IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)
ISSN: 2455-264X, Volume 7, Issue 1 (Jan. – Feb. 2021), PP 01-21
www.iosrjournals.org

Homology Modeling and Structural Analysis of theFlavanone 3-Hydroxylase (F3H) and Flavonoid 3`hydroxylase (F3`H) Genes from *Ginkgo biloba* (L.)

Mohamed ZoelfakarSayedAhmed*
(Genetic Resources Department, Desert Research Center (DRC), 1, Mathaf El-MatariyaStreet, El-Matariya B.O.P 11753 El-Matariya, Cairo, Egypt.)
**corresponding author: (Ahmed, M.Z.S.)*

Abstract:
Ginkgo biloba L. is a well-known living gymnosperm fossil that has medicinal, biologically and economically value in the worldwide. In this study, bioinformatics analysis based on Homology modeling were obtainedsuch as flavanone 3-hydroxylase (GbF3H) and flavonoid 3`-hydroxylase (GbF3`H) from Ginkgo biloba L. are key enzymes in involved pathway of plant flavonoids and the anthocyanin.The full-length cDNA of F3H *gene sequence (*GbF3H*) was isolated from G. bilobacontained a 1074 bp open reading frame(ORF) encoding a 357-amino-acid protein. The deduced GbF3H protein showed high identities to other plant F3Hs.For the full-length cDNA GbF3`H gene was contained a1674 bpopen reading frame (ORF)encoding a 556 amino acid protein.As well as, Multiple Sequence Alignment (MSAs) of selected of 30 amino acid were done using MEGA7 software with high identify and similarity.Phylogenetic analysis was performed using the amino acid sequence of GbF3H and GbF3'H with other known plant-specific F3Hs and F3'Hs to each Gymnosperm (Naked seed) and Angiosperms (Flower plant).The parameters computed by ProtParam software was obtained to the molecular weight and grand average of hydropathicity (GRAVY) for GbF3H and GbF3`H protein.The parameters computed by Protscale software was obtained to hydrophilicity scales based on different chemical and physical properties of the amino acids.We have investigated homology modelling and structure analysis to characterize of two genes (GbF3H and GbF3`H) from Ginkgo biloba to identity percentage using the SWISS-MODEL template library (SMTL). The results indicated thatmolecular characterization and bioinformatics analysis of several genes encoding key enzymes is the first step to fully understanding the regulatory mechanisms controlling flavonoid and anthocyanin biosynthesis in Ginkgo biloba. The predicted of* secondary *structure protein sequence using by SOPMA software can be used as method for further studies to understand the role of each fold of this protein in the function. Also the 3D predicted of protein sequence structure complexes can be used as a basal structure for performing site direction mutation to improve the transformation efficiency of this protein for developing new recombinant bacterial and plant species.*

Key word:*Flavanone 3-hydroxylase (GbF3H),flavonoid 3'hydroxylase (GbF3`H), Ginkgo biloba, homology modeling, Model-Template Alignment*

Date of Submission: 06-02-2021 Date of Acceptance: 20-02-2021

I. Introduction

Ginkgo biloba L. (ginkgo-maidenhair tree) is an ancient relic plant. As it is the only extant species in the division Ginkgophyta and belongs to family Ginkgoaceae, it is considered to be a "living fossil"of the plant kingdom, and has been well studied owing to the many active ingredients, such as flavonoids, contained in its leaves [1; 2].*Ginkgo biloba* L. (Gymnosperm/Naked seed) is a monotypic species native to China with great economic and medicine values. *G. biloba* is dioecious,that is, the male and female structures exist on separate trees, but the two structures can only be distinguished after the tree is around 30 years old. Leaves extract of this tree contains about 24% flavonoids, which are widely used with a plenty of pharmacological properties, the potential toxicological effects of biflavonoids remains largely unknown [3].However, the flavonoids biosynthesis and anthocyanin pathway are poorly understood in Ginkgo[4].Additionally, it is an important medicinal and biologically tree, because its leaves contain flavonoids and terpene lactones with useful pharmacological activities, strenuous efforts have been exerted to sequence its genome [5]. Flavonoids are the low molecular weight polyphenolic secondary metabolic compounds, universally distributed in green plant kingdom, located in cell vacuoles and important functions in the plant's adaptations to specific ecological niches or its responses to biotic and abiotic stresses [6].They have very high application value in medical and health care [7].They are classified into different subgroups, mainly including flavones, flavonols, flavanones,

flavanols, isoflavones, aurones, anthocyanins, and proanthocyanidins (PA, also called condensed tannins) [8].At molecular cloning level, twenty six flavonoid biosynthesis-related gene candidates were identified, of which twenty are novel. They belong to nine families potentially encoding chalcone synthase (CHS), chalconeisomerase (CHI), flavone synthase (FNS), flavanone 3-hydroxylase (F3H), flavonoid 3`-hydroxylase (F3`H), flavonoid 3`,5` hydroxylase (F3`5`H), flavonol synthase (FLS), dihydroflavonol 4-reductase (DFR), and anthocyanidin synthase (ANS), respectively.Flavonoids play a variety of physiological roles in plant growth, development, and reproduction. Flavonoids are one of the largest groups of plant secondary metabolites with a C6-C3-C6 general structural backbone [9].Both F3`H and F3'5'H (cytochromes P450) are key enzymes in the flavonoid pathway leading to the production of the colouredanthocyanins. The hydroxylation pattern is determined by two cytochromes P450, flavonoid 3'-hydroxylase (F3'H) and flavonoid 3', 5'-hydroxylase (F3'5'H) and thus they play a crucial role in the determination of flower colour. F3'H and F3'5'H mostly belong to CYP75B and CYP75A, respectively, but [10]except for the F3'5'Hs in compositae that were derived from gene duplication of CYP75B and neofunctionalization [9].In the flavonoid biosynthesis pathway, F3′H is an important enzymefor controlling the hydroxylation of naringenin, dihydrokaempferol,kaempferol, and apigenin at the 3′ position in the B-ring to generateeriodictyol, dihydroquercetin, quercetin, and luteolin, respectively(Figure 1), which are important intermediates for anthocyanin and proanthocyanidin biosynthesis in *Brassica napus* [11].There are several enzymes such as flavanone 3-hydroxylase (GbF3H) and flavonoid 3`-hydroxylase (GbF3'H) are key enzymes in involved pathway of plant flavonoids and the anthocyanin. Flavanone 3-hydroxylase (E.C. 1.14.11.9) is one of the 'core' enzymes acting at the bifurcation of the anthocyanin and flavonols branches. Flavanone-3-hydroxylase is a key enzyme acting at the flavanone branch point and is the first in the flavonol pathway, converting the flavanones (2S)-naringenin and (2S)-eriodictyol to (2R,3R)-dihydrokaempferol and (2R,3R)-dihydroquercetin, respectively [12; 13].While, flavonoid 3'-monooxygenase (EC 1.14.14.82) with alternative name flavonoid 3′-hydroxylase (Formerly F3′H: EC 1.14.13.21) or CYP75B1 is an important enzyme which determines the hydroxylation pattern of anthocyanins. The majority of several genes encoding F3Hs from many different plants species have been cloned and characterized at the chemical, genetically and enzymological levels; such as from *Arabidopsis thaliana*[14], *Medicago sativa*[15],*Ginkgo biloba* L.[16],*Medicagotruncatula*[17],*Reaumuriasoongorica*[18],*Reaumuriatrigyna* [19],*Phyllanthusemblica*[20],*Lyciumchinense*[21] and *Artiemisaannua*L.[22]and subsequently characterized.While, several studies for cloning and characterization of F3`Hs gene have been reported in PhF3`H from *Perillafrutescens* [23], AtF3`H from *Arabidopsis thaliana* [24],MdF3`H from *Malusˆ domestica* [25],IbF3`H from *Ipomoea batatas*[26],GBF3`H from *G. biloba*[27] and other plants. Despite all of the available sequence information we have, 3D structure and structure-function studies of several plants in Protein Data Base (PDB) which are 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. This subject will help in elucidating the mechanisms of protein function, since function is determined by 3D structure[28].

In this study, our major objectives were to investigated homology modelling and structure analysis predictionsto identity enzymatic activities of two cDNA genes (GbF3H and GbF3`H) from *Ginkgo biloba* L. including molecular characterization, multiple sequencealignment (MSA), phylogenetic analysis and Homology Modeling and performed some several necessary bioinformatics analysis to helped increase fully understanding and molecular mechanisms and deduce its regulatory role in flavonoid and anthocyanin biosynthesis.

II. Material and method

Plant materials:

Several young leaves collectedas from male tree type from*Ginkgo biloba* L. (plant exchange from Frankfort, Germany) were obtained from International Park in Nasr City, Cairo, Egypt. One hundred milligram of collected frozen tissue two 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 (Qiagen,RNasy mini plant 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 F3H and F3`H 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 5-flavanone 3-hydroxyrase (F3H) gene (Gb_Fne_Fwd: 5`-ATG GCT CCT GTG CAG AGC GTC-3` with Gb_Fne_Rev1: CTA TTT GGA CTC GTC TTG TTG AAG and for partial length (Gb_Fne_Fwd with Gb_Fne_Rev2: 5`-AAC CTG GAA AAT GCC CCA TTC C-3`) according to accession no. AAU93347.To obtain the full length of 4-flavonoid 3` hydroxylase-like protein (F3'H) gene (Gb_Fid_Fwd: 5`-ATG CAC TTG TTT TTG CCA CCA C-3` with Gb_ Fid_Rev1: 5`-CTA GCA ATA CAA ATG AGG GGG-3`) and for partial length (Gb_Fid_Fwd with Gb_Fid_Rev2: 5`-CTC AGG CCT AGT CTT AGG GAC-3`) according to accession no. AJO67233.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 4 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, 29.5 µl DEPC H_2O and spin for 15 Sec. Touchdown PCR program was used to amplification for F3H and F3`H cDNA 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 62 - 56ºC was decreased (▼2ºC/Cycles) and (25 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 [29].A volume of 40 µl of each sample were analyzed using 1.2% agarose gel electrophoreses with DNA ladder sizein 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 F3H and F3`H Gene from *Ginkgo biloba*:

The putative *Ginkgo biloba* for flavanone 3-hydroxyrase (F3H) gene cDNA and flavonoid 3' hydroxylase-like protein(F3`H) were analyzed by bioinformatics software. Search for GbF3Hs and GbF3`Hs -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). GbF3H protein sequence from *Ginkgo biloba* with accession no. AAU93347.1 was used for BLASTp and homology searches against other plants species. GbF3`H protein sequence from *Ginkgo biloba* with accession no. AJO67233.1 was used for BLASTp and homology searches against other plants species. Multiple Sequence Alignments (MSA) and JalViewprogram [30] with total 30 protein sequences, 15 from the putative amino-acid sequence of F3Hs and 15 from F3`Hs were to compare and performed using Clustal Omega software online(https://www.ebi.ac.uk/Tools/msa/clustalo/), Phylogenetic tree analysis involved 30 amino acid sequences of GbF3Hs and GbF3`Hs gene with other plant species were conducted in MEGA 7.0 software program by Maximum Likelihood method [31].

Primary and secondary structural prediction: In this study, secondary structure of GbF3H and GbF3`H protein from *Ginkgo biloba* were analyzed using online server (http://www.expasy.org/tools/protparam.html) based on the gene sequence and secondary protein structure of this two proteins were predicted and analyzed using in online server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?Page=/NPSA/npsa_sopma.html).

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 F3H and F3`H protein Three-Dimensional (3D) structural were predicted using Swiss-Model online server (https://swissmodel.expasy.org/interactive). The three-stage structure model is shown by Homology Modeling with SWISS-MODEL [32].

