



## **IN SILICO ANALYSIS OF ARTEMISININ BIOSYNTHETIC GENES AND SEARCH OF METABOLIC PATHWAYS IN OTHER PLANTS**

**Manprit Kaur<sup>1</sup>, Rahul Verma<sup>2</sup>, Arvind Sharma<sup>3</sup>, Ashwani Kumar<sup>4\*</sup>**

<sup>1</sup>Department of Biotechnology, Chaudhary Bansi Lal University, Bhiwani, Haryana, India

<sup>2</sup>Department of Biotechnology, Chaudhary Bansi Lal University, Bhiwani, Haryana, India

<sup>3</sup>NIMR, New Delhi

<sup>4\*</sup>Department of Biotechnology, Chaudhary Bansi Lal University, Bhiwani, Haryana, India,  
Email: ashwani.biotech@cblu.ac.in

**\*Corresponding Author:** Ashwani Kumar

\*Department of Biotechnology, Chaudhary Bansi Lal University, Bhiwani, Haryana, India,  
Email: ashwani.biotech@cblu.ac.in

### **Abstract**

One of the most significant drugs for treating malaria is artemisinin, derived from the plant *Artemisia annua*. Artemisinin has seen a sharp rise in demand over the past ten years due to the development of resistance against other antimalarial drugs. However, natural artemisinin production is meager, which results in a global supply scarcity. Chemical synthesis of it is challenging due to its complicated structure. As a result, *Artemisia annua* remains the sole plant from which artemisinin is commercially available. Therefore, it is necessary to develop new plans to enhance output or identify alternative sources. In the present study, the multiple sequence alignments, secondary structure analysis, tertiary structure prediction, functional validation and protein-protein interactions of six enzymes namely ADH1, ALDH1, AMS1, CPR, CYP71AV1 and DBR2 were examined. In-silico analysis carried out in this paper has produced promising results.

**Keywords:** *Artemisia annua*, biosynthetic pathway, ADH1, ALDH1, AMS1, CYP71AV1, protein-protein interactions.

### **1. INTRODUCTION**

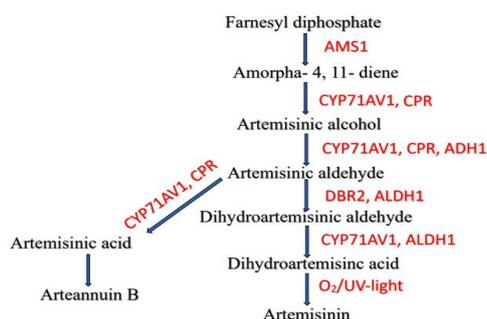
According to World Health Organisation, malaria continues to be a major global health concern with an estimated 214 million annual infections and 430,000 annual deaths, mostly among children under the age of 5. *Plasmodium* sp., in particular *Plasmodium falciparum*, which thrives in female *Anopheles* mosquitoes, are the culprits behind this lethal illness (Cox, 2010). The declines in malaria mortality and morbidity are mostly attributable to antimalarial drugs. The search for new antimalarial drugs has been centred on screening traditional medicine for many years (Simonsen et al., 2001; Adia et al., 2016; Nondo et al., 2017). *Artemisia annua* L. (sweet wormwood) is a member of the family Asteraceae and is classified under the genus *Artemisia*. It is a native of China and was first discovered in the steppes of the Chahar and Suiyuan Provinces. It is now widely cultivated throughout the world, including Vietnam, Thailand, Burma, Madagascar, Malaysia, the United States, Brazil, Australia (Tasmania), Holland, Switzerland, France, and Finland. *A. annua* produces artemisinin (a sesquiterpene lactone) which has remarkable efficacy against multi-drug resistant malaria-causing

parasite (*Plasmodium* species), however it is in short supply and not accessible to most malaria patients (Jing et al., 2009). Artemisinin and its semi-synthetic artemisinin derivatives, such as dihydro-artemisinin, artesunate, artemether, and arteether, are used in conjunction with other artemisinin derivatives even though the extracted artemisinin content of *Artemisia annua* is very low (dry weight approximately 0.01 to 1%) for the production of malaria combination medicines (Ro et al., 2006). In addition to being effective against malaria, artemisinin and its derivatives are also effective against a number of cancers and tumours (Efferth 2017), some viruses (Efferth et al., 2008; Efferth, 2018), including SARS-CoV-2 (Nair et al., 2021), tuberculosis (Martini et al., 2020; Zheng et al., 2017), and a variety of parasites, including *Schistosoma* species (Munyangi et al., 2018). In two plants, *Nicotiana benthamiana* and *Physcomitrella patens*, artemisinin may also be generated heterologously (Han et al., 2016; Wang et al., 2016).

