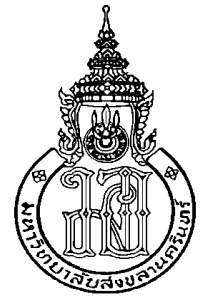


Molecular Cloning and Expression of cDNAs Encoding the Genes in Early Steps of Isoprenoid Biosynthesis *via* the 2*C*-Methyl-D-Erythritol 4-Phosphate Pathway in *Hevea brasiliensis* Müll. Arg.

Yortyot Seetang-Nun

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Biochemistry

Prince of Songkla University

2008

Thesis Title	**Molecular Cloning and Expression of cDNAs Encoding the Genes in Early Steps of Isoprenoid Biosynthesis *via* the 2*C*-Methyl-D-Erythritol 4-Phosphate Pathway in *Hevea brasiliensis* Müll. Arg.**
Author	**Mr. Yortyot Seetang-Nun**
Major Program	**Biochemistry**

Major Advisor:

……

(Assoc. Prof. Dr. Wallie Suvachittanont)

Co-advisor:

……..

(Prof. Dr. Thomas D. Sharkey)

Examining Committee:

……. Chairperson

(Prof. Dr. Prapon Wilairat)

……...Committee

(Assoc. Prof. Dr. Wallie Suvachittanont)

……...Committee

(Prof. Dr. Thomas D. Sharkey)

……...Committee

(Assoc. Prof. Dr. James R. Ketudat-Cairns)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Biochemistry

……

(Assoc. Prof. Dr. Krerkchai Thongnoo)

Dean of Graduate School

ชื่อวิทยานิพนธ์ การโคลนและการแสดงออกของยีนในช่วงต้นของการสังเคราะห์ไอโซปรีนอยด์โดยวิถี 2*C*-Methyl-D-Erythritol 4-Phosphate ในยางพารา

ผู้เขียน นายยอดยศ ศรีตังนันท์

สาขาวิชา ชีวเคมี

ปีการศึกษา 2550

บทคัดย่อ

ในพืชชั้นสูงการสังเคราะห์สาร IPP และ DMAPP ซึ่งเป็นโครงสร้างพื้นฐานของสารกลุ่ม isoprenoid มี 2 วิถีคือ วิถี mevalonic acid (MVA) ซึ่งพบได้ใน cytoplasm และมีความเกี่ยวข้องกับการสังเคราะห์สาร sesquiterpenes steroids และ ubiquinones ในขณะที่วิถี 2*C*-methyl-D-erythritol 4-phosphate (MEP) ที่พบใน plastids มีความเกี่ยวข้องกับการสังเคราะห์สาร isoprene monoterpenes diterpenes carotenoids chlorophylls และ plastoquinones สารตัวกลาง เอนไซม์ และยีนต่างๆ ในวิถี MEP ถูกค้นพบเมื่อไม่นานมานี้ ปฏิกิริยาขั้นแรกของวิถี MEP เป็นการรวมตัวระหว่าง pyruvate และ glyceraldehyde 3-phosphate (GAP) ได้สาร 1-deoxy-D-xylulose 5-phosphate (DXP) เป็นผลิตภัณฑ์ โดยอาศัยการทำงานของเอนไซม์ DXP synthase (DXS) ซึ่งถูกควบคุมการผลิตโดยยีน *DXS* ในขั้นที่สองของปฏิกิริยา DXP จะถูกเปลี่ยนไปเป็น MEP ซึ่งเป็นสารตัวแรกที่จำเพาะกับวิถี MEP โดยอาศัยการทำงานของเอนไซม์ DXP reductoisomerase (DXR) ซึ่งถูกควบคุมการผลิตโดยยีน *DXR*

การศึกษาครั้งนี้ มีวัตถุประสงค์เพื่อทำการโคลนและศึกษาคุณลักษณะของยีนที่ควบคุมการผลิตเอนไซม์ DXS และ DXR ในยางพารา โดยวิธี reverse transcription-polymerase chain reaction (RT-PCR) และวิธี rapid amplification reaction of cDNA ends (RACE) ผลการศึกษา พบว่า ยีน *DXS* มี 2 ยีน คือ ยีน *HbDXS1* และยีน *HbDXS2* ในทำนองเดียวกัน ยีน *DXR* ก็มี 2 ยีน คือ ยีน *HbDXR1* และยีน *HbDXR2* โดยที่ยีน *HbDXS1* และยีน *HbDXR1* โคลนได้จากใบ ในขณะที่ยีน *HbDXS2* และยีน *HbDXR2* โคลนได้จากน้ำยาง

ยีน *HbDXS1* มีขนาด 2163 คู่เบส สามารถถอดรหัสได้โปรตีนที่มีขนาด 720 กรดอะมิโน มีมวลโมเลกุลขนาด 78 kDa และค่า pI เท่ากับ 7.66 ในขณะที่ยีน *HbDXS2* มีขนาด 2136 คู่เบส ซึ่งสามารถถอดรหัสได้โปรตีนที่มีขนาด 711 กรดอะมิโน มีมวลโมเลกุลขนาด 76 kDa และค่า pI เท่ากับ 7.34 ยีน *HbDXS1* และยีน *HbDXS2* มีความเหมือนกัน 62% ในระดับดีเอ็นเอและ 69% ในระดับโปรตีน โปรตีน HbDXS1 และโปรตีน HbDXS2 มีส่วนประกอบที่สำคัญของเอนไซม์ DXS เช่น TPP-binding motif (GDG-X_{25}-NDN) และ transketolase motif (DRAG-X_{28}-P-X-D) รวมทั้งกรดอะมิโนที่มีความสำคัญต่อการเร่งปฏิกิริยาของเอนไซม์ DXS การศึกษาความสัมพันธ์เชิงวิวัฒนาการ พบว่า โปรตีน HbDXSs ทั้งสองจากยางพารา จัดอยู่ในกลุ่มเดียวกับโปรตีน DXS จากพืชชนิดต่างๆ แต่โปรตีน HbDXS1 และโปรตีน HbDXS2 จัดอยู่ในกลุ่มย่อยที่ต่างกัน โดยที่โปรตีน HbDXS1 มีความใกล้ชิดกับโปรตีน DXS จากพืชชนิดต่างๆที่จัดอยู่ในกลุ่มที่ 1 ในขณะที่โปรตีน HbDXS2 จะมีความใกล้ชิดกับโปรตีน DXS จากพืชชนิดต่างๆที่จัดอยู่ในกลุ่มที่ 2 โครงสร้างตติยภูมิของโปรตีน HbDXS1 และโปรตีน HbDXS2 ที่ทำนายได้ มีรูปแบบที่คล้ายคลึงกับโครงสร้างตติยภูมิของโปรตีน DXS จากแบคทีเรีย *Deinococcus radiodurans* ที่ใช้เป็นแม่แบบ

ส่วนยีน *HbDXR1* และยีน *HbDXR2* มีขนาดเท่ากันคือ 1416 คู่เบส สามารถถอดรหัสได้โปรตีนที่มีขนาด 471 กรดอะมิโน มีมวลโมเลกุลขนาด 51 kDa และค่า pI เท่ากับ 5.67 สำหรับโปรตีน HbDXR1 และเท่ากับ 6.04 สำหรับโปรตีน HbDXR2 ยีน *HbDXR1* และยีน *HbDXR2* มีความเหมือนกัน 94% ทั้งในดีเอ็นเอและในระดับโปรตีน โปรตีน HbDXR1 และโปรตีน HbDXR2 มีส่วนประกอบที่สำคัญของเอนไซม์ DXR เช่น Proline rich region [$P_{2\text{-}4}$AWPG(R/T)] และ NADPH-binding motif (GXXGXXG) รวมทั้งกรดอะมิโนต่างๆ ที่มีความสำคัญต่อการเร่งปฏิกิริยาของเอนไซม์ DXR การศึกษาความสัมพันธ์เชิงวิวัฒนาการ พบว่า โปรตีน HbDXRs ทั้งสองจากยางพารา จัดอยู่ในกลุ่มเดียวกับโปรตีน DXR จากพืชชนิดต่างๆ และโปรตีนทั้งสองมีความใกล้ชิดกันเองมากกว่าโปรตีน DXR จากพืชชนิดอื่นๆ การวิเคราะห์ดีเอ็นเอด้วยวิธี Southern blot แสดงให้เห็นว่ามียีน *DXR* อย่างน้อยสองยีนในยางพารา โครงสร้างตติยภูมิของโปรตีน HbDXR1 และโปรตีน HbDXR2 ที่ทำนายได้มีรูปแบบที่คล้ายคลึงกับโครงสร้างตติยภูมิของโปรตีน DXR จากแบคทีเรีย *Escherichia coli* ที่ใช้เป็นแม่แบบ

การศึกษาการแสดงออกของยีน *HbDXSs* และยีน *HbDXRs* โดยวิธี semi-quantitative RT-PCR พบว่า ยีนแต่ละยีนมีการแสดงออกที่แตกต่างเมื่อศึกษาในเนื้อเยื่อต่างชนิดกันทั้งในต้นกล้าและต้นที่โตเต็มวัย ยีน *HbDXS1* และ *HbDXR1* มีการแสดงออกมากในเนื้อเยื่อที่เกี่ยวข้องกับการสังเคราะห์แสง เช่น ใบ ในขณะที่ยีน *HbDXS2* และ *HbDXR2* มีการแสดงออกมากในเนื้อเยื่อที่ไม่เกี่ยวข้องกับการสังเคราะห์แสง เช่น ราก ดอก และน้ำยาง ยกเว้นยีน *HbDXS2* ซึ่งมีการแสดงออกมากในใบ การแสดงออกของยีน *HbDXS2* และยีน *HbDXR2* ในยางพารา 2 สายพันธุ์ คือ พันธุ์ RRIM 600 ซึ่งจัดเป็นพันธุ์ที่ให้ผลผลิตสูง และพันธุ์พื้นเมือง พบว่า ระดับการแสดงออกของยีน *HbDXS2* ในยางพาราทั้ง 2 สายพันธุ์ ไม่มีความแตกต่างกัน ในขณะที่ระดับการแสดงออกของยีน *HbDXR2* ในยางพารา RRIM 600 สูงกว่าพันธุ์พื้นเมือง และมีความสัมพันธ์กับเนื้อยางแห้ง เมื่อศึกษาผลของ ethephon ต่อการแสดงออกของยีน *HbDXS2* และยีน *HbDXR2* ในยางพาราพันธุ์พื้นเมือง พบว่า ระดับการแสดงออกของยีนทั้งสองเพิ่มขึ้นชั่วขณะ ในขณะที่ปริมาณน้ำยางยังคงเพิ่มขึ้นตลอดช่วงของการทดลอง

โปรตีน HbDXS1 และโปรตีน HbDXR1 สามารถแสดงออกได้ในแบคทีเรีย *E. coli* แต่พบการแสดงออกในรูปของ inclusion bodies โปรตีนทั้งสองชนิดได้ถูกทำให้บริสุทธิ์ภายใต้สภาวะ denaturation เมื่อวิเคราะห์ด้วย sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) พบว่า โปรตีน HbDXS1 มีขนาดประมาณ 70 kDa ส่วนโปรตีน HbDXR1 มีขนาดประมาณ 40 kDa เมื่อนำโปรตีน HbDXR1 ไปทำให้กลับคืนสภาพ และหาค่ากิจกรรมเอนไซม์ พบว่า โปรตีน HbDXR1 ที่ได้ไม่มีค่ากิจกรรมเอนไซม์ นอกจากนี้เมื่อนำโปรตีนทั้งสองชนิดไปผลิต polyclonal antibody ในการศึกษาเบื้องต้น พบว่า polyclonal antibody ที่ได้จากโปรตีน HbDXS1 สามารถตรวจจับโปรตีน HbDXS1 ที่ทำบริสุทธิ์ได้ แต่ไม่สามารถตรวจจับโปรตีน DXS จากยางพาราและจากพืชชนิดต่างๆ ได้ ในขณะที่ polyclonal antibody ที่ได้จากโปรตีน HbDXR1 สามารถตรวจจับได้ทั้งโปรตีน HbDXR1 ที่ทำบริสุทธิ์และโปรตีน DXR จากยางพาราและจากพืชชนิดต่างๆ ได้

for the synthesis of sesquiterpenes, steroids, and ubiquinones, while the 2*C*-methyl-D-erythritol 4-phosphate (MEP) pathway occurs in the plastids and is involved in the synthesis of isoprene, monoterpenes, diterpenes, carotenoids, chlorophylls, and plastoquinones. The MEP pathway has only been identified in the past few years and its intermediates, enzymes, and genes have recently been characterized. The first step of the MEP pathway is condensation of pyruvate and glyceraldehyde 3-phosphate (GAP) to form 1-deoxy-D-xylulose 5-phosphate (DXP), catalyzed by DXP synthase (DXS), encoded by *DXS* gene. The second step is the conversion of DXP to MEP, the first intermediate specific to the MEP pathway. This reaction is catalyzed by DXP reductoisomerase (DXR) encoding by *DXR* gene.

The present study aimed to clone and characterize cDNAs encoding DXS and DXR from *Hevea brasiliensis*. The cloning strategy was based on reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) methods. As a result, two cDNAs for DXS (*HbDXS1* and *HbDXS2*) and two cDNAs for DXR (*HbDXR1* and *HbDXR2*) were obtained. The *HbDXS1* and *HbDXR1* were isolated from leaves, while *HbDXS2* and *HbDXR2* were from the latex.

The *HbDXS1* cDNA contained an open reading frame (ORF) of 2163 bp encoding a polypeptide of 720 amino acids with a calculated molecular mass of 78 kDa and pI value of 7.66, while the *HbDXS2* cDNA contained an ORF of 2136 bp encoding a polypeptide of 711 amino acids with a predicted molecular mass of 76 kDa and pI value of

7.34. These cDNAs showed 62% and 69% identities to each other at the nucleotide and amino acid levels, respectively. The deduced amino acid sequences of both HbDXS1 and HbDXS2 contained all typical characteristics of DXS, such as a TPP-binding motif (GDG-X_{25}-NDN) and a transketolase motif (DRAG-X_{28}-P-X-D). Several invariant amino

acids important for DXS activity were also conserved in both HbDXSs. Phylogenetic analysis showed that both HbDXSs are grouped within a plant DXS family but in different groups. The HbDXS1 is more closely related to other plant DXSs belonging to DXS class I, while the HbDXS2 is close to class II of the plant DXS family. The predicted three dimensional (3D) structures of both HbDXSs were similar to that of the *D. radiodurans* DXS template.

The *HbDXR1* and *HbDXR2* cDNAs contained an ORF of 1416 bp encoding a polypeptide of 471 amino acids with a predicted molecular mass of about 51 kDa. The predicted pI values of HbDXR1 and HbDXR2 were 5.67 and 6.04, respectively. The *HbDXRs* exhibited 94% identities to each other at both nucleotide and amino acid levels. The deduced amino acid sequences of both HbDXRs contained all typical characteristics of DXR, such as a Proline-rich region [$P_{2\text{-}4}$AWPG(R/T)] and an NADPH-binding motif (GXXGXXG). Several invariant amino acids important for DXR activity were also conserved in both HbDXRs. Phylogenetic analysis showed that both HbDXRs are grouped within the plant DXR family and they are more closely related to each other than to other plant species. Southern blot analysis showed that there are at least two *DXR* genes in *H. brasiliensis*. The predicted 3D structures of both HbDXRs showed an overall folding pattern consistent with the *E. coli* DXR template.

Analysis of expression of both *HbDXS* and *HbDXR* genes by semi-quantitative RT-PCR in both seedlings and mature plants showed that both *HbDXS* and *HbDXR* mRNAs were differentially expressed in different tissues of *H. brasiliensis*. Transcripts of both *HbDXS1* and *HbDXR1* were highly expressed in photosynthetic tissues, such as leaves, while the transcripts of both *HbDXS2* and *HbDXR2* were highly expressed in non-photosynthetic tissues, such as roots, flowers, and latex, with an exception of a higher expression of *HbDXR2* transcript in leaves. In different

clones of *H. brasiliensis*, the expression level of *HbDXS2* transcripts showed no difference between clone RRIM 600 (a high-latex yielding clone) and wild type (a low-latex yielding clone), while the level of *HbDXR2* transcript showed higher expression in clone RRIM 600 than

in the wild type plant and seemed to be correlated with the latex yield as judged by dry rubber content. Upon ethephon treatment, the levels of both *HbDXS2* and *HbDXR2* transcripts were transiently induced, while the rubber yield continued to increase.

Both HbDXS1 and HbDXR1 were successfully expressed in *E. coli* but as inclusion bodies. These proteins were purified under denaturing condition using nickel-based affinity chromatography. By sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis, the purified recombinant proteins showed a major band corresponding to the molecular weight of about 70 kDa (for HbDXS1) and 40 kDa (for HbDXR1), consistent with their molecular sizes predicted from their nucleotide sequences. Refolding of the purified recombinant HbDXR1 protein failed to yield an active enzyme, as detected by an activity assay. These two purified recombinant proteins were also used for antibody production. Preliminary results showed that anti-HbDXS1 polyclonal antibody can detect neither the putative DXS protein from *H. brasiliensis* nor DXS proteins from other plants. In contrast, anti-HbDXR1 polyclonal antibody detected a putative DXR protein from *H. brasiliensis* and cross-reacted with DXR proteins from other plants.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Wallie Suvachittanont, for her continuous encouragement and never-ending optimism throughout these years. I am most grateful to my other supervisor, Prof. Dr. Thomas D. Sharkey, for his enthusiasm and providing such excellent facilities to do the research at the Department of Botany, University of Wisconsin-Madison. I would like to express my gratitude to my examination committee: Prof. Dr. Prapon Wilairat and Assoc. Prof. Dr. James R. Ketudut-Cairns for their helpful suggestions and invaluable comments throughout the examination.

I also wish to express my respect to Assoc. Prof. Dr. Porntip Prapunpoj, Assoc. Prof. Dr. Wilaiwan Chotigeat, Dr. Pluang Suwanmanee, Dr. Nualpun Sirinupong, and Dr. Tanya G. Falbel for their expert technical assistance. I would like to thank Dr. Rapiporn Sotthibandhu for reading some parts of the thesis. I also would like to thank Asst. Prof. Dr. Juraithip Wungsintaweekul for her inspiring discussion regarding the biosynthesis of isoprenoid.

All staff members of the Department of Biochemistry, Faculty of Science, the Scientific Equipment Center, Prince of Songkla University, and the members of Department of Botany, University of Wisconsin-Madison deserve my thanks for helping me. The biggest thanks go to all my friends for their unconditional friendship. Finally, I would like to express my deepest gratitude to my families for their love, understanding, and encouragement throughout the study of Ph.D. program and especially for my life.

This work was supported by the Royal Golden Jubilee Graduate Program (PHD/0115/2543) from the Thailand Research Fund, and by the Graduate Research Fund from the Department of Biochemistry, Faculty of Science, Prince of Songkla University. Work in Madison was supported by the US NSF by grants IBN-0212204 and IOB-0640853.

Yortyot Seetang-Nun

CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS AND SYMBOLS

AFLP	=	amplified fragment length polymorphism
ATP	=	adenosine-5′-triphosphate
bp	=	base pair
BLAST	=	Basic Local Alignment Search Tool
BSA	=	bovine serum albumin
cDNA	=	DNA complementary to RNA
cTP	=	chloroplast transit peptide
CDP-ME	=	4-diphosphocytidyl-2*C*-methyl-D-erythritol
CDP-Star	=	disodium 2-chloro-5-(4-methoxyspiro{1,2-dioxetane-3,2′-(5′-chiro)tricyclo[3.3.1.1$_{3,7}$]decan}-4-yl)-1-phenyl phosphate-Star
CIP	=	calf intestinal alkaline phosphate
CMK	=	CDP-ME kinase
CMS	=	CDP-ME synthase
CTAB	=	cetyl trimethyl ammonium bromide
3D	=	three dimension(al)
dATP	=	deoxyadenosine-5′-triphosphate
dNTP	=	deoxynucleotide-5′-triphosphate
DEPC	=	diethylpyrocarbonate
DMAPP	=	dimethylallyl diphosphate
DNA	=	deoxyribonucleic acid
DNase I	=	deoxyribonuclease I
DRC	=	dry rubber content
DTT	=	dithiothreitol
dUTP	=	deoxyuridine-5′-triphosphate
DX	=	1-Deoxy-D-xylulose
DXP	=	1-Deoxy-D-xylulose 5-phosphate
DXR	=	DXP reductoisomerase
DXS	=	DXP synthase

LIST OF ABBREVIATIONS AND SYMBOLS (continued)

EDTA	=	ethylenediamine-*N*,*N*,*N*′,*N*′-tetraacetic acid
Ethephon	=	2-chloroethylphosphonic acid
ESTs	=	expressed sequence tags
ExPASy	=	Expert Protein Analysis System
FCA	=	Freund's Complete Adjuvant
FIA	=	Freund's Incomplete Adjuvant
FPP	=	farnesyl diphosphate
g	=	gram
g	=	gravity
GAP	=	glyceraldehyde 3-phosphate
GFP	=	green fluorescence protein
GGPP	=	geranylgeranyl diphosphate
GSP	=	gene-specific primer
GUS	=	β-glucuronidase
h	=	α-helice
h	=	hour
HDR	=	HMBPP reductase
HDS	=	HMBPP synthase
His-Tag	=	hexahistidyl-tag
HMBPP	=	1-hydroxy-2-methyl-butenyl 4-diphosphate
HMGCoA	=	3-hydroxy-3-methylglutaryl-coenzyme A
HMGR	=	HMGCoA reductase
HMGS	=	HMGCoA synthase
IC_{50}	=	50% inhibitory concentration
IDI	=	IPP isomerase
IgG	=	immunoglobulin G
IPP	=	isopentenyl diphosphate
IPTG	=	isopropyl-β-D-thiogalactopyranoside
kb	=	kilobase

LIST OF ABBREVIATIONS AND SYMBOLS (continued)

kDa	=	kilodalton
k_{cat}	=	catalytic constant
*K*i	=	kinetic inhibitor
K_m	=	Michaelis constant
l	=	liter
LB	=	Luria-Bertani
M	=	molar
mA	=	milliampare
MCS	=	MECDP synthase
ME	=	2*C*-methyl-D-erythritol
MECDP	=	2*C*-methyl-D-erythritol 2,4-cyclodiphosphate
MEGA	=	Molecular Evolutionary Genetics Analysis
MEP	=	2*C*-methyl-D-erythritol 4-phosphate
min	=	minute
µg	=	microgram
µl	=	microliter
µM	=	micromolar
ml	=	milliliter
mM	=	millimolar
Mw	=	molecular weight
mRNA	=	messenger RNA
MVA	=	mevalonic acid
MVK	=	MVA kinase
ng	=	nanogram
nm	=	nanometer
nM	=	nanomolar
NADH	=	nicotinamide adenine dinucleotide
NADPH	=	reduced β-nicotinamide adenine dinucleotide phosphate
NCBI	=	National Center for Biotechnology Information

LIST OF ABBREVIATIONS AND SYMBOLS (continued)

Ni-NTA	=	nickle-nitrotriacetic acid
OD	=	optical density
ORF	=	open reading frame
pfu	=	plaque forming unit
pH	=	-log hydrogen ion concentration
pI	=	isoelctric point
pmole	=	picomole
PAGE	=	polyacrylamide gel electrophoresis
PBS	=	phosphate buffer saline
PBST	=	phosphate buffer saline plus Tween 20
PCR	=	polymerase chain reaction
PDB	=	Protein Data Bank
PMD	=	MVA diphosphate decarboxylase
PMK	=	phosphomevalonate kinase
PMSF	=	phenylmethanesulphonyl fluoride
rpm	=	revolution per minute
RLM-RACE	=	RNA ligase-mediated rapid amplification of 5´- and 3´-ends
RNA	=	ribonucleic acid
RNase	=	ribonuclease
RNase A	=	ribonuclease A
RNase H	=	ribonuclease H
rRNA	=	ribosomal RNA
RRIM	=	Rubber Research Institute of Malaysia
RSMD	=	root-mean-square deviation
RT-PCR	=	reversed transcription and PCR
s	=	second
s	=	β-strand
SDS	=	sodium dodecyl sulfate

LIST OF ABBREVIATIONS AND SYMBOLS (continued)

SSC	=	standard saline citrate
Tm	=	melting temperature
TAP	=	tobacco acid pyrophosphatase
TAE	=	tris acetate EDTA
TEMED	=	*N,N,N′,N′*-Tetramethylethylenediamine
TIGR	=	The Institute for Genomic Research
TPP	=	thiamine diphosphate
Tris	=	Tris(hydroxymethyl)aminomethane
U	=	unit
UTR	=	untranslated region
UV	=	ultraviolet
V	=	volt
v/v	=	volume per volume
w/v	=	weight per volume
X-gal	=	5-bromo-4-chloro-3 indolyl-β-D-thiogalactopyranoside
α	=	alpha
β	=	beta
Å	=	Ångström
ºC	=	degree Celsius
%	=	percentage

CHAPTER 1

INTRODUCTION

Introduction

Hevea brasiliensis Müll. Arg., an important economic plant originated in the Amazon basin but it is currently mainly cultivated in Southeast Asia, especially in Thailand, Malaysia, and Indonesia. *H. brasiliensis* has been recognized as a main source of natural rubber production, 99% of the global natural rubber supply, although more than 2000 species of plants can produce natural rubber (Steinbüchel, 2003).

Nowadays, Thailand has become the largest natural rubber producer and exporter country, accounting for 35% of production worldwide, ahead of Malaysia and Indonesia. In 2003, about 1.6 million hectares of land were used for plantation of *H. brasiliensis*, which produces more than 2.4 million tons of natural rubber. In 2005, Thailand exported about 3 million tons of natural rubber, accounting for around 215 billion bahts (The International Rubber Research and Development Board, 2005; Office of the Rubber Replanting Aid Funds, 2006). Because the demand for natural rubber has steadily increased since the beginning of the 20th century (Steinbüchel, 2003), the genetic improvement of *H. brasiliensis* with increased yields, disease resistance, and adaptation to different environments needs to be investigated and developed.

Natural rubber (*cis*-1,4-polyisoprene), a high-molecular weight of polyisoprene, consists of almost entirely of five-carbon (C_5) isoprene units in *cis*-configuration and is present in the latex of *H. brasiliensis* (for review see Wititsuwannakul and Wititsuwannakul, 2001). It is produced and stored in specialized cells called laticifers, which are located in the outer bark overlaying the cambium of *H. brasiliensis* (de Faÿ and Jacob, 1989). When the bark was severed, the latex flows freely. Latex contains all cell constituents plus rubber particles and the specific organelles: lutoids and Frey Wyssling particles. About 30-50% (w/w) of *cis*-1,4-polyisoprene were found in the latex of *H. brasiliensis* (Kekwick, 1989). Secretion of the latex has been hypothesized as a part of defense system against pest and/or pathogen attacks (Farrell *et al.*, 1991). However, the exact physiological roles of latex remain unclear.

The biosynthesis of natural rubber is considered to be consisted of two sections. First, isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP) are synthesised from central intermediates. Second, these two components are polymerized to a polyisoprene molecule of high molecular weight (Steinbüchel, 2003). It is now established that there are two pathways for the formation of IPP and DMAPP in plants: the mevalonic acid (MVA) pathway and the 2*C*-methyl-D-erythritol 4-phosphate (MEP) pathway (Eisenreich *et al.*, 1998; Lichtenthaler, 1999; Eisenreich *et al.*, 2001; Rodríguez-Concepción and Boronat, 2002; Eisenreich *et al.*, 2004). So far, only the MVA pathway has been considered to be responsible for the formation of IPP and DMAPP in *H. brasiliensis* (for a review see Kekwick, 1989). Since the biosynthesis of natural rubber is believed to be similar to the biosynthesis of cholesterol in animals (Balasubramaniam *et al.*, 1977), the first two enzymes of the MVA pathway: 3-hydroxy-3-methylglutaryl-CoA (HMGCoA) synthase (HMGS) and HMGCoA reductase (HMGR) have been shown to regulate the natural rubber biosynthesis (Lynen, 1969). Activities of these two enzymes were found to correlate with the latex yield (Wititsuwannakul *et al.*, 1988; Suvachittanont and Wititsuwannakul, 1995). In addition, their corresponding genes were more highly expressed in the latex than in leaves (Kush *et al.*, 1990; Suwanmanee *et al.*, 2002; Sirinupong *et al.*, 2005), and they were also more highly expressed in a high-yielding rubber clone than in the low-yielding clone (Wititsuwannakul, 1988; Suwanmanee *et al.*, 2004).

Since two pathways for IPP and DMAPP biosynthesis exist in plants, it would be of great interest to know whether the MEP pathway plays any role in isoprenoid biosynthesis related to natural rubber biosynthesis. Recently, the presence of the MEP pathway in *H. brasiliensis* was experimentally supported. Ko *et al*. (2003) used the expressed sequence tags (ESTs) and cDNA-amplified fragment length polymorphism (AFLP) analysis to study the transcriptome of *H. brasiliensis* latex, and found the transcripts of 1-deoxy-D-xylulose 5-phosphate (DXP) synthase (DXS), the first enzyme of the MEP pathway in the latex. The *DXS* transcripts were more highly expressed in the latex than in the leaves. This implies that the MEP pathway occurs in the latex of *H. brasiliensis* and therefore, it could provide an alternative means of generating IPP for the biosynthesis of natural rubber. The identified *DXS* gene is, however, available only as a partial cDNA and the relation between the MEP pathway genes and the natural rubber biosynthesis has not yet been investigated.

So far, studies of the MEP pathway in the control of isoprenoid biosynthesis in plants have been focused on DXS and DXP reductoisomerase or DXR (the second enzyme of the MEP pathway). Several experiments suggest that DXS plays a key role in a regulatory step of the MEP pathway (Bouvier *et al.*, 1998; Chahed *et al.*, 2000; Lois *et al.*, 2000; Veau *et al.*, 2000; Walter *et al.*, 2000; Estevez *et al.*, 2001; Han *et al.*, 2003; Hans *et al.*, 2004; Khemwong and Suvachittanont, 2005). However, the reaction catalyzed by DXR is actually the first step specific to the MEP pathway, because DXP is a precursor not only for isoprenoid biosynthesis but also for the formation of vitamins (B_1 and B_6) (Himmeldirk *et al.*, 1996). The role of DXR in the control of plant isoprenoid biosynthesis has been shown. For example, overexpression of *DXR* in transgenic peppermints led to increased essential oil monoterpenes synthesis (Mahmoud and Croteau, 2001).

In the present work, two cDNAs encoding DXS and DXR were cloned from *H. brasiliensis* using reverse transcription-polymerase chain reaction (RT-PCR) based techniques. There are two full-length cDNAs encoding each form of *DXS* and *DXR* gene, which were designated as *HbDXS* and *HbDXR*. The *HbDXS1* and *HbDXR1* cDNAs were isolated from leaves, while the *HbDXS2* and *HbDXR2* cDNAs were isolated from latex. The expression of these genes in different tissues and in different clones of *H. brasiliensis* was investigated. In addition, the effect of ethephon on the expression of *HbDXS2* and *HbDXR2* genes was also examined to see whether these genes play any role in rubber biosynthesis. Expression of recombinant HbDXS and HbDXR proteins in *E. coli* was also studied. The expressed HbDXS and HbDXR proteins were purified and polyclonal antibodies to the enzymes were raised to detect the presence of these proteins in samples.

Review of literatures

1. Introduction to isoprenoids

Isoprenoids (also called terpenoids) form one of the largest and most diverse groups of natural products found in all living organisms. More than 30,000 known structures of isoprenoid compounds have been reported and their number is growing steadily to date. Isoprenoids are not only numerous but also extremely variable in structure, ranging from the linear to polycyclic molecules and in size from five-carbon hemiterpenes to compound containing thousands of isoprene units such as natural rubber (Sacchettini and Poulter, 1997).

Isoprenoids have primary functions in cells such as plastoquinone and ubiquinone serve as electron transporters in photosynthesis and respiration, sterols as structural components of cytosolic membranes, and abscisic acid, cytokinins and gibberrellins as plant growth regulators. However, most of isoprenoids identified to date do not have basic roles in cells, but are known as secondary metabolites. For example, some isoprenoids are involved in interactions between environment and organisms including protection against pathogens, and attractants for pollinators and seed-dispersing animals. Several isoprenoids are considered to have potentials in medicinal application as drugs. For example, paclitaxel or taxol from *Taxus spp.* was used as anticancer, artemisinin from *Artemisia annua* as antimalarial, and plaunotol from *Croton stellatopilosus* as antiulcer agents. Some isoprenoids are nutritionally important for human health, e.g. lycopene, lutein, and carotene were recently registered as oncopreventive agents. In addition, many isoprenoids have been used in many industries as raw materials or additive for foods, perfumes, beverages, and other products.

In spite of the diversity in structural and functional, all isoprenoids are synthesized from common five-carbon building units, IPP and its allylic isomer DMAPP, the universal metabolic precursors. To date, groups of isoprenoids are classified according to the number of isoprene units present in the structures. The simplest isoprenoids contain a single C_5 unit called hemiterpene. The more complex isoprenoids are usually formed by head-to-tail or by head-to-head addition of isoprene units. For example, monoterpenes (C_{10}) consist of two isoprene units, sesquiterpenes (C_{15}) contain three isoprene units, diterpenes (C_{20}) contain four isoprene units, triterpenes (C_{30}) contain six isoprene units, tetraterpenes (C_{40}) contain eight isoprene units, and polyterpenes contain more than eight isoprene units, respectively (Table 1).

Table 1.1 Classification of isoprenoids based on the number of isoprene units in the structures

Isoprene units (n)	Carbon atoms	Names	Examples
1	5	Hemiterpenes	Isoprene
2	10	Monoterpenes	Geraniol
3	15	Sesquiterpenes	Farnesol
4	20	Diterpenes	Gibberellic acid
6	30	Triterpenes	Phytosterols
8	40	Tetraterpenes	Carotenoids
>8	>40	Polyterpenes	Natural rubber

2. Pathways of isoprenoid biosynthesis

Since the discovery of the MVA pathway in *Saccharomyces cerevisiae* (yeast) and animals in the 1950s, it has long been accepted that the MVA pathway is a universal route for the biosynthesis of IPP and DMAPP in all organisms. In many cases, however, experimental data based on the biosynthesis of specific isoprenoids in plants and in some microorganisms could not be explained in terms of the MVA pathway. For example, mevalonate itself is a very poor precursor of many isoprenoids (Charlwood and Banthorpe, 1978). However, at that time, there is no reasonable alternative to the MVA pathway prior to the discovery of a novel pathway of isoprenoid biosynthesis (Rohmer *et al.*, 1993). To date, it is now firmly established that there are two distinct pathways (the MVA pathway and the MEP pathway) used for the formation of IPP and DMAP in the living organisms (Eisenreich *et al.*, 1998; Lichtenthaler, 1999; Eisenreich *et al.*, 2001; Rodríguez-Concepción and Boronat, 2002; Eisenreich *et al.*, 2004) (Fig. 1.1).

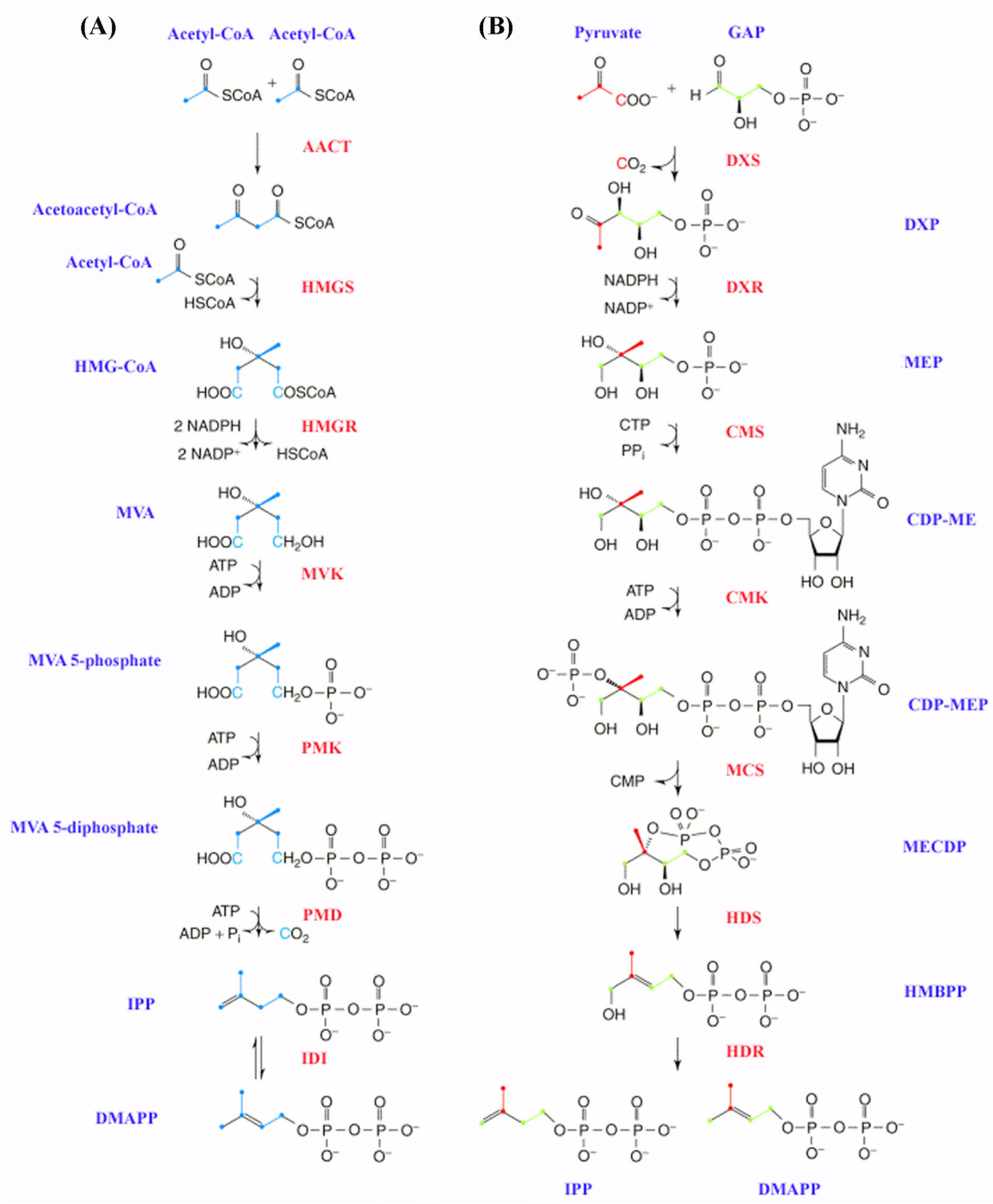


Fig. 1.1 The pathway for IPP and DMAPP formation.

(A) The MVA pathway. (B) The MEP pathway. The origins of carbon atoms from acetyl-CoA, pyruvate, and GAP are indicated in blue, red, and green, respectively. Enzymes of the pathways are labeled in red color (adapted from Eisenreich *et al.*, 2001).

2.1 The mevalonic acid (MVA) pathway

The biosynthesis of IPP and DMAPP by the MVA pathway has been elucidated by studies of yeast and animal cells (for review see Goldstein and Brown, 1990). The pathway contains six enzymatic steps (Fig. 1.1). First, a condensation of two molecules of acetyl CoA, catalyzed by acetoacetyl CoA thiolase, gives acetoacetyl CoA. Then, an aldol addition of another acetyl CoA molecule to the acetoacetyl CoA yields 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). This reaction is catalyzed by HMG-CoA synthase (HMGS). In the next step, HMG-CoA is reduced to mevalonic acid, catalyzed by HMG-CoA reductase (HMGR). Then, the two sequential phosphorylations of mevalonic acid by the enzyme MVA kinase (MVK) and phosphomevalonate kinase (PMK) produce MVA 5-phosphate and MVA 5-diphosphate, respectively. The MVA 5-diphosphate is subsequently decarboxylated and dehydrated to form IPP. The reaction is catalyzed by MVA diphosphate decarboxylase (PMD). IPP is further converted to DMAPP by IPP isomerase (IDI). Since the HMGR catalyzes the rate-limiting step in the MVA pathway, it is becoming an important target for the development of inhibitors (statins), which are widely used for the prevention and therapy of cardiovascular disease by lowering the cholesterol levels in human blood (Stancu and Sima, 2001).

2.2 The 2*C*-methyl-D-erythritol 4-phosphate (MEP) pathway

Evidence for the existence of a mevalonate-independent pathway was discovered by two independent research groups. (1) The experiment by Flesch and Rohmer (1988) found that the bacterial liopanoids, such as aminobacteriohopanetriol from *Rhodopseudomonas palustris*, supplied with the ^{13}C-acetate revealed unexpected labelling patterns in the terpenoid moiety, which were inconsistent with the MVA origin, and (2) The experiment by Arigoni and co-workers demonstrated that the administration of ^{13}C-glucose samples into the isoprenoid side chain of ubiquinone by *Escherichia coli*, and into the ginkgolides of *Gingko biloba* seedlings could not be explained in terms of the MVA pathway (for review see Eisenreich *et al.*, 1998; Dubey *et al.,* 2003). These results led to the discovery of a novel non-MVA pathway (Rohmer *et al.*, 1993). At present, this pathway was named as MEP pathway after the currently identified first committed precursor of the pathway, 2*C*-methyl-D-erythritol 4-phosphate.

The MEP pathway is well studied in *E. coli*. It consists of seven enzymatic steps (Fig. 1.1) (Rodríguez-Concepción and Boronat, 2002). This pathway starts with the formation of 1-deoxy-D-xylulose 5-phosphate (DXP) from glyceraldehydes 3-phosphate (GAP) and pyruvate, catalyzed by DXP synthase (DXS). In the next step, isomerization

and reduction of DXP gives 2*C*-methyl-D-erythritol 4-phosphate (MEP), catalyzed by DXP reductoisomerase (DXR). Since DXP can be used for the synthesis of vitamins B_1 (thiamin) and B_6 (pyridoxol) (Himmeldirk *et al.*, 1996), MEP is actually the first committed intermediate of the isoprenoid biosynthesis by the MEP pathway. Subsequently, the MEP is converted to the CDP-derivative by the following three-independent steps, catalyzed by 4-diphosphocytidyl-2*C*-methyl-D-erythritol (CDP-ME) synthase (CMS), CDP-ME kinase (CMK), and 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate (MECDP) synthase (MCS), respectively. Finally, MECDP is transformed to IPP and DMAPP, which are catalyzed by 1-hydroxy-2-methyl-butenyl 4-diphosphate (HMBPP) synthase (HDS) and HMBPP reductase (HDR), respectively. Because the MEP pathway has not been detected in animals and fungi (Disch and Rohmer, 1998), it offers an attractive target for the development of new antibacterials, antimalarials, and herbicides (Rohmer, 1998).

3. Evolution of the two isoprenoid biosynthetic pathways

The evolution of the isoprenoid biosynthetic pathways can be addressed by their distribution in different taxonomic kingdoms (Table 1.2). Animals are shown to use the MVA pathway for synthesis of their isoprenoids (Goldstein and Brown, 1990). Thus, homologues for the MVA pathway genes are found, i.e. in the human (Venter *et al.*, 2001) and *Drosophila melanogaster* genomes (Celniker and Rubin, 2003). In the entirely sequenced genomes of yeast (Giaever *et al.*, 2002) and *Schizosaccharomyces pombe* (Wood *et al.*, 2002), homologous sequences for all the MVA pathway genes are found, but there is no evidence for the presence of the DXP pathway genes are also present. By biosynthetic labeling studies, it has been shown that isoprenoid compounds from *Rhodotorula glutinis* and fungal *Aschersonia aleyrodis* are synthesized only by the MVA pathway (Disch and Rohmer, 1998). Also, archaea appear to utilize only the MVA pathway to synthesize their isoprenoids (Bult *et al.*, 1996; Klenk *et al.*, 1997; Smith *et al.*, 1997).

None of the mevalonate pathway genes are present in genomes of some eubacteria, for example, in *E. coli*, *Helicobacter pylori*, *Haemophilus influenzae*, and in *Bacillus subtilis* (Eisenreich *et al.*, 1998), suggesting the possibility that the MEP pathway is widespread in eubacteria. However, some eubacteria such as *Actinoplanes* sp. A40644 (Seto *et al.*, 1998) and *Streptomyces* sp. CL160 (Seto *et al.*, 1996) use both MVA and MEP pathways for synthesis of their isoprenoids, while some eubacteria such as *δ*-

proteobacterium *Myxococcus fulvus* (Rosa Putra *et al.*, 1998) and phototrophic eubacterium *Chloroflexus aurantiacus* (Rieder *et al.*, 1998) use only the MVA pathway for synthesis of their isoprenoids.

Most plants and algae possess both the MVA and MEP pathways but these pathways appear to be operated in different compartments. The MVA pathway is present in the cytosol, whereas the MEP pathway is present in the plastids. The MVA pathway is thought to be responsible for the production of sterols, sesquiterpenes, and ubiquinones, while the MEP pathway is involved in the biosynthesis of hemi-, mono-, and diterpenes as well as carotenoids (Lichtenthaler, 1999; Rodríguez-Concepción and Boronat, 2002). The unicellular green algae, for example *Scenedesmus obliquus*, *Chlamydomonas reinhardtii*, and *Chlorella fusca*, are the exceptions and seem to have completely lost the MVA pathway, whereas the rhodophyte *Cyanidium caldarum* and the heterokontophyte *Ochromonas danica* possess both the MEP and MVA pathways. Among the photosynthetic eukaryotes, *Euglena gracilis* uses only the MVA pathway for the synthesis of its isoprenoids (Schwender *et al.*, 1996; Disch *et al.*, 1998). In contrast, cyanobacteria have been shown to be utilized only the MEP pathway for synthesis of their isoprenoids (Lichtenthaler, 1999).

It is generally accepted that the plastids of higher plants originated by the endocytosis and subsequent the endosymbiosis of cyanobacteria (McFadden, 1999). More than 95% of cyanobacterial genes were subsequently lost or transferred into nucleus of the host cell during evolution (Martin and Herrmann, 1998). Recent studies have shown that the evolution of MEP pathway has been driven by a lateral gene transfer (Boucher and Doolittle, 2000; Lange *et al.*, 2000). Evidence for the lateral gene transfer has been inferred from the nucleotide composition, codon bias, unusual species distributions of genes, sequence similarity, and phylogenetic analysis. This event is also a major driving force of evolution of prokaryotic genomes (Doolittle *et al.*, 2003; Philippe and Douady *et al.*, 2003).

It is noteworthy that all of MEP pathway gene products contain the NH_2-terminal presequences, which are supported by both experimental evidence and computer analyses (Rodríguez-Concepción and Boronat, 2002). The MEP pathway gene products, therefore, appear to be synthesized as precursors in the cytoplasm and then post-translationally imported into plastids, which is consistent with the role of the MEP pathway in plastid isoprenoid biosynthesis. In summary, the MVA pathway is found in animals, fungi, archaea, some eubacteria, and the cytosol of plant cells, while the newly

discovered MEP pathway occurs in many eubacteria, algae, and the plastids of higher plants (Eisenreich *et al.*, 1998; Lichtenthaler, 1999; Eisenreich *et al.*, 2001; Rodríguez-Concepción and Boronat, 2002; Eisenreich *et al.*, 2004).

