IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)
ISSN: 2455-264X, Volume 7, Issue 2 (Mar. –Apr. 2021), PP 13-30
www.iosrjournals.org

FunctionalandAnalysis of Structural Prediction of the Phenylalanine Ammonia-lyase Gene(GbPAL) from *Ginkgo biloba* (L.)

Mohamed ZoelfakarSayedAhmed*
(Genetic Resources Department, Desert Research Center (DRC), 1, Mathaf El-Matariya Street, El-Matariya B.O.P 11753 El-Matariya, Cairo, Egypt.)

Abstract:
Ginkgo biloba L. (Gymnosperm) is a medicinal and nutrient-rich plant due to its high level of flavonoid content.Various bioinformatics resources were used in this study for characterization of the phenylalanine ammonia-lyase (PAL) from Ginkgo biloba L. as key enzymes in involved pathway of plant flavonoids, the anthocyanin and other important secondary metabolites in plants.The full-length cDNA of PAL *gene sequence* (*GbPAL*) *was isolated from G. bilobacontained a 2172bp open reading frame(ORF) encoding a 724 amino acid protein.As well as, multiple alignment analysis and phylogenetic analysis revealed that GbPAL has high identity with other plant PALs.In addition, each subunit is composed of an MIO domain, a core domain and an inserted shielding domain. The highly conserved Ala-Ser-Gly triad, which acts as an active site for cyclization and dehydration for MIO group formation, was also found within the MIO domain. The parameters computed by ProtParamand Protscalesoftware were obtained.We have investigated homology modeling of GbPALusing SWISS-MODEL to better understand the structure and function analysis of the GbPAL protein, the tertiary structure was predicted using SWISS-MODEL.Functional analysis was performed using NCBI-CDD, CASTpserverand Pfam server.The deduced GbPAL protein possesses the specific signature motif (GTITASGDLVPLSYIA) of PALs. Also the 3D predicted structure protein sequence of the secondary and the tertiary structure was determined by SOPMA tool.They could be used for further studies to understand the role of each fold of this protein in the function and basal structure for performing to fully understanding the regulatory mechanisms controlling flavonoid and anthocyanin biosynthesis in Ginkgo biloba.More there, they could be used in metabolic engineering and site direction mutation to improve the transformation efficiency of this protein for developing new recombinant bacterial and plant species.*

Keyword:*Ginkgo biloba L., Phenylalanine ammonia-lyase (PAL),*GbPAL*, flavonoid, anthocyanin biosynthesis,predicted structure protein.*

Date of Submission: 22-03-2021 Date of Acceptance: 06-04-2021

I. Introduction

Ginkgo biloba L., commonly known as ginkgo or gingko, also known as the maidenhair tree, is the only living species in the division Ginkgophyta, uses in traditional medicine and as a source of food[1].*Ginkgo biloba* extract possesses interesting pharmacological properties, and it becomes the most widely sold phytomedicine to treat early-stage Alzheimer's disease, vascular dementia and many others[2].*G. biloba* is a medicinal and nutrient-rich plant due to its high level of flavonoid content. The standardized preparations of *G. biloba*, Egb761, contain 24% ginkgo flavonoids, which have been elucidated to have many pharmaceutical properties for human health [3; 4].Phenylalanine ammonia-lyase (PAL) is the first step in the pathway is catalyzed by phenylalanine ammonia-lyase (PAL; EC 4.3.1.5), which is at the gateway from the primary metabolism into the important secondary phenylpropanoid metabolism in plants.In higherplants, flavonoids represent a highly diverse class of low *M*r polyphenolic compounds, synthesized via the general phenylpropanoid pathway, giving rise to three major classes: anthocyanins, proanthocyanidins (PAs), and flavonols [5]. Also, PAL catalyzes the non-oxidative elimination of ammonia from L-phenylalanine to give *trans*-cinnamic acid, [6; 7]a substrate common to the biosynthesis of different classes of phenylpropanoid products: anthocyanins, flavonoids, ultraviolet (UV) protectants, antimicrobial furanocoumarins, isoflavonoidphytoalexins, ligins and wound phenolic esters[8; 9]. PAL plays a key regulatory role in controlling biosynthesis of all phenylpropanoid products. It has been extensively studied since its discovery by Koukol and Conn [10].Due to its central role in the phenylpropanoid pathway, PAL is one of the most extensively studied plant enzymes. It is encoded by a small genes family in the plants with four members (PAL1-PAL4) in Arabidopsis[11]. Flavonoid accumulation in *G. biloba* leaves requires coordinated expression

of genes encoding enzymes in the core phenylpropanoid pathway, such as PAL and 4- coumarate CoA ligase, and enzymes in branch pathways such as chalcone synthase and chalconeisomerase[12; 13; 14; 15].

In all plants it is encoded by a multi-gene family, ranging in copy number from four in *Arabidopsis* to a dozen or more copies in some higher plants. Also, PAL is involved in the synthesis of many pharmaceutical compounds; more attention has been paid to them in several medicinal plants and many crops.So, in tomato indicates a surprisingly large family of related sequences with ~26 copies in the diploid genome [16].Three AtPAL family genes have been obtained and completely sequenced from *Arabidopsis thaliana*,Two (AtPAL1 and AtPAL2) with a single intron are structurally similar to PAL genes that have been cloned from other plant species, at a conserved position, and a long highly conserved second exon. A third AtPAL3 differs significantly from AtPAL1 and AtPAL2 and other sequenced plant PAL genes[17].On the contrary, *PAL* genes from OsPAL from*Oryza sativa*, consists of two exons and one intron encodes a polypeptide of 710 amino acids[18]. Also, LePAL gene from genomicsequence *Lycopersiconesculentum*, suggested five different classes of phenylalanine ammonia-lyase[19]and SlPAL5 cDNA from tomato (*Solanumlycopersicum* L.) exists as multiple copies in the tomato plant [20].Both *PSPAL1* and *PSPAL2* gene each contain a single intron but the sizes of the introns are different from genomicDNA *Pisumsativum*[21].While, two new PAL genes, palg2b and palg4, encodes a polypeptide of 710 amino acids were isolated from *Populuskitakamiensis*[22].The full-length *SmPAL1* was 2,827 bp and consisted of an intron and two extrons encoding a 711-amino-acid polypeptide from *Salvia miltiorrhiza***[**23].The full length of *EsPAL*2166 bp of the open reading frame (ORF) encoded 722 amino acids; were obtained from *Ephedra sinica*[24]. Xu*et al*., [25]isolated *GbPAL* from The cDNA and genomic DNA *Ginkgo biloba* with GenBank accession no. EU071050. The full-length*GbPAL* were the same with 2172bp, it call intronless. The deduced GbPAL protein consists of 724 amino acids with a predicted molecular mass of 79.1 kDa and a p*I*of 5.96showed high identities to other plant PALs.Three full-length cDNAs encoding phenylalanine ammonia-lyasethreeisoforms (*SbPAL1*, *SbPAL2*, and *SbAPL3*) and one gene encoding cinnamate 4-hydroxylase (*SbC4H*) were investigated in different organs and suspension cells obtained from *Scutellariabaicalensis*[26].Two enzymes of the general phenylpropanoid pathway, phenylalanine ammonia-lyase (PAL) and 4-coumarate: coenzyme A-ligase (4CL), were investigated in suspension cultures of lemon balm (*Melissa officinalis)*. MoPAL1 and Mo4CL1 cDNAs were the enzymes characterized[27].As well, thecDNA and genomic DNA of L EpPAL in poinsettia (*Euphorbia pulcherrima* willd.) were isolated. The full-length of cDNA was contains a 2166-bp open reading frame (ORF) encoding 721 amino acids [28].Moreover, the cDNA of LrPAL was obtained from *Lycoris radiate* with a 2,124 bp open-reading frame that encodes 708 amino acids [29].The full-length three cDNASmPALs were obtained working draft of the *Salviamiltiorrhiza* genome and cloning and identified [30].Three *Epimediumsagittatum* PALs (EsPALs) mRNA sequences, designated respectively as EsPAL1, EsPAL2 and EsPAL3 deduced to encode 708, 716, and 739 amino acids, were isolated and characterized [31].In the past 3 decades, PAL has gained considerable significance in several clinical, industrial, and biotechnological applications. Using X-ray crystallography for 3-dimensional structure of the enzyme has been characterized to understanding of the mechanism of PAL-catalyzed reactions, including the discovery of a recently described cofactor, 3,5-dihydro-5-methyldiene-4H-imidazol-4-one [32].Several programs and servers are available for homology modeling that are planned to build a complete model from query sequences. Swiss Model program was developed by[33; 34].Homology modelling of protein structures consists of four steps: template selection, target-template alignment, model building, and model evaluation. Each of these individual steps usually requires expertise in structural biology and the use of specialized computer programs[35].Homology modeling methods use the fact that evolutionary related proteins share a similar structure [36].Homology modelling has matured into an important technique in structural biology, significantly contributing to narrowing the gap between known protein sequences and experimentally determined structures[35].Modulating many secondary metabolites biosynthesis was obtained in several plants species such as: a basis for detailed structure–function studies, the enzyme from parsley (*Petroselinumcrispum*) was crystallized, and the structure was provided by [9].Protein modelling of PAL gene from three *Fagopyrum* spp. (*F. tataricum, F. esculentum*and *F. dibotrys*)was provided byThiyagarajan*et al*., [37].All of The 3D structure and structure-function studies of several plants were available sequence information in Protein Data Base (PDB) which is the best materials for studying homology modeling of protein structure complexes.Structure prediction by homology modeling (HM) can help in understanding the 3D structure of a given protein[38].

In this study, our major objectives were to identity of cDNA gene (GbPAL) from *Ginkgo biloba* L. including molecular characterization, multiple sequence alignment (MSA) and phylogenetic relationship analysis with PALs from other plant species; to investigated homology modeling, functional and structure analysis prediction (showing ligands, global quality estimate, local quality estimate, sequence identity percentage and model template alignments) and performed some several necessary bioinformatics analysis to helped increase fully understanding the enzymatic activities of GbPAL genefrom *Ginkgo biloba*, molecular mechanisms and deduce its regulatory role in flavonoid and anthocyanin biosynthesis.

II. Material and method

Plant materials:

Several young leaves were collectedas bulk from male tree type from*Ginkgo biloba* L. (plant exchange from Frankfort, Germany) and obtained from International Park in Nasr City, Cairo, Egypt. One hundred milligram of collected frozen tissue of samples were placed in sterile 2 ml eppendorf tube and immediately dipped in liquid nitrogen, are crush into fine powder using satirize mortar for homogenization to avoid browning and degradation during RNA extraction and stored at-80 °C until use for RT-PCR two step.

