Journal of Medicinal Plants Research Vol. 5(2), pp. 223-230, 18 January, 2011
Available online at http://www.academicjournals.org/JMPR
ISSN 1996-0875 ©2011 Academic Journals

Full Length Research Paper

The last gene involved in the MEP pathway of *Artemisia annua*: Cloning and characterization and functional identification

Meifang Peng[1,2], Min Chen[3], Rong Chen[1], Xiaozhong Lan[4], Minghsiun Hsieh[5], Zhihua Liao[1]*

[1]Key Laboratory of Eco-environments in Three Gorges Reservoir Region (Ministry of Education), Laboratory of Natural Products and Metabolic Engineering, Chongqing Sweet potato Research Center, School of Life Sciences, Southwest University, Chongqing 400715, China.
[2]Basic Laboratory for PLA medical center, Chengdu Military General Hospital, Chengdu 610083, China.
[3]School of Pharmaceutical Sciences, School of Life Sciences, Southwest University, Chongqing 400715, China.
[4]Agricultural and Animal Husbandry College, Tibet University, Linzhi of Tibet 860000, China.
[5]Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan.

Accepted 2 December, 2010

Hydroxymethylbutenyl 4-diphosphate reductase (HDR) catalyzes the last reaction of the methylerythritol phosphate (MEP) pathway for the biosynthesis of artemisinin precursors, a branching step that separately produces isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) in a ratio of 5:1 to 6:1. The full-length cDNA sequence of HDR was cloned and characterized from *Artemisia annua* L. for the first time. The new cDNA was designated as *AaHDR* (GenBank accession No.: GQ119345). The full-length cDNA of *AaHDR* was 1640 bp containing a 1368 bp open reading frame (ORF) encoding a polypeptide of 455 amino acids with a calculated molecular mass of 51.3 KDa and an isoelectric point of 5.63. Comparative and bioinformatic analysis revealed that AaHDR had extensive homology with HDRs from other plant species and contained a conserved transit peptide for plastids. The phylogenetic analysis indicated that all HDRs could be divided into three groups and *AaHDR* belonged to plant HDRs family. Then the homology-based structural modeling of *AaHDR* showed that *AaHDR* had the typical structure of HDR from *A. aeolicus*, which adopted a cloverleaf or trefoil-like structure with each monomer in the dimer containing three alpha/beta domains surrounding a central [Fe_3S_4] cluster ligated to Cys13, Cys96 and Cys193. Finally, *AaHDR* was transformed into the *E. coli* HDR mutant strain MG1655 ara< >HDR, which was able to rescue the lethal phenotype of the *E. coli* HDR mutant strain MG1655 ara-HDR. This confirmed that *AaHDR* had the typical function of HDR gene. The cloning and characterization of *AaHDR* will be helpful to understand more about the function of HDR at the level of molecular genetics and unveil the biosynthetic mechanism of artemisinin precursors. The present work also provides a candidate gene for metabolic engineering of the artemisinin biosynthesis pathway in *A. annua*.

Key words: *Artemisia annua* L., hydroxymethylbutenyl 4-diphosphate reductase, cloning, characterization, functional complementation.

INTRODUCTION

Malaria remains an important mortality for people in tropical region (Mercke et al., 2000). Artemisinin, a new and very potent antimalarial drug, is extracted from Chinese traditional medicinal herb *Artemisia annua* L. (Wallaart et al., 2001). It is a sesquiterpene lactone with an endoperoxide brigde and is actively against chloroquine-resistant and chloroquinine-sensitive forms of *Plasmodium falciparum*, as well as against cerebral malaria (Klayman, 1985). The supply of artemisinin is far from enough in the international market. The relative low yield (0.01 to 0.5%, DW) of artemisinin in wild-type *A.*

*Corresponding author. E-mail: zhliao@swu.edu.cn. Tel: +86-23-68367146. Fax: +86-23-68252365.

annua greatly limits the commercialization of this medicine (Wallaart et al., 2000). Attempts to obtain high-yielding strains by classical breeding and selection techniques have not improved artemisinin production significantly. Although artemisinin could be synthesized chemically, it could not be put into production due to high cost and strong toxicity. With the successful cloning of genes involved in the biosynthesis of artemisinin, genetic engineering is thought to be one of the most promising approaches to enhance the production of artemisinin.

Isopentenyl diphosphate (IPP) is the essential precursor of artemisinin biosynthesis which is provided by both mevalonic acid (MVA) pathway and methylerythritol phosphate (MEP) pathway. As is reported before, HDR is a key enzyme in MEP pathway which catalyzes the last reaction of biosynthesis of IPP. The cloning and characterization of *AaHDR* will be helpful to understand more about the function of HDR at the level of molecular genetics. It will also play an important role to unveil the biosynthetic mechanism of artemisinin precursors and provide a candidate gene for metabolic engineering of the artemisinin biosynthesis pathway in *A. annua*.

MATERIALS AND METHODS

Plant materials

All tissue materials including roots, stems, leaves, and flowers were excised from *A. annua* grown in the Medical Plants Garden of Southwest University, Chongqing, China. Plant tissues were immersed in liquid nitrogen immediately after excision and preserved in a -70℃ ultra low temperature refrigerator for RNA extraction.