Table 1.2 Distribution of isoprenoid biosynthetic pathways in nature

Organisms	Isoprenoid biosynthetic pathways	
	MVA pathway	MEP pathway
Bacteria	+	+
Archaea	+	-
Fungi	+	-
Protozoa	+	+
Animals	+	-
Algae	+	+
Plants: Cytosolic compartment	+	-
Plastidial compartment	-	+

Note:

Symbols - and + indicate the absence or presence of the genes in the organisms, respectively

4. Crosstalk between the MVA and MEP pathways in plants

Although the subcellular compartmentation allows the MVA and MEP pathways to operate independently in plants, recent studies have established that the compartmental separation between the two isoprenoid pathways is not absolute in plants. The cooperation of both pathways has been reported for the biosynthesis of some sesquiterpenes in several plants. For example, sesquiterpenes (Bisabololoxide A and Chamazulene) in *Matricaria reculica* (Adam and Zapp, 1998) are composed of two isoprenoid units formed by the MEP pathway, with a third unit being derived either from the MVA or MEP pathways. The [^{13}C]-labelling patterns of ginkgolides in *G. biloba* indicated that three isoprenoid units of the ginkgolide carried the [^{13}C]-labelling characteristics of the MVA pathway, whereas the fourth isoprenoid unit was labelled by the MEP route (for review see Eisenreich *et al.*, 1998; Dubey *et al.,* 2003). A mixed origin of the isoprene units of phytol and other diterpenes has also been observed in liverwort (*Heteroscyphus planus*) (Nabeta *et al.*, 1995) and hornwort (*Anthoceros*

punctatus) (Itoh *et al.*, 2000). In addition, other cases of the co-operation of both pathways has been reported for sesquiterpenes from *Phaseolus lunatus* (Piel *et al.*, 1998), *Antirrhinum majus* (Dudareva *et al.*, 2005), and *Daucus carota* (Hampel *et al.*, 2005). These observations suggested that the exchange of some metabolites could occur between the two pathways.

The studies of Bick and Lange (2003) showed that the isolated chloroplasts and proteoliposomes from solubilized proteins of the plastidial envelope membranes were capable of a unidirectional export of IPP and geranyl diphosphate (GPP). The process was found to be controlled by Ca^{2+} and could be stimulated by pH and membrane potential gradients. Recently, experiments using an inhibitor of the MVA pathway also provided indirect evidence for the presence of a specific transporter system for isoprenoid intermediates from plastids to the cytosol compartment in *A. thaliana* (Laule *et al.*, 2003).

In addition, McCaskill and Croteau (1995) also provided the evidence for a significant amount of the exchange of IPP or other prenyl diphosphates between subcellular compartments in the glandular trichomes of *Mentha x piperita*. Soler *et al.* (1993) have also shown that the plastids of *Vitis vinifera* are capable for uptake of cytoplasmic IPP. In addition to IPP or other prenyl diphosphates, the early intermediates of the MEP pathway, such as DXP but not the IPP, can be transported into the cytosol by the recently discovered plastidic phosphate translocator, named xylulose 5-phosphate translocator (Flügge and Gao, 2005).

The fact that the biosynthesis of isoprenoids in plants is never completely inhibited by mevinolin , an HMGR-specific inhibitor or fosmidomycin, an inhibitor of DXR, could reflect the presence of the exchange of metabolites between the cytosolic MVA and the plastidial MEP pathways across the plastid envelope membranes. In the presence of mevinolin, Hemmerlin *et al.* (2003) have shown that phytosterols, such as campesterol, sitosterol, and stigmasterol, in tobacco TBY-2 cells can be synthesized from the MEP pathway intermediates. In contrast, in the presence of fosmidomycin, the plastoquinone in tobacco TBY-2 cells was also formed from the MVA pathway intermediates. In addition, De-Eknamkul and Potduang (2003) showed that in *Croton sublyratus*, the isoprene units used for the synthesis of phytosterols have a double origin from both the MVA and MEP pathways.

Taken together, these results suggest that an active exchange between plastidial and cytosolic compartments exists in plant cells, being dependent on the species, the tissue types, the development stage of the cells, or the plant culture system (Bick and

Lange, 2003). However, the mechanisms of the metabolite exchanged between the cytosolic and plastidial compartments still remain to be established.

5. 1-Deoxy-D-xylulose 5-phosphate synthase (DXS)

DXS (EC 2.2.1.7) catalyzes the first step of MEP pathway by condensation of GAP and pyruvate to form DXP, as shown in Fig. 1.1.

5.1 Isolation of the *DXS* genes

The gene encoding DXS was first identified in *E. coli* by extensive database searches with the sequence of the E1 subunit of the pyruvate dehydrogenase complex, pyruvate decarboxylase, and transketolase as queries. The recombinant protein was shown to catalyze the formation of DXP from pyruvate and GAP (Sprenger *et al.*, 1997; Lois *et al.*, 1998). After identification of the *DXS* gene from *E. coli*, several *DXSs* from various organisms, including many other bacteria such as *Streptomyces* sp. strain CL190 (Kuzuyama *et al.*, 2000b), *Pseudomonas aeruginosa* (Altincicek *et al.*, 2000), *Streptomyces coelicolor* (Cane *et al.*, 2001), *Rhodobacter capsulatus* (Hahn *et al.*, 2001), *Mycobacterium tuberculosis* (Bailey *et al.*, 2002), *Agrobacterium tumefaciens* (Lee *et al.*, 2007), cyanobacteria (*Synechococcus leopoliensis*) (Miller *et al.*, 1999), green algae (*Chlamydomonas reinhardtii*) (Schwender *et al.*, 1999), and higher plants (for references see below), have been subsequently isolated and characterized.

The first *DXS* gene in plants was isolated from an albino *A. thaliana* mutant plant obtained by screening a collection of T-DNA mutant lines. This gene is known from that time as *CLA1*, derived from the word cloroplastos alterados or altered chloroplasts, since disruption of this gene results in an alteration of plastids, such as no thylakoid membrane proliferation and a resemblance of plastids with a distorted prolamellar body in the early stage of plastid development (Mandel *et al.*, 1996). The *CLA1* cDNA contained 2468 bp, having an open reading frame (ORF) of 2151 bp with a predicted protein of 717 amino acids and a calculated molecular mass of 77 kDa. The derived amino acid sequence of *CLA1* gene shows an extensive amino acid similarity to the published DXS sequences, suggesting that it could encode an enzyme DXS in *A. thaliana* (Mandel *et al.*, 1996).

The function of CLA1 product as DXS was experimentally confirmed in both *in vivo* and *in vitro*. Estévez *et al.* (2000) have shown that when the *CLA1*-1 plants were grown in germination media supplemented with 0.02% (w/v) DX, a non-phosphorylated version of DXP, the first true leaves of mutants develop a green

phenotype, while mutants grown in the germination medium without adding DX develop an albino phenotype. In addition, the reaction product after incubation the recombinant CLA1 protein with pyruvate and GAP and further treated with alkaline phosphatase is determined to be a DX by gas chromatography-mass spectrometric analysis (Estévez *et al.*, 2000). The *CLA1* gene is present as a single gene in the *A. thaliana* genome (Mandel *et al.*, 1996). In addition to *Arabidopsis*, a single *DXS* gene was found in *Lycopersicon esculentum* (Lois *et al.*, 2000) and *Pueraria montana* genomes (Sharkey *et al.*, 2005).

Recently, a search in *A. thaliana* genome, however, results in a finding of two more *DXS*-related genes: *DXS2* (At3g21500 or NM_113045) and *DXS3* (At5g11380 or CAB96673). The function of these genes in relation to isoprenoid biosynthesis is still unknown. The *DXS2* and *DXS3* genes are scattered in different chromosomes of *A. thaliana*. The *DXS2* and *DXS3* genes are located on chromosomes 3 and 5, respectively, while the *CLA1* or *DXS1* gene is on chromosome 4. Therefore, the *A. thaliana* should contain three *DXS* genes in its genome (Rodríguez-Concepción and Boronat, 2002).

In addition to *A. thaliana*, *M. truncatula* is the first plant in which two isoforms of *DXS* genes, designated as *MtDXS1* and *MtDXS2*, have been identified (Walter *et al.*, 2002). Both genes have been shown to productive encode DXS proteins. The deduced amino acid sequences of these cDNAs show high sequence similarity with other plant DXSs, but exhibit only 70% identity (78% similarity) at amino acid level and 64% identity at nucleotide level to each other, indicating they represent allelic variants of *DXS* gene in the *M. truncatula* genome (Walter *et al.*, 2002). Since the multimeric forms of *DXS* genes have been cloned only from *M. truncatula*, it was not clear at that time whether this is a unique feature of this plant or of legumes or other species and plant families also contain multiple *DXS* forms.

To clarify this question, data mining was performed at The Institute for Genomic Research (TIGR) database (Quackenbush *et al.*, 2000) and these allowed Walter *et al.* (2002) to discover that not only *M. truncatula*, but also *Glycine max*, *L. esculentum*, *Solanum tuberosum*, and *Zea mays* contain multiple forms of the *DXS* gene, with at least two members in each of these plants. In agreement with these findings, results of subsequent studies by Khemvong and Suvachittanont (2005), Gong *et al.* (2006), and Kim *et al.* (2006) have shown that transcripts of two *DXS* genes are also present in *Elaeis guineensis* and *G. biloba*. Recently, Phillips *et al.* (2007) isolated three full-length cDNAs of *DXS* genes from a cDNA library of methyl jasmonate (MeJA)-treated *Picea*

abies cell cultures. Based on these findings, it can be concluded that the presence of multiple forms of plant *DXS* gene is a widespread phenomenon in higher plants.

Currently, genes encoding DXS proteins have now been identified in a number of plant species. For example, the *DXS* cDNAs have been cloned from *Capsicum annuum* (Bouvier *et al.*, 1998), *M. piperita* (Lange *et al.*, 1998), *Catharanthus roseus* (Chahed *et al.*, 2000), *Tagetes erecta* (Moehs *et al.*, 2001), *Artemisia annua* (Souret *et al.*, 2002), *Morinda citrifolia* (Han *et al.*, 2003), *Stevia rebaudiana* (Totté *et al.*, 2003), *Antirrhinum majus* (Dudareva *et al.*, 2005), and *Chrysanthemum morifolium* (Kishimoto and Ohmiya, 2006).

5.2 Primary structure of the DXS proteins

A comparison of the primary amino acids sequences of DXS from various organisms revealed that there are a number of conserved structural features characteristic of DXS proteins: the thiamin diphosphate (TPP)-binding motif, a transketolase motif, and several invariant amino acids, such as histidine, glutamic acid, and arginine, which are important for the enzymatic activity (Mandel *et al.*, 1996; Lange and Croteau, 1998; Lois *et al.*, 1998; Chahed *et al.*, 2000; Querol *et al.*, 2001; Hahn *et al.*, 2001; Souret *et al.*, 2002; Khemvong and Suvachittanont, 2005; Gong *et al.*, 2006; Lee *et al.*, 2007; Phillips *et al.*, 2007). The arrangement of these motifs and important amino acid residues in the DXS sequences is shown in Fig. 1.2.

The TPP-binding motif and a transketolase-like motif within the DXS proteins are found at the N-terminus and found near the C-terminus, respectively (Fig. 1.2). The consensus sequences for the TPP-binding motif and a transketolase-like motif are designated as GDG-X_{25}-NDN and DRAG-X_{28}-P-X-D, respectively, where X denotes any amino acid (Mandel *et al.*, 1996; Querol *et al.*, 2001). Both motifs have previously shown to be the signatures of proteins belonging to the transketolases family (Hawkins *et al.*, 1989; Schenk *et al.*, 1997).

In addition to these two conserved motifs, the histidine and glutamic acid residues that were previously shown to be involved in the catalysis of transketolase in yeast are also found to be highly conserved in all DXS proteins identified so far (Fig. 1.2) (Bouvier *et al.*, 1998; Querol *et al.*, 2001; Gong *et al.*, 2006; Lee *et al.*, 2007; Xiang *et al.*, 2007). The role of the histidine residue has been shown to be in the transfer of a proton during catalysis and in a stabilization of the covalent intermediate between the substrate and TPP molecules (Linvsqvist *et al.*, 1992; Wikner *et al.*, 1997; Schneider and Linvsqvist, 1998). The glutamic acid residue has been reported to form a hydrogen bond

to the N1 nitrogen atom of the pyrimidine ring of the TPP cofactor (Wikner *et al.*, 1994). Recently, the determination of three-dimensional (3D) structures of *Deinococus radiodurans* and *E. coli* DXSs allows the identification of two additional important arginine residues (R398 and R478), which are essential for the DXS activity (Fig. 1.2) (Xiang *et al.*, 2007).

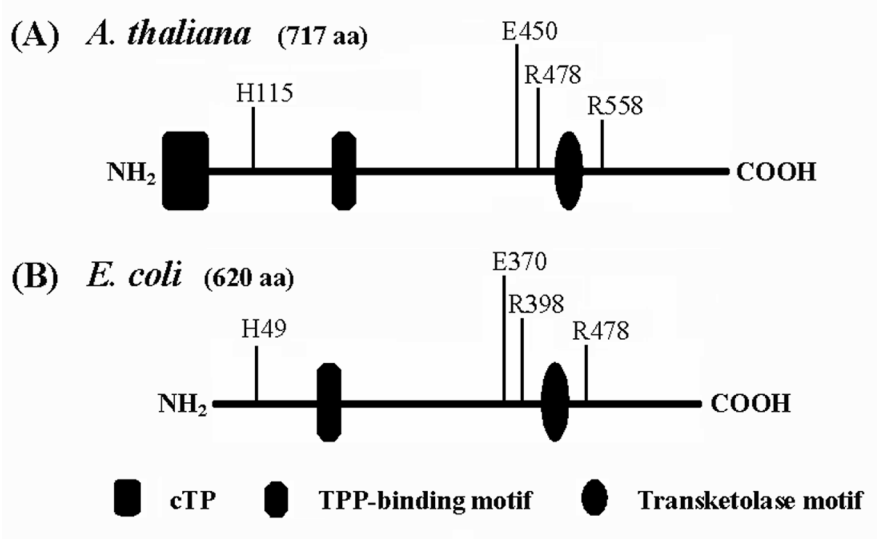


Fig. 1.2 Diagrammatic representation of DXS domain structures.

(A) Structural drawing of *A. thaliana* DXS. (B) Structural drawing of *E. colli* DXS. The predicted chloroplast transit peptide (cTP) sequence, TPP-binding motif, and transketolase motif are shown. Positions of important residues for the activity of DXS are also indicated.

As the MEP pathway is localized in plastids, the amino acid sequences of all plant DXSs identified to date are found to contain a putative chloroplast transit peptide (cTP) sequence in the extra amino acids at the N-terminus region (Fig. 1.2). These regions vary considerably in different plant species ((Mandel *et al.*, 1996; Lange and Croteau, 1998; Lois *et al.*, 1998; Chahed *et al.*, 2000; Querol *et al.*, 2001; Hahn *et al.*, 2001; Souret *et al.*, 2002; Khemvong and Suvachittanont, 2005; Gong *et al.*, 2006; Phillips *et al.*, 2007). However, they are rich in hydroxylated amino acids such as serine and threonine, and contain some basic residues. These are common features of cTP sequence in plants (Gavel and von Heijne, 1990; von Heijne and Nishikawa, 1991; Emanuelsson *et al.*, 1999). Furthermore, an immunoblot and transient expression

analyses also confirm that DXS proteins from plants are localized in plastids (Bouvier *et al.*, 1998; Lois *et al.*, 2000). The cTP sequence is important for translocation of the protein from the cytosol to the chloroplast compartment in plants (Mori and Cline, 2001).

5.3 Evolutionary analysis of the DXS sequences

A comparative analysis presents an ideal opportunity to investigate the dynamics of the DXS protein family, particularly to expand understanding of DXS evolution. As shown by Lange *et al.* (2000), the phylogenetic tree inferred from amino acid sequences of DXS from various organisms by maximum likelihood revealed that all DXS sequences share a common ancestor with TPP-dependent enzymes, such as transketolases, pyruvate:ferredoxin oxidoreductase, the E1β-subunit of pyruvate degydrogenase, and branched-chain α-ketoacid dehydrogenases. Among the DXS group, however, all plant DXSs tend to branch with the DXS homologue from the α-proteobacterium *R. capsulatus*, while the cyanobacterium *Synechocystis* sp. PCC6803 is branched with the DXS homologue from *Bacillus subtilis*. These results suggested that DXS sequences in plants were acquired by a lateral gene transfer from α-proteobacterium to assemble the MEP pathway for synthesis of their isoprenoid compounds (Lange *et al.*, 2000).

In 2002, when the first second *DXS* gene of from *M. truncatula* was available, the phylogenetic tree inferred from the neighbor joining method revealed that plant DXSs can be separated into two distinct classes after being rooted with bacterial DXS sequences (Walter *et al.*, 2002). The first DXS class named class I contains DXS sequences from *A. thaliana, A. annua*, *L. esculentum*, and *C. annuam*. Another class of DXS called class II contains DXS sequences from *C. roseus*, *M. piperita*, *T. erecta*, and *N. pseudonarcissus* (Walter *et al.*, 2002). In addition to *M. truncatula*, the presence of the two different classes of *DXS* genes in plants was also found in subsequent studies in various plants, such as in *E. guineensis*, *G. biloba*, and *Picea abies* (Khemvong and Suvachittanont, 2005; Gong *et al.*, 2006; Kim *et al.*, 2006; Phillips *et al.*, 2007). It has been suggested that the presence of two different classes of DXS sequence in plants might be derived from a gene duplication event (Walter *et al.*, 2002; Krushkal *et al.*, 2003). This event might have occurred prior to the divergence of flowering plants or angiosperms, since the two distinct classes of *DXS* genes are observed in the gymnosperms: Ginkgoales i.e. *G. biloba* (Gong *et al.*, 2006; Kim *et al.*, 2006) and in Coniferales i.e. *P. abies* (Phillips *et al.*, 2007).

5.4 Expression of the *DXS* transcripts or proteins

The expression of *DXS* genes in different tissues of various plant species has been reported. For example, the *DXS* gene (*CLA1*) from *A. thaliana* was analyzed by RNA-gel blot hybridization and β-glucuronidase (GUS) reporter gene to show that the *CLA1* gene is widely expressed throughout the plants, with the higher expression levels in young tissues (Estévez *et al.*, 2000; Carretero-Paulet *et al.*, 2002) The expression of RNA is in good agreement with the levels of protein as detected by the immunoblot analysis (Estévez *et al.*, 2000).

In *L. esculentum*, high levels of *DXS* mRNA are also found in photo-synthetic tissue, while very low levels are detected in roots (Lois *et al.*, 2000). In addition, the levels of *DXS* mRNA are correlated with the accumulation carotenoids during *L. esculentum* fruit development, suggesting that the induction of the *DXS* gene is associated with the activation of carotenoid biosynthesis at the onset of ripening (Lois *et al.*, 2000). The expression of *DXS* gene in correlation to carotenoid accumulation during ripening is also found in other plants such as *C. annuum* and *E. guineensis* fruits (Bouvier *et al.*, 1998; Khemvong and Suvachittanont, 2005).

Since the discovery of the first two *DXS* genes in *M. truncatula* by Walter *et al.* (2002), the transcripts of these genes in different tissues of this plant were also studied. The authors found that the *MtDXS1* transcripts are abundant in all above-ground tissues, while the *MtDXS2* transcripts show very low to non-detectable levels in the above-ground tissues but they are abundant in roots that interact with mycorrhizal fungi (Walter *et al.*, 2002). Upon mycorrhization, the transcripts of *DXS* in *Oryza sativa*, *Zea mays*, *Triticun aestivum*, and *Hordeum vulgare* are also increased in relation to the accumulation of apocarotenoids i.e. mycorradicin and glycosylated cyclohexenone derivatives (Walter *et al.*, 2000; Walter *et al.*, 2002).

In *G. biloba*, another plant from which two classes of *DXS* genes have been cloned, Gong *et al.* (2006) reported that the *GbDXS1* gene is expressed in all tissues analyzed including roots, stems, leaves, pericarps of fruits, and seeds. Another study by Kim *et al.* (2006) revealed that the expression of *GbDXS1* is abundant in leaves, while the transcripts of *GbDXS2* are abundant in roots. Also, they reported that the *GbDXS* transcripts are correlated with the production of ginkgolides in the embryo culture of *G. biloba* (Kim *et al.*, 2006).

Based on these results, it can be concluded as a general idea that the expression of *DXS* class I is required for the biosynthesis of primary isoprenoids, such as chlorophylls and carotenoids (Mandel *et al.*, 1996; Bouvier *et al.*, 1998; Estévez *et al.*, 2000; Lois *et al.*, 2000; Walter *et al.*, 2002; Khemvong and Suvachittanont, 2005; Kim *et al.*, 2006). In contrast, the expression of *DXS* class II might be involved in secondary metabolite biosynthesis, such as that of apocarotenoids and ginkgolides (Walter *et al.*, 2002; Kim *et al.*, 2006). Other secondary isoprenoids, such as monoterpene-alkaloids from *M. piperita* and *C. roseus*, also seem to be correlated with the expression of their *DXS* genes belonging to the plant *DXS* class II (Lange *et al.*, 1998; Veau *et al.*, 2000; Chahed *et al.*, 2000; Burlat *et al.*, 2004). To date, there is no information of *DXS* sequence belonging to *DXS* class I identified in these plants.

Several external stimuli such as light, fungal and chitosan treatments, elicitors (ceric ammonium sulfate, jasmonic acid, methyl jasmonate, arachidonic acid, acetylsalicylic acid, and methyl salicylate), and wounding have been shown to up-regulate the expression of *DXS* genes from various plants (Mandel *et al.*, 1996; Estévez *et al.*, 2000; Souret *et al.*, 2002; Walter *et al.*, 2000; Walter *et al.*, 2002; Han *et al.*, 2003; Dudareva *et al.*, 2005; Hsieh and Goodman, 2005; Gong *et al.*, 2006; Kim *et al.*, 2006; Phillips *et al.*, 2007). In addition to these stimuli, injection of DX or fosmidomycin, a specific inhibitor of DXR, has been found to induce *DXS* transcripts in *L. esculentum* (Lois *et al.*, 2000). The absence of 2,4 dichloro-phenoxyacetic acid is also known to induced monoterpene production in *C. roseus*, and hence increase the *DXS* transcripts (Veau *et al.*, 2000).

5.5 Catalytic properties of the DXS enzymes

DXS is a homodimeric enzyme consisting of two identical subunits. It has a molecular mass of 68-71 kDa in each subunit. The DXS enzyme requires TPP and Mg^{2+} as cofactors. Studies of the steady-state kinetic indicates that DXS catalyzes a ping-pong mechanism, in which the pyruvate binds first, followed by releasing of CO_2 and binding of GAP, and finally the releasing of DXP (Eubanks and Poulter, 2003). Besides GAP, it is found in *E. coli* that the DXS enzyme can use other sugar phosphates as acceptor substrates and the short-chain aldehydes as donors also (Querol *et al.*, 2002; Schuermann *et al.*, 2002).

5.6 Inhibitors of the DXS enzymes

It has been shown that fluoropyruvate covalently binds to the active site of DXS enzymes in a similar way as described for pyruvate dehydrogenase E1 (Flournoy and Frey, 1989; Altincicek *et al.*, 2000). The fluoropyruvate inhibits the DXS enzyme from *P. aeruginosa* and *E. coli* with IC_{50} values of 400 μM and 80 μM, respectively (Altincicek *et al.*, 2000). In addition, ketoclomazone has been shown to inhibit the *C. reinhardtii* DXS with the IC_{50} value of 0.1 mM (Müller *et al.*, 2000).

5.7 Crystal structure of DXS

The crystal structures of DXS have recently been determined from *D. radiodurans* and *E. coli* by selenomethionyl single-wavelength anomalous diffraction (Xiang *et al.*, 2007). The 3D structures of these bacterial DXSs are depicted in Fig. 1.3. Based on their 3D structures, there are three domains: I, II, and III detected in each subunit of DXS (Fig. 1.3). Domain I (residues 1-319) contains a five-strand parallel β-sheet, domain II of *D. radiodurans* DXS (residues 320-495) contains a six-stranded parallel β-sheet, and domain III (residues 496-629) contains a five-stranded β-sheet, but the first strand is anti-parallel to the other four strands (Fig. 1.3). The active site of DXS has been shown to be located at the interface between domains I and II of the same monomer. The amino acid residues in the active site of DXS are highly conserved and some of these residues are also conserved in transketolases (Xiang *et al.*, 2007). There are two regions in the active site of DXS: the TPP-binding region and the GAP-binding region.

In the 3D structure of *D. radiodurans* DXS in complex with TPP (Fig. 1.4), the TPP molecule is bound to DXS in a V-shaped conformation. The amino-pyrimidine ring of TPP appears to interact with domain I, while its pyrophosphate moiety is interacted with domain II. The amino-pyrimidine ring of TPP forms a hydrogen bond with the side chain of phenylalanine residue (F398) and van der Waals interactions with the side chain of isoleucine residues (I371 and I187) (Fig. 1.4). In addition, the nitrogen atom N1 in the amino-pyrimidine ring forms a hydrogen bond with the side chain of a glutamic acid residue (E373) (Fig. 1.4). Since it has been shown that the DXS requires Mg^{2+} ion as cofactor, the Mg^{2+} ion in the *D. radiodurans* DXS is found to bind between two phosphate groups of TPP molecule, which are further coordinated by the side chains of aspartic acid and asparagines residues (D154 and N183, respectively), and are also coordinated by the main chain carbonyl of methionine residue (M185) (Fig. 1.4) (Xiang *et al.*, 2007). In addition to the TPP-binding region, there is no information about the

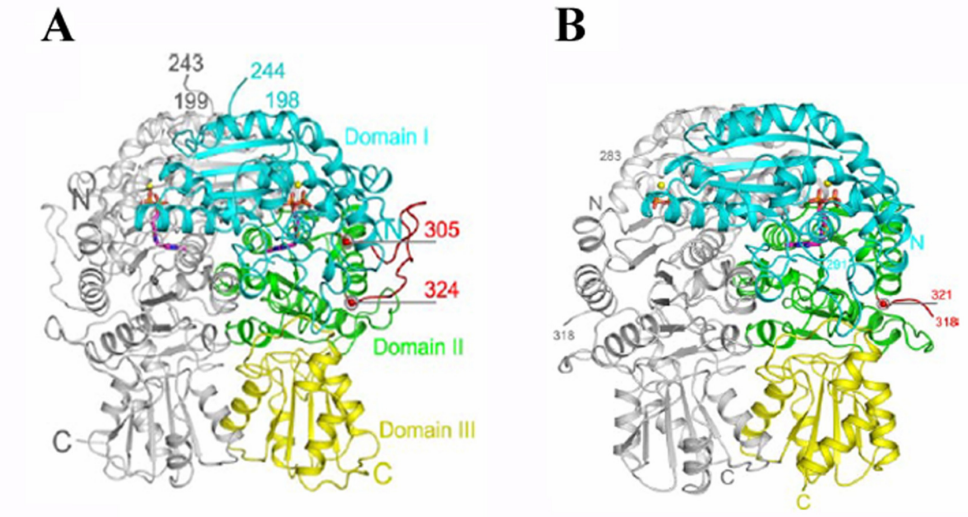


Fig. 1.3 Schematic drawing of the overall structure of DXSs.

(A) Ribbon diagram of *D. radiodurans* DXS dimer. (B) Ribbon diagram of *E. coli* DXS dimer. Domains I, II, and III are colored cyan, green, and yellow, respectively. The linker between domains I and II is colored red. One monomer is shown in grey color (Xiang *et al.*, 2007).

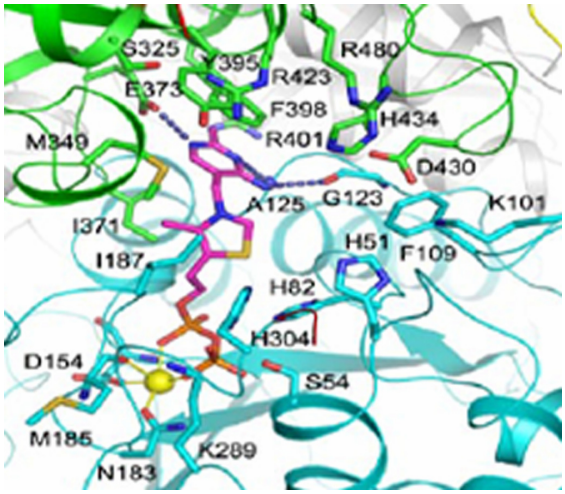


Fig. 1.4 Sterio view of the active site of *D. radiodurans* DXS.

The TPP molecule is shown by magenta and the Mg^{2+} ion is yellow-colored. Hydrogen bond between the N1 atom of amino-pyrimidine ring of TPP and the side chain of E373 is indicated by the dotted line (Xiang *et al.*, 2007).

binding mode of GAP in the 3D structure of bacterial DXSs studied by Xiang *et al.* (2007). So, the binding mode of GAP in DXS is obtained by modeling using the structure of yeast transketolase in complex with erythrose 4-phosphate as a template (Nilsson *et al.*, 1997). The results showed that GAP could interact with the side chain of amino acids: histidine (H51 and H304), tyrosine (Y395), arginine (R423 and R480), and aspartic acid (D430) residues in the bacterial DXS enzymes (Xiang *et al.*, 2007).

5.8 Mutagenesis studies

The result obtained from multiple alignments of amino acid sequences of DXS and transketolase sequences, a TPP-dependent enzyme, from *E. coli* and yeast shows a significant homology with several highly conserved residues between the sequences. One of these residues (H30), which was previously reported to be involved in proton transfer during catalysis of yeast transketolase (Lindqvist *et al.*, 1992), is also found in *E. coli* DXS where it corresponds to the amino acid position 49 (H49). Mutation of this residue to glutamine (H49Q) in the *E. coli* DXS showed that the mutant proteins cannot rescue the growth of *E. coli* defective in the *DXS* gene (Querol *et al.*, 2001). In *A. tumefaciens*, change of this residue to alanine (H48A) results in a loss of enzyme activity, exhibiting below 0.1% of specific activity (Lee *et al.,* 2007). Based on the 3D structure, this residue has been shown to be involved in GAP binding (Xiang *et al.*, 2007). Moreover, roles of other residues localized in the active site of *E. coli* DXS were also determined. Xiang *et al.* (2007) have been found that glutamic acid (E370) and arginine (R398 and R478) residues in the *E. coli* DXS play crucial roles during catalysis. Mutation of these residues to alanine (E370A, R398A, and R478A) causes loss of all detectable DXS enzyme activities. As shown in Fig. 1.4, the side chains of E370 and R398 interact with each other and with the TPP molecule. The residue R478 has been proposed to recognize the GAP molecule (Xiang *et al.*, 2007).

6. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase (DXR)

DXP reductoisomerase or DXR (EC 1.1.1.267) is a second enzyme of the MEP pathway, but it catalyzes the first reaction specific to the pathway. The enzymatic reaction of DXR consists of an intramolecular rearrangement of DXP followed by the NADPH-dependent reduction of the resulting product, 2*C*-methyl-D-erythritol 4-phosphate (Fig. 1.1).

6.1 Isolation of the *DXR* genes

Screening of *E. coli* mutants that contained a metabolic block between DXP and MEP allowed the isolation of the *yaeM* gene from the *E. coli* mutant strains that can survive in growth medium without supplementation of ME, a non-phosphorylated version of MEP (Takahashi *et al.*, 1998). *YaeM* is a single gene, thus disruption of this gene causes a lethal phenotype of *E. coli*. This gene was later named *DXR*. After the discovery of the *E. coli DXR* gene, several orthologous sequences are identified and characterized from various bacteria (Grolle *et al.*, 2000; Cane *et al.*, 2001; Altincicek *et al.*, 2000), malaria parasite (Jomaa *et al.*, 1999), cyanobacteria (Miller *et al.*, 2000; Fernandes *et al.*, 2005; Fernandes and Proteau, 2006), and higher plants (for references see below).

In plants, first *DXR* gene was isolated from *A. thaliana* by two different research groups. First, it was identified based on the nucleotide sequence of the *E. coli DXR* as query, allowing Schwender *et al.* (1999) isolate a cDNA fragment encoding DXR from *A. thaliana* by polymerase chain reaction cloning method. Sequence analysis revealed that the cloned cDNA has an ORF of 1242 bp encoding a protein of 406 amino acids with a predicted molecular mass of 44 kDa. The cDNA represents a part of the *DXR* gene, since its 5´ region is not complete. The predicted protein of this clone shared 44% and 66% identical amino acid sequences to DXRs from *E. coli* and *Synechocystis* sp. PCC6803. Although the cDNA sequence was incomplete, the recombinant protein overexpressed in *E. coli* showed DXR activity (Schwender *et al.*, 1999). In the same year, Lange and Croteau (1999) also succeeded in the isolation of the *A. thaliana DXR* gene, but its nucleotide sequence represented a truncated version of the *DXR* gene as well. Then, the complete sequence of the *A. thaliana DXR* gene was first identified by Carretero-Paulet *et al.* (2002) using the RT-PCR approach. The full-length gene contains an ORF of 1434 bp encoding a protein of 477 amino acids with a predicted molecular mass of 52 kDa. A databank search showed that the *DXR* gene is mapped on chromosome 5 of *A. thaliana* (Carretero-Paulet *et al.*, 2002). The results from Southern blot analysis revealed that the *A. thaliana DXR* is encoded by a single gene (Schwender *et al.*, 1999; Carretero-Paulet *et al.*, 2002). The function of the cloned cDNA as DXR was confirmed by complementation in an *E. coli* mutant defective in the *DXR* gene (Carretero-Paulet *et al.*, 2002).

Following the isolation of the *A. thaliana DXR* gene, use of the *A. thaliana* and *O. sativa* genes as heterologous probes allowed Sharkey *et al.* (2005) and Hans *et al.* (2004) to isolate *DXR* genes from cDNA libraries of *P. montana* and *Z. mays*, respectively. Also, the homologous *DXR* probe from *M. piperita* was used to screen for the full-length *DXR* clone from an oil gland cells cDNA library (Lange and Croteau, 1999). Furthermore, a single copy of the *DXR* gene has been isolated from *L. esculentum* by searching against EST databases at the National Center for Biotechnology Information (NCBI) and TIGR using the *M. piperita DXR* as a query (Rodríguez-Concepción *et al.*, 2001). In addition to the cDNA library screening, using the RT-PCR based approach with primers designed based on the conserved sequences among *DXR* genes, allowed several workers to clone the gene from various plant species such as *C. roseus* (Veau *et al.*, 2000), *A. annua* (Souret *et al.*, 2002), *S. rebaudiana* (Totté *et al.*, 2003), *A. majus* (Dudareva *et al.*, 2005), *Coleus forskohlii* (Engprasert *et al.*, 2005), *Taxus cuspidata* (Jennewein *et al.*, 2004), *G. biloba* (Gong *et al.*, 2005; Kim *et al.*, 2006), *Camptotheca acuminate* (Yao *et al.*, 2007), and *Rauvolfia verticillata* (Liao *et al.*, 2007).

6.2 Primary structure of the DXR proteins

Multiple alignments of amino acid sequences of all DXRs identified so far reveal two highly conserved regions (Fig. 1.5). First, a structural motif showing a high level of similarity to the consensus sequence of a reduced β-nicotinamide adenine dinucleotide phosphate (NADPH)-binding site (GXXGXXG) is found at the N-terminus of the DXR sequences (Fig. 1.5). Second, a cluster of amino acid sequences called Proline-rich region ($P_{2\text{-}4}$AWPGR/T) is also found at the N-terminus of the DXRs (Fig. 1.5), but its biological function in plants is still unclear, since this sequence was not found in the prokaryotic DXRs. It has been, however, suggested that this region might be required for protein-protein interactions in plants (Carretero-Paulet *et al.*, 2002). Besides these two important regions, the highly conserved residues forming the active site of DXR protein are found in all DXR sequences. These include amino acids D150, E152, H209, E231, E243, and H257 (numbers are assigned according to the *E. coli* DXR) (Fig. 1.5) (Kuzuyama *et al.*, 2000; Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; MacSweeney *et al.*, 2005; Fernandes and Proteau, 2006). Consistent with the plastid localization of MEP pathway in plants, the N-terminal region of plant DXRs is predicted to contain the cTP sequence (Fig. 1.5). The presence of a cTP sequence was confirmed in *A. thaliana DXR* using immunoblot analysis and transient expression. Interestingly, a computer analysis by the ChloroP program predicts that most plant DXRs have a highly

conserved cleavage site as Xaa↓Cys-Ser, where Xaa means any amino acid (Carretero-Paulet *et al.*, 2002; Hans *et al.*, 2004; Gong *et al.*, 2005; Yao *et al.*, 2007; Liao *et al.*, 2007).

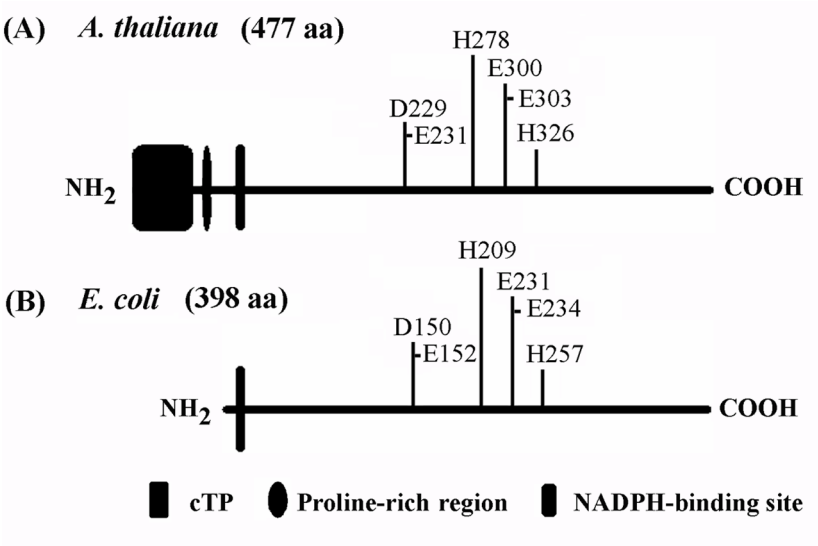


Fig. 1.5 Diagrammatic representation of DXR domain structures

(A) Structural drawing of *A. thaliana* DXR. (B) Structural drawing of *E. colli* DXR. The predicted cTP sequence, Proline-rich region, and NADPH-binding site are shown. Positions of important residues forming the active site of DXR are also indicated.

6.3 Evolutionary analysis of the DXR sequences

A phylogenetic analysis of DXR sequences from a wide-ranging of organisms including eubacteria, protists, cyanobacteria, and higher plants using maximum likelihood revealed that all DXRs shared a common ancestor of origin (Lange *et al.*, 2000). In plants, DXRs form a monophyletic group and show resemblance to known cyanobacteria counterparts, for example, *Synnechocystis* sp. PCC6803, suggesting that the nuclear-encoded DXR enzymes in higher plants are of plastid origin and might have been required by a lateral gene transfer process from cyanobacteria (Lange *et al.*, 2000; Gong *et al.*, 2005; Yao *et al.*, 2007; Liao *et al.*, 2007). It is believed that nucleus-encoded chloroplast-targeted proteins were originally encoded in an ancestral chloroplast genome, and then transferred to host nuclear genome, where they acquired the proper expression and targeting signals to allow the encoded proteins to be synthesized and imported to the

organelles with the help of a transit peptide during integration of endosymbiosis (Martin and Herrmann, 1998; McFadden, 1999).

6.4 Expression of the *DXR* transcripts or proteins

Recent studies have shown that plant DXRs are differentially regulated in various tissues and under growth development like those found in plant DXSs. For example, Carretero-Paulet *et al.* (2002) have shown in *A. thaliana* that the *DXR* transcripts and its encoded proteins are found to be expressed in all organs analyzed, but at different levels. The *DXR* transcripts also show higher expression levels in seedlings grown in light than those in the dark, and they are rapidly accumulated upon de-etiolation (Carretero-Paulet *et al.*, 2002), suggesting that its expression is under control of light/dark cycles. The highest levels of *DXR* transcripts are found in the late period of the light cycle (Hsieh and Goodman, 2006). In transgenic plants transformed with a GUS-*DXR* chimeric gene, the chimeric transcripts also show similar expression results, which are consistent with the distribution of *DXR* transcripts and proteins in intact *A. thaliana* as previously studied (Carretero-Paulet *et al.*, 2002).

The expression of the *DXR* gene from *L. esculentum*, unlike *DXS*, is not induced during fruit development or treatment with DX. In addition, the levels of DXR protein show similar levels throughout the developmental stages of *L. esculentum* fruit (Rodríguez-Concepción *et al.*, 2001). However, injection of fosmidomycin into fruit at the mature green stage shows an increase in the expression of both *DXS* and *DXR* genes, but the accumulation of fruit carotenoids is inhibited (Rodríguez-Concepción *et al.*, 2001), suggesting that DXS but not DXR plays a rate-limiting step for carotenoid biosynthesis in *L. esculentum* fruit. In addition, seeds from fruits treated with fosmidomycin are observed to germinate earlier, suggesting that the reduced seed dormancy might occur due to the reduced levels of abscisic acid, a carotenoid-derived plant hormone produced during development of seeds (Rodríguez-Concepción *et al.*, 2001). As described earlier for the *DXS* gene, the expression of the *DXR* gene in roots of some monocots has been shown to be increased after colonization by mycorrhizal fungi (Walter *et al.*, 2000). In *Z. mays* roots, the accumulation of DXR protein is found in cortex cells and is correlated with arbuscule development (Hans *et al.*, 2004).

6.5 Catalytic properties of the DXR enzymes

Like DXS, DXR is also a homodimeric enzyme consisting of two identical subunits with a molecular weight of 39-54 kDa for each subunit. The DXR enzyme requires a divalent cation, such as Mg^{2+}, Mn^{2+}, or Co^{2+}, as a cofactor and NADPH as an

electron donor. Replacement of NADPH by NADH showed a decrease in activity to 1% in the *E. coli* DXR (Takahashi *et al.*, 1998). *M. tuberculosis* DXR, however, can use both NADPH and NADH as cofactors *in vitro*, but prefers NADPH *in vivo* (Argyrou and Blanchard, 2004). Only DXP can be utilized as a substrate (Takahashi *et al.*, 1998). In-depth studies revealed that DXR is a reversible enzyme (Hoeffler *et al.*, 2002; Koppisch *et al.*, 2002; Argyrou and Blanchard, 2004).

6.6 Inhibitors of the DXR enzymes

Fosmidomycin [3-(N-formyl-N-hydroxyamino) propylphosphonic acid], a natural product from *S. lavendulae* (Okuhara *et al.*, 1980a) was reported as a potent inhibitor of the DXR enzyme (Kuzuyama *et al.*, 1998). The fosmidomycin is a structural analogue of 2C-methyl-D-erytrose 4-phosphate. The fosmidomycin was determined to be a mixed inhibitor (competitive and non-competitive with DXP) against *E. coli* and *Synechocystis* sp. PCC6803 DXRs (Kuzuyama *et al.*, 1998, Woo *et al.*, 2006) but a competitive inhibitor against *Z. mobilis* DXR (Grolle *et al.*, 2000). It was also found to inhibit DXR enzymes from *P. falciparum* and *M. tuberculosis* (Jomaa *et al.*, 1999; Kuzuyama *et al.*, 1998; Lell *et al.*, 2003; Dhiman *et al.*, 2005). Recent studies have shown that fosmidomycin is an uncompetitive inhibitor against NADPH and gives a pattern representative of slow, tight-binding competitive inhibition against DXP, suggesting the order of mechanism in which NADPH binds first followed by DXP (Koppisch *et al.*, 2002; Argyrou and Blanchard, 2004). In addition to fosmidomycin, FR-900098 and fosfoxacin, both are natural products from *S. rubellomurinus* (Okuhara *et al.*, 1980b) and *P. fluorescens* PK-52, respectively also known to be effective inhibitors of DXRs (Woo *et al.*, 2006). Moreover, several chemically-synthesized fosmidomycin analogues were shown to have an inhibitory activity against DXR (Woo *et al.*, 2006). Two DXP analogs have also been found to be inhibitors of *E. coli* DXR, but at much higher concentrations than the fosmidomycin (Hoeffler *et al.*, 2002).

6.7 Crystal structure of DXR

Five X-ray crystallographic structures of DXR have recently been determined from *E. coli* (Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; Yajima *et al.*, 2004; MacSweeney *et al.*, 2005), one from *Z. mobilis* (Ricagno *et al.*, 2004), and one from *Mycobacterium tuberculosis* (Henriksson *et al.*, 2006). The 3D structure of *E. coli* DXR is illustrated in Fig. 1.6, as an example. DXR contains two identical monomers with three distinct domains in each monomer: an N-terminal domain, a central catalytic domain, and a C-terminal domain (Fig. 1.6). The N-terminal domain

contains a Rossmann fold involved in binding of NADPH. The central (or catalytic) domain can be sub-divided into the putative binding site for substrates and metal ions, and the catalytic loop. The C-terminal domain appears to have a structural role in supporting the catalytic domain and is connected to the catalytic domain by a linker region. These three domains are arranged in a V shape molecule, where the N-terminal domain and the C-terminal domain form two arms and the catalytic domain lies at the apical region (Fig. 1.6).

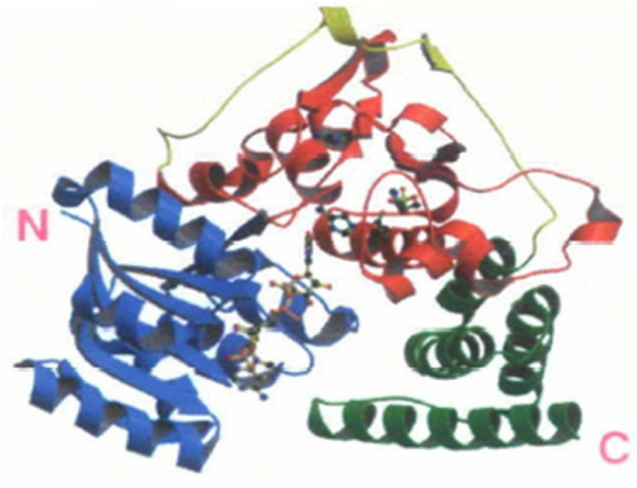


Fig. 1.6 Schematic drawing of the overall structure of *E. coli* DXR.

A ribbon diagram of the *E. coli* DXR monomer. The N-terminal domain is colored blue, a central domain is red, and the C-terminal domain is green, respectively. The linker region is shown in yellow color (Yajima *et al.*, 2002).

As shown in Fig. 1.7, the *E. coli* DXR structure in complex with DXP substrate showed that the phosphate moiety of DXP appears to form hydrogen bonds with serine (S186 and S222), asparagine (N227), and lysine (K228) residues, and with the catalytically essential histidine residue (H209). The DXP backbone interacts with the β-indole of tryptophan residue (W212) and the carbonyl group of carbon atom C2 is hydrogen bonded to glutamic acid residue (E152) and the backbone nitrogen of serine residue (S151). The hydroxyl group of carbon atom C3 is shown to be involved in hydrogen bonds with lysine (K125) and glutamic acid (E231) residues, while the hydroxyl group of carbon atom C4 is hydrogen bonded to glutamic acid (E152), asparagine (N227),

and lysine (K228). The active site of *E. coli* DXR contains highly conserved aspartic acid (D150), glutamic acid (E152, E231, and E234), and histidine (H209 and H257) residues. A cluster of aspartic and glutamic acid residues are also thought to be involved in the binding of the divalent cation (Reuter *et al.*, 2002; Steinbacher *et al.*, 2003; MacSweeney *et al.*, 2005). In addition to substrate binding, a ternary structure of *E. coli* DXR in complex with co-factor Mn^{2+} cation and inhibitor fosmidomycin was determined (Steinbacher *et al.*, 2003). As shown in Fig. 1.8, the phosphonate moiety of fosmidomycin molecule is bound in a similar fashion to the phosphate group of DXP and its hydroxamic acid moiety is coordinated with a divalent cation (Mn^{2+}) that binds with the side-chains of aspartic acid (D150) and glutamic acid (E152 and E231) residues (Steinbacher *et al.*, 2003).

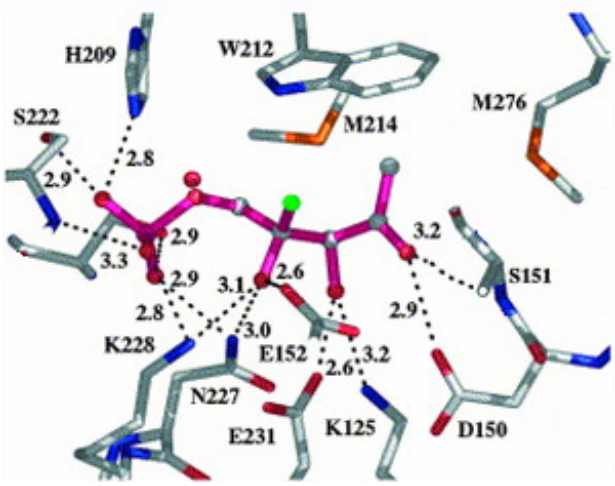


Fig. 1.7 Structure of DXP binding in the active site of *E. coli* DXR.

