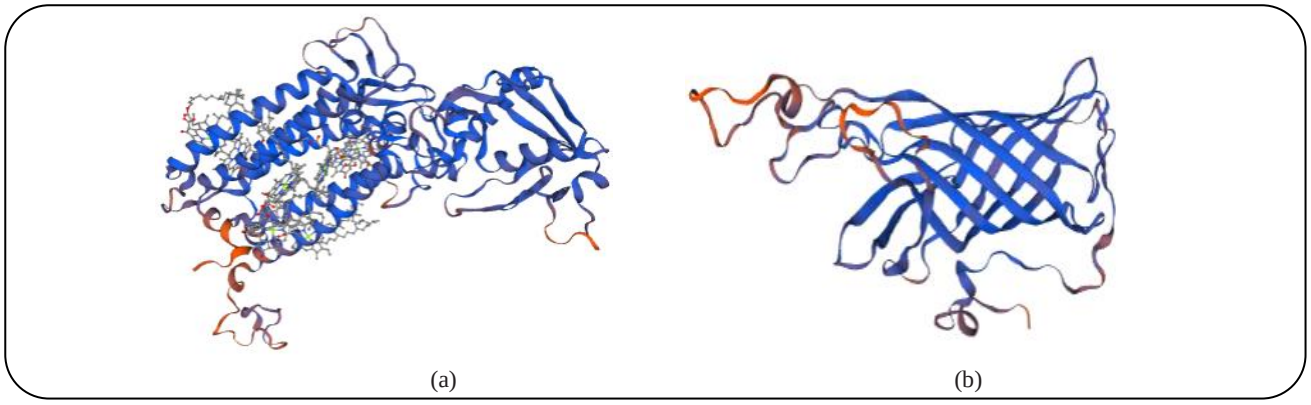

Fig. 1: Modeled three-dimensional (3D) structures of CP47 (a) and PsbO (b) based on known crystal structures of CP47 and PsbO.

of 360-440 is in contact with the region of 190-214 of PsbO. These two regions are active residues according to 5V2C_B, O, and *Komenda J. et al.* (2019) [4].

RESULTS AND DISCUSSION

Homolog modeling and Quality assessment of the models

The tertiary structure of a protein is needed to investigate and understand the ability of the protein to interact with other molecules and function. There is no crystal structure of native CP47 and PsbO protein in *Synechocystis* sp. PCC 68.8 is available in the RCSB PDB. So, after sequence alignment through the Blastp program, 5V2C_B and 5V2C_O are selected as templates, respectively. After sequence alignment by BlastP, the E_value between query and subject for CP47 and PsbO sequences was 0 and 5-e104, respectively.

SWISS-MODEL was used to generate model structures. The best models for CP47 and PsbO according to the least QMEAN and GMQE parameters were selected. Then, in order to improve the quality of initial modeled structures, energy minimization was performed by CHIMERA. The primary 3D model of CP47 and PsbO resulting from the structural modeling is shown in Fig. 1.

The quality of the generated PDB structure of CP47 and PsbO was evaluated by the Ramachandran plot that was obtained from PROCHECK (Fig. 2). The analysis of the Ramachandran plot of modeled structures for CP47 and PsbO showed the selected model found for each of them good (Table 1).

99.8% and 99.5% Φ/Ψ angles of modeled CP47 and PsbO residues are in the core and allowed regions, respectively (Table 1). In addition, protein models with a total quantity of G-factor value of more than -0.5 are considered the ideal model [38]. Overall G-factors for the developed models are -0.04 and -0.24, which indicates the designed model, has good quality (Table 1). ProSA score evaluated the interaction energy per residue of structure using a distance-based pair potential. Residues with negative ProSA energies confirm the reliability of the model. The energy profiles of generated models using the ProSA method are shown in Fig. 3. Evaluation of 3D template CP57 and PsbO, and modeled CP47 and PsbO with the ProSA server revealed a compatible Z-score of -6.04, -5.25, and, -5.68, -5.49, respectively (Table 1).

The energy profile of the structures, the geometric quality of the backbone conformation, and, residue interaction are all suitable. In addition, the average RMSD between the target and template structures was calculated and they were found to be 0.725 Å and 1.816 Å for target-template CP47 and PsbO, respectively (Fig. 4).

The low overall RMSD shows high structural conservatively, which makes good structures for homology modeling [39]. Therefore, the predicted models, which proved to be good in geometries and energy profiles, can be reasonable and reliable for further molecular docking analysis.

After modeling, native and mutant modeled proteins separately were embedded in a realistic representation of the thylakoid membrane, according to step 2.5 of materials and methods shown in Fig. 5.

Molecular dynamics simulation

We next performed 30ns MD simulations to investigate the structural consequences of mutating residue F363 to Arg, which is a naturally occurring mutation observed in CP47.

Table 1: Quality of structures checked by PROCHECK and ProSA for targets and templates.