RNA extraction, primers design and RT-PCR amplification:

Fine powder of 100 mg of each samples were subjected to RNA extraction following the manufacturer's procedure according (RNasy plant Mini Kit Cat No: 74904). RNA were suspended in 30 µl in RNase free water and stored in -80°C for further analysis. Purified RNA samples were measured using NanoDrop spectrophotometer (NanoDrop, Technologies Inc.). The integrity of total RNA was verified using 1.2% non-denaturing agarose gel electrophoresis. With 1 µg of isolated total RNA as the template and oligo (dT_{16}) as the primer, first-strand cDNA was synthesized using the first strand cDNA synthesis kits (SuperScript III Reverse Transcriptase) according to the manufacturer's instructions (Invitrogen, Cat No. 18080-085). The cDNA synthesis reaction was stored at -20 ºC to be used for second step PCR. The second step of PCR amplification for the full length and partial length of GbPAL gene were obtained. Polymerase chain reaction (PCR) was carried out in a 50 ml reaction mixture using gene specific primers to obtain the full length of phenylalanine ammonia-lyase (PAL) gene (Gb_PAL_Fwd ATG GTT GCA GGA GCA GAA AGG with Gb_ PAL_Rev1 TCA GAA TGG ACC AGG AGT TCC) and for partial length (Gb_PAL_FwdwithGb_ PAL_Rev2 CCT CAA CAC GAG ATT TGG CAG) according to *Ginkgo biloba*with accession no.ABZ04127.1 and ABU49842.1.

The High-Fidelity DNA polymerase, Phusion®*Taq*(Thermo Scientific, Product codes: F-530L, 500 Unit) with the ability to perform proof reading was used to amplify the cDNA. It generates blunt ends in the amplification products. Reaction was done in a 50 µl total volume. Reaction contained 2 µl cDNA, 10 µl 5X Phusion HF Buffer, 1 µl 10mM dNTP mix, 2.5 µl primer 1 (10 µM), 2.5 µl primer 2 (10 µM), 0.5 µl PhusionDNA polymerase, 31.5 µl DEPC H_2O and spin for 15 Sec. Touchdown PCR program was used to amplification for PALcDNA genes.The PCR conditions were one cycle60 sec of preheated at 98ºC, (10 cyclesfor 30 sec of denaturation at 98ºC, 30 sec for annealing at 64 - 58ºC was decreased (▼2ºC/Cycles) and (30 cycles; 30 sec. of denaturation at 98ºC, 30 sec. of annealing 56ºC, 1 min of extension at 72ºC) and followed by final extension at 72ºC for 7-10 min[39].A volume of 40 µl of each sample were analyzed using 1.2% agarose gel electrophoreses with DNA ladder size in range (100-3000bp) and stained with ethidium bromide (Eth-Br). The PCR fragments of each sample were excised and purified from the agarose gel with a clean, sharp scalpel. The gel slice was weighed in a colorless tube and the QIAquick® Gel Extraction Kit (Qiagen, cat. no. 28706) was used according to the manufacturer's procedure to elute the PCR product from the gelfor sequence.

Bioinformatics analysis of the phenylalanine ammonia-lyase (GbPAL) from *Ginkgo biloba*:

The putative *Ginkgo biloba* for phenylalanine ammonia-lyase (GbPAL) gene cDNA was analyzed by bioinformatics software. Search for PALs -related sequences was retrieved through Basic Local Alignment Tool (BLAST), homology, and domain searches in public domains, namely GenBank (www.ncbi.ncbi.nlm.nih.gov). GbPAL protein sequence from *Ginkgo biloba* with accession no. ABZ04127.1 and ABU49842.1were used for BLASTp and homology searches against other plants species. Multiple Sequence Alignments (MSA) and JalView program [40] with total 30 protein sequences were used to compare and performed using software online Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), Phylogenetic tree analysis of 30 amino acid sequences of involved GbPAL gene with other plant species were conducted in MEGA 7.0 software program by Maximum Likelihood method [41; 42].

Primary and secondary structural prediction and functional analysis:

In this study, secondary structure of GbPAL protein from *Ginkgo biloba* were analyzed using online server based on the gene sequence.For secondary protein structure of this protein was predicted and analyzed using in online SOPMA server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat. pl?Page= /NPSA/npsa_sopma.html).For hydrophilicity prediction analysis: Protscale software online (https://web.expasy.org/protscale/) was used to represent the profile produced by GbPAL amino acid scale on a selected protein[43].ProtParam software online (https://web.expasy.org/protparam/) were used as a tool to allow the computation of various physical and chemical parameters for (GbPAL) query user entered target protein and a given protein stored in Swiss-Prot[44] or TrEMBL or for a sequence.Template Selection and Search with BLAST and HHBlits has been performed against the SWISS-MODEL template library (SMTL, last update: 2021-01-27, last included PDB release: 2021-01-22). For each identified template, the template's quality has been predicted from features of the target-template alignment. The templates with the highest quality have then

been selected for model building.Models Building was built based on the target-template alignment using ProMod3. Coordinates which are conserved between the target and the template are copied from the template to the model. The prediction of GbPAL protein Three-Dimensional (3D) structural was predicted using Swiss-Model online server (https://swissmodel.expasy.org/interactive). The three-stage structure model is shown by Homology Modeling with SWISS-MODEL [35]. (Functional analysis was performed using NCBI-CDD, CASTp 3.0 server(http://sts.bioe.uic.edu/castp/index.html?1w27) and Pfam server(http://pfam.xfam.org/family/PF00221).

III. Results and discussion

Phenylalanine ammonia-lyase (PAL) is the first step in the pathway is catalyzed by phenylalanine ammonia-lyase (PAL; EC 4.3.1.5), which is at the gateway from the primary metabolism into the important secondary phenylpropanoid metabolism in plants (Figure 1).In all plants it is encoded by a multi-gene family, is involved in the synthesis of many pharmaceutical compounds; more attention has been paid to them in several medicinal plants and many crops.Good total RNA quality (i.e., A260/A230 and A260/A280 absorbance ratios within the range 1.9 - 2.4) and extraction yield (25-40ng/µl) were obtained from all the leaves material samples using nanodrop.The candidate gene exhibited high PCR success and the obtained purifiedPCR products were successfully sequenced with high-quality bidirectional sequences. The obtained two PCR products of the full length and partial cDNA of GbPAL gene were 2172bp and 310bp, respectively(Figure 2).The results showed that cDNA of *Ginkgo biloba* (GbPAL gene), contain an open reading frame of 2172bp open reading frame (ORF) encoding a 724-amino-acid protein with a calculated molecular weight of about 79.1kDaand isoelectric point (*pI*) of 5.96. The result showed that alignment of amino acid sequence data of the cDNAGbPAL genes from *Ginkgo biloba* (Gymnosperm species) with accession no. ABZ04127.1and accession no. ABU49842.1 about ~100% maximum identity within the same plant and less than 100% with other angiosperms species.Similarly, it was found that the PtPAL gene of *Pinustaeda*(GenBank, PTU39792) did not contain introns. PALgene in gymnosperm might be unique in that it does not contain introns.Our result agreement withXu*et al*., [25] were isolated and characterized the full-length cDNA and compared with genomic DNA sequences of GbPAL gene from *G. biloba*with same size, it call interonesswith contained a 2172bp open reading frame encoding a 724amino acid protein. The deduced GbPAL protein showed high identities to other plant PALsand Gymnosperm which were great difference with other flavones biosynthetic pathway gene. Our result agreement with Xu*et al*., [26]isolated three full-length cDNAs encoding phenylalanine ammonia-lyase isoforms (SbPAL1, SbPAL2, and SbAPL3) and one gene encoding cinnamate 4-hydroxylase (*SbC4H*) from different organs and suspension cells *Scutellariabaicalensis.*Meanwhile, Xu*et al*., [45]obtained the full-length cDNA of the JrPAL gene from*Juglansregia*with contained a 1935bp open reading frame encoding a 645-amino-acid protein with a calculated molecular weight of about 70.4 kD and isoelectric point (pI) of 6.7. The deduced JrPAL protein showed high identities with other plant PALs. Also, our result agreement with[23]obtained the full-length SmPAL1 from *Salviamiltiorrhiza* using genome walking technology about 2,827 bp in size and consisted of an intron and two extrons encoding a 711-amino-acid polypeptide. Sequence alignment revealed that SmPAL1 shared more than 80% identity with the PAL sequences reported in *Arabidopsis thaliana* and other plants. Also, Hou*et al*., [30]identified three SmPALs through searching the recently obtained working draft of the *Salvia miltiorrhiza* genome and full-length cDNA cloning.Bioinformatics and phylogenetic analyses showed that SmPAL1 and SmPAL3 clustered in a sub-clade of dicot PALs, whereas SmPAL2 fell into the other one.

Analysis of phenylalanine ammonia-lyase (GbPAL)from *Ginkgo biloba:*

For search PALscDNA amino acid sequence of *Ginkgo* (ABZ04127.1),*Ginkgo biloba* (ABU49842.1),*Piceaasperata* (QFI56873.1), *Larixkaempferi* (AHA44840.1), *Pinusstrobus* (QOW08349.1),*Pinusmassoniana* (ACS28225.2),*Pinustaeda* (AHX74218.2),*Pinustabuliformis* (AFR79235.1),*Pinuspinaster* (AAT66434.1),*Ephedra sinica* (BAG74770.1),*Ephedra sinica* (BAG74772.1),*Piceasitchensis* (ABK24709.1), *Selaginellamoellendorffii*(EFJ17024.1), *Physcomitrella patens* (XP_001760495.1),*Sesamumindicum*(XP_011094662.1), *Vitisriparia* (XP_034710329.1), *Nicotianatabacum* (NP_001311946.1), *S. lycopersicum-X1* (XP_004246649.1),*Solanumpennellii* (XP_015088044.1), *Rosa chinensis* (XP_024187208.1), *Gossypiumhirsutum* (NP_001313878.1), *Gossypiumaustrale* (KAA3464974.1), *Phoenix dactylifera* (XP_008805217.1), *Lactuca sativa-2* (AAO13347.1), *Rhuschinensis* (AGH13333.1), *Vitisvinifera* (XP_002268181.1), *Cucumismelo* (XP_008449237.1), *Bambusaoldhamii* (AAR24505.1), *Zea mays* (NP_001151482.1) and *Oryza sativa* (CAA34226.1)were downloaded from genbank database.These sequences were stored in a FASTA file including PALcDNA sequence of*Ginkgo biloba* **(**ABZ04127.1 and ABU49842.1).As well as,Multi Sequence Alignment (MSA) of thededuced polypeptide sequence of GbPAL and other selectedPALsfrom several plant species were carried out. It was found that GbPALpresented84.97, 84.97, 84.55, 84.55, 84.12, 84.00, 83.36, 77.82, 77.69, 84.83, 72.12 and 71.72% with high identity of different