The hydrogen bonding interactions between the substrate DXP (pink) and key amino acids in the active site of *E. coli* DXR are shown by dotted lines with their distances in the model in Ångström written next to them (MacSweeney *et al.*, 2005).

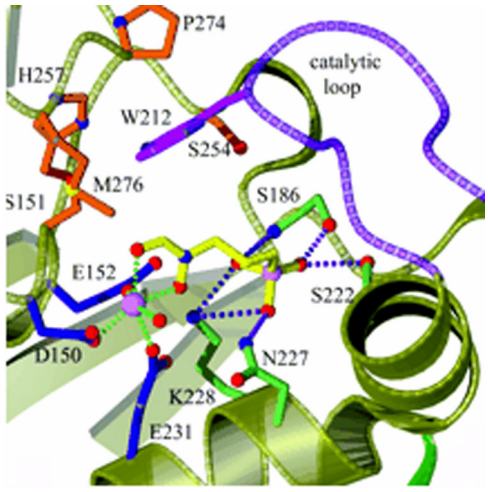


Fig. 1.8 Structure of fosmidomycin binding in the active site of *E. coli* DXR.

The hydrogen bonding interactions between the hydroxamic acid moiety of the inhibitor fosmidomycin (yellow) and key amino acids in the active site of *E. coli* DXR are shown by dotted lines. The Mn^{2+}cation is shown as a purple ball (Steinbacher *et al.*, 2003).

6.8 Mutagenesis studies

Before the publication of crystal structures, several mutants of *E. coli* DXR were evaluated. For example, mutation of glutamic acid to lysine (E231K) shows a lower k_{cat} value in the mutant enzyme (450-fold less than the wild-type) (Kuzuyama *et al.*, 2000a). This residue has been shown to be a part of the Mn^{2+} binding, based on the *E. coli* DXR structure in complex with Mn^{2+} and fosmidomycin (Steinbacher *et al.*, 2003). Mutation of this residue to histidine (E223H) also shows a low specific activity in the *Synechocystis* sp. PCC6803 DXR, as compared to the wild type enzyme (Fernandes and Proteau, 2006).

The other two important amino acid residues: aspartic acid (D150) and glutamic acid (E152) have also been shown to be metal ion-binding residues (Steinbacher *et al.*, 2003). Mutation of aspartic acid to alanine (D152A) results in an inactive enzyme, but change of this residue to asparagines (D152N) shows a 40-fold decrease in activity in the *Synechocystis* sp. PCC6803 DXR, respectively (Fernandes and Proteau, 2006). Moreover, a catalytic efficiency of glutamic acid has been shown in the *Synechocystis* sp. PCC6803 DXR. Fernandes and Proteau (2006) found that mutation of this residue to aspartic acid (E154D) shows a loss of catalytic efficiency (to 0.28% of wild type enzyme),

while the change to glutamine (E154Q) shows a low enzymatic activity (0.008%). These findings suggest the important role of these acidic residues during DXR catalysis by forming ligands to the metal ion (Reuter *et al.*, 2002; Steinbacher *et al.*, 2003).

In addition to metal ion-binding residues, the role of other important residues in DXR are also shown. For example, Kuzuyama *et al.* (2000b) found that change of histidine residue (H153) in the *E. coli* DXR to glutamine (H155Q) shows a 36-fold decrease in the catalytic efficiency, suggesting that this residue is involved in the catalysis. In contrast, change of this residue to alanine (H155A) in the *Synechocystis* sp. PCC6803 DXR yields a similar catalytic efficiency in both wild type and the mutant enzymes, suggesting that this residue is not critical for the *Synechocystis* sp. PCC6803 DXR (Fernandes and Proteau, 2006). It has been proposed that this residue plays a structural role, since it is not directly in contact with the substrate but is surrounded by conserved aspartic and glutamic acid residues (Yajima *et al.*, 2002). Moreover, two additional mutations of histidine residues, H209Q and H257Q, have even more dramatic effects on the activity, since these residues have been shown to be involved in a hydrogen-bonding to the phosphate group of DXP molecule (MacSweeney *et al.*, 2005). Elimination of hydroxyl group of serine (153) in the *Synechocystis* sp. PCC6803 DXR by mutation to alanine (S153A) results in a 9000-fold decrease in the catalytic efficiency of mutant enzymes, but increases in K_m values for DXP and NADPH, suggesting an important role of this residue for obtaining optimal arrangement of active site residues for catalysis (Fernandes and Proteau, 2006). Based on the crystal structures, the hydroxyl group of S153 forms a hydrogen bond with the C2 carbonyl oxygen of DXP (Steinbacher *et al.*, 2003; MacSweeney *et al.*, 2005).

The other important amino acids found in the flexible loop, such as methionine and tryptophan residues that act in the binding of DXP, have also been studied. When the methionine residue (M206) in the *Synechocystis* sp. PCC6803 DXR (or M214 in the *E. coli* DXR) was changed to valine (M206V), the mutant enzymes showed no detectable activity, while a 25-fold increase in the K_m values for DXP of the mutant enzyme (M206A) was observed (Fernandes and Proteau, 2006). In addition, mutants of tryptophan residue (W204) in the *Synechocystis* sp. PCC6803 DXR (or W212 in the *E. coli* DXR), such as W204L, W204V, and W204A, show an increase in K_m for DXP, and a decrease in the catalytic efficiency. These results confirm that these two residues play an important role in the binding of substrate (Fernandes *et al.*, 2005), which is similar to the results obtained based on the X-ray structure (MacSweeney *et al.*, 2005).

7. *Hevea brasiliensis*

7.1 Taxonomic classification

H. brasiliensis (Willd. *ex* Adr. Juss.) Müll. Arg. was discovered in 1824 by French botanist named Antoine de Jussieu. The previous scientific name of *H. brasiliensis* was *Siphonia brasiliensis*. In 1865, Johannes Müller (Argoviensis), a botanist from Switzerland replaced it under the genus *Hevea*, belonging to family Euphorbiaceae. Common names of *H. brasiliensis* are known as: rubber tree, arbre de para, Brasilian rubber tree, and siringa (http://www.rain-tree.com/rubber.htm). The taxonomic classification of rubber tree is shown as follows:

Kingdom: Plantae
Division: Magnoliophyta
Class: Rosopsida
Order: Euphorbiales
Family: Euphorbiaceae
Genus: *Hevea*
Species: *Hevea brasiliensis*

7.2 Botanical description

H. brasiliensis (2n = 36) is a tropical tree found in the tropical forest of the Amazon basin, Brazil and adjoining countries in South America, such as Bolivia, Ecuador, Peru and Venezuela. *Hevea* trees can grow up to 40 m tall, with a foliose crown. The trunk is cylindrical with an uncommonly swollen shape at the base. Latex is a white or creamy color and is produced from the laticifers, which are uniquely differentiated cells of the longitudinal elements of the phloem component of the vascular system. Leaves are compound consisting of three epilliptical leaflets with 10-15 cm long and 3-6 cm broad (Fig. 1.9). Flowers are unisexual with male and female reproductive parts found on separate flowers. The male flower is 5 mm in length with 10 sessile anthers (Fig. 1.9) and the female flower is 8 mm in length with a green disc at base (Fig. 1.9). Fruits have a 3-lobed capsule with 1 seed per capsule (Fig. 1.9). Seeds are variable in size with a hard outer coat (Fig. 1.9). The embryo is surrounded by cotyledon and oily endosperm.

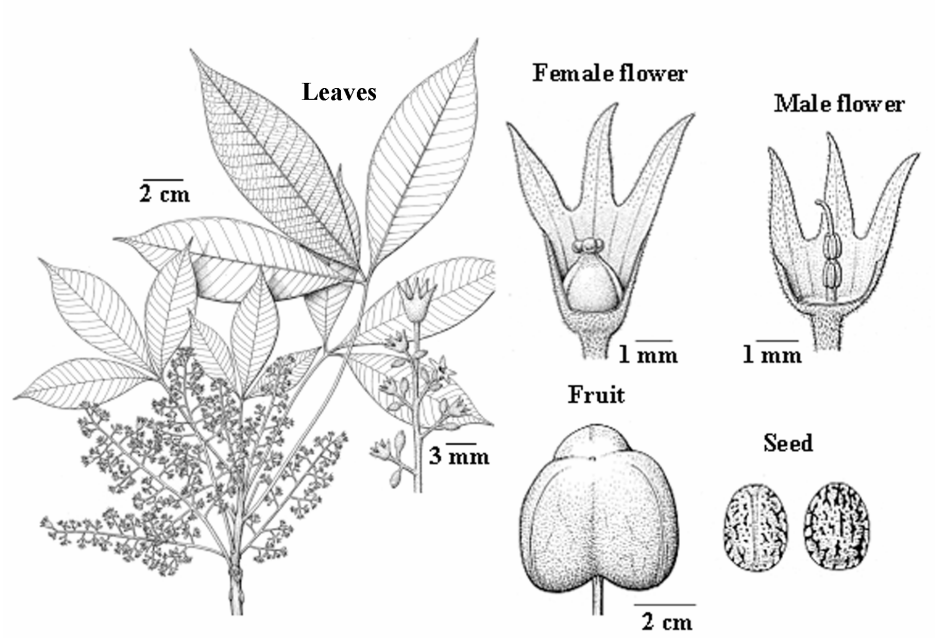


Fig. 1.9 ***Hevea brasiliensis*** **(Willd.** ***ex*** **Adr. Juss.) Müll. Arg.**

This picture was obtained and modified from the http://www. nationalherbarium.nl/Thai Euph/ThSearchVT/ThSearchVTT.

7.3 Latex of *H. brasiliensis*

Latex is a specialized cytoplasm of the laticifers obtained by cutting the bark of *H. brasiliensis* tree. Latex contains 30-40% of rubber particles along with other non-rubber compounds, such as proteins, carbohydrates, lipids, nucleic acid, and phospholipids. By using a high speed centrifugation, latex can be separated into four fractions (d'Auzac and Jacob, 1989). The first fraction is a white fraction containing most of rubber particles. The second fraction is an orange or yellow fraction consisting of Frey-Wyssling particles. Frey-Wyssling particles are chromoplast-liked organelles that are rich in carotenes, resulting in the orange or yellow color of this part (de Faÿ *et al*., 1989). The third fraction is a colorless serum called clear serum or C-serum. It contains proteins, nucleic acids (DNA and RNA), and other organic substances. The last fraction called bottom fraction contains mainly lutoid particles. This fraction has been shown to play an important role in the biosynthesis of rubber (Wititsuwannakul *et al.*, 2003; Wititsuwannakul *et al.*, 2004).

7.4 Biosynthesis of natural rubber

The biosynthesis of natural rubber is a side-branch of the ubiquitous isoprenoid pathway. Natural rubber is made almost entirely of isoprene units derived from IPP and its isomer DMAPP. Also, *trans*-allylic diphosphates substrates, such as GPP, farnesyl diphosphate (FPP), geranylgeranyl diphosphate (GGPP), and others, are essential for rubber formation as they are used to initiate all new rubber molecules. The biosynthesis of natural rubber is considered to consist of two parts (Steinbüchel, 2003). First, the biochemical step essential for the biosynthesis of natural rubber is the isomerization of IPP to DMAPP by the enzyme IPP isomerase (Fig. 1.10). Second, the polymerization of IPP in the *trans*-configuration to DMAPP yields C_{10} (GPP), C_{15} (FPP), and C_{20} (GGPP) allylic pyrophosphates, catalyzed by the enzyme prenyltransferase. Once the initiator is in place, rubber transferease can then begin the *cis*-elongation of isoprene units from IPP to form a high-molecular weight of *cis*-1,4-polyisoprene (Fig. 1.10) (Kekwick, 1989). The biosynthesis of natural rubber take place on the surface of rubber particles suspended in the latex. The highest quality rubber has a molecular weight of around 1.5 million (Steinbüchel, 2003).

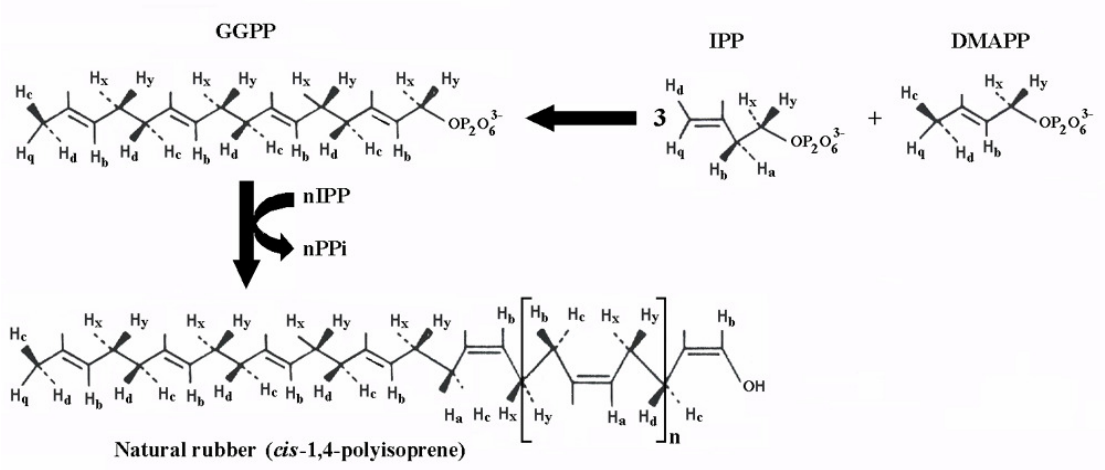


Fig. 1.10 Schematic representation of the biosynthesis of natural rubber.

Natural rubber is synthesized by the polymerization of IPP to allylic diphosphates substrates i.e. GGPP (adapted from Kekwick, 1989).

The MVA pathway has been conventionally studied as a major pathway for the biosynthesis of natural rubber, since it is believed that the pathway for the biosynthesis of natural rubber is similar to the biosynthesis of cholesterol in animals (Balasubramaniam *et al.*, 1977). Evidence supporting the MVA pathway participated in the biosynthesis of natural rubber is derived from a high incorporation level of radiolabelled intermediates such as mevalonate (Skilleter and Kekwick, 1971) and HMG CoA (Hepper and Audley, 1969) into rubber molecules. Further support is provided by the finding that activities of the first two enzymes of the MVA pathway, HMGS and HMGR, are correlated well with the latex yield (Wititsuwannakul *et al.*, 1988; Suvachittanont and Wititsuwannakul, 1995) and their corresponding genes are more abundantly expressed in the latex than in other tissues (Kush *et al.*, 1990; Chye *et al.*, 1992; Suwanmanee *et al.*, 2002; Suwanmanee *et al.*, 2004; Sirinupong *et al.*, 2005). As natural rubber is a polymer of isoprene, the plastidic MEP pathway has recently been considered as a possible alternative route for the biosynthesis of natural rubber. In 2003, Ko *et al.* discovered the *DXS* in latex of *H. brasiliensis*, suggesting that the MEP pathway might provide an alternative means for generating IPP molecule for the biosynthesis of natural rubber together with the well-known MVA pathway.

7.5 Enhanced rubber production by ethephon

Biosynthesis of natural rubber, like other secondary metabolites, is affected by plant hormones. One of the extensively studied plant hormones in *H. brasiliensis* in relation to rubber biosynthesis and latex production is ethylene, which is applied as ethephon (2-chloroethylphosphonic acid). Ethephon is widely used to date to enhance commercially rubber production in *H. brasiliensis*. Treatment of *H. brasiliensis* bark with ethephon increases volume of the exported latex around 1.5- to 2-fold and also stimulates the latex generation between tappings (Coupé and Chrestin, 1989; Pujade-Renaud *et al.*, 1994; Suwanmanee *et al.*, 2004). Although the exact mechanism of action of ethylene on *H. brasiliensis* is not clearly understood, several physiological and biochemical evidences have been shown that ethylene increases the duration of latex flow but delays the plugging process. The delay in the coagulation process is thought to be due to the stabilization of lutoids, hence reducing the plugging of latex (Coupé and Chrestin, 1989).

A number of investigations have been undertaken to study the expression pattern of various rubber biosynthesis related genes, such as HMGS, HMGR, FPP synthase, and rubber elongation factor under ethephon treatments. It was found that the ethylene can induce the transcripts of these genes (Chye *et al.*, 1992; Suwanmanee *et al.*,

2004; Priya *et al.*, 2006; Priya *et al.*, 2007), but not induce the FPP synthase (FDS) gene (Adiwilaga and Kush, 1996). It is presumed that ethylene affects the genes involved in rubber biosynthesis directly or indirectly by modifying general nitrogen metabolism in the latex. In this context, nitrogen metabolism is thought to be involved in protein and nucleic acid synthesis (Coupé and Chrestin, 1989; Pujade-Renaud *et al*., 1994).

Objectives

The objectives of this study were to:

(1) Establish the cDNA sequences of *HbDXSs* and *HbDXRs* genes from *H. brasiliensis*,

(2) Characterize the deduced amino acid sequences of HbDXSs and HbDXRs,

(3) Study the expression pattern of *HbDXSs* and *HbDXRs* genes in different tissues of *H. brasiliensis*,

(4) Identify the expression pattern of *HbDXS2* and *HbDXR2* transcripts in different *H. brasiliensis* clones (RRIM 600 and wild type), and in wild type plants under the ethephon treatment,

(5) Express the HbDXS1 and HbDXR1 recombinant proteins in *E. coli*,

(6) Preliminarily characterize the HbDXS1 and HbDXR1 polyclonal antibodies.

CHAPTER 2

MATERIALS AND METHODS

Materials

1. Plant materials

H. brasiliensis clones RRIM 600 and wild type, collected from local rubber plantation and from Prince of Songkla University, Thailand, respectively, were used in this study. Clone RRIM 600 was cut every second day, while the wild type plants were cut for a few days before performing the experiments to ensure that the amount of latex from each tree was constant. Seedlings of RRIM 600 cultivars were also collected from the same local rubber plantation, Hat Yai, Thailand. Various tissues of seedlings (roots, stems, green leaves, and etiolated leaves) and mature plants (flowers, normal leaves, and naturally *Oidium heveae* infected leaves) were harvested and stored at -80ºC for RNA isolation.

2. Bacterial strains

There are four strains of *E. coli* used in this study: (1) the strain XL1-Blue MRF´ [Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*)173 *endA*1 *supE*44*thi-1 recA*1 *gyrA*96 *relA*1 *lac*[*F´proAB lacI*^{q} *Z*ΔM15 Tn10(tet^{r})] was purchased from Stratagene (USA), (2) the strain DH5α [F^{-}, Φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(rk^{-}, mk^{+}), *phoA*, *supE*44, λ^{-}, *thi*-1, *gyrA*96, *relA*1] was from Invitrogen (USA), (3) the strain TOP10 [F^{-} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*Z.M15Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^{R}) *end*A1 *nup*G] was from Invitrogen, and (4) the strain BL21 (DE3) [*E. coli* B F^{-} *ompT gal* (*E. coli* B is naturally *dcm and lon*) $hsdS_B$ with DE3, a λ prophage carrying the T7 RNA polymerase gene and lacI^{Q})] was from Novagen (USA).

3. Vectors

The T-overhang vectors used for cloning of RT-PCR and PCR products were: pDrive (QIAGEN, Germany), pCR4-TOPO (Invitrogen), and pGEM-T Easy vectors (Promega, USA). The expression vectors used for recombinant protein production were: pET28a (Novagen, USA) and pQE30 (QIAGEN).

4. Primers

The degenerate and gene-specific primers were synthesized by either Gibco BRL (USA), QIAGEN, or IDT (Integrated DNA Technology, Inc., USA).

5. Enzymes

The restriction enzymes were purchased from New England Biolabs (USA). *Taq* DNA polymerase was obtained from Promega. Proofreading *Taq* DNA polymerases (*Pfu* and platinum *Pfx*) were from Promega and Invitrogen, respectively. T4 DNA ligase was from New England Biolabs. RNase A and DNase A were purchased from Sigma-Aldrich (USA) and QIAGEN, respectively.

6. Kits

Kits for total RNA purification (RNeasy Plant Mini Kit), for RT-PCR and PCR amplification (One-step RT-PCR Kit and *Taq* PCR Master Mix Kit), for PCR product purification (QIAquick PCR Purification Kit and QIAquick Gel Extraction Kit), and for plasmid DNA isolation (QIAprep Spin Miniprep Kit) were purchased from QIAGEN. The GeneRacer kit for 5´- and 3´-ends of cDNA amplification was from QIAGEN. The ABI PRISM Big Dye dideoxy nucleotide chain Terminator v 1.1 cycle sequencing kit was from PE Applied Biosystems (USA). Probe labeling (Gene Images Random Prime Labeling Module) and detection (Gene Images CDP-Star Detection Module) kits were from GE Healthcare (UK). The ECL Western blotting detection system was from GE Healthcare. The ImProm-II Reverse Transcription System was from Promega.

7. Chemicals and reagents

All chemicals and reagents used in this study were of analytical grade and were purchased from several companies, such as Fermentas (USA), GE Healthcare, Fluka (USA), Merck (Germany), Pierce (USA), Quantum Bioprobe (Canada), and Sigma-Aldrich.

8. Instruments

Instrument	Model	Company
Analytical balance	Precisa 300A	Precisa balance
Autoclave	ES-315	Tomy
Automated DNA sequencer	377	PE Applied Biosystems, USA
Bacteria incubator	70	Enlab Limited, USA
Centrifuge (refrigerated)	5804R	Eppendorf, Germany
Gel document	VisiDoc-It system	UVP Inc., USA
Heat box	Multi-Block	LAB-Line, USA
Electrophoresis (horizontal)	i-Mupid	Cosmo Bio Co. LTD, Japan
Electrophoresis (ertical)	XCell SureLock Mini-Cell	Invitrogen, USA
Hybridization incubator	Shack in Stack	Hybaid, UK
Mastercycler-PCR	5333	Eppendorf, Germany
Microcentrifuge (refrigerated)	TJ-6	Eppendorf, Germany
Microcentrifuge (non-refrigerated)	Spectrafuge 16M	Labnet, USA
Protein blotter	Novex Semi-Dry Blotter	Invitrogen, USA
Power supply	1000/500	Bio-Rad, USA
Spectrophotometer	UV 160A	Shimadzu, Japan
UV transilluminator		Vilber Lourmat, France
Vortex-mixture	G-560E	Scientific Industries, USA
Water bath	EcoTempTW20	Julabo, USA
X-ray film processor	-	Kodak, USA

Methods

1. Ethephon treatment

Bark of wild type *H. brasiliensis* plants were scraped off with a tapping knife in a semicircular direction. Then, the bark area underneath the tapping cut was spread with 2.5% of ethephon (2-chloroethylphosphonic acid) in water with a brush.

2. Collection of latex from *H. brasiliensis* and determination of rubber content

First few drops of latex from each tapping were discarded in order to eliminate the cellular debris and other contaminants. The latex from each tapping was collected separately in an ice-cooled beaker, kept on ice, and then transferred to the laboratory for determination of the dry rubber content and isolation of total RNA. In the ethephon treatment, latex was collected from the trees before treatment (day 0) and after treatment (day 1, day 3, and day 5, respectively). To determine the rubber content, aliquots of latex (exact volume of 5-10 ml) from each tapping were poured in a 90 mm Petri dish and dried at 80ºC for 2 days to a constant weight. Rubber content was determined as previously described by Suvachittanont and Wititsuwannakul (1995). It was assumed that total solid content after drying reflects the rubber content of latex since rubber is the major component of the latex.

3. Isolation of total RNA

Total RNA was isolated from various tissues of *H. brasiliensis* by using the phenol/SDS method previously described by Suwanmanee *et al.* (2004) with some modifications as described below.

3.1 Isolation of total RNA from latex

About 5-10 ml of fresh latex from each tapping were transferred from an ice-cooled beaker (section 2) into a 50 ml conical tube containing equal volumes of RNA extraction buffer [100 mM Tris-HCl (pH 9.0), 100 mM LiCl, 10 mM EDTA and 1% SDS (w/v)], and 50% saturated phenol (pH 8.0). The mixture was mixed thoroughly with a glass rod and the resultant coagulated rubber particles were discarded. Then, equal volumes of phenol, chloroform and isoamyl alcohol (25:24:1, v/v/v) was added into the mixture, and it was vigorously shaken by inverting the tube for 20 min, and centrifuged in a refrigerated microcentrifuge (model TJ-6, Eppendorf, Germany) at 7000 x *g* at 4ºC for 20 min. The upper-aqueous phase was collected and transferred into a new 15 ml conical

tube, mixed with equal volumes of phenol, chloroform and isoamyl alcohol (25:24:1, v/v/v), vigorously shaken, and centrifuged as described above. The upper-aqueous phase was then collected and extracted with equal volumes of chloroform and isoamyl alcohol (24:1, v/v) twice. After centrifugation, the upper-aqueous phase was collected and transferred into a 1.5 ml microcentrifuge tube, mixed with 0.25 volumes of 10 M LiCl, and kept at 4ºC overnight to precipitate total RNA. Total RNA was pelleted by centrifugation using a high speed refrigerated centrifuge (model 5804R, Eppendorf, Germany) at 16000 x *g* at 4ºC for 20 min, washed with 70% ethanol, and dried at room temperature for 15 min. The RNA pellet was further dissolved in diethylpyrocarbonate (DEPC)-treated water, and re-precipitated with 300 mM sodium acetate (pH 5.2) and 2.5 volumes of ethanol at -20ºC overnight. After centrifugation, the RNA pellet was washed twice with 70% ethanol, dried, and re-suspended in DEPC-treated water. The quality and concentration of total RNA was determined by agarose gel electrophoresis and by spectrophotometric analysis (section 4.1 and 4.2, respectively). Total RNA was stored at -80ºC until use.

3.2 Isolation of total RNA from leaves and other tissues

About one g fresh weight of various tissues of *H. brasiliensis* such as leaves, stems, roots, and flowers was ground into a fine powder in liquid nitrogen with a pre-cooled mortar and pestle. The powder was transferred into a 50 ml conical tube. Three milliliters of hot RNA extraction buffer (80ºC) containing 50% of saturated phenol (pH 8.0) was added into the powder, mixed thoroughly by inverting the tube for 20 min by hand, and then centrifuged at 7000 x *g* for 20 min at 4ºC. The upper-aqueous phase was collected and extracted with phenol, chloroform and isoamyl alcohol (25:24:1, v/v/v) as described in section 3.1. Precipitation of total RNA was also done as described in section 3.1. The quality and concentration of total RNA was determined as described in section 4.1 and 4.2, respectively. Total RNA was stored at -80ºC until use.

3.3 Purification of total RNA with the RNeasy Plant Mini Kit

Total RNA isolated from section 3.1 and 3.2 was further purified using the RNeasy Plant Mini Kit (QIAGEN) according to the manufacturer's instructions. Aliquots of total RNA (100 µl) were mixed with 350 µl of buffer RLT containing guanidine isothiocyanate and β-mercaptoethanol and 225 µl of ethanol (95%) by pipetting. The mixture was then applied into the RNeasy mini column, and centrifuged using microcentrifuge (Spectrafuge 16M, Labnet, USA) at 8000 x *g* at room temperature for 30 s. The flow-through was discarded, and the column was washed once with 700 µl of

buffer RW1 containing ethanol. The contaminating genomic DNA in the RNA samples was removed by adding DNase I solution (3 Kunitz units, QIAGEN) directly onto the RNeasy mini column, and incubated at room temperature for 20 min followed by washing once with 350 µl of buffer RW1. The column was transferred into a new microcentrifuge tube, and then washed twice with 500 µl of buffer RPE containing 80% ethanol. Total RNA was eluted from the column with 50 µl of RNase-free water by centrifugation at 12000 x *g* at room temperature for 2 min. The quality and concentration of total RNA was determined as described in section 4.1 and 4.2, respectively. The purified RNA was stored at -80ºC until use.

4. Analysis of nucleic acids

The quality and concentration of nucleic acids, such as total RNA, genomic DNA and PCR products as well as plasmid DNA were determined by agarose gel electrophoresis and by spectrophotometric analysis as described below.

4.1 Agarose gel electrophoresis

Most agarose gels used in this study were made between 1.0% and 1.5% (w/v) of agarose. The gel was prepared by mixing an appropriate amount of agarose and 50 ml of 1X TAE buffer [40 mM Tris-acetate and 1 mM EDTA (pH 8.0)] in a 250 ml conical flask and boiled on a hot plate to dissolve the agarose. The molten agarose was then cooled down to 55ºC, and mixed with 1 µl of ethidium bromide (10 mg/ml). The agarose solution was poured into a gel tray, inserted a comb, and left at room temperature for at least 30 min (preferably 1 h) to set the gel. After removing the comb, the gel was transferred into an electrophoresis chamber (i-Mupid, Cosmo Bio Co. LTD, Japan) and submerged in 1X TAE buffer. Aliquots of samples (5-10 µl) mixed with 1 µl of 6X loading dye [0.25% (w/v) Bromophenol Blue and 4% (w/v) sucrose], were loaded into the gel, and then run at 75 V for 45 min to 1 h. The expected RNA or DNA bands were visualized under ultraviolet (UV) light using either the UV transilluminator (Vilber Lourmat, France) or the gel documentation system (VisiDoc-It system, UVP Inc., USA).

4.2 Spectrophotometric analysis

The absorption of UV light of the ring structure of purine and pyrimidine bases can be used to measure the amount of total RNA and DNA using a spectrophotometer. Aliquots of samples (5 µl) were diluted with 1 ml of sterile water. The absorbance at 260 and 280 nm was read in a spectrophotometer (model UV 160A, Shimadzu, Japan). The concentrations of total RNA or DNA were calculated based on an

OD_{260} of one corresponding to approximately 40 µg/µl for RNA and 50 µg/µl for double-stranded DNA. The ratios of OD_{260}/OD_{280} for total RNA and DNA were taken as 2.0 and 1.8, respectively. Contamination by protein or phenol in the samples will cause the ratio to be significantly lower than 1.8. The RNA samples with A_{260}/A_{280} ratio from 1.6 to 2.0 were used in the cDNA synthesis reaction.

5. One-step RT-PCR amplification of the genes of interest

Partial cDNA fragments containing *DXS* and *DXR* were amplified from *H. brasiliensis* by using the RNA-based RT-PCR methods and the degenerate primers, which were designed based on the conserved regions of the nucleotide sequences of *DXS* and *DXR* from various plant species. The designs of degenerate primers and PCR protocols are described as follows.

5.1 Designing the degenerate primers

The nucleotide sequences encoding DXS from GenBank database were selected from *A. thaliana* (Mandel *et al.*, 1996), *C. annuum* (Bouvier *et al.*, 1998), and *L. esculentum* (Lois *et al.*, 2000), and those for DXR were from *M. piperita* (Lange and Croteau, 1999), *A. thaliana* (Schwender *et al.*, 1999), *A. annua* (Souret *et al.*, 2002), and *Z. mays* (Hans *et al.*, 2004). The selected nucleotide sequences were subsequently aligned using ClustalX software with default parameters (Thompson *et al.*, 1997) to find the most highly conserved regions of these two genes. From these sequence data, the amino acid sequences WDVGHQ (WDV-DXS1) and IAEQHA (IAE-DXS2) were selected to synthesize degenerate primers for amplification of *DXS* isoform I (*HbDXS1*), and amino acid sequences PIHMKN (PIH-DXS3), YFAEAL (YFA-DXS4), MISGSG (MIS-DXS5) and ITVEEG (ITV-DXS6) were selected for amplification of *DXS* isoform II (*HbDXS2*), respectively. The amino acid sequences GSIGTQ (GSI-DXR1) and GWPDMR (GWP-DXR2) were selected for amplification of both *DXR* isoform I (*HbDXR1*) and isoform II (*HbDXR2*). The nucleotide sequences corresponding to these amino acid sequences including their melting temperature, Tm (ºC) are shown in Table 2.1. All of these degenerate primers were commercially synthesized by Gibco BRL and IDT.

Table 2.1 Nucleotide sequences of degenerate primers for amplification of partial *DXS* and *DXR* cDNAs from *H. brasiliensis*.

Primer names	Nucleotide sequence (5´→3´)	Tm (ºC)
WDV-DXS1	TGGGATGTTGGTCATCAGTC	54.5
IAE-DXS2	ACTGC(G/A)TGTTGTTC(C/T)GCTAT	54.7
PIH-DXS3	CC(A/T)(G/A)TTCACATGAA(G/A)AA(C/T)CT	49.1
YFA-DXS4	ATCAGAGCTTCAGCAAAGTA	51.3
MIS- DXS5	ATGAT(C/T)AGTGGTTCTGGATC	50.4
ITV-DXS6	GA(C/T)CCTTCTTC(G/A/T)AC(G/C/A)GTGAT	52.9
GSI-DXR1	GG(C/T)TC(C/T)ATTGG(A/C)AC(C/T)CAGAC	54.3
GWP-DXR2	AACCGCATATC(G/A)GGCCA(A/T)CC	59.1

5.2 Amplification of target cDNAs

Amplification of partial cDNA fragments containing *DXS* and *DXR* sequences was performed using the One-step RT-PCR Kit (QIAGEN) according to the manufacturer's instructions. The reaction mixture consisted of 10 µl of 5X QIAGEN OneStep RT-PCR buffer containing 12.5 mM $MgCl_2$, 2 µl of dNTP mix (10 mM of each dNTP), 5 µl of each primer (10 pmole/µl), 2 µl of QIAGEN OneStep RT-PCR enzyme mix containing Omniscript RT, Sensiscript RT, and HotStarTaq DNA polymerase, 1 µg of total RNA isolated from leaves and latex as templates, and RNase-free water to a total volume of 50 µl. RT-PCR was performed on Mastercycler (model 5333, Eppendorf, Germany) using the following conditions: reverse transcription at 50ºC for 30 min, initial PCR activation step at 95ºC for 15 min followed by 35 cycles of 94ºC for 1 min, 55ºC for 1 min, 72ºC for 1 min, and a final extension at 72ºC for 7 min. The resulting RT-PCR products were resolved on 1.5% agarose gel electrophoresis and visualized under the UV light (section 4.1).

6. Purification of PCR products

The amplified cDNA fragments obtained from RT-PCR or PCR reactions can be further purified directly from the reaction mixture or after running the agarose gel electrophoresis using the QIAquick PCR Purification Kit (QIAGEN), and the QIAquick Gel Extraction Kit (QIAGEN), respectively, as described below.

6.1 QIAquick PCR Purification Kit

Aliquots of RT-PCR or PCR products (100 µl) were mixed with 500 µl of buffer PB containing guanidine hydrochloride by pipetting. The mixture was then applied into the QIAquick spin column and centrifuged at 10000 x *g* at room temperature for 1 min. The flow-through was discarded and the column was washed with 750 µl of buffer PE containing 80% ethanol twice. The DNA fragments were eluted from the column with 50 µl of buffer EB [10 mM Tris-HCl (pH 8.5)] by centrifugation at 12000 x *g* at room temperature for 2 min. The concentration of DNA was determined as described in section 4.2 and stored at -20ºC until use.

6.2 QIAquick Gel Extraction Kit

The expected cDNA bands in agarose gel after the electrophoresis were excised using a sterile razor blade and transferred into a 1.5 ml microcentrifuge tube. The gel was dissolved in 500 µl of buffer QG containing guanidine thiocyanate (pH 7.5), and incubated at 50ºC for 15 min with vortexing every 5 min during incubation. The mixture was then applied into the QIAquick spin column and centrifuged at 10,000 x *g* at room temperature for 1 min. The flow-through was discarded and the column was washed once with 500 µl of buffer QG and washed twice with 750 µl of buffer PE containing 80% ethanol. The DNA fragments were eluted from the column as described above (section 6.1). After determination of the concentration of DNA (section 4.2), the DNA samples were stored at -20ºC.

7. Cloning of PCR products into T-overhang vectors

All PCR products amplified using *Taq* DNA polymerase and other non-proofreading DNA Polymerases contain additional A-overhang at the 3´-ends. The advantage of having the A-overhang at the 3´-terminal ends was to clone all of these PCR products directly into the T-overhang vectors. The T-overhang vectors used in this study were pDrive (QIAGEN), pCR4-TOPO (Invitrogen), and pGEM-T Easy vectors (Promega). The protocols for the ligation of these PCR products into T-overhang vectors were according to the manufacturer's instructions.

7.1 Ligation of purified PCR products into pDrive vector

Ligation of the purified cDNA into pDrive vector was carried out in a total volume of 10 µl, containing 1 µl of pDrive vector (50 ng/µl), 4 µl of purified cDNA, and

5 µl of ligation master mix. The reaction was incubated at 16ºC for 1 h or 4ºC overnight, and then transformed into *E. coli* competent cells (section 8.2).

7.2 Ligation of purified PCR products into pCR4-TOPO vector

Ligation of the purified cDNA into pCR4-TOPO vector was carried out in a total volume of 6 µl, containing 1 µl of pCR4-TOPO vector (10 ng/µl), 4 µl of purified cDNA, and 1 µl of salt solution (200 mM NaCl and 10 mM $MgCl_2$). The reaction was incubated at room temperature for 5 min, and then transformed into *E. coli* competent cells (section 8.2).

7.3 Ligation of purified PCR products into pGEM-T Easy vector

Ligation of the purified cDNA into pGEM-T Easy vector was carried out in a total volume of 10 µl, containing 1 µl of pGEM-T Easy vector (50 ng/µl), 3 µl of purified cDNA, 1 µl of T4 DNA ligase, 5 µl of 2X Rapid ligation buffer [60 mM Tris-HCl (pH 7.8), 20 mM $MgCl_2$, 20 mM DTT, 2 mM ATP and 10% polyethylene glycol, MW 8000], and T4 DNA ligase. The reaction was incubated at 16ºC for 1 h or 4ºC overnight, and then transformed into *E. coli* competent cells (section 8.2).

8. Transformation of recombinant plasmid DNA into *E. coli*

Methods for preparation of competent *E. coli* cells and transformation of the plasmid DNA into the competent cells were performed as described by Sambrook *et al.* (1989) and described below.

8.1 Preparation of *E. coli* competent cells

E. coli competent cells were prepared either from strains XL1-Blue MRF´ (Stratagene) or strain DH5α (Invitrogen) or TOP10 (Invitrogen) using $CaCl_2$ method. A single colony of *E. coli* was inoculated into 5 ml of LB medium (10 g Bacto-tryptone, 5 g yeast extract and 10 g NaCl) containing appropriate antibiotics and grown overnight at 37ºC with vigorous shaking at 250 rpm (Bacteria incubator, model 70, Enkab Limited, USA). The culture was diluted in 50 ml of fresh LB medium (1: 50 dilutions) and incubated further at 37ºC with vigorous shaking to an OD_{600} of 0.4. The culture was cooled on ice for 30 min followed by centrifugation at 2000 x *g* at 4ºC for 10 min to collect cells. Cells were gently resuspended in 10 ml of iced-cold 0.1 M $CaCl_2$, incubated further on ice for 30 min, and then centrifuged at 2000 x *g* at 4ºC for 10 min. Cells were gently resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$ containing 15% (v/v) glycerol, aliquoted (100 µl) into a 1.5 ml microcentrifuge tube, and stored at -80ºC.

8.2 Transformation of plasmid DNA

About five µl of the ligation mixture prepared from section 7.1, 7.2, or 7.3 was gently mixed with 100 µl of *E. coli* competent cells and placed on ice for 20 min to allow the adsorption of the recombinant plasmid DNA to cells. The cells were transferred to a heat box (42ºC) (Multi-Block, LAB-Line, USA) for 90s and immediately cooled on ice for 3 min. Warm LB medium (37ºC), 400 µl was then added to the transformed cells and incubated further at 37ºC for 1 h with horizontal shaking at 150 rpm. Aliquots of culture (150 µl) were spread on LB agar plate containing 100 µg/ml of ampicillin, 100 µl of 100 mM isopropyl-β-D-thiogalacto-pyranoside (IPTG) (Fermentas) and 20 µl of 50 mg/ml 5-bromo-4-chloro-3 indolyl-β-D-thiogalactopyranoside (X-gal) (Quantum Bioprobe), and incubated at 37ºC overnight.

9. Selection of transformed cells

Colonies of the transformed cells containing the inserted cDNA were white color when grown on LB-plates containing X-Gal. Recombinant plasmid DNA from white colonies was isolated using the QIAprep Spin Miniprep Kit (QIAGEN) according to the manufacturer's instructions, followed by restriction digestion analysis to confirm that all colonies analyzed contained an insertion of the expected cDNA size.

9.1 Isolation of recombinant plasmid DNA

A single white colony was inoculated into 5 ml of fresh LB medium supplemented with 100 µg/ml of amplicilin and cultured overnight at 37ºC with vigorous shaking at 250 rpm. Cell pellets were collected by centrifugation at 4000 x *g* for 5 min at 4ºC and resuspended in 250 µl of P1 buffer [50 mM Tris-HCl (pH 8.0) and 10 mM EDTA] containing 100 mg/ml of RNase A in a 1.5 ml microcentrifuge tube. The bacterial membrane was then broken with 250 µl of P2 lysis buffer containing chaotropic salts at room temperature for 5 min by inverting the tube several times until the solution becomes viscous and slightly clear. Subsequently, N3 buffer (350 µl) containing chaotropic salts was added to precipitate proteins and followed by centrifugation at 10000 x *g* at room temperature for 10 min. The sample was applied into the QIAprep spin column and centrifuged at 10000 x *g* at room temperature for 1 min. The flow-through was discarded. The column was washed twice with 750 µl of PE buffer containing 80% ethanol. Recombinant plasmid DNA was eluted from the column with 50 µl of EB buffer, followed by centrifugation at 12000 x *g* at room temperature for 2 min. The concentration of recombinant plasmid DNA was determined as described in section 4.2 and stored at -20ºC.

9.2 Restriction enzyme digestion analysis

Recombinant plasmid from section 9.1 was digested with restriction enzyme *Eco*RI (New England Biolabs). The digestion reaction contained 2 µl of recombinant plasmid DNA (500 ng/µl), 2 µl of *Eco*RI buffer [50 mM NaCl, 100 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and 0.025% Triton X-100], 0.5 µl of *Eco*RI (20,000 units/ml), and sterile water to a total volume of 20 µl. The reaction mixture was incubated at 37ºC for 5 h and aliquots of samples (5 µl) were resolved on 1.5% agarose gel electrophoresis and visualized under UV light (section 4.1).

10. Automated DNA sequencing

Recombinant plasmid containing the cDNA of interest was sequenced by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) using the ABI PRISM Big Dye dideoxy nucleotide chain Terminator v 1.1 cycle sequencing kit (PE Applied Biosystems) according to the manufacturer's instructions. PCR reactions were carried out in a total volume of 20 µl, consisting of 1 µl of Big Dye sequencing buffer, 2 µl of Big Dye Terminator ready reaction mix, 5 µl of plasmid DNA (100 ng/µl), 1 µl of primer (1.6 pmole), and deionized water using the following profiles: 96ºC for 30 s and 25 cycles of 96ºC for 10 s, 55ºC for 5 s, and 70ºC for 4 min. The PCR products were purified by ethanol precipitation. Aliquots of PCR products (10 µl) were mixed with 2.5 µl of 125 mM EDTA and 30 µl of absolute ethanol (96%). The mixture was left standing at room temperature for 15 min and centrifuged at 16000 x *g* at 4ºC for 20 min. After washing with 70% ethanol, the DNA pellet was dried at 90ºC for 1 min, and immediately analyzed on ABI PRISM-System (model 377, PE Applied Biosystems) or stored at -20ºC for further analysis.

11. Amplification of 5´- and 3´-ends of cDNAs

The 5´- and 3´-ends of the cDNAs of interest were obtained using the RNA ligase-mediated and oligo-capping rapid amplification of cDNA (RLM-RACE), according to the method described in the GeneRacer Kit (Invitrogen) (Fig. 2.1).

11.1 Construction of first-strand cDNA

11.1.1 Dephosphorylation of uncapped RNA

Aliquots of total RNA (5 µg) prepared in section 3.3 were treated with Calf Intestinal Phosphatase (CIP) to remove 5´-terminal phosphates of any RNA (i.e. mRNA, truncated mRNA and non-mRNA) at 50ºC for 1 h. The reaction contained 1 µl of 10X CIP buffer [0.5 M Tris-HCl (pH 8.5) and 1 mM EDTA], 1 µl of RNaseOut (40 units/µl), 1 µl of CIP [10 units/µl in 25 mM Tris-HCl (pH 7.6), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, and 50% glycerol (w/v)], and DEPC-treated water to give a total volume of 10 µl. The mixture was then brought to 100 µl with DEPC-treated water and extracted once with an equal volume of phenol, chloroform and isoamyl alcohol (25:24:1, v/v/v), vortexed vigorously for 30 s, and centrifuged at 16000 x *g* at 4ºC for 5 min. The upper-aqueous phase (100 µl) was transferred into a new 1.5 ml tube, mixed with 2 µl of 10 mg/ml mussel glycogen, 10 µl of 3 M sodium acetate, pH 5.2 and 220 µl of 96% ethanol, and put on dry ice for 10 min. The RNA was pelleted by centrifugation at 16000 x *g* at 4ºC for 20 min and washed with 500 µl of 70% ethanol by inverting the tube several times followed by centrifugation at 16000 x *g* at 4ºC for 2 min. The RNA pellet was air-dried for 2 min at room temperature, and then re-suspended in 7 µl of DEPC-treated water.

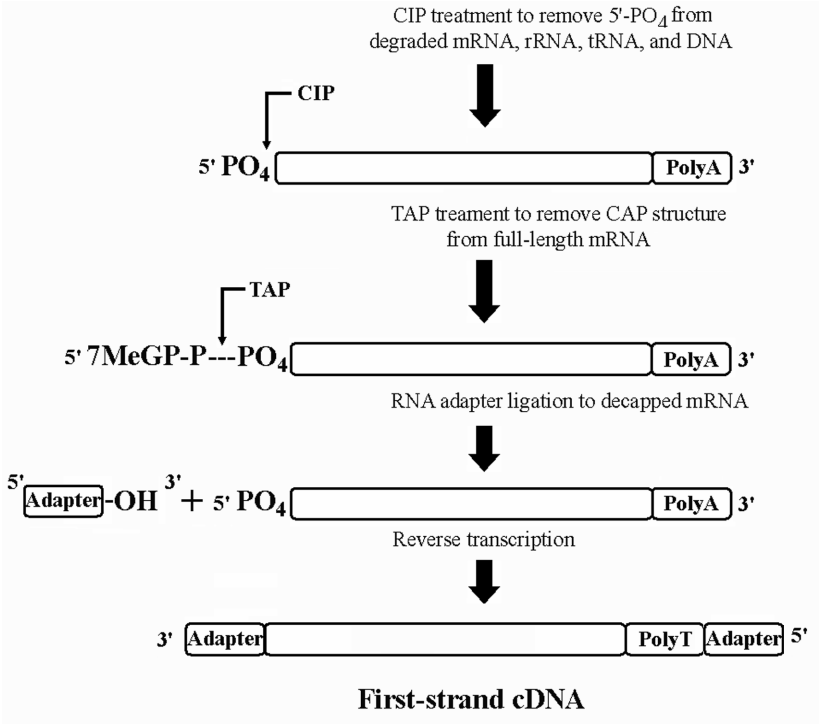


Fig. 2.1 Diagrammatic representation of first-strand cDNA synthesis for amplification of 5´- and 3´-ends of cDNA.

First-strand cDNA for amplification of 5´- and 3´-ends was synthesized according to the RLM-RACE protocol (Invitrogen).

11.1.2 Decapping of RNA molecules

The 7-methyl guanisine triphosphate cap-structure of the de-phosphorylated RNA was removed by Tobacco Acid Pyrophosphatase (TAP). The reaction contained 1 µl of 10X TAP buffer [0.5 M sodium acetate (pH 6.0), 10 mM EDTA, 1% β-mercaptoethanol, and 0.1% Triton X-100], 1 µl of RNaseOut (40 units/µl), 1 µl of TAP [0.5 units/µl in 10 mM Tris-HCl, pH 7.5, 0.1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 0.01% Triton X-100, and 50% glycerol (w/v)], and DEPC-treated water to give a total volume of 10 µl. The mixture was incubated at 37ºC for 1 h followed by phenol/chloroform extraction and ethanol precipitation as described in section 11.1.1. The RNA pellet was resuspended in 7 µl of DEPC-treated water.