PROCHECK[a]		Ramachandran Plot quality (%)			Goodness factor[b]			ProSA Z-score[c]
	Core	Allowed	General	Disallowed	Dihedral	Covalent	Overall	
Template_CP47	92	8	0.0	0.0	-0.02	0.16	0.05	-6.04
Modeled_ CP47	92.3	7.5	0.2	0.0	-0.11	0.03	-0.04	-5.68
Template_PsbO	87.7	11.3	0.5	0.5	-0.32	0.2	-0.12	-5.25
Modeled_ PsbO	85.6	13.9	0.5	0.0	-0.35	-0.11	-0.24	-5.49

[a] *Ramachandran plot qualities show the percentage (%) of residues belonging to the core, allowed, generally allowed, and disallowed region of the plot;* [b] *goodness factors show the quality of covalent and overall bond/angle distance; these scores should be above -0.5 for a reliable model.* [c] *ProSA z-score indicates overall model quality.*

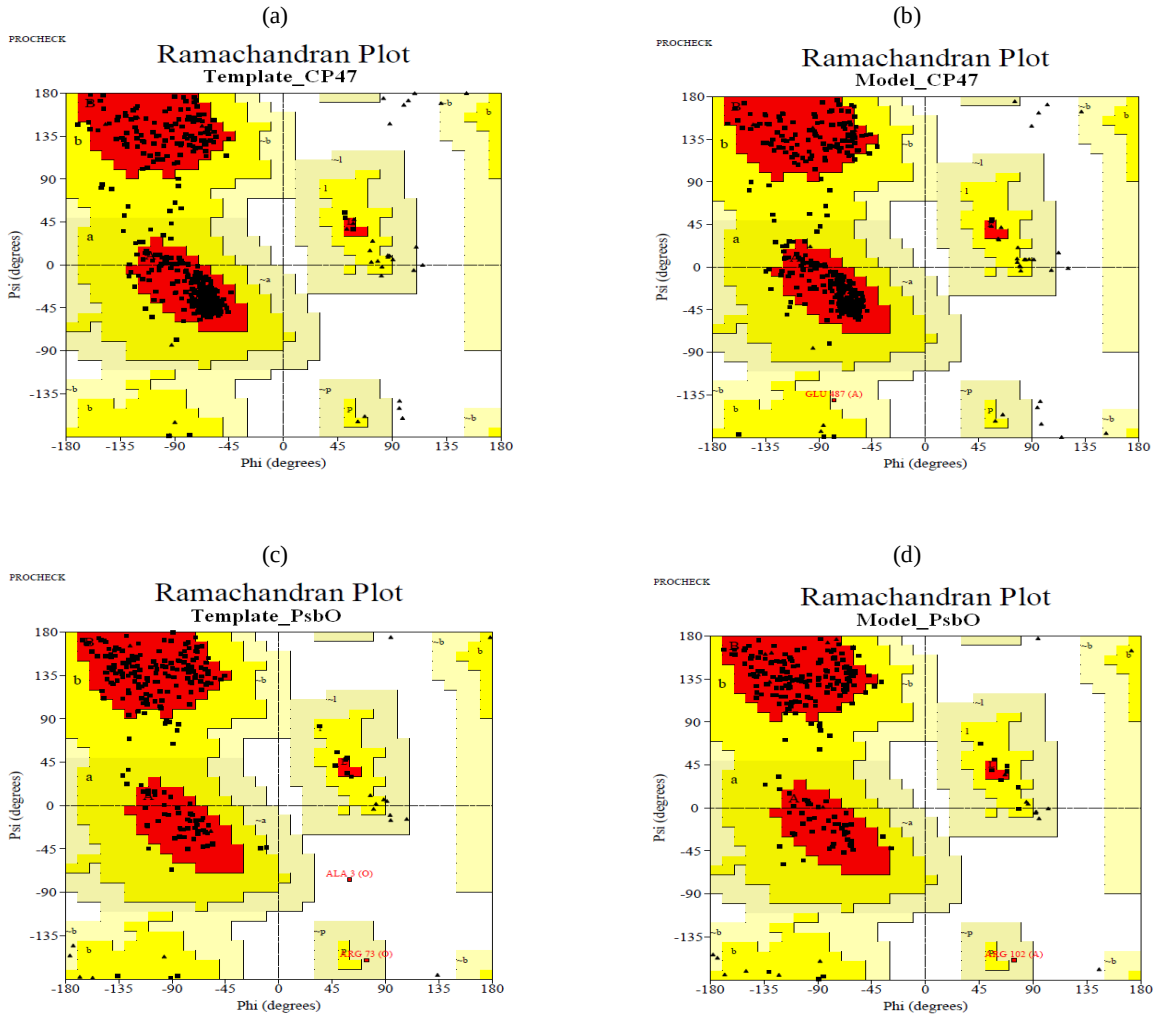

Fig. 2: Ramachandran plot of CP47 crystal structure (template) (a), modeled CP47 (target) (b), PsbO crystal structure (template) (c), and modeled PsbO (target). The favored and most favored region is yellow and red respectively, pale yellow is the generally allowed and disallowed regions are white.

