

An *in-silico* study of Polymerase Epsilon catalytic subunit proteins in *Arabidopsis thaliana*

Mahamat Baraka Adoum¹, Kanita Šabanović², Mohamed Ragab Abdel Gawwad^{3*}

^{1,2,3} Genetics and Bioengineering, International University of Sarajevo

*Corresponding author: mragab@ius.edu.ba

© The Author
2020.
Published by
ARDA.

Abstract

Arabidopsis thaliana genome encodes two POLE2 homologs known as polymerase epsilon catalytic subunit A (POLE2A) and polymerase epsilon catalytic subunit B (POLE2B). They play a very important role in DNA repair mechanisms. In this study, bioinformatics tools were used to understand DNA repair mechanisms in *A. thaliana* in which POLE2A and POLE2B proteins are involved. Through interactome analysis of POLE2A and POLE2B homolog proteins in *A. thaliana*, their additional roles in DNA repair were explored. The most important proteins that are participating in DNA repairs, like MSH2, MSH5, PCNA1, PCNA2, PRL, and CDC45 were identified as interactors of both POLE2A and POLE2B. The three-dimensional structure of POLE2 proteins was identified to decipher the complexity of NER, GG-NER, MMR, TFIIH, and TC-NER repair mechanisms through the identification of docking sites. The interaction complex of POLE2A and POLE2B with six proteins was confirmed and found to have a significant role in DNA repair processes and UV-B tolerance. The interactome analysis of POLE2A and POLE2B performed here once again confirms the complexity of the DNA repair mechanism in plants.

Keywords: POLE2; 3-D structure; Interactome; Docking sites; *Arabidopsis thaliana*

1. Introduction

1.1. *Arabidopsis thaliana* as a plan model organism

Arabidopsis thaliana emerged a model organism for research in plant biology. A consensus was reached about the need to focus on a single organism to integrate the classical disciplines of plant science with the expanding fields of genetics and molecular biology. Ten years after publication of its genome sequence, *A. thaliana* remains the standard reference plant for all of fields of biology. We reflect here on the major advances and shared resources that led to the extraordinary growth of the *A. thaliana* research community. We also underscore the importance of continuing to expand and refine our detailed knowledge of *A. thaliana* while seeking to appreciate the remarkable diversity that characterizes the plant kingdom [1].

Maintenance of the stability and integrity of plant genome when affected with UV radiation is crucial for proper functioning of the plant and, indirectly, the rest of the ecosystem which is connected to it. *A. thaliana* specifically, has been proven to release volatile compounds when irradiated by UVC radiation (mostly methyl salicylate and methyl jasmonate), that cause genomic instability in neighboring plants, which even belongs to other species, like *Nicotiana tabacum*. Thus, to analyze and understand the specificities of plant DNA repair mechanisms, starting from well-researched model higher plant like *A. thaliana* is an inevitable step of plant science research, with far-reaching consequences [2].

1.2. DNA mutations

A deoxyribonucleotide polymer, is known as the primary genetic material of most cells. Any injuries to DNA which introduce alterations in chemical structure of DNA may result in a mutation or block of DNA replication, if left unrepaired [3].

Some mutations are inherited parentally or occur during cell division, whereas some mutations are caused by environmental factors, including chemicals, viruses, and ultraviolet (UV) radiation [4]. Ultraviolet radiation introduces noxious effects in all living organisms ranging from prokaryotic bacteria to lower and higher eukaryotic plants, animals, and humans. UV radiation is generally divided into three classes: UV-C, UV-B and UV-A. DNA is highly sensitive to UV-B (280–315 nm) radiation for which it has been shown to be harmful to living organisms, causing photo-transformation, which results in direct and indirect DNA damage [5], [6].

1.3. UV radiation and irradiation repairing

In plants, exposing high doses of UV-B irradiation causes morphological and physiological damage that result in decreasing plant development rate, nitrogen productivity, interrupting processes including photosynthesis, secondary metabolism, responses to stress and photomorphogenesis [7], [8]. To prevent irradiation from UV lights, plants have been developed two main mechanisms, light, and dark repair. Light repair mechanism is based on visible light, and enzyme known as photolyases. Photolyase is involved in repairing UV induced damages in plants, which binds complementary DNA strands and break certain types of pyrimidine dimers [7]. Pyrimidine dimers are seen when a pair of thymine or cytosine bases are on the same strand of DNA and are covalently linked, these dimers are result in DNA lesion. This DNA lesion is recovered by photoreactivation treatment with light of longer wavelength. Photoreactivation is the direct correction of the pyrimidine-pyrimidine ring back into the normal sequence by using the enzyme photolyase with the energy of blue visible light. This repair mechanism does not appear in human, as proved before [9].

1.4. Repair Mechanism

Nucleotide excision repair (NER) is a mechanism to recognize and repair bulky DNA damage caused by compounds, environmental carcinogens, and exposure to UV light. The repair of damaged DNA involves at least 30 polypeptides within two different sub-pathways of NER, known as transcription-coupled repair (TCR-NER) and global genome repair (GGR-NER) [10]. Base excision repair (BER) is the predominant DNA damage repair pathway for the processing of small base lesions, derived from oxidation and alkylation damages [2], [11]. Another mechanism involved in DNA repair is DNA mismatch repair (MMR) and defects in this mechanism have been found in skin cancer [12].

1.5. Polymerase epsilon

During DNA replication, polymerase epsilon is a polymerase responsible for the forming leading strand. DNA polymerase epsilon, in yeast, has four subunits. In *A. thaliana*, DNA polymerase epsilon consists of a following units: for the regulatory subunits: DPB2, DPB4, and DBP3; and for the catalytic subunit: POL2a and POL2b [13].

In *A. thaliana*, POL2A, together with DPB2, is essential for early stages of embryogenesis. Also, it is main catalytic subunit that is active in meristems and during embryogenesis. The POL2B catalytic subunit is mainly expressed during stress conditions. If knockout is done in POL2B, no morphological phenotypes are exhibited, however if null mutation occurs in DPB2 or POL2A, it is lethal. Mutants which have lost function of POL2B, have cells which are larger, longer cell cycles and delayed embryo development [13,14].

In the *Arabidopsis* genome, there is only one copy of the TIL gene, but some studied monocots have shown that in several, two homologs, TIL1 and TIL2 can be found. In *A. thaliana*, polymerase epsilon is encoded by the TILED (TIL1) locus [15].

1.6. Aim of the study

The aim of this study is to decipher the DNA repair mechanism in *A. thaliana*, focusing on POLE2A and POLE2B proteins, through a detailed *in silico* analysis of the functional interactome of these proteins.

2. Materials and methods

2.1. Retrieving POLE2a and POLE2b sequences

The sequences for AtPOLE2A and AtPOLE2B proteins were obtained from the National Center for Biotechnology Information (NCBI) database (The Arabidopsis Information Resource) database. The accession number for POLE2a and POLE2b subunits are shown in table 1.

Table 1. The accession number of the proteins

POLE2 subunits	TAIR ID	NCBI References
POLE2a (TIL1)	At1g08260	NP_172303.5
POLE2b (TIL2)	At2g27120	NP_180280.2

2.2. Phylogenetic tree construction

To infer the evolutionary relationship between the POLE2A and POLE2B proteins, a phylogenetic tree was constructed using Phylogeny.fr web service for phylogenetic analysis of molecular sequences. The “One Click” option on phylogeny.fr was used, where from the Newick format, the tree was obtained [16].

2.3. Prediction of 3-D structure models and Ramachandran plot confirmation

The expasy-swiss model is structural bioinformatics software for homology modeling of three-dimensional (3D) structure of proteins. It was used to generate the PDB files online, whereas Pymol was used to visualize 3D structure of AtPOLE2A and AtPOLE2B. The 3D structure of each subunit was generated and confirmed using Ramachandran plots using the RAMPAGE program. Domain analysis was made using SMART Server [17].

2.4. Prediction of the interactome

STRING is a tool where the protein-protein interactions can be easily seen and analyzed to show the interactions between the POLE2A and POLE2B subunits with other related proteins, as well as to determine how strong these interactions are [18].

2.5. Domain identification

The identification of domains in the six POLE2A and POLE2B proteins was performed using the online Simple Modular Architecture Research Tool (SMART) tool, on the website of the European Molecular Biology Laboratory (EMBL). The tool can detect more than 500 domain families from chromatin-associated, extracellular and signaling proteins. These domains are comprehensively interpreted with regards to functional class, important residues, and tertiary structures [19].

Domain analysis was made using SMART Server. The tool can detect more than 500 domain families from chromatin-associated, extracellular and signaling proteins. These domains are comprehensively interpreted with regards to functional class, important residues, and tertiary structures [20].

2.6. Subcellular localization of POLE2A and POLE2B proteins

To check the subcellular localization of the proteins, online tool Plant Subcellular Localization Integrative predictor (PSI) was used. It uses the group voting strategy and machine learning to combine the results of 11 independent subcellular localization tools: cello, mPloc, Predotar, mitoProt, MultiLoc, TargetP, WolfPSORT, subcell Predict, iPsort, Yloc and PTSI [21].

For the subnuclear localization prediction, we used the Subnuclear Compartments Prediction System (Version 2.0), an SVM-based system developed at Laboratory of Computational Functional Genomics, Department of Bioengineering, PML body, nuclear lamina, nuclear speckles, chromatin, nucleoplasm (nuclear diffuse) and nucleolus [22].

Alternatively, the same proteins were also submitted to the SUB cellular localization database for Arabidopsis proteins (SUBA), which comprises of large proteomic and GFP datasets of localizations for this model plant, but also provides access to several bioinformatics prediction tools, giving an integrated approach to localization in cellular compartments of *A. thaliana* [23].

2.7. Docking site prediction

Docking is a computational modeling of protein-protein complexes in quaternary structure by interacting at least two or more biological macromolecules. ClusPro 2.0 server is used to predict docking sites for AtPOLE2A and AtPOLE2B proteins. Before using Clus Pro algorithm, PDB files were needed and generated for each ligand proteins by ExPasy-Swiss tool and then run into algorithm to get results done individually and visualized by PyMol [23].

3. Results

3.1. Phylogenic tree construction

Proteins of POLE2a and POLE2b in species were obtained and used to construct the phylogenetic tree as cladogram shown in Fig. 2.