Hydrophilicity prediction analysis: Protscale software online (https://web.expasy.org/protscale/) was used to represent the profile produced by any amino acid scale on a selected protein [33].ProtParam software online (https://web.expasy.org/protparam/) were used as a tool to allow the computation of various physical and chemical parameters for query user entered target protein and a given protein stored in Swiss-Prot[34] or TrEMBL or for a sequence.

III. Results and discussion

Good total RNA quality (i.e., A260/A230 and A260/A280 absorbance ratios within the range 1.9 - 2.5) and extraction yield (25-40 ng/µl) were obtained from all the leaves material samples using nanodrop.The two candidate genes exhibited high PCR success and the obtained PCR products of full length were successfully sequenced with high-quality bidirectional sequences. The results showed that cDNA of *Ginkgo biloba* (GbF3H

gene), contain an open reading frame of 1074 bp open reading frame (ORF) encoding a 357-amino-acid protein with a calculated molecular weight of about 40 kDa and isoelectric point (*pI*) of 5.57. For the full-length cDNAGbF3'H gene was 1671bpopen reading frame (ORF)encoding a 556 amino acid proteinwith a predicted molecular weight of about 63.47 kDa and isoelectric point (*pI*) of 7.71. The obtained PCR products of partial lengthcDNAGbF3H and GbF3'H genewere240bp and 203bp, respectively.Both fragments represent the full length and partialcDNA of the two genes (F3H and F3`H) as shown in (Figure 2).The result showed that alignment of sequence data the cDNA F3H and F3`H genes from *Ginkgo biloba* (gymnosperm species) with accession no. AAU93347.1 and accession no. AJO67233.1 about ~100% maximum identity within the same plant and less than 100% with other angiosperms species. Our result agreement with Shen*et al*. [16] wereisolated the full-length cDNA and compared with genomic DNA sequences of GbF3H gene from *G. biloba.* The conserved amino acids were found in GbF3H at the similar positions like other F3Hs. Meanwhile, Jin *et al*.[35]identified of flavonoid gene (SmF3H) from *Saussurea medusa* as traditional Chinese medicinal plantwhich contained 1032 bp open reading frame (ORF) encoding a protein of 343 amino acid residues and its similarities to metabolic enzymes from other plants.Also, our result agreement with Li *et al*.[27],they suggested that the GbF3'H gene has close relation with the formation of flavonoids and anthocyanin biosynthesisin *Ginkgo biloba*and the highest expression in stamens the next in mature leaves and gynoecium which were great difference with other flavones biosynthetic pathway gene. Also, Zhou *et al*. [26]isolated and characterized the full-length cDNA and genomic DNA of F3'H from the purple-fleshed sweet potato (*Ipomoea batatas*). IbF3'H was 1,789 bpcontaining a 1,554 bp open reading frame (ORF) encoding 518 amino acids. Comparative and bioinformatics analysis revealed that IbF3'H was highly homologous with F3'Hs from other plant species.

Analysis of flavanone 3-hydroxyrase (GbF3H)from*Ginkgo biloba*:

For search F3H cDNA amino acid sequence of *Pinustaeda*(QBI90549.1), *Pinustaeda*(AGY80772.1),*Pinus radiate*(QBI90547.1), *Piceaabies*(QBI90546.1), *Ipomoea batatas* (ACT31918.1), *Camptotheca acuminate* (ARO92271.1),*Vitisvinifera*(RVW42566.1), *Verniciafordii*(ARV78456.1), *Nicotianatabacum* (NP_001312012.1), *Curcuma alismatifolia*(QPZ56413.1), *Actinidiarufa* (GFY93965.1), *Gossypiumhirsutum* (ABM64799.1), *Rosa rugos*a (AKT71853.1) and *Strelitziareginae* (AGC74052.1) were downloaded from genbank database. These sequences were stored in a FASTA file including F3H cDNA sequence of *Ginkgo biloba* (AAU93347.1).As well as, Multi Sequence Alignment (MSA) of thededuced polypeptide sequence of GbF3H and other selected F3`Hsfrom several plant species were carried out. It was found that GbF3Hpresented 79.33, 78.69, 75.54, 74.30, 71.56, 70.64, 70.20, and 70.43% identity with E-value= zero to *Pinustaeda, Pinustaeda*, *Pinus radiate, Piceaabies, Ipomoea batatas, Camptotheca acuminate, Vitisvinifera*, and*Verniciafordii*, respectively as 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 (-3319.73) is shown in(Figure3). The phylogenetic tree analysis showed two branch, the first branch contain *Ginkgo biloba* (AAU93347.1), *Pinustaeda*(QBI90549.1), *Pinustaeda*(AGY80772.1),*Pinus radiate*(QBI90547.1) and *Piceaabies*(QBI90546.1) and other branch. The results revealed that *Ginkgo biloba* (GbF3H) cDNA in this investigation was closelyto*Pinustaeda, Pinustaeda, Pinus radiate*and *Piceaabies*. The genetic relationship between the F3H cDNA is consistent with the phylogenetic tree.

Analysis of flavonoid 3` hydroxylase-like protein (F3`H) from *Ginkgo biloba*:

For search F3`H cDNA amino acid sequence of*Amborellatrichopoda, Lupinusalbus,Lupinusangustifolius, Ipomoea triloba, Ipomoea nil,Nymphaescolorata, Glycine max,Panicumhallii, Nicotianatabacum, Solanumlycopersicum,Chenopodium quinoa,Ricinuscommunis* were downloaded from genbank database. These sequences were stored in a FASTA file including F`3H cDNA sequence of *Ginkgo biloba* (AJO67233.1).As well as, Multi Sequence Alignment (MSA) of thededuced polypeptide sequence of GbF3`H and other selected F3`Hsfrom several plant species were carried out. It was found that GbF3`Hpresented to 75.44, 75.415, 74.09, 73.06, 73.62 and 72.82% identity with E-value= zero to*Amborellatrichopoda,Lupinusalbus,Lupinusangustifolius, Ipomoea triloba, Ipomoea nil, Nymphaescolorata, Glycine max* and*Panicumhallii*, respectively as shown in Table 2.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 (-5359.61) as shown in (Figure 4).The phylogenetic tree analysis showed two branch, the first branch contain *Ginkgo biloba* (AJO67233.1), *Amborellatrichopoda* (XP_006836296.1),*Nymphaescolorata* (XP_031500353.1) and other branch. The results revealed that *Ginkgo biloba* (GbF3`H) cDNA in this investigation was closelyto*Amborellatrichopoda,Nymphaescolorata*. The genetic relationship between the F3`H cDNA is consistent with the phylogenetic tree.

The phylogenetic tree analysis was performed using the amino acid sequence of GbF3H and GbF3′H from *Ginkgo biloba*with other known plant-specific F3′Hs and F3′5′Hs. Based on the phylogenetic tree, F3H

and F3′H were separated into two clades (CYP75A and Cytochrome P450,respectively), which were highly supported with 100% bootstrap values [36; 37].GbF3H was grouped into the F3H clade, suggesting that the GbF3Hgene belongs to the F3H family and GbF3`H gene was grouped into other clade F3`H family. 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 (-6772.41) is shown(Figure 5).Phylogenetic tree analysis revealed that GbF3H and GbF3`H from *Ginkgo biloba*(Gymnosperm species) were shared the same ancestor in evolution with other F3Hs and F3Hs 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 GbF3H and GbF3`H gene had considerable high homology withother plant F3Hs and F3`Hs gene families.

The high similarity among flavanone 3-hydroxylase (GbF3H) proteins from *G. biloba* was observed from residues 25 to 350, with variability in length and composition being found in the N-terminal and C-terminal regions(Figure 6). According to Britsch*et al*. [12],five similar motifs were foundMotif-1 (A71CE/SEWGIFQVVD/HHGV85), Motif-2 (W154PQ/V156), Motif-3 (Y210PKCP214), Motif-4 (H225TDPGTITLLLQDQVGGLQA244), Motif-5 (H281QAVVNSNSSRLSITF298)**,** among which the motif 3 and motif 4 of GbF3H were the same number of motif with the other species, but there wereseveral amino acid differences in 1, 2 and 5 motifs among the F3Hs. The reason for thisdifference might be explained that *G. biloba* wasgymnosperm species, from which few F3Hgenes had been cloned previously. The conserved amino acids ligating ferrous iron and residues participating in 2-oxoglutarate binding (R-X-S) were found in GbF3H at the similar like other F3Hs. Three prolines were strictly conserved in motif-2 and motif-3, which were predicted to have important roles in the folding process of the polypeptide.It was also worth noting that the amino acid residues His (H) 258, Asp (D) 260 and His (H)315 for ligating ferrous iron, and Arg (R)326 and Ser (s) 328 participating in the 2-oxoglutarate binding (RXS motif) were the same conserved at the other positions among F3Hs [12; 38]. All the observed conservation of these amino acids in all the aligned sequences especially in *G. biloba*, suggested the existence of function of GbF3H protein.

The high similarity among flavonoid 3'hydroxylase(GbF3`H) proteins from *G. biloba* was observed from residues 68 – 556 amino acid, with variability in length and composition being found in the N-terminal and C-terminal regions(Figure 7).The deduced amino acid sequence of GbF3′H contained the proline-rich "hinge" region domain (Motif (P70PGPKF/GWP) may act as a hinge motif necessary for the optimal orientation of the P450 enzyme [26].The motif (A/G) GX (D/E) T (T/S) forms a binding pocket for oxygen molecules required for catalytic activity [39], and the EXXR motif (E405TFR408) stabilizes the core structure[40].The P450 consensus contained with heme-binding domain (F485xxGxRxCxG494) or (F485(G/S)AG(R/K)RIC(A/P)G494), which is responsible for carbon monoxide-binding ability [40],was found conserved at the other positions among GbF3′H. Also, the binding pocket motif for oxygen molecules was found (A349TDTS353).Zhou *et al*.[26] revealed that conserved domain IbF3'H was a cytochrome P450 dependent enzyme. Phylogenetic analysis revealed that IbF3'H was clustered into the same subgroup with the homologues from *Ipomoea purpurea*, *Ipomoea tricolor* and *Ipomoea nil*. In order to better understand the deduced GbF3H and GbF3`H protein, a comparative modeling of 3D model of F3Hs and F3`Hs was performed at *ExPASy* using SWISS-MODEL [30; 41].