### 1.1 Biosynthetic pathway:

IPP (isopentenyl diphosphate) and its isomer, DMAPP (dimethylallyl diphosphate), which were produced by the MVA (mevalonate) pathway in the cytoplasm and the MEP (methylerythritol phosphate) pathway in the plastid, were used as raw materials to generate artemisinin from the plant terpenoid biosynthesis pathway. Farnesyl diphosphate synthase (FPPS) catalysed the conversion of two molecules of IPP and one molecule of DMAPP into 15-carbon FPP, which entered the artemisinin biosynthetic pathway. The production of artemisinin increased when FPS was overexpressed in *Artemisia annua* (Han et al., 2006; Banyai et al., 2010), confirming the importance of FPS and the availability of substrates in the control of artemisinin biosynthesis (Simonsen et al., 2013). The enzyme amorpha-4,11-diene synthase (ADS) transforms FPP into amorpha-4,11-diene (**Fig.1**). Amorpha4,11-diene is hydroxylated in the following two stages of oxidation to produce artemisinic alcohol, which is then further oxidised to form artemisinic aldehyde in the presence of the cytochrome P450 enzyme amorphadiene monooxygenase (CYP71AV1). Artemisinic aldehyde is further reduced by the trichome-resident enzyme ADH1 to dihydroartemisinic aldehyde, which is then further oxidised by the trichome-resident enzyme artemisinic aldehyde 11 (13) reductase (DBR2) to dihydroartemisinic acid (Liu et al., 2016). Both the oxidation of dihydroartemisinic aldehyde to acid and the oxidative conversion of artemisinic aldehyde to acid in plants (catalysed by CYP71AV1) are known to be facilitated by the enzyme ALDH1 (Teoh et al., 2006; Teoh et al., 2009).

This research concentrates on the search for enzymes involved in artemisinin biosynthesis pathways in metabolic pathways of other plants because *Artemisia annua* has a low artemisinin content and poor availability against high demand.



**Fig. 1.** Artemisinin biosynthetic pathway in *Artemisia annua*. The intermediates of pathways are defined as AMS1, amorpha-4,11-diene synthase; CYP71AV1, amorpha-4,11-diene C-12 oxidase; CPR, cytochrome P450 reductase; ADH1, alcohol dehydrogenase 1; ALDH1, aldehyde dehydrogenase 1, DBR2, artemisinic aldehyde reductase.

### 2. Methodology

Protein sequences of amorpha-4,11-diene synthase (AMS1), amorpha-4,11-diene C-12 oxidase (CYP71AV1), alcohol dehydrogenase 1 (ADH1), aldehyde dehydrogenase 1 (ALDH1), cytochrome P450 reductase (CPR) and artemisinic aldehyde reductase (DBR2) from various plant sources were

retrieved from NCBI (<http://www.ncbi.nlm.nih.gov>) in FASTA format for computational analysis. Using the free CLC Sequence Viewer version 8.0 (<https://clc-sequence-viewer.software.informer.com/8.0/>), the alignment of all chosen sequences was examined. The helix, sheet, turn, and coil of amino acid sequences were predicted by Chou and Fasman secondary structure prediction (CFSSP) server (<http://cho-fas.sourceforge.net/index.php>). SWISS-MODEL (<https://swissmodel.expasy.org/interactive>) was used to design the 3D model of target proteins by selecting the best template. Further functional annotations were made to the resultant 3D structure. From the UCLA-DOE LAB— SAVES v6.0 server (<https://saves.mbi.ucla.edu/>), the predicted protein model was examined and validated. PROCHECK was used to evaluate the Ramachandran plot from SAVES. The STRING v10.0 server (<http://string-db.org/>), which creates a functional protein association network, was used to determine how the target proteins interact with other proteins that are closely related to them.

### 3. RESULTS AND DISCUSSION

#### 3.1 Sequence analysis:

The amino acid sequences from six different enzymes which are involved in the biosynthesis of artemisinin were retrieved from the NCBI database and were taken into consideration for further computational analysis. The list of the enzymes with accession numbers and their plant sources is shown in **Table 1**.

**Table 1.** List of the enzymes involved in biosynthesis of artemisinin and their plant sources taken for the study.

Accession No.	Name of Enzyme	Organism
AEI16475.1	alcohol dehydrogenase 1 (ADH1)	<i>Artemisia annua</i>
XP_022008917.1	alcohol dehydrogenase 1 (ADH1)	<i>Helianthus annuus</i>
ACR61719.1	aldehyde dehydrogenase 1 (ALDH1)	<i>Artemisia annua</i>
XP_035836614.1	aldehyde dehydrogenase 1 (ALDH1)	<i>Helianthus annuus</i>
AFA34434.1	amorpha-4,11-diene synthase (AMS1)	<i>Artemisia annua</i>
ABB82944.1	amorpha-4,11-diene C-12 oxidase (CYP71AV1)	<i>Artemisia annua</i>
ABL09938.1	cytochrome P450 reductase (CPR)	<i>Artemisia annua</i>
AHM95445.1	cytochrome P450 reductase (CPR)	<i>Matricaria chamomilla</i> <i>var. recutita</i>
GEV10977.1	cytochrome P450 reductase (CPR)	<i>Tanacetum cinerariifolium</i>
BAU61367.1	artemisinic aldehyde reductase (DBR2)	<i>Artemisia annua</i>
BAU61366.1	artemisinic aldehyde reductase (DBR2)	<i>Artemisia absinthium</i>

#### 3.2. Multiple sequence alignment (MSA):

The amino-acid sequences from four enzymes- alcohol dehydrogenase 1 (ADH1), aldehyde dehydrogenase 1 (ALDH1), cytochrome P450 reductase (CPR) and artemisinic aldehyde delta (11(13)) reductase (DBR2) were subjected to multiple sequence alignments using CLC Sequence Viewer 8 and compared for the recognition of conserved residues (**Fig. 2-5**).