Strain and plasmids

E. coli DH5α, M15 and plasmid pQE30 were used for the general construction of recombinant plasmids. *E. coli* HDR mutant strain MG1655 ara-HDR, which was used to confirm the typically function of HDR gene was kindly provided by Ming-Hsiun Hsieh.

Total RNA extraction

The RNA isolating kit purchased from TianGen (Beijing, China) was used for total RNA extraction. The quality and concentration of the RNA was checked by agarose gel electrophoresis and spectrophotometer (Shanghai, China) analysis. RNA samples were stored in -70℃ ultra low temperature refrigerator for future using.

Cloning of the core fragment of *AaHDR*

Firstly, Single-strand cDNAs were synthesized from 5 μg of total RNA with an oligo (dT)17 primer and reversely transcribed according to the manufacturer's protocol (PowerScriptTM, Clontech, USA). After RNase H treatment, the single-strand cDNA mixtures were used as templates for PCR amplification of the conserved region of HDR from *A. annua*. Then a core fragment of *AaHDR* was isolated with a pair of degenerate primers (FC: 5`-GT(C/T)GAGCG(C/T)GC(A/T/C)GT(G/T/C)CAGAT(G/T)G C-3`; RC: 5`-CATC(C/T)TC(A/C)AC(G/A/T)(A/G)CCTT(A/G)TC(A/G/C/T)GG-3`) from the cDNA of *A. annua* by standard gradient PCR amplification (from 50 to 60℃) on BioRad My Cycler (USA). The core fragment was amplified and subcloned into pGEM T-easy vector (Promega, USA), then transformed into *Escherichia coli* strain DH5a followed by sequencing. The core fragment was subsequently used to design the gene-specific primers for the cloning of the full-length cDNA of *AaHDR* by the technology of rapid amplification of cDNA ends (RACE).

Cloning of the full-length cDNA of *AaHDR* by RACE

Then RACE-ready cDNA for 5`-RACE of *A. annua* was acquired by the method supplied by BD SMART™ RACE cDNA Amplification Kit (CLONTECH, USA). Universal Primer A Mix (UPM, provided by Clontech) and HDR5-1 (5`-ATCTTGTCATCAGGGAACTG-3`) were used as primers for upstream sequence isolation. The reactions consisted of 5 cycles of 30 s at 94℃ and 3 min at 72℃, followed by 5 cycles of 94℃ for 30 s, 70℃ for 30 s and 72℃ for 3 min, then 25 cycles of 30 s at 94℃, 30 s at 68℃ and 3 min at 72℃ and then 6 min of extension at 72 ℃. And RACE-Ready cDNA for 3`-RACE of *A. annua* was acquired by the method supplied by TaKaRa RNA PCR Kit (AMV) Ver.3.0. M13 Primer M4 and HDR3-1 (5`-GCAAGAAAACTGGCTACCAA-3`) were used as primers for downstream sequence isolation. The PCR was conducted as the following procedures: 2 min at 94℃, followed by 32 cycles of 30 s at 94℃, 30 s at 56℃ and 1 min at 72℃, then 8 min of extension at 72℃. The first amplified PCR products (50-fold dilution) were respectively used as templates for the nested PCR amplification of the specific sequence of *AaHDR*. Nested Universal Primer (NUP, provided by Clontech) and HDR5-2 (5`-TCCTAGCTTCATAAGCAATCT-3`) were used as the nested primers for the nested PCR amplification of the upstream specific sequence. The nested PCR was carried out under the condition bellow: 1 min at 95℃, followed by 30 cycles of 30 s at 94℃, 30 s at 68℃ and 1 min at 72℃ and then 8 min of extension at 72℃. M13 Primer M4 and HDR3-2 (5`-GTTACATCTGGTGCTTCTAC-3`) were used as the nested primers for the nested PCR amplification of the downstream specific sequence. The nested PCR was carried out under the condition bellow: 2 min at 94℃, followed by 30 cycles of 30 s at 94℃ for 30 s at 56℃ and 1 min at 72℃, follow by 8 min of extension at 72℃. By 3`-RACE and 5`-RACE, both ends of *AaHDR* were respectively obtained. The PCR products were purified and subcloned into pGEM T-easy vector followed by sequencing.

After assembling the core fragment, the upstream and downstream sequences, the full-length cDNA sequence of *AaHDR* was deduced. According to the deduced *AaHDR* cDNA sequence, two gene-specific primers: Ffaahdr(5`-CGCGGGACACACAAAAACAC-3`) and Rfaahdr (5`-GTAACAAGTCGTTTATAGCAACC-3`) were used to amplified the full-length of *AaHDR* from 5`-RACE-ready cDNA samples through proof-reading PCR. All the PCR amplifications and sequencings for the cDNA of *AaHDR* were repeated three times to avoid PCR errors. Finally *AaHDR* was submitted to GenBank to be assigned with an accession number.

Bioinformatic analysis

Comparative and bioinformatics analysis of *AaHDR* were carried out online at the websites (http://www.ncbi.nlm.nih.gov and http://www.expasy.org). The nucleotide sequence, deduced amino acid sequence and ORF (open reading frame) encoded by *AaHDR* were analyzed and the sequence comparison was conducted through a database search using the BLAST program (Altschul et al., 1997).