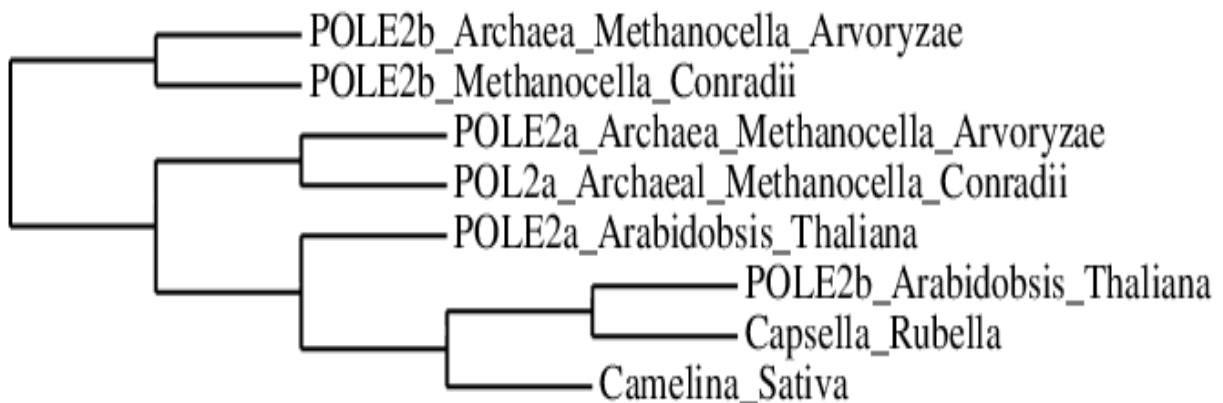


Figure 1. Phylogenetic tree of polymerase epsilon in different species; constructed by Phylogeny.fr

3.2. Predicted 3-D structure models and Ramachandran plot validation

For AtPOLE2a and AtPOLE2b, the 3-D structure predictions were visualized by using Pymol and the confidence of the prediction was confirmed by a Ramachandran plot.

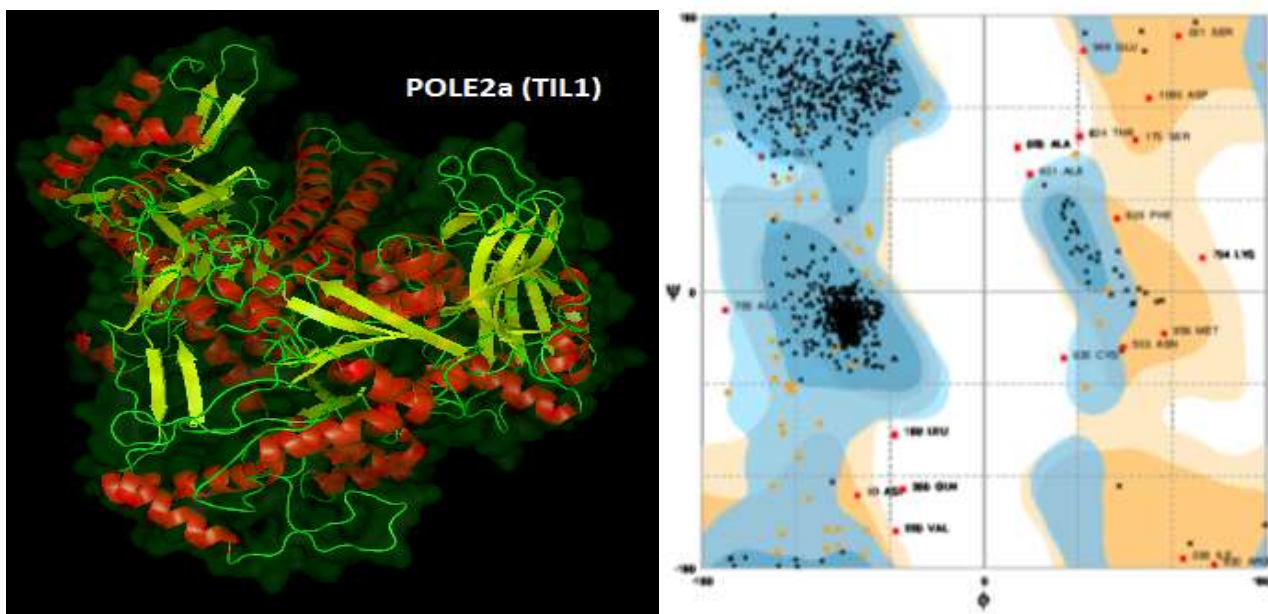


Figure 2. POLE2a (TIL1) in 3-dimensional structure as viewed by Pymol and confirmed via Ramachandran plot. Secondary structure: helix is shown in red; sheets are shown in yellow and loop are shown in green

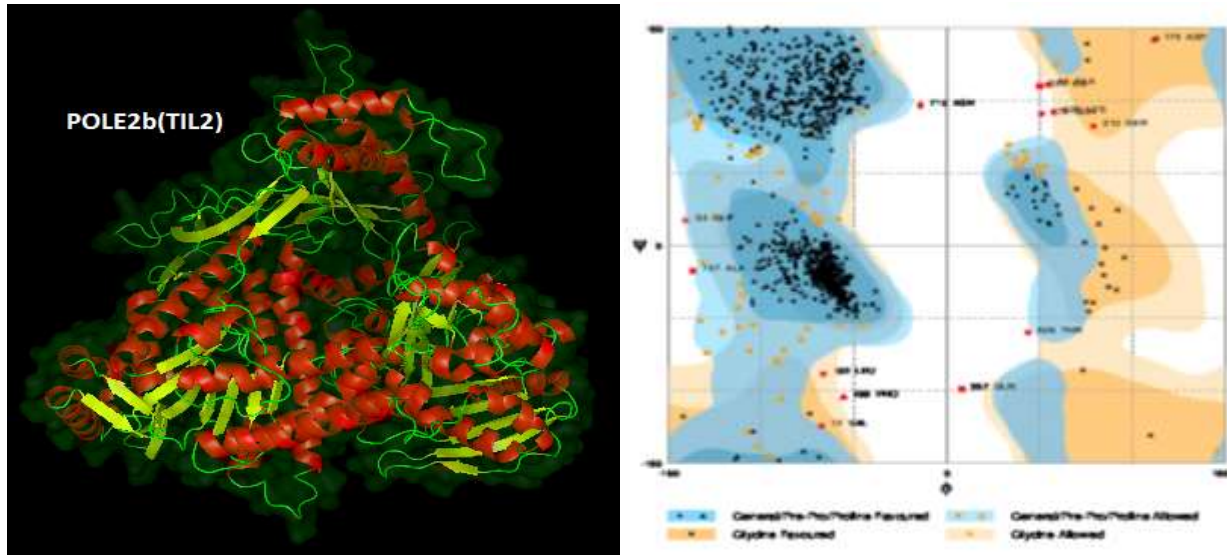


Figure 3. POLE2b (TIL2) in 3-dimensional structure as viewed by Pymol and confirmed via Ramachandran plot. Secondary structure: helix is shown in red; sheets are shown in yellow and loops are shown in green

3.3. Interactome of POLE2A (TIL1) and POLE2B (TIL2)

Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) has revealed interactions for both POLE2a and POLE2b. Only those which have high confidence interactions were chosen, whereas lower confidence proteins were discarded for this work and shown in table 2, indicates proteins subcellular localization and interactome annotations. Fig. 4 shows that both POLE2A and POLE2B are interacting with MSH2, MSH5, PCNA1, PCNA2, PRL, CD45.

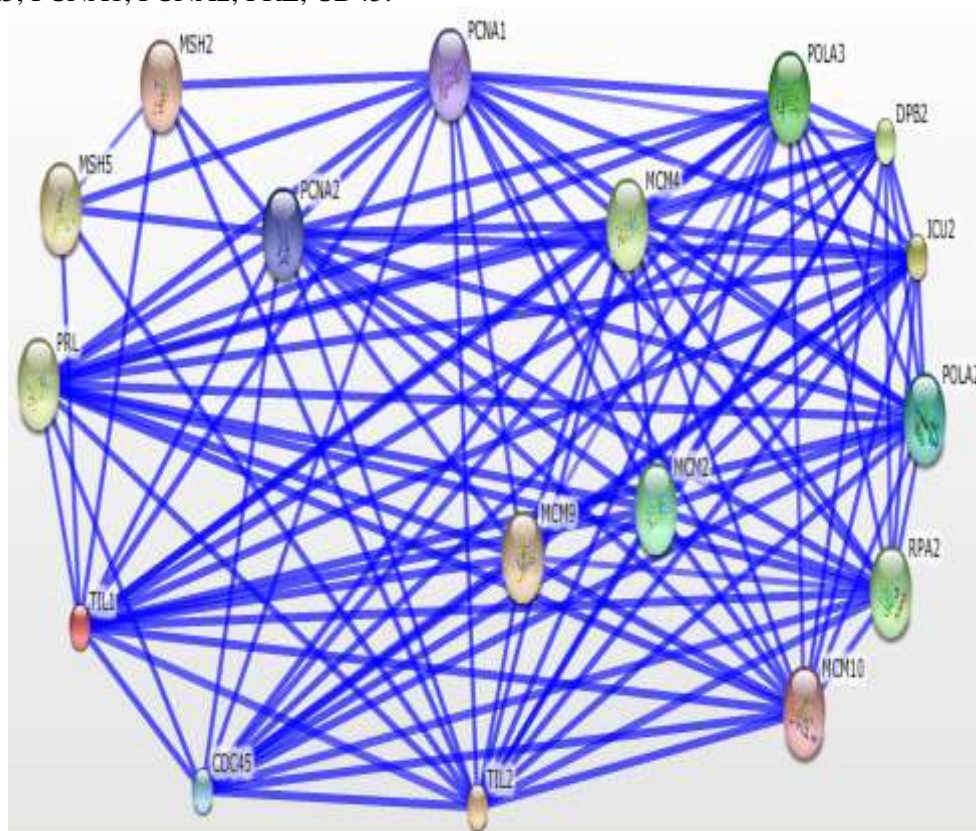


Figure 4. Interactome of TIL1 and TIL2. Interactome visualization obtained using STRING and adjusted only for high interactions

3.3.1. Prediction of the interactome of TIL1 (POLE2a)

STRING tool was used for obtaining the interactome of proteins of POLE2A. Fifteen interaction of highest confidence are presented in table 2.