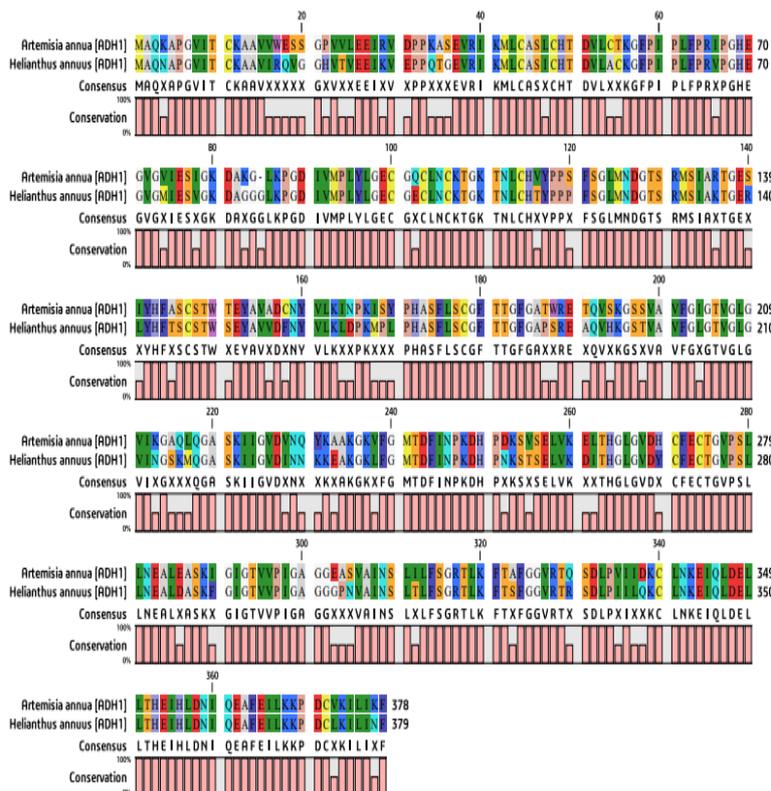


Fig. 2. MSA of the amino acid sequence of ADH1 of *Artemisia annua* with the corresponding sequence of *Helianthus annuus*.

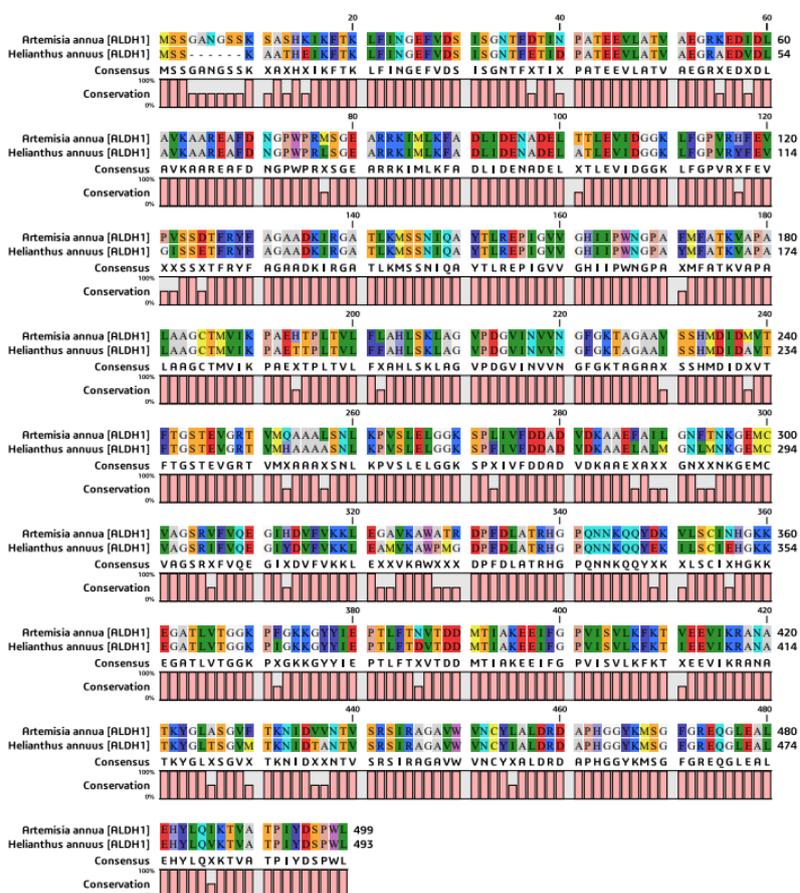
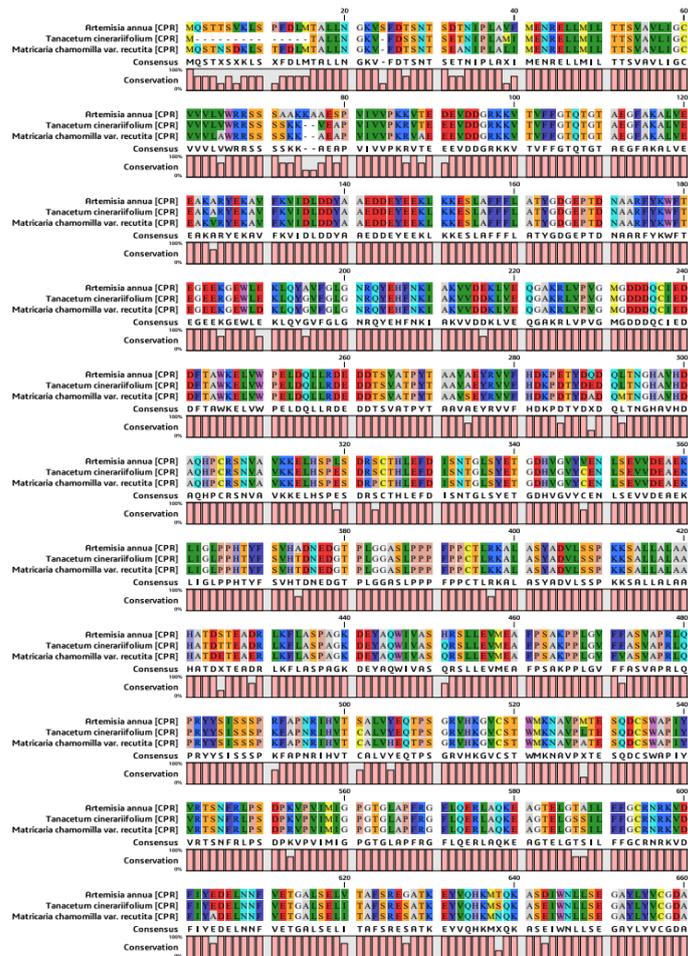
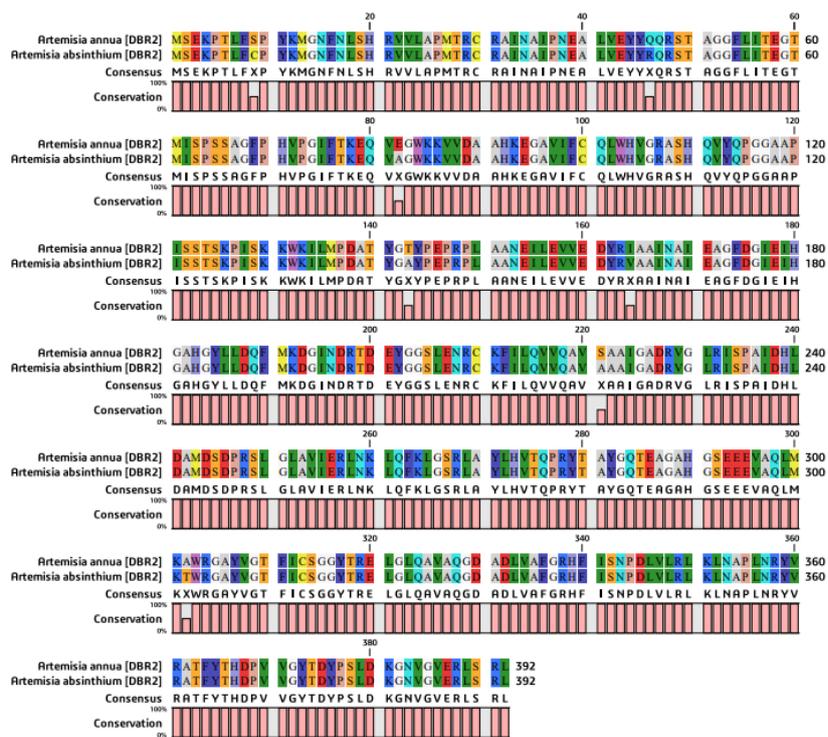