11.1.3 Ligation of an RNA oligo to 5´-end of mRNA

The 5´-ends of the full-length mRNA prepared in section 11.1.2 was ligated to the specific GeneRacer RNA oligo (5´-CGACUGGAGCACGAGGAC ACUGACAUGGACUGAAGGAGUAGAAA-3´, 44 bases) with T4 RNA ligase. The reaction contained 1 µl of 10X ligase buffer [330 mM Tris-Acetate (pH 7.8), 660 mM potassium acetate, 100 mM magnesium acetate, and 5 mM DTT], 1 µl of 10 mM ATP, 1 µl of RNaseOut (40 units/µl), 1 µl of T4 RNA ligase [5 units/µl in 50 mM Tris-HCl (pH 7.5), 0.1 M NaCl, 0.1 mM EDTA, 1 mM DTT, 0.1% Triton X-100, and 50% glycerol (w/v)], and DEPC-treated water to a total volume of 10 µl. The ligation reaction was performed at 37ºC for 1 h followed by phenol/chloroform extraction and ethanol precipitation, as described in section 11.1.1. The RNA pellet was resuspended in 10 µl of DEPC-treated water.

11.1.4 Synthesis of first-strand cDNA

First-strand cDNA was synthesized using GeneRacer RNA oligo dT [5´-GCTGTCAACGATACGCTACGTAACGGCATGACAGTG$(T)_{24}$-3´, 60 bases] as primer. Ten µl of the ligated mRNA prepared in section 11.1.3 was mixed with 1 µl of GeneRacer RNA oligo dT primer (50 µM), 1 µl of dNTP Mix (10 mM each), and 1 µl of DEPC-treated water. The reaction was incubated at 65ºC for 5 min and chilled on ice for 2 min to remove the secondary structure of RNA. Then, the mixture was mixed with 4 µl of 5X RT buffer [250 mM Tris-HCl (pH 8.3), 375 mM KCl, and 15 mM $MgCl_2$], 1 µl of

0.1 M DTT, 1 µl of SuperScript III RT [200 units/µl in 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 0.01% Nonidet P-40 (v/v), and 50% glycerol (w/v)], and 1 µl of RNaseOut (40 units/µl) to a total volume of 10 µl. The mixture was incubated at 50ºC for 1 h, followed by heating at 70ºC for 15 min to inactivate the enzyme SuperScript III RT. Then, 1 µl of RNase H [2 units/µl in 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT, 50 µg/ml BSA, and 50% glycerol (w/v)] was added to the mixture and it was incubated further at 37ºC for 20 min to digest the RNA templates. The first-strand cDNA was used as a template for subsequent PCR reactions.

11.2 Amplification of 5´-end of cDNAs

The 5´-termini of cDNA ends were amplified using *Taq* PCR Master Mix Kit (QIAGEN), according to the manufacturer's instructions. The reaction consisted of 50 µl of *Taq* PCR master mix [1X QIAGEN PCR buffer containing 1.5 mM $MgCl_2$, 2.5 units of *Taq* DNA polymerase and 200 µM of each dNTP], 5 µl of each gene-specific primer 1 (GSP1) (10 pmole/µl) and GeneRacer 5´ primer, 1 µl of first-strand cDNA prepared in section 11.1, and sterile water to a total volume of 100 µl. PCR was carried out using a touch down PCR protocol of three steps (5 cycles of 94ºC for 30 s and 72ºC for 1 min, 5 cycles of 94ºC for 30 s and 70ºC for 1 min, and 25 cycles of 94ºC for 30 s, 62ºC for 30 s and 72ºC for 1 min), and a final extension at 72ºC for 5 min. The first PCR products (5 µl) were subjected to a second PCR using GSP2 of each gene and GeneRacer 5´ nested primer as primers. In case of *HbDXR1*, a second PCR was performed using GSP1 and GeneRacer 5´ nested primer as primers. The GSP for amplification of the 5´-end of each gene (Table 2.2) were designed based on the sequence information obtained from the original One-step RT-PCR clones and synthesized by QIAGEN and IDT. The GeneRacer 5´ and GeneRacer 5´ nested primers were supplied from the GeneRacer Kit.

The 5´-RACE PCR products were resolved on 1.5% agarose gel electrophoresis and visualized under the UV light (section 4.1). The expected cDNA fragments were purified from agarose gel (section 6.2), ligated into pCR4-TOPO (section 7.2) or pGEM-T Easy vectors (section 7.3), and transformed into *E. coli* competent cells strain DH5α or TOP10 (section 8.2). Recombinant plasmids were isolated from positive clones followed by restriction enzyme digestion analysis (section 9.1 and 9.2). The nucleotide sequence was determined using the automated DNA sequencer (section 10).

Table 2.2 Nucleotide sequences of gene-specific primers and GeneRacer primers used in 5´-RACE experiments.

Primer names	Nucleotide sequence (5´→3´)	Tm (ºC)
GSP1-HbDXS1	ACTGGTGGAATGGGTCCATCAAG	64.5
GSP2-HbDXS1	ATTCACTCTCCGACCGCTTCGTG	66.3
GSP1-HbDXS2	AGGCATCATAAACACTCTCATCTC	61.9
GSP2-HbDXS2	CTGAACACATGATGCAGAGCCACTG	66.2
GSP1-HbDXR1	ATCCCTGAAAGCTCCACCTGAAG	64.5
GSP1-HbDXR2	TTGCAGCCACCGTAGGCTTTAAG	64.5
GSP2-HbDXR2	CAGCTGCGAGTGCCACAACTCTG	68.1
GeneRacer 5´	CGACTGGAGCACGAGGACACTGA	68.8
GeneRacer 5´ nested	GGACACTGACATGGACTGAAGGAGTA	66.4

11.3 Amplification of 3´-end of cDNAs

The 3´-termini of cDNA ends were amplified using the first-strand cDNA (section 11.1) as templates. PCR components and cycling parameters were the same as described in section 11.2. The first PCR was performed using GSP1 of each gene and GeneRacer 3´ primer as primers and then aliquots of first PCR products (5 µl) were subjected to a second PCR using GSP2 of each gene and GeneRacer 3´ nested primer as primers. In case of *HbDXS1*, a second PCR was performed using GSP1 and GeneRacer 3´ nested as primers, while the third PCR of *HbDXR2* was carried out using GSP3 and GeneRacer 3´ nested as primers. The GSP for amplification of 3´-end of each gene (Table 2.3) were designed based on the sequence information obtained either from the original One-step RT-PCR clones or 5´-RACE PCR products. The GeneRacer 3´ and GeneRacer 3´ nested primers were also included in the GeneRacer Kit (Invitrogen).

The 3´-RACE PCR products were resolved on 1.5% agarose gel electrophoresis and visualized under the UV light (section 4.1). The expected cDNA fragments were purified from agarose gel (section 6.2), ligated into pCR4-TOPO (section 7.2) or pGEM-T Easy vectors (section 7.3) and transformed into *E. coli* competent cells strain DH5α or TOP10 (section 8.2). Recombinant plasmids were isolated from positive

clones followed by restriction enzyme digestion analysis (section 9.1 and 9.2). The nucleotide sequence was determined using the automated DNA sequencer (section 10).

Table 2.3 Nucleotide sequences of gene-specific primers and GeneRacer primers used in 3´-RACE experiments.

Primer names	Nucleotide sequence (5´→3´)	Tm (ºC)
GSP1-HbDXS1	CATGCTGCAATGGGAGGTGGAAC	66.3
GSP1-HbDXS2	TTGCTTGCCCAACATGGTGGTCA	65.2
GSP2-HbDXS2	GTCAGGAGATTGGCTAAGGAGCA	65.2
GSP1-HbDXR1	CTGCTTCAGGTGGAGCTTTCAGG	66.3
GSP2-HbDXR1	CAGTGGACTCCGCTACCCTTTTC	66.3
GSP1-HbDXR2	CCTTCTTCTGCCATTTTCACCTGCT	64.5
GSP2-HbDXR2	ATGGCGCTCAATTTGCTTTCCCCTGCT	68.1
GSP3-HbDXR2	GGATGGCCCAAAGCCTATTTCAGTC	66.2
GeneRacer 3´	GCTGTCAACGATACGCTACGTAACG	66.2
GeneRacer 3´ nested	CGCTACGTAACGGCATGACAGTG	67.0

12. Generation of full-length open reading frames (ORFs)

By comparing and aligning of the nucleotide sequences obtained from the initial One-step RT-PCR, and the 5´- and 3´-RACE clones, gene-specific primers of each *DXS* and *DXR* isoform were designed. Each gene was amplified using the help of proofreading thermostable DNA polymerase, for example, *Pfu* or platinum *Pfx* DNA polymerases. PCR protocols are described below.

12.1 Amplification of genes of interest

PCR was performed in a 50 µl reaction system consisting of 0.25 µl of *Pfu* DNA polymerase (5 units/µl) (Promega) or 0.5 µl of platinum *Pfx* DNA polymerase (2.5 units/µl) (Invitrogen), 5 µl of 10X PCR buffer, 1 µl of dNTP mixture (10 mM each), and 5 µl of each gene-specific primer (10 pmole/µl). The 10X PCR buffer used for *Pfu* DNA polymerase was [200 mM Tris-HCl (pH 8.8), 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mM $MgSO_4$, 1.0% Triton X-100, and 1 mg/ml nuclease-free BSA] and that for platinum *Pfx* DNA polymerase contained 50 mM $MgSO_4$. Cycling parameters were 94ºC for 3 min, 35 cycles of 94ºC for 30 s, 55ºC (for *HbDXS1*) or 58ºC (for *HbDXS2*) or 65ºC (for *HbDXR1* and *HbDXR2*) for 30 s and 72ºC (for *Pfu* DNA polymerase) or 68ºC (for platinum *Pfx* DNA polymerase) for 2 min, and 72ºC for 5 min in a final extension step. The gene-

specific primers used for amplification of *HbDXS1* were GSP1/2-HbDXS1, for *HbDXS2* were GSP1/2-HbDXS2 and GSP3/4-HbDXS2, for *HbDXR1* were GSP1/2-HbDXR1, and for *HbDXR2* were GSP1/2-HbDXR2, respectively. All of these primers were synthesized by QIAGEN and IDT. The nucleotide sequences and Tm (ºC) of primers were shown in Table 2.4.

The resulting PCR products were resolved on 1.5% agarose gel electrophoresis and visualized under the UV light (section 4.1). The expected cDNA fragments were purified from agarose gel (section 6.2), A-tailed (section 12.2), ligated into pDrive (section 8.1) or pGEM-T Easy vectors (section 7.3) and transformed into *E. coli* competent cells strain DH5α or TOP10 (section 8.2). Recombinant plasmid DNAs were isolated from positive clones followed by restriction enzyme digestion analysis (section 9.1 and 9.2). The nucleotide sequence was determined using the automated DNA sequencer (section 10).

Table 2.4 Nucleotide sequences of gene-specific primers for amplification of full-length genes.

Primer names	Nucleotide sequence (5´→3´)	Tm (ºC)
GSP1-HbDXS1	CCGTGGTTCTCTATTTGCA	58.3
GSP2-HbDXS1	TCCACTTTCTCTATGATGAC	56.3
GSP1-HbDXS2	TCTCACCTTATTTCTCTCACTACAT	59.6
GSP2-HbDXS2	TTGCTTCCCAGTTTGAACATCA	59.6
GSP3-HbDXS2	AGGACCAGTTCTGATTCACAT	59.4
GSP4-HbDXS2	CTTCCAGGAATAGCAAAAGCCAT	61.7
GSP1-HbDXR1	ATGGAGCTTAATTTGCTTTCCCCTGCTGA	66.2
GSP2-HbDXR1	GGAAACACATTTGGGTTGCCCATCTTGCTA	68.0
GSP2-HbDXR2	ATGGCGCTCAATTTGCTTTCCCCTGCT	68.1

12.2 A-tailing protocol for blunt-end PCR fragments

Since the proof reading enzymes do not contain the 3´→5´ exonuclease activity, the PCR products amplified by these enzymes can be successfully cloned into all of the T-overhang vectors as followed. The method was first described by Promega. The reaction consisted of 7 µl of PCR products, 1 µl of 10X reaction buffer without $MgCl_2$ [500 mM KCl, 100 mM Tris-HCl (pH 9.0), 1% Triton X-100], 1 µl of 2 mM dATP, and 1

µl of *Taq* DNA polymerase (5 units/µl, Promega). The mixture was incubated at 70ºC for 30 min followed by ligation into T-overhang vectors (section 7) immediately.

13. Bioinformatics analysis

Analysis of the nucleotide and deduced amino acid sequences were performed online as described below. Homology searches were carried out using Basic Local Alignment Search Tool (BLAST) program at the NCBI server (*http://www.ncbi.nlm.nih.gov/BLAST*) (Altschul *et al.*, 1997). Organelle localization and cleavage site were predicted by the ChloroP program (*http://www.cbs.dtu.dk*) (Emanuelsson *et al.*, 1999). Molecular weights and pI values were calculated by the ProtParam tool available at the Expert Protein Analysis System (ExPASy) server (*http://www.expasy.org*) (Gasteiger *et al.*, 2003). Alignment was carried out using the ClustalX software with default parameters (*ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX)* (Thompson *et al.*, 1997) or with ClustalW program (Thompson *et al.*, 1994) and visualized with the GeneDoc program (*http://www.psc.edu/biomed/genedoc*) (Nicholas *et al.*, 1997).

Phylogenetic tree was constructed by the neighbor-joining method with 10000 bootstrap replicates using the Molecular Evolutionary Genetics Analysis (MEGA) program version 3.1 (*http://www.megasoftware.net*) (Kumar *et al.*, 2004). The deduced amino acid sequences used for alignment and phylogenetic tree construction were collected from GenBank through the NCBI server, from gene indices available at the Institute for Genomic Research (TIGR) (*http://www.tigr.org/tdb*) (Quackenbush *et al.*, 2000), from the *Populus trichocarpa* database (*http://www.populus.db.umu.se*) (Sterky *et al.*, 2004), from the *Physcomitrella patens* database (*http://www.brc.riken.go.jp-/lab/epd/blast/index .shtml*) (Nishiyama *et al.*, 2003), from the *Cyanidioschyzon merolae* Genome Project (*http://merolae.biol.s.u-tokyo.ac.jp/*) (Matsuzaki *et al.*, 2004), and from the Joint Genome Institute (JGI; *http://www.jgi.doe.gov/*).

Secondary and three-dimensional (3D) structures were predicted by using the CPHmodels 2.0 (*http://www-.cbs.dtu.dk/services/CPHmodels/*) (Lund *et al.*, 2002), or the SWISS-MODEL programs (*http://swissmodel.expasy.org/*) (Schwede *et al.*, 2003). The 3D models were visualized with the Swiss-Pdb Viewer program (*http://au.expasy.org/spdbv/*) (Guex and Peitsch, 1997).

14. Southern blot analysis

The number of genes encoding HbDXRs in the *H. brasiliensis* genome was estimated by Southern blot analyses using the genomic DNA isolated from young leaves of mature rubber trees and further detected with a homologous *HbDXR* probe. Details for the isolation of genomic DNA, cutting with restriction enzymes, transferring onto the membrane and probe preparation as well as the hybridization and detection protocols are described below.

14.1 Isolation of genomic DNA

Genomic DNA was extracted from young leaves of *H. brasiliensis* by CTAB method according to Ausubel *et al.* (1995). Two grams of young leaves were ground in liquid nitrogen with a pre-cooled mortar and pestle. The powder was transferred to a 50 ml conical tube containing 10 ml of CTAB buffer [20 g CTAB/l, 1.4 M NaCl, 0.1 M Tris-HCl (pH 8.0), and 20 mM EDTA], mixed thoroughly and incubated at 65ºC for 30 min. After centrifugation at 7000 x *g* at 4ºC for 20 min, the supernatant was transferred into a new tube and extracted twice with equal volumes of phenol: chloroform: isoamyl alcohol (25: 24: 1, v/v/v) followed by extraction twice with equal volumes of chloroform: isoamyl alcohol (24: 1, v/v). Genomic DNA was precipitated by adding two volumes of CTAB precipitation buffer (5 g CTAB/l and 0.04 M NaCl) to the supernatant, incubated at room temperature for 1 h and then centrifuged at 7000 x *g* at 4ºC for 20 min. The resulting DNA pellet was dissolved in 1.2 M NaCl, mixed with 0.7 volume of isopropanol and incubated at -20°C for 1 h followed by centrifugation at 7000 x *g* at 4ºC for 20 min. The DNA pellet was washed with ice-cold 70% ethanol, dried, and resuspended in TE buffer [10 mM Tris-HCl (pH 8.0) and 1 mM EDTA]. The concentration of DNA was determined (section 4.2) and stored at -20ºC for future use.

14.2 Digestion of genomic DNA and transferring to membrane

Aliquots of DNA (20 µg) were separately digested with the restriction enzymes *Bam*HI, *Bgl*II, *Eco*RI, *Eco*RV, *Sac*I, *Xba*I, and *Xho*I (New England Biolabs) at 37ºC overnight and separated on 1.0% agarose gel electrophoresis (section 4.1) for 16 h. After electrophoresis, the gel was stained with 0.1 mg/ml of ethidium bromide and destained in sterilized water for 30 min each. Then, the gel was soaked in 0.2 M HCl (250 ml) with gentle agitation for 10 min to facilitate the transferring efficiency of large DNA fragments. The double-strand DNA was denatured in 250 ml of denaturing buffer (1.5 M NaCl and 0.5 N NaOH) followed by neutralization in 250 ml of neutralization buffer [1.5 M NaCl and 1 M Tris-HCl (pH 7.5)] for 30 min each with gentle agitation. The DNA

fragments were transferred onto Hybond-N^+ membrane (GE Healthcare) by capillary blotting (Sambrook *et al.*, 1989).

The gel was placed on the platform containing the Whatman 3 MM paper (Whatman International Ltd., UK) wet in 20X SSC buffer [3 M NaCl and 0.3 M CH_3COONa (pH 7.0)] and surrounded by the plastic wrap. The membrane was moistened in 20X SSC buffer briefly and then placed on top of the gel followed by two pieces of the wet Whatman 3 MM paper on top of the membrane. Any air bubbles were removed by rolling a glass rod back and forth over the surface several times in each step. Then, a 15-20 cm stack of dry paper towels, glass plate and weight were placed on top of the Whatman 3 MM paper, respectively, and left for 18-24 h. After the transferring process was completed, DNA fragments were fixed onto the membrane by baking the membrane at 80ºC for 2 h. The membrane was used for hybridization immediately or stored at room temperature in a plastic wrap until further analysis.

14.3 Preparation of the *DXR* probe

Recombinant plasmid DNA containing the partial *HbDXR1* gene was digested with restriction enzyme *Eco*RI (section 9.2) at 37ºC for 5 h and purified from agarose gel as described in section 6.2. The purified *HbDXR1* fragment (about 0.7 kb) were denatured by boiling in a heat box for 10 min and rapidly chilled on ice for 5 min. The *HbDXR1* probe was then labeled with a fluorescein-11dUTP using the Gene Images Random Prime Labeling Module (GE Healthcare) according to the manufacturer's instructions using denatured *HbDXR1* fragment as template. The labeling reaction contains 2 µg of DNA template, 10 µl of nucleotide mix, 5 µl of random primer, 1 µl of Klenow fragment (5 units/µl), and sterile water to give a total volume of 50 µl. The mixture was incubated at 37ºC for 2 h.

The efficiency of the labeling reaction was analyzed as follows. The nucleotide mix was diluted in TE buffer to a final concentration of 1/5, 1/10, 1/25, 1/50, and 1/100. Five microliters of each diluent were subsequently dotted onto the Hybond-N^+ membrane and 5 µl of fluorescein-labeled *DXR* probe was also dotted onto another membrane. Both membranes were washed in 2X SSC buffer [0.3 M NaCl and 0.03 M CH_3COONa (pH 7.0)] at 60ºC for 15 min, placed on the Whatman 3 MM paper moistened with buffer TE and visualized under the UV light to estimate the concentration of the probe.

14.4 Hybridization and detection

The membrane was pre-hybridized in 20 ml of pre-hybridization buffer [5X SSC, 0.1% (w/v) SDS, 5% (w/v) dextran sulphate (Mw 500000, Sigma) and 20-fold dilution of liquid block] at 60ºC in hybridization incubator (model Shake& Stack, Hybaid, UK) for 1 h according to the manufacturer's instructions (GE Healthcare). The membrane was then allowed to hybridize with the fluorescein-labeled *HbDXR1* probe (3-5 ng/ml) at 60ºC overnight. After hybridization, the membrane was washed with 1X SSC and 0.1% (w/v) SDS once followed by washing with 0.1X SSC and 0.1% (w/v) SDS once at 60ºC for 15 min in each step.

The hybridization signals were detected by chemiluminescent based method using the Gene Images CDP-Star Detection Module (GE Healthcare) according to the manufacturer's instructions. The membrane was incubated in blocking reagent containing 5X of liquid block in buffer A [300 mM NaCl and 100 mM Tris-HCl (pH 9.0)] at room temperature for 1 h with gentle agitation followed by incubation with the anti-fluorescein alkaline phosphatase-conjugate for 1 h at the same conditions. Then, the membrane was washed with buffer A containing 0.3% of Tween-20 at room temperature three times (10 min in each step) to remove the unbound conjugate. Subsequently, the membrane was immersed in 5 ml of the CDP-Star detection reagent (disodium 2-chloro-5-(4-methoxyspiro {1,2-dioxetane-3,2'-(5'-chloro) tricyclo-[3.3.1.1] decan}-4-yl)-1-phenyl phosphate) and incubated further at room temperature for 5 min. After removing the excess detection reagent, the membrane was wrapped in plastic bag and exposed to X-ray film (Hyperfilm ECL, GE Healthcare) for 1 h. The film was developed by using an automatic X-ray film processor (Kodak, USA).

15. Semi-quantitative analysis

The expression pattern of the *HbDXSs* and *HbDXRs* genes in various tissues of seedlings, mature plants, latex of high-yielding (RRIM 600) and low-yielding (wild type) rubber clones under natural growth conditions were examined by semi-quantitative RT-PCR. The effect of ethephon on the expression of both *HbDXSs* and *HbDXRs* genes in rubber latex of wild type rubber clone were also analyzed. Protocols of first-strand cDNA synthesis and amplification of the target genes in semi-quantitative RT-PCR are described below.

15.1 First-strand cDNA synthesis

First-strand cDNA was synthezied using the ImProm-II Reverse Transcription System (Promega) according to the manufacturer's instructions. Aliquots of total RNA (1 µg) from various tissues were mixed with 1 µl of random primer (0.5 µg/µl) and DEPC-treated water to a total volume of 5 µl. The reaction was incubated at 70ºC for 5 min and chilled on ice for 2 min to remove the secondary structure of RNA. The mixture was mixed with 4 µl of ImProm-II 5X reaction buffer [375 mM KCl, 50 mM DTT, 25 mM $MgCl_2$, and 250 mM Tris-HCl (pH 8.3)], 1.2 µl of 25 mM $MgCl_2$, 1 µl of dNTP (10 mM of each dNTP), 0.5 µl of recombinant RNasin ribonuclease inhibitor (40 units/µl), 1 µl of ImProm-II RT, and DEPC-treated water to give a total volume of 20 µl. After incubation at room temperature for 5 min, the mixture was incubated further at 42ºC for 1 h followed by heating at 70ºC for 15 min to inactivate the enzyme ImProm-II RT. The first-strand cDNA was used as a template for subsequent PCR reactions.

15.2 Amplification of target cDNAs

Aliquots of first-strand cDNA (1 µl) synthesized in section 15.1 were used as a template in a subsequent PCR reaction to amplify genes of interest. PCR was performed in a 50 µl reaction system consisting of 2.5 units of *Taq* DNA polymerase (QIAGEN), 1X QIAGEN PCR buffer containing 1.5 mM $MgCl_2$, 200 µM of each dNTP, 10 µM of each gene-specific primer, and 10 or 2.5 µM of each 18S ribosomal RNA primers. PCR was carried out under the following conditions: 94ºC for 3 min, 34 cycles of 94ºC for 30 s, 58ºC for 30 s, and 72ºC for 30 s, and a final extension at 72ºC for 5 min. The PCR primers used were: HbDXS1f/r, HbDXS2f/r, HbDXR1f/r and HbDXR2f/r for the amplification of *HbDXS1*, *HbDXS2*, *HbDXR1* and *HbDXR2*, respectively. The expected cDNA sizes were 468, 416, 252 and 245 bp, respectively. The expression of both *HbDXSs* and *HbDXRs* transcripts were normalized to 18S ribosomal RNA with the primers 18S-0.2f/r primers (245 bp) or 18S-0.5f/r primers (531 bp). Equal volumes of each PCR product were loaded on 1.0% (w/v) agarose gels for electrophoresis (section 6.1), and the intensities of PCR products were quantified using the VisiDoc-It System (UVP Inc.). PCR reactions for each primer pairs were carried out at least twice. The nucleotide sequences of primers and their Tm (ºC) are shown in Table 2.5. These primers were synthesized by QIAGEN or Integrated DNA Technology, Inc.

Table 2.5 Nucleotide sequences of gene-specific primers for *HbDXS1*, *HbDXS2*, *HbDXR1*, *HbDXR2*, and *18S rRNA* cDNAs.

Primer names	Nucleotide sequence (5´→3´)	Tm (ºC)
HbDXS1f	TACTTGCTTGCTCCCGTT	58.4
HbDXS1r	ATGACCAACATCCCACAG	58.4
HbDXS2f	GTCTAGCTCTTCGTTTGTATC	59.4
HbDXS2r	GGTATGCCTGATGACCAACA	60.4
HbDXR1f	GAGTGCTTAGTGCTGCTA	58.4
HbDXR1r	AGAGACTACTTTGGCATGAG	58.3
HbDXR2f	TGCTTAGTGCTGCGAATG	58.4
HbDXR2r	AGAATCTCTGTAGTTCCAGG	58.3
18S-0.2f	GGCGGATGTTGATTATAGGA	58.3
18S-0.2r	AACTAAGAACGGCCATGCAC	60.4
18S-0.5f	CAAAGCAAGCCTACGCTCTG	62.4
18S-0.5r	CGCTCCACCAACTAAGAACG	62.4

16. Expression of *HbDXS1* and *HbDXR1* genes in *E. coli*

16.1 Generation of *HbDXS1* and *HbDXR1* expression constructs

cDNA fragments containing the putative mature sequence of *HbDXS1* (approximately 1.9 kb) and *HbDXR1* (approximately 1.2 kb) were amplified by PCR using gene-specific primers: HbDXS1f/r for *HbDXS1* and HbDXR1f/r for *HbDXR1* (Table 2.6) and vectors harboring the full-length *HbDXS1* or *HbDXR1* cDNAs as template. The primers HbDXS1f/r contain *Eco*RI and *Xho*I restriction sites, respectively and those for HbDXR1f/r contain *Sph*I and *Pst*I restriction sites, respectively. PCR was done in a 50 µl reaction containing 0.25 µl of *Pfu* DNA polymerase (5 units/µl) (Promega, USA), 5 µl of 10X PCR buffer, 1 µl of dNTP mixture (10 mM each), 100 ng/µl of vectors, 5 µl of each gene-specific primer (10 pmole/µl), and sterile water to a total volume of 50 µl. Cycling parameters were 94ºC for 3 min, 30 cycles of 94ºC for 30 s, 65ºC for 30 s and

72ºC for 2 min (*HbDXS1*) or 1 min (*HbDXR1*), and 72ºC for 5 min in a final extension step.

Table 2.6 Nucleotide sequences of gene-specific primers for amplification of putative mature *HbDXS1* and *HbDXR1* cDNAs.

Primer names	Nucleotide sequence (5´→3´)	Enzymes	Tm (ºC)
HbDXS1f	CCGGAATTCGCATCGCTGTCAGAG	*Eco*RI	68.8
HbDXS1r	TCCGCTCGAGCTATGATGACATTATCTG	*Xho*I	66.9
HbDXR1f	GAGGAAAGCATGCTGTTCGGCCCAGCCT	*Sph*I	72.7
HbDXR1r	CAGAATTCTGCAGTCATGCAAGAACAGG	*Pst*I	66.9

Note: Recognition sites for the restriction enzymes used are underlined.

After amplification, the resulting PCR products were resolved on 1.5% agarose gel electrophoresis and visualized under the UV light (section 4.1). The expected cDNA fragment was purified from an agarose electrophoresis gel (section 6.2) and digested with appropriate restriction enzymes at 37ºC for 12 h. The digested cDNA fragment was further purified by QIAquick PCR Purification Kit (section 6.1) and ligated into expression vector pET28a (for *HbDXS1*) or pQE30 (for *HbDXR1*), which were previously digested with the same restriction enzymes. Ligation reaction was performed by using 0.5 µg of digested cDNAs, 0.5 µg of digested vectors, 1 unit of T4 DNA ligase (New England Biolabs), 1X ligation buffer [50 mM Tris-HCl (pH 7.6), 1 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, and 5% (v/v) polyethylene glycol-8000], and sterile water to a total volume of 20 µl. The reaction was incubated at 4ºC overnight and then the gene (*HbDXS1 or HbDXR1*) was transformed into *E. coli* competent cells (section 8.2). *E. coli* strain BL21 (DE3) (Novagen) competent cells were prepared by the $CaCl_2$ method (section 8.1) and used as host for the expression of HbDXS1 recombinant protein. For expression of HbDXR1 recombinant protein, the *E. coli* competent cells strain XL1-Blue MRF´ was used. Then, recombinant plasmid DNAs were isolated from the positive

clones (section 9.1) and sequenced (section 10) to confirm whether the orientation of the insertion of either *HbDXS1* or *HbDXR1* cDNAs into the expression vector was corrected.

16.2 Culturing *E. coli* cells harboring the recombinant plasmids

The *E. coli* cells harboring the appropriated recombinant plasmids were cultured in 10 ml LB medium containing appropriate antibiotics (50 µg/ml of kanamycin for *HbDXS1* and 100 µg/ml of ampicillin for *HbDXR1*) at 37ºC with vigorous shaking to an OD_{600} of 0.6. IPTG was then added to the bacterial culture to a final concentration of 0.5 mM to induce the expression of the recombinant proteins. After 4 h of induction, the cells were harvested by centrifugation at 11000 x *g* for 20 min at 4ºC, and stored at -20ºC until used.

16.3 Monitoring the recombinant proteins

16.3.1 Preparation of the *E. coli* lysates

The *E. coli* cells were washed with 1/10 volume of wash buffer [50 mM NaH_2PO_4 (pH 8.0) and 300 mM NaCl], and resuspended in 250 µl of lysis buffer [50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole, pH 8.0] in the presence of lysozyme (1 mg/ml). Samples were incubated on ice for 30 min and then centrifuged at 11000 x *g* for 10 min at 4ºC. The supernatant was collected and mixed with equal volume of 2X SDS-PAGE sample buffers [0.5M Tris-HCl (pH6.8), 4.4% (w/v) SDS, 2% (v/v) β-mercaptoethanol, 1.5% (w/v) bromophenol blue, and 20% (v/v) glycerol]. The membrane fraction was resuspended in 300 µl of 2X of SDS-PAGE sample buffer by vortexing. All samples were boiled for 10 min followed by centrifugation at 11000 x *g* for 10 min and then analyzed by denaturing polyacrylamide gel electrophoresis (SDS-PAGE).

16.3.2 SDS-PAGE and Coomassie Brilliant Blue R-250 staining

SDS-PAGE was performed according the method of Leammli (1970). The gels were prepared with the composition shown in Table 2.7. Electrophoresis was carried out with a constant current of 24 mA for 90 min in an electrode buffer [1 M Tris-HCl (pH 8.8), 50 mM EDTA, and 0.1% (w/v) SDS]. After electrophoresis, the SDS-gel was immersed in a Coomassie Blue solution [0.08% (w/v) Coomassie Brilliant Blue R-250 dissolved in 50% (v/v) methanol and 7.5% (v/v) acetic acid] for 2 h and fixed in a fixing solution [50% (v/v) methanol and 7.5% (v/v) acetic acid] for 30 min with gentle agitation. Background staining of the SDS-gel was removed by soaking the gel in a destaining solution [5% (v/v) methanol and 7.5% (v/v) acetic acid] for 30 min. All steps were carried out at room temperature.

17. Purification of recombinant proteins

17.1 Preparation of inclusion bodies

The cell pellets (section 16.2) were resuspended in lysis buffer containing 1 mg/ml of lysozyme to a concentration of 4 mg of cell paste per ml, incubated on ice for 30 min, and then lysed by sonication for 1 min. The inclusion bodies were separated by centrifugation at 11000 x *g* and washed with lysis buffer containing 2% Triton X-100 to remove the cellular debris.

Table 2.7 Composition of SDS-polyacrylamide gel.

Solutions	Stacking gel (3%)	Separating gel (12%)
30% Acrylamide (ml)	0.30	2.40
1.5 M Tris-HCl, pH 8.8 (ml)	-	1.50
0.5 M Tris-HCl, pH 6.8 (ml)	0.75	-
10% SDS (µl)	30	60
10% Ammonium persulfate (µl)	30	60
H_2O (ml)	1.86	1.97
TEMED (µl)	3	3
Total volume (ml)	3.00	6.00

17.2 Purification of recombinant proteins

The inclusion bodies were resuspended in a denaturating buffer [100 mM NaH_2PO_4, 10 mM Tris-HCl and 8 M urea, pH 8.0] and non-solubilized materials were removed by centrifugation (11000 x *g* for 60 min at 4ºC). The recombinant proteins were purified under denaturing conditions on Ni-NTA (QIAGEN) column according to the manufacturer's instructions. In brief, a solubilized-inclusion body was mixed with 50% Ni-NTA slurry in a ratio of 4:1 at 4ºC for 60 min. Then, the mixture was applied onto polypropylene column (GE Healthcare) and the flow-through was discarded. The column was then washed twice with buffer C [100 mM NaH_2PO_4, 10 mM Tris-HCl and 8 M urea, pH 6.3] to remove other proteins and the recombinant proteins were eluted with buffer D [100 mM NaH_2PO_4, 10 mM Tris-HCl and 8 M urea, pH 5.9] four times and buffer E [100 mM NaH_2PO_4, 10 mM Tris-HCl and 8 M urea, pH 4.5] four times. The purified recombinant proteins were stored at -20ºC until used.

17.3 Assay of the expressed HbDXR activity

The reaction mixture for the assay of HbDXR activity contained 100 mM Tris-HCl (pH 8.0), 1 mM $MnCl_2$, 0.3 mM DXP, 0.15 mM NADPH and protein in a total

volume of 1 ml. The reaction was initiated by adding the HbDXR enzyme, which was previously refolded in a refolding buffer [100 mM Tris-HCl (pH 8.0), 0.4 M L-Arginine, 0.2 mM PMSF, 0.5 mM GSSG (oxidized glutathione), and 5 mM GSSH (reduced glutathione)], to the assay mixture. Refolding of HbDXR was performed by 1:2 serial dilutions with the refolding buffer until proteins start to precipitate. Other refolding methods, such as coexpression with chaperone plasmids, and optimization of growth media by varying concentration of IPTG and varying temperature of the cell culture were also attempted. The HbDXR activity was measured by monitoring the oxidation of NADPH at 340 nm using a Beckman DU-640 spectrophotometer (Beckman Coulter, USA).

18. Preparation of polyclonal antibodies

The purified recombinant proteins (section 17.2) were concentrated by a Centricon device (YM-10, Millipore, USA) followed by analysis with SDS-PAGE (section 16.3.2). After destaining the SDS-gels (section 16.3.3), the gel fragments containing both HbDXS1 and HbDXR1 proteins (approximately 1 mg) were sent to the Covance Research Products, Inc. (USA) for the production of polyclonal antibodies using a white New Zealand rabbit (female) as host. The immunogen was emulsified in Freund's Complete Adjuvant (FCA) for the initial injection and a Freund's Incomplete Adjuvant (FIA) was used for all subsequent injections. The injection was done three times (3-week intervals each boost): first with 250 µg immunogen with FCA, second with 125 µg immunogen with FIA, and final with 125 µg immunogen with FIA. Test bleeds are taken approximately 10 days after the boosts. The rabbit serum from the third bleed (about 50 ml) was used as primary antibodies in the immunoblot analysis.

19. Immunoblot analysis

19.1 Preparation of crude protein extracts and transferring to membranes

Leaf punctures were ground in a mortar containing the SDS-PAGE sample buffer and then boiled for 5 min. Samples of about 10-20 µl were analyzed by SDS-PAGE (section 16.3.2). After electrophoresis, the gel was rinsed in a transfer buffer (0.192 M glycine, 25 mM Tris base, 20% methanol, and 0.005% SDS) for 5 min, and the gel foot and the stacking gel were cut off. Then, the gel was placed on a sheet of Whatman 3 MM paper pre-wetted in transfer buffer. A layer of nitrocellulose membrane, pre-wetted in a methanol, followed by the transfer buffer was applied to the gel. Air

bubbles between gel and membrane were removed by smoothing with a glass rod. The membrane was covered with another layer of pre-wetted Whatman 3 MM paper. The sandwich was placed in a Novex Semi-Dry Blotter (Invitrogen), filled with the transfer buffer, and run with a constant current of 90 mA for 1 h. All steps were carried out at room temperature.

19.2 Immunoreaction and detection

After the transferring process was completed, the membrane was removed from the transfer apparatus, stained with the Ponceau S for 1 min, and destained with water for several times. The membrane was blocked in 2% nonfat dried milk in Tris-buffered saline [TBS, 0.02 M Tris-HCl (pH 7.6) and 0.137 M NaCl] containing 0.1% Tween 20 (TBST) for 30 min with gentle agitation. The membrane was washed briefly for two times in TBS-T, followed by incubation with primary antibodies (diluted 1:500 in TBST) for 1 h with gentle agitation. After removing excess primary antibodies by soaking the membrane in TBST (repeated two times, 10 min each), the membrane was incubated further with a secondary antibody for 30 min. The anti-rabbit IgG-horseradish peroxidase conjugated diluted 1:5000 in TBST was used as the secondary antibody (GE Healthcare). The membrane was washed twice in TBST (10 min each) to remove the excess secondary antibody.

After washing, signals were detected by using an ECL Western blotting detection system (GE Healthcare). The membrane was covered with mixtures of the detection solution 1 and detection solution 2 (1:1) for 1 min and the excess detection reagents were drained off. The membrane was wrapped with the Saran-wrap and exposed to X-ray film (Hyperfilm ECL, GE Healthcare) for a few minutes. The film was developed by four steps as follows: (1) immersing the film in Kodak GBX Developer and Replenisher for 3 min, followed by (2) soaking the film in water for 30 s, (3) immersing the film in Kodak GBX Fixer and Replenisher for 1 min, and (4) soaking the film in water for 2 min and drying. All steps were carried out at room temperature with gentle agitation.

CHAPTER 3

RESULTS

In order to isolate the genes that are involved in the biosynthesis of isoprenoids by the MEP pathway in *H. brasiliensis*, a latex UniZAP-XR phage cDNA library that previously prepared by Suwanmanee *et al.* (2002) was screened using the cDNAs encoding the first two enzymes of the MEP pathway: DXS and DXR from *P. montana* (Sharkey *et al.*, 2005) as heterologous probes. By this method, neither cDNA clone homologous to *DXS* nor *DXR* genes were identified (data not shown). This may because that the transcripts of both *DXS* and *DXR* in *H. brasiliensis* or the concentration of the library used in this study were be low. The concentration of the original library was found to be 1.94 x 10^8 pfu/ml (Suwanmanee *et al.*, 2002), while the concentration of the library used in this study was accounted to be 3 x 10^5 pfu/ml.

In recent years, the information of DNA sequence of various genes in the MEP pathway of various plants has been increasingly published. This information made it possible to isolate the *DXS* and *DXR* genes from *H. brasilienis* using the sensitive method called RNA-based RT-PCR amplification. The method was performed using primers designed based on the conserved regions of both *DXS* and *DXR* genes from different plant species. The flanking sequences at both the 5´- and 3´-ends were further identified by RACE-based methods. Therefore, these methods offer a good chance for a rapid and specific amplification of unknown sequences as compared to the classical cDNA library screening approach. Using these methods, new cDNA clones encoding for DXS and DXR have been successfully isolated from *H. brasilienis*. The cDNA clones were named *HbDXS* and *HbDXR*, respectively. Each clone contains two isoforms. The *HbDXS1* and *HbDXR1* cDNAs were isolated from leaves, while the *HbDXS2* and *HbDXR2* cDNAs were isolated from the latex. Details of results for the isolation and characterization of the representative cDNA clones are described below.

1. Molecular cloning and sequence analysis of cDNA encoding HbDXS

1.1 A partial sequence of *HbDXSs*

The nucleotide and amino acid sequences of *DXS* from various plants such as *A. thaliana* (Mandel *et al.*, 1996), *C. annuum* (Bouvier *et al.*, 1998), and *L. esculentum* (Lois *et al.*, 2000) were obtained from GenBank database. They were aligned using the ClustalX software (Thompson *et al.*, 1997) to find the most highly conserved regions among these sequences. Two conserved amino acid sequences: WDVGHQ and IAEQHA were selected to synthesize degenerate primers. The primer WDV-DXS1 was used as forward primer and the primer IAE-DXS2 was used as reverse primer (Table 2.1). Using these two primers and total RNA isolated from leaves and latex of *H. brasiliensis* as templates in One-step RT-PCR, the cDNA fragments with the size of about 0.9 kb in length were amplified from both leaves (Fig. 3.1, lane 1) and latex (data not shown). These fragments were purified, cloned into pDrive or pGEM-T Easy, and sequenced.

Analysis of the nucleotide sequences using the BLAST program (Altschul *et al.*, 1997) indicated that both cDNAs shared high sequence homology to the corresponding region of DXS from various plants (>68% identity). The highest sequence homology was found to *MtDXS1* from the *M. truncatula* (92% identity) (Walter *et al.*, 2002). When the nucleotide sequences of these two clones were compared, each cDNA had 100% identical to each other, implying that both of them might be derived from the same gene, even though they were from distinct organs of *H. brasiliensis*. Therefore, both cDNAs were designated as *HbDXS1*.

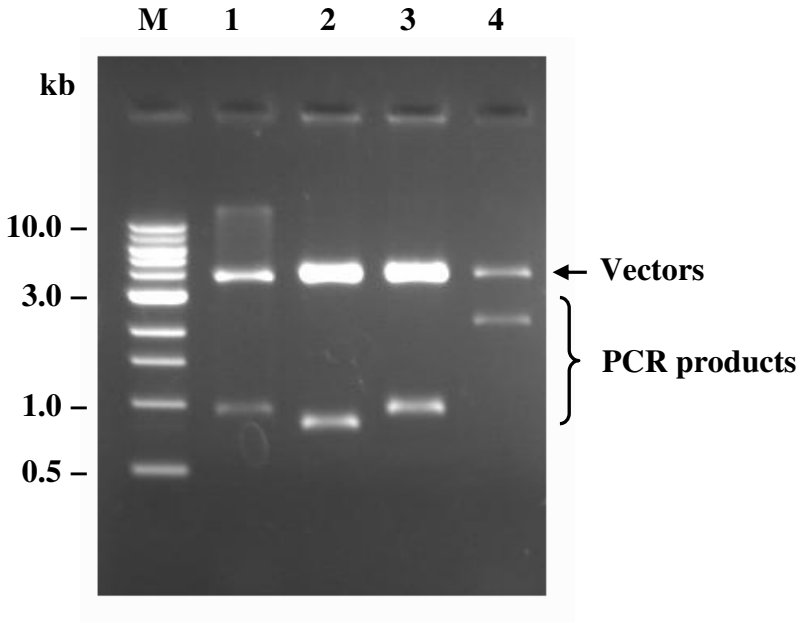


Fig. 3.1 Agarose gel analysis of the *HbDXS1* cDNA.

Plasmid DNAs containing the expected cDNAs of *HbDXS1* was digested with restriction enzyme *Eco*RI at 37ºC for 1 h and separated by electrophoresis on a 1.0% agarose gel containing 0.1 mg/ml of ethidium bromide. Lanes 1 and 4 are pDrive vectors containing 0.9 kb RT-PCR product from One-step RT-PCR and 2.1 kb ORF of *HbDXS1*, respectively. Lanes 2 and 3 are pCR4-TOPO vectors containing 0.8 kb and 1.1 kb RACE-PCR products from 5´- and 3´-RACE, respectively. Lane M is a 1 kb DNA Ladder (New England Biolabs).

1.2 A full-length *HbDXS1* cDNA

According to the internal nucleotide sequence originally obtained from the One-step RT-PCR, gene-specific primers were designed to amplify the missing sequences at the 5´- and 3´-ends by the RLM-RACE method. The first-strand cDNA was synthesized from total RNA isolated from leaves as template. The 5´-end of *HbDXS1* cDNA was amplified with primers GSP1-HbDXS1/GeneRacer 5´ in the first round of PCR and primers GSP2-HbDXS1/GeneRacer 5´ nested in the second round of PCR, respectively (Table 2.2). The 3´-end of *HbDXS1* cDNA was amplified with primers GSP1-HbDXS1/GeneRacer 3´ in the first round of PCR and the same gene-specific primer together with the GeneRacer 3´ nested primer in the second round of PCR, respectively (Table 2.3). The 5´- and 3´-RACE PCRs yielded PCR products of about 0.8 and 1.1 kb in length (Fig. 3.1, lane 2 and lane 3, respectively). These PCR products were purified from gel and cloned into pCR4-TOPO vector for sequencing. Analysis of the nucleotide sequences of both 5´- and 3´-RACE products revealed that these two cDNAs encompassed part of the 5´-untranslated region (5´-UTR) and the 3´-UTR, as well as a portion of coding regions.