Table 2. Interactome of POLE2a/b subunits, information extracted by STRING

Proteins	Interacting proteins	Functions	Score
TIL1 (POLE2a) and TIL2 (POLE2b)	MSH2	Component of the post-replicative DNA mismatch repair system	0.991
	CD45	Cell division cycle function	0.994
	MSH5	Involved in meiotic recombination in association with MSH4	0.988
	PRL	Required for megagametophyte and embryo development	0.988
	PCNA1	Proliferating cellular nuclear antigen 1	0.993
	PCNA2	Proliferating cell nuclear antigen 2	0.994
	ICU2	Polymerase alpha in a complex with DNA primase	0.999
	DPB2	DNA polymerase epsilon subunit B2	0.999
	POLA3	DNA primase small subunit	0.999
	MCM2	Minichromosome Maintenance protein 2	0.998
	POLA2	DNA polymerase alpha 2	0.996
	MCM10	Minichromosome maintenance 10	0.992
	MCM9	Minichromosome maintenance 9	0.988
	MCM4	Minichromosome maintenance 4	0.987
	RPA2	Component of the replication protein A complex (RP-A)	0.986

3.4. Domain identification

Tables 3 and 4 represent identified domains of POLE2a and POLE2b, respectively.

Table 3. Confidently predicted domains (POLE2a)

Protein	Domain	Start	End	E-Value
POLE2a	POLBc	239	843	8.5e-72
	DUF1744	1488	1872	9.74e-193
MSH2	Pfam: MutS_I	22	129	5.5e-18
	Pfam: MutS_II	142	284	4.4e-17
	MUTSd	314	642	1.05e-108
	MUTSac	659	855	4.32e-119
MSH5	MUTSd	211	547	1.2e-44
	MUTSac	562	757	1.74e-50
CDC45	Pfam: CDC45	27	595	1.3e-156
PLR	MCM	139	638	0
	AAA	368	521	0.00378
PCNA1	Pfam: PCNA_N	1	125	1.9e-61
	Pfam: Rad9	12	242	4.1e-9
	Pfam: PCNA_C	127	254	8.4e-63
PCNA2	Pfam: PCNA_N	1	125	2.9e-61
	Pfam: Rad9	12	241	3.2e-9
	Pfam: PCNA_C	127	254	1.1e-62

Table 4. Confidently predicted domains (POLE2b)

Protein	Domain	Start	End	E-Value
POLE2b	POLBc	231	835	3.05e-73
	DUF1744	1471	1848	5.15e-183
MSH2	Pfam: MutS_I	22	129	5.5e-18
	Pfam: MutS_II	142	284	4.4e-17
	MUTSd	314	642	1.05e-108
	MUTSac	659	855	4.32e-119
MSH5	MUTSd	211	547	1.2e-44
	MUTSac	562	757	1.74e-50
CDC45	Pfam: CDC45	27	595	1.3e-156
PLR	MCM	139	638	0
	AAA	368	521	0.00378
PCNA1	Pfam: PCNA_N	1	125	1.9e-61
	Pfam: Rad9	12	242	4.1e-9
	Pfam: PCNA_C	127	254	8.4e-63
PCNA2	Pfam: PCNA_N	1	125	2.9e-61
	Pfam: Rad9	12	241	3.2e-9
	Pfam:PCNA_C	127	254	1.1e-62

3.5. Subcellular localization of POLE2A and POLE2B proteins

The subcellular localization prediction was done by using PSI predictor of AtPOLE2A and AtPOLE2B. The score indicates the confidence of the localization of the protein in a cellular compartment, end score between 0 and 1 is shown as level of presence of the protein in a cellular compartment [24]. Table 6 shows subcellular localization of POLE2a and POLE2b proteins.

Table 6. Subcellular localization of POLE2a and POLE2b

Proteins	Predicted subcellular localization	Score	p-Value
POLE2a	Nucleus	0.759	0
POLE2b	Nucleus	0.759	0

3.6. Docking site prediction

A total of six interacting proteins were selected for docking site investigation and prediction. The interacting proteins were chosen for CD45, PCNA1, PCNA2, PRL, MSH5 and MSH2 and have been selected for both AtPOLE2A and AtPOLE2B for further docking site analysis. Additionally, AtPOLE2A and AtPOLE2B share high level of sequence similarity domains, so both have been docked to see their relationship. Figure 5 to 16 show surface and cartoon structural view of the proteins that interacted with AtPOLE2A and AtPOLE2B, respectively, as tables 4 and 5 illustrates the docking residues for each of the investigated proteins. The sequences are aligned from N-terminus to C-terminus.

Six interacted proteins have been selected for docking site prediction to see orientation to a chosen interactome molecule: CD45, PCNA1, PCNA2, PRL, MSH5 and MSH2 were chosen for POLE2a.

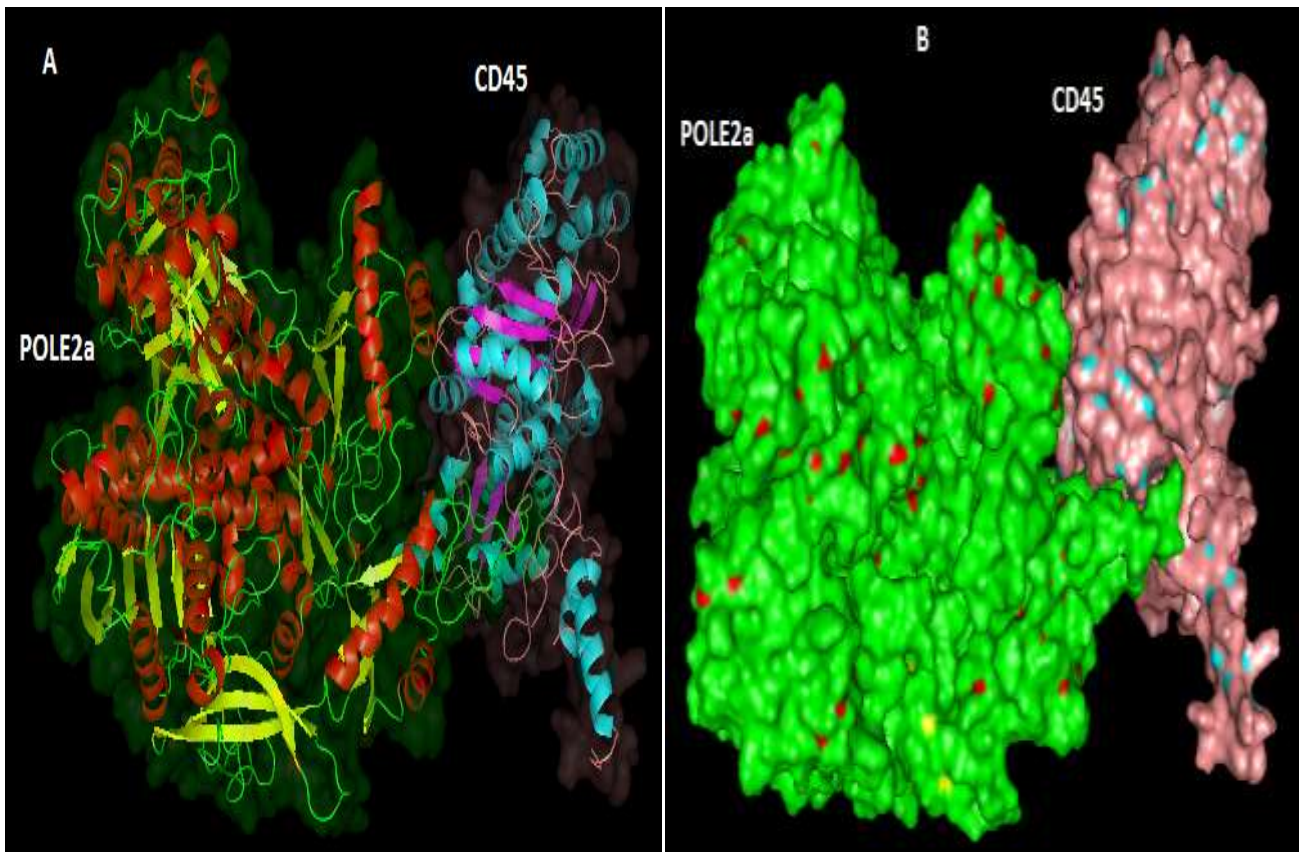


Figure 5. POLE2a-CD45 docking site. (A) a cartoon structure of POLE2a –CD45 interaction, (B) surface structure of POLE2a-CD45 interaction.