The multiple alignments of AaHDR and HDRs from other plant species were aligned with CLUSTAL X (Thompson et al., 1997) using default parameters. A phylogenetic tree was constructed using MEGA version 3.0 (Kumar et al., 2004) from CLUSTAL X alignments. The neighbor-joining method (Saitou and Nei, 1987) was used to construct the tree. The homology-based 3-D structural modeling of AaHDR was accomplished by Swiss-Modeling. WebLab ViewerLite was used for 3-D structure displaying (Schwede et al., 2003; Arnold et al., 2006; Guex et al., 1997).The basic character of the HDR protein and the analysis of the transit peptide were analyzed using the bioinformatics software supplied by http://www.expasy.org..

Functional complementation

The endogenetic HDR gene of *E. coli* HDR mutant strain MG1655 ara-HDR was replaced by a kanamycin-resistant cassette and the whole chromosome just contained a single copy of HDR under the control of the P_{BAD} promoter (McAteer et al., 2001). The MEP pathway resided in *E. coli*, and the HDR gene was imperative for the survival of *E. coli* (Hsieh and Goodman, 2005). MG1655 ara-HDR was maintained on Luria-Bertani (LB) medium containing 50 mg·L^{-1} Kanamycin (Kan) and 0.2% (w/v) Arabinose (Ara) (McAteer et al., 2001), but not able to form colonies on LB medium containing 0.2% (w/v) Glucose (Glc) in the absence of Ara (McAteer et al., 2001). In order to detect the catalytic activity of AaHDR, we used the *E. coli.* MG1655 ara-HDR to make a complementation experiment.

The coding sequence of *AaHDR* was amplified by PCR using primers F-cdsaahdr (5`-CCGGATCCATGGCGTCTTTGCAGCTAAC-3`) and R-cdsaahdr (5`-CCGTCGACCTACACCAATTGCAGGGC-3`). Both of the fragment of *AaHDR* and the plasmid pQE30 were digested with *Bam*H I and *Sal* I for 10 h to construct recombinant expression vector pQE30-*AaHDR,* which was subsequently transformed into the *E. coli* HDR mutant strain MG1655 ara-HDR. As a control, the empty pQE30 vector was also transformed into the MG1655 ara-HDR. Then MG1655 ara-HDR containing the empty pQE30 vector and the pQE30-*AaHDR* were respectively inoculated on LB solid medium containing 50 mg·L^{-1}Kan, 50 mg·L^{-1} Amp, 0.2% Glc and 0.5 mM IPTG, and their growth was observed after being incubated for 12 h at 37°C.

RESULTS AND DISCUSSION

Cloning of the full-length cDNA of *AaHDR*

Based on the conserved fragment of other plants *HDR* sequences such as *Adonis palaestina* (AF270978), *Hevea brasiliensis* (AB294708), *Solanum tuberosum* (DQ252518), *Stevia rebaudiana* (DQ269451), *Ginkgo biloba* (DQ364231) and etc, two degenerate primers (FC and RC) were designed and used for amplification of the core fragment of *HDR* from *A. annua*. Following PCR amplification, an approximately 1000 bp product was obtained and sequenced. The BLAST search revealed that the 976 bp cDNA core fragment had high homologous with HDR genes from plant species such as *Nicotiana langsdorffii x Nicotiana sanderae*, *Hevea brasiliensis*, *Adonis palaestina* and etc. These strongly suggested that the core fragment of *AaHDR* had been obtained. Then, this fragment was used to design gene specific primers for both 3`-RACE and 5`-RACE. By nested 3`-RACE and 5`-RACE, the 331 bp 3`-end and 411 bp 5`-end of *A. annua* were respectively obtained. By aligning and assembling the sequences of 3`-RACE, 5`-RACEand the core fragment on Contig Express (Vector NTI Suite 8.0), the full-length cDNA sequence of *AaHDR* with 1640 bp was deduced. Finally the physical full-length *AaHDR* cDNA was amplified and confirmed by sequencing. Then, the full-length RvIspH sequence was submitted to GenBank and assigned an accession number: GQ119345.

Comparative and bioinformatic analysis of *AaHDR*

The (ORF) finding analysis on NCBI (http://www.ncbi.nlm.nih.gov/gorf.html) showed that the *AaHDR* contained a 1368-bp coding sequence encoding a 455-amino-acid polypeptide with a calculated molecular mass of 51.3 kDa and an isoelectric point of 5.63 (http://www.expasy.org/tools/protscale.html). The *AaHDR* contained a 35-bp 5`-UTR, 194-bp 3`-UTR and 17-bp polyA, and its termination codon was TAG (Figure 1). The deduced amino acid sequence of *AaHDR* was submitted to NCBI for BLAST searching and the results showed that *AaHDR* had high similarities with HDRs from other plant species, such as *Stevia rebaudiana* (82% identities), *Vitis vinifera* (80% identities), *Catharanthus roseus* (78% identities). Thus, the BLAST analysis results indicated that AaHDR belonged to the HDR family. The subcellular prediction analysis by TargetP and ChloroP suggested that *AaHDR* had a 34-amino acid sequence with characteristics of plastidial targeting sequences at its N-terminal end, which was consistent with the fact that MEP pathway located in the plastid (Lichtenthaler et al., 1997). Secondary structure prediction of *AaHDR* was done on expasy network station using GOR4 (Garnier et al., 1996), which indicated that *AaHDR* contained 33.85% α-helixes, 41.62% random coils and 21.54% extended strands.