Bioinformatics analysis of F3H and F3`H genes were important which involved in the biosynthetic pathways of flavonoids from *Ginkgo biloba* L. based on Homology modeling. The secondary structure of the GbF3H protein (357aa) was predicted by the SOPMA tool. The results indicated that GbF3H consists mainly of α-helices (Hh) (121 is 33.89%) and random coils (Cc) (151 is 42.30%) as well as a few extended strands (Ee) (64 is 17.93%) and beta turns (Tt) (21 is 5.88%) as shown in figure (8).Also, the secondary structure of the GbF3`H protein (556aa) was predicted. The results indicated that GbF3`H consists mainly of α-helices (Hh) (251 is 45.14%) and random coils (Cc) (199 is 35.79%) as well as a few extended strands (Ee) (78 is 14.03%) and beta turns (Tt) (28 is 5.04%) as shown in figure (9).Three-dimensional structure modeling showed that GbF3H had a jerry roll in the enzyme core consisted of β-sheet, a typical structure shared by all 2-oxoglutarate-dependent dioxygenases including F3Hs (Figure 8&9).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 theamino acids [33].Hydrophilicity prediction of GbF3H Proteinfrom *Ginkgo biloba*by using the software of Computer pI/Mw Tool at (https://web.expasy.org/compute_pi/) the deduced for GbF3H protein had atheoretical *p*I at 5.57 and a calculated molecular weight ofabout 63.47 kDaaccording to Gasteiger*et al*. [34].Hydrophilicity of *Ginkgo biloba* GbF3H protein was predicted with 357 amino acid utilizing program of ProtScale according toKyte and Doolittle, [33].The results showed that most sites of *Ginkgo biloba* GbF3H protein with score: 2.022 to -2.700 in the hydrophilic region as showed in (Figure 10). It was concluded that the GbF3H protein is a hydrophilic protein. The parameters computed by ProtParam software online

(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 [9].The parameters computed of GbF3H Protein from *Ginkgo biloba* by ProtParam were obtained to the molecular weight, theoretical *pI*, amino acid composition: (24 (A) alanine-Ala6.7%,18 (R) arginine-Arg 5.0%, 11 (N) asparagine-Asn 3.1%, 19 (D) aspartic acid-Asp 5.3%, 5 (C) cysteine-Cys 1.4%, 20(Q) glutamine-Gln 5.6%, 32 (E) glutamic acid-Glu 9.0%, 21 (G) glycine-Gly 5.9%,10 (H) histidine-His 2.8%, 14 (I) isoleucine-Ile 3.9%, 33 (L) leucine-Leu 9.2%, 23 (K) lysine-Lys 6.4%,11 (M) methionine-Met 3.1%,14 (F) phenylalanine-Phe 3.9%,20 (P) proline-Pro 5.6%, 24 (S) serine-Ser 6.7%, 13 (T) threonine-Thr 3.6%,5 (W) tryptophan-Trp 1.4%,10 (Y) tyrosine-Tyr 2.8%, 30 (V) valine-Val 8.4%, zero (O) pyrrolysine-Pyl 0.0%, andzero (U) selenocysteine-Sec 0.0%), Atomic composition: C: 1797, H: 2820, N: 490, O: 538, S: 16, Extinction coefficient: 42650, Estimated half-life: 30 hours (mammalian reticulocytes, in vitro), The instability index (II) is computed to be 56.88, Aliphatic index: 82.44 and grand average of hydropathicity (GRAVY): -0.408 asaccording to Gasteiger*et al*., [34].

The findings suggest that the F3H protein under study werehydrophobic in nature due to presence of high non-polar residues content. GbF3H protein has high percentage of alanine (6.7%), leucine (9.2%) and serine (6.7%). Results also showed that the maximum number of amino acid present in the sequence was found to be leucine (9.2%) and the least was for cysteine (1.4%) andtryptophan (1.4%). Total number of negatively charged residues (Asp + Glu): 51 and total number of positively charged residues (Arg + Lys): 41.

Hydrophilicity prediction of *Ginkgo biloba* (GbF3`H) Proteinby using the software of Computer *pI*/Mw Tool athttp://www.expasy.org/, the deduced for GbF3`H protein had atheoretical *p*I at 7.71 and a calculated molecular weight ofabout 63.47 kDaaccording to Gasteiger*et al*., [34]. On other hand, Hydrophilicity of *Ginkgo biloba* GbF3`H protein was predicted with 556 amino acid utilizing program of ProtScale. The results showed that most sites of *Ginkgo biloba* F3`H protein with score: 3.189 to -3.100 in the hydrophilic region as showed in (Figure 11). It was concluded that the *Ginkgo biloba* GbF3`H protein is a hydrophilic protein. The parameters computed GbF3H Protein from *Ginkgo biloba*by ProtParam were obtained include the molecular weight, theoretical pI, Amino acid composition: (40 (A) alanine-Ala 7.2%, 42 (R) arginine-Arg 7.6%, 16 (N) asparagine-Asn 2.9%, 34 (D) aspartic acid-Asp 6.1%, 10 (C) cysteine-Cys 1.8%, 16 (Q) glutamine-Gln 2.9%, 35 (E) glutamic acid-Glu 6.3%, 32 (G) glycine-Gly 5.8%, 17 (H) histidine-His 3.1%, 33 (I) isoleucine-Ile 5.9%,59 (L) leucine-Leu 10.6%, 28 (K) lysine-Lys 5.0%, 22 (M) methionine-Met 4.0%, 25 (F) phenylalanine-Phe4.5%, 37 (P) proline-Pro 6.7%, 25 (S) serine-Ser 4.5%, 28 (T) threonine-Thr5.0%, 9 (W) tryptophan-Trp 1.6%, 15 (Y) tyrosine-Tyr 2.7%, 33 (V) valine-Val 5.9%, zero (O) pyrrolysine-Pyl 0.0%, and zero (U) selenocysteine-Sec 0.0%),Atomic composition: C: 2849, H: 4483, N: 785, O: 795, S: 32, Extinction coefficient: 8944, Estimated half-life: 30 hours (mammalian reticulocytes, in vitro), The instability index (II) is computed to be 43.27, Aliphatic index: 88.94 and grand average of hydropathicity (GRAVY): -0.225 as according to Gasteiger*et al*. [34].The findings suggest that the F3H protein under study werehydrophobic in nature due to presence of high non-polar residues content. GbF3H protein has high percentage of leucine (Leu) 10.6%, arginine (Arg) 7.6%, alanine (Ala) 7.2%, and glutamic (Glu) 6.3%. Results also showed that the maximum number of amino acid present in the sequence was found to be leucine (Leu) 10.6% and the least was for tryptophan (Trp)1.6% and cysteine (Cys) 1.8%. Total number of negatively charged residues (Asp + Glu): 69, Total number of positively charged residues (Arg + Lys): 70.

Advanced Structure of *Ginkgo biloba* (GbF3H and GbF3`H) Protein:

The structure prediction from primary to advanced structure is an important task in the field of protein research. The three-dimensional structure model of GbF3H and GbF3`H protein from *Ginkgo biloba* were predicted by the Swiss-Model server, by homology modeling based on the available structures [32].Several different databases provided functional analysis of proteins by classification of protein families and predicting domains and important sites. Template search in either FASTA or Clustal format with the highest quality for model building have then been selected from BLAST [42]and HHBlits database [43] 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 [44].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 [46]. For Model Quality Estimation: The global and per-residue model quality has been assessed using the QMEAN scoring function [46]. 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 [47]. 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 (GbF3H) protein from *Ginkgo biloba* with more than 50 templates available in databases by named using the PDB ID format such as:SMTL-ID: 1gp4.1A (for anthocyanidin synthase from *Arabidopsis thaliana*complexed with trans-dihydroquercetin) with biounitoligomeric state: monomer, QMEAN: -2.60, GMQE: 0.63, sequence identify: 33.02%; sequence similarity: 0.38% ; SMTL-ID: 2brt.1 (for anthocyanidin synthase from *Arabidopsis thaliana* with Naringenin) with biounitoligomeric state: monomer, QMEAN: -2.60, sequence identify: 32.61%, sequence similarity: 0.37% and other species with high homology and three-dimensional structure were one model built successfully as template alignment (Table 3). Because the C terminal of *Ginkgo biloba* F3H protein is poor homology to *Arabidopsis thaliana* less than 37 amino acids, the template of matching 20 - 339 to *Arabidopsis thaliana* F3H was selected for homology modeling (Fig. 12). The results were close to the protease real space conformation. On other hand, the homologous sequence of GbF3`H protein from *Ginkgo biloba* with more than 50 templates available in databases by named using the PDB-ID format such as:SMTL-ID: 5ylw.1A (for Ferruginol synthase from *Salvia miltiorrhiza*complexed with(CYP76AH1) with Biounitoligomeric state: monomer, none ligands, QMEAN: -1.81, GMQE: 0.56, sequence identify: 31.07% and sequence similarity: 0.37-0.38%. Also with SMTL-ID: 6vby.1.A (for Cinnamic acid 4-hydroxylase or C4H1 (Cytochrome P450-73A33) from *Sorghum bicolor* with Biounitoligomeric state: monomer, none ligands, QMEAN: -1.81, GMQE: 0.59, sequence identify: 29.62%; sequence similarity: 0.36% and other species with high homology and three-dimensional structure were one model built successfully as template alignment (Table 4). Because the C terminal of *Ginkgo biloba* F3`H protein is poor homology to *Salvia miltiorrhiza*less than 103 amino acids, the template of matching 68 - 556 to *Salvia miltiorrhiza*F3`H was selected for homology modeling (Figure 13). 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 GbF3H and GbF3`H 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 GbF3H and GbF3`H 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 [30; 41].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 F3H and F3`H genes were important which involved in the biosynthetic pathways of flavonoids from *Ginkgo biloba* L. Phylogenetic tree analysis revealed that GbF3H shared the same ancestor in evolution with other F3Hs and had a further relationship with other angiosperms species.Bioinformatics analysis show that GbF3'H have a signal recognition peptide and belong to a microsomal cytochrome P450-dependent monooxygenasesmultigene families. This result agreement to Li *et al*.[27]were suggested that the expression model of GbF3'H gene has close relation with the formation of flavonoids and anthocyanin biosynthesiswhich indicated that the F3'H regulate the flavonoids transferring to anthocyanin in *G. biloba*. They showed that GbF3'H have a signal recognition peptide and belong to a microsomal cytochrome P450-dependent monooxygenasesmultigene families using bioinformatics analysis.

IV. Conclusion

We have isolated and sequenced cDNAGbF3H and GbF3′H genesfrom *Ginkgo biloba* L.in this study. As well as, Multiple Sequence Alignment (MSA) involved 30 amino acid sequences of F3Hs and F3′Hs geneswere done with each gene families with high identify and similarity. Phylogenetic analysis was performed using the amino acid sequence of GbF3H and GbF3′H with other known plant-specific F3Hs and F3′Hs. We have investigated homology modelling and structure analysis to characterize enzymatic activities of two genes (GbF3H and GbF3`H) from *Ginkgo biloba*. The GbF3H and GbF3`H theoretical 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 modeling andstructure analysis predictions 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

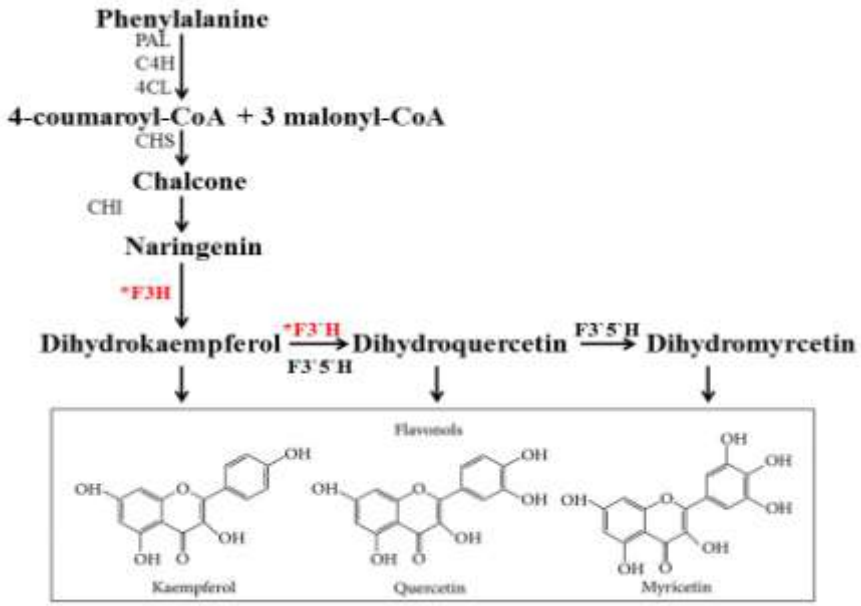

Figure (1): Flavonol biosynthesis in plants (redrawn from Czemmel*et al*.,[48]). The red letter and black box indicate the enzyme and compound analyzed in this study. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaroyl CoA ligase; CHS, chalcone synthase; CHI, chalconeisomerase; *F3H, flavone 3-hydroxylase; *F3`H, flavonoid 3`-hydroxylase; F3`5`H, flavonoid 3`5`-hydroxylase; FLS, flavonol synthase.