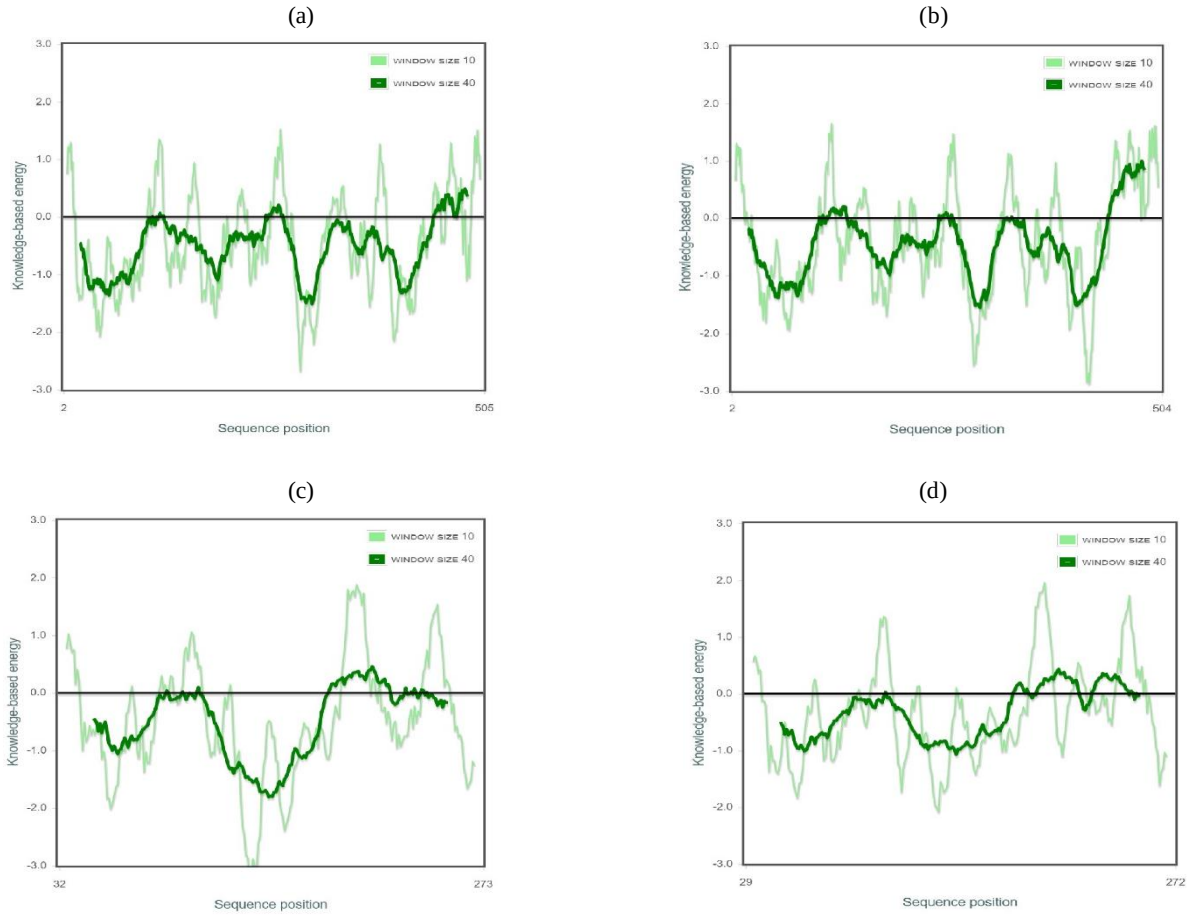

Fig. 3: ProSA energy profiles calculated for CP47 crystal structure (Template) (a), modeled CP47 (b), PsbO crystal structure (Template) (c), and modeled PsbO (d).

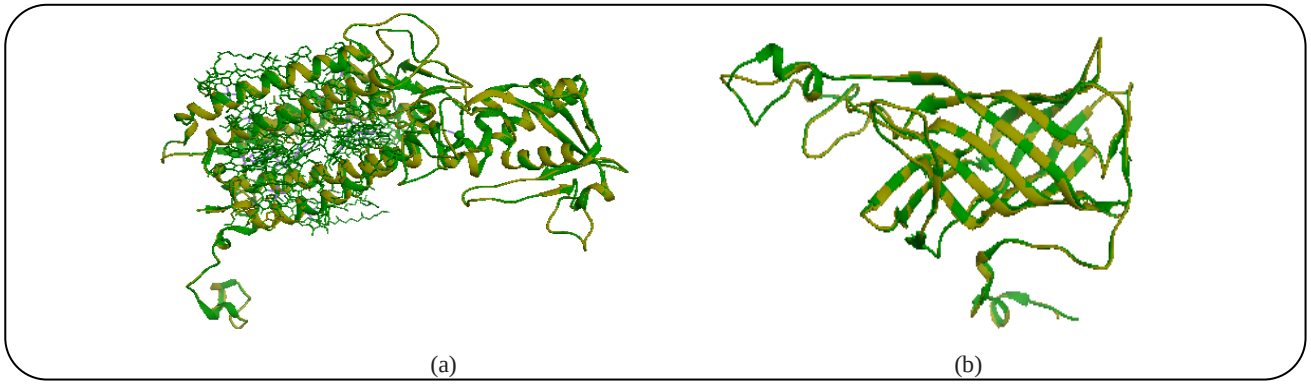

Fig. 4: Superimposition of modeled CP47 (target) on the crystal structure of CP47 (template) (a), modeled PsbO (target) on the crystal structure of PsbO (template) (b) using the 3d-SS tool. In these ribbon diagrams, yellow represents the target and the green represents the template.

Fig. 6a shows the solute potential energy of the 30ns trajectory. After the ~15ns, the solute potential energy ended its decreasing trend and fluctuated around -1905 MJ/mol and -1934 MJ/mol for native and mutant proteins, respectively. This is our reason for discarding the first 15ns of the trajectory and starting the equilibrium statistics thereafter.