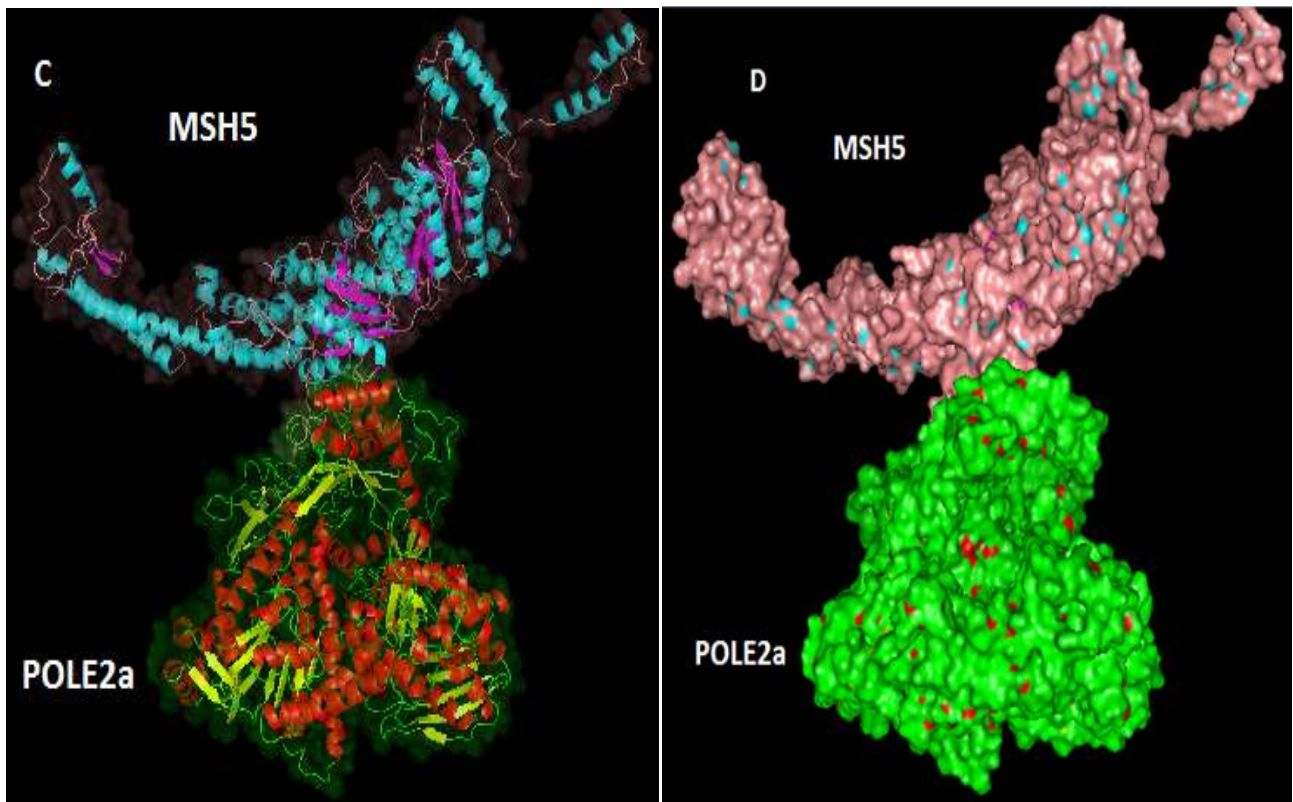


Figure 6. POLE2a-MSH5 docking site. (C) A cartoon structure of POLE2a–MSH5 interaction, (D) surface structure of POLE2a–MSH5 interaction

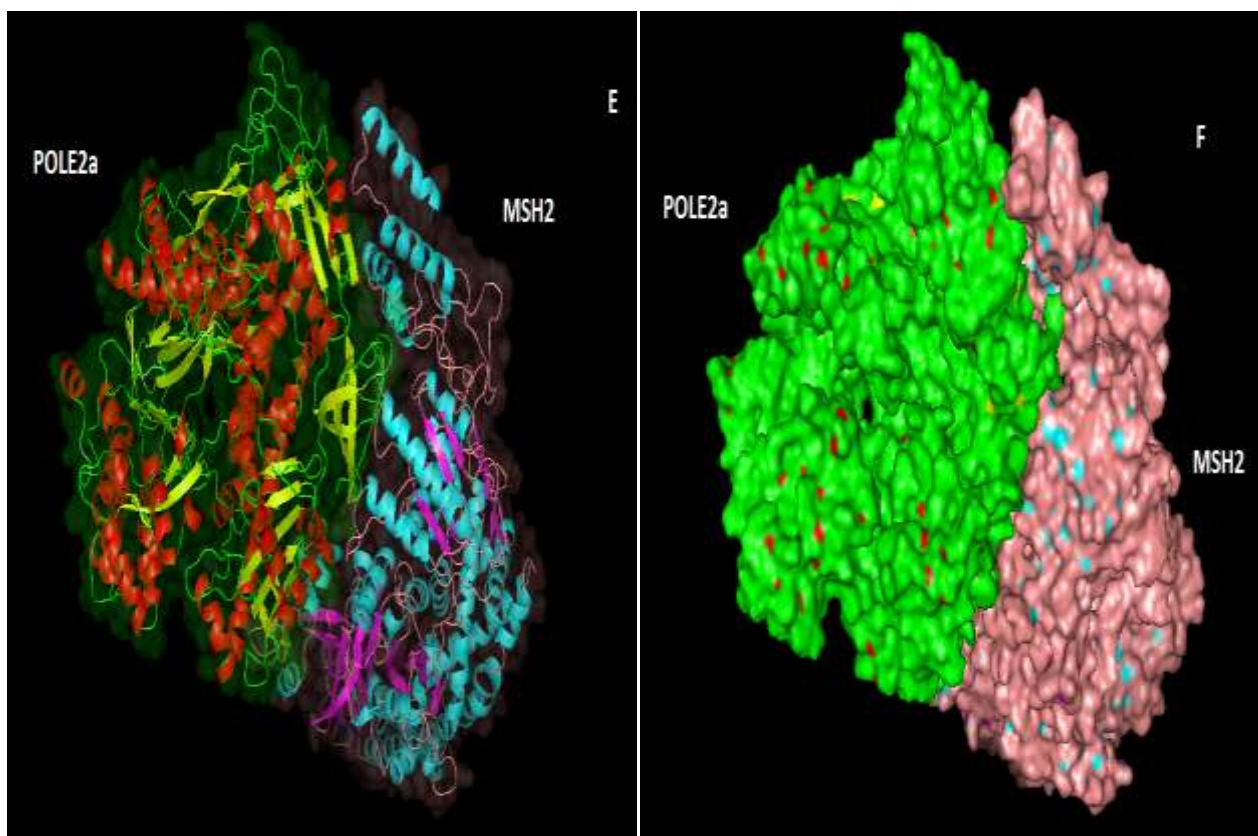


Figure 7. POLE2a-MSH2 docking site. (E) A cartoon structure of POLE2a-MSH2 interaction, (F) surface structure of POLE2a-MSH2 interaction.

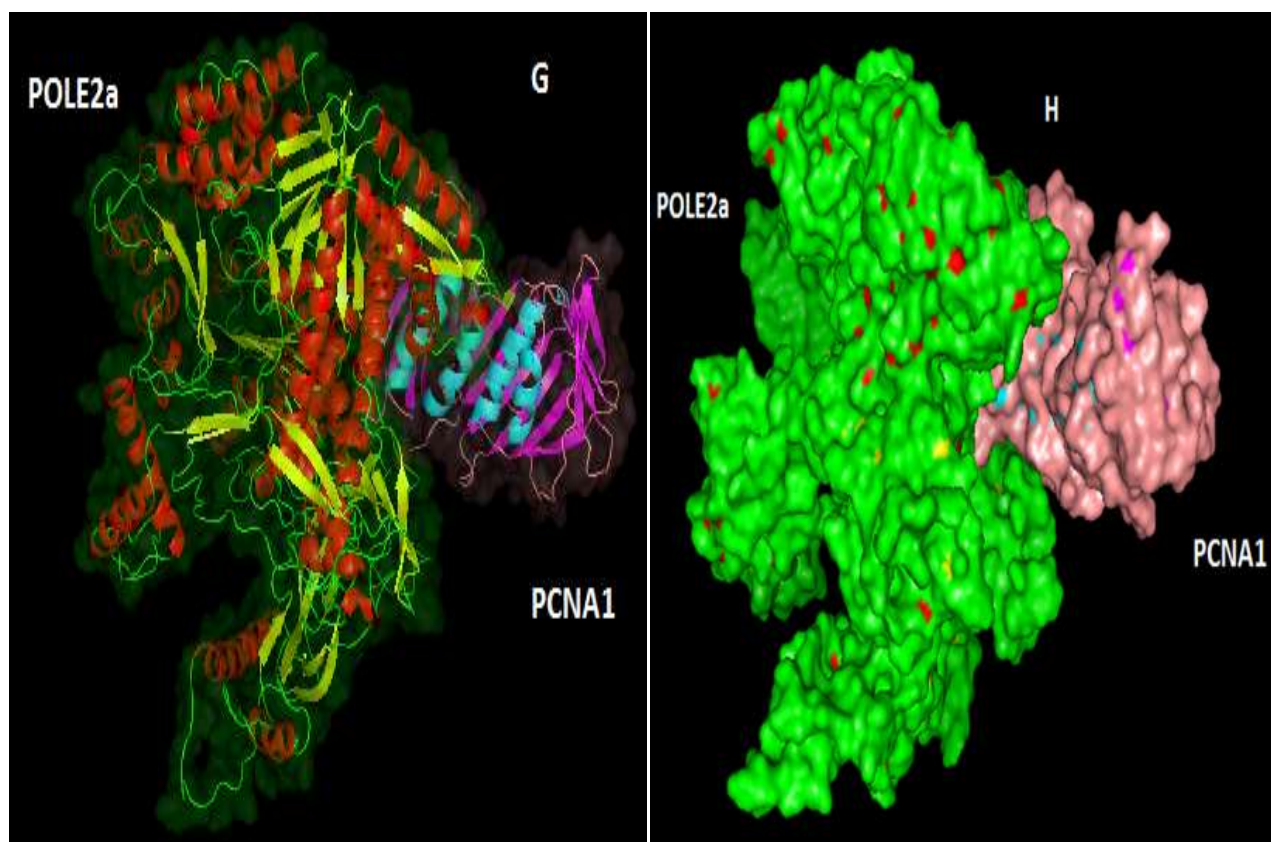


Figure 8. POLE2a-PCNA1 docking site. (G) A cartoon structure of POLE2a-PCNA1 interaction, (H) surface structure of POLE2a-PCNA1 interaction