A sequence comparison was made among *AaHDR* with HDR from other plants and *E. coli*. It indicated that all the HDRs of plants had a plastidial transit peptide at the N terminus, but the HDR protein of *E. coli* lacked the N-terminal extension (Figure 2).

Further analysis indicated that all the plant HDRs had four conserved cysteine residues which were presumed to play an important role in the coordination of the iron-sulfur bridge proposed to be involved in catalysis (Seemann et al., 2002; Wolff et al., 2003). The position of one of these cysteine residues was not conserved in the *E. coli* protein, so it was conjectured that the third active site just took part in the iron-sulfur bridge's coordination in plants.

The phylogenetic tree was constructed by using MEGA 3.0 which based on CLUSTAL X alignments. The phylogenetic analysis showed that HDRs were derived from an ancestor gene and evoluted into three groups

1 cgcgggacacacaaaaacacacgcagcataattca

36

ATGGCGTCTTTGCAGCTAACACCTCTGTCAACTCGCACAGACTACCTCTCCTTACC

TGCAGACATAAAAGTATTCCGGTGCCGGAAGCCG

M A S L Q L T P L S T R T D Y L S L P A D I K V F R C R K P

126

TTAACAGTCCGATGCTCCGGCGGTGACACGTCATCGTCAACGCAATTTGATGCGA

AGGTGTTCAGGCATAATTTGACAAGGAGCGAGAAT

L T V R C S G G D T S S S T Q F D A K V F R H N L T R S E N

216

TATAATAGGAAAGGATTTGGTCATAAGAAGGAGACTCTTGAGCTCATGAGTCAGGA

GTATTTTAGCGACATTATAAAGACTTTGAAGGAG

Y N R K G F G H K K E T L E L M S Q E Y F S D I I K T L K E

306

AATAACTACGAATATACATGGGGAAATGTCACTGTAAAGCTTGCAGAAGCTTTTGG

TTTTTGTTGGGGTGTTGAGCGTGCCGTCCAGATT

N N Y E Y T W G N V T V K L A E A F G F C W G V E R A V Q I

396

GCTTATGAAGCTAGGAAACAGTTCCCTGATGACAAGATATGGATCACAAATCAAAT

TATTCACAACCCTACTGTTAACAAGAGGCTAGAA

A Y E A R K Q F P D D K I W I T N Q I I H N P T V N K R L E

486

GAGATGGAAGTTACGGATATCCCCATTGACGGCGGAGAGAAACAGTTTGATGTTG

TTGACAAGGGCGATGTTGTAATTCTGCCTGCCTTT

E M E V T D I P I D G G E K Q F D V V D K G D V V I L P A F

576

GGAGCTGCAGTAGACGAGATGCGGATTTTGAGTAACAAAGAAGTACAAATAGTCG

ATACAACATGCCCATGGGTGACTAAGGTGTGGAAT

G A A V D E M R I L S N K E V Q I V D T T C P W V T K V W N

Figure 1. The full-length cDNA of *AaHDR*. The coding sequence and the deduced amino acid sequence were shown in capital lettersin bold font and the stop codon was marked with a star; the UTR was shown in small letters. The plastidial transit peptide was underlined.

including plants, algaes and bacteria HDR group (Figure 3), and *AaHDR* belongs to plant HDR branch which originated from the same ancestor. As to the evolution of bacterial HDRs, the evolution of blue-green algae was so conservative that it formed an independent branch. The previous investigation had indicated that *DXR* is the only gene which originates from blue-green algae among all the genes involved in the MEP pathway (Bick and Lange, 2003). While another research supposed that the HDR gene and *DXR* gene had the same origin in plants (Guevara-Garcia et al., 2005). In our present research, we had poor proof to indicate that the HDR gene of plants was evolved from blue-green algae or other bacteria due to the sophistication of evolution of species. However, our analysis indicated that *AaHDR* comes from plants and it was an important enzyme of the MEP pathway.

The homology-based 3-D structural modeling of *AaHDR* was analyzed by Swiss-Modeling and displayed by WebLab ViewerLite, using *Aquifex aeolicus*'s HDR tertiary protein structure as a control (Figure 4). The result indicated that the tertiary protein structure of HDR from *A. annua* and *A. aeolicus* were extremely similar.