The RMSD value for all protein backbone atoms that were calculated relative to the starting structure during MD

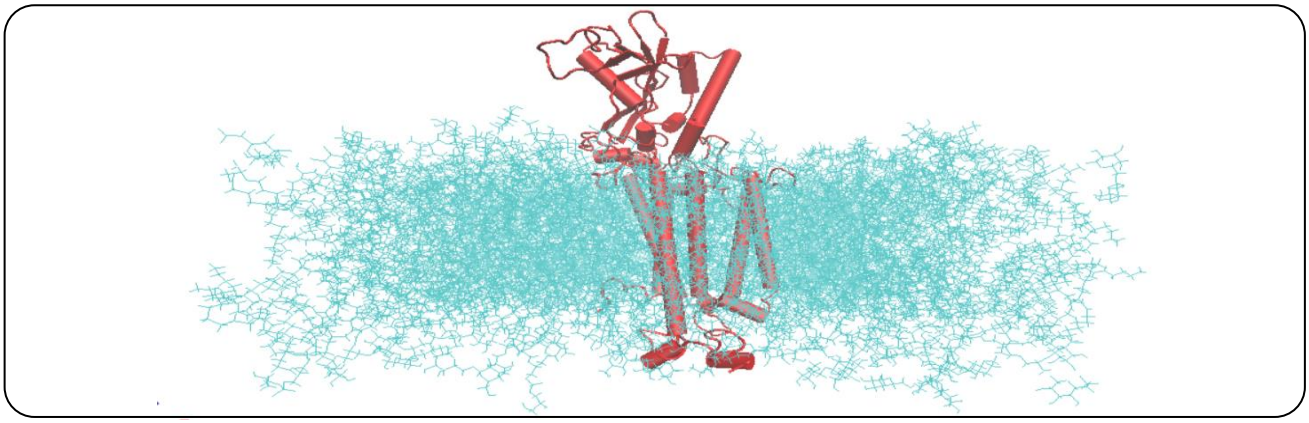

Fig.5: Snapshot of the simulation system embedded in a thylakoid membrane. Phospholipids groups are shown in cyan; protein is shown in turquoise red.