Fig. 3. MSA of the amino acid sequence of ALDH1 of *Artemisia annua* with the corresponding sequence of *Helianthus annuus*.



**Fig. 4.** MSA of the amino-acid sequence of CPR of *Artemisia annua* with the corresponding sequences of *Tanacetum cinerariifolium* and *Matricaria chamomilla* var. *recutita*.



**Fig 5.** MSA of the amino-acid sequence of DBR2 of *Artemisia annua* with the corresponding sequence of *Artemisia absinthium*.

### 3.3 Secondary Structure prediction:

The secondary structure prediction outcome displayed here showed that the percentage of alpha helix was significantly higher than the percentage of other types of protein conformations, such as sheet and turn (Fig. 6). The percentages of helix, sheet and turn are 63.7, 71.4 and 8.2 in ADH1 of *Artemisia annua*, 56.2, 69.7 and 8.2 in ADH1 of *Helianthus annuus*, 79.2, 62.7 and 12.0 in ALDH1 of *Artemisia annua*, 81.9, 62.79 and 11.4 in ALDH1 of *Helianthus annuus*, 72.7, 71.2 and 11.9 in AMS1 of *Artemisia annua*, 75.7, 58.1 and 13.6 in CPR of *Artemisia annua*, 71.0, 56.3 and 13.8 in CPR of *Matricaria chamomilla* var. *recutita*, 73.0, 56.3 and 13.8 in CPR of *Tanacetum cinerariifolium*, 59.4, 38.3 and 12.2 in DBR2 of *Artemisia annua*, 57.1, 40.1 and 12.2 in DBR2 of *Artemisia absinthium*, and 77.2, 64.6 and 12.5 in CYP71AV1 of *Artemisia annua*. Our results showed that the percentage of helix is higher in all predicted secondary structures but exceptions were for ADH1 of *Artemisia annua* and *Helianthus annuus*. In both, the percentage of sheet (71.4 in *Artemisia annua* and 69.7 in *Helianthus annuus*) is more than helix (63.7 in *Artemisia annua* and 56.2 in *Helianthus annuus*).