species from Gymnosperm, Bryophytes, and Lucophyteswith E-value= zero to *Piceaasperata,Larixkaempferi*, *Pinusstrobus*, *Pinusmassoniana*, *Pinustaeda*, *Pinustabuliformis*, *Pinuspinaster*, *Ephedra sinica*, *Ephedra sinica*, *Piceasitchensis*,*Selaginellamoellendorffii*and *Physcomitrella patens*,respectively but low identity in range 70.36- 62.25% with different species fromAngiospermsas shown in Table 1.The neighbor-joining phylogenetic tree was constructed with the FASTA file by software MEGA 7.0 using the Maximum Likelihood method and the tree with the highest log likelihood (-10015.70) asshown inFigure 3. The phylogenetic tree analysis showed two branches, the first branch contain *Ginkgo biloba* (ABZ04127.1 and ABU49842.1), several species belonging to Gymnosperm, Bryophytes, and Lucophytesand other branch contained several plant speciesbelonging toAngiosperms. The results revealed that *Ginkgo biloba* (GbPAL) cDNA in this investigation was closelyto*Piceaasperata,Larixkaempferi*, *Pinusstrobus*, *Pinusmassoniana*, *Pinustaeda*, *Pinustabuliformis*, *Pinuspinaster*and*Piceasitchensis*. The genetic relationship between the PALcDNA is consistent with the phylogenetic tree.Phylogenetic tree analysis revealed that GbPAL from *Ginkgo biloba*(Gymnosperm species) were shared the same ancestor in evolution with other PALs and had a further relationship with other angiosperms species.A database searchwith (http://www.ncbi.nlm.nih.gov/) and themulti alignment sequences of amino acid showed that thededuced GbPAL gene had considerable high homology withother plant PALs gene families.The high similarity among Phenylalanine ammonia-lyase (GbPAL)proteins from *G. biloba* wasobserved from residues 2 to 726, with variability in length and composition being found in the N-terminal and C-terminal regionsasshown inFigure 4.Multiple alignment analysisshowed that the GbPAL protein sequence has high similarity with that of other reported PALs proteins,indicating that the PAL protein has been highly conserved during evolution. The conserved motif(G208TITASGDLVPLSYIA223), which is a specific signature sequence of phenylalanine ammonia-lyase[32;46], was detected at position 32 of the GbPAL protein.In addition, the active site sequence(Ala212-Ser213-Gly214),which is considered to be key for the formation of a 3,5-dihydro-5-methylidene-4*H*imidazol-4-one (MIO) group, which areaccording to JcPAL protein sequences from*Jatrophacurcas*L.byGao*et al*.,[47], is present in theconserved motif of GbPAL. Conserved deamination sites (L216, V217, L265 and A266) and catalyticactive sites (N269, G270, NDN391-393, H405 and HNQDV495-499) are thought to play an important rolein the function of the PAL proteinwas reported byJin*et al*., [48] andSangsil*et al*., [49].These sites were also found in the GbPAL protein, suggestingthat GbPAL has a similar function to that of other PALs in protein data base (PDB). GbPAL contains five types of modificationsites, i.e., N-myristoylation, casein kinase II phosphorylation, protein kinase C phosphorylation,N-glycosylation and tyrosine kinase phosphorylation sites, which are also reported and according toObPALfrom *Ocimumbasilicum*L.[50]as showed in Figure (5&6).

Advanced Structure of *Ginkgo biloba* GbPAL Protein:

Amino acid scale is defined by a numerical value assigned to each type of amino acid using Protscale software online (https://web.expasy.org/protscale/) and the most frequently used scales are the hydrophobicity or hydrophilicity scales and the secondary structure conformational parameters scales, but many other scales exist which are based on different chemical and physical properties of the amino acids [43].The parameters computed by ProtParam softwareonline (http://web.expasy.org/protparam/) include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) were obtained. Molecular weight and theoretical pI are calculated as in Compute pI/Mw [44]. Hydrophilicity prediction of Phenylalanine ammonia-lyase (GbPAL) protein from *Ginkgo biloba* by using the software of Computer pI/Mw Tool at(https://web.expasy.org/compute_pi/)thededuced forGbPALprotein had atheoretical *p*I at 5.96 and a calculated molecular weight ofabout 79.108kDaaccording to (Schomburg and Salzmann, [51].Hydrophilicity of GbPAL protein form *Ginkgo biloba* was predicted with 724 amino acid utilizing program of ProtScale–expasy [43]. The results showed that most sites of GbPAL protein is in the hydrophilic region (Score: 2.489 to -2.767) as showed in Figure (7). It was concluded that GbPAL protein from *Ginkgo biloba* is a hydrophilic protein. The parameters computed by ProtParamwas obtained to the molecular weight, theoretical *pI*, amino acid composition: (63Ala (A) 8.7%, 31Arg (R) 4.3%, 35Asn (N) 4.8%, 32Asp (D) 4.4%, 10Cys (C) 1.4%, 37Gln (Q) 5.1%, 50Glu (E) 6.9%, 55Gly (G) 7.6%, 14His (H) 1.9%, 32Ile (I) 4.4%, 76Leu (L) 10.5%, 42Lys (K) 5.8%, 19Met (M) 2.6%, 21Phe (F) 2.9%, 33Pro (P) 4.6%, 53Ser (S) 7.3%, 45Thr (T) 6.2%, 5Trp (W) 0.7%, 20Tyr (Y) 2.8%, 51Val (V) 7.0%, zero (O) pyrrolysine-Pyl 0.0%, and zero (U) selenocysteine-Sec 0.0%), Atomic composition: C: 3480, H: 5570, N: 964, O: 1079, S: 29,Extinction coefficient: 57300, Estimated half-life: 30 hours (mammalian reticulocytes, in vitro), The instability index (II) is computed to be 32.45, Aliphatic index: 87.31 and grand average ofhydropathicity (GRAVY): -0.257 as according to Gasteiger*et al*., [44].The findings suggest that the GbPAL protein under study werehydrophobic in nature due to presence of high non-polar residues content. GbPAL protein has high percentage of alanine (8.7%), leucine (10.5%), Gly (7.6%) and serine (7.3%). Results also showed that the maximum number of amino acid present in the sequence was found to be leucine (10.5%) and the least was for cysteine (1.4%) andtryptophan (0.7%). Total number of negatively charged residues (Asp

+ Glu): 82, Total number of positively charged residues (Arg + Lys): 73.Recently a new method called the self-optimized prediction method (SOPM) has been described to improve the success rate in the prediction of the secondary structure of proteins [52]. The improved SOPM method (SOPMA) correctly predicts 69.5% of amino acids for a three-state description of the secondary structure (alpha-helix, beta-sheet and coil) in a whole database containing 126 chains of non-homologous (less than 25% identity) proteins. The secondary structure of the GbPAL protein was predicted by the SOPMA tool (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html). The results indicated that GbPAL consists mainly of α-helices (Hh) (399 is 55.11%) and random coils (Cc) (232 is 32.04%) as well as a few extended strands (Ee) (53 is 7.32%) and beta turns (Tt) (40 is 5.52%)as shown in figure (8). The structure prediction from primary to advanced structure is an important task in the field of protein research. The three-dimensional structure model of phenylalanine ammonia-lyase (GbPAL) protein from *Ginkgo biloba* were predicted by the Swiss-Model server, by homology modeling based on the available structures [41; 53].CASTp server (Computed Atlas of Surface Topography of proteins) is a web server that provides online services for locating, delineating and measuring these geometric and topological properties of protein structures [54].CASTpwas obtained for the positions of some feature such as; active site 110 (Proton donor/acceptor), binding site 354 (Substrate), modified residue 203 (2,3didehydroalanine/Ser), mutagenesis site 203, 210 (S-A. complete loss of activity or no loss activity), cross-link 203 (5-imidazolinone/Ala-Gly). These topographic features provide structural basis and micro-environments for proteins to carry out their functions such as ligand binding, DNA interaction and enzymatic activity.Several different databases provided functional analysis of proteins by classification of protein families and predicting domains and important sites. Pfamserver(http://pfam.xfam.org/family/PF00221)version 33.1was obtained. The result showed that the phenylalanine aminomutase are members of family Aromatic amino acid lyase from taxol biosynthesis. The family gene aromatic amino acid lyase in InterPro web (Classification of protein families) which are includes proteins with phenylalanine ammonia-lyase, EC:4.3.1.24 (PAL), histidine ammonia-lyase, EC:4.3.1.3 (HAL), and tyrosine aminomutase, EC:5.4.3.6, activities. The PAL enzyme has the MIO prosthetic group (4-methylideneimidazole-5-one), derived from an Ala-Ser-Gly motif. Other MIO enzymes include Phe, Tyr, and His ammonia-lyases. This entry serves as an exception to overrule assignments by IPR005922 for phenylalanine ammonia-lyaseaccording to Apper*et al*., [55].Template search in either FASTA or Clustal format with the highest quality for model building have then been selected from BLAST, Camacho *et al*., [56] and HHBlitsdatabase,Steinegger*et al*., [57]has been performed against the SWISS-MODEL Template Library (SMTL-ID) for evolutionary related structures matching the target sequence. HHblits (a database of HMMs) first converts the query sequence (or MSA) to an HMM. This is conventionally done by adding pseudocounts of amino acids that are physicochemically similar to the amino acid in the query [58].For each identified template, the template's quality has been predicted from features of the target-template alignment. Models were built based on the target-template alignment using ProMod3. In case loop modelling with ProMod3 fails, an alternative model is built with PROMOD-II [59].For Model Quality Estimation: The global and per-residue model quality has been assessed using the QMEAN scoring function[59]. Ligands present in the template structure are transferred by homology to the model. For Oligomeric State Conservation: The quaternary structure annotation of the template is used to model the target sequence in its oligomeric form[60].The method is based to other template features to provide a Quaternary Structure Quality Estimate (QSQE). The QSQE score is a number between 0 and 1, reflecting the expected accuracy of the interchain contacts for a model built based a given alignment and template. Higher numbers indicate higher reliability. This complements the Global Model Quality Estimation (GMQE) score which estimates the accuracy of the tertiary structure of the resulting model.The homologous sequence of phenylalanine ammonia-lyase (GbPAL) protein from*Ginkgo biloba* with more than 28 (PALs) templates were available in databases by named using the PDB ID format.The result showed the top 19 filtered templates in the table (2). A further 28 (PALs) templates were found which were considered to be less suitable for modelling than the filtered list.(SMTL –ID:1eb4.1.A, 1gk2.1.A, 1gk3.1.A, 1gkj.1.A, 1gkm.1.D, 1t6j.1.A, 1y2m.1.A, 2nyf.1.A, 2nyn.1.A, 2o6y.1.A, 2o7f.1.A, 2ohy.1.A, 2qve.1.A, 2rjr.1.A, 3czo.1.A, 3kdy.1.A, 3kdz.1.A, 3nz4.1.A, 3unv.1.A, 4baa.1.A, 4bab.1.D, 4c5s.1.A, 4c5u.1.A, 4v2q.1.A, 4v2r.1.A, 5ltm.1.A, 6s7j.1.A, 6v6h.1.A). The result showed with high homologous sequence with (GbPAL) protein from*Ginkgo biloba* such as: SMTL-ID: 6f6t.1.A (Phenylalanine ammonia-lyase 1 (PAL-1) from *Petroselinumcrispum*complexed with S-APPA) with biounitoligomeric state: homo-tetramer, QMEAN: -0.56, GMQE: 0.81, QSQE: 0.97, sequence identify: 68.29%; sequence similarity: 0.50% and Ligand: CV2. Because the C terminal of GbPAL protein from*Ginkgo biloba* is poor homology to *Petroselinumcrispum*less than 35 amino acids, the template of matching 35-724 to *Petroselinumcrispum*PAL. SMTL-ID: 1w27.1.A (for phenylalanine ammonia-lyase (PAL) from *Petroselinumcrispum*) with biounitoligomeric state: homo-tetramer, QMEAN: -1.65, GMQE: 0.80, QSQE: 0.98, sequence identify: 68.44%, sequence similarity: 0.50% % and Ligand: DDT as shown inFigure 9. Because the C terminal of GbPAL protein from*Ginkgo biloba* is poor homology to *Taxuschinensis* less than 50 amino acids, the template of matching 50-718 to *Taxuschinensis* PAL. OtherSMTL-ID: 4c5s.1.A (for phenylalanine ammonia-lyase structural