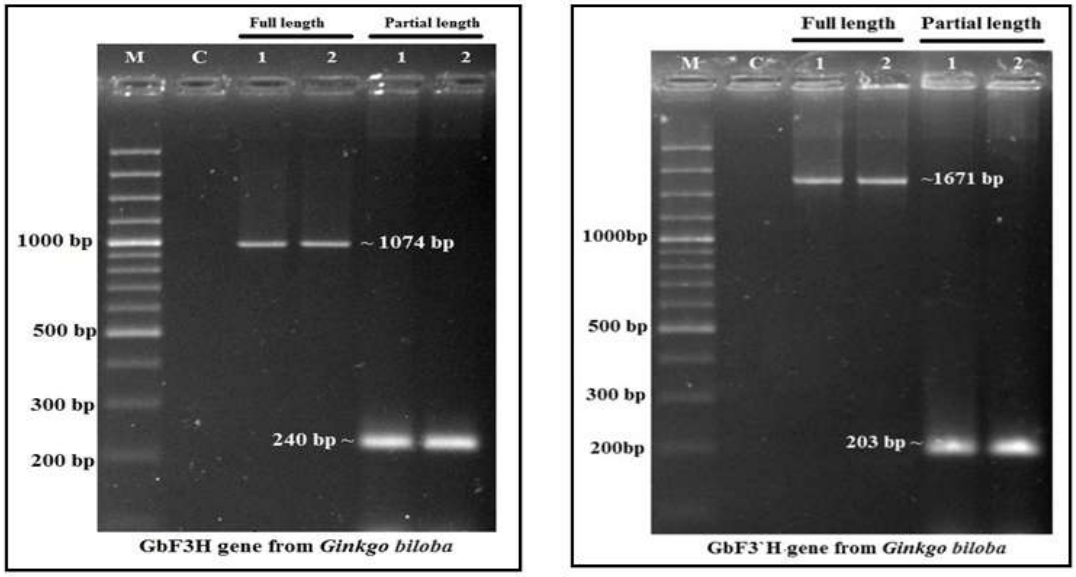

Figure 2: RT-PCR product of full length and partial cDNA using specific primer pair to amplify GbF3H and GbF3`H ORF. 1 and 2 leave samplesfrom*Ginkgo biloba*,M: DNA size marker(100bp DNA Ladder), C: Negative control.

Table (1): Homology of amino acid sequences for 15 selected accession lists and its related *Ginkgobiloba*for flavanone 3-hydroxyrase (GbF3H) sequenced in this study, BLAST top hits against GenBank protein database, similarity score, accession length and accession no.

	Scientific Name	Accession	Per. Identify	Max Score	Total Score	Query Cover	Acc. Length
1	*Ginkgo biloba*	AAU93347.1	~100.00	745	745	100	357
2	*Pinustaeda*	QBI90549.1	79.33	590	590	99	363
3	*Pinustaeda*	AGY80772.1	78.69	580	580	97	363
4	*Pinusradiata*	QBI90547.1	75.54	574	574	99	365
5	*Piceaabies*	QBI90546.1	74.30	557	557	99	359
6	*Ipomoea batatas*	ACT31918.1	71.56	515	515	92	368
7	*Camptothecaacuminata*	ARO92271.1	70.64	520	520	95	368
8	*Vitisvinifera*	RVW42566.1	70.20	517	517	96	394
9	*Verniciafordii*	ARV78456.1	70.43	517	517	95	364
10	*Nicotianatabacum*	NP_001312012.1	69.74	519	519	95	369
11	*Curcuma alismatifolia*	QPZ56413.1	69.60	516	516	96	376
12	*Actinidiarufa*	GFY93965.1	69.57	517	517	96	363
13	*Gossypiumhirsutum*	ABM64799.1	69.57	516	516	95	368
14	*Rosa rugosa*	AKT71853.1	69.03	515	515	97	364
15	*Strelitziareginae*	AGC74052.1	68.18	515	515	96	373

Table (2): Homology of amino acid sequences for 15 selected accession lists and its related *Ginkgo biloba* for flavonoid 3' hydroxylase-like protein (GbF3`H) sequenced in this study, BLAST top hits against GenBank protein database, similarity score, accession length and accession no.

	Scientific Name	Accession	Per. Identify	Max Score	Total Score	Query Cover	Acc. Length
1	*Ginkgo biloba*	AJO67233.1	100.00	1158	1158	100	556
2	*Amborellatrichopoda*	XP_006836296.1	75.44	832	832	92	516
3	*Lupinusalbus*	KAE9620790.1	75.15	832	832	92	522
4	*Lupinusangustifolius*	XP_019465190.1	74.09	769	769	88	521
5	*Ipomoea triloba*	XP_031090613.1	73.52	748	748	88	517
6	*Ipomoea nil*	XP_019158316.1	73.06	741	741	87	518
7	*Nymphaescolorata*	XP_031500353.1	73.60	764	764	92	532
8	*Glycine max*	XP_003541057.1	73.62	761	761	87	523
9	*Panicumhallii*	XP_025822223.1	72.82	766	766	88	503
10	*Nicotianatabacum*	XP_016495342.1	71.60	766	766	91	542
11	*Cajanuscajan*	XP_020224177.1	71.29	768	768	91	523
12	*Sesamumindicum*	XP_011093925.1	71.46	771	771	92	526
13	*Solanumlycopersicum*	XP_004248085.1	71.54	769	769	87	556
14	*Chenopodium quinoa*	XP_021748114.1	70.08	766	766	92	518
15	*Ricinuscommunis*	XP_002523334.1	70.76	764	764	92	515

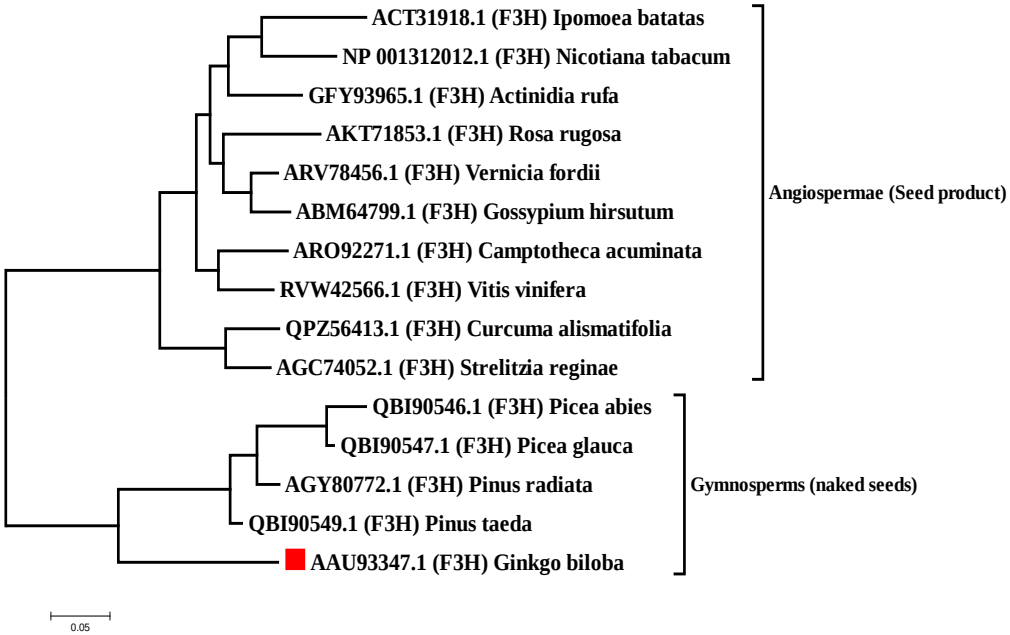

Figure (3): A phylogenetic tree showing the relationship of 15 amino acid sequences F3`Hs protein from several plants species included (AAU93347.1) GbF3H from *Ginkgo biloba*. Evolutionary analyses were conducted in MEGA7 using the Maximum Likelihood method and the tree with the highest log likelihood (-3319.73) is shown [16].

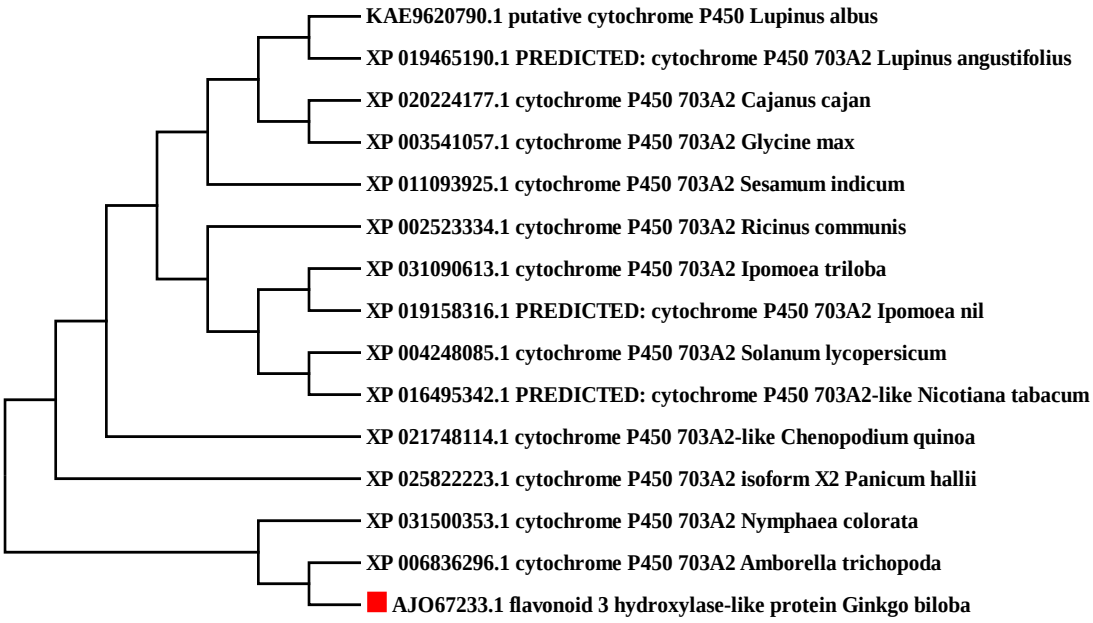

Figure (4): A phylogenetic tree showing the relationship of 15 amino acid sequences (F3`Hs) from several plants species included (AJO67233.1) GbF3`H from *Ginkgo biloba*. Evolutionary analyses were conducted in MEGA7 using the Maximum Likelihood method and the tree with the highest log likelihood (-5359.61) is shown [16].