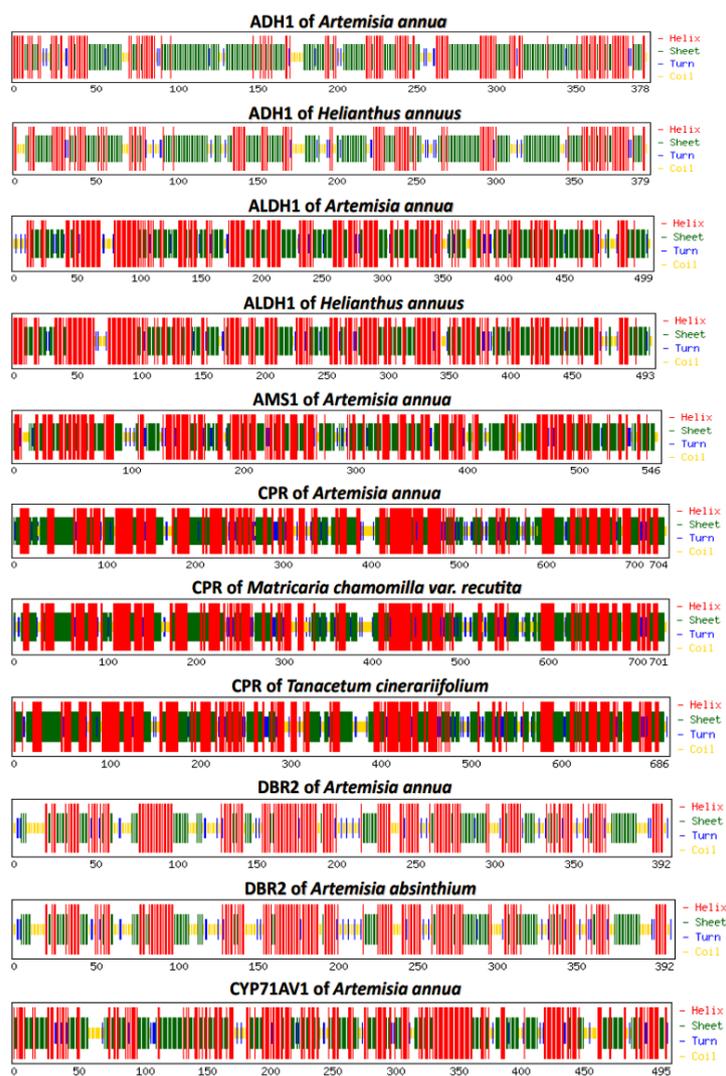


Fig. 6. Secondary structure prediction of amino-acid sequences of enzymes involved in artemisinin production.

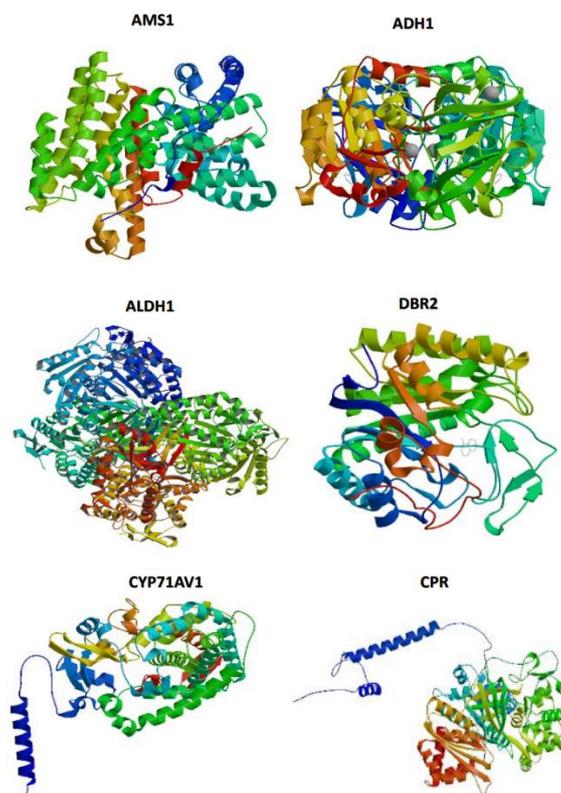
### 3.3 Tertiary Structure prediction:

As demonstrated in Table 2, SWISS-MODEL was successfully used to implement homology-based 3-D modelling of the target enzymes based on their corresponding best templates. Here, the selected template sequence for AMS1 was 4gax.1. A, a crystal structure of an alpha-Bisabolol synthase mutant that shared 82.23% sequence identity with query sequence. The chosen template for ADH1 was 6ljh.1.A, a crystal structure of Alcohol dehydrogenase 1 from *Artemisia annua* in complex with

NAD<sup>+</sup> that shares 100% sequence identity with the query sequence. For ALDH1, the structure of Zm ALDH2-3 (RF2C) in complex with NAD was picked as a suitable template sequence with SMTL ID 4pxl.1.A. The template sequence and query sequence shared 64.88% of the sequence identity. The crystal structure of DBR2 with mutation M27L was chosen as the best template sequence for implementing the three-dimensional protein model of DBR2 and shared 95.29% sequence identity with the query sequence. The SMTL ID of the selected template sequence for DBR2 was 5dy2.1.A. An AlphaFold DB model of GAO\_HELAN (gene: GAO, organism: *Helianthus annuus* (Common sunflower)) with SMTL ID, D5JBX0.1.A was chosen as the best template sequence for CYP71AV1 that shared 84.22% sequence identity with the target protein sequence. The template sequence selected for CPR was A0A2U1LIM9.1.A, an AlphaFold DB Model of NCPR1\_ARTAN (gene: CPR1, organism: *Artemisia annua* (Sweet wormwood)). The sequence identity between the template sequence and query sequence for CPR was 99.57%. **Fig. 7.** depicts the three-dimensional structures of the target proteins that were modelled.