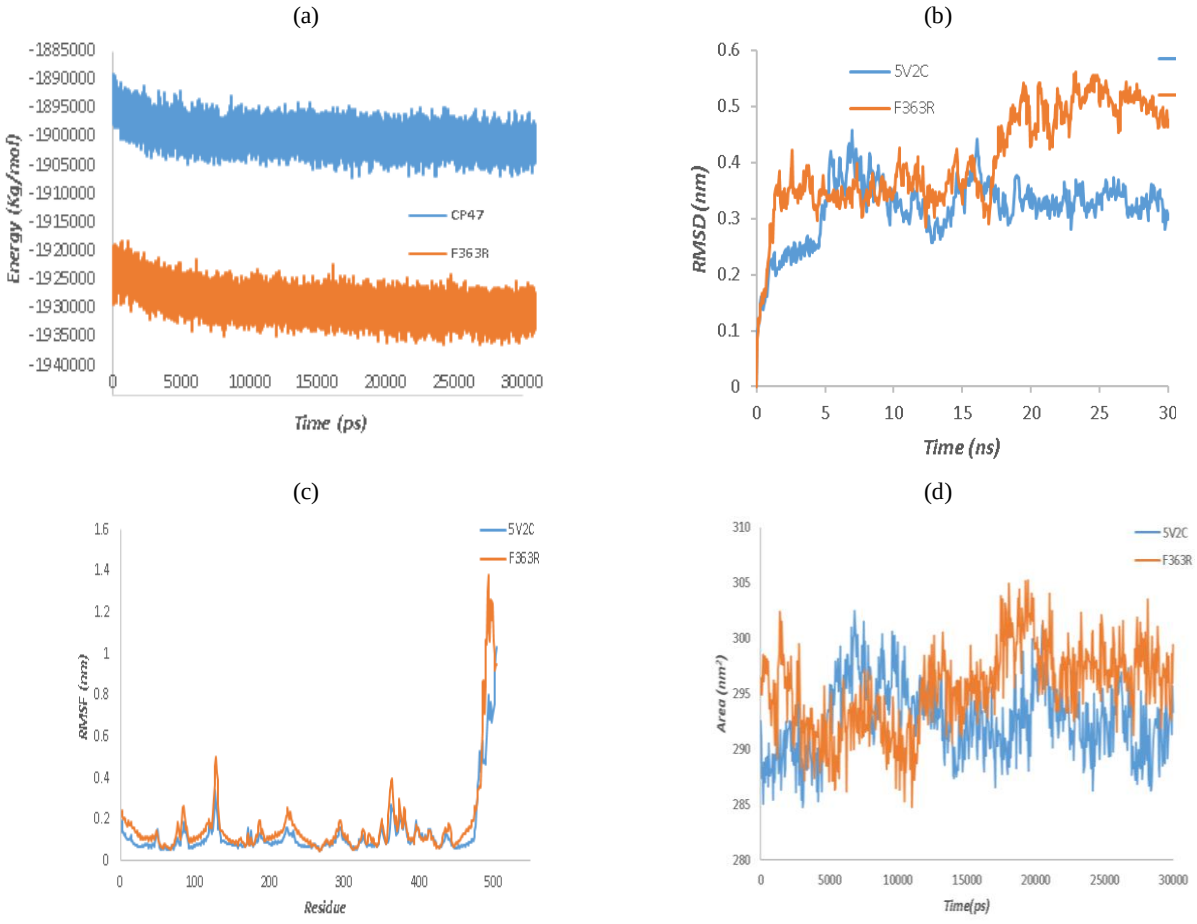

Fig. 6: The solute potential energy (a), Root-mean-square deviation (RMSDs) (b), Root-mean-squared fluctuation (RMSF) (c), and the Solvent Accessible Surface Area (SASA) during MD simulation of native and mutant structures.

simulation is shown in Fig. 6b. The native and F363R mutant CP47 show similar deviations from their starting structures during the first 17 ns of the trajectory. After 18.3 ns, the structure of the F363R mutant shows a significant increase in deviation to ~0.5 nm, which at this point represents a non-native structure.

The average value of fluctuations in backbone RMSD during the last 13 ns is depicted in Table 2. The larger

Table 2: The average value of RMSD, RMSF, and SASA during the last 15 ns for native and mutant structures.

Parameters	CP47	F363R
RMSD	0.033 ±0.019 (nm)	0.69 ± 0.041 (nm)
RMSF[1]	0.101 ± 0.048 (nm)	0.130 ± 0.062 (nm)
SASA	292.04 ± 2.80	297.79 ± 2.67

1) The average value is amino acids between 25- 480.

RMSD values for the mutant structure indicate that the overall topology of the mutant structure has changed and was under high structural instability during the simulations.

In addition, to compare the effect of mutation on the dynamics of the backbone atoms, the Root-mean-squared fluctuation (RMSF) values were computed for backbone atoms at each time point of the trajectories of the native and mutant structures. As shown in Fig. 6c, F363R mutation led to an increase in the RMSF values to those of the native structure. The average number of RMSF for amino acid sequences between 25 and 480 is 0.101± 0.048nm and 0.130 ± 0.062nm for native and mutant proteins, respectively (Table 2). These data and the inspection of fluctuation values (Fig. 6c) indicate that this mutation occurring in native protein led to an increase of structural flexibility as an effect of the entire structure and eventually the mutant structure became looser.

In particular, greater flexibility is observed for residues 357-368. These results suggest that the high flexibility of Arg than Phe can be attributed to increasing protein flexibility at the binding region and inducing conformational change. The larger RMSF values indicate increased random motions of this residue.

The Solvent-Accessible Surface Area (SASA) during MD simulation of native and mutant structures is shown in Fig. 6d. Solvent accessibility surface area accounts for the biomolecular surface area that is accessible to solvent molecules. After ~15ns from the initial structure, the mutant structure showed a greater value of SASA than the native. The average SASA value during the last 15ns simulation period was 292.04 ± 2.80 in native, whereas the F363R mutant structure showed 297.79 ± 2.67, as depicted in Table 2. The increased value of SASA in mutant structure denotes its conformational Change and relatively swollen nature as compared to the native structure.

Another important structural parameter, which can be obtained from molecular dynamics simulations and provides important information about the structural plasticity of proteins, is the time-dependent secondary structures which are calculated using DSSP classification of secondary structure elements. As shown in Fig. 7, an almost different pattern is seen in the plots of native and mutant structures. At the binding site region of native CP47 with PsbO, that is the residues between 360—440 appeared α-helix, β-sheets, Turn, Bend, and Coil conformation until the end of the simulation. During the simulation of the mutant structure, α-helix, and Coil are almost stable and do not change, but showed β-sheet and Bend with little and Turn with more traces than the native structure. In Table 3, regions of mutant protein structure whose secondary structure elements have changed compared to native are summarized. Overall, we observed that α-helix, β-sheets, and Bend decreased but Turn and Coil increased. In other regions, almost no change in secondary structure conformation was observed.

HADDOCK

After running HADDOCK, the best cluster according to cluster size and its Z-score was selected. Cluster size and Z-score indicate the number of conformers in a cluster and how many standard deviations from the average this cluster is located in terms of the score, respectively (the more negative the better). The statistics of the best cluster for the native and mutant complexes are indicated in Table 4.

To find out the affinity level of the biological partner, total binding energy which is the summation of Van der Waals and electrostatic energy are computed [40]. According to the results obtained from the ducking analysis, the total binding energy of the native complex was -457.95 kcal/mol. On the other hand, the mutant complex showed relatively less amount of total binding energy -429.70kcal/mol, when compared to the native.

In addition, the Buried Surface Area (BSA) of the mutant complex after docking is less than the native complex

Table 3: Changes in mutant secondary structure elements compared to native protein during simulation, according to DSSP algorithm.