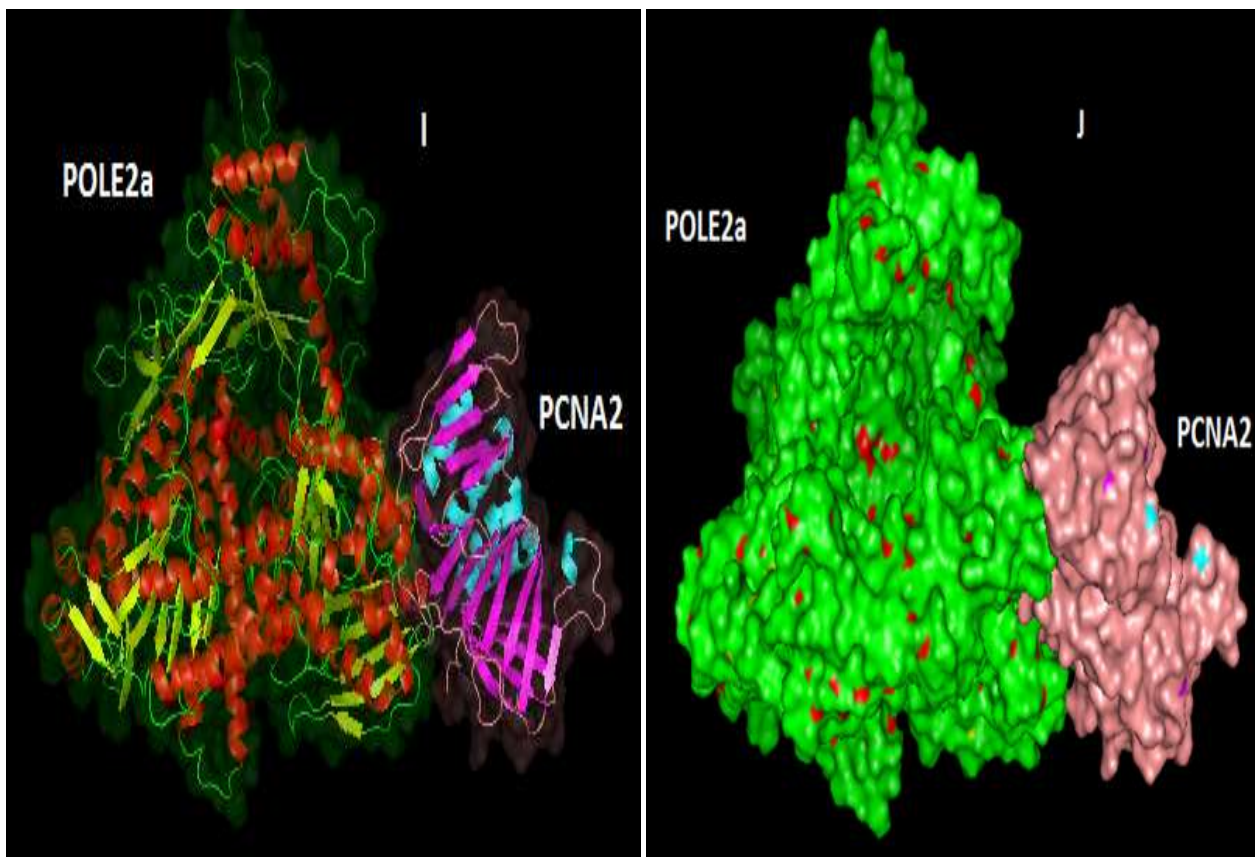


Figure 9. POLE2a-PCNA2 docking site. (I) A cartoon structure of POLE2a –PCNA2 interaction, (J) surface structure of POLE2a–PCNA2 interaction.

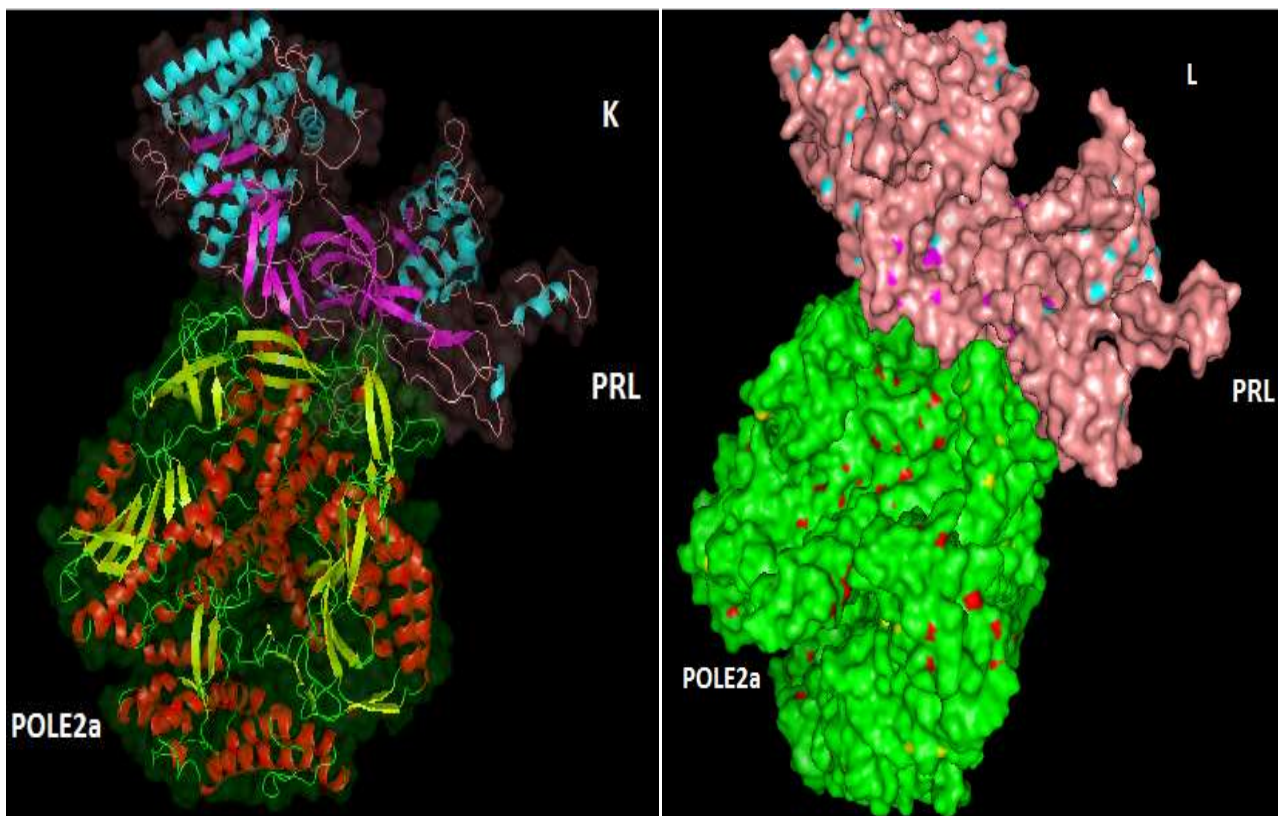


Figure 10. POLE2a-PRL docking site. (K) A cartoon structure of POLE2a–PRL interaction, (L) surface structure of POLE2a–PRL interaction.

Six interacted proteins had been selected for docking site prediction to see orientation to a chosen interactome molecule: CD45, PCNA1, PCNA2, PRL, MSH5 and MSH2 were chosen for POLE2b.

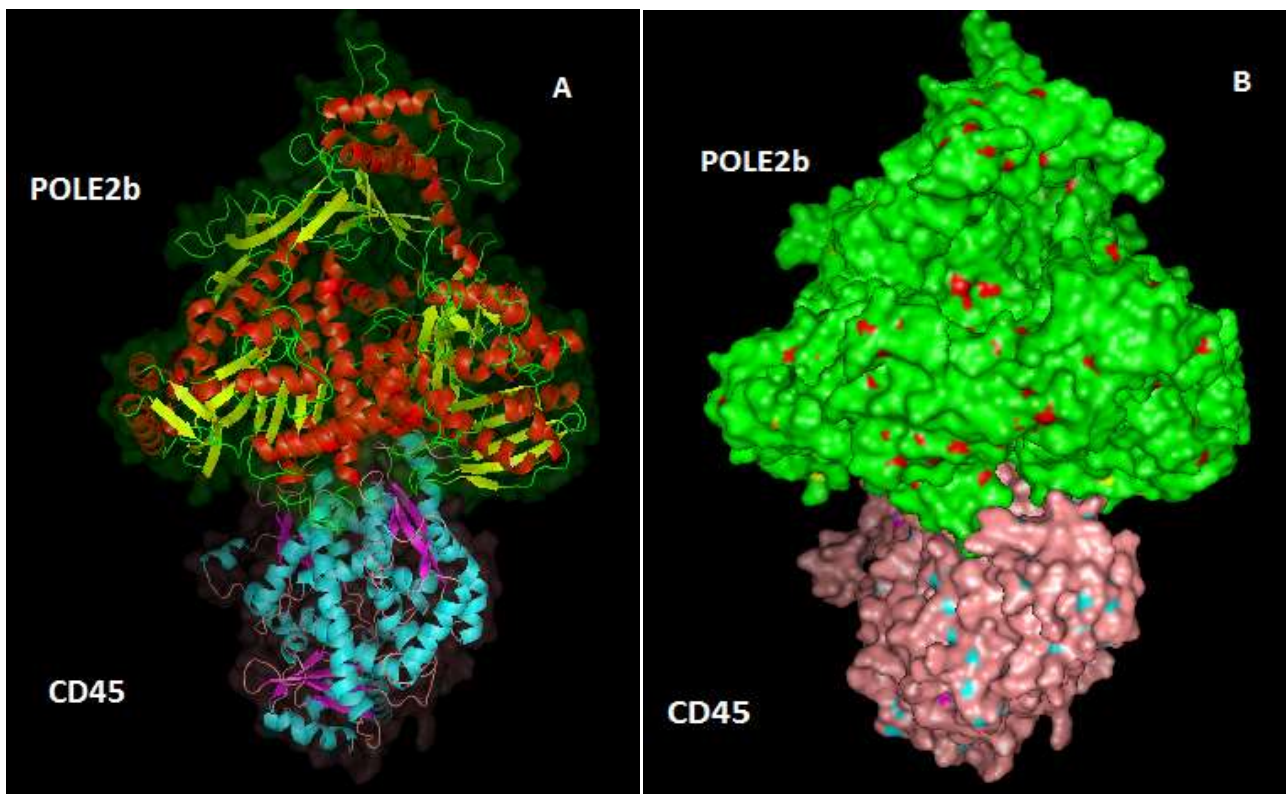


Figure 11. POLE2b-CD45 docking site. (A) A cartoon structure of POLE2b-CD45 interaction, (B) surface structure of POLE2b-CD45 interaction.