**Table 2:** Template search result of target proteins by SWISS MODEL.

Target protein	Template							
	SMTL ID	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Coverage
AMS1	4gax.1.A	82.23	Monomer	HHblits	X-rays	1.99Å	0.56	1
ADH1	6ljh.1.A	100.00	Homo-dimer	HHblits	X-rays	1.80Å	0.62	1
ALDH1	4pxl.1.A	64.88	Homo-tetramer	HHblits	X-rays	2.25Å	0.50	0.97
DBR2	5dy2.1.A	95.29	Monomer	-	X-rays	1.57Å	0.60	0.97
CYP71AV1	D5JBX0.1.A	84.22%	Monomer	AFDB search	AlphaFold v2	-	0.57	0.99
CPR	A0A2U1LIM9.1.A	99.57	Monomer	AFDB search	AlphaFold v2	-	0.61	1

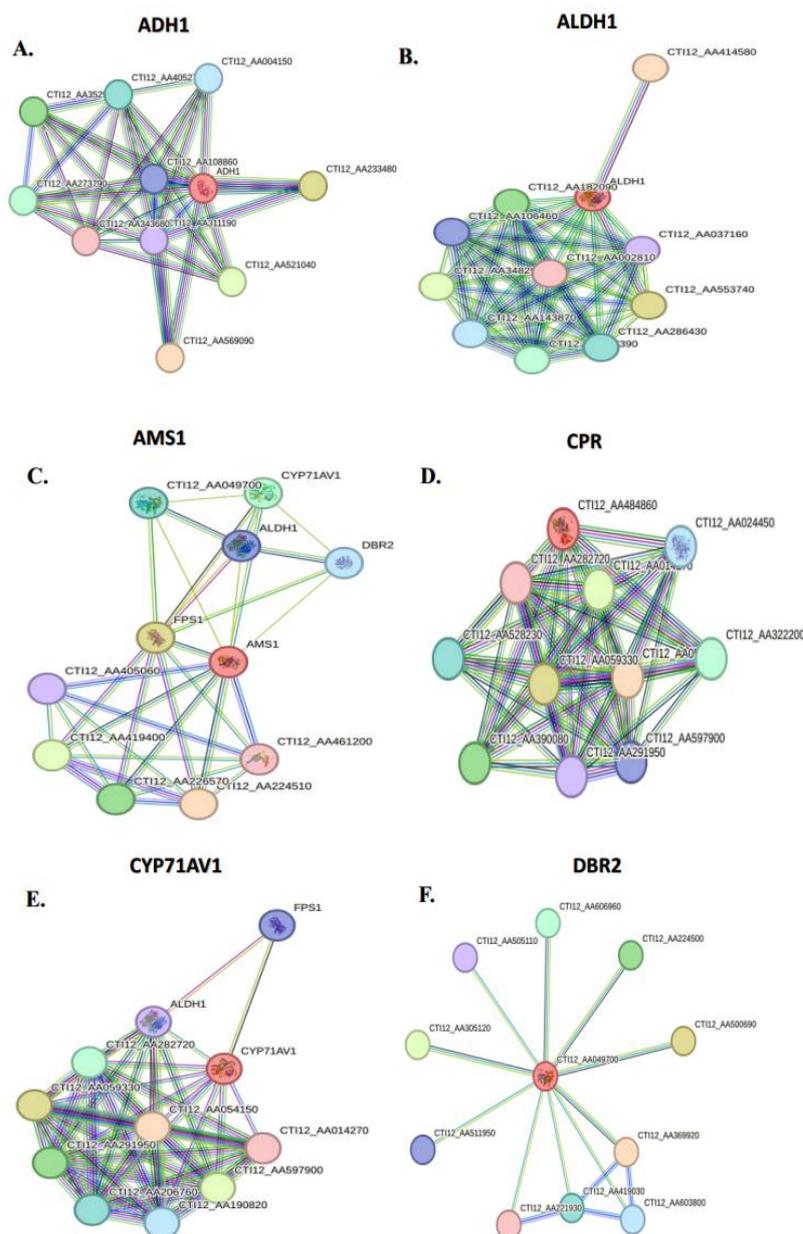


**Fig. 7.** Protein structure of the representative target proteins involved in the artemisinin biosynthetic pathway modelled using SWISS-MODEL.