After determination of the initiation start codon (ATG) and the stop codon (TAG), a 2193 bp fragment containing the coding region of *HbDXS1* (Fig. 3.1, lane 4) was finally obtained by PCR using the primers GSP1/2-HbDXS1 (Table 2.4). The PCR product was further purified from gel, cloned into pDrive vector, and sequenced. Analysis of the nucleotide sequence revealed that the cloned cDNA contained an ORF of 2163 bp. It was predicted to encode a polypeptide of 720 amino acids with a molecular mass and pI values of 78 kDa and 7.66, respectively. The complete nucleotide and the predicted amino acid sequences of *HbDXS1* have been deposited to GenBank database under the accession number AY502939 (Fig. 3.2).

```
1    atggctctctctgcatgttcatttcctgcccatgtaaacaggaacacaatctcagatctt
     M  A  L  S  A  C  S  F  P  A  H  V  N  R  N  T  I  S  D  L    20
61   caaaagtacagctatgtttcttctcatttcttatcgagaaaaaatccgttggcccaatct
     Q  K  Y  S  Y  V  S  S  H  F  L  S  R  K  N  P  L  A  Q  S    40
121  ctacacagacttaatcaggcaaagagcaagagaaggccagaacgggtttgtgcatcgctg
     L  H  R  L  N  Q  A  K  S  K  R  R  P  E  R  V  C  A  S  L    60
181  tcagagagagaagaatatcactcacagagaccaccaacacctctcttggacaccataaac
     S  E  R  E  E  Y  H  S  Q  R  P  P  T  P  L  L  D  T  I  N    80
241  tatccaattcatatgaaaaatctatcaatcaaggaactaaaacaactggcagatgagctg
     Y  P  I  H  M  K  N  L  S  I  K  E  L  K  Q  L  A  D  E  L   100
301  cggtcagatgttatcttcaatgtttctggaactgggggtcacttaggatcaagtcttggt
     R  S  D  V  I  F  N  V  S  G  T  G  G  H  L  G  S  S  L  G   120
361  gttgttgagctcactgtggctcttcactatgttttcaatgctcctcaagacaagattctg
     V  V  E  L  T  V  A  L  H  Y  V  F  N  A  P  Q  D  K  I  L   140
421  tgggatgttggtcatcagtcctacccacacaaaatcctgactgggagaagagacaagatg
     W  D  V  G  H  Q  S  Y  P  H  K  I  L  T  G  R  R  D  K  M   160
481  cacacaatcagacagacaaatggactttctggtttcacgaagcggtcggagagtgaatat
     H  T  I  R  Q  T  N  G  L  S  G  F  T  K  R  S  E  S  E  Y   180
541  gattgctttgggactggccatagctctaccactatttctgcaggcttggggatggcagtg
     D  C  F  G  T  G  H  S  S  T  T  I  S  A  G  L  G  M  A  V   200
601  gggagagatttgaaaggaagaaaaaacaacgtagttgctgttataggtgatggtgccatg
     G  R  D  L  K  G  R  K  N  N  V  V  A  V  I  G  D  G  A  M   220
661  acagcaggacaagcttatgaagctatgaacaatgcagggtacctggactctgatatgatt
     T  A  G  Q  A  Y  E  A  M  N  N  A  G  Y  L  D  S  D  M  I   240
721  gttattctcaatgacaacaaacaagtttctttaccaactgctactcttgatggacccatt
     V  I  L  N  D  N  K  Q  V  S  L  P  T  A  T  L  D  G  P  I   260
781  ccaccagtgggagctttgagcagtgctcttagtaggttgcaatcaaacaggcctctcagg
     P  P  V  G  A  L  S  S  A  L  S  R  L  Q  S  N  R  P  L  R   280
841  gaactaagagaggttgctaagggtgtcacaaagcggattggtggatccatgcatgaactg
     E  L  R  E  V  A  K  G  V  T  K  R  I  G  G  S  M  H  E  L   300
901  gcagcaaaggttgatgaatatgctcgtgggatgatcagtggttctggatcaaccctcttc
     A  A  K  V  D  E  Y  A  R  G  M  I  S  G  S  G  S  T  L  F   320
961  gaagagcttggattatattatattggtcctgttgatggccacaacatagatgatcttata
     E  E  L  G  L  Y  Y  I  G  P  V  D  G  H  N  I  D  D  L  I   340
1021 gctattctcaaagaggttaagggtactaaaacaactggtccagtcttgatacatgttgtc
     A  I  L  K  E  V  K  G  T  K  T  T  G  P  V  L  I  H  V  V   360
1081 actgagaaaggtcgggggtatccatatgctgagaaagctgcagacaagtaccatggggtt
     T  E  K  G  R  G  Y  P  Y  A  E  K  A  A  D  K  Y  H  G  V   380
1141 accaagtttgatcctgcaactggaaaacaattcaagggcagtgctattacacagtcttac
     T  K  F  D  P  A  T  G  K  Q  F  K  G  S  A  I  T  Q  S  Y   400
1201 actacatactttgcagaggctttgattgcagaagcagaagtggacaaggatattgttgca
     T  T  Y  F  A  E  A  L  I  A  E  A  E  V  D  K  D  I  V  A   420
1261 attcatgctgcaatgggaggtggaacaggcttaaatctcttccttcgccgtttcccaaca
     I  H  A  A  M  G  G  G  T  G  L  N  L  F  L  R  R  F  P  T   440
```

Fig. 3.2 Nucleotide and deduced amino acid sequences of the *HbDXS1* cDNA.

A translated amino acid sequence is shown in a single-letter code below the nucleotide sequence. Nucleotides are numbered on the left margin and the amino acid positions are indicated on the right side. A translation termination codon is marked with a dash (-).

```
1321 agatgctttgatgttggaatagcggaacagcatgcagttacatttgctgcaggattagcc
     R  C  F  D  V  G  I  A  E  Q  H  A  V  T  F  A  A  G  L  A   460
1381 tgtgaaggccttaaaccattttgtgcaatctactcatctttcatgcagagggcttatgac
     C  E  G  L  K  P  F  C  A  I  Y  S  S  F  M  Q  R  A  Y  D   480
1441 caggtagtccatgatgtggatttgcagaagctgccagtaagatttgcaatggacagagct
     Q  V  V  H  D  V  D  L  Q  K  L  P  V  R  F  A  M  D  R  A   500
1501 ggactggttggagcagatggtcccacacattgtggagcttttgatgtcacttttatggca
     G  L  V  G  A  D  G  P  T  H  C  G  A  F  D  V  T  F  M  A   520
1561 tgtctccctaacatggttgtgatggctccttctgatgaggcagaacttttttcacatggtt
     C  L  P  N  M  V  V  M  A  P  S  D  E  A  E  L  F  H  M  V   540
1621 gccaccgctgccgccatagatgatcgtcctagctgcttccgatatccaagaggtaacggt
     A  T  A  A  A  I  D  D  R  P  S  C  F  R  Y  P  R  G  N  G   560
1681 gttggtgttcagctgccaccaggaaacaaaggcattcctcttgaggttggaaaaggcagg
     V  G  V  Q  L  P  P  G  N  K  G  I  P  L  E  V  G  K  G  R   580
1741 atattgattgaaggggaaagagtggcactcttgggttatgggacagcagttcagagctgt
     I  L  I  E  G  E  R  V  A  L  L  G  Y  G  T  A  V  Q  S  C   600
1801 ttggctgctgcctctttagtggaaccccatggcttgcttataacagtagcagatgcgaga
     L  A  A  A  S  L  V  E  P  H  G  L  L  I  T  V  A  D  A  R   620
1861 ttctgtaaacctttggatcacaccctcattcgaagcctagcaaaaccacatgaagttttg
     F  C  K  P  L  D  H  T  L  I  R  S  L  A  K  P  H  E  V  L   640
1921 ataacggttgaagaaggatcaattgggggctttggatctcatgttgcacattttctggcc
     I  T  V  E  E  G  S  I  G  G  F  G  S  H  V  A  H  F  L  A   660
1981 cttgatggtcttcttgatggcaaactgaagtggcggccactcgttcttccagataggtat
     L  D  G  L  L  D  G  K  L  K  W  R  P  L  V  L  P  D  R  Y   680
2041 attgaccatggatccccgtctgtccagttgatagaggctggtctaacgccatctcacgtt
     I  D  H  G  S  P  S  V  Q  L  I  E  A  G  L  T  P  S  H  V   700
2101 gcagcaacagtactcaacatacttggaaataaaagagaagctctgcagataatgtcatca
     A  A  T  V  L  N  I  L  G  N  K  R  E  A  L  Q  I  M  S  S   720
2161 tag
     -
```

Fig. 3.2 (continued)

1.3 A full-length *HbDXS2* cDNA

Two distinct isoforms of cDNA encoding DXS have recently been identified from many plants, such as *M. truncatula* (Walter *et al.*, 2002), *E. guineensis* (Khemvong and Suvachittanont, 2005), and *G. biloba* (Gong *et al.*, 2006; Kim *et al.*, 2006). In *H. brasiliensis*, a partial cDNA of *DXS* has also been cloned from latex of the plant (Ko *et al.*, 2003). A pairwise analysis of its nucleotide sequence and the *HbDXS1* cDNA showed 60% identity, indicating that there are at least two representatives of *DXS* genes present in *H. brasiliensis* genome.

In order to search for more *DXS* genes in *H. brasiliensis*, the nucleotide and amino acid sequences of *DXS* from various plants were re-aligned to design a new set of degenerate primers. Alignment was performed using the ClustalX software as described above (Thompson *et al.*, 1997). Based on the new sequence alignment, two sets of degenerate primers were designed. The primers in set 1 were PIH-DXS3 (sense orientation) and YFA-DXS4 (antisense orientation). These primers correspond to the amino acid sequences PIHMKN and YFAEAL, respectively (Table 2.1). The primers in set 2 were MIS-DXS5 (sense orientation) and ITV-DXS6 (antisense orientation), corresponding to the sequences MISGSG and ITVEEG, respectively (Table 2.1). Using these two sets of degenerate primers and total RNA isolated from latex as template, the One-step RT-PCR yielded two PCR products of about 0.9 (using primer set 1, Fig. 3.3A, lane 1) and 1.1 kb in length (using primer set 2, Fig. 3.3A, lane 2), respectively. These two fragments were purified, cloned into pGEM-T Easy vector, and sequenced.

Comparison of these nucleotide sequences using the BLAST program (Altschul *et al.*, 1997) showed that these two cDNAs showed high homology to *DXS* from various plants, such as *S. rebaudiana* (88% identity for 0.8 kb fragment and 90% identity for 1.1 kb fragment). Since the two degenerate primer sets used in this study overlapped, analysis of these nucleotide sequences further showed that the overlapping region between these two sequences were identical (data not shown). Interestingly, these two cDNAs exhibited 70% identity to the *HbDXS1* cDNA, suggesting that both of them represent a new *DXS* gene from *H. brasiliensis*. So, they were named *HbDXS2*.

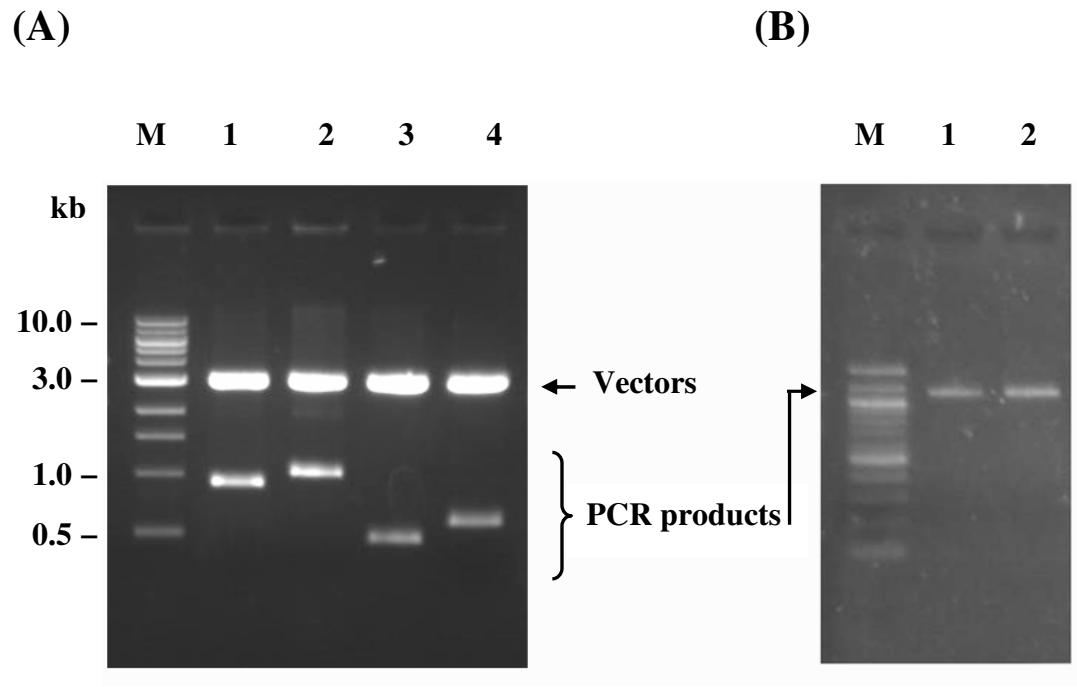


Fig. 3.3 Agarose gel analysis of the *HbDXS2* cDNA.

(A) Plasmid DNAs containing the expected cDNAs of *HbDXS2* were digested with restriction enzyme *Eco*RI at 37ºC for 1 h and separated by electrophoresis on a 1.0% agarose gel containing 0.1 mg/ml of ethidium bromide. Lanes 1 and 2 are pGEM-T Easy vectors containing 0.9 and 1.1 kb RT-PCR products from One-step RT-PCR. Lanes 3 and 4 are pGEM-T Easy vectors containing 0.4 and 0.5 kb RACE-PCR products from 5´-RACE and 3´-RACE, respectively. Lane M is a 1 kb DNA Ladder (New England Biolabs).

(B) PCR products of *HbDXS2* cDNA. Lanes 1 and 2 are 1.2 kb PCR products amplified with primers GSP1/2-HbDXS2 and primers GSP3/4-HbDXS2, respectively. Lane M is a 100 bp DNA Ladder (New England Biolabs).

The 5´- and 3´-ends of *HbDXS2* cDNA were amplified using the RLM-RACE method, as described above, but the first-strand cDNA used as template were prepared from total RNA isolated from latex. The 5´-end was amplified with primers GSP1/2-HbDXS2 together with the GeneRacer 5´ and GeneRacer 5´ nested primers in the first and second rounds of PCR, respectively (Table 2.2). The 3´-end was amplified by primers GSP1-HbDXS2/GeneRacer 3´ in the first PCR, and primers GSP2-HbDXS2/GeneRacer 3´ nested in the second round of PCR, respectively (Table 2.3). The 5´- and 3´-RACE PCRs gave PCR products of about 0.4 and 0.5 kb in length (Fig. 3.3A, lane 3 and lane 4, respectively). These PCR products were then purified, cloned into pGEM-T Easy vector, and sequenced. Analysis of nucleotide sequences revealed that these cDNAs spanned the coding region, and the 5´- and 3´-UTRs of the *HbDXS2* cDNA. Finally, a full-length cDNA of *HbDXS2* was obtained using two step-wise PCRs and two sets of gene-specific primers. The primers set 1 was GSP1/2-HbDXS2 and the primers set 2 was GSP3/4-HbDXS2 (chapter 2: Table 2.4). PCR amplified cDNA fragments of about 1.2 kb in length in both set of primers (Fig. 3.3B). These fragments were further purified, cloned into the pGEM-T Easy vector, and sequenced. Several positive clones were selected for sequencing.

Assembling of nucleotide sequences from these two fragments using the CAP3 program (Huang and Madan, 1999) yielded a completion of a single cDNA sequence of 2222 bp in length. The overlapping nucleotide sequences of about 115 bp between these two fragments showed 100% identical. The ORF of *HbDXS2* cDNA was 2136 bp in length. It was predicted to encode a polypeptide of 711 amino acids. The calculated molecular mass and a predicted pI value of the ORF were 76 kDa and 7.34, respectively. The *HbDXS2* cDNA shared 62% identity to the *HbDXS1*, considering only the coding region. The complete nucleotide and the predicted amino acid sequences of *HbDXS2* have been submitted to GenBank database under the accession number DQ473433 (Fig. 3.4).

```
1    atggcggtgtctagctcttcgtttgtatcaaatcaatctttctccccattcttgaaagca
     M  A  V  S  S  S  S  F  V  S  N  Q  S  F  S  P  F  L  K  A    20
61   ccaagatcaaacctctgtggcaggaaacagttctgtttacgagcatctgccggccaccca
     P  R  S  N  L  C  G  R  K  Q  F  C  L  R  A  S  A  G  H  P    40
121  gatgaagaagggaaaatgatgataaggaaagaaaaagatggctggaaaattgatttctca
     D  E  E  G  K  M  M  I  R  K  E  K  D  G  W  K  I  D  F  S    60
181  ggggagaaaccacccacaccattgctagatacaatcaattaccctgttcacacgaagaac
     G  E  K  P  P  T  P  L  L  D  T  I  N  Y  P  V  H  T  K  N    80
241  ctatccacacaggatcttgaacaactagcagcagagcttagagcagatattgtatacagt
     L  S  T  Q  D  L  E  Q  L  A  A  E  L  R  A  D  I  V  Y  S   100
301  gtatcgaagacaggagggcatctgagttcaagcttaggagttgtcgagctagcagtggct
     V  S  K  T  G  G  H  L  S  S  S  L  G  V  V  E  L  A  V  A   120
361  ctgcatcatgtgttcagcacacctgatgataaaatcatatgggatgttggtcatcaggca
     L  H  H  V  F  S  T  P  D  D  K  I  I  W  D  V  G  H  Q  A   140
421  tacccacataaaatttaacaggaagaaggtctaggatgcacaccataaggaaaacttca
     Y  P  H  K  I  L  T  G  R  R  S  R  M  H  T  I  R  K  T  S   160
481  gggcttgcaggatttcctaaaagagatgagagtgtttatgatgcctttggcgccggacat
     G  L  A  G  F  P  K  R  D  E  S  V  Y  D  A  F  G  A  G  H   180
541  agttccacaagcatctctgctggtcttggaatggcagttgcaagggacctactagggaag
     S  S  T  S  I  S  A  G  L  G  M  A  V  A  R  D  L  L  G  K   200
601  aacaacaatgtcatttctgtaattggagatggagccatgacagctggacaagcatatgag
     N  N  N  V  I  S  V  I  G  D  G  A  M  T  A  G  Q  A  Y  E   220
661  gccatgaacaatgcaggattccttgatgctaatctaattgttatattaaatgacaataag
     A  M  N  N  A  G  F  L  D  A  N  L  I  V  I  L  N  D  N  K   240
721  caagtatctttacccacagctactcttgatggccctgcaactcctgtaggagctctcagt
     Q  V  S  L  P  T  A  T  L  D  G  P  A  T  P  V  G  A  L  S   260
781  agtgctttggccaagatccaagcaagcacccagttccgcaaacttcgtgaagctgcaaaa
     S  A  L  A  K  I  Q  A  S  T  Q  F  R  K  L  R  E  A  A  K   280
841  agcatcacaaagcaaattggtggtaaaacgcatcaagttgcagcaaaggtagatgaatat
     S  I  T  K  Q  I  G  G  K  T  H  Q  V  A  A  K  V  D  E  Y   300
901  gcaagaggaatgattagtgcttctgggtctactctctttgaggagttggggttatattat
     A  R  G  M  I  S  A  S  G  S  T  L  F  E  E  L  G  L  Y  Y   320
961  atcggtccagtggatgggcacaacattgaagatctagtgaccatatttcaaaaagtaaaa
     I  G  P  V  D  G  H  N  I  E  D  L  V  T  I  F  Q  K  V  K   340
1021 gcaatgcctgccccaggaccagttctgattcacattgtgacagagaaagggaagggctat
     A  M  P  A  P  G  P  V  L  I  H  I  V  T  E  K  G  K  G  Y   360
1081 cccccagctgaggcagcagctgataaaatgcatggtgttgtcaagtttgatgttcaaact
     P  P  A  E  A  A  A  D  K  M  H  G  V  V  K  F  D  V  Q  T   380
1141 gggaagcaattcaagccaaaatcccctacacttttcatatacacagtactttgctgaagct
     G  K  Q  F  K  P  K  S  P  T  L  S  Y  T  Q  Y  F  A  E  A   400
1201 ctgatcaaagaagctgagacagacaataagattgtagccattcatgctgcaatgggtggt
     L  I  K  E  A  E  T  D  N  K  I  V  A  I  H  A  A  M  G  G   420
1261 gggactggtctcaattatttccagaagaggtttccagatcgctgctttgatgtgggcatt
     G  T  G  L  N  Y  F  Q  K  R  F  P  D  R  C  F  D  V  G  I   440
```

Fig. 3.4 Nucleotide and deduced amino acid sequences of the *HbDXS2* cDNA.

A translated amino acid sequence is shown in a single-letter code below the nucleotide sequence. Nucleotides are numbered on the left margin and the amino acid positions are indicated on the right side. A translation termination codon is marked with a dash (-).

```
1321 gctgagcaacatgctgttacatttgcagctggtttagccactgaaggactcaagcccttc
     A  E  Q  H  A  V  T  F  A  A  G  L  A  T  E  G  L  K  P  F   460
1381 tgtgcaatttactcatcattcctgcaacgaggatatgatcaggtggtacatgatgttgat
     C  A  I  Y  S  S  F  L  Q  R  G  Y  D  Q  V  V  H  D  V  D   480
1441 cttcaaaagttaccagtcagatttgccatggatcgagctggtttggttggtgcagatgga
     L  Q  K  L  P  V  R  F  A  M  D  R  A  G  L  V  G  A  D  G   500
1501 cccacccattgtggagcatttgatattgcatacatggcttgcttgcccaacatggtagtc
     P  T  H  C  G  A  F  D  I  A  Y  M  A  C  L  P  N  M  V  V   520
1561 atggctccatctgatgaagctgagctgatgcacatggttgccactgcagcagccatagat
     M  A  P  S  D  E  A  E  L  M  H  M  V  A  T  A  A  A  I  D   540
1621 gacagacccagctgcttcaggttcccaaggggcaacggaattggagcagctcttcctcct
     D  R  P  S  C  F  R  F  P  R  G  N  G  I  G  A  A  L  P  P   560
1681 aataataaaggaaccccacttgagattgggaaagggagaatactgatggaaggcaacaga
     N  N  K  G  T  P  L  E  I  G  K  G  R  I  L  M  E  G  N  R   580
1741 gttgccattttgggatatggttctatagttcaacaatgtgtagaagcagcaagtatgctc
     V  A  I  L  G  Y  G  S  I  V  Q  Q  C  V  E  A  A  S  M  L   600
1801 agaacccaaggcatttctgtgacagtagctgatgcgagattttgcaaacctttagataca
     R  T  Q  G  I  S  V  T  V  A  D  A  R  F  C  K  P  L  D  T   620
1861 gatcttatcaggcaattggccaaggagcatgagttcctcattacagtggaagagggttct
     D  L  I  R  Q  L  A  K  E  H  E  F  L  I  T  V  E  E  G  S   640
1921 attggaggtttttcttcccatgtatctcacttcctgagtttaagtggcattctggatgga
     I  G  G  F  S  S  H  V  S  H  F  L  S  L  S  G  I  L  D  G   660
1981 cccctaaagttgagagcaatggttctccctgacagatacattgaccatggatcgcctcaa
     P  L  K  L  R  A  M  V  L  P  D  R  Y  I  D  H  G  S  P  Q   680
2041 gaccaaattcaagaagcaggcatctcatcaaaccatatcactgccacagttttatctctc
     D  Q  I  Q  E  A  G  I  S  S  N  H  I  T  A  T  V  L  S  L   700
2101 ttggggaagccaaaagaagctcttcaattcaagtga
     L  G  K  P  K  E  A  L  Q  F  K  -
```

Fig. 3.4 (continued)

1.4 Comparison of the deduced *HbDXSs* amino acid sequences

A comparison of the deduced amino acid sequences of HbDXS1 and HbDXS2 showed that they have 69% identity to each other with a variation of 218 amino acids between the two sequences (Fig. 3.5). Database searching using the BLASTP program (Altschul *et al.*, 1997) revealed that both HbDXSs proteins showed a high sequence homology with the DXS sequences from various plants. The HbDXS1 showed highest sequence identity with DXS from *P. montana* (87%), followed by *A. annua* (86%), *A. paniculata* (85%), *C. annuum* (85%), *E. guineensis* (85%), and *L. esculentum* (85%) (Table 3.1), while the HbDXS2 showed highest sequence identity with DXS from *C. roseus* (84%), followed by *M. truncatula2* (83%), *S. rebaudiana* (83%), *O. sativa2* (79%), *M. citrifolia* (78%), and *G. biloba2* (76%) (Table 3.1). In addition, moderate and lower homologies of both HbDXSs sequences were observed in bacteria, cyanobacteria, and malaria parasite, with the sequence identity ranging from 30 to 65% (data not shown).

Table 3.1 Percent identity of the deduced amino acid sequences of HbDXSs with some representative plant DXSs.

The DXS sequences from various plants that show the highest identity to HbDXS1 and HbDXS2 sequences were shown. Data were analyzed using the BLAST program (Altschul *et al.*, 1997).

Species	% identity	Species	% identity
H. brasiliensis1 (HbDXS1)	100	*H. brasiliensis2* (HbDXS2)	100
P. montana	87	*C. roseus*	84
A. annua	86	*M. truncatula2*	83
A. paniculata	85	*S. rebaudiana*	83
C. annuum	85	*O. sativa2*	79
E. guineensis	85	*M. citrifolia*	78
L. esculentum	85	*G. biloba2*	76

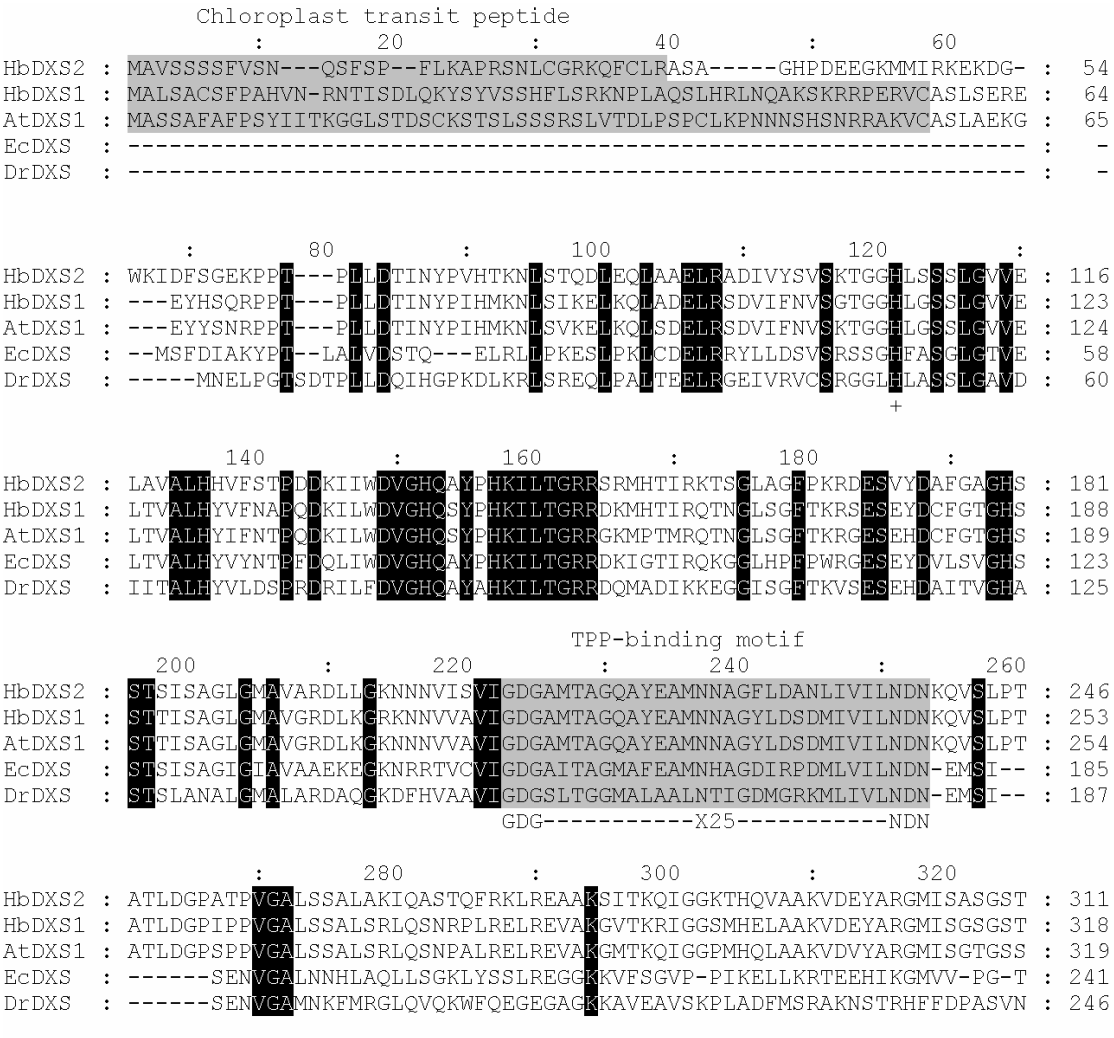


Fig. 3.5 Alignment of the deduced amino acid sequences of HbDXSs and some representative plant DXSs.

Multiple alignments were performed using the ClustalX software (Thompson *et al.*, 1997). Identical residues are shaded in black. Dashes (-) indicate gaps inserted to optimize the alignment. The putative cTP sequences, the TPP-binding motif, and a transketolase motif are indicated. Several residues important for activity of DXS are shown by plus symbols (+). The sequences used were *H. brasiliensis* (HbDXS1, AY502939 and HbDXS2, DQ473433), *A. thaliana* (AtDXS, U27099), *E. coli* (EcDXS, AF035440), and *D. radiodurans* (DrDXS, Q9RUB5).

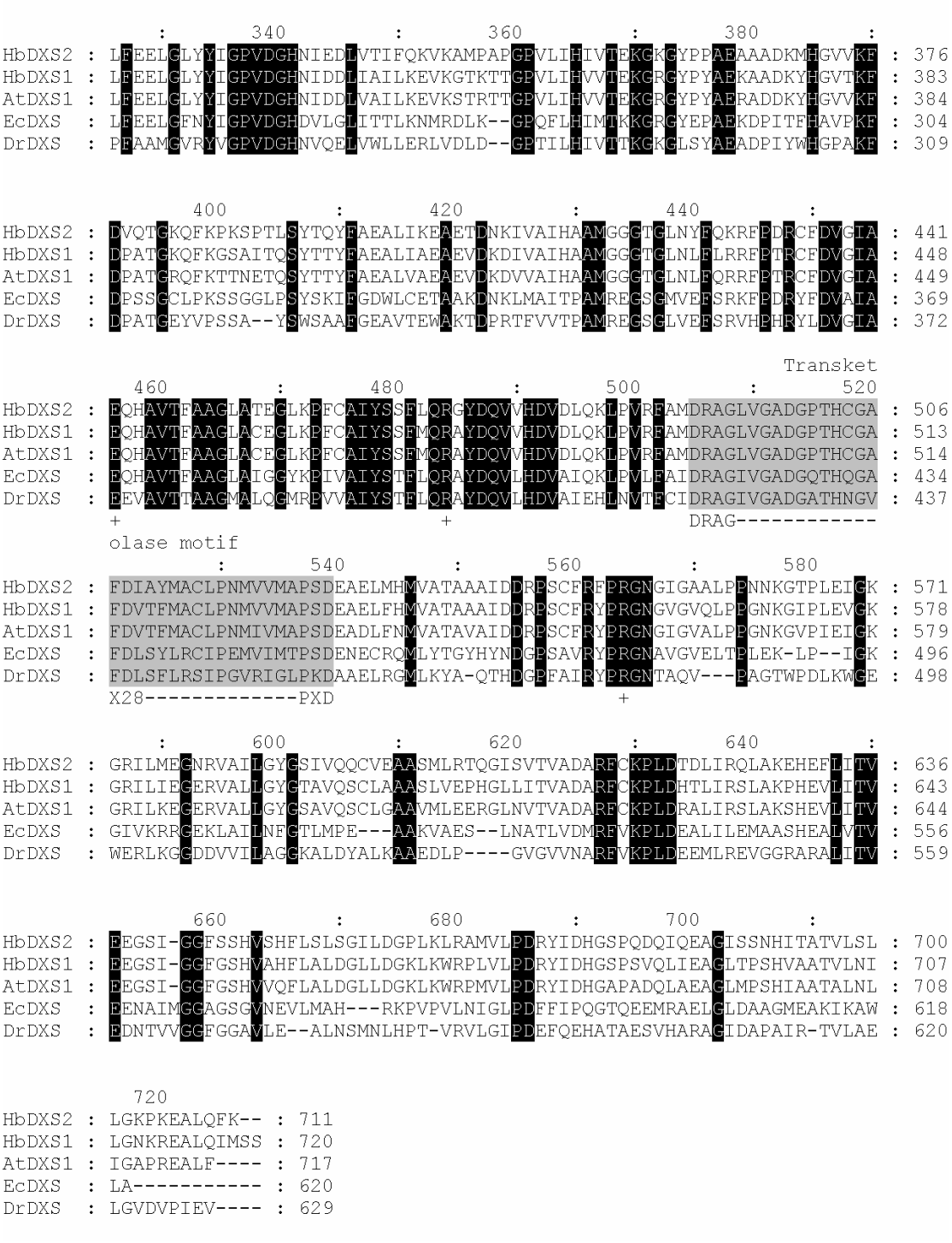


Fig. 3.5 (continued)

As shown in Fig. 3.5, multiple alignments of the deduced amino acid sequences of both HbDXSs with those from various higher plants, such as *A. thaliana*, and bacterial DXSs, such as *E. coli* and *D. radiodurans*, revealed two highly conserved regions and amino acids of some biological significance (Fig. 3.5). First, a highly conserved region with a high level of similarity to the consensus TPP-binding motif 'GDG-X_{25}-NDN' (Mandel *et al.*, 1996; Querol *et al.*, 2001), was found in both HbDXSs, corresponding to amino acids 216 to 246 in the HbDXS1 and amino acids 209 to 239 in the HbDXS2 sequences, respectively (Fig. 3.5). Second, a cluster of amino acids known to be important for structure and function of transketolases as a transketolase motif 'DRAG-X_{28}-P-X-D' (Querol *et al.*, 2001) was also present in both HbDXSs (Fig. 3.5). This motif was found at amino acids 498 to 532 in the HbDXS1 and amino acids 491 to 525 in the HbDXS2 sequences, respectively (Fig. 3.5). The letter X denotes any amino acid in both cases of TPP-binding motif and transketolase motif. In addition to these two conserved regions, several invariant residues, such as histidine (H49), glutamic acid (E370) and arginine (R398 and R478) residues in *E. coli* DXS, and glutamic acid residue (E449) in *C. annum* DXS, which previously shown to be important for the activity of DXS (Bouvier *et al.*, 1998; Querol *et al.*, 2001; Xiang *et al.*, 2007), were also conserved in both HbDXSs. These residues correspond to H114, E449, R477, and R557 in the HbDXS1 sequence, and amino acids H107, E442, R470, and R550 in the HbDXS2 sequence, respectively (Fig. 3.5).

As the MEP pathway is localized in plastid compartment of higher plants (Eisenreich *et al.*, 1998; Lichtenthaler, 1999; Eisenreich *et al.*, 2001; Rodríguez-Concepción and Boronat, 2002; Eisenreich *et al.*, 2004), both HbDXS sequences were found to contain extra amino acids at the NH_2-terminal region, suggesting that these extra amino acids could encode a cTP sequence in plants. These regions were not found in bacterial DXS sequences (Fig. 3.5), and they seemed to be poorly conserved in different plant species, but rich in hydroxylated amino acids, such as serine and threonine residues, and positively charge amino acids, such as arginine, lysine, and histidine residues. These are common features of the cTP sequence in plants (Gavel and von Heijne, 1990; von Heijne and Nishikawa, 1991; Emanuelsson *et al.*, 1999). In accordance with this notion, further analysis of the deduced amino acid sequences of both HbDXSs revealed the presence of ten hydroxylated amino acids (9 serines and 1 threonine) and thirteen positively charged amino acids (6 arginines, 4 lysines, and 3 histidines), accounting for 17.54% and 22.81%, respectively, of the amino acids in the cTP of the HbDXS1

sequence, while eight hydroxylated amino acids (8 serines, 23.52%) and five positively charged amino acids (3 arginines and 2 lysines, 14.71%), respectively, were found in the HbDXS2 sequence (Fig. 3.5 and Fig. 3.6).

Analysis of NH_2-terminal extension regions using the neural network ChloroP program (Emanuelsson *et al.*, 1999) showed that most plant DXSs analyzed, including both HbDXSs, were predicted to be localized in chloroplasts, with reasonable confidence values greater than 0.5 (Table 3.2). The value >0.5 is considered as an indicative of chloroplast localization (Emanuelsson *et al.*, 1999). As shown in Fig. 3.6, sixteen DXS sequences (about 67%) out of 24 sequences analyzed showed a nearly complete conservation surrounding the putative cTP cleavage sites as a consensus motif 'Xaa↓Ala-Ser-Xaa', where Xaa denotes any amino acid. Based on the ChloroP program and these observations, the putative cTP cleavage sites in both HbDXSs were predicted to be recognized between amino acids 57 and 58 in the HbDXS1 and 34 and 35 in the HbDXS2 sequences, respectively (Fig. 3.5 and Fig. 3.6). After cleavage the cTP sequences, the putative mature HbDXS1 and HbDXS2 sequences were predicted to contain 663 amino acids with a molecular weight of 71 kDa and a pI value of 6.4 for HbDXS1, and to contain 677 amino acids with a predicted molecular weight of 72 kDa and a theoretical pI value of 6.56 for HbDXS2.

Table 3.2 Prediction of the localization and length of cTP sequences in HbDXSs and in some representative plant DXSs.

The localization and length of cTP sequences in HbDXSs and in various plant DXSs were predicted by the ChloroP program (Emanuelsson *et al.*, 1999).

Species	Accession nos.	ChloroP		
		Localization	cTP score	cTP length
H. brasiliensis1 (HbDXS1)	AY502939	Chloroplast	0.541	57
H. brasiliensis2 (HbDXS2)	DQ473433	Chloroplast	0.527	34
A. paniculata	AY254390	-	0.443	31
A. thaliana	U27099	Chloroplast	0.585	58
A. annua	AF182286	Chloroplast	0.524	51
A. majus	AY770407	Chloroplast	0.561	47
C. annuum	Y15782	-	0.477	57
C. morifolium	BAE79547	-	0.429	56
C. roseus	AJ011840	Chloroplast	0.509	35
E. guineensis	AY583783	-	0.462	44
G. biloba1	AY505128	Chloroplast	0.507	52
G. biloba2	AY494185	Chloroplast	0.504	45
L. esculentum	AF143812	-	0.474	56
M. truncatula1	AJ430047	Chloroplast	0.518	54
M. truncatula2	AJ430048	Chloroplast	0.521	36
M. piperita	AF019383	Chloroplast	0.539	41
M. citrifolia	AF443590	Chloroplast	0.516	55
N. pseudonarcissus	AJ279019	Chloroplast	0.536	45
O. sativa1	AAT58851	-	0.498	51
O. sativa2	XM_476952	Chloroplast	0.524	30
P. montana	AY315652	Chloroplast	0.515	45
S. rebaudiana	AJ429232	Chloroplast	0.537	34
T. erecta	AF251020	Chloroplast	0.515	44
T. media	AY505129	Chloroplast	0.572	63
T. aestivum	BT009346	-	0.485	45

Note:

Number after the scientific name refers to DXS sequence 1 and DXS sequence 2, respectively.

Dashes (-) indicate no prediction of cTP sequences to be localized in the chloroplast.

1.5 Phylogenetic analysis of HbDXSs sequences

The evolutionary relationship among DXS proteins was analyzed using the phylogenetic program, MEGA (version 3.1) (Kumar *et al.*, 2004). Forty-two DXS sequences from different species including plants, green and red alga, cyanobacteria, and bacteria were aligned using the ClustalW software (Thomson *et al.*, 1994). The alignment was then used to construct a phylogenetic tree by the neighbor-joining method implemented in the MEGA program. A significant branching of the tree was verified by bootstrapping analysis with 10000 replicates.

Considering the tree as a whole, the branching pattern of the DXS trees is in agreement with the organismic evolution. The tree can be classified into three major groups: bacteria, cyanobacteria, and plantae with the red alga *Cyanidioschyzon merolae* as sister, when the DXS sequence from malaria parasite *P. falciparum* was used as the out group (Fig. 3.7 and Fig. 3.8). Bacterial DXSs analyzed included: *E. coli*, *Haemophilus influenzae*, *P. aeruginosa*, *Vibrio cholerae*, and *Yersinia mollaretii*, and cyanobacterial DXSs characterized included: *Nostoc* sp., *Prochloro-coccus marinus*, *Synechocystis* sp., *Synechococcus elongatus*, and *Thermosynecho-coccus elongatus* (Fig. 3.7). Plantae DXSs analyzed included streptophytes (or land plants) and chlorophytes such as *C. reinhardii*, *Ostreococcus tauri*, and *Volvox carteri*. Both DXSs are located in different braches of the tree (Fig. 3.8). Rhodophyte *C. merolae* is branched basally to the green plastids (Fig. 3.8).

Specifically to land plants, the DXS sequences can be separated into two subgroups: class I and class II with a high bootstrap value support of 90% (Fig. 3.8). The name class I and class II were assigned according to Walter *et al.* (2002). As Fig. 3.8 shows, the two HbDXS sequences were grouped with other plant DXSs, but they are located in different branches. The HbDXS1 was closed to the DXS sequences from plants belonging to the DXS class I, while the HbDXS2 was grouped with other plant DXSs belonging to the DXS class II. The plant DXS class I contains *A. thaliana1*, *A. thaliana2*, *A. annua*, *A. paniculata*, *C. annuum*, *E. guineensis*, *G. biloba1*, *L. esculentum*, *M. truncatula1*, *O. sativa1*, *P. montana*, and *T. aestivum*, while the plant DXS class II contains *A. majus*, *C. roseus*, *C. morifolium*, *G. biloba2*, *M. truncatula2*, *M. piperita*, *M. citrifolia*, *N. pseudonarcissus*, *O. sativa2*, *S. rebaudiana*, *T. erecta*, and *T. media* (Fig. 3.8).

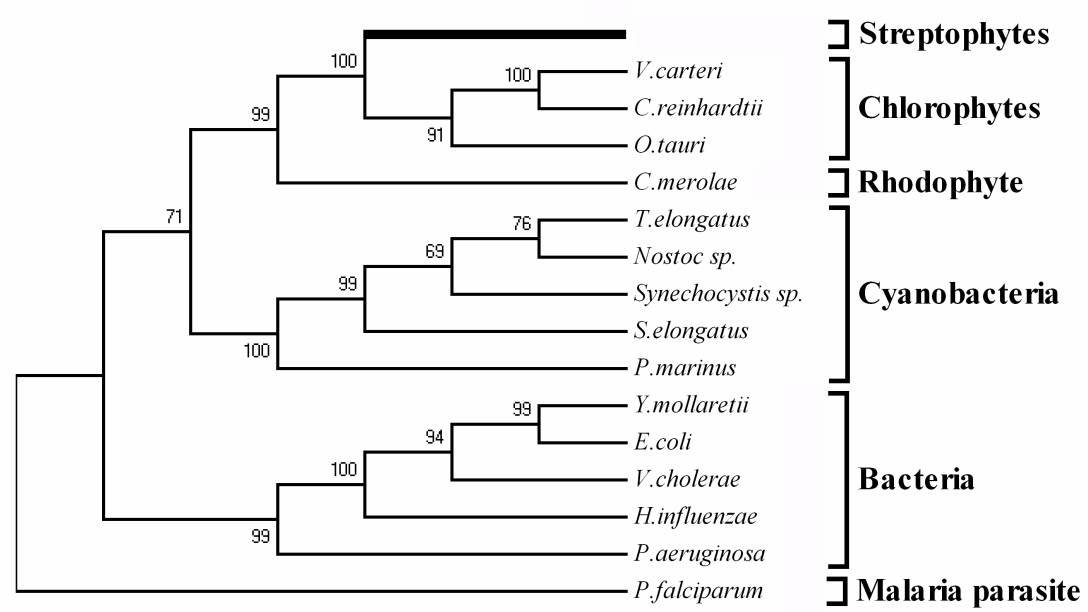


Fig. 3.7 Phylogenetic tree among DXS sequences from different taxa.

The tree was constructed by the MEGA program (version 3.1) (Kumar *et al.*, 2004) using the neighbor-joining method based on amino acid sequences alignment from the ClustalW (Thompson *et al.*, 1994). Numbers on branches represent percent bootstrap values support for 10000 replicates. The DXS sequence from *P. falciparum* (AF111814) was used as the out group. The DXS sequences from eubacteria: *E. coli* (AF035440), *H. influenzae* (ZP_00157279), *P. aeruginosa* (ZP_01296842), *V cholerae* (ZP_01485715), and *Y mollaretii* (ZP_00826763) and from cyanobacteria: *P. marinus* (ZP_01005433), *Nostoc* sp. (NP_484643), *Synechocystis* sp. (NP_440409), *S. elongatus* (YP_399449), and *T. elongatus* (NP_681412) were obtained from the GenBank database. The DXS sequences from algae were from the JGI database and *C. merolae* Genome Project (Matsuzaki *et al.*, 2004).

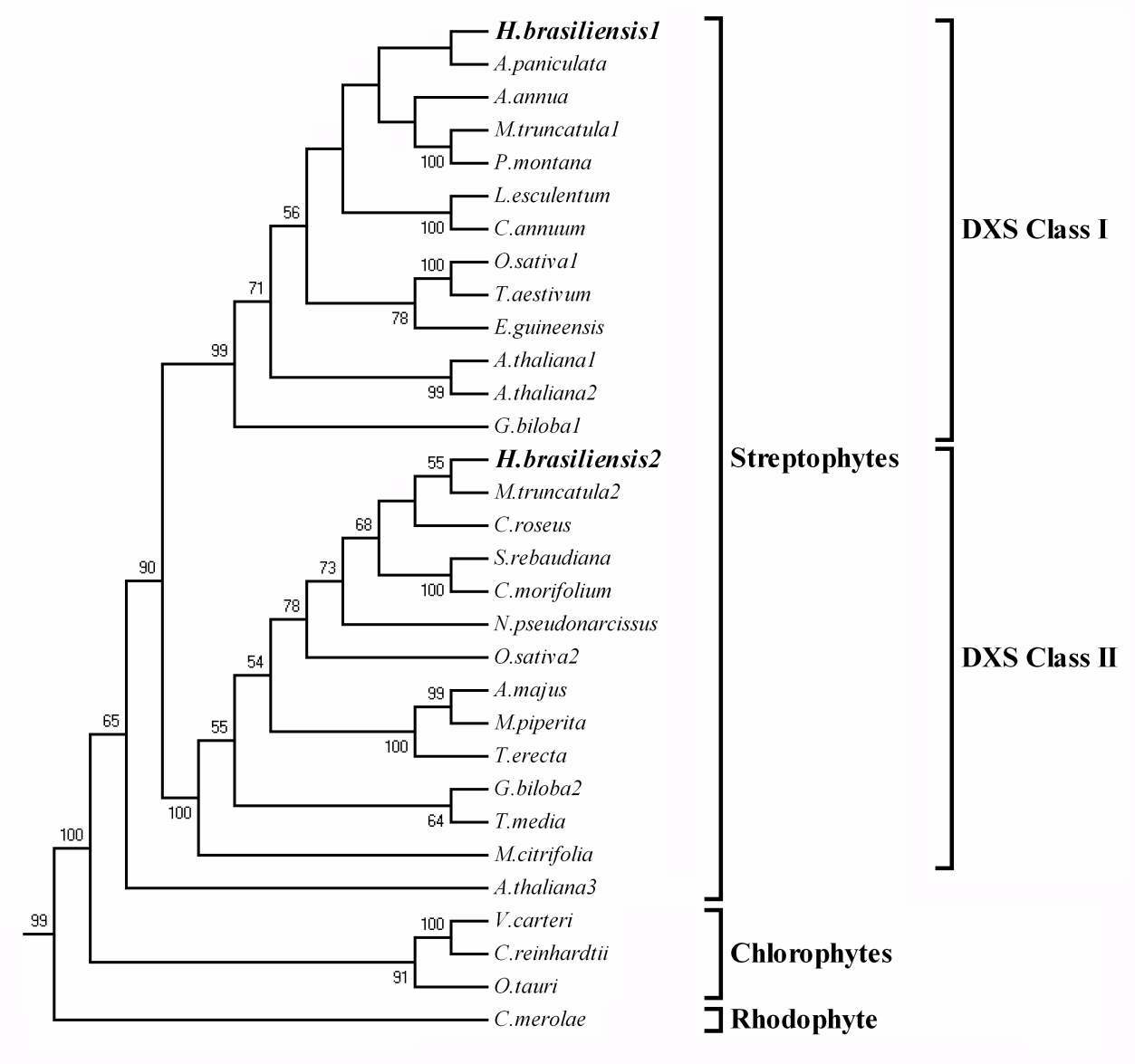


Fig. 3.8 Phylogenetic tree among DXS sequences from different plant species.

The tree was constructed by the MEGA program (version 3.1) (Kumar *et al.*, 2004) using the neighbor-joining method and bootstrapped with 10000 replicates, as described in Fig. 3.7. The species of plants and their GenBank accession nos. used were described in Table 3.2.

1.6 Prediction of secondary and 3D structures of HbDXSs

Recently, the crystal structures DXS from *E. coli* and *D. radiodurans* have been solved (Xiang *et al.*, 2007). Using the crystallographic data from *D. radiodurans* DXS (PDB code 2O1X) as template, the secondary and 3D structure of HbDXS sequences were modeled by the CPH-models 2.0 server (Lund *et al.*, 2002). Amino acids corresponding to cTP sequences in both HbDXSs were deleted during predictions. The average sequence similarity between HbDXSs and *D. radiodurans* DXS was about 39%. The resulting predicted 3D structures of HbDXSs were displayed with the Swiss-Pdb Viewer (Guex and Peitsch, 1997).