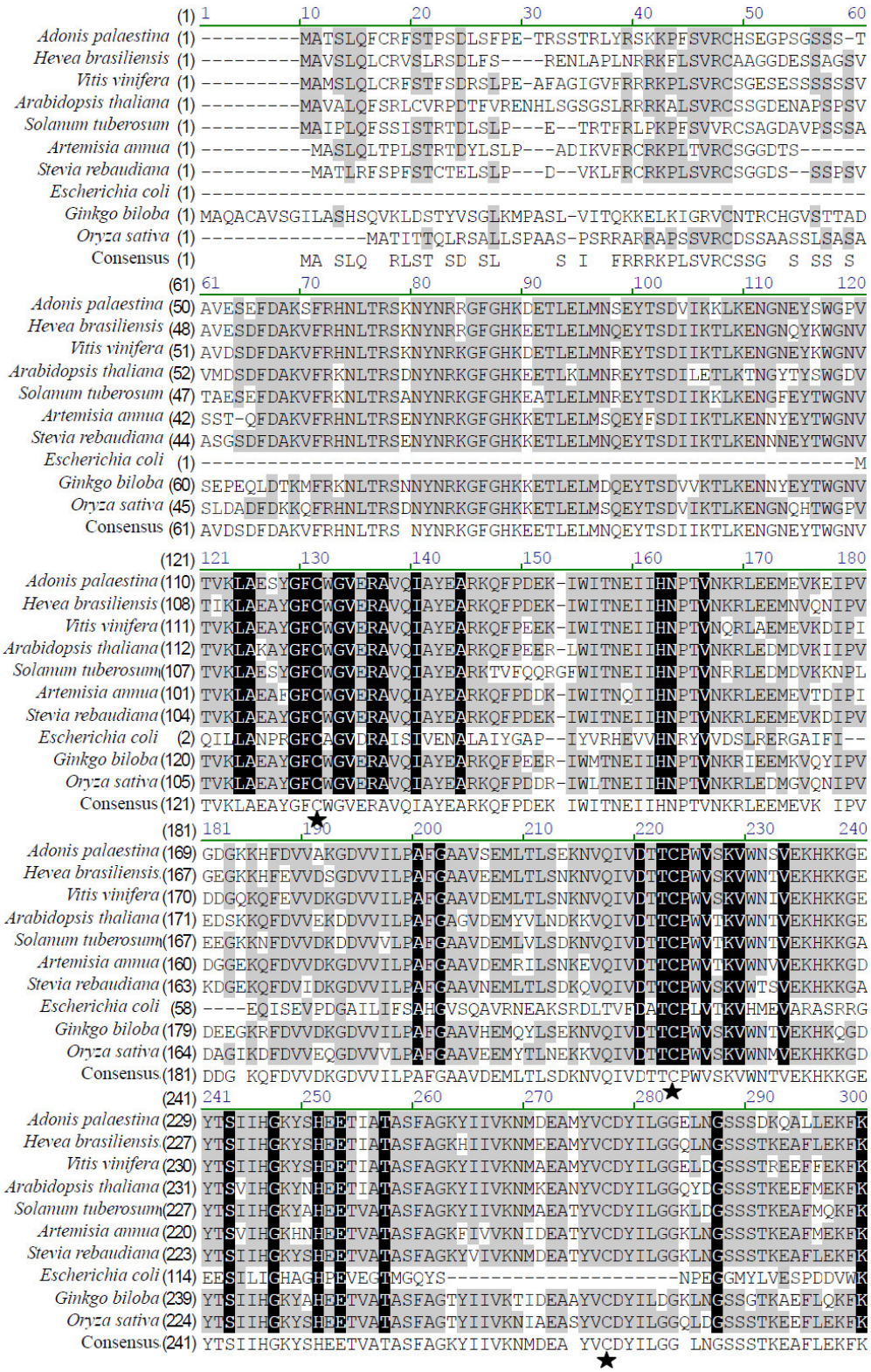

Figure 2. Multi-alignment of amino acid sequences of HDR from plants and *E. coli* The identical amino acids were showed in white with black background and the conserved amino acids were showed in black with gray background, other amino acids were showed in black with white background. Stars mark the position of conserved cysteine residues. The sequences used were listed bellow with Accession number: *Adonis palaestina*, AAG21984.1; *Hevea brasiliensis*, BAF98297.1; *Vitis vinifera*, CAO47671.1; *Arabidopsis thaliana*, AAN87171.1; *Solanum tuberosum*, ABB55395.1; *Stevia rebaudiana*, ABB88836.2; *Escherichia coli*, AAL38655.1; *Ginkgo biloba*, ABC84344.1; *Oryza sativa*, NP_001051167.1.