**3.5 Protein-protein interaction study:** After that, we used the STRING database to conduct a protein-protein interaction network analysis (PPI). Individual proteins are represented as nodes and predicted functional associations between the proteins are represented by edges in the STRING output. The PPI analysis of ADH1 represents 11 nodes and 37 edges (**Fig. 9A**). Ten interacting proteins were CTI12\_AA569090, CTI12\_AA233480, CTI12\_AA004150 and CTI12\_AA108860, S-formylglutathione hydrolase; CTI12\_AA521040, CTI12\_AA352590 and CTI12\_AA273790, Aldehyde dehydrogenase; CTI12\_AA405270, Delta 1-pyrroline-5-carboxylate synthetase; CTI12\_AA311190, Alcohol dehydrogenase-like 7 and CTI12\_AA343680, Glucose/ribitol dehydrogenase. The PPI analysis of ALDH1 represents 11 nodes and 46 edges (**Fig. 9B**). The interacting proteins were CTI12\_AA414580, Acyl-CoA synthetase 5; CTI12\_AA553740, Ketose-bisphosphate aldolase class-II family protein; CTI12\_AA348260, Ketose-bisphosphate aldolase class-II family protein; CTI12\_AA182090 and CTI12\_AA002810, Glutamate/acetyl glutamate kinase; CTI12\_AA307390, CTI12\_AA286430 and CTI12\_AA286430, Delta-1-pyrroline-5-carboxylate synthase; CTI12\_AA143870 and CTI12\_AA037160, Aldehyde dehydrogenase. The PPI analysis of AMS1 represents 11 nodes and 33 edges (**Fig. 9C**). The interacting proteins were CTI12\_AA224510, Geranylgeranyl pyrophosphate synthase; FPS1, Farnesyl pyrophosphate synthase; CTI12\_AA419400, Geranylgeranyl reductase; CTI12\_AA226570, Uncharacterized protein; Belongs to the FPP/GGPP synthase family; CYP71AV1, Amorpha-4,11-diene 12-monooxygenase; CTI12\_AA049700, DBR2; DBR2, Artemisinic aldehyde Delta(11(13)) reductase; ALDH1, Aldehyde dehydrogenase 1; CTI12\_AA405060, Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid and CTI12\_AA461200, E-beta-farnesene synthase.

The PPI analysis of CPR represents 11 nodes and 49 edges (**Fig. 9D**). The predicted functional partners were CTI12\_AA054150, CTI12\_AA059330 and CTI12\_AA014270, APS reductase 3; CTI12\_AA390080 and CTI12\_AA322200; Nitrite/Sulfite reductase ferredoxin-like domain-containing protein; CTI12\_AA528 230, Uncharacterized protein; CTI12\_AA024450, Ferredoxin-nitrite reductase protein and CTI12\_AA597900, CTI12\_AA291950 and CTI12\_AA282720, Nitrate reductase. The PPI analysis of CYP71AV1 represents 11 nodes and 46 edges (**Fig. 9E**). The interacting proteins were CTI12\_AA054150, CTI12\_AA014270 and CTI12\_AA059330, APS reductase 3; CTI12\_AA597900, CTI12\_AA291950, CTI12\_AA282720, CTI12\_AA206760 and CTI12\_AA190820, Nitrate reductase; FPS1, Farnesyl pyrophosphate synthase and ALDH1 Aldehyde dehydrogenase 1. The PPI analysis of DBR2 represents 11 nodes and 14 edges (**Fig. 9F**). The predicted functional partners were CTI12\_AA369920, OPC-8:0 CoA ligase1; CTI12\_AA500690, CTI12\_AA305120, CTI12\_AA224500 and CTI12\_AA606960, Allene oxide cyclase; CTI12\_AA419030, AMP-dependent synthetase/ligase; CTI12\_AA603800, CTI12\_AA511950, CTI12\_AA505110 and CTI12\_AA221930 AMP-binding, conserved site- containing protein.



**Fig. 9 A-F.** Protein-protein interaction network for the enzymes involved in artemisinin biosynthesis.

#### 4. Discussion:

Over the past few years due to the development of resistance against other antimalarial drugs, artemisinin has seen a sharp rise in demand. However, natural artemisinin production is very low and it is difficult to synthesize chemically which results in a global supply scarcity. Thus, *Artemisia annua* remains the sole plant from which artemisinin is commercially available. Modern researches mainly target the enzymes involved in artemisinin biosynthetic pathways. The present work dealt with multiple sequence alignments, secondary structure analysis, tertiary structure analysis and protein-protein interactions of enzymes responsible for artemisinin synthesis. Our results showed that the percentage of helix is higher in all predicted secondary structures but exceptions were for ADH1 of *Artemisia annua* and *Helianthus annuus*. In both, the percentage of sheet is more than helix. The percentage of sheet was 71.4 for ADH1 of *Artemisia annua* which is more than the percentage of helix i.e., 63.7. Similarly, the percentage of sheet was 69.7 for ADH1 of *Helianthus annuus* which is more than the percentage of helix i.e., 56.2. The three-dimensional protein structure analysis showed the amino-acid sequences of enzymes shared 0.50% to 0.6% sequence similarity with the template

sequences. The stereochemical quality of the predicted structure (s) was evaluated using the PROCHECK server, and the results were validated using a Ramachandran plot. For the six enzymes under investigation, the Ramachandran plot results (**Fig. 8**) revealed a range of 89.1%–93.2% of amino acid residues inside the region that was most favoured.

### Conclusions:

In conclusion part of this study, template plants namely *Helianthus annuus*, *Matricaria chamomilla* var. *recutita* and *Tanacetum cinerariifolium* and *Artemisia absinthium* were used for comparative analysis along with *Artemisia annua*. The percentage of helix is higher in all predicted secondary structures but exceptions were for ADH1 of *Artemisia annua* and *Helianthus annuus*. In both, the percentage of sheeth is more than helix. An excellent quality model has more than 90% of its residues in the favoured region. In this paper, three-dimensional protein model of AMS1 showed 93.2% residues in the most favoured region. The protein-protein interaction outcome revealed a strong association with interacting proteins. This research might be a significant contribution in the field of Bioinformatics research and could aid other researchers in understanding the protein structures, physicochemical characteristics, and protein-protein interactions of enzymes involved in artemisinin production.