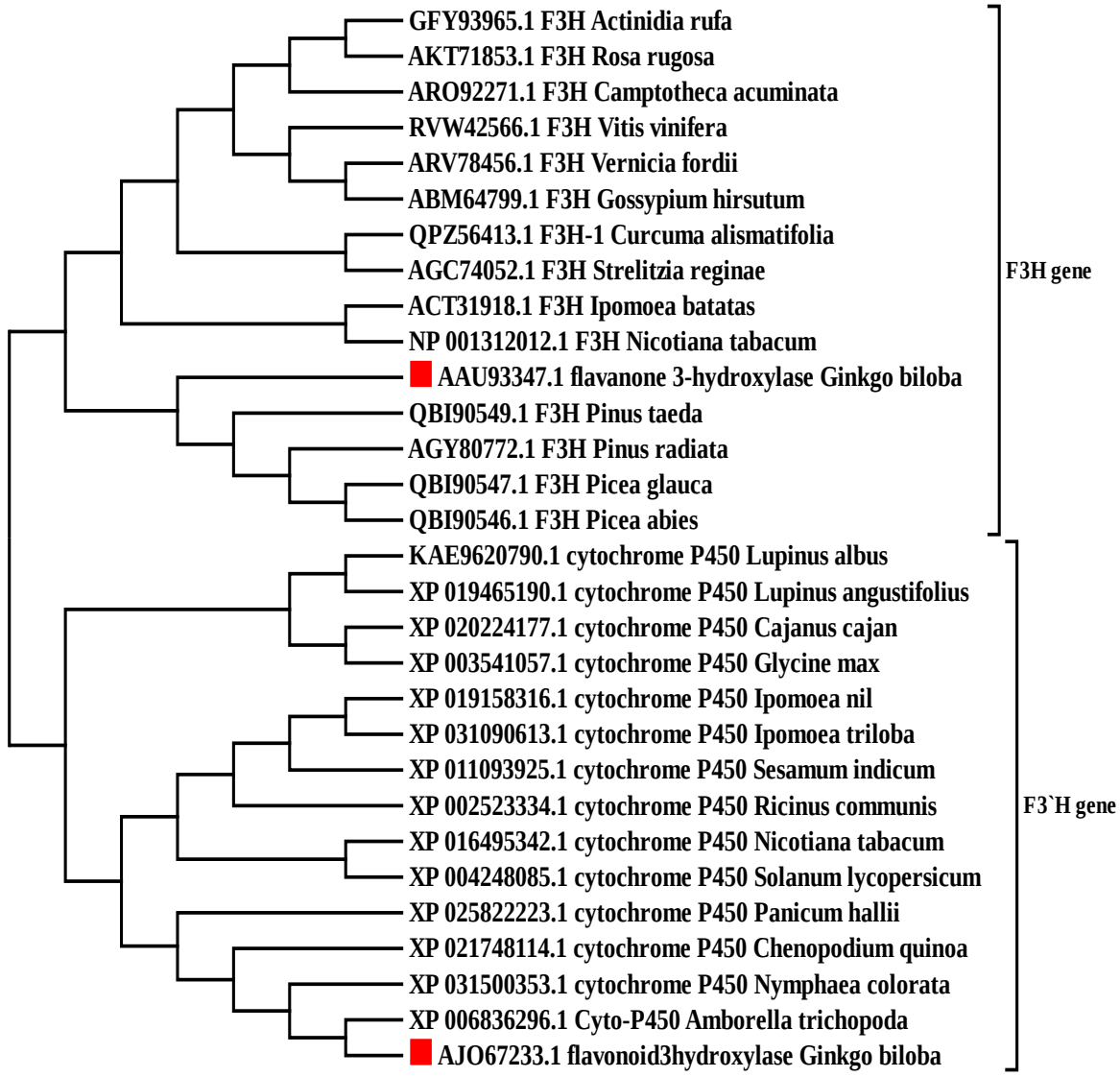

Figure (5): Molecularphylogenetic analysis involved 30 amino acid sequences (15flavanone 3-hydroxylase (F3Hs) gene and 15 flavonoid 3'hydroxylase (F3`Hs) gene) from several plant species (Gymnosperm and Angiosperms)included GbF3H and GbF3`H gene sequences were conducted in MEGA 7.0 software program by Maximum Likelihood method. The tree with the highest log likelihood (-6772.41) is shown [16].

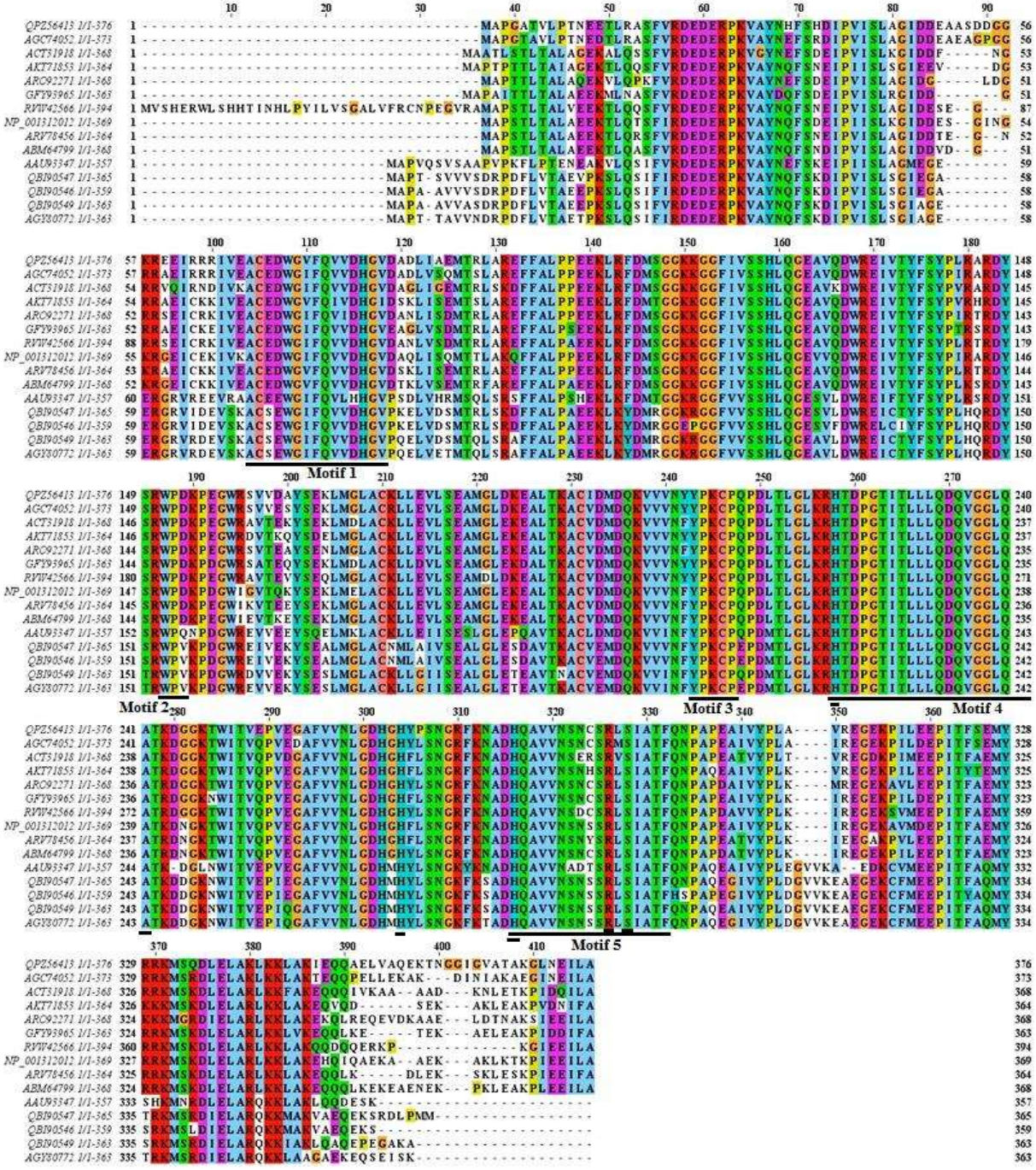

Figure (6): The amino acid sequence alignment of 15 flavanone 3-hydroxylase (F3Hs) gene sequences included GbF3H gene sequence with other plant species were used in this study (www.https://www.ebi.ac.uk/Tools/msa/clustalo/).Five motifs were obtained with dark underline. Genbank accession numbers for the F3H proteinsin the alignment are as follow:*Ginkgo biloba*(AAU93347.1), *Pinustaeda*(QBI90549.1),*Pinustaeda*(AGY80772.1), *Pinus radiate*(QBI90547.1), *Piceaabies*(QBI90546.1),*Ipomoea batatas*(ACT31918.1), *Camptotheca acuminate*(ARO92271.1),*Vitisvinifera*(RVW42566.1),*Verniciafordii*(ARV78456.1),*Nicotianatabacum*(NP_001 312012.1), *Curcuma alismatifolia*(QPZ56413.1),*Actinidiarufa*(GFY93965.1),*Gossypiumhirsutum*(ABM64799.1), *Rosa rugosa*(AKT71853.1)and *Strelitziareginae*(AGC74052.1).

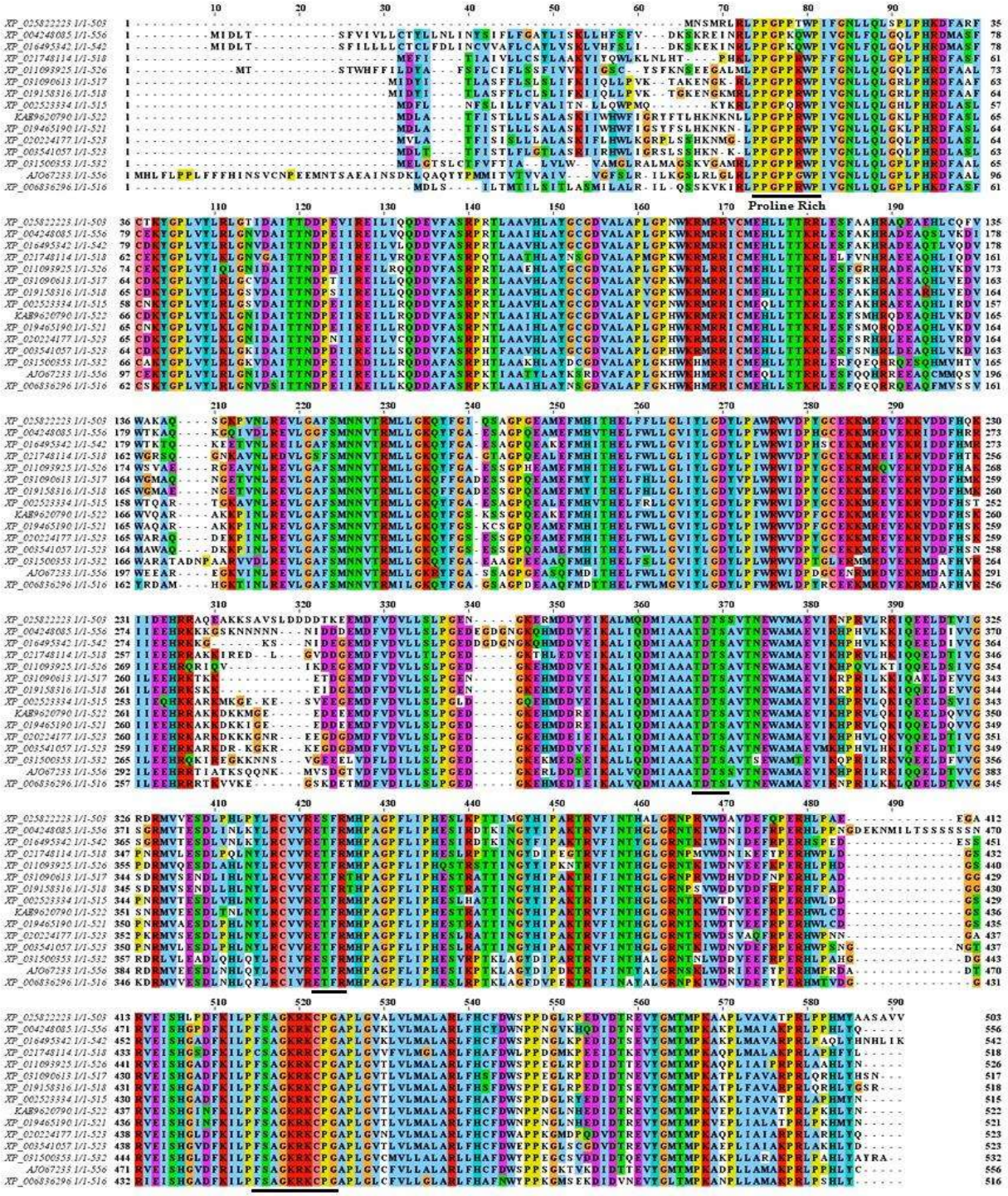

Figure (7): The amino acid sequence alignment of 15 flavonoid 3'hydroxylase (F3`Hs) gene sequences included GbF3`H gene sequence with other plant species were used in this study (www.https://www.ebi.ac.uk/Tools/msa/clustalo/).Several motifs were obtained with dark underline. Genbank accession numbers for the F3`H proteinsin the alignment are as follow:*Ginkgo biloba*(AJO67233.1),*Amborellatrichopoda*(XP_006836296.1),*Lupinusalbus*(KAE9620790.1),*Lupinusangustifolius*(XP_019465190.1), *Ipomoea triloba*(XP_031090613.1),*Ipomoea nil*(XP_019158316.1),*Nymphaescolorata*(XP_031500353.1), *Glycine max*(XP_003541057.1), *Panicumhallii*(XP_025822223.1),*Nicotianatabacum*(XP_016495342.1),*Cajanuscajan*

(XP_020224177.1),*Sesamumindicum*(XP_011093925.1),*Solanumlycopersicum*(XP_004248085.1), *Chenopodium quinoa*(XP_021748114.1)and *Ricinuscommunis*(XP_002523334.1).