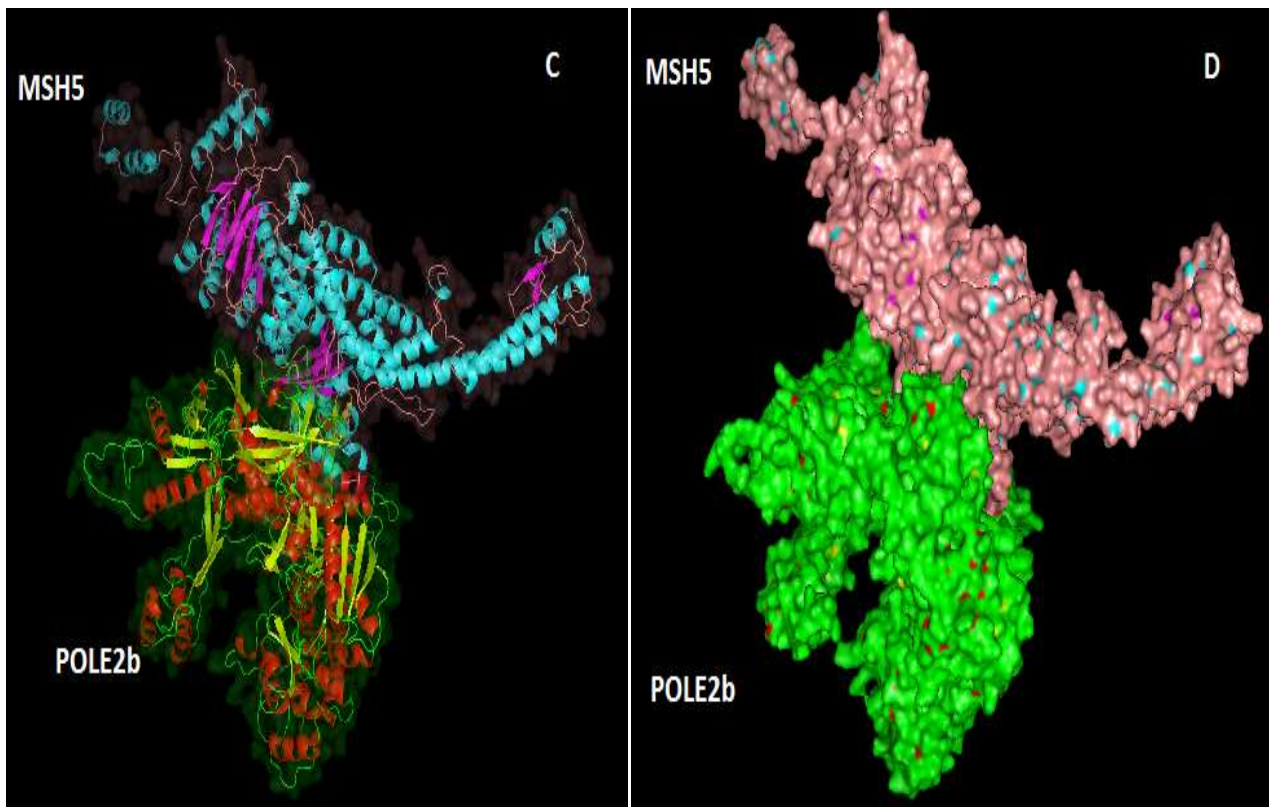


Figure 12. POLE2b-MSH5 docking site. (C) A cartoon structure of POLE2b-MSH5 interaction, (D) surface structure of POLE2b-MSH5 interaction.

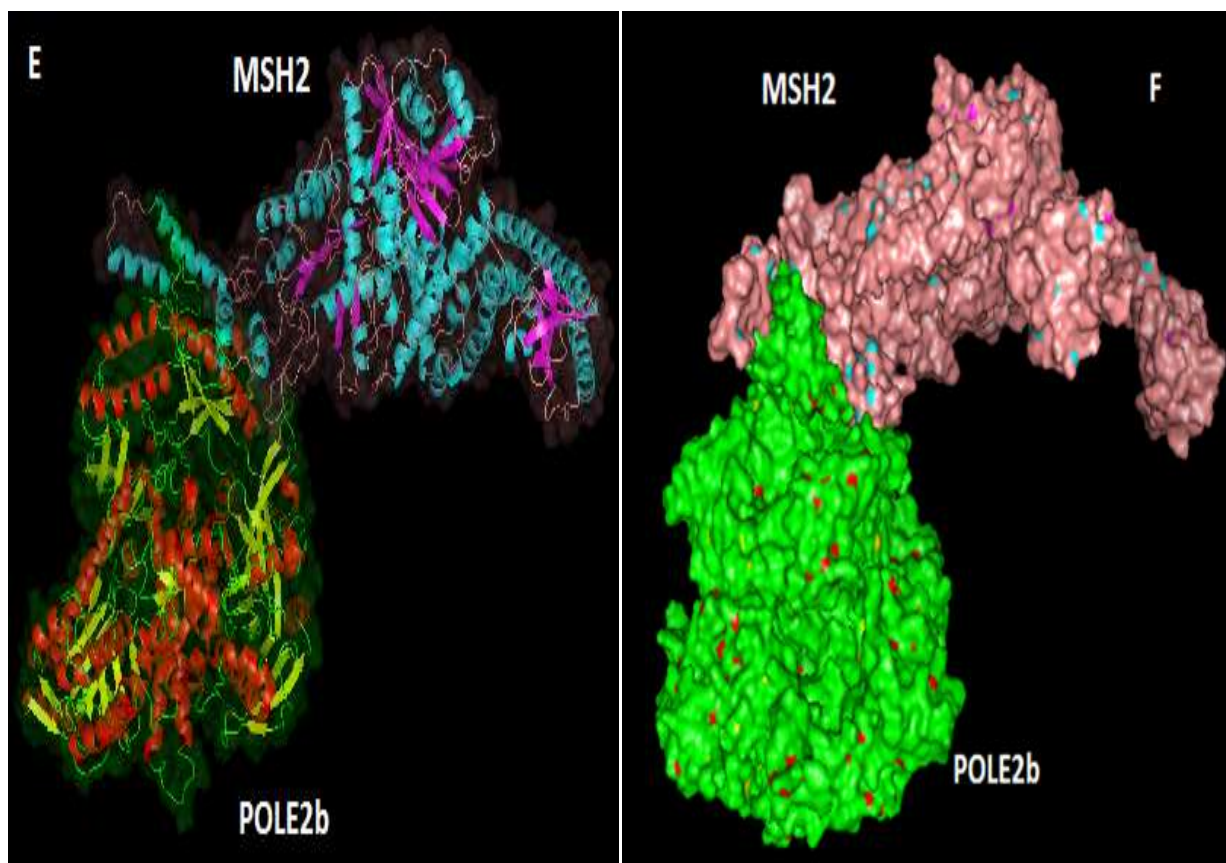


Figure 13. POLE2b-MSH2 docking site. (E) A cartoon structure of POLE2b-MSH2 interaction, (F) surface structure of POLE2b-MSH2 interaction.

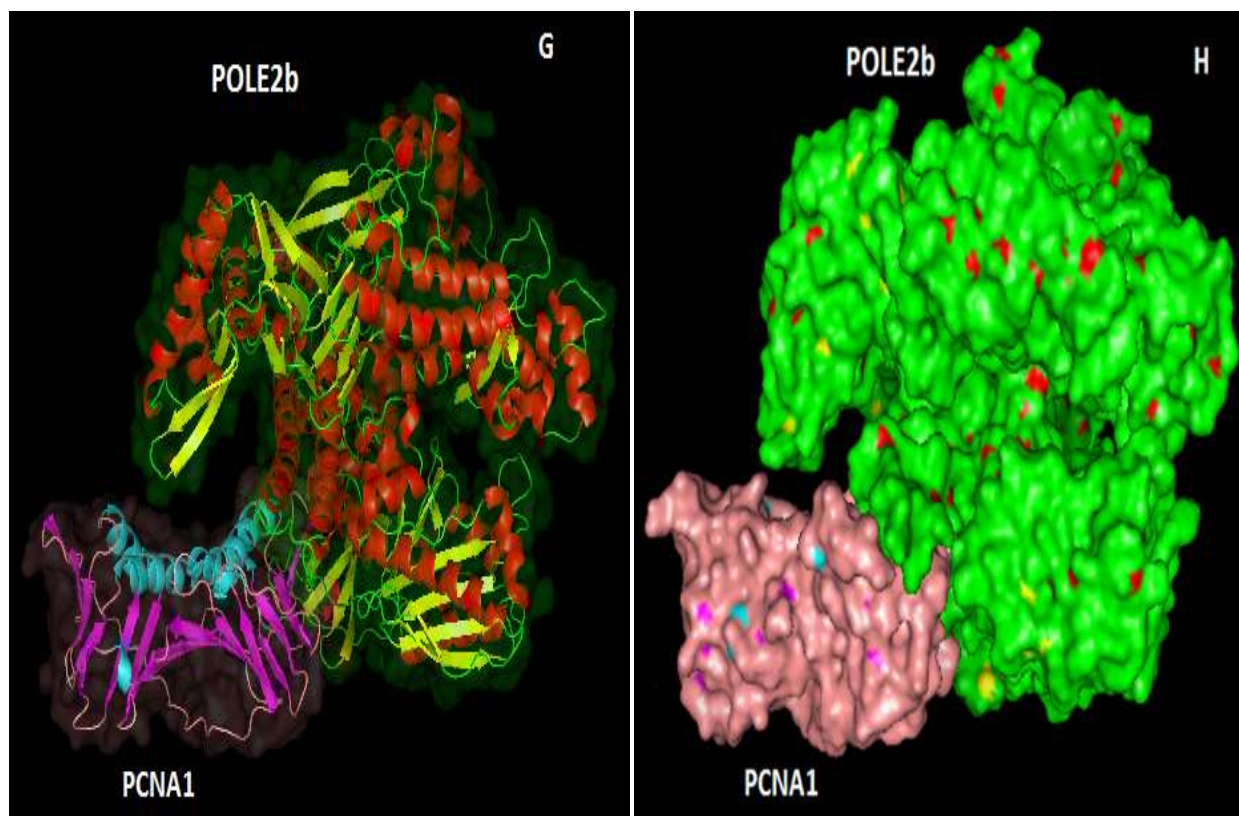


Figure 14. POLE2b-PCNA1 docking site. (G) A cartoon structure of POLE2b-PCNA1 interaction, (H) surface structure of POLE2b-PCNA1 interaction.