### References:

1. Adia MM, Emami SN, Byamukama R, Faye I, Borg-Karlson AK. (2016). Antiplasmodial activity and phytochemical analysis of extracts from selected Ugandan medicinal plants. *Journal of Ethnopharmacology*, 186, 14-19.
2. Banyai W, Kirdmanee C, Mii M, Supaibulwatana K. (2010). Overexpression of farnesyl pyrophosphate synthase (FPS) gene affected artemisinin content and growth of *Artemisia annua* L. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 103, 255-265.
3. Cox FE. (2010). History of the discovery of the malaria parasites and their vectors. *Parasites & vectors*, 3(1), 1-9.
4. Efferth T, Romero MR, Wolf DG, Stamminger T, Marin JJ, Marschall M. (2008). The antiviral activities of artemisinin and artesunate. *Clinical infectious diseases*, 47(6), 804-811.
5. Efferth T. (2017). From ancient herb to modern drug: *Artemisia annua* and artemisinin for cancer therapy. In *Seminars in cancer biology* (Vol. 46, pp. 65-83). Academic Press.
6. Efferth T. (2018). Beyond malaria: The inhibition of viruses by artemisinin-type compounds. *Biotechnology advances*, 36(6), 1730-1737.
7. Han J, Wang H, Kanagarajan S, Hao M, Lundgren A, Brodelius PE. (2016). Promoting artemisinin biosynthesis in *Artemisia annua* plants by substrate channeling. *Molecular Plant*, 9(6), 946-948.
8. Han JL, Liu BY, Ye HC, Wang H, Li ZQ, Li GF. (2006). Effects of overexpression of the endogenous farnesyl diphosphate synthase on the artemisinin content in *Artemisia annua* L. *Journal of Integrative Plant Biology*, 48(4), 482-487.
9. Jing F, Zhang L, Li M, Tang Y, Wang Y, Wang Y, Tang K. (2009). Abscisic acid (ABA) treatment increases artemisinin content in *Artemisia annua* by enhancing the expression of genes in artemisinin biosynthetic pathway. *Biologia*, 64, 319-323.
10. Liu M, Shi P, Fu X, Brodelius P E, Shen Q, Jiang W, Tang K. (2016). Characterization of a trichome-specific promoter of the aldehyde dehydrogenase 1 (ALDH1) gene in *Artemisia annua*. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 126, 469-480.
11. Martini MC, Zhang T, Williams JT, Abramovitch RB, Weathers PJ, Shell SS. (2020). *Artemisia annua* and *Artemisia afra* extracts exhibit strong bactericidal activity against *Mycobacterium tuberculosis*. *Journal of ethnopharmacology*, 262, 113191.
12. Munyangi J, Cornet-Vernet L, Idumbo M, Lu C, Lutgen P, Perronne C, Weathers P. (2018). Effect of *Artemisia annua* and *Artemisia afra* tea infusions on schistosomiasis in a large clinical trial. *Phytomedicine: international journal of phytotherapy and phytopharmacology*, 51, 233.

13. Nair MS, Huang Y, Fidock DA, Polyak SJ, Wagoner J, Towler M, Mishra S. (2021). *Artemisia annua* L. extracts prevent in vitro replication of SARS-CoV-2 (preprint).
14. Nondo RSO, Moshi MJ, Erasto P, Masimba PJ, Machumi F, Kidukuli AW, Zofou D. (2017). Anti-plasmodial activity of Norcaesalpin D and extracts of four medicinal plants used traditionally for treatment of malaria. *BMC complementary and alternative medicine*, 17, 1-8.
15. Pramanik K, Soren T, Mitra S, Maiti TK. (2017). In silico structural and functional analysis of Mesorhizobium ACC deaminase. *Computational Biology and Chemistry*, 68, 12-21.
16. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Keasling JD. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 440(7086), 940-943.
17. Simonsen HT, Nordskjold JB, Smitt UW, Nyman U, Palpu P, Joshi P, Varughese G. (2001). In vitro screening of Indian medicinal plants for antiplasmodial activity. *Journal of Ethnopharmacology*, 74(2), 195-204.
18. Simonsen HT, Weitzel C, Christensen SB. (2013). Guaianolide sesquiterpenoids: pharmacology and biosynthesis. *Natural products*, 5, 3069-3098.
19. Teoh KH, Polichuk DR, Reed DW, Covello PS. (2009). Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. *Botany*, 87(6), 635-642.
20. Teoh KH, Polichuk, DR, Reed DW, Nowak G, Covello PS. (2006). *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS letters*, 580(5), 1411-1416.
21. Wang B, Kashkooli AB, Sallets A, Ting HM, De Ruijter NC, Olofsson L, Van der Krol AR. (2016). Transient production of artemisinin in *Nicotiana benthamiana* is boosted by a specific lipid transfer protein from *A. annua*. *Metabolic Engineering*, 38, 159-169.
22. Zheng H, Colvin CJ, Johnson BK, Kirchhoff PD, Wilson M, Jorgensen-Muga K, Abramovitch RB (2017). Inhibitors of Mycobacterium tuberculosis DosRST signaling and persistence. *Nature chemical biology*, 13(2), 218-225.