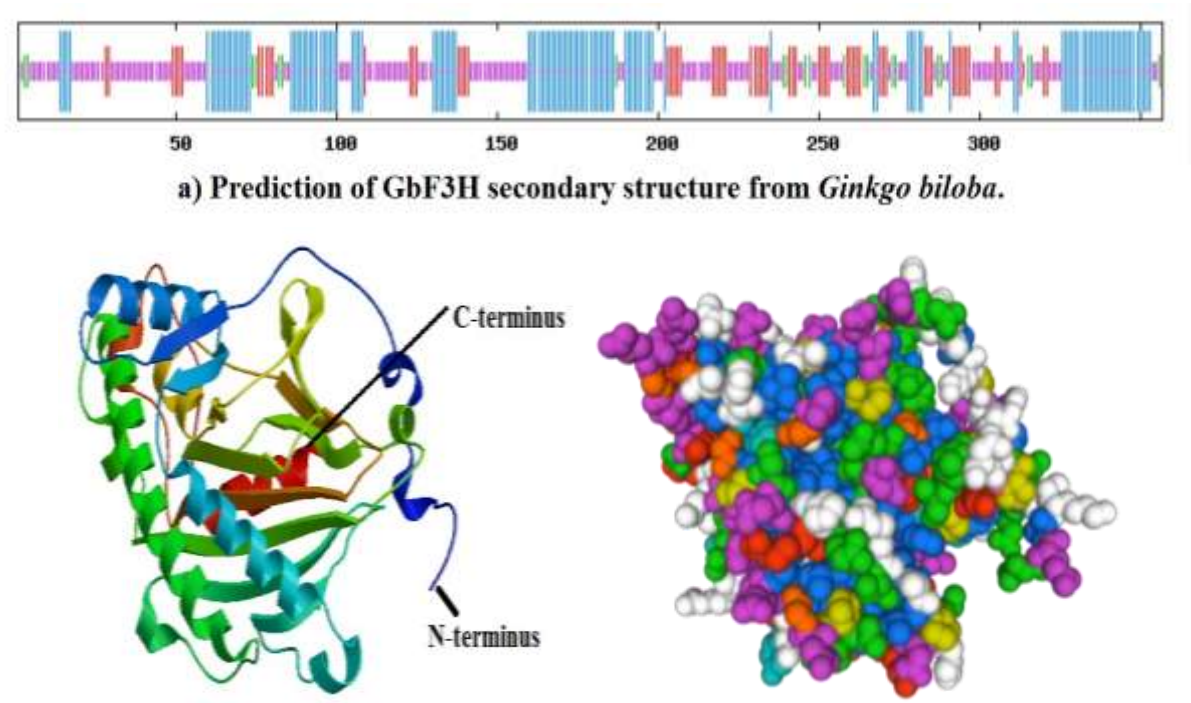

Figure (8): The three-dimensional model of flavanone 3-hydroxyrase GbF3H protein from *Ginkgo biloba*. a) Prediction of GbF3H secondary structure: ά–helicase in red and green and β-sheets are indicated by patches in blue by SOMPA program. Turns and loops are indicated by lines. (https://swissmodel.expasy.org/interactive).

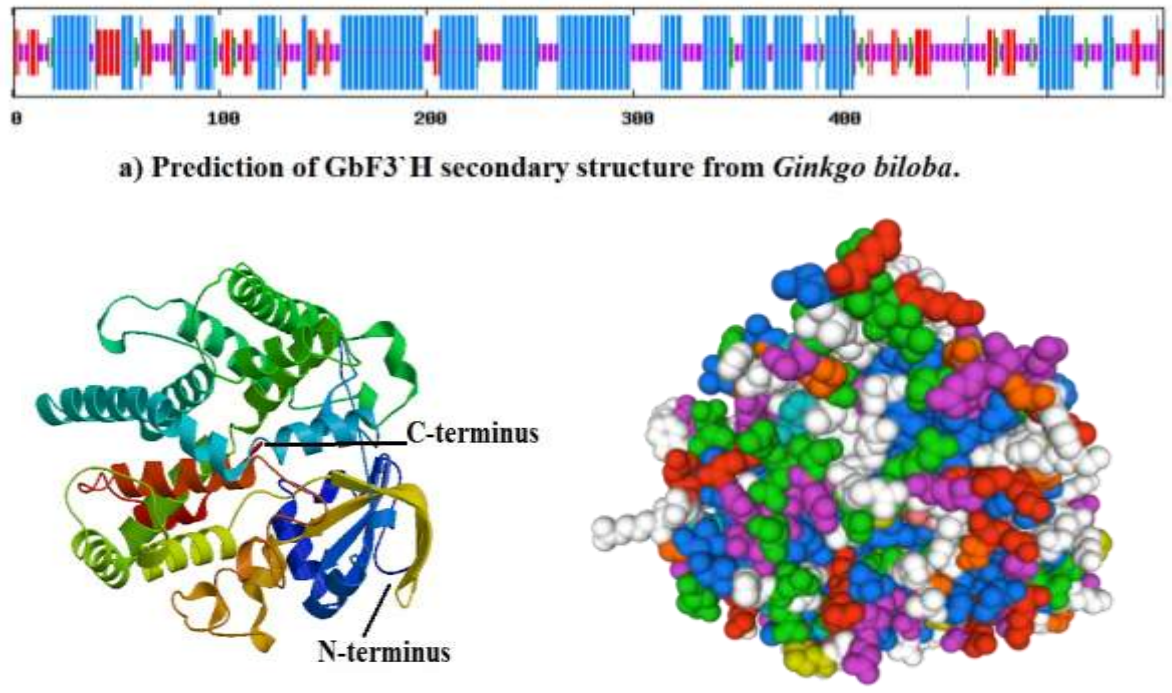

Figure (9): The three-dimensional model of flavonoid 3' hydroxylase-like protein (GbF3`H)from *Ginkgo biloba*. a) Prediction of GbF3H secondary structure:ά–helicase in red and green and β-sheets are indicated by patches in blueby SOMPA program. Turns and loops are indicated by lines (https://swissmodel.expasy.org/interactive).

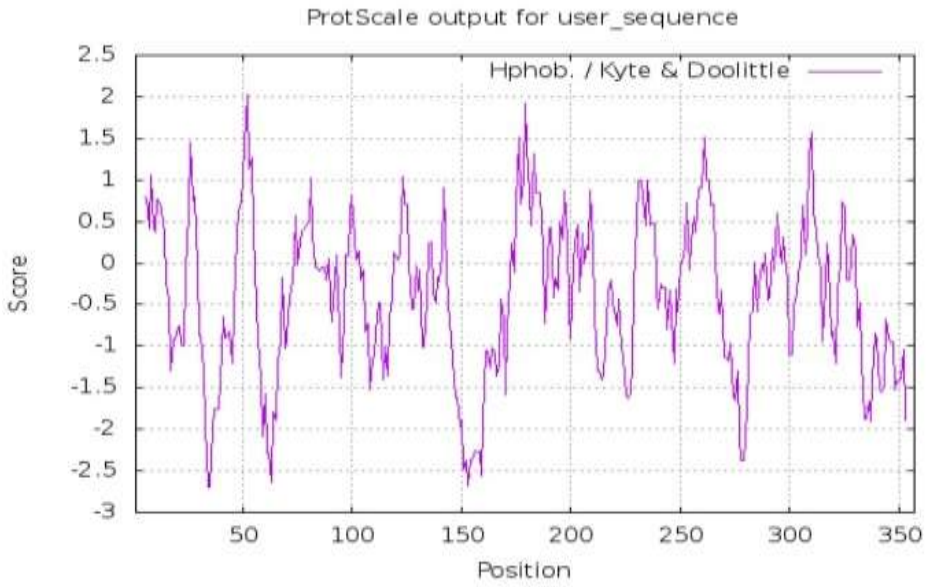

Figure (10): Hydrophilicityprofile of flavanone 3-hydroxyrase protein (GbF3H)from *Ginkgo biloba* L. using (https://web.expasy.org/protscale/).

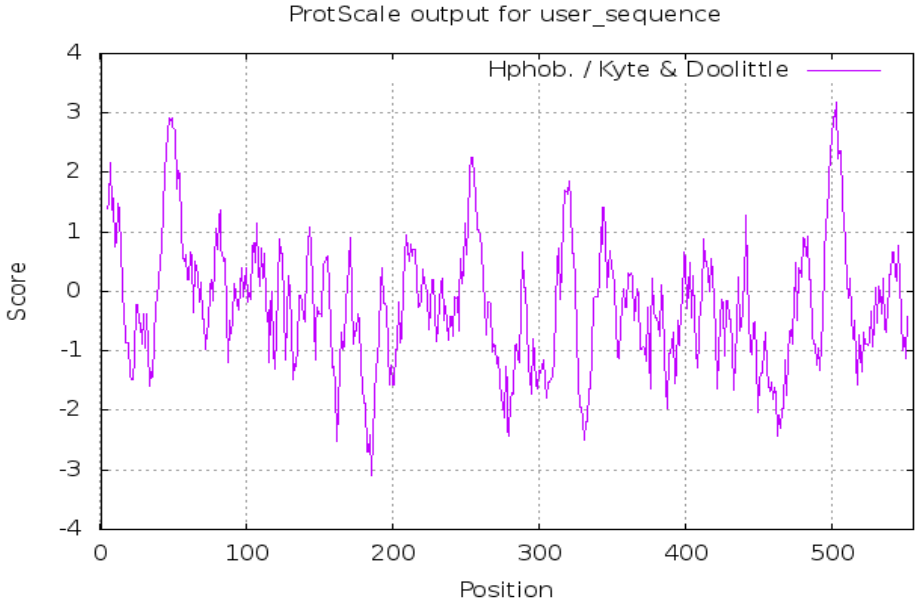

Figure (11): Hydrophilicityprofile of flavonoid 3' hydroxylase-like protein (GbF3`H)from *Ginkgo biloba* L. using (https://web.expasy.org/protscale/).

Table (3): As shown the top 20 filtered templates for protein from (SMTL) with high quality for flavanone 3-hydroxylase (F3Hs) model building with GbF3H protein 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 98 templates were found which were considered to be less suitable for modelling than the filtered list.