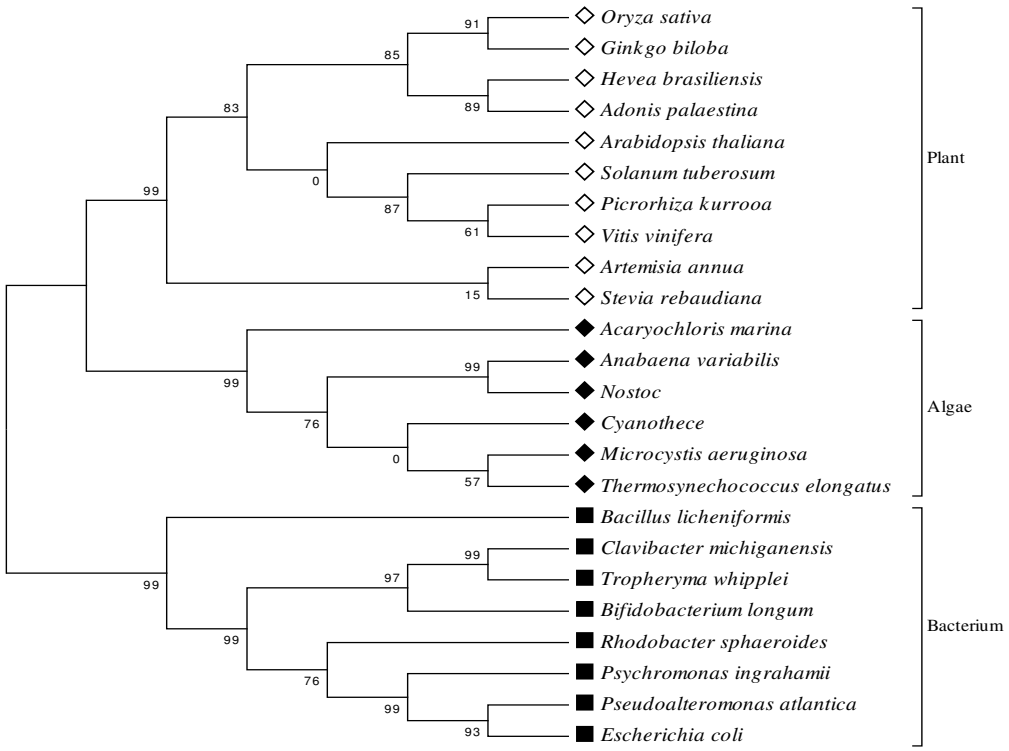

Figure 3. A phylogenetic tree of HDR from different organisms constructed by neighbor-joining method on MEGA 3.0.Plant-derived HDSs were marked with◇, Algae-derived HDSs were marked with◆, Bacterium-derived HDSs were marked with■. The sequences used were listed bellow with Accession number: *Acaryochloris marina*, YP_001519239.1; *Thermosynechococcus elongatus*, NP_681832.1; *Microcystis aeruginosa*, CAO90213.1; *Cyanothece*, ZP_01731309.1; *Anabaena variabilis*, YP_323455.1; *Nostoc*, NP_485028.1; *Bacillus licheniformis*, YP_079844.1; *Clavibacter michiganensis*, YP_001222973.1; *Tropheryma whipplei*, CAD67323.1; *Bifidobacterium longum*, NP_696525.1; *Rhodobacter sphaeroides*, ABN75415.1; *Psychromonas ingrahamii*, ABM04955.1; *Pseudoalteromonas atlantica*, ABG41681.1.

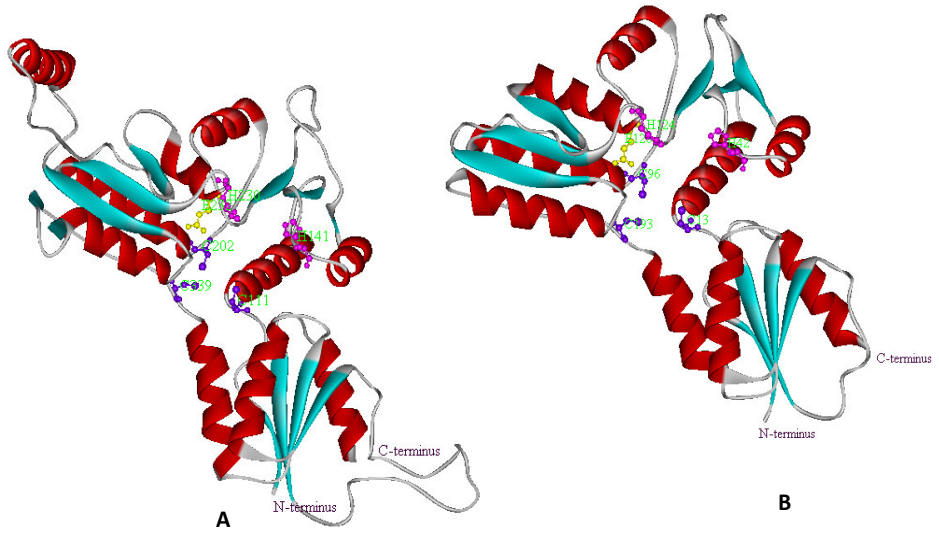

Figure 4. The 3-D structure of AaHDR and Aquifex aeolicus HDR established by homology-based modeling (Swiss-Model) The α-helix, β-sheet and random coil were showed in red (helix-shaped), blue (sheet-shaped), grey (rope-shaped) respectively;the conserved and activated sites including Cys13-Cys96 Cys193、His42-His124、Glu126 were showed in different color balls.