	60-90[1]	120-135	317-337	355-440
β–Sheet	-	-	-	Decrease
Bend	-	Decrease	Decrease	-
Turn	-	Increase	Increase	Increase
α–Helix	Decrease	-	Decrease	-
3-Helix	-	-	-	-
Coil	increase	-	-	-

1) Region of residues

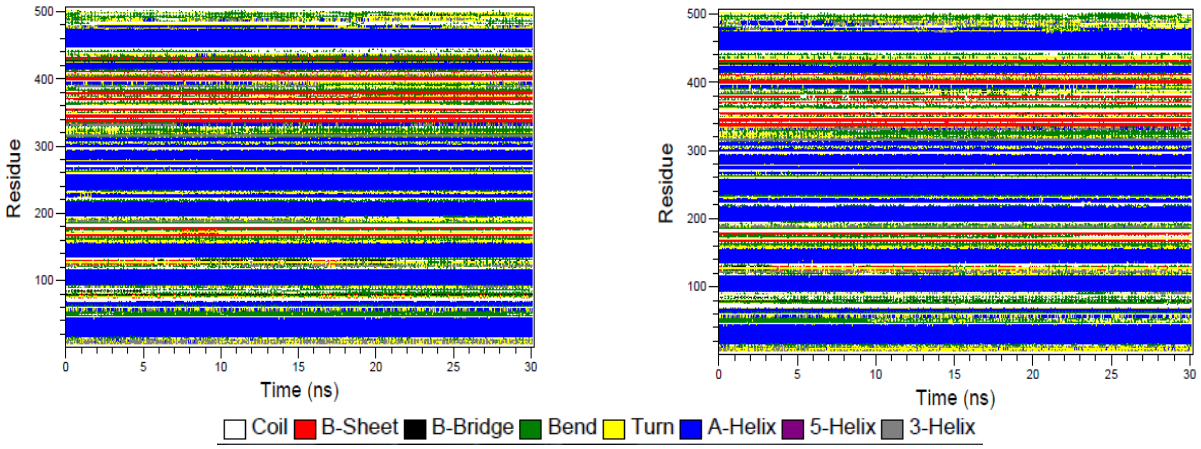

Fig. 7: Secondary structure as a function of the simulation time for native and mutant structures. Figure panels from up to down represents the secondary structure plot for native CP47 and F363R mutants, respectively.

(2379.0 ± 67.7Å^2 and 2793.6 ± 432.8Å^2, respectively). Accordingly, the difference between docking score, total binding energy, and BSA shows a better interaction between native CP47 and its partner (PsbO) in comparison with a mutant complex.

According to *Erijman et al.* [41], two biophysical features that showed the highest correlation with binding affinity are the geometric complementarity measured by the Van der Waals energy and side-chain conformational changes.

In this investigation, HADDOCK results show that the Van der Waals energy plays a major role in reducing the total binding energy of the mutant complex while the electrostatic energy is almost unchanged (Although it is only a little more) (Table 4). A comparison of the side chain in Phe and Arg shows that Arg has a highly flexible side chain with a positive charge and is extremely hydrophilic. However, Phe has a large aromatic side chain with moderate flexibility, quite hydrophobic (https://www.ncbi.nlm.nih.gov/Class/Structure/aa/aa_explorer.cgi). It seems that amino aside structural features namely side-chain conformational flexibility, at position 363 of the E loop binding domain has an important role in the interaction of CP47 and PsbO. Thus, due to the significant differences in structural features of Phe and Arg, we observed a decrease in BSA and Vander Waals energy at the interface of the mutant complex.

CONCLUSIONS

At present, due to the key role of cyanobacteria in ecosystems as well as the biotechnological development of these organisms, the study of mutations and the selection of strong samples for use in agriculture, medicine, etc. have received much attention. Most research is on the role of mutations that affect the photosynthetic apparatus.

Table 4: The statistics of the best clusters for native and mutant CP47 in interaction with PsbO.

Haddock criteria	native Complex	Mutant Complex
HADDOCK score	-61.9± 11.9	-57.9 ± 17.5
Cluster size	58	46
RMSD from the overall lowest-energy structure	22.1 ± 5.6	18.6 ± 2.4
Van der Waals energy	-98.2 ± 16.9	-72.0 ± 6.3
Electrostatic energy	-322.9 ± 56.7	-329.2 ± 50.7
Desolvation energy	-5.3 ± 7.9	-23.0 ± 2.2
Restrains violation energy	1103.1 ± 86.84	988.9 ± 145.21
Buried Surface Area (Å^2)	2793.6 ± 432.8	2379.0 ± 67.7
Z-Score	-1.2	-1.5
Total Interaction energy	-457.95	-429.70

This investigation aimed to carry out a computational study to show the effect of CP47-F363R mutation on the interaction between CP47 and PsbO as the essential components of PSII in the thylakoid membrane of cyanobacteria.

First, we developed a homology model of CP47 and PsbO based on the crystal structure of 5V2C-B, O. Then the variant form of CP47 (F363R) was constructed. After minimizing energy, MD simulation was carried out for the modeled structures of the native and mutant CP47. Based on RMSD and RMSF analysis, we confirmed that the variant showed higher deviation and fluctuation as compared to native. In order to investigate the effect of point mutations on the structure and function of the protein, molecular docking was performed between CP47 and PsbO for native and variant complexes. Docking analysis shows that point mutation of F363R can have some local impacts on the protein interaction and decrease the binding energy of CP47 with PsbO.

Generally, in agreement with each other, the results obtained from MD simulation and docking analysis, show that conformational change and the instability of mutant CP47 may play the main role in their being susceptible to reducing CP47-PsbO binding affinity. It seems that amino aside structural features, namely side-chain conformational flexibility, at position 363 of the E loop binding domain have an important role in the interaction of CP47 and PsbO.