	Template from (SMTL - ID)	SeqIdentity	Seq Identity	QSQ	Found by	Resolution	Seq Similarity	Cover	Description
1	1gp4.1.A	33.02	monomer	-	HHblits	2.10Å	0.38	0.89	Anthocyanidinsynthasefrom *Arabidopsis thaliana*
2	2brt.1.A	32.61	monomer	-	HHblits	2.20Å	0.37	0.90	Leucoanthocyanidindioxygenasefrom *Arabidopsis thaliana*
3	5o7y.1.A	32.09	monomer	-	HHblits	1.97Å	0.37	0.90	Thebaine 6-O-demethylasefrom*Papaversomniferum*
4	4xae.1.A	31.33	monomer	-	HHblits	2.77Å	0.37	0.89	Feruloylcoaortho-hydroxylase 1 (F6H) from *Arabidopsis thaliana*
5	4xae.2.A	31.33	monomer	-	HHblits	2.77Å	0.37	0.89	
6	6ttm.1.A	30.09	monomer	-	HHblits	1.91Å	0.36	0.89	Hyoscyamine 6 beta-hydroxylase
7	5gja.1.A	28.97	homo-octamer	0.11	HHblits	2.10Å	0.36	0.81	1-aminocyclopropane-1-carboxylate oxidase 2 Crystal structure of *Arabidopsis thaliana* ACO2 in complex with POA
8	5gj9.1.A	28.97	monomer	-	HHblits	2.10Å	0.36	0.81	
9	6ttn.1.A	32.31	monomer	-	HHblits	1.12Å	0.37	0.82	Hyoscyamine 6 beta-hydroxylase
10	4xae.2.A	36.51	monomer	-	BLAST	2.77Å	0.39	0.85	Feruloylcoaortho-hydroxylase 1(F6H) from *Arabidopsis thaliana*
11	4xae.1.A	36.51	monomer	-	BLAST	2.77Å	0.39	0.85	
12	1wa6.1.A	31.51	homo-tetramer	0.17	HHblits	2.55Å	0.37	0.82	1-aminocyclopropane -1-carboxylate oxidase 1 The structure of ACC oxidase
13	6jyv.1.A	30.87	monomer	-	HHblits	1.65Å	0.35	0.83	Probable iron/ascorbateoxidoreductasefrom *Pseudomonas aeruginosa*PAO1
14	6ku3.1.A	24.15	homo-tetramer	0.04	HHblits	2.15Å	0.33	0.82	Gibberellin 2-beta-dioxygenase 3Crystal structure of gibberellin 2-oxidase3 (GA2ox3)inrice
15	5c3p.1.A	21.96	monomer	-	HHblits	2.10Å	0.31	0.83	Thymine dioxygenase Crystal structure of the full-length *Neurosporacrassa* T7H in complex with alpha-KG
16	5c3p.2.A	21.96	monomer	-	HHblits	2.10Å	0.31	0.83	
17	5c3p.3.A	21.96	monomer	-	HHblits	2.10Å	0.31	0.83	
18	5o7y.1.A	36.09	monomer	-	BLAST	1.97Å	0.39	0.75	Thebaine 6-O-demethylase from *Papaversomniferum*
19	1gp4.1.A	37.69	monomer	-	BLAST	2.10Å	0.40	0.73	Anthocyanidin synthase from *Arabidopsis thaliana*

20	**2brt.1.A**	**37.69**	**monomer**	**-**	**BLAST**	**2.20Å**	**0.40**	**0.73**	**Leucoanthocyanidin dioxygenase from *Arabidopsis thaliana***

Table (4): As shown the top 20 filtered templates for protein from (SMTL) with high quality for flavonoid 3' hydroxylase (F3`Hs) model building with GbF3`H protein 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 98 templates were found which were considered to be less suitable for modelling than the filtered list.

	Template from (SMTL-ID)	SeqIdentity	Seq Identity	QSQ	Found by	Resolution	Seq Similar	Cover	Description
1	**5ylw.1.A**	**31.07**	**monomer**	**-**	**HHblits**	**1.70Å**	**0.37**	**0.82**	**Ferruginolsynthasefrom *Salvia miltiorrhiza***
2	**5ylw.1.A**	**32.37**	**monomer**	**-**	**BLAST**	**1.70Å**	**0.38**	**0.81**	
3	**6vby.1.A**	**29.62**	**monomer**	**-**	**HHblits**	**1.70Å**	**0.36**	**0.81**	**Cinnamic acid 4-hydroxylase (Cytochrome P450-73A33) from Sorghum bicolor**
4	**2hi4.1.A**	**27.87**	**monomer**	**-**	**HHblits**	**1.95Å**	**0.34**	**0.85**	***Human Microsomal* CytochromeP450 1A2**
5	**6b82.1.B**	**26.84**	**homo-dimer**	**0.06**	**HHblits**	**3.03Å**	**0.33**	**0.83**	**Cytochrome P450, family 17, subfamily A, polypeptide 1 *Zebra Fish* CYP-450**
6	**6b82.1.A**	**26.84**	**homo-dimer**	**0.05**	**HHblits**	**3.03Å**	**0.33**	**0.83**	
7	**4i8v.1.A**	**28.44**	**monomer**	**-**	**HHblits**	**2.60Å**	**0.35**	**0.80**	**Human Cytochrome P450 1A1**
8	**6udm.2.A**	**28.44**	**monomer**	**-**	**HHblits**	**3.08Å**	**0.35**	**0.80**	
9	**6oyu.2.A**	**25.58**	**monomer**	**-**	**HHblits**	**2.95Å**	**0.34**	**0.85**	**Structure of an ancestral-reconstructed Cytochrome P450 1B1**
10	**3c6g.1.A**	**23.64**	**monomer**	**-**	**HHblits**	**2.80Å**	**0.32**	**0.83**	**Cytochrome P450 2R1 Crystal structure of CYP2R1 in complex with vitamin D3**
11	**4nkv.3.A**	**24.03**	**monomer**	**-**	**HHblits**	**2.65Å**	**0.33**	**0.83**	**Steroid 17-alpha-hydroxylase/17,20 lyase *Human steroidogenic cytochrome P450***
12	**4nkv.1.A**	**24.03**	**monomer**	**-**	**HHblits**	**2.65Å**	**0.33**	**0.83**	
13	**4nkw.1.A**	**24.03**	**monomer**	**-**	**HHblits**	**2.50Å**	**0.33**	**0.83**	**Steroid 17-alpha-hydroxylase/17,20 lyase Human steroidogenic cytochrome P450**
14	**4nkw.3.A**	**24.03**	**monomer**	**-**	**HHblits**	**2.50Å**	**0.33**	**0.83**	
15	**3c6g.2.A**	**23.64**	**monomer**	**-**	**HHblits**	**2.80Å**	**0.32**	**0.83**	**Cytochrome P450 2R1 Crystal structure of CYP2R1 in complex with vitamin D3**
16	**6oyu.1.A**	**25.58**	**monomer**	**-**	**HHblits**	**2.95Å**	**0.34**	**0.85**	**Structure of an ancestral-reconstructed cyto-P4501B1**
17	**6udl.4.A**	**28.44**	**monomer**	**-**	**HHblits**	**2.85Å**	**0.35**	**0.80**	**Structure of Human Cytochrome P450 1A1**
18	**6o5y.3.A**	**28.44**	**monomer**	**-**	**HHblits**	**3.17Å**	**0.35**	**0.80**	

19	**6oyv.2.A**	**25.58**	**monomer**	**-**	**HHblits**	**3.10Å**	**0.34**	**0.85**	
20	**6udl.3.A**	**28.44**	**monomer**	**-**	**HHblits**	**2.85Å**	**0.35**	**0.80**	

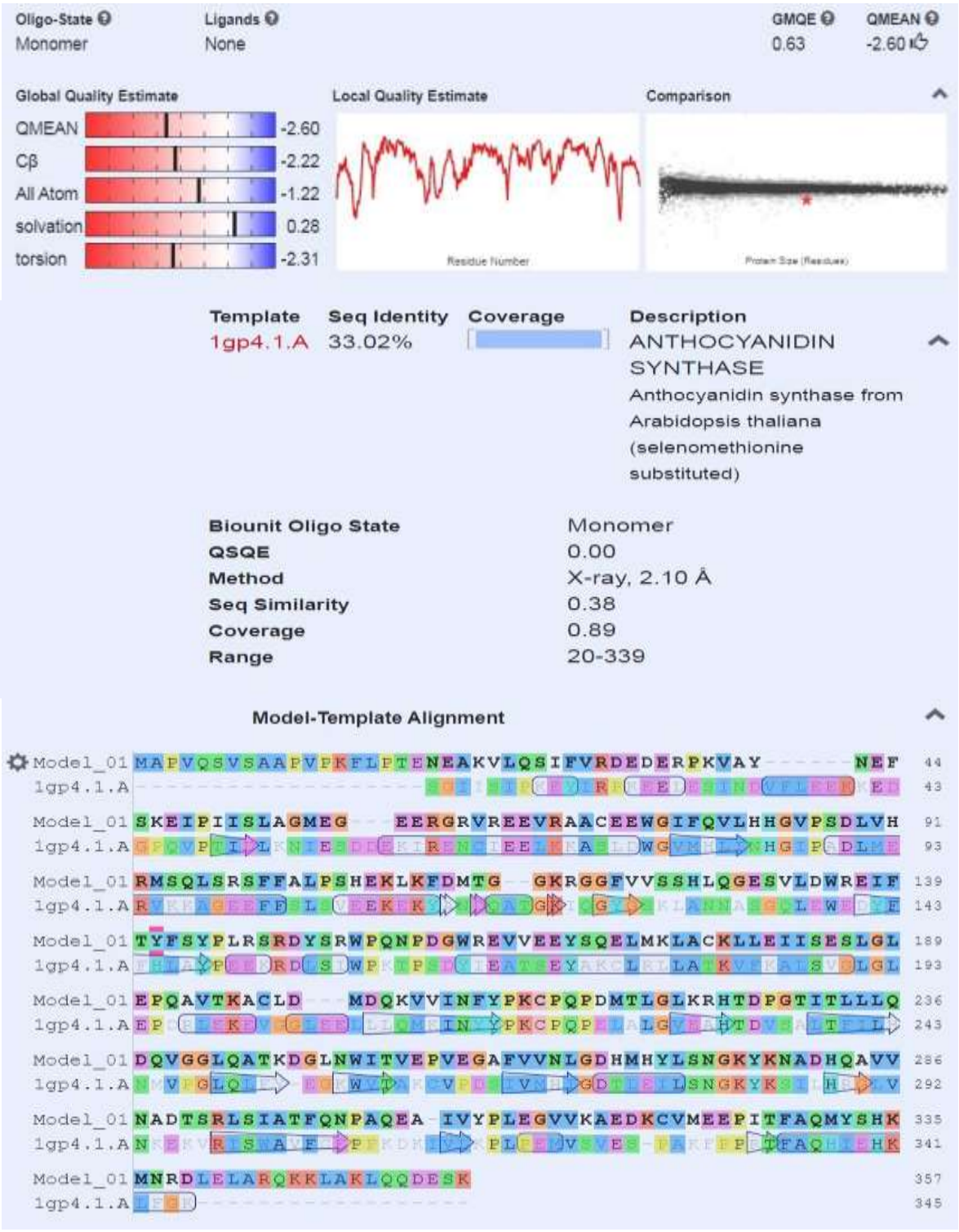

Figure (12): Homology model of flavanone 3-hydroxylase (GbF3H) from *Ginkgo biloba* showing quality estimate (global quality estimate, local quality estimate, comparison non-redundant set of PDB structures) and model-template alignment with STML-ID: 1gp4.1.A.