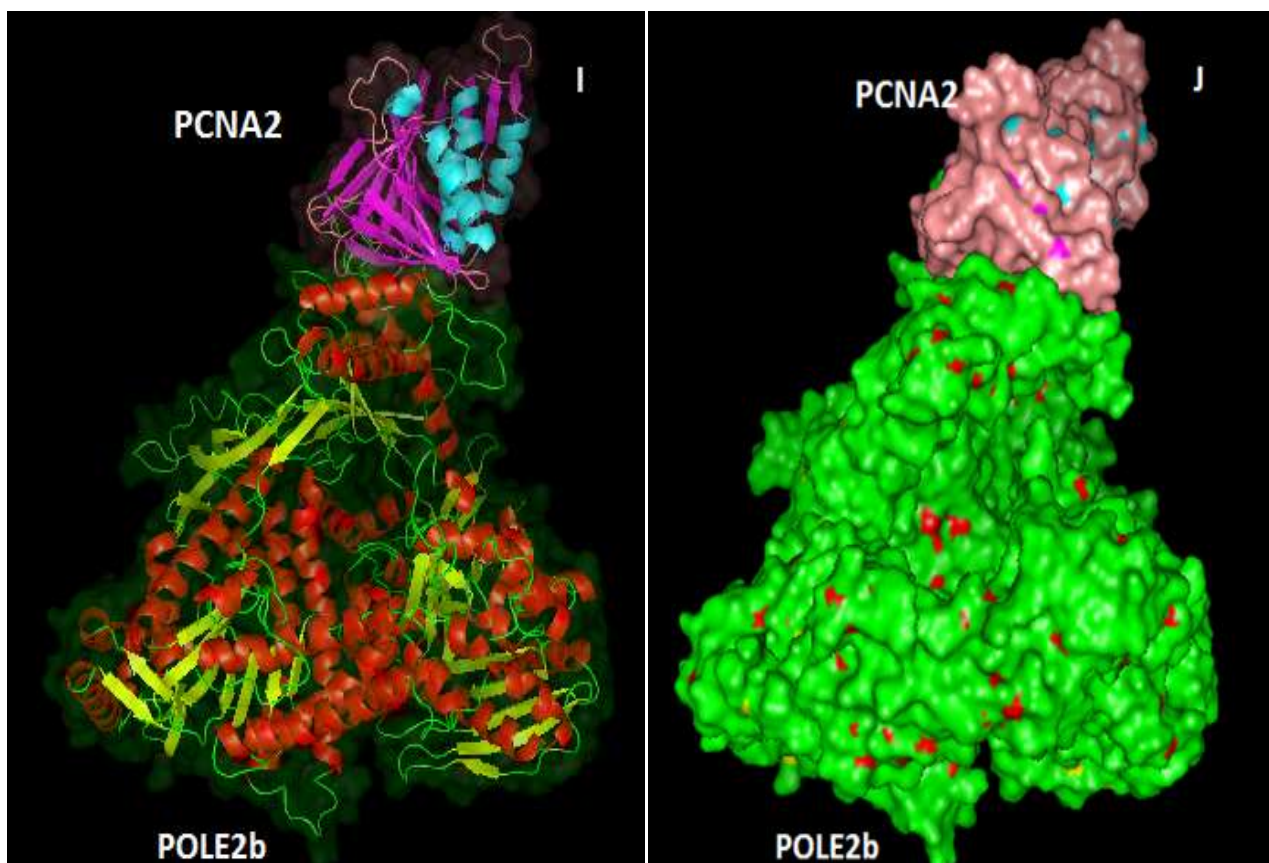


Figure 15. POLE2b-PCNA2 docking site. (I) A cartoon structure of POLE2b-PCNA2 interaction, (J) surface structure of POLE2b-PCNA2 interaction.

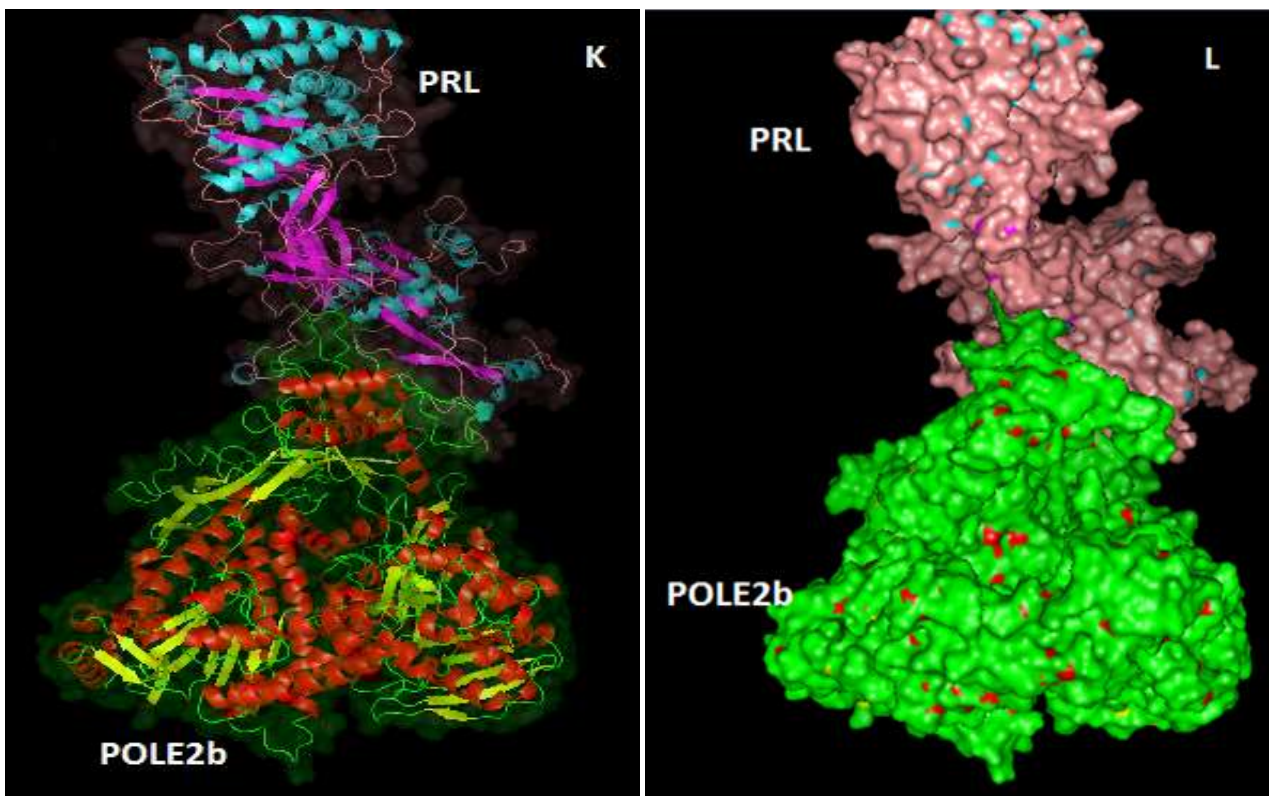


Figure 16. POLE2b-PRL docking site. (K) A cartoon structure of POLE2b-PRL interaction, (L) surface structure of POLE2b-PRL interaction.

4. Discussion

After generating the interactome of AtPOLE2A and AtPOLE2B proteins, the results showed that proteins of interest interact with various proteins. As the number of interactions was numerous, after generating interactome using STRING online tool, in this study, only the highest confidence interactions were chosen for further analysis, while medium and low confidence were omitted. The results are summarized in Fig. 4 and in table 2.

There are several DNA damage repair pathways reported. In mammals, the nucleotide excision repair (NER) is involved in both global genome repair (GGR) and transcription coupled repair (TCR) [25]. POLE2 and Damage-specific DNA Binding Protein 2 (DDB2) are major factors of classical (GGR) and (TCR) repair. The predicted amino acid sequence reveals that POLE2 is a member of DUF repeat protein family as human-POLE2 protein. DUF repeats are not involved in any catalytic activity, but they maintain protein-protein interactions. The protein in *Saccharomyces cerevisiae* is also known as NER proteins that are required for inactivation of NER. Deletion of POLE2 (POLE2A and POLE2B) showed that it is not essential for survival of yeast cells after exposing UV light unless it depends on few other radiation sensitive proteins, whereas both human POLE (POLE2A and POLE2B) and yeast TILED C is specific for transcription-coupled repair; and missing one of these proteins block either global repair mechanism of DNA or TCR. Even though scientists still do not completely understand which mechanisms are involved in the regulation of repair of UV induced damage and how the UV response functions in living organisms, studies have showed that GGR is active usually during early development, while TCR activates later in development of an organism [26,27].

Domain analysis of AtPOLE2 protein revealed that it consists of DUF-repeat domains [28]. DUF-repeat proteins consist of six or more repeating units that contain a conserved core which is usually approximately 40 amino acids long, and end with tryptophan-aspartic acid (Trp-Asp). DUF-repeat proteins form a circularized beta propeller structure [29]. When tandem copies of DUF repeats fold together they form a circular solenoid domain called the DUF1744 domain. Through the interblade loops of the DUF repeat region many of the interaction with DNA are regulated. These proteins have an important role in many processes such as signal transduction, regulation of transcription, cell death and even in some human diseases [30]. Tables 4 and 5 represent identified domains of POLE2a and POLE2b, respectively.

Accurate DNA replication is one of the most important events in the life of a cell. To perform this task, the cell utilizes several DNA polymerase complexes. Previous research investigates the role of DNA polymerase epsilon during gametophyte and seed development using forward and reverse genetic approaches. In *Arabidopsis*, the catalytic subunit of this complex is encoded by two genes, AtPOL2a and AtPOL2b, whereas the second largest regulatory subunit AtDPB2 is present as a unique complete copy [31]. Disruption of AtPOL2a or AtDPB2 resulted in a sporophytic embryo-defective phenotype, whilst mutations in AtPOL2b produced no visible effects. Loss of AtDPB2 function resulted in a severe reduction in nuclear divisions, both in the embryo and in the endosperm. Mutations in AtPOL2a allowed several rounds of mitosis to proceed, often with aberrant planes of division. Moreover, AtDPB2 was not expressed during development of the female gametophyte, which requires three post-meiotic nuclear divisions [13,14].

Our results showed the structure of AtPOLE2A and confirmed that 93.8% of the residues are in favored region, 4.4% are in allowed region and 1.8% are in outlier region, whereas AtPOLE2B showed that 93.0% of the residues are in favored region, 5.8% are in allowed region and 1.2% are in outlier regions are predicted 3D Structure Models/Ramachandran Plot Validation. Those two proteins POLE2A and POLE2B displayed key role in DNA repair. One of the aims of this study has been to indicate how do these two homologues interact and perform biological roles in DNA repair. These proteins, POLE2A and POLE2B, have a very important role in DNA repair mechanism. In this study, we provided additional roles polymerase epsilon catalytic subunit A and B in DNA repair and, through interactome analysis of POLE2A and POLE2B proteins in *Arabidopsis thaliana*, we identified other most important proteins that are participating in DNA repair. The 3D structure of POLE2 proteins was identified to decipher the complexity of NER, GG-NER and TC-NER mechanisms through identification of docking sites. We found that both POLE2A and POLE2B are interacting with MSH2, MSH5, PCNA1, PCNA2, PRL and CDC45 and we confirmed the interaction complex, which has a significant role in DNA repair processes and UV-B tolerance. The interactome analysis of POLE2A and POLE2B revealed the complexity of DNA repair mechanism in plants, also, by using STRING. Finally, the

crucial role of POLE2A and POLE2B in DNA repair, maintenance of genome integrity, regulation of protein degradation and thus cellular stability and cycle is undeniable.