Received : Sep. 12, 2022 ; Accepted : Dec. 19, 2022

REFERENCES

[1] Croce R., Amerongen H.V., Light-harvesting and Structural Organization of Photosystem II: From Individual Complexes To Thylakoid Membrane, *J. Photobiol. Photobiol. B: Biology.*, **104(1-2)**: 142-153 (2011).

[2] Blankenship R.E., "Molecular Mechanisms of Photosynthesis", Wiley Blackwell, United States, (2014).

[3] Knoppová J., Sobotka R., Tichy M, Yu J., Konik P., Halada P., Nixon PJ., Komenda.J., Discovery of a Chlorophyll Binding Protein Complex Involved in the Early Steps of Photosystem II Assembly in Synechocystis, *Plant Cell*., **26(3)**:1200-1212(2014).

[4] Komenda J., Roman S., Chlorophyll-Binding Subunits of Photosystem I and II: Biosynthesis, Chlorophyll Incorporation and Assembly, *Advances in Botanical Research.*, **91**:195-223(2019).

[5] Bučinská L., Kiss É., Koník P., Knoppová .J, Komenda J., Sobotka R., The Ribosome-Bound Protein Pam68 Promotes Insertion of Chlorophyll into the CP47 Subunit of Photosystem II, *Plant Physiol.*, **176(4)**:2931-2942(2018).

[6] Lubitz W., Chrysina M., Cox N., Water Oxidation in Photosystem II, *Photosynth Res*, **142(1)**:105-125(2019).

[7] Bondarava N., Beyer P., Krieger-Liszkay A., Function of the 23 kDa Extrinsic Protein of Photosystem II as a Manganese Binding Protein and its Role in Photoactivation, *Biochim Biophys Acta.*, **1708(1)**:63-70 (2005).

[8] Popelkova H., Charles F.Y, PsbO, the Manganese-Stabilizing Protein: Analysis of the Structure–Function Relations that Provide Insights into its Role in Photosystem II, *Journal of Photochemistry and Photobiology B: Biology.*, **104(1-2)**:179-190(2011).

[9] Nagao R., Suzuki T., Okumura A., Niikura A., Iwai M., Naoshi D., Tomo T., Ren Shen J., Ikeuchi M., Enami I., Topological Analysis of the Extrinsic PsbO, PsbP and PsbQ Proteins in a Green Algal PSII Complex by Cross-Linking with a Water-Soluble Carbodiimide, *Plant and Cell Physiology*, **51(5)**:718-727(2010).

[10] Roose L.J., Frankel K.L., Mummadisetti P.M., Bricker T.M., The Extrinsic Proteins of Photosystem II: Update, *Planta.*, **243(4)**:889-908(2016).

[11] Cindy P.E., Robert B., Jituo W.u., John W., Terry M.B., Site-Directed Mutagenesis of the CP 47 Protein of Photosystem II: Alteration of Conserved Charged Residues in the Domain ^{364}E–^{444}R, *Biochemistry***.,** **35(13)**: 4046–4053 (1996)**.**

[12] Becker K., Cormann U.K., Nowaczyk M.M., Assembly of the Water-Oxidizing Complex in Photosystem II, *J. Photochem. Photobiol. B.*, **104(1-2)**:204-11(2011).

[13] Debus R.J., Amino Acid Residues that Modulate the Properties of Tyrosine Y(Z) and the Manganese Cluster in the Water Oxidizing Complex of Photosystem II, *Biochim Biophys Acta.*, **1503(1-2)**:164-86(2001).

[14] Shannon M., Clarke and Julian J., Eaton R., Mutation of Phe-363 in the Photosystem II Protein CP47 Impairs Photoautotrophic Growth, Alters the Chloride Requirement, and Prevents Photosynthesis in the Absence of either PSII-O or PSII-V in *Synechocystis* sp. PCC 6803, *Biochemistry.*, **38(9)**: 2707-2715 (1999).

[15] Stewart A. A., Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins, *Chem Rev.*, **106(5)**:1589–1615 (2006).

[16] Scott A.H., Ron O.D., Molecular Dynamics Simulation for All, *Neuron.*, **99(6)**: 1129-1143 (2018).

[17] Cyril D., Rolf B., Alexandre M.J.J., HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information, *J. Am. Chem. Soc.*, **125(7)**: 1731-1737 (2003).

[18] MacDougall I.J., Lewis P.J., Griffith R., Homology Modelling of RNA Polymerase and Associated Transcription Factors from Bacillus Subtilis, *J. Mol. Graph. Model.*, **23(4)**:297-303 (2005).

[19] Sefidbakht Y., Ranaei S.O., Taheri F., Homology Modeling And Molecular Dynamics Study on Schwanniomyces Occidentalis Alpha-Amylase, *J. Biomol. Struct. Dyn.*, **35(3)**:574-584 (2017).

[20] Vuister G.W , Fogh R.H., Hendrickx P.M.S., Doreleijers J.F., Gutmanas A., An Overview Of Tools For The Validation Of Protein NMR Structures, *J. Biomol. NMR.*, **58(4)**:259-85(2014).