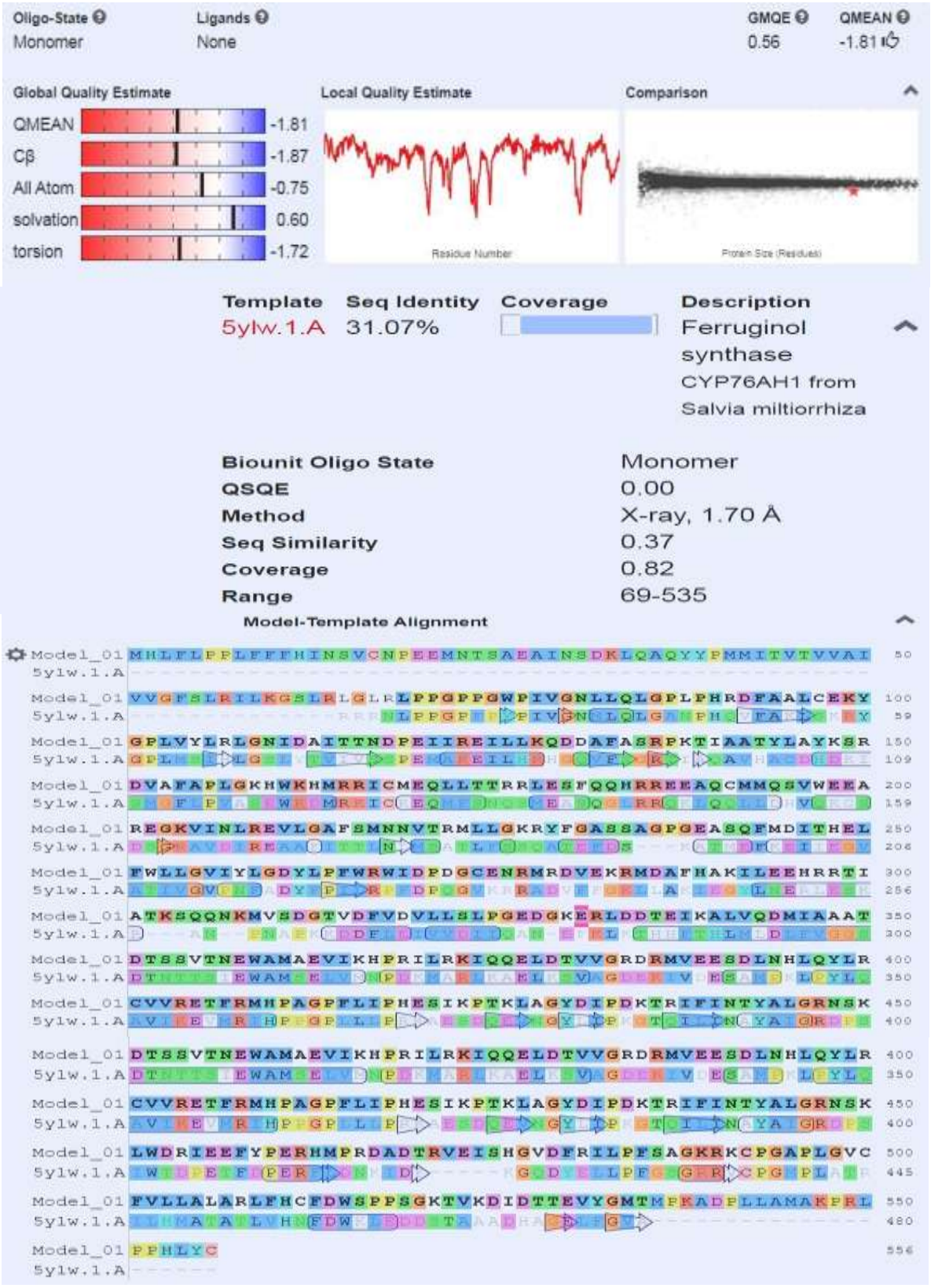

Figure (13): Homology model of flavonoid 3' hydroxylase-like protein (GbF3`H) from *Ginkgo biloba* showing quality estimate (global quality estimate, local quality estimate, comparison non-redundant set of PDB structures) and model-template alignment with STML-ID: 5y1w.1A.

Reference

[1]. Jacobs BP, Browner WS. *Ginkgo biloba*: a living fossil. Am. J Med. 2000; 108:341–342
[2]. Singh B, Kaur P, Gopichand; Singh RD, Ahuja PS. Biology and chemistry of *Ginkgo biloba*. Fitoterapia, 2008;79(6): 401–418.
[3]. Van Beek TA. Chemical analysis of *Ginkgo biloba* leaves and extracts. J Chromatography. 2002; 967:21–55.
[4]. Strømgaard K, Nakanishi K. Chemistry and Biology of TerpeneTrilactones from Ginkgo Biloba. Angew. Chem. Int. Ed. 2004; 43:1640 – 1658.
[5]. Liu H, Cao F, Yin T, Chen YA. Highly dense genetic map for *Ginkgo biloba* constructed using sequence-based markers. 2017; Front. Plant Sci., *8*, 1041.
[6]. Samanta A, Das G, Das SK. Roles of flavonoids in plants. Int. J. Pharm. Sci. Tech., (2011); 6(1):12-35.
[7]. Kumar S, Pandey AK. Chemistry and biological activities of flavonoids: An overview. Sci. World J., 2013; 162750.
[8]. Deng Y, Lu SF. Biosynthesis and regulation of phenylpropanoids in plants. Crit. Rev. Plant Sci., 2017; 36:1–34.
[9]. Tanaka Y, Brugliera F. Flower colour and cytochromes P450. Philos. Trans. R. Soc. Lond B. Biol. Sci., 2013; 368 (1612): 20120432. doi: 10.1098/rstb.2012.0432
[10]. Santos-Buelga C, Feliciano AS, Flavonoids: From structure to health issues. Molecules. 2017; 22 (3): 477. https://doi.org/10.3390/molecules22030477
[11]. Xu BB, Li JN, Zhang XK, Wang R, Xie LL, Chai YR. Cloning and molecular characterization of a functional flavonoid 3`-hydroxylase gene from *Brassica napus*. J Plant Physiol. 2007; 164:350–363. doi:10.1016/j.jplph.2006.03.001
[12]. Britsch L, Grisebach H. Improved preparation and assay of chalcone synthase. Phytochemistry 1985; 24:1975-1976.
[13]. Pelt JL, Downes WA, Schoborg RW, McIntosh CA. Flavanone 3-hydroxylase expression in *Citrus paradisi*and*Petunia hybrida*seedlings. Phytochemistry. 2003; 64:435-444.
[14]. Pelletier MK, Winkel-Shirley B. Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings. Plant Physiol. 1996; 111:339–345.
[15]. Charrier B, Coronado C, Kondorosi A, Ratet P. Molecular characterization and expression of alfalfa (*Medicago sativa* L.) flavanone-3-hydroxylase and dihydroflavonol-4-reductase encoding genes. Plant Mol. Biol. 1995; 29:773–786.
[16]. Shen G, Pang Y, Wu W, Deng Z, Zhao L, Cao Y, et al. Cloning and characterization of a flavanone 3-hydroxylase gene from *Ginkgo biloba*. Bioscience Rep. 2006; 26(1):19–29.
[17]. Shen X, Martens S, Chen M, Li D, Dong J, Wang T. Cloning and characterization of a functional flavanone-3-bhydroxylase genes from *Medicagotruncatula*. Mol. Biol. Rep. 2010; 37:3283–3289.
[18]. Liu M, Li X, Liu Y, Cao B. Regulation of flavanone 3-hydroxylase gene involved in the flavonoid biosynthesis pathway in response to UV-B radiation and drought stress in the desert plant, *Reaumuriasoongorica*. Plant Physiol. Bioch.; 2013; 73:161–67.
[19]. Zhang H, Zhao L, Wang J, et al. Cloning and functional analysis of two flavanone-3-hydroxylase genes from *Reaumuriatrigyna*. ActaPhysiol Plant. 2014; 36(5):1221–29.
[20]. Kumar A, Singh B, Singh K. Functional characterization of flavanone 3-hydroxylase gene from *Phyllanthusemblica* (L.). J Plant Biochem. Biotechnol., 2015; 24(4):453–460.
[21]. Song X, Diao J, Ji J, Wang G, et al. Molecular cloning and identification of a flavanone 3-hydroxylase gene from *Lyciumchinense*, and its overexpression enhances drought stress in tobacco. Plant Physiol. Biochem. 2016; 98: 89–100.
[22]. Xiong S, Tian N, Long J, Chen Y, et al. Molecular and characterization of a flavnone 3-Hydroxylase gene from *Artimisaannua* L. Plant physiology and Biochemistry. 2016; 105:29-36.
[23]. Brugliera F, Barri-Rewell G, Holton TA, Mason JG. Isolation and characterization of a flavonoid 30-hydroxylase cDNA clone corresponding to the Ht1 locus of Petunia hybrida. Plant J 1999; 19:441–451. doi:10.1046/j.1365-313X.1999.00539.x
[24]. Schoenbohm c, Martens s, Eder c, Forkmann g, Weisshaar b. Identification of the *Arabidopsis thaliana* flavonoid 3′-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol. Chem. 2000; 381:749-753.
[25]. Han YP, Vimolmangkang S, Soria-Guerra RE. et al. Ectopic expression of apple F3`H genes contribute to anthocyanin accumulation in the Arabidopsis tt7 mutant grown under nitrogen stress. Plant Physiol., 2010; 153, 806–820.
[26]. Zhou W, Gong, YF, Lu X, Huang, CT, Gao F. Molecular cloning and characterization of a flavonoid 31-hydroxylase gene from purple-fleshed sweet potato (*Ipomoea batatas*). Molecular Biology Reports. 2012; 39 (1):295– 302.
[27]. Li L-l, Cheng H, Chen X-l, Cheng S-Y. Molecular cloning, characterization and expression of Flavonoid 3' Hydroxylase-like protein gene from *Ginkgo biloba*. ActaHorticulturaeSinica. 2015; 42(4): 643-654. doi: 10.16420/j.issn.0513-353x.2014-0822
[28]. 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.
[29]. Korbie JD, Mattick JS. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat. Protoc., 2008; 3(9):1452-1456.
[30]. 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
[31]. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 2016; 33:1870-1874.
[32]. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucl. Acids Res., 2003; 31:3381–3385.
[33]. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 1982; 157:105-132.
[34]. Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on ExPASy server, (In) Johan M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. 2005; pp. 571-607.
[35]. Jin Z, Grotewold E, Qu W, Fu G, Zhao D. Cloning and characterization of a flavanone 3-hydroxylase gene from *Saussurea medusa*. DNA Sequence, April 2005; 16(2): 121–129.
[36]. Winkel-Shirley B. 2001a. It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. Plant Physiol 12:1399–1404.
[37]. Winkel-Shirley B. 2001b. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126:485–493.
[38]. Lukacin R, Britsch L. Identification of strictly conserved histidine and arginine residues as part of the active site in Petunia hybridaflavanone 3beta-hydroxylase. Eur. J Biochem., 1997.249:748–757.
[39]. Chapple, C., Molecular-genetic analysis of plant cytochrome P450- dependent monooxygenases. Annu. Rev. Plant Physiol. Plant Mol. Biol. 1998. 49:311-343.
[40]. Werck-Reichhart D, Feyereisen R, Cytochromes P450: a success story. Genome Biol. 2000; 1, REVIEWS3003.

[41]. Waterhouse A, Bertoni M, Bienert S, Studer G, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018; 46, W296-W303.
[42]. Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421-430.
[43]. Steinegger M, Meier M, Mirdita M, et al. HH-suite3 for fast remote homology detection and deep protein annotation. BMC Bioinformatics. 2019; 20, 473.
[44]. Remmert M, Biegert A, Hauser A, SödingJ.HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nature Methods 2011; 9:173-175.
[45]. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis. 2009; 30, S162-S173.
[46]. Studer G, Rempfer C, Waterhouse AM, et al. QMEANDisCo - distance constraints applied on model quality estimation. Bioinformatics. 2020; 36:1765-1771.
[47]. 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.
[48]. 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.

(Ahmed, M.Z.S.), et. al. "Homology Modeling and Structural Analysis of theFlavanone 3-Hydroxylase (F3H) and Flavonoid 3`hydroxylase (F3`H) Genes from *Ginkgo biloba* (L.)." *IOSR Journal of Biotechnology and Biochemistry (IOSR-JBB)*, 7(1), (2021): pp. 01-21.