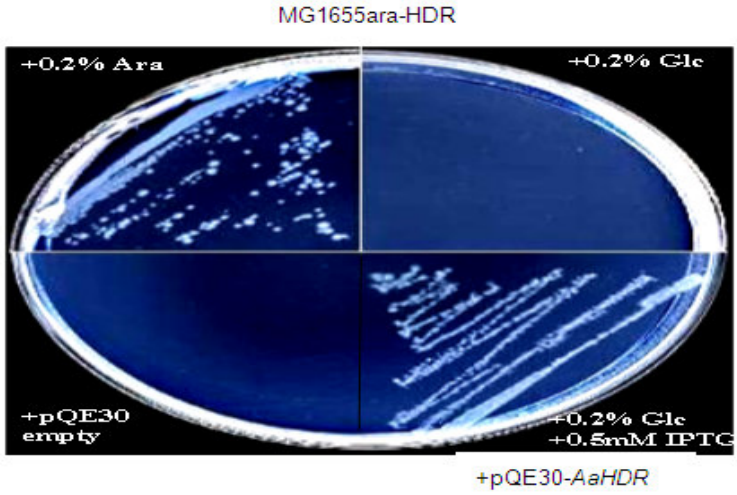

Figure 5. Functional demonstration of *AaHDR* activity. The *E. coli* HDR mutant strain MG1655ara-HDR was able to grow on LB media containing 0.2% Ara, but not on media containing 0.2% Glc (above). After transformation with the *AaHDR* coding region (pQE30-*AaHDR*) and, as a control, with the empty vector (pQE30) alone, the resulting strains were tested for growth on media containing 0.2% Glc and 0.5 mM IPTG (below). Expression of AaHDR protein successfully restored the growth of the *E. coli* HDR mutant (below right).

AaHDR had 10 α-helixes and 10 β-sheets, while the HDR of *A. aeolicus* only had nine α-helixes. *AaHDR* had the typical structure of HDR from *A. aeolicus* (Rekittke et al., 2008) which adopted a cloverleaf or trefoil-like structure with each monomer in the dimer containing three alpha/beta domains surrounding a central [Fe_3S_4] cluster ligated to Cys13, Cys96, and Cys193. Two highly conserved His (His 42 and His 124) and a totally conserved Glu (Glu126) are located in the same central site and are proposed to be involved in ligand binding and catalysis. Substrate access is proposed to occur from the front-side face of the protein, with the HMBPP diphosphate binding to the two His and the 4OH of -4-Hydroxy-3-methyl-but-2-enyl pyrophosphate (HMBPP) binding to the fourth iron thought to be present in activated clusters, while Glu126 provides the protons required for isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) formation. All the bioinformatic analysis proved that *AaHDR* was a functional gene encoding the HDR enzyme.

AaHDR Complements the *E. coli* HDR Mutant

To test whether the HDR protein of *A. annua* has similar enzymatic activity to its *E. coli* counterpart, we used a complementation assay with an *E. coli HDR* mutant strain MG1655 ara-HDR, which could only grow on the media containing Ara, while it could not grow on the media containing Glc. The prokaryotic expression vector of *AaHDR* was constructed successfully, and had been checked by digested with *Bam*H I and *Sal* I. When recombinated expression plasmid pQE30-*AaHDR* introduced the *AaHDR* gene into the *E. coli* MG1655 ara-HDR, it could grow on the media containing Glc, while negative control strain (with empty pQE30 vector) could not grow (Figure 5). So, the genetic complementation assay demonstrated that *AaHDR* is able to recover the function of HDR gene in *E. coli* MG1655 ara-HDR and indeed encodes a functional enzyme of HDR.

Conclusions

The last step of the MEP pathway is catalyzed by HDR (formerly designated as LytB or IspH), which converts HMBPP into IPP or DMAPP. Now little more work is still ahead to analyze the HDR's role in the control of the flux of intermediates through the pathway and the supply of IPP and DMAPP for the synthesis of plastid isoprenoid end products. The *Synechocystis* HDR gene and a HDR cDNA from the flowering plant *Adonis aestivalis* were each found to significantly enhance accumulation of carotenoids (Cunningham et al., 2000). Using tobacco rattle virus (TRV) to post-transcriptionally silence the expression of HDR in the *Nicotiana benthamiana* plants had albino leaves that contained less than 4% of the chlorophyll and carotenoid pigments of control leaves (Page et al., 2004). Double transgenic Arabidopsis plants overproducing both tomato HDR and Arabidopsis taxadiene synthase (TXS) showed a 13-fold higher increase in the accumulation of taxadiene. All the studies showed that HDR was an important enzyme and may be an ideal target for metabolic engineering of the isoprenoid biosynthetic pathway. In our present study, we have

successfully isolated and characterized the *AaHDR* cDNA from *A. annua* for the first time.

The full-length cDNA of *AaHDR* was 1640 bp containing a 1368 bp ORF encoding a polypeptide of 455- amino acid with a calculated molecular mass of 51.3 KDa and an isoelectric point of 5.63. The further Comparative and bioinformatic analysis revealed that *AaHDR* had extensive homology with HDRs from other plant species and contained a conserved transit peptide for plastids with a 34-amino acid at the N terminus, which was consistent with the fact that MEP pathway located in the plastid (Lichtenthaler et al., 1997). *AaHDR* and all the plant HDRs had four conserved cysteine residues, which were presumed to play an important role in the coordination of the iron-sulfur bridge proposed to be involved in catalysis (Seemann et al., 2002; Wolff et al., 2003). The phylogenetic analysis indicated that all HDRs could be divided into three groups and *AaHDR* belonged to plant HDRs family. Then the homology-based structural modeling of *AaHDR* showed that *AaHDR* had the typical structure of HDR from *A. aeolicus*, which adopted a cloverleaf or trefoil-like structure with each monomer in the dimer containing three alpha/beta domains surrounding a central [Fe_3S_4] cluster ligated to Cys13, Cys96, and Cys193. Two highly conserved His (His 42 and His 124) and a totally conserved Glu (Glu126) was located in the same central site and was proposed to be involved in ligand binding and catalysis. Finally, the *AaHDR* was been introduced into *E.coil* HDR mutant strain MG1655 ara-HDR and overexpressed, and then the mutant was rescued by *AaHDR* in the media. The result of genetic complementation demonstrated that *AaHDR* gene reported here did encode the active enzyme HDR.In summary, a functional gene encoding *HDR* (*AaHDR*) was cloned, characterized and functionally identified from *A. annua*. It will facilitate the unveiling of the biosynthesis of Artemisinin at the level of molecular biology and provide a candidate gene for metabolic engineering of Artemisinin.