investigations into the stereochemistry and activity of a phenylalanine-2, 3-aminomutase from *Taxuschinensis*) with biounitoligomeric state: homo-tetramer, QMEAN: -1.39, GMQE: 0.73, QSQE: 0.86, sequence identify: 47.68%, sequence similarity: 0.43% and Lagand: BQ7. Three-dimensional structure was some models built successfully as template alignment. The results were close to the protease real space conformation. Local estimates of the model quality based on the QMEAN scoring function are shown as a per-reside plot and as a global score in relation to a set of high-resolution PDB structures (Z-score). Based on the results was obtained, homology model can be considered a reliable model. The high similarity was observed for GbPAL protein with model template alignment, but the N terminal and C terminal regions showed some variability in length andcomposition. It was clear from the multi sequence alignment that GbPAL protein from gymnosperm plants were moresimilar to each other than to those of angiosperm plants,as confirmed by the phylogenetic analysis. The stringentconservation among evolutionary diverse plant speciesmay indicate the functional significance of these aminoacids.Homology modeling was used as useful tool for the prediction of protein structure when the model protein (with a known sequence and an unknown structure) is related with high/identify to at least one other protein with both a known sequence and a known structure. Structural information is often more valuable than sequence alone for determining protein function[40; 35]. The quality was obtained for the predicted structure by homology modeling depends on the degree of similarity between the model and template sequences. If the similarity was very low, homology modeling of the query protein does not yield a meaningful result.Homology modeling and bioinformatics analysis of GbPAL gene was important which involved in the biosynthetic pathways of flavonoids from *Ginkgo biloba* L. Multiple sequence alignment and phylogenetic tree analysis based on templates modeling of SMTL-ID: 4c6g.1A, 4c5u.1A, 4c6g.1A, 5f5t.1A and 4c5s.1a were revealed that GbPAL shared the same ancestor in evolution with other PALs and had a further relationship with other angiosperms species as shown in figure (10).This result agreement to Xu*et al*., [25]were suggested molecular modeling of JrPAL showed that the 3D model of JrPAL was similar to that of PAL protein from *Petroselinumcrispum* (PcPAL), implying that JrPAL may have similar functions with PcPAL. Phylogenetic tree analysis revealed that JrPAL shared the same evolutionary ancestor of other PALs and had a closer relationship with other angiosperm species. Bioinformatics analysis show that the expression model of GbPAL gene has close relation with the formation of flavonoids and anthocyanin biosynthesiswhich indicated that the PAL regulate the flavonoids transferring to anthocyanin in *G. biloba*. Wybenga*et al*., [61]elucidated Phenylalanine-2,3-aminomutase (PAM) from *Taxuschinensis*, three-dimensional (3D) structures with a bound (R)-β-phenylalanine analogue and with bound trans-cinnamic acid. In addition, 3D structures of the (inactive) Y322A and N231A mutants of PAM were elucidated, which were found to be MIO-less as shown in figure (10).This study provides fundamental information for the PALs gene family and lays the foundation of potential resources using bioinformatics analysis for enhancing the content of flavonoids in *Ginkgo biloba* L. by metabolic engineering.

IV. Conclusion

We have isolated and sequenced cDNAGbPALgenesfrom*Ginkgo biloba* L.in this study. As well as, Multiple Sequence Alignment (MSA) involved 30 amino acid sequences ofPALs genewere done with each gene familiesbelonging to Gymnosperms, *Bryophytes*, Fern (*Lycophytes*) and Angiospermsincluded GbPALgeneswith high identify and similarity. Multiple sequence alignment and phylogenetic analysis was performed using the amino acid sequence of GbPAL geneswith other known plant-specificPALs gene.The results revealed that *Ginkgo biloba* (GbPAL) cDNA in this investigation was closelytoPALsspecies fromGymnosperms. We have investigated functionaland structure analysis prediction to characterize enzymatic activities of GbPAL genefrom *Ginkgo biloba* using bioinformatics tool. The GbPALtheoretical 3D model were predicted using homology modeling to showing ligands, global quality estimate, local quality estimate, sequence identity percentage and model template alignments. Our results indicated that molecular identification,phylogenetic analysis, homology modelingandstructure analysis prediction of several genes encoding key enzymes are the first step to fully understanding the regulatory mechanisms controlling flavonoid and anthocyanin biosynthesis in *Ginkgo biloba*.

V. Figure and Table

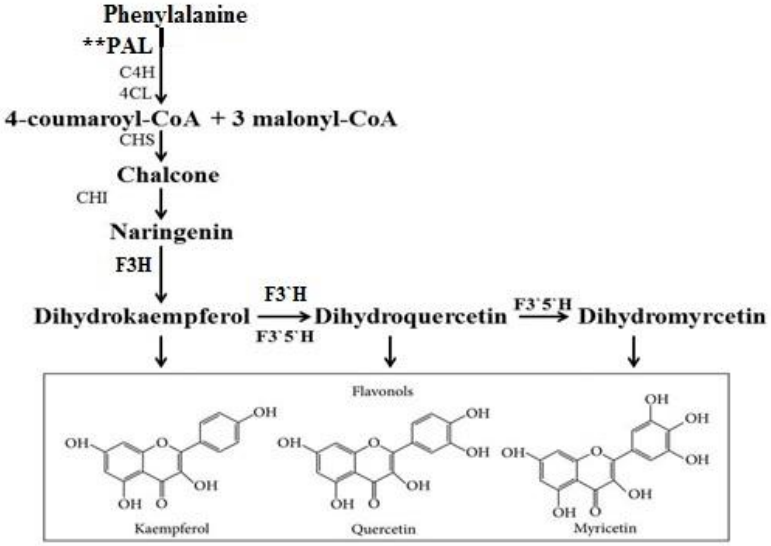
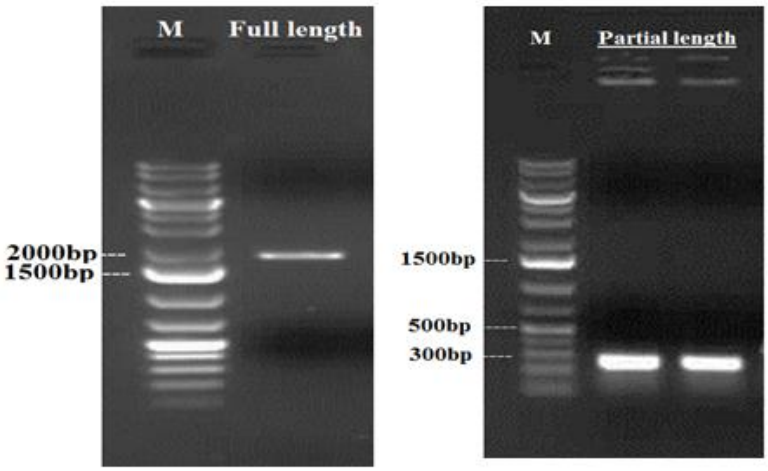


Figure (1): Flavonol biosynthesis in plants, redrawn fromCzemmel, *et al*., [5].The Two stars indicate the enzyme and compound analyzed in this study. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaroyl CoA ligase; CHS, chalcone synthase; CHI, chalconeisomerase; *F3H, flavone 3-hydroxylase; *F3`H, flavonoid 3`-hydroxylase; F3`5`H, flavonoid 3`5`-hydroxylase; FLS, flavonol synthase.



Figure(2): The PCR amplification product of the *GbPAL* gene. M: DNA molecular weight standards. The PCR product full-length cDNA*GbPAL* and partial length *GbPAL* from *Ginkgo biloba* were obtained.

Table (1): Homology of amino acid sequences for 30 selected accession lists of phenylalanineammonia-lyase proteinsand its related with (GbPAL)from *Ginkgo biloba*in this study, for Code 1-12 plant species fromGymnosperm, for code 13 from*Bryophtes*(fern), for code 14 from *Lycophytes* and 15- 30 plant species from Angiosperms, BLAST top hits against GenBank protein database, similarity score, accession no. and accession length.

		Scientific Name	Accession no.	Per. Identify	Max Score	Total Score	Query Cover	Acc. Len
Gymnosperm	1	*Ginkgo*	ABZ04127.1	100.00	1504	1504	100	724
	2	*Ginkgo biloba*	ABU49842.1	99.96	1504	1504	100	724
	3	*Picea asperata*	QFI56873.1	84.97	1280	1280	100	720
	4	*Larix kaempferi*	AHA44840.1	84.97	1278	1278	100	719
	5	*Pinus strobus*	QOW08349.1	84.55	1270	1270	100	718
	6	*Pinus massoniana*	ACS28225.2	84.55	1270	1270	100	718
	7	*Pinus taeda*	AHX74218.2	84.12	1266	1266	100	709
	8	*Pinus tabuliformis*	AFR79235.1	84.00	1256	1256	100	718
	9	*Pinus pinaster*	AAT66434.1	83.36	1243	1243	99	754
	10	*Picea sitchensis*	ABK24709.1	84.83	1277	1277	100	720
	11	*Ephedra sinica*	BAG74770.1	77.82	1171	1171	100	722
	12	*Ephedra sinica*	BAG74772.1	77.69	1170	1170	100	722
Bryoph. (fern)	13	*Selaginella moellendorffii*	EFJ17024.1	72.12	1021	1021	93	676
Lycoph.	14	*Physcomitrella patens*	XP_001760495.1	71.72	1054	1054	95	710
Angiosperms	15	*Sesamum indicum*	XP_011094662.1	70.36	1000	1000	95	711
	16	*Vitis riparia*	XP_034710329.1	69.66	999	999	96	710
	17	*Nicotiana tabacum*	NP_001311946.1	68.21	1000	1000	97	717
	18	*S. lycopersicum-X1*	XP_004246649.1	68.27	1001	1001	96	721
	19	*Solanum pennellii*	XP_015088044.1	68.41	1003	1003	96	724
	20	*Rosa chinensis*	XP_024187208.1	66.90	1004	1004	99	727
	21	*Gossypium hirsutum*	NP_001313878.1	68.67	1009	1009	99	721
	22	*Gossypium australe*	KAA3464974.1	68.81	1010	1010	99	721
	23	*Phoenix dactylifera*	XP_008805217.1	68.73	1009	1009	97	704
	24	*Lactuca sativa-2*	AAO13347.1	67.82	976	976	95	713
	25	*Rhus chinensis*	AGH13333.1	68.41	986	986	96	707
	26	*Vitis vinifera*	XP_002268181.1	69.52	1001	1001	96	710
	27	*Cucumis melo*	XP_008449237.1	69.11	1001	1001	97	717
	28	*Bambusa oldhamii*	AAR24505.1	65.59	940	940	97	712
	29	*Zea mays*	NP_001151482.1	67.00	934	934	95	718
	30	*Oryza sativa*	CAA34226.1	62.25	861	861	95	701