As assigned in the 3D structure of *D. radiodurans* DXS (Xiang *et al.*, 2007), both HbDXSs also contain three important domains: domain I, II, and III in each monomer. The predicted secondary and 3D structures of each domain in HbDXSs are shown in Fig. 3.9 and Fig. 3.10, respectively. In the domain I, the HbDXS1 (residues 73-393) was predicted to contain 12 α-helices and 10 β-strands, while the HbDXS2 (residues 66-386) contained 13 α-helices and 11 β-strands (Fig. 3.9 and Fig. 3.10). The TPP-binding is located in this domain (Xiang *et al.*, 2007). In the domain II, the HbDXS1 (residues 394-575) was predicted to contain 5 α-helices and 8 β-strands, while the HbDXS2 (residues 387-568) contained 8 α-helices and 6 β-strands (Fig. 3.9 and Fig. 3.10). In the domain III, the HbDXS1 (residues 576-709) was predicted to contain 6 α-helices and 5 β-strands, while the HbDXS2 (residues 569-702) contained 8 α-helices and 5 β-strands (Fig. 3.9 and Fig. 3.10).

In addition, the overall folding patterns of the predicted 3D structures of both HbDXSs (Fig. 3.11), and their structures after superimposition onto the 3D structure of *D. radiodurans* DXS are also shown in Fig. 3.12. Superimposition of the HbDXSs modeled structures to the *D. radiodurans* DXS template showed very similar foldings (Fig. 3.12). By comparison to the *D. radiodurans* DXS, amino acids forming the active site of both HbDXSs were highly conserved. The active site of DXS contained two regions: TPP- and GAP-binding sites. The amino acids that were predicted to participate in both the TPP- and GAP-binding regions in both HbDXSs are shown in Table 3.3 and the localizations of these amino acids in the 3D models of HbDXSs are shown in Fig. 3.13 and Fig. 3.14, respectively.

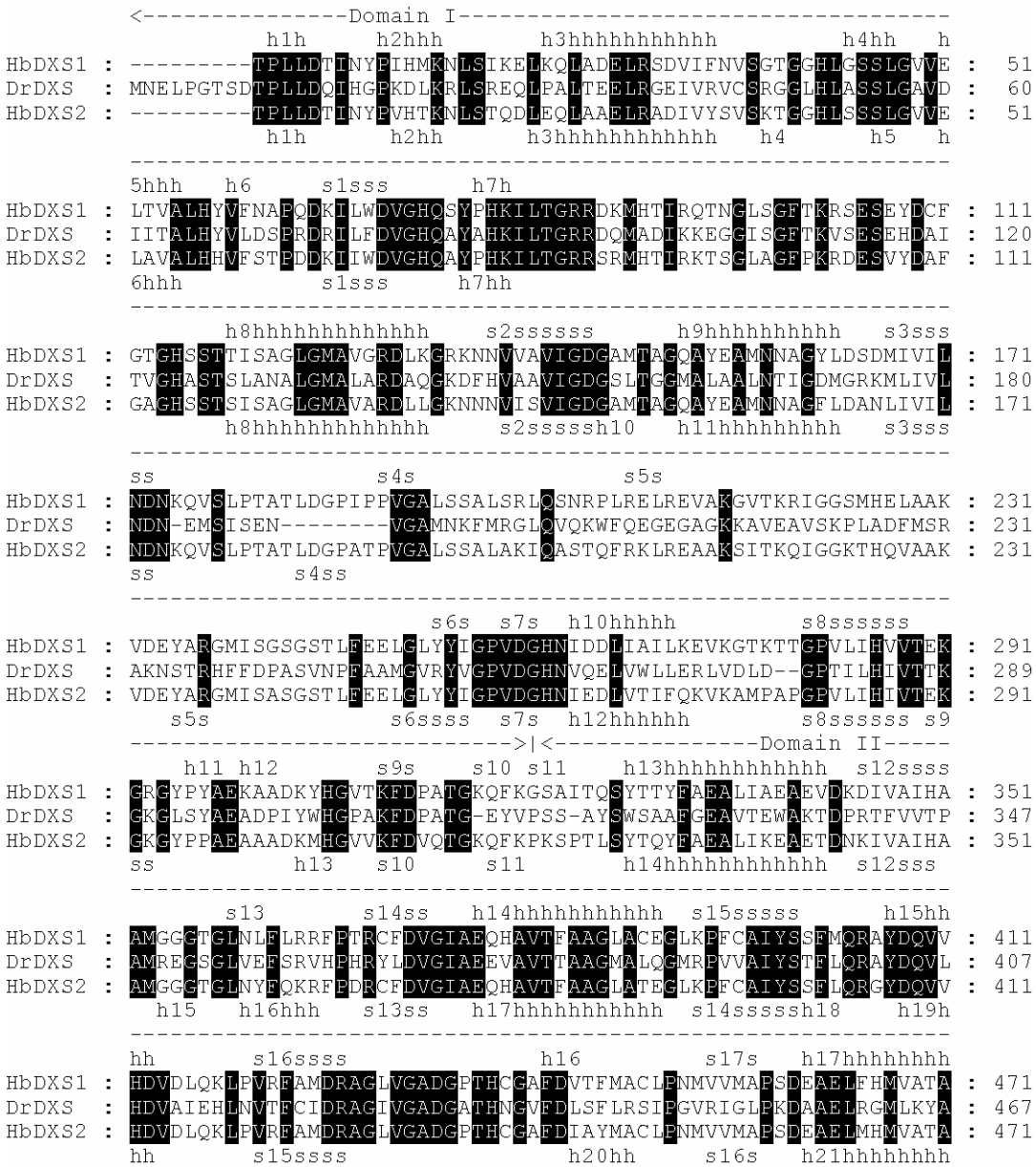


Fig. 3.9 Alignment of the predicted secondary structures of HbDXSs.

The secondary structures of both HbDXSs were predicted using the CPHmodels 2.0 server (Lund *et al.*, 2002) and they are given above (HbDXS1) and below (HbDXS2) the aligned sequences. Alpha-helices are indicated by "h" and β-strands by "s" and numbered from the NH_2-terminus. Domains are assigned according to the *D. radiodurans* DXS structure (PDB code 2O1X) (Xiang *et al.*, 2007).

```
          ------------------------------>|<---------------Domain III--
          h     s18sss                          s19    s20sssss  h18hhhhh
HbDXS1  : AAIDDRPSCFRYPRGNGVGVQLPPGNKGIPLEVGKGRILIEGERVALLGYGTAVQSCLAA : 531
DrDXS   : QTHDG-PFAIRYPRGN---TAQVPAGTWPDLKWGEWERLKGGDDVVILAGGKALDYALKA : 523
HbDXS2  : AAIDDRPSCFRFPRGNGIGAALPPNNKGTPLEIGKGRILMEGNRVAILGYGSIVQQCVEA : 531
          h     s17sss                          s18s     s19s   h22hh
          ------------------------------------------------------------
                          s21 s22    h19h   h20  s23ss        h21hhhhhhhhh
HbDXS1  : ASLVEPHGLLITVADARFCKPLDHTLIRSLAKPHEVLITVEE-GSIGGFGSHVAHFLALD : 590
DrDXS   : AEDLP----GVGVVNARFVKPLDEEMLREVGGRARALITVEDNTVVGGFGGAVLEALNSM : 579
HbDXS2  : ASMLRTQGISVTVADARFCKPLDTDLIRQLAKEHEFLITVEE-GSIGGFSSHVSHFLSLS : 590
                        s20sssss    h23hhhhhh   s21ss        h24hhhhhhhhh
          ------------------------------------------------------------>|
                                    h22hhhhh   h23hhhhhhhhh
HbDXS1  : GLLDGKLKWRPLVLPDRYIDHGSPSVQLIEAGLTPSHVAATVLNILGNKREALQIMSS : 648
DrDXS   : NLHP---TVRVLGIPDEFQEHATAESVHARAGIDAPAIR-TVLAELG-VDVPIEV--- : 629
HbDXS2  : GILDGPLKLRAMVLPDRYIDHGSPQDQIQEAGISSNHITATVLSLLGKPKEALQFK-- : 646
                      s22           h25hhhh   h26h   h27hh
```

Fig. 3.9 (continued)

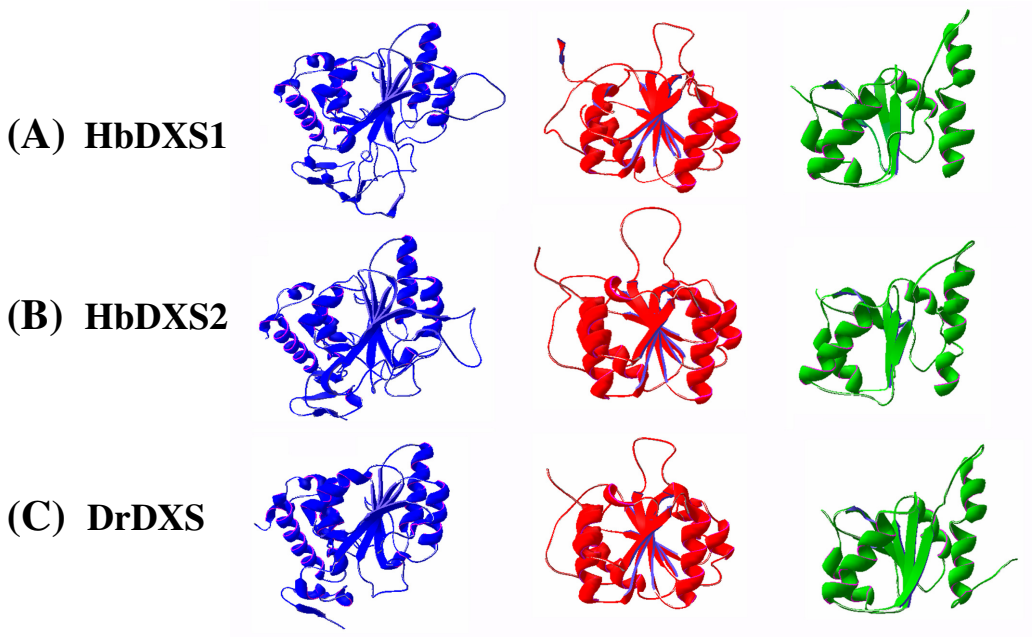


Fig. 3.10 The predicted 3D structures of domains in HbDXSs.

The 3D structure models of both HbDXSs were predicted using the CPHmodels 2.0 server (Lund *et al.*, 2002) and displayed with the Swiss-Pdb Viewer (Guex and Peitsch, 1997). (A) HbDXS1. (B) HbDXS2. (C) *D. radiodurans* DXS. Domain I, II, and III are colored in blue, red, and green, respectively.

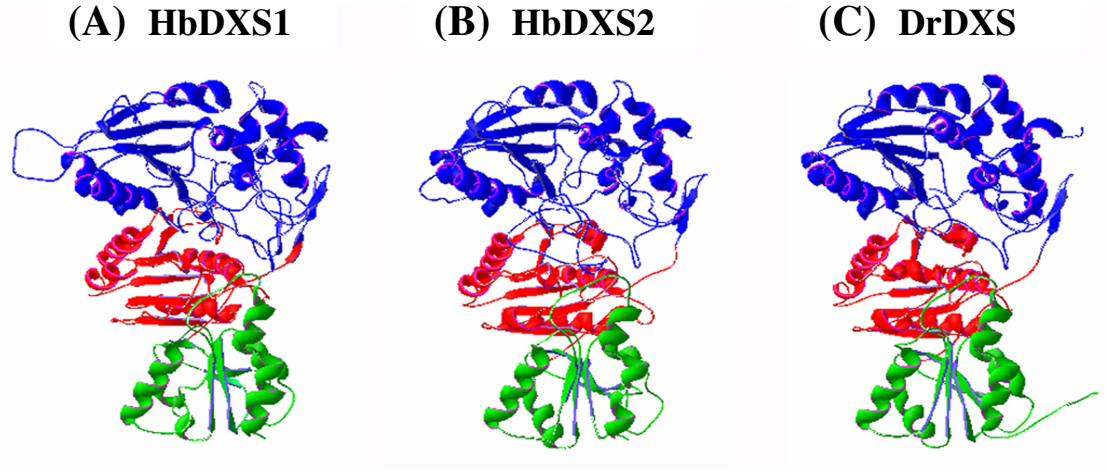


Fig. 3.11 The overall predicted 3D structures of HbDXSs.

(A) HbDXS1. (B) HbDXS2. (C) *D. radiodurans* DXS. Domains are colored in a similar fashion to Fig. 3.10.

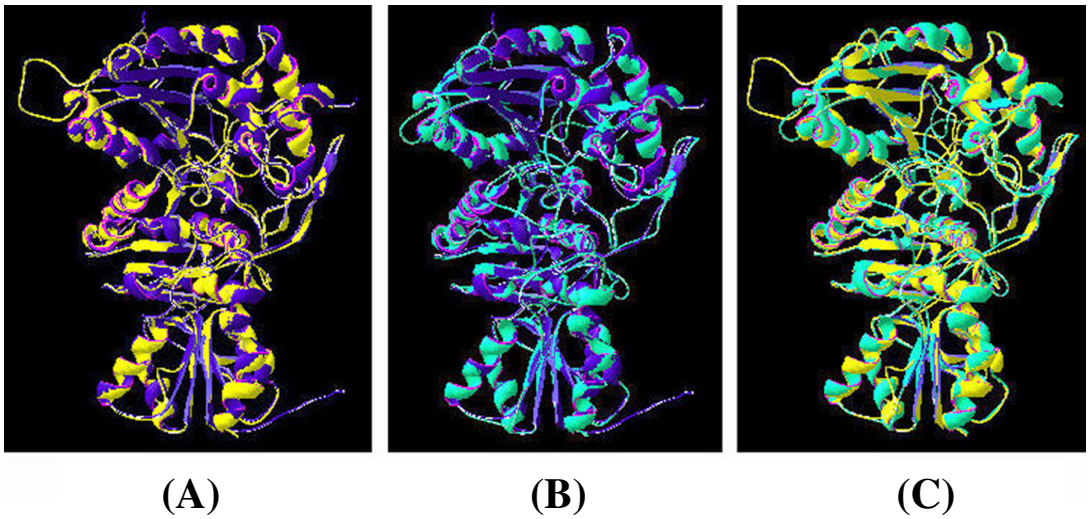


Fig. 3.12 Superimposition of the predicted 3D structures of HbDXSs.

(A) Superimposition of HbDXS1 (yellow) and *D. radiodurans* DXS (blue). (B) Superimposition of HbDXS2 (green) and *D. radiodurans* DXS. (C) Superimposition of HbDXS1 and HbDXS2.

Table 3.3 Amino acids involved in the TPP- and GAP-binding sites in HbDXSs.

The putative amino acids that might be involved in the TPP- and GAP-binding sites in HbDXSs were deduced based on the alignment of amino acid sequences of HbDXSs and *D. radiodurans* DXS (Xiang *et al.*, 2007).

TPP-binding site			GAP-binding site		
DrDXS	HbDXS1	HbDXS2	DrDXS	HbDXS1	HbDXS2
Ser54	Ser117	Ser110	His51	His114	His107
His82	His145	His138	Lys101	Arg164	Arg157
Gly123	Gly186	Gly179	Phe109	Phe172	Phe165
Ala125	Ser188	Ser181	Ile187	Leu251	Leu244
Asp154	Asp217	Asp210	Ser325	Thr401	Thr394
Asn183	Asn246	Asn239	Met349	Met425	Met418
Met185	Val249	Val242	Gly352	Gly428	Gly421
Lys289	Lys363	Lys356	Tyr395	Tyr471	Tyr464
Ile371	Ile447	Ile440	Arg423	Arg499	Arg492
Glu373	Glu449	Glu442	Asp430	Asp506	Asp499
Phe398	Phe474	Phe467	His434	His510	His503
Arg401	Arg477	Arg470	Arg480	Arg557	Arg550

Note:

DrDXS means the DXS sequence from *D. radiodurans*.

The amino acids of HbDXSs that found in the same position but showed different residues when alignments with the DXS sequence from *D. radiodurans* are shown in grey-labeled.

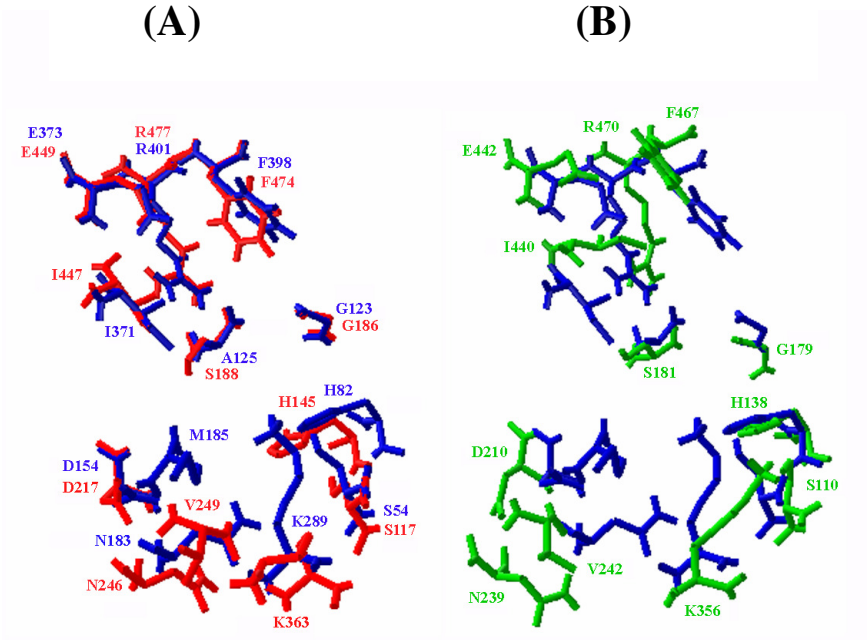


Fig. 3.13 Localization of amino acids involved in the TPP-binding site.

(A) Comparison between *D. radiodurans* DXS (blue) and HbDXS1 (red). (B) Comparison between *D. radiodurans* DXS and HbDXS2 (green).

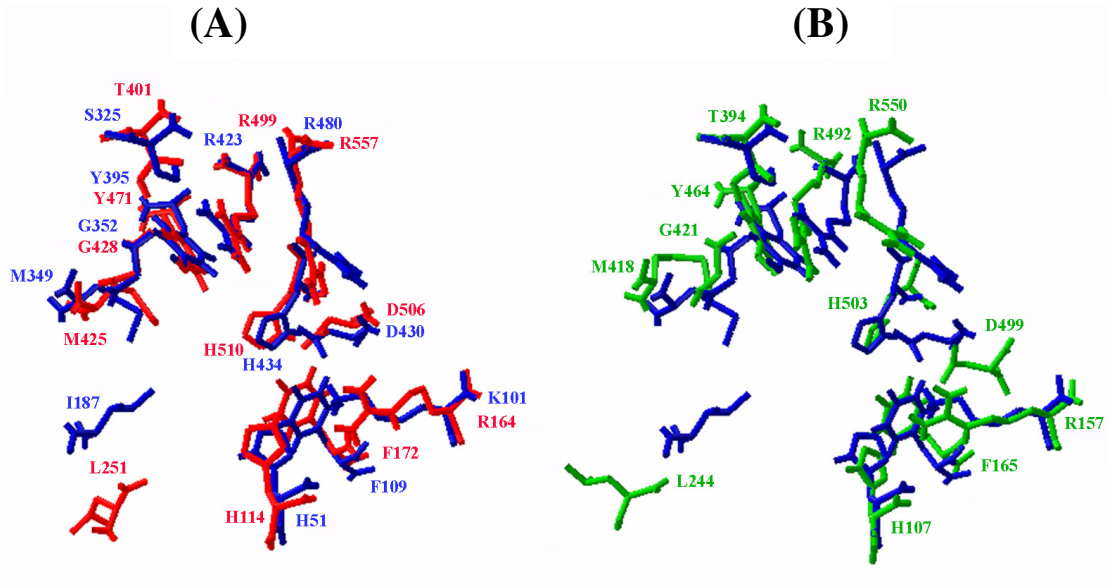


Fig. 3.14 Localization of amino acids involved in the GAP-binding site.

(A) Comparison between *D. radiodurans* DXS (blue) and HbDXS1 (red). (B) Comparison between *D. radiodurans* DXS and HbDXS2 (green).

2. Molecular cloning and sequence analysis of cDNA encoding HbDXR

2.1 A partial sequence of *HbDXRs*

Beside *DXS* gene, a remarkable conservation of both nucleotide and amino acid sequences among DXRs from *M. piperita* (Lange and Croteau, 1999), *A. thaliana* (Schwender *et al.*, 1999), *A. annua* (Souret *et al.*, 2002), and *Z. mays* (Hans *et al.*, 2004) also made it possible to design degenerate primers for cloning of a *DXR* gene from *H. brasiliensis*. Two degenerate primers, GSI-DXR1 (sense orientation) and GWP-DXR2 (antisense orientation) (Table 2.1), corresponding to amino acid sequences GSIGTQT and GWPDMRL, respectively, were synthesized. Using these primers in One-step RT-PCR, cDNA fragments of about 0.7 kb in length were obtained, when total RNA isolated from leaves (Fig. 3.15, lane 1) and latex (Fig. 3.16, lane 1) were used as templates. These cDNAs were purified and cloned into pDrive or pGEM-T Easy vector for sequencing.

Sequence analysis by searching against GenBank database using the BLAST program (Altschul *et al.*, 1997) revealed that these two cDNAs showed high homology to the conserved region of plant DXRs with >80% identity. These results suggested that the two cDNAs could encode a *DXR* gene in *H. brasiliensis*. Surprisingly, a pairwise analysis revealed that these cDNAs exhibited 94% identity to each other, implying that both of them encode the different forms of *DXR* gene in *H. brasiliensis*. Therefore, the *DXR* cDNA derived from *H. brasiliensis* leaves was named *HbDXR1* and that from the latex of *H. brasiliensis* was *HbDXR2*.

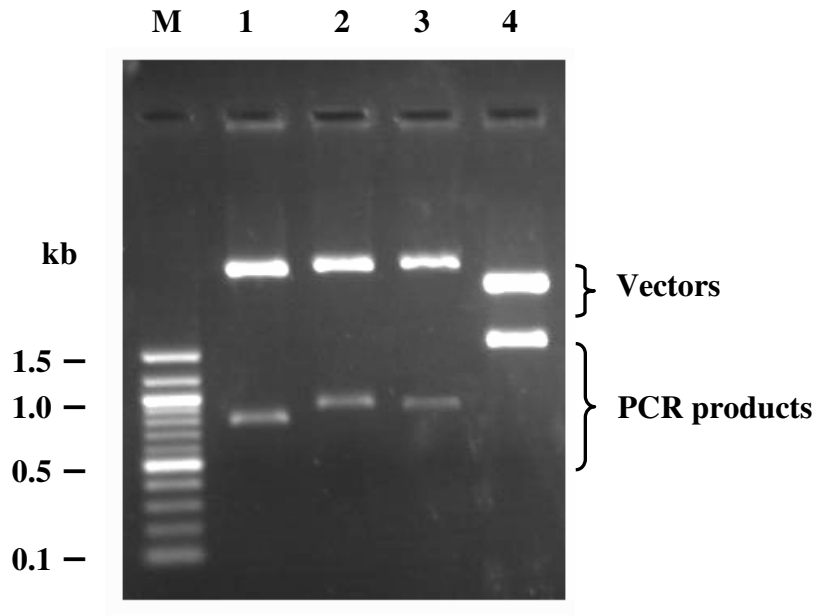


Fig. 3.15 Agarose gel analysis of the *HbDXR1* cDNA.

Plasmid DNAs containing the expected cDNAs of *HbDXR1* were digested with restriction enzyme *Eco*RI at 37ºC for 1 h and separated by electrophoresis on a 1.0% agarose gel containing 0.1 mg/ml of ethidium bromide. Lane 1 is a pDrive vector containing 0.7 kb RT-PCR product from One-step RT-PCR. Lanes 2 and 3 are pCR4-TOPO vectors containing 0.8 kb RACE-PCR products from 5´- and 3´-RACE, respectively. Lane 4 is a pGEM-T Easy vector containing 1.5 kb ORF of *HbDXR1*. Lane M is a 100 bp DNA Ladder (New England Biolabs).

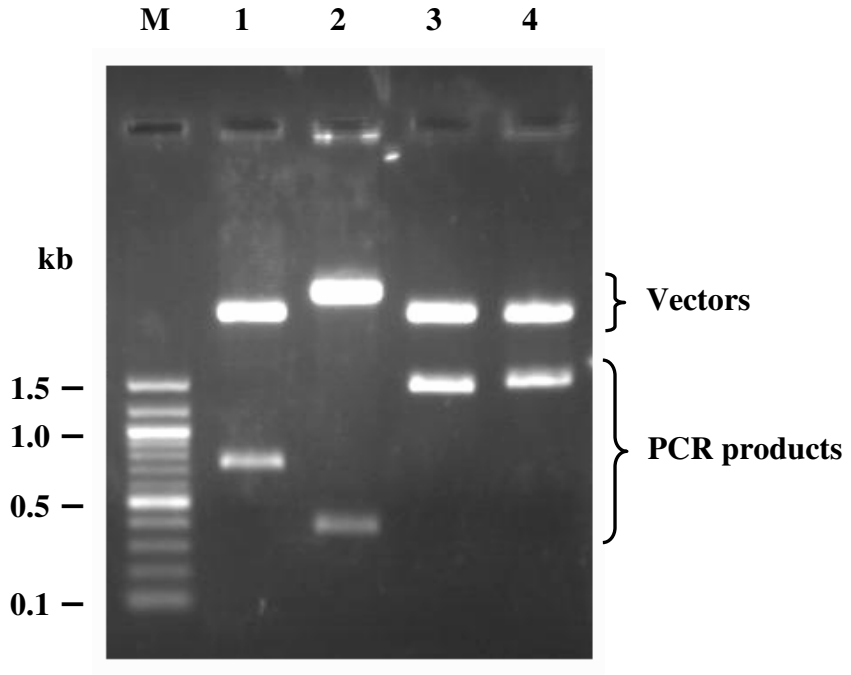


Fig. 3.16 Agarose gel analysis of the *HbDXR2* cDNA.

Plasmid DNAs containing the expected cDNAs of *HbDXR2* were digested with restriction enzyme *Eco*RI at 37ºC for 1 h and separated by electrophoresis on a 1.0% agarose gel containing 0.1 mg/ml of ethidium bromide. Lanes 1, 3, and 4 are pGEM-T Easy vectors containing 0.7 kb RT-PCR product from One-step RT-PCR, 1.4 kb RACE-PCR product from 3´-RACE, and 1.5 kb ORF of *HbDXR2*, respectively. Lane 2 is a 0.4 kb RACE-PCR product from 5´-RACE inserted into pCR4-TOPO vector. Lane M is a 100 bp DNA Ladder (New England Biolabs).

2.2 A full-length *HbDXR1* cDNA

The nucleotide sequence obtained from One-step RT-PCR as described above was also used to design gene-specific primers for amplification and cloning of the 5´- and 3´-ends of *HbDXR1* gene. The RLM-RACE was performed using the first-strand cDNA prepared from total RNA isolated from leaves as template. The 5´-end of *HbDXR1* was amplified with primers GSP1-HbDXR1/-GeneRacer 5´ in the first round of PCR. The same gene-specific primer together with the GeneRacer 5´ nested primer was used in the second round of PCR (Table 2.2). The 3´-end of *HbDXR1* was obtained with primers GSP1-HbDXR1/GeneRacer 3´ in the first round of PCR and primers GSP2-HbDXR1/GeneRacer 3´ nested in the second round of PCR, respectively (Table 2.3). Agarose gel electrophoresis of RACE-PCR products showed cDNA fragments of about 0.8 kb in length in both 5´- and 3´-RACE experiments, respectively (Fig. 3.15, lane 2 and lane 3, respectively). The cDNAs were purified, cloned either into pDrive or pGEM-T Easy vector, and sequenced. Sequence analysis revealed that both cDNAs contained a portion of coding regions, and the 5´-UTR and 3´-UTR in the 5´- and 3´-RACE-cDNA clones, respectively.

By combining and aligning of nucleotide sequences of these cDNAs, a full-length cDNA of *HbDXR1* was derived and it was further confirmed by PCR using the primers GSP1/2-HbDXR1 (Table 2.4). The final PCR product was purified and cloned into pGEM-T Easy vector (Fig. 3.15, lane 4), followed by DNA sequencing. Sequence analysis revealed that the full-length *HbDXR1*cDNA contained an ORF of 1416 bp in length encoding a polypeptide of 471 amino acids with a predicted molecular mass of about 51 kDa and a pI value of 5.67. The complete nucleotide and the deduced amino acid sequences of *HbDXR1* have been deposited to GenBank database under the accession number AY502937 (Fig. 3.17).

```
 1   atggagcttaatttgctttcccctgctgaaatcaaggccatttccttcttagattccacc
     M  E  L  N  L  L  S  P  A  E  I  K  A  I  S  F  L  D  S  T    20
61   aagtccaaccaccttcctaagtttccaggtgttatcagtctgaagaggaaggattttggg
     K  S  N  H  L  P  K  F  P  G  V  I  S  L  K  R  K  D  F  G    40
121  gcagcatttgggaggaaagtgcagtgttcggcccagcctcctccaccagcctggccagga
     A  A  F  G  R  K  V  Q  C  S  A  Q  P  P  P  P  A  W  P  G    60
181  agagcttttccagagccaggtcgtaagacttgggatggtccaaagcctatttcagtcgtt
     R  A  F  P  E  P  G  R  K  T  W  D  G  P  K  P  I  S  V  V    80
241  ggatccactggctccatcgggactcagactttggacatcgtagcagagaatccaggtaaa
     G  S  T  G  S  I  G  T  Q  T  L  D  I  V  A  E  N  P  G  K   100
301  ttcagagttgtagcactcgcagctggttcaaatgttactcttcttgctgatcaggtgaag
     F  R  V  V  A  L  A  A  G  S  N  V  T  L  L  A  D  Q  V  K   120
361  actttcaaacctcaacttgttgctgttaggaatgagtcgttagctgatgaactcaaggac
     T  F  K  P  Q  L  V  A  V  R  N  E  S  L  A  D  E  L  K  D   140
421  gctttggctgatcttgaagaaaagcctgagattattcctggggaggaaggtgttgttgag
     A  L  A  D  L  E  E  K  P  E  I  I  P  G  E  E  G  V  V  E   160
481  gttgctcgccatccagatgctgtcagtgtagttacgggaatagtaggttgtgcaggctta
     V  A  R  H  P  D  A  V  S  V  V  T  G  I  V  G  C  A  G  L   180
541  aagcctactgttgctgccatagaagcaggaaaagatatatgcttggccaataaagagacg
     K  P  T  V  A  A  I  E  A  G  K  D  I  C  L  A  N  K  E  T   200
601  ttaatcgctggaggtccttttgtccttcctctggctcacaaatataacgtaaaaattctt
     L  I  A  G  G  P  F  V  L  P  L  A  H  K  Y  N  V  K  I  L   220
661  ccagctgattcagaacattctgccgtatttcagtgtattcaaggcctgccagatggtgca
     P  A  D  S  E  H  S  A  V  F  Q  C  I  Q  G  L  P  D  G  A   240
721  ctgcggcgtattatttaactgcttctggtggagctttcagggattggcctgttgataaa
     L  R  R  I  I  L  T  A  S  G  G  A  F  R  D  W  P  V  D  K   260
781  ttgaaagaagttaaagtagctgatgctttaaagcatcctaattggaatatggggaaaaag
     L  K  E  V  K  V  A  D  A  L  K  H  P  N  W  N  M  G  K  K   280
841  attactgtggactctgctacccttttcaataagggcttagaagtcattgaagcccattat
     I  T  V  D  S  A  T  L  F  N  K  G  L  E  V  I  E  A  H  Y   300
901  ttgtttggagctgactatgataatattgagatagtaattcatccacaatctataatacat
     L  F  G  A  D  Y  D  N  I  E  I  V  I  H  P  Q  S  I  I  H   320
961  tcaatggttgaaacacaggattcatctgttcttgcacagctggggtggcctgatatgcgc
     S  M  V  E  T  Q  D  S  S  V  L  A  Q  L  G  W  P  D  M  R   340
1021 ctaccaattctatgcactatgtcatggcccgagagaatatactgctctgaaataacctgg
     L  P  I  L  C  T  M  S  W  P  E  R  I  Y  C  S  E  I  T  W   360
1081 cctcgccttgatctttgcaagcttgggtctctaacatttaaagctcctgacaatgtaaaa
     P  R  L  D  L  C  K  L  G  S  L  T  F  K  A  P  D  N  V  K   380
1141 tacccatccatggatcttgcctatgctgctggtcgggctggaggcaccatgaccggagtg
     Y  P  S  M  D  L  A  Y  A  A  G  R  A  G  G  T  M  T  G  V   400
1201 cttagtgctgctaatgagaaagctgtagagatgttcatagatgaaaagatcagctatctt
     L  S  A  A  N  E  K  A  V  E  M  F  I  D  E  K  I  S  Y  L   420
1261 gatattttcaagattgtggagctaacttgtgataaacatagggcagaattggtgacctca
     D  I  F  K  I  V  E  L  T  C  D  K  H  R  A  E  L  V  T  S   440
1321 ccctctctcgaggaaattatacattatgacttgtgggcacgagactatgctgctagttta
     P  S  L  E  E  I  I  H  Y  D  L  W  A  R  D  Y  A  A  S  L   460
1381 caaccctcttctggtctaagccctgttcttgcatga
     Q  P  S  S  G  L  S  P  V  L  A  -                          471
```

Fig. 3.17 Nucleotide and deduced amino acid sequences of the *HbDXR1* cDNA.

A translated amino acid sequence is shown in a single-letter code below the nucleotide sequence. Dash (-) indicates a translation stop codon.

2.3 A full-length *HbDXR2* cDNA.

A partial *HbDXR2* cDNA was amplified when total RNA isolated from latex was used as temple. To obtain the missing sequences of the 5´- and 3´-ends of the cDNA, gene-specific primers were designed and also used in the RLM-RACE amplification. The first-strand cDNA for the 5´- and 3´-ends amplification prepared from total RNA that was isolated from latex was used as template. The 5´-end of *HbDXR2* cDNA was amplified with primers GSP1-HbDXR1/GeneRacer 5´ and primers HbDXR1/GeneRacer 5´ nested in the first round and second rounds of PCR, respectively (Table 2.2). The 3´-end of *HbDXR2* cDNA was obtained using three rounds of PCR with different sets of the primers. The first round of PCR was performed with primers GSP1-HbDXR2/GeneRacer 3´, the second round of PCR with primers GSP2-HbDXR2/GeneRacer 3´ nested, and the third round of PCR with primers GSP3-HbDXR2/GeneRacer 3´ nested, respectively (Table 2.3). The amplified products of the 5´- and 3´-ends were about 0.4 kb and 1.4 kb in length, respectively, as shown in 1.0% agarose gel electrophoresis (Fig. 3.16, lane 2 and lane 3, respectively). The PCR products were purified, cloned into either pDrive or pGEM-T Easy vector, and sequenced. Nucleotide sequence analysis showed that the 5´-end of cDNA contained the 5´-UTR and a portion of the coding region of the *HbDXR2* gene and that for the 3´-end spanned the 3´-UTR and part of the ORF.

By combining and aligning of nucleotide sequences of these three cDNAs, a full-length cDNA of *HbDXR2* was obtained. Finally, the cDNA was amplified by PCR using the primers GSP2-HbDXR1/GSP2-HbDXR2 (Table 2.4) to confirm that all fragments were derived from the same gene. The resulting PCR product was purified, cloned into pGEM-T Easy vector (Fig. 3.16, lane 4), and sequenced. Sequence analysis revealed that the ORF of *HbDXR2* contained 1416 bp, which is predicted to encode a putative protein of 471 amino acids. The predicted molecular mass and a pI value of the deduced amino acid sequence of *HbDXR2* were 51 kDa and 6.04, respectively. In addition, the nucleotide sequence of *HbDXR2* showed 94% identity to the *HbDXR1* cDNA. The complete nucleotide and the deduced amino acid sequences of *HbDXR2* have been submitted to GenBank database under the accession number DQ473432 (Fig. 3.18).

```
1    atggcgctcaatttgctttcccctgctgaaatcaaggctatctccttcttagattccacc
     M  A  L  N  L  L  S  P  A  E  I  K  A  I  S  F  L  D  S  T    20
61   aagtccagccaccttactaagcttccaggtggtttcagtttaaagaggaaggattttggg
     K  S  S  H  L  T  K  L  P  G  G  F  S  L  K  R  K  D  F  G    40
121  gcagcatttgggaagaaagtgcagtgttcggcccagcctcctccaccagcctggccagga
     A  A  F  G  K  K  V  Q  C  S  A  Q  P  P  P  P  A  W  P  G    60
181  agagcttttccagatttaggccgtaagacttgggatggcccaaagcctatttcagtcgtt
     R  A  F  P  D  L  G  R  K  T  W  D  G  P  K  P  I  S  V  V    80
241  ggatccactggctccattgggactcagacattggacatcgtggcagagaatccagataaa
     G  S  T  G  S  I  G  T  Q  T  L  D  I  V  A  E  N  P  D  K   100
301  ttcagagttgtggcactcgcagctggttcaaatgttactcttcttgcagatcaggtgaag
     F  R  V  V  A  L  A  A  G  S  N  V  T  L  L  A  D  Q  V  K   120
361  actttcaaacctcaacttgttgctgttagaaatgagtcattagttcatgaactcagagaa
     T  F  K  P  Q  L  V  A  V  R  N  E  S  L  V  H  E  L  R  E   140
421  gctttggctgatgttgaagaaaaacctgagattattcctggggagcaaggagttgttgag
     A  L  A  D  V  E  E  K  P  E  I  I  P  G  E  Q  G  V  V  E   160
481  gttgctcgccatccagatgctgtcagtgtagttacaggaatagtaggttgtgcaggctta
     V  A  R  H  P  D  A  V  S  V  V  T  G  I  V  G  C  A  G  L   180
541  aagcctacggtggctgcaatagaagctggaaaagacatatgcttggccaataaagagaca
     K  P  T  V  A  A  I  E  A  G  K  D  I  C  L  A  N  K  E  T   200
601  ttaattgctggagggccctttgtccttcctcttgctcacaaatataatgtgaaaattctc
     L  I  A  G  G  P  F  V  L  P  L  A  H  K  Y  N  V  K  I  L   220
661  ccggctgattcagaacattctgctatatttcagtgtattcaaggcctgccagatggtgca
     P  A  D  S  E  H  S  A  I  F  Q  C  I  Q  G  L  P  D  G  A   240
721  ctgcggcgtattattttaactgcttcaggtggagctttcagggattggcctgttgataaa
     L  R  R  I  I  L  T  A  S  G  G  A  F  R  D  W  P  V  D  K   260
781  ttgaaagaagttaaagtagctgatgctttaaagcatcctaactggaatatggggaaaaag
     L  K  E  V  K  V  A  D  A  L  K  H  P  N  W  N  M  G  K  K   280
841  attacagtggactccgctacccttttcaataagggtttagaagtcattgaagcccattat
     I  T  V  D  S  A  T  L  F  N  K  G  L  E  V  I  E  A  H  Y   300
901  ttgtttggagctgagtatgataatattgagatagtaattcatccacaatctataatacat
     L  F  G  A  E  Y  D  N  I  E  I  V  I  H  P  Q  S  I  I  H   320
961  tcaatggttgaaacacaggattcatctgttcttgcacagttggggtggcccgatatgcgt
     S  M  V  E  T  Q  D  S  S  V  L  A  Q  L  G  W  P  D  M  R   340
1021 ttaccaattctatatactatgtcatggcctgacagaatatactgctctgaaataacctgg
     L  P  I  L  Y  T  M  S  W  P  D  R  I  Y  C  S  E  I  T  W   360
1081 cctcgccttgacctttgcaagcttgggtctctaacatttaaagctcctgacaatgtaaag
     P  R  L  D  L  C  K  L  G  S  L  T  F  K  A  P  D  N  V  K   380
1141 taccccttctatggatcttgcctatgctgctggacgggctggaggcaccatgactggagtg
     Y  P  S  M  D  L  A  Y  A  A  G  R  A  G  G  T  M  T  G  V   400
1201 cttagtgctgcgaatgagaaagctgttgagatgttcatcaatgaaaagatcggctatctt
     L  S  A  A  N  E  K  A  V  E  M  F  I  N  E  K  I  G  Y  L   420
1261 gatattttcaagattgtggagctaacgtgtgataaacataggtcagaactggtggcgtca
     D  I  F  K  I  V  E  L  T  C  D  K  H  R  S  E  L  V  A  S   440
1321 ccctctctcgaggaaattatacattatgacttgtgggcacgagactatgctgctagtttg
     P  S  L  E  E  I  I  H  Y  D  L  W  A  R  D  Y  A  A  S  L   460
1381 caacccacttctggtctaagccctgttcttgcatga
     Q  P  T  S  G  L  S  P  V  L  A  -                          471
```

Fig. 3.18 Nucleotide and deduced amino acid sequences of the *HbDXR2* cDNA.

A translated amino acid sequence is shown in a single-letter code below the nucleotide sequence. Dash (-) indicates a translation stop codon.

2.4 Comparison of the deduced *HbDXRs* amino acid sequences

A comparison of the deduced amino acid sequences of both HbDXRs showed 94% identity to each other but had 25 amino acid differences (Fig. 3.19). An extensive search against the GenBank database revealed that both HbDXRs had high similarity with those of other DXR sequences from various organisms. The highest sequence similarity was found in plants, ranging from 76% identity with a DXR from *Pinus* sp. to 88% identity with a DXR from *L. esculentum* (Rodríguez-Concepción *et al.*, 2001) (Table 3.4). In addition, a reduced level of similarity was observed with DXR sequences from cyanobacteria (56-69% identity), bacteria (40-50% identity), and from the malaria parasite (40% identity) counterparts, respectively (data not shown).

Table 3.4 Percent identity of the deduced amino acid sequences of HbDXRs with some representative plant DXRs.

The DXR sequences from various plants obtained from GenBank and TIGR databases were compared using the ClustalW program (Thompson *et al.*, 1994).

Species	% identity	
	H. brasiliensis1 (HbDXR1)	*H. brasiliensis2* (HbDXR2)
L. esculentum	87	88
C. roseus	86	87
A. thaliana	84	85
Z. mays	83	83
S. rebaudiana	80	81
A. annua	80	80
M. piperita	79	80
G. biloba	78	78
T. cuspidata	77	76
Pinus sp.	76	76

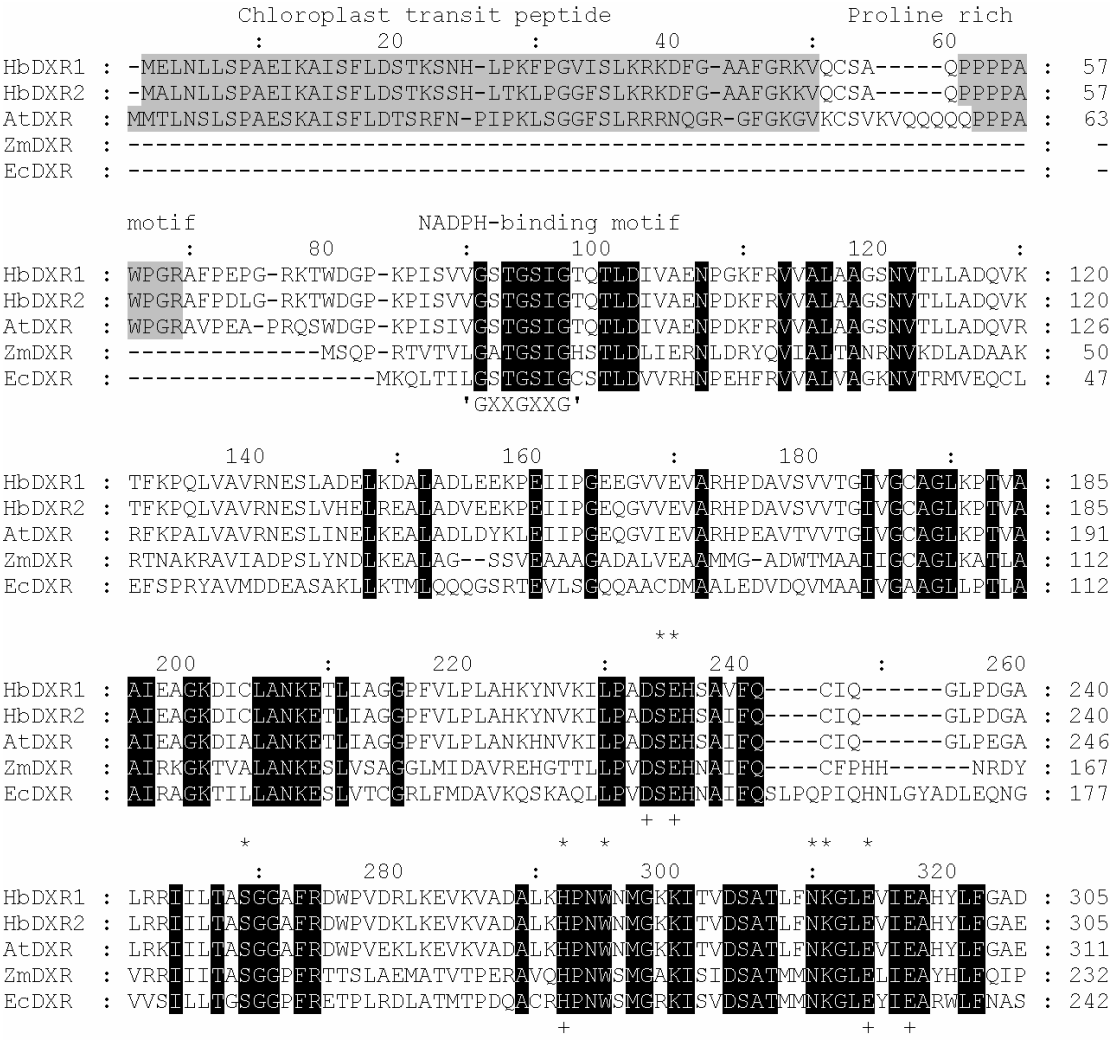


Fig. 3.19 Alignment of the deduced amino acid sequences of HbDXRs and some representative plant DXRs.

Multiple alignments were performed using the ClustalX software (Thompson *et al.*, 1997). Identical residues are shaded in black. Dashes (-) indicate gaps inserted to optimize the alignment. The putative cTP sequences, a proline rich motif, and the NADPH-binding motif are indicated. Several residues important for activity of DXR are shown by plus symbols (+). Residues for DXP binding are shown by stars. The sequences used were *H. brasiliensis* (HbDXR1, AY502937 and HbDXR2, DQ473432), *A. thaliana* (AtDXR, AF148852), *Z. mobilis* (ZmDXR, AAD29659), and *E. coli* (EcoDXR, AB013300).