[21] Wiederstein M., Sippl M.J., ProSA-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins, *Nucleic Acids Res.*, **35(2)**:407-410 (2007).

[22] Sumathi K., Ananthalakshmi P., Roshan M.N.A., Sekar K., 3dSS: 3D Structural Superposition, *Nucleic Acids Research.*, **34**:128–132 (2006).

[23] Sakurai I., Shen J.R., Leng J., Ohashi S., Kobayashi M., Wada H., Lipids in Oxygen-Evolving Photosystem II Complexes of Cyanobacteria and Higher Plants, *J. Biochem.*, **140(2)**:201-209 (2006).

[24] Emilia L.W.u., Olof E., Sunhwan J.o., Danielle S., Min S.Y., Jeffery B.K., Göran W., Wonpil I., Molecular Dynamics and NMR Spectroscopy Studies of E. Coli Lipopolysaccharide Structure and Dynamics, *Biophys. J.*, **105(6)**: 1444–1455 (2013).

[25] Emilia L.W.u., Patrick J.F., Min S.Y., Göran W., Jeffery B.K., Karen G.F., Wonpil I., Ecoli Outer Membrane and Interactions with OmpLA, *Biophys J.*, **106(11)**: 2493–2502 (2014).

[26] Jo S., Wu E.L., Stuhlsatz D., Klauda J.B., MacKerell A.D Jr., Widmalm G.I.m. W., Lipopolysaccharide Membrane Building and Simulation, *Methods Mol. Biol.*, **1273**:391-406 (2015).

[27] Sunhwan J.o., Taehoon K., and Wonpil I., Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations, *PLoS One.*, **2(9)**:1-9 (2007).

[28] Sunhwan J.o., Joseph B.L., Jeffery B.K., Wonpil I., CHARMM-GUI Membrane Builder for Mixed Bilayers and its Application to Yeast Membranes, *Biophys J.*, **97(1)**: 50–58 (2009).

[29] Jo S., Kim T., Iyer V.G., Im W., CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM, *J. Comput. Chem.*, **29(11)**:1859-65(2008).

[30] Van Eerden F.J., de Jong D.H., de Vries A.H., Wassenaar T.A., Marrink S.J., Characterization of Thylakoid Lipid Membranes From Cyanobacteria and Higher Plants by Molecular Dynamics Simulations, *Biochim. Biophys. Acta, Biomembr.*, **1848(6)**:1319-1330 (2015).

[31] Brooks B.R., Brooks C.L., Mackerell A.D Jr., Nilsson L., Petrella R J., Roux B., Won Y., Archontis G., Bartels C., Boresch S., Caflisch A., Caves L., Cui Q., Dinner AR., Feig M., Fischer S., Gao J., Hodoscek M., Kuczera K., CHARMM: the Biomolecular Simulation Program, *Journal of Computational Chemistry*, **30(10)**:1545–1614 (2009).

[32] Jeffery B.K., Richard M.V., Alfredo F., Joseph W.O., Update of the CHARMM all-Atom Additive Force Field for Lipids: Validation on Six Lipid Types, *J. Phys. Chem. B.*, **114(23)**:7830–7843(2010).

[33] Rezazadeh Mofradnia S., Ashouri R., Abtahi N., Yazdian F., Rashedi H., Sheikhpour M., Ashrafi F., Production and Solubility of Ectoine: Biochemical and Molecular Dynamics Simulation Studies, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **39(6)**:259-269(2020).

[34] Ibeji C.U., Ujam O.T., Chukwuma Chime C., Akpomie K G., Anarado C.J.O., Odewole O.A., Dehydroacetic Acid-Phenylhydrazone as a Potential Inhibitor for Wild-Type HIV-1 Protease: Structural, DFT, Molecular Dynamics, 3D QSAR and ADMET Characteristics, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **40(1)**: 215-230(2021).

[35] Jumin L.X., Jason M.S., Min S.Y., Peter K.E., Justin A.L., Shuai W., Joshua B., Jong C.J., CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field, *J. Chem. Theory Comput.*, **12(1)**:405–413 (2016).

[36] Elber R., Ruymgaart A.P., Hess B., SHAKE Parallelization, *Eur. Phys. J. Spec. Top.*, **200(1)**:211-223(2011).

[37] Kaya H., Hardy D.J., Skeel R.D., Multilevel Summation For Periodic Electrostatics Using B-splines, *J.Chem. Phys.*, **154(14)**:144105(2021).

[38] Azizian H., Bahrami H., Pasalar P., Amanlou M., Molecular Modeling of Helicobacter Pylori Arginase and the Inhibitor Coordination Interactions, *J. Mol. Graph. Model.*, **28(7)**: 626-35 (2010).

[39] Umamaheswari A., Pradhan D., Hemanthkumar M., Virtual Screening for Potential Inhibitors of Homology Modeled Leptospira Interrogans Murd Ligase, *Journal of Chemical Biology.*, **3(4)**:175–187 (2010).

[40] Kumar A., Purohit R., Use of Long Term Molecular Dynamics Simulation in Predicting Cancer Associated SNPs, *PLoS Computational Biology.*, **10(4)**: 1-14 (2014).

[41] Erijman A., Rosenthal E., Shifman J.M., How Structure Defines Affinity in Protein-Protein Interactions, *PLoS One.*, **9(10)**: 1-10 (2014).