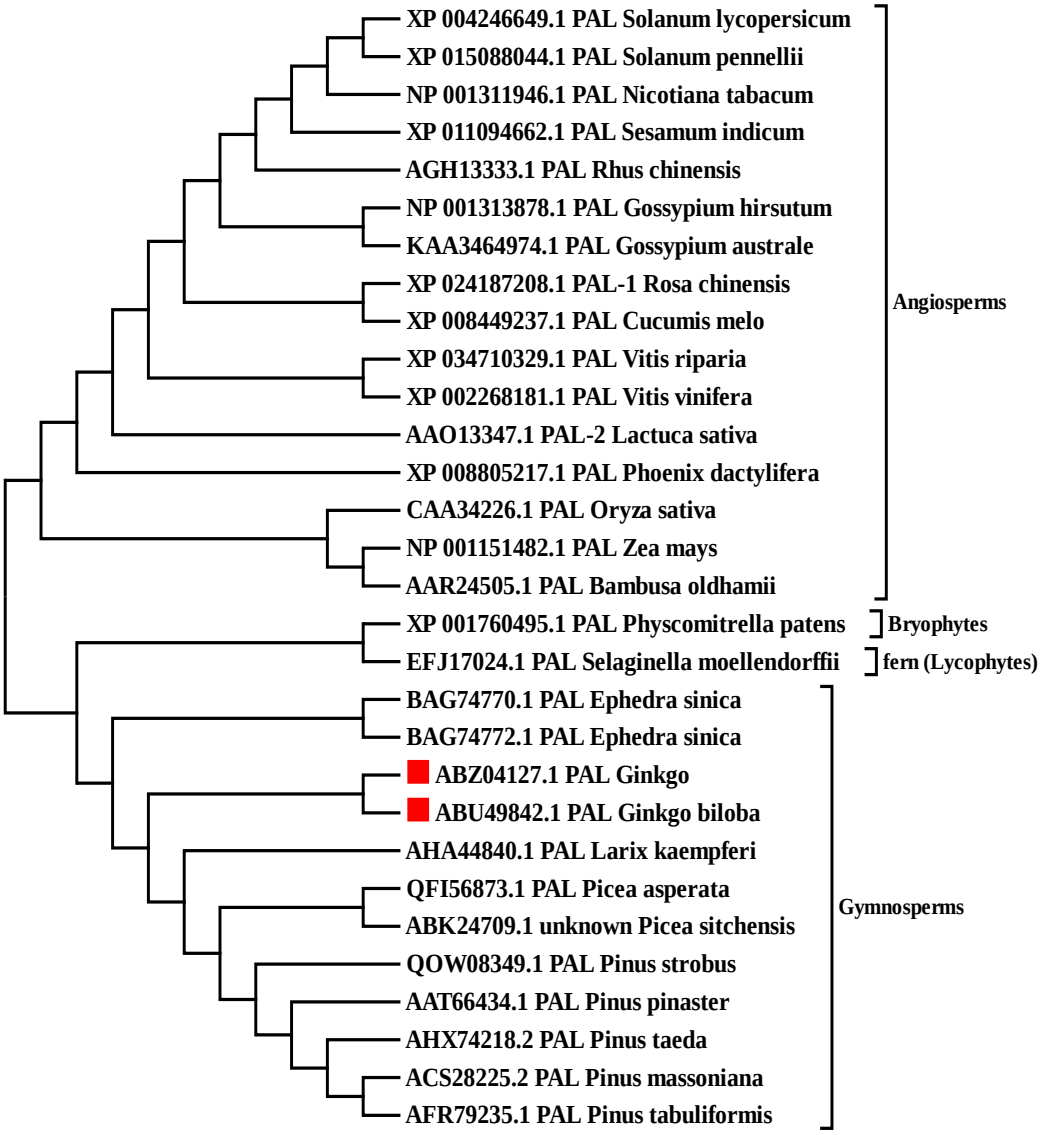


Figure (3): Molecular phylogenetic analysis involved 30 amino acid sequences of phenylalanine ammonia-lyase (PALs) gene from different plant species belonging Gymnosperms, *Bryophytes*, Fern (*Lycophytes*)) and Angiospermsincluded *Ginkgo biloba* GbPAL were conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-10015.70) is shown[42].

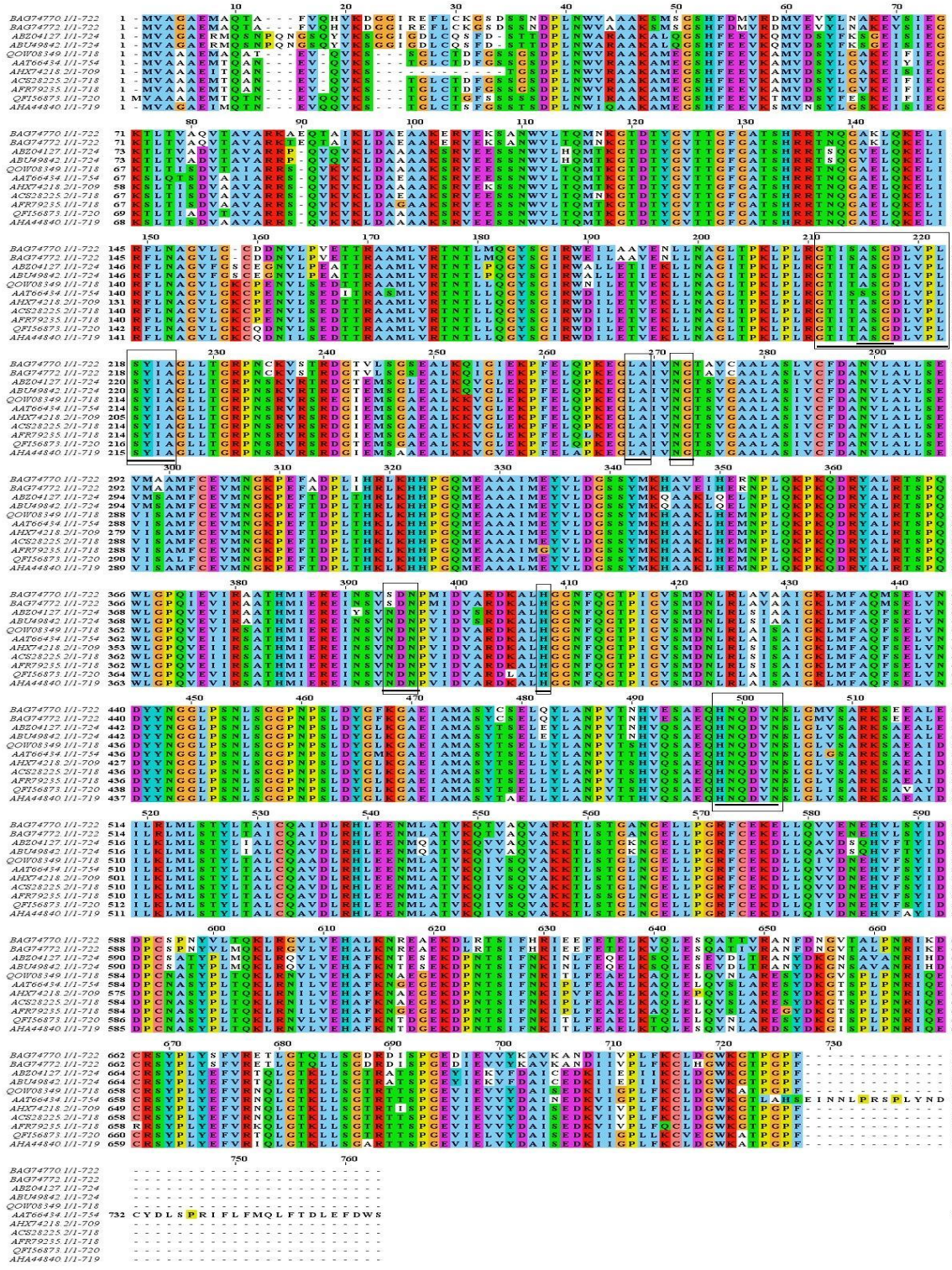


Figure (4): The amino acid sequence alignment of phenylalanine ammonia-lyase (PALs) gene from 11 PALs sequences included GbPAL gene sequence *Ginkgo biloba* with other plant species (Gymnosperm) were used in this study (www.https://www.ebi.ac.uk/Tools/msa/clustalo/).Genbank accession numbers for the PALs proteins in the alignment are as follow:*Ephedra sinica* (BAG74770.1),*Ephedra sinica* (BAG74772.1),*Ginkgo* (ABZ04127.1),*Ginkgo biloba* (ABU49842.1),*Pinusstrobus* (QOW08349.1),*Pinuspinaster* (AAT66434.1),*Pinustaeda* (AHX74218.2),*Pinusmassoniana* (ACS28225.2),*Pinustabuliformis* (AFR79235.1),*Piceaasperata* (QFI56873.1) and *Larixkaempferi* (AHA44840.1).

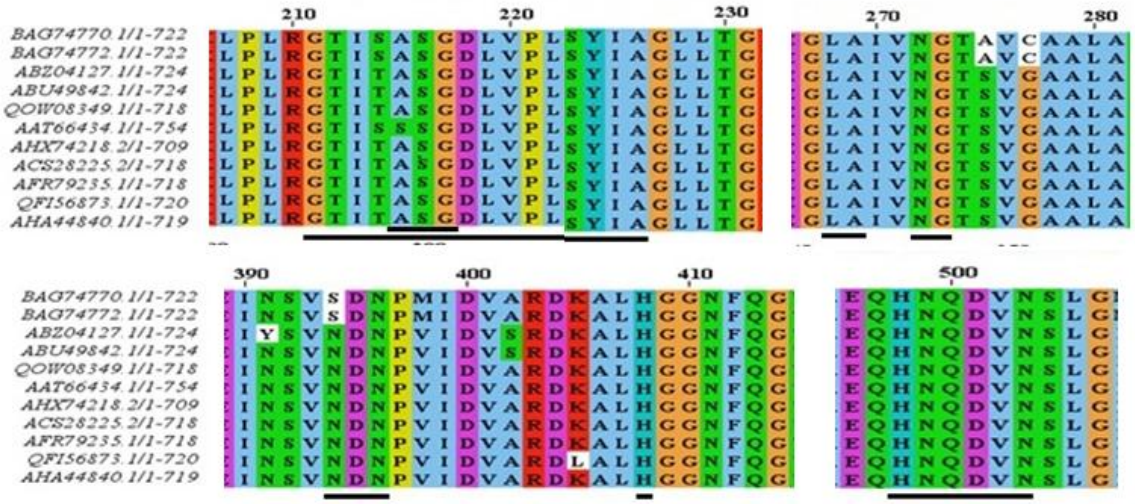


Figure (5): Sequence alignment of amino acid with the highly conserved catalytic domains in various MIO-enzymes with (Ala-Ser-Gly)was marked by double underline.