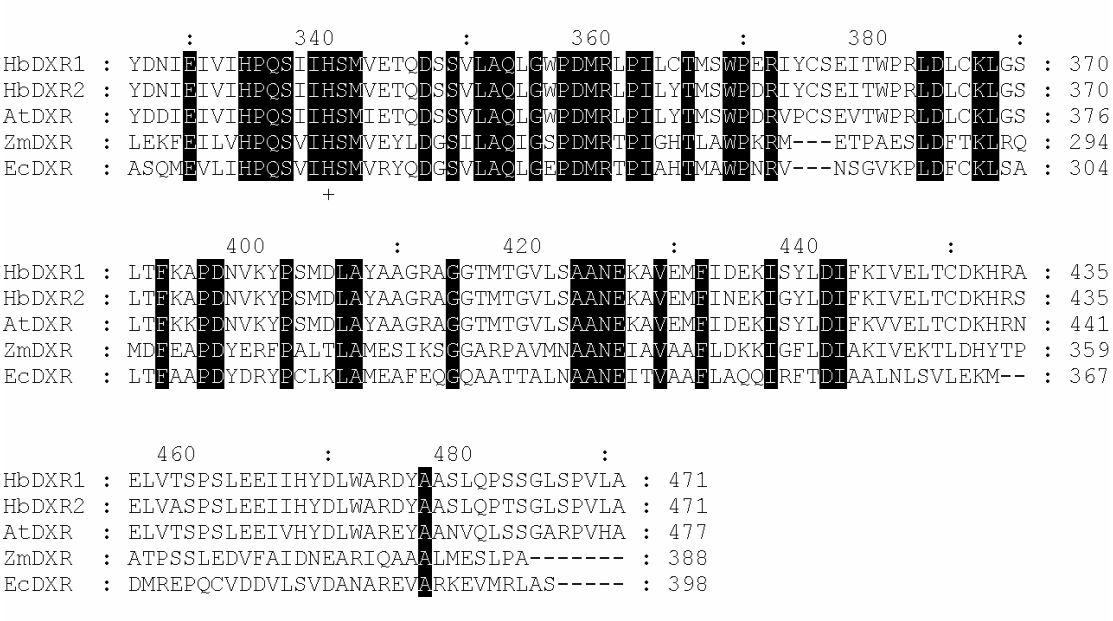


Fig. 3.19 (continued)

As shown in Fig. 3.19, multiple amino acid sequence alignments of both HbDXRs with those of other DXRs from various organisms such as plants (*A. thaliana*) and bacteria (*E. coli* and *Z. mobilis*) revealed two highly conserved motifs. The first motif is the NADPH-binding motif 'GXXGXXG' and the second motif is the Proline rich region 'P_{2-4}AWPG(R/T)'. The NADPH-binding motif was found at amino acids 81-87 at the NH_2-terminal regions of both HbDXRs (Fig. 3.19). The Proline rich region was also found at the NH_2- terminal regions but at the amino acids 53-61 in both HbDXRs (Fig. 3.19). This region was not found in prokaryotic DXRs such as *Z. mobilis* and *E. coli* (Fig. 3.19). Therefore, its biological function in plants is still unknown.

Beside these two conserved regions, amino acids that previously shown to be important for the activity of DXR, such as aspartic acid (D150), glutamic acid (E152, E231, and E234), and histidine (H209 and H257) with residue numbers assigned according to the *E. coli* DXR (Kuzuyama *et al.*, 2000; Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; MacSweeney *et al.*, 2005; Fernandes and Proteau, 2006), were also found to be highly conserved in both HbDXRs. These amino acids correspond to amino acids D223, E225, E294, E297, H272, and H320 in both HbDXRs sequences (Fig. 3.19). In addition, residues for the DXP-binding site, such as serine (S151 and S186), asparagine (N227), lysine (K228), histidine (H209), tryptophane (W212), and glutamic acid (E152 and E231) with residue numbers assigned according to the *E. coli* DXR (MacSweeney *et al.*, 2005), were conserved in both HbDXRs. These amino acids were found at positions S224, S249, N290, K291, H272, W275, E225, and E294 in both HbDXRs (Fig. 3.19).

Like DXS, all DXR proteins contain extensions of about 80 amino acids at the NH_2-terminus. These regions were also found to be rich in hydroxylated and basic amino acids, accounting for six hydroxylated amino acids (5 serines and 1 threonine, 12.50%) and nine basic amino acids (2 arginines, 6 lysines, and 1 histidine, 18.75%) in the HbDXR1 sequence, and eight hydroxylated amino acids (6 serines and 2 threonines, 16.67%) and nine basic amino acids (1 arginine residue, 7 lysines, and 1 histidine, 18.75%) in the HbDXR2 sequence. Computer analysis using the ChloroP program (Emanuelsson *et al.*, 1999) revealed that most plant DXRs (8 sequences out of 12 sequences determined) were predicted to be localized in the chloroplasts (Table 3.5). As suggested by the ChloroP program, the possible cleavage sites for cTP sequences of both HbDXRs were predicted to be localized between amino acids 48 and 49 (Fig. 3.19 and

Fig. 3.20). In addition, multiple alignments of all plant DXRs exhibited a similar conserved motif for the cTP sequence cleavage site 'Xaa↓Cys-Ser-Xaa', where Xaa means any amino acid (Fig. 3.19 and Fig. 3.20). When the HbDXRs sequences were cleaved at these cleavage sites, the putative mature proteins of both HbDXRs were predicted to contain 423 amino acids with a predicted molecular mass of about 46 kDa. The predicted pI values of both HbDXR1 and HbDXR2 proteins were 5.24 and 5.45, respectively.

Table 3.5 Prediction of the localization and length of cTP sequences in HbDXRs and in various plant DXRs.

The localization and length of cTP sequences in HbDXRs and in various plant DXRs were predicted by the ChloroP program (Emanuelsson *et al.*, 1999).

Species	Accession nos.	ChloroP		
		Localization	cTP score	cTP length
H. brasiliensis1 (HbDXR1)	AY502937	-	0.485	48
H. brasiliensis2 (HbDXR2)	DQ473432	-	0.500	48
L. esculentum	AF331705	Chloroplast	0.509	51
C. roseus	AF250235	Chloroplast	0.509	51
A. thaliana	AF148852	Chloroplast	0.539	49
Z. mays	AJ297566	Chloroplast	0.511	48
S. rebaudiana	AJ429233	-	0.491	48
A. annua	AF182287	Chloroplast	0.531	46
M. piperita	AF116825	Chloroplast	0.514	51
G. biloba	AY494186	Chloroplast	0.513	55
T. cuspidata	AY575140	Chloroplast	0.567	55
Pinus sp. *	TC53644	-	0.488	57

Note:

Number after the scientific name refers to DXR sequence 1 and DXR sequence 2, respectively.

Dashes (-) indicate no prediction of cTP sequences to be localized in the chloroplast.

* The DXR sequence from *Pinus* sp. was obtained from the TIGR database (Quackenbush *et al.*, 2000).

2.5 Phylogenetic analysis of HbDXRs sequences

A better understanding of evolutionary relationship between both HbDXRs and other DXR proteins from different organisms in databases was obtained from a phylogenetic tree. Twenty-six DXR sequences (12 from plants, 3 from algae, 6 from cyanobacteria, 4 from eubacteria, and 1 from malaria parasite) were used for phylogenetic tree construction using the neighbor-joining method in the MEGA program (version 3.1) (Kumar *et al.*, 2004).

The neighbor-joining tree is shown in Fig. 3.21. Similar to the DXS tree, the topology of the DXR tree also reflects the taxonomic classification of the species with minor exceptions, when the DXR sequence from malaria parasite *P. falciparum* was used as the out group (Fig. 3.21). The tree can be organized into three main groups. The first group contained all eubacteria, such as *E. coli*, *H. influenzae*, *V. cholerae*, and *Z. mobilis* (Fig. 3.21). The second group contained all cyanobacteria, such as *Crocosphaera watsonii*, *N. punctiforme*, *Nostoc* sp., *S. leopoliensis*, and *T. elongatus* (Fig. 3.21). It was surprising to find that whether the red alga *C. merolae* was grouped with the cyaobacterial DXRs in this study (Fig. 3.21).

The third group contained all green plants, which were further separated into two groups: streptophytes and chlorophytes (*C. reinhardii* and *O. lucimarinus*) (Fig. 3.21). Again, streptophytes can be separated into two groups: gymnosperms and angiosperms, with a bootstrap value of 100% (Fig. 3.21). The gymnosperms included *Pinus* sp., *T. cuspidata*, and *G. biloba*, and the angiosperms included *L. esculentum*, *C. roseus*, *M. piperita*, *A. thaliana*, *Z. mays*, *A. annua*, and *S. rebaudiana* (Fig. 3.21). As expected, the two HbDXRs were grouped with other plant DXRs but they were more similar to the angiosperms than the gymnosperms. Within the angiosperms clade, both HbDXRs were more closely related to *L. esculentum* than to those of other plant DXRs (Fig. 3.21). In addition, the two HbDXRs formed a cluster together with 100% bootstrap value support (Fig. 3.21), suggesting that both of them might evolve from a recent gene duplication event.

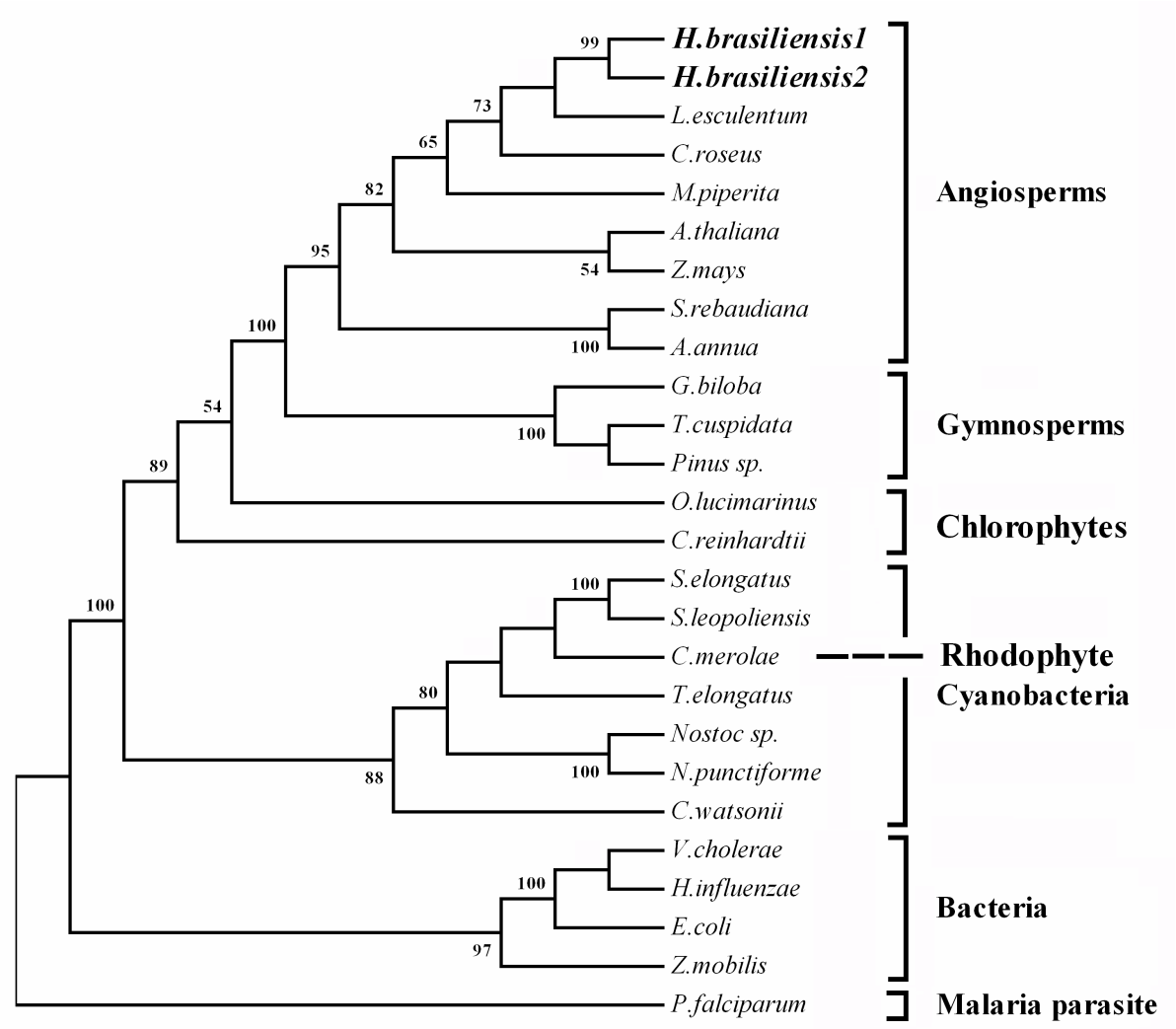


Fig. 3.21 Phylogenetic tree among DXR sequences from different taxa.

The tree was constructed by the MEGA program (version 3.1) (Kumar *et al.*, 2004) using the neighbor-joining method. Numbers represent bootstrap values support for 10000 replicates. The DXR sequence from the malaria parasite *P. falciparum* (NP_702530) was used as the out group. The DXR sequences from eubacteria: *H. influenzae* (NP_438967), *E. coli* (AB013300), *V. cholerae* (NP_231885), and *Z. mobilis* (AAD29659) and from cyanobacteria: *Nostoc* sp. (NP_488391), *N. punctiforme* (ZP_00111307), *S. leopoliensis* (YP_173208), *T. elongatus* (NP_681831), and *C. watsonii* (ZP_00515416) were obtained from the GenBank database. The DXR sequences from algae were from the JGI database and *C. merolae* Genome

Project (Matsuzaki *et al.*, 2004). For plant DXR sequences, please see Table 3.5 for more information.

2.6 *DXR* genes in *H. brasiliensis* belong to a small gene family

The number of *HbDXRs* genes in the *H. brasiliensis* genome was estimated by Southern blot analysis. Genomic DNA of *H. brasiliensis* was isolated from young leaves of mature plants, digested with different restriction enzymes, and run on 1.0 % agarose gel electrophoresis. After electrophoresis, the digested DNA fragments were transferred to Hybond-N^+ membrane and allowed to hybridize with a 0.7 kb *HbDXR1* fragment from the initial One-step RT-PCR clone as probe. The probe was labeled with fluorescein.

The result of Southern blot analysis is shown in Fig. 3.22. It was found that multiple hybridizing signals ranging from about 2 to 20 kb were detected in all restriction enzymes used with strong and weak intensities. Two strong hybridizing bands were observed when the genomic DNA was digested with *Bam*HI and *Eco*RV. Three strong signals were observed in *Sac*I digestion. Five and seven strong bands were observed in *Eco*RI and *Bgl*II digestion, respectively. Only one strong signal was detected when the genomic DNA was digested with *Xba*I and *Xho*I. In addition, several weakly hybridized fragments were also detected in all these samples. These results demonstrate that the *DXR* gene in *H. brasiliensis* is most likely to be encoded by a small gene family, with at least two members. However, it is not clear whether the weakly hybridized bands represent an additional *DXR* gene or genes encoding closely related proteins in the *H. brasiliensis* genome.

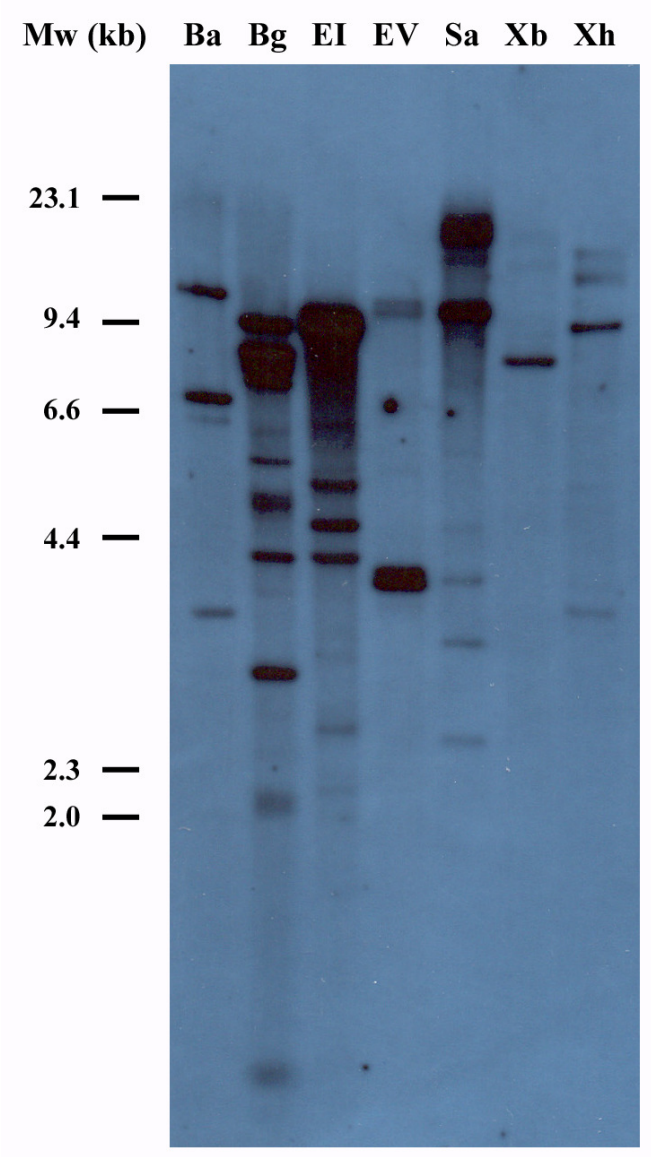


Fig. 3.22 Southern blot analysis of the *DXR* gene in *H. brasiliensis*.

The digested genomic DNA was transferred to a Hybond-N^+ nylon membrane and hybridized with a fluorescein-labeled *HbDXR1* probe. The position of Lambda DNA/*Hin*dIII molecular weight marker (Mw) is shown in kb on the left. Ba: *Bam*HI, Bg: *Bgl*II, EI: *Eco*RI, EV: *Eco*RV, Sa: *Sac*I, Xb: *Xba*I, and Xh: *Xho*I.

2.7 Secondary and 3D structures of HbDXRs

As another method of the prediction, secondary assignments of the deduced HbDXRs protein sequences were obtained using the SWISS-MODEL program (Schwede *et al.*, 2003). In this study, the recent crystal structure of *E. coli* DXR, determined by the single isomorphous replacement with anomalous scattering (Reuter *et al.*, 2002), obtained from the PDB database (code 1K5H), was used as reference for a prediction of the 3D structure of HbDXRs. The average sequence similarities between HbDXRs and their *E. coli* homologue were about 48%. The cTP sequences were not used in the predictions.

The predicted secondary structures of HbDXRs are shown in Fig. 3.23, and their predicted 3D models are displayed in Fig. 3.24, Fig. 3.25 and Fig. 3.26, respectively. The Swiss-Pdb Viewer program (Guex and Peitsch, 1997) was used to visualize the models. According to the models, the HbDXR monomer can be divided into three domains: an NH_2-terminal NADPH-binding domain, a central catalytic domain, and a COOH-terminal domain, as previously assigned in the *E. coli* DXR (Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; Yajima *et al.*, 2004; MacSweeney *et al.*, 2005). The NH_2-terminal domain of both HbDXRs (residues 74-223) was predicted to consist of 7 α-helices and 7 β-strands (Fig. 3.23 and Fig. 3.24). This domain exhibited an α/β topology, which is a typical NADPH-binding site. The central catalytic domain (residues 224-348) consisted of 6 α-helices and 4 β-strands, and the COOH-terminal domains (residues 378-471) consisted of 3 α-helices (Fig. 3.23 and Fig. 3.24). The COOH-terminal domain was connected to the central catalytic domain by an extended loop, which is further assigned at positions 349-377 in both HbDXRs (Fig. 3.23 and Fig. 3.24). Residues corresponding to the flexible loop, which is function as a lid of the active site of *E. coli* DXR, can also be assigned. They correspond to amino acids 249-279 in both HbDXRs (Fig. 3.23 and Fig. 3.24). The overall folding patterns of the 3D models of HbDXRs and upon superimposition to the X-ray structure of *E. coli* DXR are shown in Fig. 3.25 and Fig. 3.26, respectively, and the localization of amino acids that form the active site geometry of both HbDXRs models is shown in Fig. 3.27. These revealed very similar spatial arrangements of important domains and residues between plants and bacterial DXR.

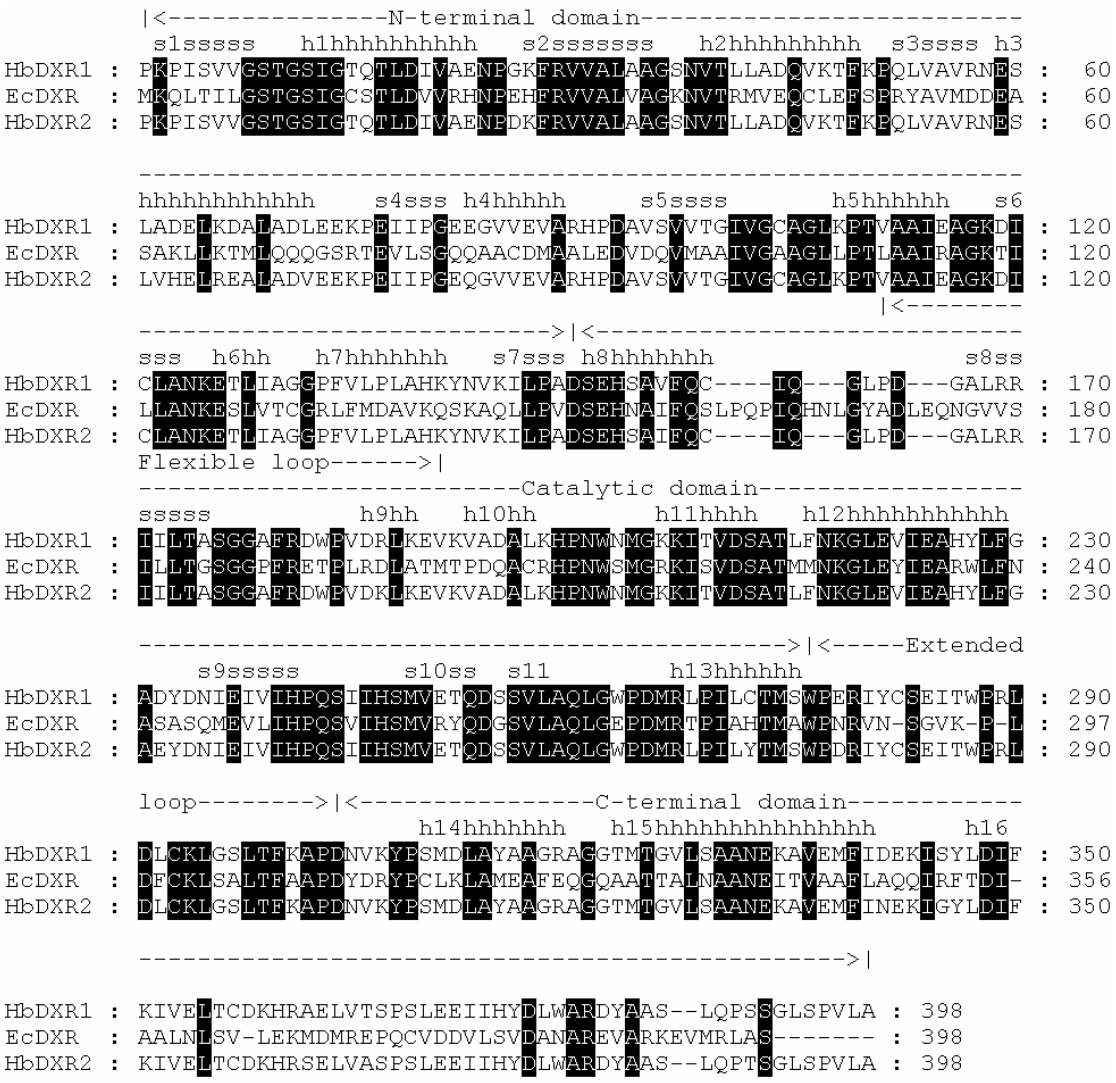


Fig. 3.23 Alignment of the predicted secondary structures of HbDXRs.

Secondary structures of both HbDXRs were predicted using the SWISS-MODEL program (Schwede *et al.*, 2003) and they are given above the aligned sequences. Alpha-helices are indicated by "h" and β-strands by "s" and numbered from the NH_2-terminus. Domains are assigned according to the *E. coli* DXR structure (PDB code 1K5H) (Reuter *et al.*, 2002).

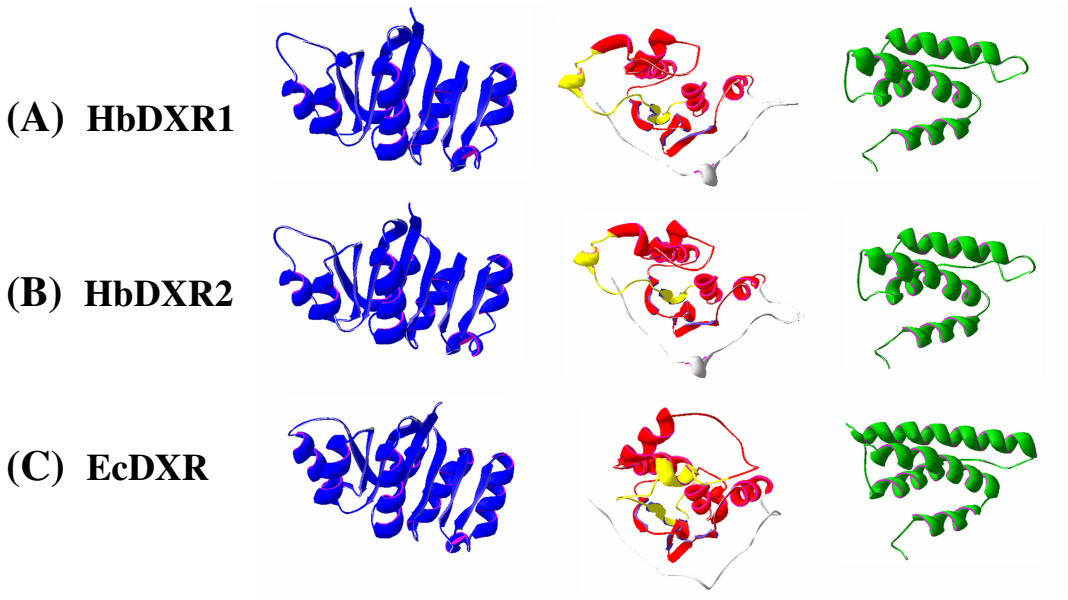


Fig. 3.24 The predicted 3D structures of domains in HbDXRs.

The 3D structure models of both HbDXRs were predicted using the SWISS-MODEL program (Schwede *et al.*, 2003) and displayed with the Swiss-Pdb Viewer (Guex and Peitsch, 1997). (A) HbDXR1. (B) HbDXR2. (C) *E. coli* DXR. The NH_2-terminal, a catalytic, and the COOH-terminal domains are colored in blue, red and green, respectively. Regions of extended and flexible loops are colored in grey and yellow, respectively.

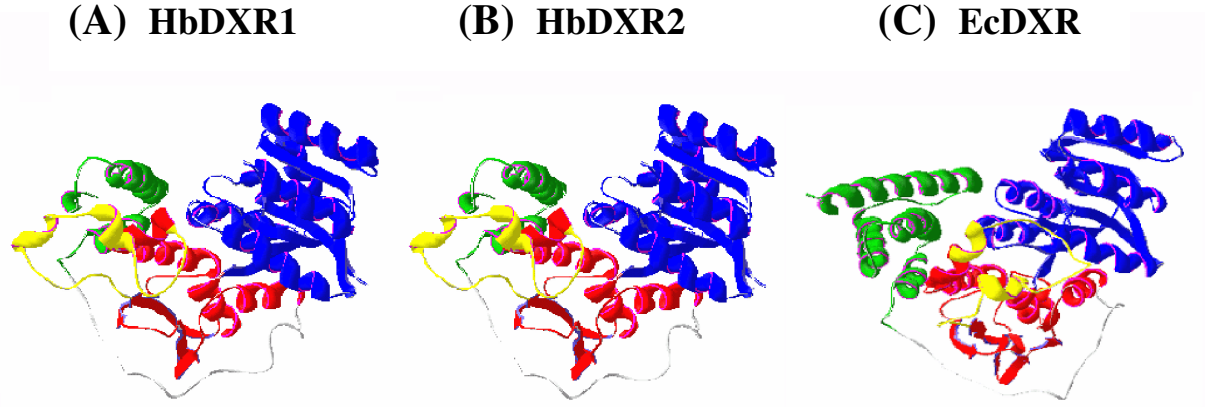


Fig. 3.25 The overall predicted 3D structures of HbDXRs.

(A) HbDXR1. (B) HbDXR2. (C) *E. coli* DXR. Domains are colored in a similar fashion as described in Fig. 3.24.

Fig. 3.26 Superimposition of the predicted 3D structures of HbDXRs.

(A) Superimposition of HbDXR1 (cyan) and *E. coli* DXR (brown). (B) Superimposition of HbDXR2 (gold) and *E. coli* DXR. (C) Superimposition of HbDXR1 and HbDXR2.

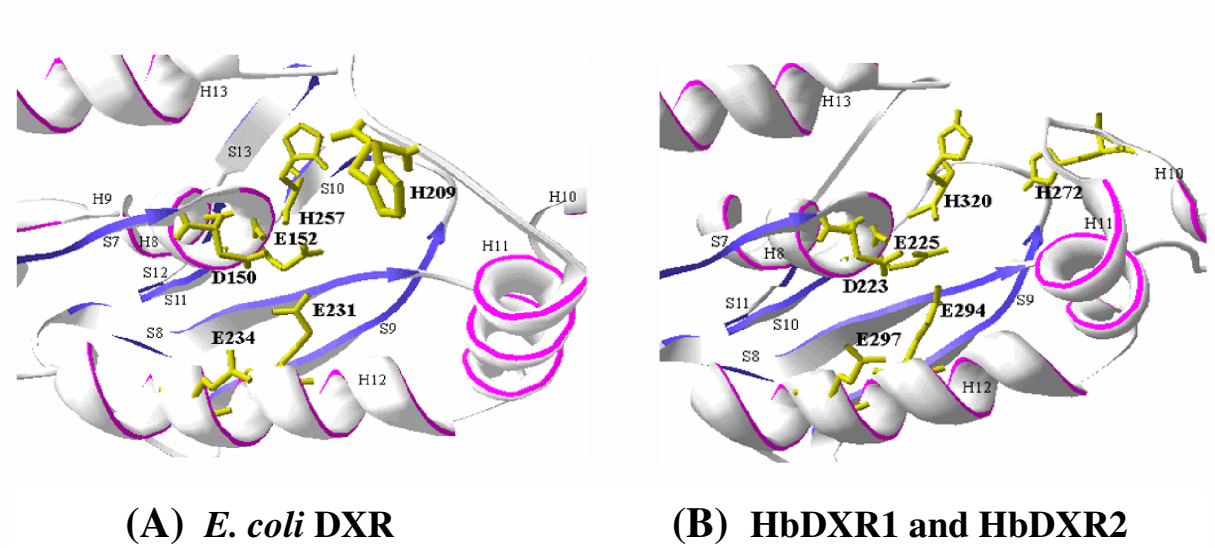


Fig. 3.27 3D structures representation of active site architectures of DXRs.

(A) Active site architecture of *E. coli* DXR. (B) Active site architecture of HbDXRs. Numbers of α-helices and β-strands are identical to Fig. 3.23. Three conserved acidic residues (aspartic and glutamic acids) and two invariant histidine residues are shown.

3. Expression analysis of *HbDXSs* and *HbDXRs* mRNAs

The expression pattern of *HbDXSs* and *HbDXRs* mRNAs in various tissues of *H. brasiliensis* including in the latex of different cultivars and the effect of ethephon treatment on their expression was analyzed by the semi-quantitative RT-PCR method. RT-PCR is a sensitive amplification procedure that has been used to detect the expression of genes in various organisms. In general, the quantification was obtained by measuring the abundance of gene of interest relative to that of the internal control gene. The internal control used in this study was the *18S rRNA* gene.

3.1 Nucleotide sequence of *18S rRNA* gene

Since there is no sequence information about the *18S rRNA* gene in *H. brasiliensis*, this gene was cloned from *H. brasiliensis* using gene-specific primers. The gene-specific primers for the *18S rRNA* gene were designed based on the nucleotide sequences of *18S rRNA* in different plant species that were available in the GeneBank database. The forward primer was 5´-CAAAGCAAGCCTACGCTCTG-3´ and the reverse primer was 5´-CGCTCCACCAACTAAGAACG-3´. Using these primers in One-step RT-PCR amplification, a cDNA fragment of about 0.5 kb was amplified. The cDNA was then purified, cloned into pDrive vector (Fig. 3.28, lane 1), and sequenced.

Analysis of the nucleotide sequence using the BLAST program (Altschul *et al.*, 1997) revealed that the cDNA showed 100% identity to the *18S rRNA* gene from various plants belonging to Euphorbiaceae family, such as *Oldfieldia dactylophylla* (AB233611), *Hura crepitans* (AB233574), *Excoecaria cochinchinensis* (AB233573), *Schinziophyton rautanenii* (AB233570), and *Dalechampia spathulata* (AB233551). The result suggested that the cloned cDNA could encode a partial *18S rRNA* gene in *H. brasiliensis*. The nucleotide sequence of the *18S rRNA* cDNA has been submitted in GenBank database under the accession number AY496880 (Fig. 3.29).

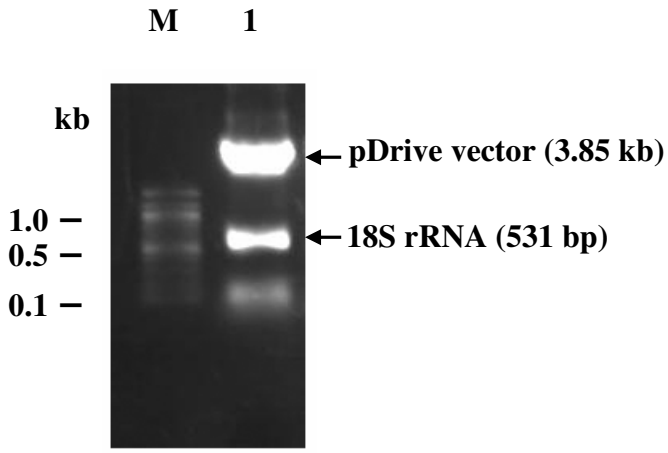


Fig. 3.28 Agarose gel analysis of the *18S rRNA* cDNA clone.

pDrive vector containing a 531 bp of *18S rRNA* cDNA (lane 1) was digested with restriction enzyme *Eco*RI at 37ºC for 1 h and separated by electrophoresis on a 1.0% agarose gel containing 0.1 mg/ml of ethidium bromide. Lane M is a 100 bp DNA Ladder (New England Biolabs).

```
  1  CAAAGCAAGC CTACGCTCTG GATACATTAG CATGGGATAA CATCATAGGA
 51  TTTCGGTCCT ATTGTGTTGG CCTTCGGGAT CGGAGTAATG ATTAACAGGG
101  ACAGTCGGGG GCATTCGTAT TTCATAGTCA GAGGTGAAAT TCTTGGATTT
151  ATGAAAGACG AACAACTGCG AAAGCATTTG CCAAGGATGT TTTCATTAAT
201  CAAGAACGAA AGTTGGGGGC TCGAAGACGA TCAGATACCG TCCTAGTCTC
251  AACCATAAAC GATGCCGACC AGGGATCGGC GGATGTTGCT TTTAGGACTC
301  CGCCGGCACC TTATGAGAAA TCAAAGTCTT TGGGTTCCGG GGGGAGTATG
351  GTCGCAAGGC TGAAACTTAA AGGAATTGAC GGAAGGGCAC CACCAGGAGT
401  GGAGCCTGCG GCTTAATTTG ACTCAACACG GGGAAACTTA CCAGGTCCAG
451  ACATAGTAAG GATTGACAGA CTGAGAGCTC TTTCTTGATT CTATGGGTGG
501  TGGTGCATGG CCGTTCTTAG TTGGTGGAGC G
```

Fig. 3.29 Nucleotide sequence of the partial *18S rRNA* gene from *H. brasiliensis*.

3.2 Expression pattern of *HbDXSs* in various tissues

The expression of the *HbDXS1* and *HbDXS2* genes in various tissues was analyzed in both seedlings and mature plants. The gene-specific primers used for *HbDXS1* and *HbDXS2* genes were expected to amplify the corresponding cDNA of 468 bp and 416 bp in lengths, respectively. The results showed that both *HbDXS1* and *HbDXS2* genes were found to be differentially expressed in different tissues and in different developmental stages of *H. brasiliensis* (Fig. 3.30). In seedlings, the transcripts of *HbDXS1* were found to be expressed higher in stems and leaves than in roots (Fig. 3.30A), while in mature plants, the transcripts of *HbDXS1* were found to be expressed higher in leaves than in flowers and in leaves that were naturally infected with the fungus *O. heveae* (Fig. 3.30B). No expression of *HbDXS1* mRNA was detected in the latex sample (Fig. 3.30B).

In contrast, the expression of *HbDXS2* transcripts was found in all tissues examined, but at different levels. In seedlings, the transcripts of *HbDXS2* were found to be expressed higher in roots than in leaves and stems (Fig. 3.30A). In mature plants, the *HbDXS2* was found to be expressed higher in leaves than in latex and flowers (Fig. 3.30B). Leaves that were naturally infected with the fungus *O. heveae* also showed a low level of *HbDXS2* mRNA when compared to normal leaves (Fig. 3.30B).

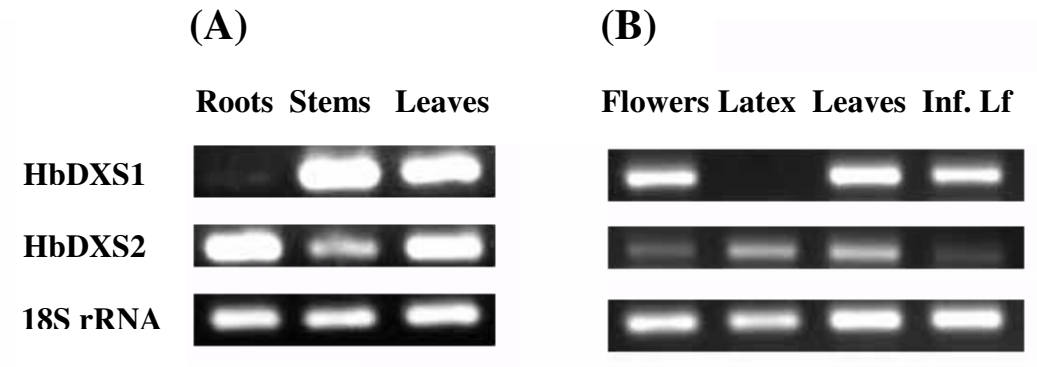


Fig. 3.30 Expression of *HbDXS* mRNAs in various tissues.

(A) Seedlings. (B) Mature plants. The expression of *18S rRNA* gene was used as an internal control. Inf. Lf: leaves from mature plants infected with fungal *O. heveae*.

3.3 Expression pattern of *HbDXR s* in various tissues

Since the nucleotide sequences of the two *HbDXR* genes were highly similar (data not shown), gene-specific primers that were used for semi-quantitative analysis were carefully designed and the specificity of these PCR primers was tested using the full-length cDNA clones of each *HbDXR* gene as templates. As shown in Fig. 3.31, the major cDNA band of about 252 bp in length was amplified when the gene-specific primers and the full-length cDNA of *HbDXR1* were used (Fig. 3.31, lane 1), while the cDNA band of about 245 bp was amplified when the gene-specific primers and the full-length cDNA of *HbDXR2* were used (Fig. 3.31, lane 3). These results suggested that the primers designed were specific to each *HbDXR* gene, although a tiny amount of reaction products were observed with *HbDXR1* primers and *HbDXR2* template and with the *HbDXR2* primers and *HbDXR1* template (Fig. 3.31, lane 2 and lane 4, respectively).

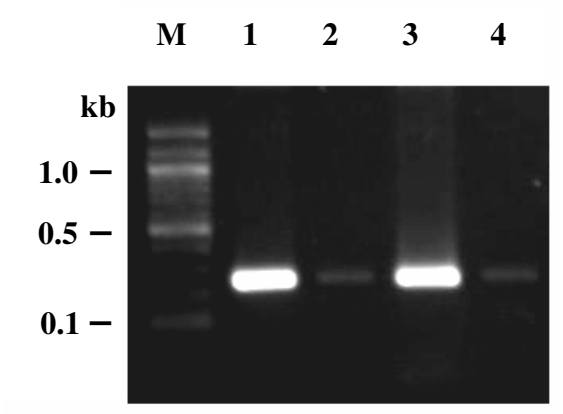


Fig. 3.31 Specificity analysis of the *HbDXR1* and *HbDXR2* primers.

Plasmid DNAs (100 ng/lane) containing either the full-length cDNA of *HbDXR1* (lane 1 and 2) or *HbDXR2* (lane 3 and 4) were used as templates in PCR reactions with different primers. The primers used in lane 1 and lane 4 were HbDXR1f/r, which were specified to the *HbDXR1* gene, and those for lane 2 and 3 were HbDXR2f/r, which were specified to the *HbDXR2* gene, respectively. PCR products were analyzed on a 1.0% agarose gel electrophoresis containing 0.1 mg/ml of ethidium bromide. Lane M is a 100 bp DNA Ladder (New England Biolabs).

Analysis of the expression of both *HbDXR1* and *HbDXR2* mRNAs in different tissues and in different developmental stages of *H. brasiliensis* is shown in Fig. 3.32. Both *HbDXR1* and *HbDXR2* mRNAs showed differential expression in all tissues analyzed. The level of *HbDXR1* transcripts was found to be higher in stems and leaves than in roots, while the level of *HbDXR2* transcripts was found to be higher in roots than in other tissues of seedlings (Fig. 3.32A). In mature plants, both *HbDXR1* and *HbDXR2* mRNAs were found to be expressed higher in leaves than in flowers and latex (Fig. 3.32B). A distinct expression pattern between the *HbDXR1* and *HbDXR2* was found in the latex sample, in which the *HbDXR2* transcripts were more abundantly expressed in latex than the *HbDXR1* transcripts (Fig. 3.32B). In addition, the *HbDXR2* transcripts showed lower expression in fungus-infected leaves than in normal leaves, while the level of *HbDXR1* transcripts was not affected by fungal infection (Fig. 3.32B).

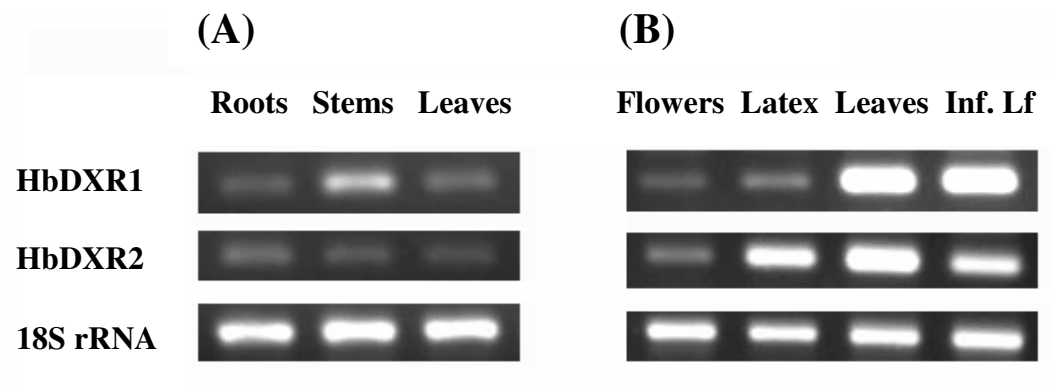


Fig. 3.32 Expressiom of *HbDXR* mRNAs in various tissues.

(A) Seedlings. (B) Mature plants. The expression of *18S rRNA* gene was used as internal control. Inf. Lf: leaves from mature plants infected with fungal *O. heveae*.

3.4 Expression of *HbDXS2* and *HbDXR2* mRNAs in different cultivars

It has been known that *H. brasiliensis* clone RRIM 600 produces more rubber yields than wild type plants (Suwanmanee *et al.*, 2004). The latex yield produced from clone RRIM 600 was about 105.5±4.9 ml, while the latex yield from wild type plants was about 40.8±2.1 ml (Fig. 3.33A). Dry rubber content in clone RRIM 600 was calculated to be 39.0±2.3 g tree^{-1} tapping^{-1} and in wild type plants was 21.0±1.9 g tree^{-1} tapping^{-1}, respectively (Fig. 3.33B). The results showed that *H. brasiliensis* clone RRIM 600 produced approximately two-fold higher rubber yield than the wild type (Fig. 3.33B). Therefore, the clone RRIM 600 was referred as a high-latex yielding clone and the wild type plants was a low-latex yielding clone.

Since the expression of only *HbDXS2* gene was found to be expressed in latex but not the *HbDXS1* gene (Fig. 3.30B), and the transcripts of *HbDXR2* were more abundant in latex than the transcripts of *HbDXR1* (Fig. 3.32B), the expression of both *HbDXS2* and *HbDXR2* mRNAs in relation to the latex yield was examined in these two *H. brasiliensis* clones by RT-PCR analysis. The results are shown in Fig. 3.33C. The levels of *HbDXS2* transcripts showed no significant difference between clone RRIM 600 and wild type plants, although some of the wild type plants showed a very low expression levels of *HbDXS2* gene. In contrast to *HbDXS2*, the expression level of *HbDXR2* transcripts was found to be higher in clone RRIM 600 than in the wild type plants (Fig. 3.33C). The results suggest that the expression of the *HbDXR2* gene might contribute in some way to the biosynthesis of natural rubber.

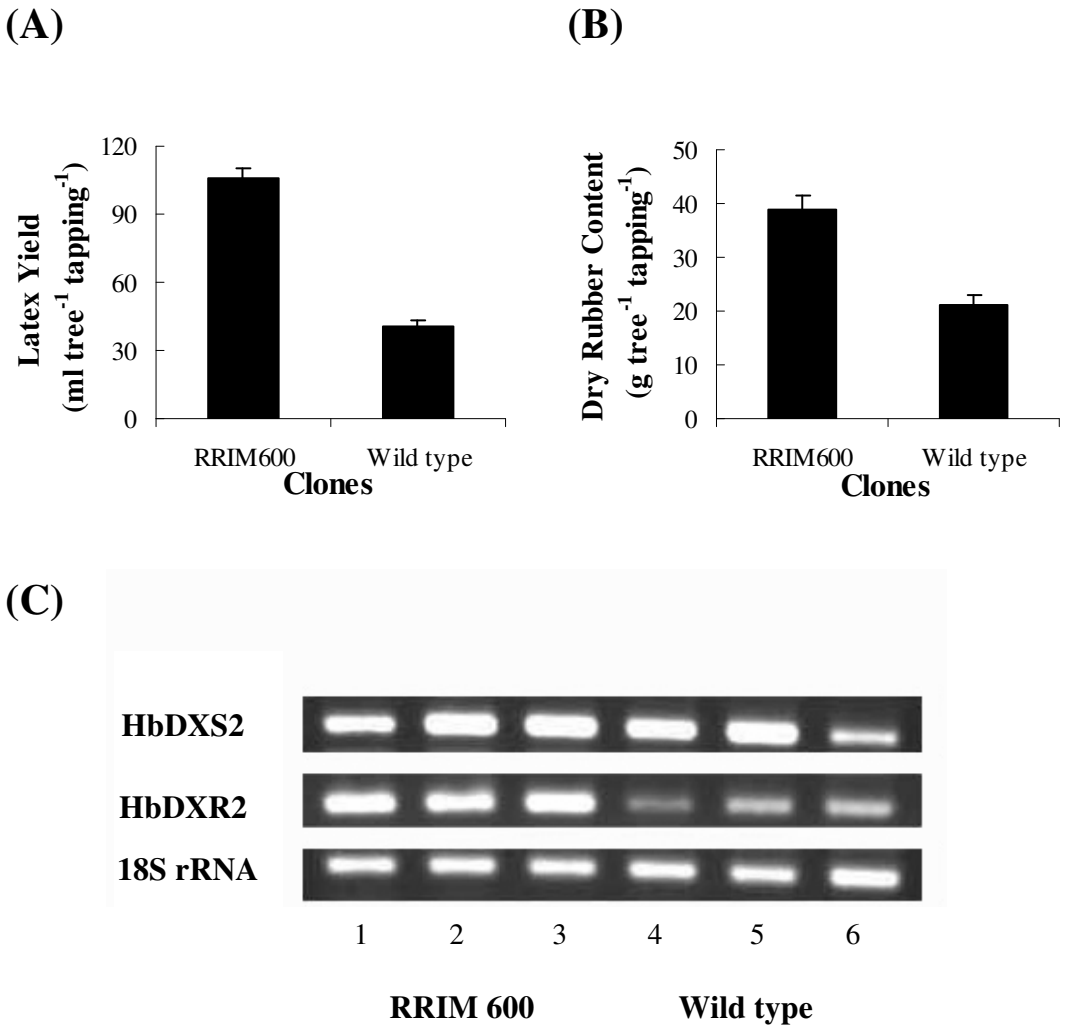


Fig. 3.33 Comparison of the expression levels of *HbDXS2* and *HbDXR2* genes in different clones of *H. brasiliensis*.