5. Conclusion

In *Arabidopsis thaliana*, TIL1 and TIL2 locus encodes for POLE2A and POLE2B, respectively. POLE2A and POLE2B have a very important role in DNA repair mechanism. In this study, we have provided additional roles of POLE2A and POLE2B in *A. thaliana* through interactome analysis, and we have identified other important proteins that participate in DNA repair. The 3D structure of POLE2A and POLE2B proteins were identified to decipher the complexity of NER, GG-NER and TC-NER repair mechanisms through identification of docking sites. Also, we have confirmed that POLE2A and POLE2B are interacting with MSH2, MSH5, PCNA1, PCNA2, PLR and CD45; and that those complexes have a significant role in DNA repair processes and UV-B tolerance.

Through bioinformatics analysis, we have identified and structurally predicted six homologues of POLE2A and POLE2B in *A. thaliana*. Since they are similar in but are not identical in structure and differentiate in group's phylogenetic analysis, further inference about the functions and localization of the homologues was required. For that purpose, localization tools were used to predict subcellular locations, and the trend of differences between the homologues continued. Furthermore, the interactome analysis have shown differing function and associations with many crucial processes responsible for cellular integrity and stability maintenance. From the presented results, it can be concluded that the cellular functions of the six homologues are not completely the same, both in terms of the actual role in a cellular mechanism and possibly in the location where the homologue performs the role.

Finally, the crucial role of POLE2A and POLE2B in DNA repair, maintenance of genome integrity, regulation of protein degradation, and thus, cellular stability and cell cycle, was found to be undeniable. Heavily exposed to UV radiations, as plants are; to properly understand and analyze the DNA repair mechanism is at the forefront of genomic and proteomic research in plant science; and, *in silico* analysis is an excellent starting point for further developments.

6. References

- [1] L. S. Leutwiler, B. R. Hough-evans, and E. M. Meyerowitz, "The DNA of Arabidopsis Thaliana," *Mol Gen Genet*, vol. 194, pp. 15–23, 1984.
- [2] S. Kimura, Y. Tahira, T. Ishibashi, Y. Mori, and T. Mori, "DNA repair in higher plants; photoreactivation is the major DNA repair pathway in non-proliferating cells while excision repair (nucleotide excision repair and base excision repair) is active in proliferating cells," *Nucleic Acid Res.*, vol. 32, no. 9, pp. 2760–2767, 2004, doi: 10.1093/nar/gkh591.
- [3] H. Lodish *et al.*, *Molecular Cell Biology*. New York: w.h.freeman Macmillin Learning, 2016.
- [4] "Mutations and Disease," 2014. <https://genetics.thetech.org/about-genetics/mutations-and-disease>.
- [5] H. Frohnmeyer and D. Staiger, "Ultraviolet-B Radiation-Mediated Responses in Plants. Balancing Damage and Protection," *Plant Physiol.*, vol. 133, no. December, pp. 1420–1428, 2003, doi: 10.1104/pp.103.030049.deeper.
- [6] V. F. Zenoff, F. Siñeriz, M. E. Farías, and V. Ferna, "Diverse Responses to UV-B Radiation and Repair Mechanisms of Bacteria Isolated from High-Altitude Aquatic Environments Diverse Responses to UV-B Radiation and Repair Mechanisms of Bacteria Isolated from High-Altitude Aquatic Environments □," *Appl. Environ. Microbiol.*, vol. 72, no. 12, 2006, doi: 10.1128/AEM.01333-06.
- [7] R. P. Sinha and D. Häder, "UV-induced DNA damage and repair: a review," *Photochem. Photobiol. Sci.*, vol. 1, pp. 225–236, 2002, doi: 10.1039/b201230h.
- [8] R. P. Rastogi, A. Kumar, M. B. Tyagi, and R. P. Sinha, "Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair," *J. Nucleic Acids*, vol. 2010, 2010, doi: 10.4061/2010/592980.
- [9] M. S. Reagan and E. C. Friedberg, "Recovery of RNA polymerase II synthesis following DNA damage in mutants of *Saccharomyces cerevisiae* defective in nucleotide excision repair," *Nucleic Acid Res.*, vol. 25, no. 21, pp. 4257–4263, 1997.
- [10] M. Fousteri and L. H. F. Mullenders, "Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects," *Cell Res.*, vol. 18, pp. 73–84, 2008, doi:

- 10.1038/cr.2008.6.
- [11] J. B. Pawley, *HANDBOOK OF BIOLOGICAL CONFOCAL MICROSCOPY*, Third. Springer, 2006.
- [12] K. A. Henning *et al.*, “The Cockayne Syndrome Group A Gene Encodes a WD Repeat Protein That Interacts with CSB Protein and a Subunit of RNA Polymerase II TFIIH,” *Cell*, vol. 82, pp. 555–564, 1995.
- [13] H. Yin *et al.*, “Epigenetic Regulation, Somatic Homologous Recombination, and Abscisic Acid Signaling Are Influenced by DNA Polymerase ϵ Mutation in Arabidopsis,” *Plant Cell*, vol. 21, no. February, pp. 386–402, 2009, doi: 10.1105/tpc.108.061549.
- [14] N. A. Eckardt, “Functions of DNA Polymerase ϵ ,” *Plant Cell*, vol. 21, no. February, 2009, doi: 10.1105/tpc.109.210212.
- [15] P. D. Jenik, R. E. J. Jurkuta, and M. K. Barton, “Interactions between the Cell Cycle and Embryonic Patterning in Arabidopsis Uncovered by a Mutation in DNA Polymerase ϵ ,” *Plant Cell*, vol. 17, no. December, pp. 3362–3377, 2005, doi: 10.1105/tpc.105.036889.1.
- [16] J. López-bucio *et al.*, “*Bacillus megaterium* Rhizobacteria Promote Growth and Alter Root-System Architecture Through an Auxin- and Ethylene-Independent Signaling Mechanism in Arabidopsis thaliana,” *MPMI*, vol. 20, no. 2, pp. 207–217, 2007.
- [17] P. P. Pang and E. M. Meyerowitz, “Arabidopsis Thaliana: A Model System for Plant Molecular Biology,” *Nature*, vol. 5, pp. 1177–1181, 1987.
- [18] D. Szklarczyk *et al.*, “STRING v10: protein – protein interaction networks, integrated over the tree of life,” *Nucleic Acid Res.*, vol. 43, no. October 2014, pp. 447–452, 2015, doi: 10.1093/nar/gku1003.
- [19] P. Artimo *et al.*, “ExpASY: SIB bioinformatics resource portal,” *Nucleic Acid Res.*, vol. 40, no. May, pp. 597–603, 2012, doi: 10.1093/nar/gks400.
- [20] I. Letunic, T. Doerks, and P. Bork, “SMART 7: recent updates to the protein domain annotation resource,” *Nucleic Acid Res.*, vol. 40, no. November 2011, pp. 302–305, 2012, doi: 10.1093/nar/gkr931.
- [21] L. Liu, Z. Zhang, Q. Mei, and M. Chen, “PSI: A Comprehensive and Integrative Approach for Accurate Plant Subcellular Localization Prediction,” *PLoS One*, vol. 8, no. 10, 2013, doi: 10.1371/journal.pone.0075826.
- [22] R. Kumar, S. Jain, B. Kumari, and M. Kumar, “Protein Sub-Nuclear Localization Prediction Using SVM and Pfam Domain Information,” *PLoS One*, vol. 9, no. 6, 2014, doi: 10.1371/journal.pone.0098345.
- [23] S. Jo, M. Vargyas, J. Vasko-Szedlar, B. Roux, and W. Im, “PBEQ-Solver for online visualization of electrostatic potential of biomolecules,” *Nucleic Acid Res.*, vol. 36, no. May, pp. 270–275, 2008, doi: 10.1093/nar/gkn314.
- [24] E. Secic, J. Sutkovic, and M. R. Abdelgawwad, “Interactome analysis and docking sites prediction of radiation sensitive 23 (RAD 23) Proteins in Arabidopsis thaliana,” *Curr. Proteomics*, vol. 23, no. May 2016, 2015.
- [25] P. K. Bhatia, R. A. Verhage, and J. Brouwer, “Molecular Cloning and Characterization of *Saccharomyces cerevisiae* RAD28, the Yeast Homolog of the Human Cockayne Syndrome A (CSA) Gene,” *J. Bacteriol.*, vol. 178, no. 20, pp. 5977–5988, 1996.
- [26] J. M. S. Wong and C. J. Ingles, “A compromised yeast RNA Polymerase II enhances UV sensitivity in the absence of global genome nucleotide excision repair,” *Mol Gen Genet*, vol. 264, pp. 842–851, 2001, doi: 10.1007/s004380000374.
- [27] B. C. O. Connell and J. W. Harper, “Ubiquitin proteasome system (UPS): what can chromatin do for you?,” *Curr. Opin. Cell Biol.*, vol. 19, pp. 206–214, 2007, doi: 10.1016/j.ceb.2007.02.014.
- [28] M. Dasso, “RCC1 in the cell cycle: the regulator of chromosome condensation takes on new roles,” *TIBS*, vol. 18, pp. 96–101, 1993.
- [29] J. M. Avis and P. R. Clarke, “Ran, a GTPase involved in nuclear processes: its regulators and effectors,” *J. Cell Sci.*, vol. 109, pp. 2423–2427, 1996.
- [30] E. J. Neer, C. J. Schmidt, R. Nambudripad, and T. F. Smith, “The Ancient Regulatory-protein Family of WD-repeat proteins,” *Nature*, vol. 371, no. 22, pp. 297–300, 1994.
- [31] D. Li and R. Roberts, “Cellular and Molecular Life Sciences Human Genome and Diseases : Review WD-repeat proteins : structure characteristics , biological function , and their involvement in human diseases,” *C. Cell. Mol. Life Sci.*, vol. 58, pp. 2085–2097, 2001.