Phenylalanine ammonia-lyase from *Ginkgo biloba* (GbPAL):

```
        10         20         30         40         50         60
MVAGAERMQS NPQNGSQYVK SGGIGDLCQS FDSTTDPLNW ARAAKALQGS HFEEVKQMVD

        70         80         90        100        110        120
SYFKSGEISI EGKTLTVADV TAVARRPQVQ VKLDAAAAKS RVEESSNWVL HQMTKGTDTY

       130        140        150        160        170        180
GVTTGFGATS HRRTSQGVEL QKELIRFLNA GVFGSCEGNV LPEATTRAAM LVRTNTLPQG

       190        200        210        220        230        240
YSGIRWALLE TIEKLLNAGI TPKLPLRGTI TASGDLVPLS YIAGLLTGRP NSKVRTRDGT

       250        260        270        280        290        300
EMSGLEALKQ VGLEKPFELQ PKEGLAIVNG TSVGAALASI VCFDANVLAV LSEVMSAMFC

       310        320        330        340        350        360
EVMNGKPEFT DPLTHRLKHH PGQMEAAAIM EYVLDGSSYM KQAAKLQELN PLQKPKQDRY

       370        380        390        400        410        420
ALRTSPQWLG PQVEVIRAAT HMIEREIYSV NDNPVIDVSR DKALHGGNFQ GTPIGVSMDN

       430        440        450        460        470        480
LRLSIAAIGK LMFAQFSELV NDYYNGGLPS NLSGGPNPSL DYGLKGAEIA MASYTSELEY

       490        500        510        520        530        540
LANPVTNHVQ SAEQHNQDVN SLGLVSARKS AEALEILKLM LSTYLIALCQ AVDLRHLEEN

       550        560        570        580        590        600
MQATVKQVVA QVAKKTLSTG KNGELLPGRF CEKDLLQAVD SQHVFTYIDD PCSATYPLMQ

       610        620        630        640        650        660
KLRQVLVEHA FKNTESEKDP NTSIFNKINL FEQELKSQLE SEVDLTRANY DKGNSAVANR

       670        680        690        700        710        720
IHDCRSYPLY EFVRTQLGTK LLSGTRATSP GEYIEKVFDA ICEDKIIEPI IKCLDGWKGT

  724
PGPF
```

Figure (6): The deducedamino acid sequence of the full-length cDNA ofGbPALwith ORF 1-724 aafrom *Ginkgo biloba* with catalytic domains in various MIO-enzymes (Ala212-Ser213-Gly214). The conserved motif was (G208TITASGDLVPLSYIA223), conserved deamination sites (L216, V217, L265 and A266) and catalytic active sites (N269, G270, NDN391-393, H405 and HNQDV495-499) were marked by double underline.

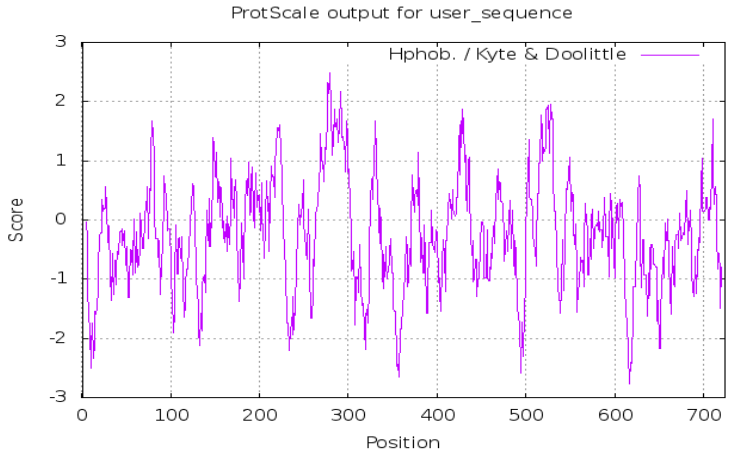


Figure (7): Hydrophilicityprofi**l**e of Phenylalanine ammonia-lyase (GbPAL)protein from *Ginkgo biloba* L. using (https://web.expasy.org/protscale/).

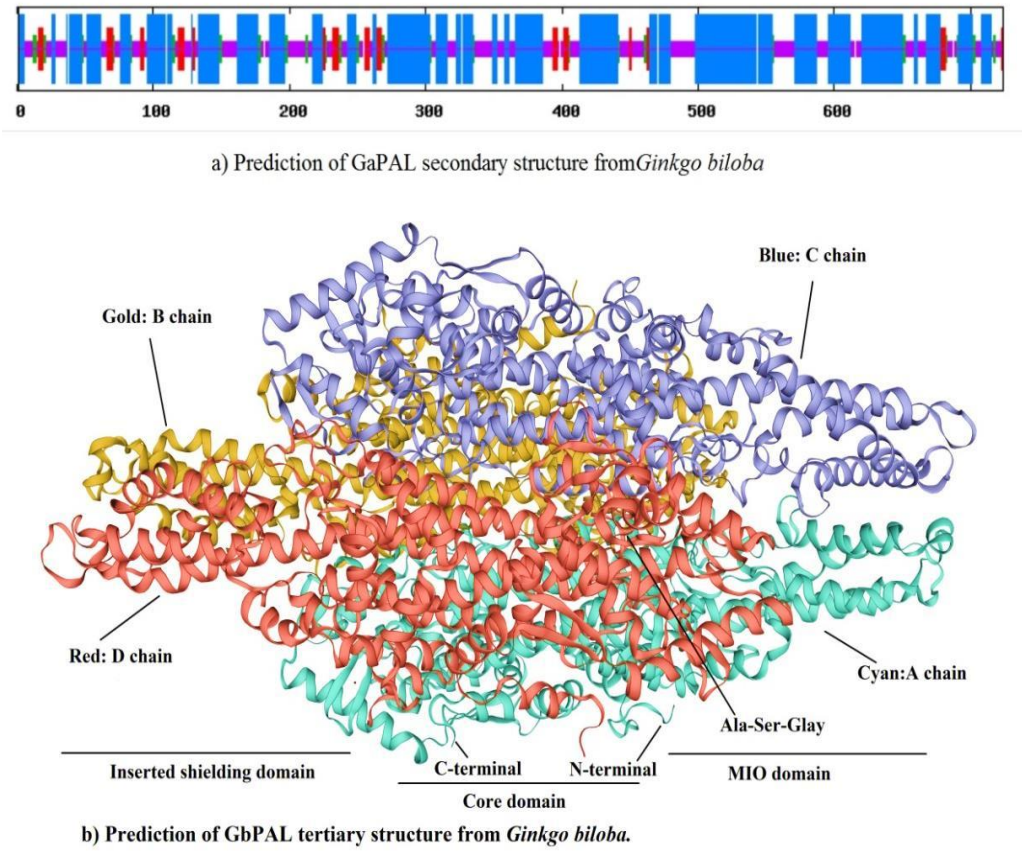


Figure (8): Prediction of structure with three-dimensional model of Phenylalanine ammonia-lyase (GbPAL) protein from*Ginkgo biloba*. a)Prediction of GbPAL secondary structure;ά–helicase in red and green and β-sheets are indicated by patches in bluewas obtained by SOPMA tool.Turns and loops are indicated by lines. (https://swissmodel.expasy.org/interactive).b)Prediction of GbPAL tertiary structure.

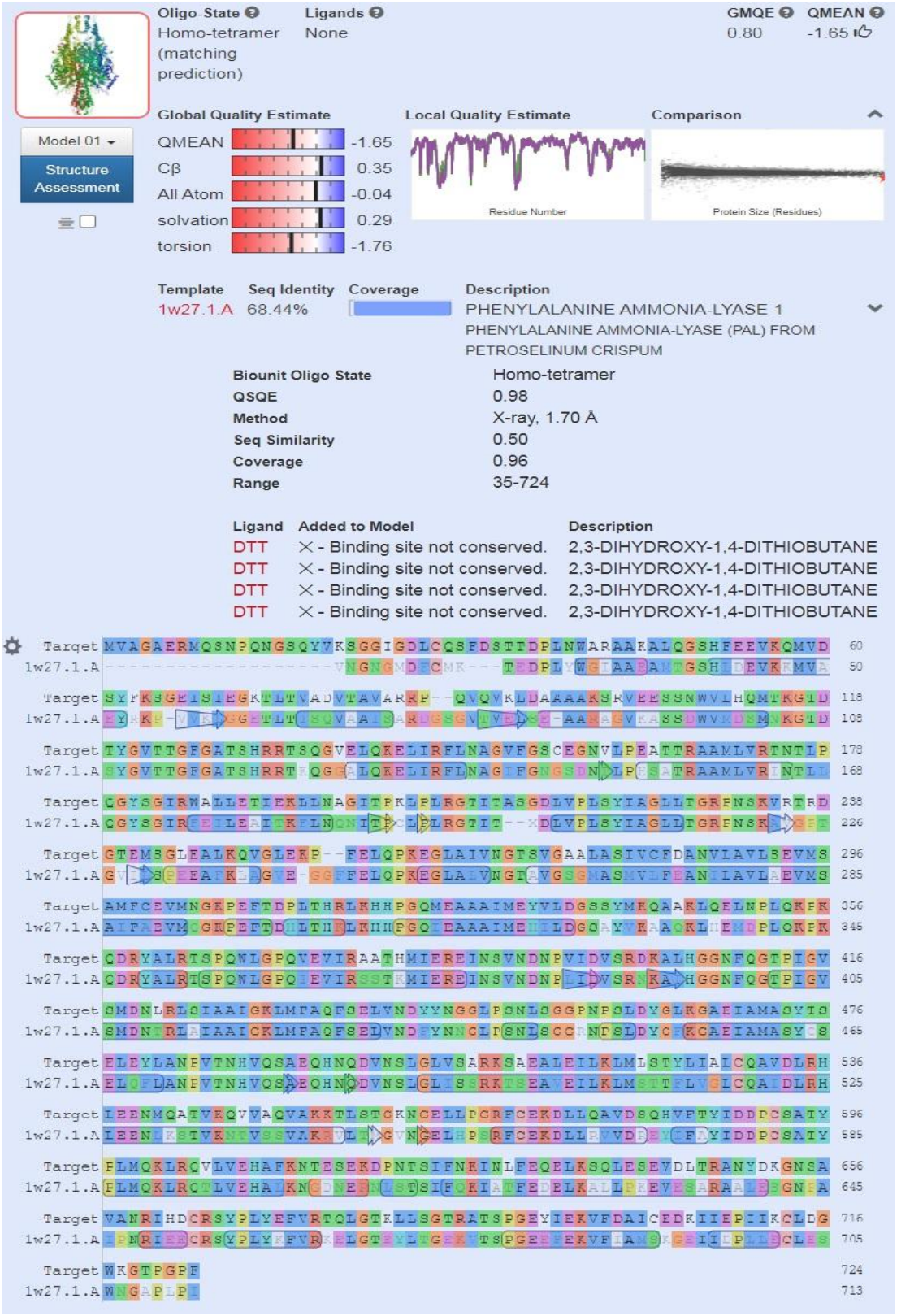


Figure (9): Homology model of protein structures complexes Phenylalanine ammonia-lyase (GbPAL) protein from *Ginkgo biloba* showing global, local quality and target-template alignments (www.swissmodel.expasy.org).