ACKNOWLEDGEMENTS

This research was financially supported by the China National 863 High-Tech program (Project No. 2010AA100503) and NSFC project (31070266).

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 25: 3389-3402.

Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. Bioinformat., 22: 195-201.

Bick JA, Lange BM (2003). Metabolic crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. arch. biochem. biophy., 415: 146-154.

Cunningham FX, Jr Lafond TP, Gantt E (2000). Evidence of a role for LytB in the nonmevalonate pathway of isoprenoid biosynthesis. J.Bacteriol., 182: 5841-5848.

Garnier J, Gibrat JF, Robson B (1996). GOR secondary structure prediction method version IV. Methods Enzymol., 266: 540-553.

Guevara-Garcia A, San Roman C, Arroyo A, Cortes ME, Gutierrez-Nava M de la Luz, Leon P (2005). Characterization of the Arabidopsis clb6 mutant illustrates the importance of post transcriptional regulation of the methyl-D-erythritol 4-phosphate pathway. Plant Cell., 17: 628-643.

Guex N, Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis, 18: 2714-2723.

Hsieh MH, Goodman HM (2005). The arabidopsis IspH homolog is involved in the plastid nonmevalonate pathway of isoprenoid biosynthesis. Plant physiol., 138: 641-653.

Klayman DL (1985). Qinghaosu (artemisinin): an antimalarial drug from China. Science, 228: 1049-1055.

Kumar S, Tamura K, Nei M (2004). MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. Brie Bioinform, 5:150-163.

Lichtenthaler HK, Rohmer M, Schwender J (1997). Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. Plant Physiol., 101: 643-652.

McAteer S, Coulson A, McLennan N (2001). Masters M, The lytB gene of *Escherichia coli* is essential and specifies a product needed for isoprenoid biosynthesis. J Bacterio., 183 7403–7407

Page JE, Hause G, Raschke M, Raschke M, Gao WY, Schmidt J, Zenk MH, Kutchan TM (2004). Functional Analysis of the Final Steps of the 1-Deoxy-D-xylulose 5-phosphate (DXP) Pathway to Isoprenoids in Plants Using Virus-Induced Gene Silencing. Plant Physiol., 134: 1-13.

Mercke P, Bengtsson M, Bouwmeester HJ, Maarteen AP, Peter EB (2000). Molecular Cloning, Expression, and Characterizating of Amorpha-4,11-diene Synthase, a Key Enzyme of Artemisinin Biosynthesis in *Artemisia annua* L. Archives Biochem. Biophys., 381: 173-180.

Rekittke I, Wiesner J, Demmer U, Warkentin E, Xu W, Troschke K, Hintz M, No JH, Duin EC, Oldfield E, Jomaa H, Ermler U (2008). Structure of (E)-4-hydroxy-3-methyl-but-2-enyl diphosphate reductase, the terminal enzyme of the non-mevalonate pathway. J. Am. Chem. Soc., 130: 17206-17207.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol., 4: 406-425.

Schwede T, Kopp J, Guex N, Peitsch MC (2003). SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res., 31: 3381-3385.

Seemann M, Bui BT, Wolff M, Tritsch D, Campos N, Boronat A, Marquet A, Rohmer M (2002). Isoprenoid biosynthesis through the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) is a [4Fe-4S] protein [J]. Angew. Chem. Int. Ed. Engl., 41: 4337-4339.

Wallaart TE, Bouwmeester HJ, Hille J, Poppinga L, Maijers NCA (2001). Amopha-4, 11-diene sythase: clonging and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalatial drug artemisinin. Planta., 212: 460-465.

Thompson LD, Gibson T J, Plewniak F, Jeanmougin F, Higgins DG (1997). The Clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res., 24: 4876-4882.

Wallaart TE, Pras N, Quax WJ (2000). Seasonal variations of artemisinin and its biosynthetic precursors in tetraploid *Artemisia annua* plants compared with the diploid wild-type. Planta., 66: 57-62.

Wolff M, Seemann M, Tse Sum Bui B, Frapart Y, Tritsch D, Estrabot AG, Rodriguez-Concepcion M, Boronat A, Marquet A, Rohmer M (2003). Isoprenoid biosynthesis via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from Escherichia coli is a [4Fe-4S] protein, FEBS Lett., 541: 115–120.