Table (2): as shown the top 19 filtered templatesfor protein with high qualityfor phenylalanine ammonia-lyase (PALs)model building using X-ray according to parameter (Sequence Identity, Oligo-state (matching prediction), QSQE, found by (HHblits or BLAST), Resolution, Sequence, Similarity, Coverage and Description). A further 28 templates were found which were considered to be less suitable for modelling than the filtered list.

	ID Template from (SMTL)	Seq Identity	Oligo-state	QSQE	Found by	Resolution	Seq Similarity	Cover	Description
1	6f6t.1.A	68.29	homo-tetramer	0.97	HHblits	1.90Å	0.50	0.96	(PAL) from *Petroselinum crispum complexed with S-APPA*
2	6f6t.1.A	67.42	homo-tetramer	0.97	BLAST	1.90Å	0.50	0.98	
3	1w27.1.A	68.44	homo-tetramer	0.98	HHblits	1.70Å	0.50	0.96	Phenylalanine ammonia-lyase (PAL-1) from *Petroselinum crispum* PAL- 1
4	1w27.1.A	67.56	homo-tetramer	0.98	BLAST	1.70Å	0.50	0.98	
5	6rgs.1.A	68.15	homo-tetramer	0.93	HHblits	2.42Å	0.50	0.96	Crystal Structure of Phenylalanine Ammonia Lyase (PAL) from *Petroselinum crispum*
6	6rgs.1.A	67.28	homo-tetramer	0.93	BLAST	2.42Å	0.50	0.98	
7	6at7.1.A	64.19	homo-dimer	0.72	HHblits	2.49Å	0.49	0.95	Phenylalanine Ammonia-Lyase (PAL) from *Sorghum bicolor*
8	6at7.1.A	64.62	homo-dimer	0.71	BLAST	2.49Å	0.49	0.94	
9	6at7.1.B	64.19	homo-dimer	0.71	HHblits	2.49Å	0.49	0.95	
10	6at7.1.B	64.62	homo-dimer	0.70	BLAST	2.49Å	0.49	0.94	
11	4c5s.1.A	47.68	homo-tetramer	0.86	HHblits	1.85Å	0.43	0.92	PAL from *Taxus chinensis*
12	4baa.1.A	47.68	homo-tetramer	0.85	HHblits	2.50Å	0.43	0.92	Redesign of a PAL into a beta- PAL
13	4c6g.1.A	47.84	homo-tetramer	0.85	HHblits	2.10Å	0.43	0.93	Structural Investigations PAL from *Taxus chinensis*
14	4c5u.1.A	47.84	homo-tetramer	0.83	HHblits	2.19Å	0.43	0.93	
15	4v2r.1.A	47.98	homo-tetramer	0.84	HHblits	2.20Å	0.43	0.92	Dissecting the structural determinants PAL
16	4bab.1.D	47.83	homo-tetramer	0.84	HHblits	2.56Å	0.43	0.92	Redesign of a PAL into a beta- PAL
17	4v2q.1.A	48.13	homo-tetramer	0.84	HHblits	1.95Å	0.43	0.92	Dissecting the structural determinants PAL
18	4c6g.1.A	49.09	homo-tetramer	0.83	BLAST	2.10Å	0.44	0.91	Structural Investigations PAL from *Taxus chinensis*
19	3nz4.1.A	50.38	homo-dimer	0.60	BLAST	2.38Å	0.44	0.90	Crystal Structure of a Taxus *Phenylalanine Aminomutase*

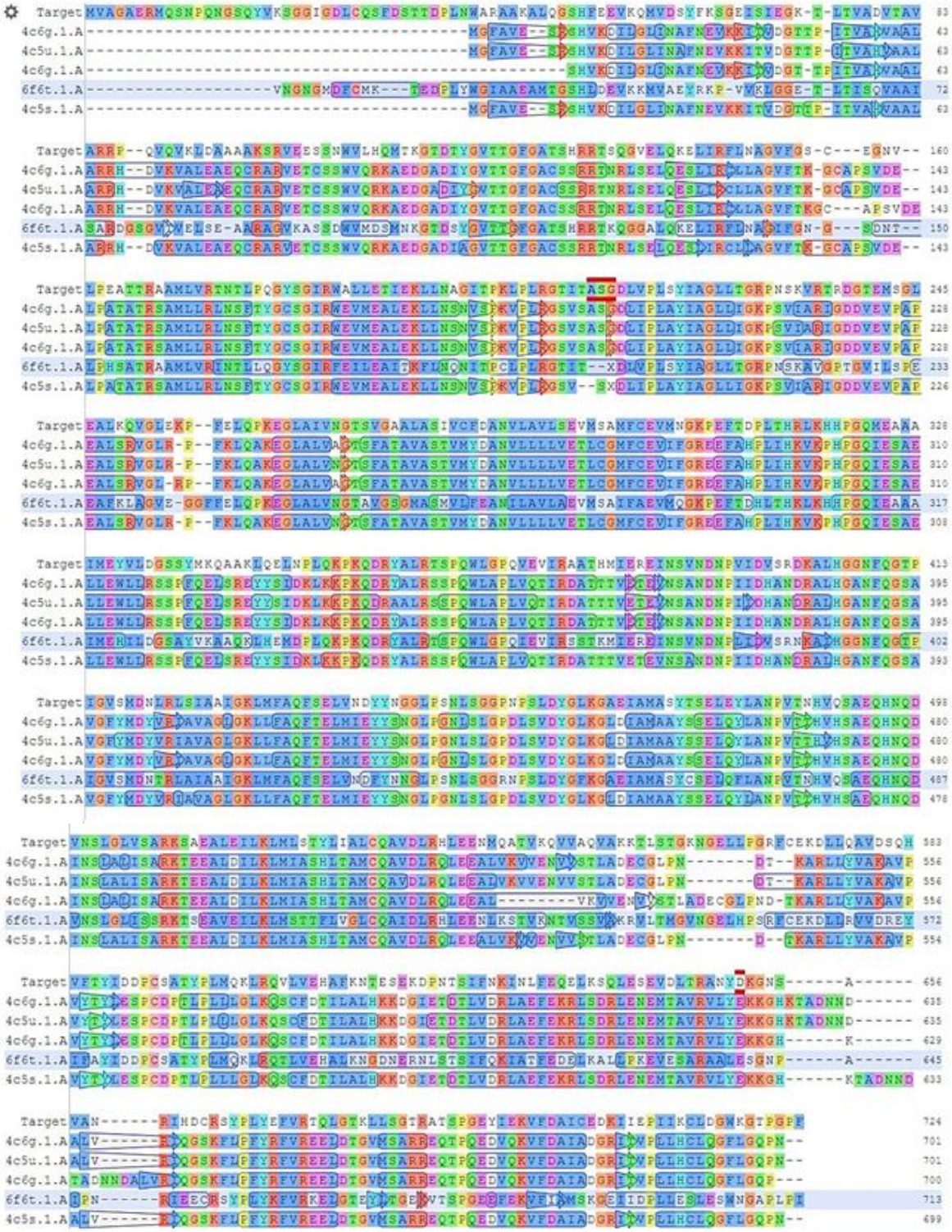


Figure (10): Multiple sequence alignmentof the deduced GbPAL protein withother PALs. Primary and secondary structures of PAL from SMTL-ID: 4c6g.1.A, 4c5u.1.A, 4c6g.1.A, 4c5s.1.A from (*Taxuschinensis*) and 6f6t.1.A from **(***Petroselinumcrispum***).** The conservedactive site motif (Ala-Ser-Gly) can be converted into a MIO prostheticgroup and is indicated by Red under the sequences in location (212-214 aa).

Reference

[1]. Hori T, Tulecke RW, Del W,et al.: In *Ginkgo Biloba* Treasure: From Biology to Medicine, Springer, 1997; pp.444. DOI 10.1007/978-4-431-68416-9

[2]. Sierpina VS, Wollschlaeger B, Blumenthal M.*Ginkgo biloba*. AmericanFamily Physician,2003;68(5): 923-927.www.aafp.org/afp.

[3]. vanBeek TA. Chemical analysis of *Ginkgo biloba* leaves and extracts. J. Chromatogr. A. 2002;967(1): 21-55.doi: 10.1016/s0021-9673(02)00172-3.
[4]. Smith JV, Luo Y. Studies on molecular mechanisms of *Ginkgo biloba* extract. Appl. Microbiol. Biotechnol. 2004;64(4): 465-472.
[5]. Czemmel S, Stracke R, Weisshaar B, et al. The grapevine R2R3-MYB transcription factor VvMYBF1 regulates flavonol synthesis in developing grape berries. Plant Physiology. 2009; 151(3): 1513–1530.
[6]. Hyun MW, Yun YH, Kim JY, Kim SH. Fungal and Plant Phenylalanine Ammonia-lyase. Mycobiology. 2011; 39(4): 257-265.
[7]. Ferrer JL, Austin MB, Stewart C, Noelb JP. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiol.Bioch., 2008; 46(3): 356-370.
[8]. Dixon RA, Paiva NL. Stress-induced phenylpropanoidmetabolism. Plant Cell.1995; 7(7):1085–1097.doi: 10.1105/tpc.7.7.1085
[9]. Ritter H, Schulz GE. Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. Plant Cell. 2004; 16(12):3426-3436.doi: 10.1105/tpc.104.025288. Epub 2004 Nov 17.
[10]. Koukol J, Conn EE. The metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of *Hordeumvulgare*. J. Biol. Chem.,1961; 236(10): 2692–2698.doi.org/10.1016/S0021-9258(19)61721-7
[11]. Huang JL, Gu M, Lai ZB, et al. Functional analysis ofthe *arabidopsis* PAL gene family in plant growth, development, and response to environmental stress.Plant Physiol.,2010; 153(4):1526–1538.DOI: 10.1104/pp.110.157370
[12]. Pang Y, Shen G, Wu W, Liu X, et al. Characterization and expression of chalcone synthase gene from *Ginkgo biloba*. Plant Science. 2005;168(6):1525-1531.https://doi.org/10.1016/j.plantsci.2005.02.003
[13]. Shen G, Pang Y, Wu W, Deng Z, et al. Cloning and characterization of a flavanone 3-hydroxylase gene from *Ginkgo biloba*. Bioscience Reports. 2006a;26(1):19-29.doi: 10.1007/s10540-006-9007-y
[14]. Shen G, Pang Y, Wu W, Liu X,et al. Isolation and characterization of a putative anthocyanindinreductase gene from *Ginkgo biloba*. J. Plant Physiol. 2006b; 163(2): 224-227.doi: 10.1016/j.jplph.2005.06.012
[15]. Xu F, Cheng SY, Cheng SH, Wang Y, Du HW. Time course of expression of chalcone synthase gene in *Ginkgo biloba*. J. Plant Physiol. and Molecular Biology. 2007;33(4): 309-317.
[16]. Chang A, Lim MH, Lee SW, Robb EJ, Nazar RN. Tomato phenylalanine ammonia-lyase gene family, highly redundant but strongly underutilized. J. BiologicalChemistry. 2008; 283(48): 33591- 33601.doi: 10.1074/jbc.M804428200
[17]. Wanner LA, LiG, WareD, SomssichIE, Davis KR.The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*.Plant Mol. Biol., 1995; 27(2):327-338. doi: 10.1007/BF00020187.
[18]. Minami E, Ozeki Y, Matsuoka M, Koizuka N, Tanaka Y. Structure and some characterization of the gene for phenylalanine ammonialyase from rice plants. Eur. J. Biochem., 1989; 185(1):19-25.doi: 10.1111/j.1432-1033.1989.tb15075.x.
[19]. Lee SW, Robb J, Nazar RN. Truncated phenylalanine ammonialyase expression in tomato (*Lycopersiconesculentum*). J. Biol. Chem., 1992; 267(17):11824-11830. doi.org/10.1016/S0021-9258(19)49773-1
[20]. Guo J, Wang M-H. Characterization of the phenylalanine ammonia-lyase gene (SlPAL5) from tomato (*Solanumlycopersicum* L.).Mol. Biol. Rep., 2009; 36(6):1579–1585. DOI 10.1007/s11033-008-9354-9
[21]. Yamada T, Tanaka Y, Sriprasertsak P, Kato H, et al. Phenylalanine ammonia-lyase genes from *Pisumsativum*: structure, organ-specific expression and regulation by fungal elicitor and suppressor. Plant and Cell Physiology.1992; 33(6): 715-725.doi.org/10.1093/oxfordjournals.pcp.a078310
[22]. Osakabe Y, Osakabe K, Kawai S, Katayama Y, Morohoshi N. Characterization of the structure and determination of mRNA levels of the phenylalanine ammonia-lyase gene family from *Populuskitakamiensis*. Plant Mol. Biol.,1995; 28: 1133-1141.
[23]. Song J, Wang Z. Molecular cloning, expression and characterization of a phenylalanine ammonia-lyasegene (SmPAL1) from *Salvia miltiorrhiza*. MolBiol Rep. 2009; 36(5): 939-952. doi.org/10.1007/s11033-008-9266-8
[24]. Okada T, Mikage M, Sekita S. Molecular characterization of the phenylalanine ammonia-lyase from *Ephedra sinica*. Biologicaland Pharmaceutical Bulletin. 2008; 31(12):2194–2199.doi.org/10.1248/bpb.31.2194
[25]. Xu F, Cai R, Cheng SY, Du HW, Wang Y, Cheng SH. Molecular cloning, characterization and expression of phenylalanine ammonia-lyase gene from *Ginkgo biloba*. Afr. J. Biotechnology. 2008; 7(6):721–729.
[26]. Xu H, Park NI, Li X,et al. Molecular cloning and characterization of phenylalanine ammonia-lyase, cinnamate 4-hydroxylase and genes involved in flavone biosynthesis in *Scutellariabaicalensis*. Bioresource Technology.2010; 101(24):9715– 9722. doi.org/10.1016/j.biortech.2010.07.083
[27]. Weitzel C. Petersen M. Enzymes of phenylpropanoid metabolism in the important medicinal plant *Melissa officinalis*L. Planta2010, 232(3):731–742.doi: 10.1007/s00425-010-1206-x.
[28]. Tao J, Cao C, Zhao D, Zhou C. Liang G. Molecular analysis and expression of phenylalanine ammonia-lyase from poinsettia (*Euphorbia pulcherrima*willd). Afr. J Biotech., 2011; 10(2):126–135.
[29]. Jiang YM, Xia N, Li, XD, et al. Molecular cloning and characterization of a phenylalanine ammonia-lyase gene (*LrPAL*) from *Lycorisradiata*. Mol. Biol. Rep.2011; 38:1935–1940.https://doi.org/10.1007/s11033-010-0314-9
[30]. Hou X, Shao F, Ma Y, Lu S. The phenylalanine ammonia-lyase gene family in *Salvia miltiorrhiza*: Genome-wide characterization, molecular cloning and expression analysis. Mol. Biol. Rep.,2013; 40(7):4301–4310.https://doi:10.1007/s11033-013-2517-3
[31]. Zeng S, Liu Y, Zou C, Huang W, Wang Y. Cloning and characterization of phenylalanine ammonia-lyase in medicinal *Epimedium* species. Plant Cell Tissue and Organ Culture (PCTOC),2013;113:257-267. https://rdcu.be/cfWba
[32]. MacDonald MJ,D'Cunha GB. A modern view of phenylalanine ammonia lyase. Biochem.Cell Biol. 2007; 85(3): 273–282.https://doi: 10.1139/o07-018.
[33]. PeitschMC.ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. Biochem. Soc. Trans. 1996; 24:274–279.doi: 10.1042/bst0240274
[34]. Peitsch,M.C. Large scale protein modelling and model repository. Proc. Int. Conf. Intell. Syst. Mol. Biol., 1997; 5, 234-236.
[35]. Waterhouse A, Bertoni M, Bienert S, Studer G, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018; 46(W1): W296-W303.https://doi: 10.1093/nar/gky427.
[36]. Martí-Renom MA, Stuart AC, Fiser A,et al. Comparative protein structure modeling of genes and genomes. AnnualReview BiophysicsBiomolecular Structure. 2000;29(1):291–325. https://DOI: 10.1146/annurev.biophys.29.1.291
[37]. ThiyagarajanK,Vitali F, Tolaini V, Galeffi P, et al. Genomic characterization of phenylalanine ammonia lyase gene in Buckwheat. PLoS ONE2016; 11(3): e151187. https://doi.org/10.1371/journal.pone.0151187
[38]. Roy S, Maheshwari N, Chauhan R, Sen NK, Sharma A. Structure prediction and functional characterization of secondary metabolite proteins of Ocimum. Bionformation. 2011; 6: 315-31914.
[39]. Korbie JD, Mattick JS. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat. Protoc., 2008; 3(9):1452-1456.
[40]. Waterhouse AM, Procter JB, Martin DMA, et al. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009; 25: 1189-1191. Pubmed: 19151095 DOI: doi:10.1093/bioinformatics/btp033

[41]. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res., 2003; 31(13):3381–3385.doi: 10.1093/nar/gkg520.

[42]. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 2016; 33(7):1870-1874.doi: 10.1093/molbev/msw054.

[43]. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 1982; 157(1):105-132.doi: 10.1016/0022-2836(82)90515-0.

[44]. Gasteiger E, Hoogland C, Gattiker A, et al. In: Protein identification and analysis tools on ExPASy server, (In) Johan M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. 2005; pp. 571-607.

[45]. Xu F, Deng G, Cheng S, Zhang W, et al. Molecular cloning, characterization and expression of the phenylalanine ammonia-lyase gene from *JuglansRegia*. Molecules.2012; 17(7):7810–7823.doi: 10.3390/molecules17077810.

[46]. Mahesh V, Rakotomalala JJ, le Gal L. Vigne H, et al. Isolation and genetic mapping of a *Coffeacanephora*phenylalanine ammonia lyase gene (*CcPAL1*) and its involvement in the accumulation of caffeoylquinic acids. Plant Cell Rep. 2006; 25(9): 986–992.doi: 10.1007/s00299-006-0152-3.

[47]. Gao JH, Zhang SW, Cai F, Zheng XJ, et al. Characterization and expression profile of a phenylalanine ammonia lyase gene from *Jatrophacurcas*L. Mol. Biol. Rep. 2012; 39, 3443–3452.

[48]. Jin Q, Yao Y, Cai YP, Lin Y. Molecular cloning and sequence analysis of a phenylalanine ammonia-lyase gene from *Dendrobium*. PLoS ONE.2013; 8(4): e62352.https://doi.org/10.1371/journal.pone.0062352

[49]. Sangsil, P.; Nualsri, C.; Woraathasin, N.; Nakkanong, K. Characterization of the phenylalanine ammonia lyase gene from the rubber tree (*Heveabrasiliensis*Müll. Arg.) and differential response during *Rigidoporusmicroporus*infection. *J. Plant* Prot. Res.2016, 56(4): 380–388.DOI: https://doi.org/10.1515/jppr-2016-0056

[50]. Khakdan F, Alizadeh H, Ranjbar M. Molecular cloning, functional characterization and expression of a drought inducible phenylalanine ammonia-lyase gene (*ObPAL*) from *Ocimumbasilicum*L. Plant Physiol. Biochem. 2018, 130, 464–472.DOI: 10.1016/j.plaphy.2018.07.026

[51]. Schomburg D, Salzmann M. In: Class 4: lyase, phenylalanine ammonia-lyase. In: Enzyme Handbook 1, Schomburg D, Salzmannn M, (eds.), (1990). pp. 1-5, Springer-Verlag, Germany, Berlin.

[52]. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Appl. Biosci. 1995 Dec;11(6):681-684.

[53]. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. Nucleic Acids Res., 2009;37(Database issue):D387-392.

[54]. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: Computed atlas of surface topography of proteins. Nucleic Acids Res. 2018; 2(w1): w363-w367. doi: 10.1093/nar/gky473.

[55]. Appert C, Logemann E, Hahbrock K, et al. Structural and catalytic properties of the phenylalanine ammonia-lyaseisoenzymes from parsley (*Petroslinumcrispum*Nym.). European J. of Biochemstry.1994; 255(1):491-499.

[56]. Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421-430.doi: 10.1186/1471-2105-10-421.

[57]. Steinegger M, Meier M, Mirdita M, et al. HH-suite3 for fast remote homology detection and deep protein annotation. BMC Bioinformatics. 2019; 20(1): 473.doi: 10.1186/s12859-019-3019-7.

[58]. Remmert M, Biegert A, Hauser A, Söding J.HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nature Methods. 2012; 9(2):173-175.doi: 10.1038/nmeth.1818.

[59]. Studer G, Rempfer C, Waterhouse AM, et al. QMEANDisCo - distance constraints applied on model quality estimation. Bioinformatics. 2020; 36(8):1765-1771. doi: 10.1093/bioinformatics/btz828

[60]. Bertoni M, Kiefer F, Biasini M, et al. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Scientific Reports. 2017; 7: 10480. doi: 10.1038/s41598-017-09654-8

[61]. Wybenga GG, Szymanski W, Wu B, Feringa BL, et al. Structural Investigations into the Stereochemistry and Activity of a Phenylalanine-2,3-aminomutase from *Taxuschinensis*. Biochemistry. 2014; 53 (19): 3187-3198. DOI: 10.1021/bi500187a

Mohamed ZoelfakarSayedAhmed.“Functionaland Analysis of Structural Prediction ofthe Phenylalanine Ammonia-lyase Gene(GbPAL) from*Ginkgo biloba* (L.).”*IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)*, 7(2), (2021): pp. 13-30.