(A) Latex yield. (B) Dry rubber content. (C) Electrophoresis of *HbDXS2* and *HbDXR2* RT-PCR products. The expression level of the *18S rRNA* gene was used as internal control. Data points represent the mean values ± S.E. of three replicates.

3.5 Effect of ethephon on the expression of the *HbDXS2* and *HbDXR2* genes

Ethephon (2-chloroethylphosphonic acid) is widely used to enhance the rubber production in *H. brasiliensis* tree (Coupé and Chrestin, 1989; Pujade-Renaud *et al*., 1994). In this study, when 2.5% ethephon was applied to the bark area underneath the tapping cut of *H. brasiliensis* tree, it was found that the volume of exported latex was around three-fold higher in a treated *H. brasiliensis* tree than in an untreated tree (Fig. 3.34A). The volume of the exported latex increased from 32. 7±2.1 ml in day 0 (the day before treatment, and also served as the control) to 100.3±5.4 ml within 24 h (day 1 after treatment), and still increased up to day 5 of the experiment (235.7±4.4 ml) (Fig. 3.34A). The rubber content was also increased from 18.9±1.8 g tree^{-1} tapping^{-1} in day 0 to 50.6±3.9 g tree^{-1} tapping^{-1} after 24 h treatment and also still increased to 109.5±8.5 g tree^{-1} tapping^{-1} in day 5 of the ethephon treatment (Fig. 3.34B).

In order to study the effect of ethephon on the expression of *HbDXS2* and *HbDXR2* genes, total RNA was extracted from the latex obtained from *H. brasiliensis* trees, reverse-transcribed, and subjected to semi-quantitative RT-PCR analysis. The results showed that both *HbDXS2* and *HbDXR2* genes were found to be transiently induced after 24 h treatment with ethephon, and then they were decreased (Fig. 3.34C). The results suggest that the induced accumulation of natural rubber was preceded by ethylene, which is generated in *vivo* from ethephon, but was not by increased either in the accumulation of *HbDXS2* or *HbDXR2* transcripts.

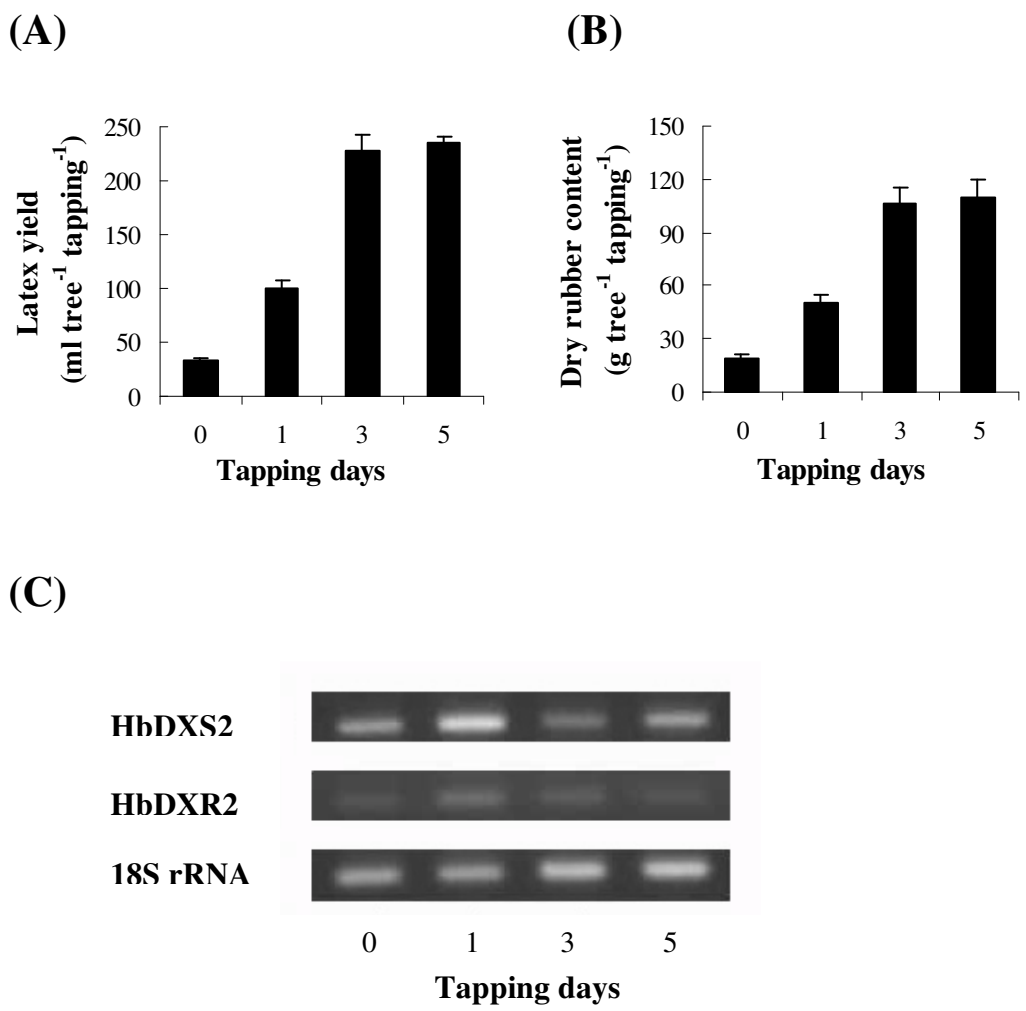


Fig. 3.34 Effect of ethylene treatment on the expression levels of *HbDXS2* and *HbDXR2* genes

Latex was collected from the *H. brasiliensis* trees before treatment (day 0) and after treatment with 2.5% ethephon (day 1, day 3, and day 5). (A) Latex yield. (B) Dry rubber content. (C) Electrophoresis of *HbDXS2* and *HbDXR2* RT-PCR products. The expression level of the *18S rRNA* gene was used as an internal control. Data points represent the mean values ± S.E. of three replicates.

4. Expression of recombinant proteins and Western blot analysis

In order to analyze whether that the cloned cDNAs can encode the interested proteins, the prokaryotic expression system in *E. coli* was used. The putative mature sequences, excluding the cTP sequences of both *HbDXS1* and *HbDXR1*, were generated by PCR and fused with the expression vectors. The commercial vectors used in this study contain the IPTG-inducible promoter and also encode hexa-histidyl (His-Tag) leader sequence at the NH_2-terminus. The presence of His-Tag helps a facilitating purification of fusion proteins by nickel-based affinity chromatography.

4.1 Construction, expression and purification of recombinant proteins

The cDNA fragments of the predicted mature sequences of HbDXS1 or HbDXR1 were generated by PCR using primers for each gene (Table 2.6) and plasmid DNA containing either the full-length sequences of *HbDXS1* or *HbDXR1* as template. The PCR products yielded single DNA fragments of about 1.9 and 1.2 kb for the *HbDXS1* and *HbDXR1* cDNAs, respectively. Then, the cDNAs were purified, digested with restriction enzymes, and cloned into the expression vectors, which are pre-digested with the same restriction enzymes. The *HbDXS1* cDNA was cloned into the *Eco*RI/*Xho*I sites of the pET-28a vector, yielding a construct named pET-28a/HbDXS1, and the *HbDXR1* cDNA was cloned into the *Sph*I/*Pst*I sites of the pQE30 vector to give a construct named pQE30/HbDXR1. Both constructions were verified by restriction enzymes analysis (Fig. 3.35) and sequenced to prove that the cloned cDNAs were inserted into the expression vectors with a correct orientation.

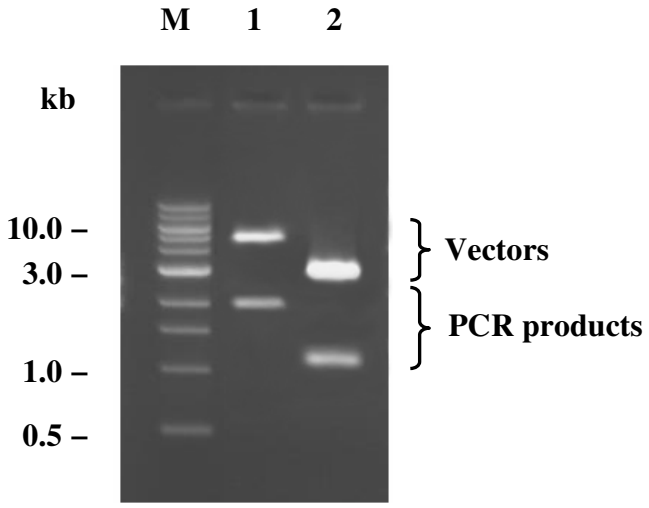


Fig. 3.35 Identification of recombinant plasmid DNAs construction.

Plasmid DNAs were digested with *Eco*RI and *Xho*I (for pET-28a/HbDXS1, lane 1) or with *Sph*I and *Pst*I (for pQE30/HbDXR1, lane 2) at 37ºC and separated by electrophoresis on a 1.0% agarose gel containing 0.1 mg/ml of ethidium bromide. Lane M is a 1 kb DNA Ladder (New England Biolabs).

The confirmed constructs of both pET-28a/HbDXS1 and pQE30/-HbDXR1 were then transformed into *E. coli* strain BL21 (DE3) and XL1-Blue MRF´, respectively. The small-scale cultures of both cells were subjected to IPTG induction to identify the capacity of expression at 37°C. By SDS-PAGE analysis, both recombinant HbDXS1 and HbDXR1 proteins were not expressed as soluble proteins (Fig. 3.36A). To improve proteins solubility, various growth conditions, for example, reduction of temperature or IPTG concentration were used, but the expressed proteins were still not soluble (data not shown). Thereafter, insoluble aggregates of inclusion bodies fraction were solubilized in a buffer containing urea and recombinant proteins were further purified under denaturating conditions using the Ni-NTA column according to the manufacturer' protocols (Qiagen, Germany), as described in section 17.2.

The purified recombinant proteins were resolved on SDS-PAGE (Fig. 3.36). The major protein band of *E. coli* cells harboring the plasmid pET-28a/HbDXS1 was about 70 kDa (Fig. 3.36B, lane 1), and that for *E. coli* cells transformed with the construct pQE30/HbDXR1 was about 40 kDa on SDS-gel (Fig. 3.36B, lane 2). These molecular sizes were almost similar to their molecular weights predicted from the nucleotide sequences (71 kDa for HbDXS1 and 46 kDa for HbDXR1). In an attempt to test the activity of purified proteins by enzymatic assay, only recombinant HbDXR1 was refolded in a re-denaturating buffer containing glutathione and tested for activity. The result showed that the recovered recombinant HbDXR1 protein did not exhibit the DXR activity (data not shown).

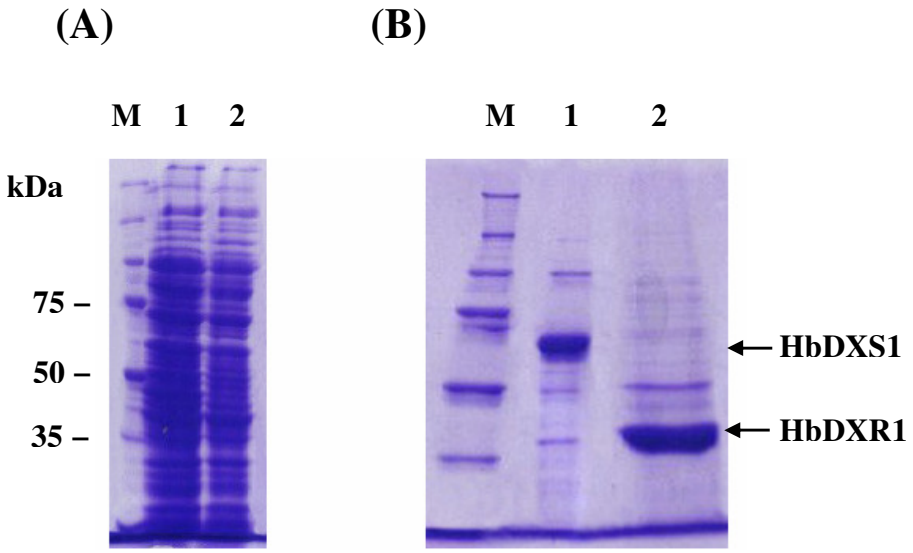


Fig. 3.36 SDS-PAGE analysis of recombinant proteins.

(A) Soluble fraction of *E. coli* harboring pET-28a/HbDXS1 vector (lane 1) and pQE30/HbDXR1 vector (lane 2). (B) The purified recombinant proteins. Lane 1 is a recombinant HbDXS1 protein. Lane 2 is a recombinant HbDXR1 protein. Lane M is a protein marker (Promega).

4.2 Production of polyclonal antibodies and Western blot analysis

To produce antibodies against HbDXS1 and HbDXR1, the proteins bands corresponding to HbDXS1 and HbDXR1 were recovered from SDS-gels and used as immunogens. Sera from third bleed of immunized-rabbit were used as source for polyclonal antibodies. The specificity of both HbDXS1 and HbDXR1 antisera to recognize the DXS and DXR proteins, respectively, was investigated using the western blot analysis. The results are shown in Fig. 3.37 and Fig. 3.38. The concentration of both HbDXS1 and HbDXR1 antisera used in this study was 1:500 in TBST and the pre-immunized rabbit sera were used as negative controls (Fig. 3.37B). As shown in Fig. 3.37A, the purified recombinant HbDXR1 (lane 1) and HbDXS1 (lane 1) proteins were easily detected using their antisera, although some degradation products were found. The detected sizes of these proteins were almost similar to those predicted from nucleotide sequences and on the SDS-gel (Fig. 3.36B).

In addition, the HbDXR1 antiserum was able to recognize the putative DXR protein in *H. brasiliensis* (Fig. 3.37A, lane 3), while the HbDXS1 antiserum detected proteins having size lower than that the purified recombinant HbDXS1 protein (Fig. 3.37A, lane 3). Moreover, the level of the putative DXR protein in leaves of *H. brasiliensis* collected at different time points (daytime and nighttime) gave similar results (Fig. 3.38A, lane 3 and 4, respectively). In addition to *H. brasiliensis*, the HbDXR1 antiserum can cross-react with the putative DXR protein from other plants, such as *A. thaliana* (Fig. 3.37A, lane 2 or Fig. 3.38A, lane 2) and *P. trichocarpa* (Fig. 3.38A, lane 1), indicating that the HbDXR1 antiserum can serve as a good tool to characterize a DXR protein in plants.

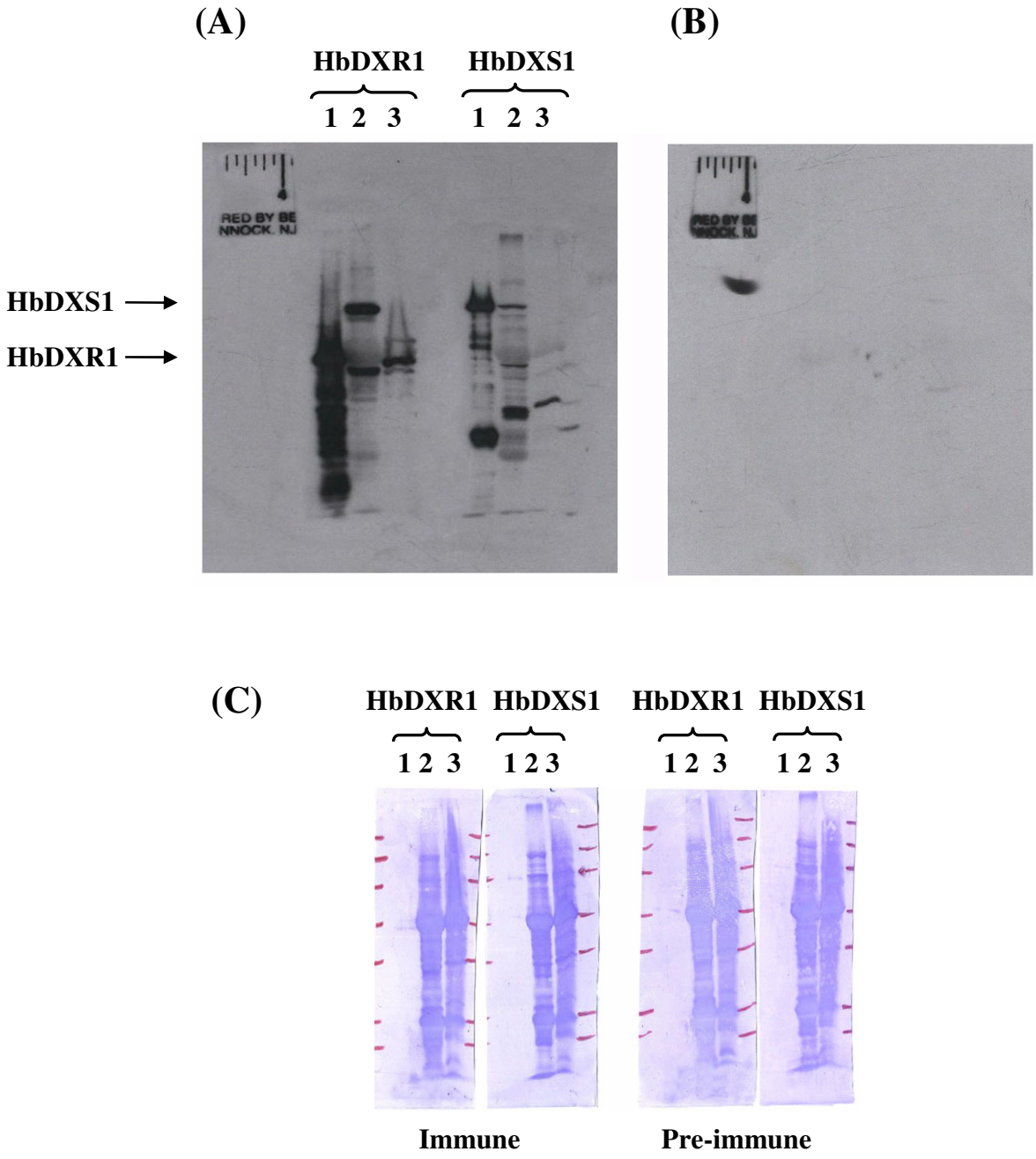


Fig. 3.37 Immunoreaction of HbDXS1 and HbDXR1 antisera.

(A). Immunoreaction signals using HbDXR1 (lane 1-3) and HbDXS1 (lane 1-3) antisera. (B). Immunoreaction signals using pre-immune sera. (C). Coomassie Blue-stained membranes. Lane 1 is the putative purified

recombinant HbDXR1 and HbDXS1 proteins. Lane 2 is crude leaf extracts from *A. thaliana*. Lanes 3 is crude leaf extracts from *H. brasiliensis*. Exposure time is 1 min. Arrows indicate the putative DXR and DXS proteins.

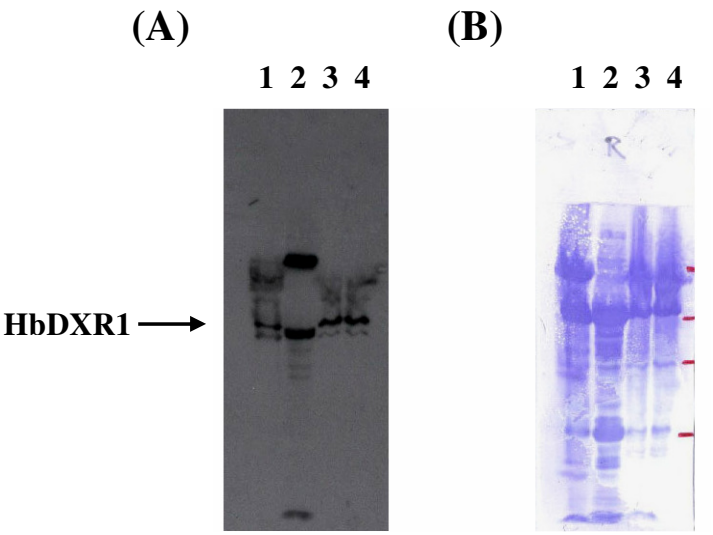


Fig. 3.38 Immunoreaction of HbDXR1 antiserum with protein extracts from leaves of different plants.

(A). Immunoreaction signals. (B) Coomassie Blue-stained membranes. Lane 1 is crude leaf extracts from *P. trichocarpa*. Lane 2 is crude leaf extracts from *A. thaliana*. Lanes 3 and 4 are crude leaf extracts from *H. brasiliensis* collected at daytime and nighttime, respectively. Exposure time is 1 min. Arrow indicates putative DXR protein.

CHAPTER 4

DISCUSSION

1. Molecular cloning and sequence analysis of cDNA encoding HbDXS

1.1 Cloning of two *HbDXSs* genes in *H. brasiliensis*

PCR amplification of leaf and latex-derived *HbDXS* cDNAs was done with degenerate primers designed based on the highly conserved sequences of various plant DXSs gave two PCR products that shared high sequence identity with other known DXSs (>68% identity). However, these two PCR products shared only 60% identity to each other, suggesting that they could both encode DXS, but were derived from different genes. Therefore, the *DXS* cDNA from leaves was named *HbDXS1* and that from the latex was *HbDXS2*. Then, the missing sequences of both *HbDXSs* at 5′ and 3′ ends were identified by RLM-RACE method, and full-length ORF of each gene was amplified by DNA polymerase with proof reading ability to confirm the integrity and identity of the sequences. The nucleotide and deduced amino acid sequences of both *HbDXS1* and *HbDXS2* cDNAs are shown in Fig. 3.2 and Fig. 3.4, respectively. The lengths of the nucleotide and amino acid sequences of both *HbDXS* cDNAs are in good agreement with the previously reported size of other plant DXSs (Table 4.1).

Recently, Ko *et al.* (2003) reported the presence of *DXS* mRNA in latex of *H. brasiliesis*. The sequence (HbAFLP01D08) contained 204 nucleotides encoding 28 amino acids and showed 94% identity with the *HbDXS2* sequence. Their finding together with the present study indicates that multiple forms of *DXS* are present in *H. brasiliesis*, and at least two full-length isogenes have been cloned. In agreement with these results, several plants such as *M. truncatula* (Walter *et al.*, 2002), *E. guineensis* (Khemvong and Suvachittanont, 2005), and *G. biloba* (Gong *et al.*, 2006 and Kim *et al.*, 2006) contain two cDNAs corresponding to the *DXS* gene, whereas *O. sativa* (Kim *et al.*, 2005) and *Picea abies* (Phillips *et al*., 2007) contain three members of the *DXS* gene. In addition, Han *et al.* (2003) proposed that the *M. citrifolia* may contain up to four *DXS* genes, based on the results of Southern blot analysis, however, only one *DXS* gene has been analyzed in this plant.

Table 4.1 Properties of the predicted *HbDXSs* proteins and some representative plant DXSs.

	ORF (bp)	Residues	Mw (kDa)	pI	References
HbDXS1	2163	720	78	7.66	This study
HbDXS2	2136	711	76	7.34	This study
A. thaliana	2154	717	77	6.86	Mandel *et al.*, 1996
C. annum	2160	719	78	6.73	Bouvier *et al.*, 1998
C. roseus	2151	716	77	6.93	Chahed *et al.*, 2000
L. esculentum	2160	719	78	6.32	Lois *et al.*, 2000
M. truncatula1	2154	717	73	6.45	Walter *et al.*, 2002
M. truncatula2	2136	711	77	6.96	Walter *et al.*, 2002
S. rebaudiana	2148	715	77	7.36	Totté *et al.*, 2003
P. montana	2154	717	78	7.36	Sharkey *et al.*, 2005

Data mining in several plant EST databases also revealed the presence of multigenic sequences of *DXS* gene in various plant species. For example, two *DXS* genes are identified in *P. trichocarpa* (fgenesh4_pg.C_LG_VIII001771 and estExt_Genewise1_v1.C_LG_XVIII1471) (Sterky *et al.*, 2004), and four *DXS* genes are found in *Physcomitrella patens* at the JGI database (e_gw1.139.67.1, estExt_gwp_gw1.C_110156, estExt_fgenesh1_pm.C_220042, and fgenesh1_pg.scaffold_ 42000033). Additional partial *DXS* sequences from *L. esculentum* (TC94650, TC 85474, and TC85473), *Solanum tuberosum* (TC16033 and TC17144), and *Glycine max* (TC69708, TC62913, and TC63865) have also been identified from the TIGR EST databases (Quackenbush *et al.*, 2000) as previously denoted by Walter *et al.* (2002). To date, an additional *DXS* sequence from *Z. mays* is submitted in the GenBank database (EF507248). Further analysis of the nucleotide sequences revealed that these *DXS* cDNAs could be derived from different independent clones, since their sequences, especially in the coding regions, are quite different (data not shown). The completion sequence of *A. thaliana* genome also reveals the presence of three *DXS* genes in its genome (Rodríguez-Concepción and Boronat, 2002), but only one gene was proven to function as DXS genetically and biochemically (Mandel *et al.*, 1996; Estévez *et al.*, 2000; Estévez *et al.*, 2001). The precise role of these paralogous genes in isoprenoid biosynthesis of *A.thaliana* is not clear and needs to be confirmed in the future. In addition to higher plants, multiple copies of *DXS* genes have been identified in many bacteria such

as *Rhodobacter capsulatus* (Hahn *et al.*, 2001) and *Streptomyces* sp. (Li *et al.,* 2001). Taken together, the cloning of two *HbDXSs* cDNAs in this study and above data strongly supports the existence of multigenic family of the *DXS* gene in *H. brasiliensis*. Therefore, it can be concluded that the occurrence of multiple *DXSs* in higher plants is a widespread phenomenon.

1.2 Comparison of the deduced HbDXSs amino acid sequences

A comparison of the deduced amino acid sequences of both HbDXSs along with those of other DXS sequences from different organisms in the databases using the BLAST program revealed that the two HbDXSs from *H. brasiliensis* shared highest sequence similarity with those from higher plants (Table 3.1). In addition, moderate and lower identities were observed with those from cyanobacteria, bacteria, and malaria parasite (30-65% identity). The extensive sequence similarities between both HbDXSs and other DXS from different organisms suggest that the cloned HbDXSs represent new members of the plant DXS protein family. Based on the results of multiple amino acid sequences alignment, the two highly conserved motifs: the TPP-binding motif and a transketolase motif, both of which are characteristic of DXS, are presented in the HbDXSs proteins (Fig. 3.5). These motifs are considered as signature domains for proteins belonging to transketolases family (Hawkins *et al.*, 1989; Schenk *et al.*, 1997; Mandel *et al.*, 1996; Querol *et al.*, 2001). Thus, the presence of these two motifs in both HbDXSs sequences confirms that both HbDXSs from *H. brasiliensis* are new TPP-dependent enzyme belonging to the DXS family.

The TPP-binding motif was firstly identified in *A. thaliana* by comparing the CLA1 sequence with several transketolases (Mandel *et al.*, 1996) followed by the alignment of *E. coli* DXS with transketolases from both *E. coli* and yeast (Querol *et al.*, 2001). The TPP-binding motif in the DXS sequences starts with amino acids GDG and ends by amino acids NDN. Other residues between these sequences vary considerably, depending on the species. However, the consensus for the TPP-binding motif in the DXS sequences could be assigned as GDG-X_{25}-NDN. Another conserved region, a transketolase motif, was also found during the alignment of *E. coli* DXS with those of transketolases from *E. coli* and yeast (Querol *et al.*, 2001). The transketolase motif begins with the cluster of amino acids DRAG and concludes with the cluster of amino acids P-X-D. The DRAG motif substitutes the amino acids THDS in the transketolase motif as found in other proteins in the transketolases family. The structural transketolase motif

common to the DXS proteins is, therefore, depicted as DRAG-X_{28}-P-X-D. These motifs are found to be highly conserved in all DXS sequences identified so far.

In addition to these conserved motifs, the important histidine residue that previously shown to be involved in catalysis of yeast transketolase (Linvsqvist *et al.*, 1992; Wikner *et al.*, 1997; Schneider and Linvsqvist, 1998) appear to be highly conserved in the DXS sequences (Querol *et al.*, 2001). This residue was found at position 49 in *E. coli* DXS and its function was studied. Querol *et al.* (2001) have found that the mutation of this residue to glutamine (H49Q) in the *E. coli* DXS cannot recover the growth of an *E. coli* mutant defective in the *DXS* gene. However, the mutant cells can grow in media supplemented with DX. Disruption of *DXS* gene has been shown to cause a lethal phenotype of *E. coli* due to a block in the biosynthesis of essential DXP-derived compounds (Charon *et al.*, 2000). Moreover, the role of this residue has also been studied in *A. tumefaciens*. Upon site-directed mutagenesis of histidine residue to alanine (H48A), the mutant DXS protein from *A. tumefaciens* shows a lower activity (0.1%) as compared to the wild type enzyme (Lee *et al.*, 2007). These results suggest an essential role of the histidine residue, probably in the catalysis of DXS, as found in other transketolases.

In addition, an important role of an invariant glutamic acid residue in DXS has also been studied. It was found that mutation of this acid residue to alanine (E449A) shows a loss of activity in *C. annum* DXS (0.1% in mutant enzyme versus 100% of wild type enzyme (Bouvier *et al.*, 1998), and no enzyme activity was detected in a mutant *E. coli* DXS after changing glutamic acid to alanine (E370A) (Xiang *et al.*, 2007). This residue has been shown to be involved in a catalysis of yeast transketolase by forming a hydrogen bond to N1 nitrogen atom of a pyrimidine ring of the TPP cofactor (Wikner *et al.*, 1994), and part of the TPP-binding site in the *D. radiodurans* DXS (Xiang *et al.*, 2007).

Functional roles of other residues such as arginine corresponding to amino acids 398 and 478 in the *E. coli* DXS (R398 and R478), were also shown to have crucial roles during the catalysis of DXS based on the known structure of the DXS protein (Xiang *et al.*, 2007). Mutation of these residues to alanine (R398A and R478A) in the *E. coli* DXS results in a loss of enzyme activity, confirming the important role of these residues during the catalysis of DXS (Xiang *et al.*, 2007). These residues have been shown to be part of the TPP- and GAP-binding residues, respectively (Xiang *et al.*, 2007). These essential residues are also found in both HbDXSs (Fig. 3.5 and Table 3.3), and they also appear to be highly conserved in all DXSs identified so far. Thus, the presence of TPP-

binding and transketolase motifs, and essential amino acids important for an enzymatic activity of DXS suggest that the two HbDXSs proteins could have similar catalytic properties and mechanism to those from bacteria and other DXSs.

Unlike bacteria, all plant DXSs contain extra amino acids at the NH_2-terminal region that have the common features of cTP sequences in plants. They contain a high amount of hydroxylated and positively charged amino acids, but contain very few acidic amino acids (Gavel and von Heijne, 1990; von Heijne and Nishikawa, 1991; Emanuelsson *et al.*, 1999). These NH_2-terminal extensions vary considerably (Fig. 3.6). The cTP regions in plants serve as an important part for translocation of the protein from cytoplasm to the chloroplast (Mori and Cline, 2001). In addition, computer analysis using the ChloroP program (Emanuelsson *et al.*, 1999) confirmed the presence of cTP sequence in most plant DXSs including both HbDXSs, except DXS sequences from *A. paniculata*, *C. annuum*, *C. morifolium*, *E. guineensis*, *L. esculentum*, *O. sativa1*, and *T. aestivum*, that are not predicted to be localized in the chloroplasts. In some plants, such as *C. annuum* and *L. esculentum*, the computer program prediction does not agree with the recent experimental data. Bouvier *et al.* (1998) have shown that antibody raised against the recombinant *C. annum* DXS recognize plastidial proteins i.e. chloroplasts and chromoplasts, but not mitochondrial proteins by the immunoblot analysis. Similar result has also been shown in *L. esculentum*. When 480 amino acids of the NH_2-terminal region of *L. esculentum* DXS was fused with green fluorescence protein (GFP) under the control of cauliflower mosaic virus 35S promoter, the DXS proteins are shown to be exclusively expressed in chloroplasts *in vivo* (Lois *et al.*, 2000).

Although amino acids within the cTP sequences and around the cleavage motif in different plants vary and cannot be predicted precisely by the computer program, they could have similar properties and function as a chloroplast-targeting signal, consistent with the previously established compartments of the biosynthesis of plant isoprenoids, in which monoterpenes, diterpenes, and tetra-terpenes are biosynthesized in plastids using the MEP pathway (Lichtenthaler, 1999). As suggested by the ChloroP program, most plant DXSs (16 out of 24 DXS sequences or about 67%) shares a similar cleavage site motif 'Xaa↓Ala-Ser-Xaa' (Fig. 3.6). The differences of cTP cleavage site motif in other plant DXSs, which do not fit this pattern, may possible indicated that these sequences were cleaved in more than one step by different types of stromal proteases. This motif is reasonably fit to the consensus cTP cleavage site motif in plants (Gavel and von Heijne, 1990). It is interesting to note that a pair-wise comparison of mature

sequences of HbDXSs after cleavage of the cTP sequences shows higher sequence identity (76%), suggesting that the COOH-terminal parts seem to be crucial for the activity.

1.3 Phylogenetic analysis of the HbDXSs sequences

A phylogenetic tree constructed based on the amino acid sequences of DXS from different organisms including higher plants, green algae, cyanobacteria, and bacteria using the neighbor-joining method in the MEGA program (version 3.1) (Kumar *et al.*, 2004), with the amino acid distances adjusted by the Poisson correction, revealed that the topology of tree present in this study is basically in agreement with the classic taxonomic classification of species (Fig. 3.7 and Fig. 3.8). Bacterial DXSs fall into the same branch, cyanobacterial DXSs are grouped together, and DXSs from plantae are also robustly formed a monophyletic branch. These suggest that the *DXS* genes from different groups of organisms share a common ancestral gene during the course of evolution. The phylogeny of the DXS tree presented in this study is almost identical to that described by Krushkal *et al.*, (2003), Khemvong and Suvachittanont (2005), Gong *et al.*, (2006), Kim *et al.*, (2006) and Phillips *et al.*, (2007), except that the DXS sequences from rhodophyte *C. merolae* and chlorophytes i.e. *C. reinhardii* and *V. carteri* are included in the present analysis.

It is generally accepted that the plantae or green plastids are divided into two mains groups: streptophytes and chlorophytes (Lewis and McCourt, 2004). Streptophytes include all land plants, and chlorophytes include prasinophyceans (i.e. *O. tauri*) and chlorophyceans (i.e. *C. reinhardii* and *V. carteri*). As shown in Fig. 3.8, these two chlorophytes form a sister clade to the streptophytes, well-supported by a high bootstrap probability of 100%. The rhodophyte *C. merolae* is branched basally to the green plastids, with a bootstrap probability support of 99%. These phylogenetic positions are consistent with previous analysis that place chlorophytes at the base of streptophytes, and that for rhodophytes at the base of the green lineage. These analyses are based on fifty sequences of plastid-encoded proteins (Rodríguez-Ezpeleta *et al.*, 2005) and eighty-three genes for plastid-targeted proteins (Archibald *et al.*, 2003), or a single elongation factor Tu gene (Ishida *et al.*, 1997).

It is noteworthy that DXS sequences from streptophytes, chlorophytes, and the rhodophyte are more closely related to cyanobacteria than to those from bacteria, supported by a bootstrapping value of 71% (Fig. 3.7), suggesting that these DXSs are orthologs of the cyanobacterial DXS. It has been proposed that plastids from the present

eukaryotes are incorporated from cyanobacteria progenitors through the endosymbiotic gene transfer (McFadden, 1999; Martin and Herrmann, 1998). Therefore, the closely relationships of DXS between green and red plastids, and cyanobacteria suggest that these DXSs are endosymbiont-derived from the cyanobacteria (Boucher and Doolittle, 2000; Lange *et al.*, 2000).

With respect to higher plants, *DXS* gene forms two distantly groups: class I and class II, with a 90% bootstrap support (Fig. 3.8). The identification of these two divergent classes of plant DXSs was first described by Walter *et al.* (2002). In the present study, twelve plant DXSs are grouped in class I i.e. *A. thaliana*, *A. annua*, *A. paniculata*, *C. annuum*, *E. guineensis*, *G. biloba1*, *L. esculentum*, *M. truncatula1*, *O. sativa1*, *P. montana*, and *T. aestivum*, including HbDXS1 (or *H. brasiliensis1*), and thirteen sequences from i.e. *A. majus*, *C. roseus*, *C. morifolium*, *G. biloba2*, *M. truncatula2*, *M. piperita*, *M. citrifolia*, *N. pseudonarcissus*, *O. sativa2*, *S. rebaudiana*, *T. erecta*, and *T. media*, including HbDXS2 (or *H. brasiliensis2*) are grouped in the DXS class II (Fig. 3.8). Comparison of the DXS sequences from some plants such as *H. brasiliensis*, *G. biloba*, *M. truncatula*, and *O. sativa*, where the DXS class I and class II have been identified, revealed that the members of DXS class I are more closely related to each other (78-83% identity) than they are to the DXS class II in the same species (66-71% identity). For example, the HbDXS1 is more closely related to *G. biloba1* (80% identity), *M. truncatula1* (83% identity), and *O. sativa1* (82% identity) than to the HbDXS2 (69% identity), and vice-versa (Table 4.2). These results are in agreement with the finding that orthologous genes are often more similar to each other than they are to the paralogous genes (Koonin, 2000).

Orthologous genes are genes in different species that evolved from the same gene in the last common ancestor, and paralogous genes are genes that are duplicated from a single gene on the same genome (Koonin, 2000). Thus, in the case of *DXS* gene, all genes belonging to the DXS class II are classified as paralogous genes to the sequences from DXS class I. Genes belonging in the same class (either class I or class II) but present in different species are classified as orthologous genes to each other. The presence of two classes of *DXS* gene in plants suggests that they might diverge from an ancient gene duplication event, which occurs relatively early in the evolution, at least in seed plants. To date, two members of *DXS* gene have been identified from *G. biloba* (Gong *et al.*, 2006; Kim *et al.*, 2006), and three from *Picea abies* (Phillips *et al.*, 2007), both are gymnosperms. For the angiosperms, two members of *DXS* gene have also been

identified from *M. truncatula* (Walter *et al.*, 2002) and *H. brasiliensis* (as presented in this study).

It is already known that gene duplication is an important source of new genes and it plays a vital role during the evolution of organisms. This event may arise through three principal mechanisms: segmental duplication, tandem duplications, and retroposition. After gene duplication, one copy still maintains the ancestor's function, and the other one can either evolve to gain new functions or carry out the same functions of ancestral gene (Gu *et al.*, 2003).

Table 4.2 Percent identity of the deduced amino acid sequence between class I and class II of some representative plant DXSs.

Pair-wise comparison was carried out using the ClustalW program (Thompson *et al.*, 1994). The species used were: *Gb*, *G. biloba*; *Hb*, *H. brasiliensis*; *Mt*, *M. truncatula*; *Os*, *O. sativa*, and their accession numbers are the same as described in Table 3.2. Number 1 means DXS class I and that for number 2 means DXS class II.

	HbDXS1	*GbDXS1*	*MtDXS1*	*OsDXS1*	*HbDXS2*	*GbDXS2*	*MtDXS2*	*OsDXS2*
Hb1	100	80	83	82	69	69	66	69
Gb1		100	78	78	70	69	67	69
Mt1			100	81	67	67	66	67
Os1				100	70	69	68	71
Hb2					100	76	83	79
Gb2						100	72	75
Mt2							100	76
Os2								100

In case of the *DXS* gene, the occurrence of a gene family in plants suggests that they might have different functions particularly in isoprenoid biosynthesis. It has been proposed that genes belonging to the plant DXS class I have important roles in biosynthesis of isoprenoid with primary functions such as chlorophyll and carotenoid which are involved in photosynthesis, while the plant DXS class II may have different functions, without affecting the primary functions (Walter *et al.*, 2002). For example, Lange *et al.* (1998) and Chahed *et al.* (2000) have shown that both *DXS* genes from *M. piperita* and *C. roseus* accumulate at high levels in the terpenoid-accumulating organs/cells and correlate with monoterpenes contents. Hence, these genes could have a

function associated with the synthesis of specific terpenoid rather than other isoprenoids. In addition, the transcripts of *MtDXS2*, a class II DXS from *M. truncatula*, have been found to be correlated nicely with the accumulation of apocarotenoids: mycorradicin and cyclohexenone derivatives. Both are produced in roots after symbiosis of arbuscular mycorrhizas. It is believed that these compounds are also involved in the development of arbuscular mycorrhizal fungi (Walter *et al.*, 2002). Based on these observations, it could be suggested that the *HbDXS1* and *HbDXS2* genes in *H. brasiliensis* might have different functions in the biosynthesis of isoprenoid with different end products as found in other plants.

1.4 Prediction of secondary and 3D structures of HbDXSs

So far, there is no report on the X-ray structure of a plant DXS yet, whereas crystal structures of DXS from bacteria (*E. coli* and *D. radiodurans*) are currently available. Both structures are determined by the selenomethionyl single-wavelength anomalous diffraction (Xiang *et al.*, 2007). Thus, the possible 3D structure models of HbDXSs can be predicted using such structures as templates. It is believed in general that the 3D structure of the two proteins will be similar if the amino acid sequence identity between the target and the template protein is greater than 30%. Amino acid sequences alignment between the HbDXSs (without the cTP sequence) and *D. radiodurans* DXS shows about 39% identity, suggesting that the X-ray structure of *D. radiodurans* DXS (PDB code 2O1X) would be one of the best templates for homology modeling of the structures of HbDXSs. In this study, the CPH-models 2.0 server (Lund *et al.*, 2002) was used to predict the 3D structures of HbDXSs. As seen in Fig. 3.11 and Fig. 3.12, the predicted 3D model structures of HbDXSs and their superimposition onto the structure of *D. radiodurans* DXS revealed that they share common overall folding patterns consistent with the template, implying that the putative function between HbDXSs models and the structure of *D. radiodurans* DXS could be similar.

Inspecting the multiple alignments (Fig. 3.5), there are two important regions of the active site of DXS: the TPP- and GAP-binding sites. These motifs appear to be highly conserved in both HbDXSs, although alteration of some amino acids is found (Table 3.3). For example, residue M185 in *D. radiodurans* DXS is substituted with valine residues in the HbDXSs (V249 in HbDXS1 and V242 in HbDXS2). In addition, substitutions of another four amino acids are found in the HbDXSs sequences compared to the *D. radiodurans* DXS (Table 3.3). It seems likely that these substitutions are conserved in other plant DXSs such as *A. thaliana* (Mandel *et al.*, 1996). Based on the

crystallographic study, Xiang *et al.* (2007) have shown that the main chain carbonyl of M185 is involved in TPP-binding, specifically to the phosphate moiety of the TPP molecule. Therefore, the catalytic mechanism of HbDXSs including the essential role of these amino acids in substrate binding or catalytic activity should be further investigated. To date, there is no information about the kinetic study of DXS enzymes from plants.

2. Molecular cloning and sequence analysis of cDNA encoding HbDXR

2.1 Cloning of two *HbDXRs* genes in *H. brasiliensis*

On the basis of highly conserved sequence of all known plant DXRs published in Genbank database, two closely related *DXR* cDNA clones, designated *HbDXR1* and *HbDXR2*, were identified from *H. brasiliensis*. The *HbDXR1* cDNA was identified from leaves and the *HbDXR2* cDNA was from the latex. The nucleotide and deduced amino acid sequences of both the *HbDXR1* and *HbDXR2* cDNAs are shown in Fig. 3.17 and Fig. 3.18, respectively. Both the numbers of nucleotide and amino acids, as well as molecular sizes predicted for HbDXRs, coincide well with those of other reported DXR in plants identified so far (Table 4.3).

Table 4.3 Properties of the predicted *HbDXRs* proteins and some representative plant DXRs.

	ORF (bp)	Residues	Mw (kDa)	pI	References
HbDXR1	1416	471	51	5.67	This study
HbDXR2	1416	471	51	6.04	This study
C. roseus	1425	474	51	6.10	Veau *et al.*, 2000
L. esculentum	1428	475	52	5.94	Rodríguez-Concepción *et al.*, 2001
A. thaliana	1434	477	52	6.58	Carretero-Paulet *et al.*, 2002
A. annua	1419	472	51	6.15	Souret *et al.*, 2002
S. rebaudiana	1422	473	51	5.97	Totté *et al.*, 2003
Z. may	1419	472	51	6.44	Hans *et al.*, 2004
P. montana	1401	466	51	5.83	Sharkey *et al.*, 2005

Analysis of the nucleotide sequences revealed that the *HbDXR1* and *HbDXR2* cDNAs share 94% identity at the nucleotide level with each other. There are 83 nucleotide mismatches causing 25 amino acid substitutions in the coding regions of these cDNA clones. Twelve of these nucleotide mismatches are found at the first position, nine at the second position, and seven at the third position of the respective codons. The other 55 mismatches observed at the first (5 nucleotides) and third (55 nucleotides) positions do not alter the encoded amino acids. Thus, these results indicate that the two *DXR* cDNAs cloned in this study are likely to represent the allelic variants of one *DXR* gene or products of two different *DXR* genes in *H. brasiliensis*. They are not artifacts of cloning or sequencing errors.

Coincident with these results, Southern blot analysis using the genomic DNA of *H. brasiliensis*, digested with restriction enzymes and probed with the *HbDXR1* fragment, indicates that there is more than one *DXR* genes in *H. brasiliensis* genome (Fig. 3.22). Analysis of the nucleotide sequences of probe and the full-length cDNA clones of *HbDXRs* revealed that there are no detectable of cleavage sites for the enzymes used, except *Bam*HI and *Xho*I sites. However, these sites are found outside the probe at nucleotide positions 241-245 and 1327-1332 in both the *HbDXR1* and *HbDXR2* sequences, respectively (data not shown). This information could support the cloning of two *HbDXRs* genes in *H. brasiliensis*. Thus, the *DXR* gene found in *H. brasiliensis* is part of a small gene family, consisting of at least two members. The presence of two copies of *DXR* genes in *H. brasiliensis* might result from a gene duplication event, as described for the *DXS* gene. At present, the existence of the additional *HbDXRs* genes in the *H. brasiliensis* genome cannot be excluded, since there are some weak hybridizing bands observed in the X-ray film (Fig. 3.22). It is possible to obtain further information about the other possible *DXR*-related genes in *H. brasiliensis* in the future.

To our knowledge, this is the first study on cloning of two *DXR* genes in plants. Thus, further investigation in more plant species is required to determine whether the presence of more than one *DXR* gene is also widespread in other different plant species. A search in the TIGR database (Quackenbush *et al.*, 2000) using the *HbDXR1* cDNA as a query revealed that there are also two *DXR* ESTs in *G. max* (TC204480 and TC204482) and *M. truncatula* (TC76992 and TC78572). An extensive search against the *P. trichocarpa* database also revealed the presence of two *DXR* genes (estExt_Genewise1_v1.C_LG_XII0355 and estExt_Genewise1_v1.C_LG_XV1512) (Sterky *et al.*, 2004). In addition to higher plants, two EST clones corresponding to two

DXR genes were also found in the *P. patens* (pph19g08 and pph20p16) but their sequences were incomplete (Nishiyama *et al.*, 2003). The occurrence of multiple forms of *DXR* in these plants may reflect the diversity of the *DXR* genes that catalyze reactions for the synthesis of a wide range of primary and secondary metabolites, as suggested for the case of the *DXS* gene (Walter *et al.*, 2002). In contrast to these findings, previous studies have shown that a single *DXR* gene was isolated from a number of plants such as *A. thaliana* (Schwender *et al.*, 1999; Carretero-Paulet *et al.*, 2002), *L. esculentum* (Rodríguez-Concepción *et al.*, 2001), *O. sativa* and *Z. may* (Hans *et al.*, 2004), and *G. biloba* (Gong *et al.*, 2005; Kim *et al.*, 2006).

Despite the MEP pathway genes, a gene family for enzymes involved in isoprenoid biosynthesis, such as *HMGS* and *HMGR* genes, were also found to be encoded by a gene family in *H. brasiliensis*. There are two *HMGS* genes (Suwanmanee *et al.*, 2002; Sirinupong *et al.*, 2005) and three *HMGR* genes (Chye *et al.*, 1991; Chye *et al.*, 1992) identified in *H. brasiliensis*. These genes encode the enzymes HMGS and HMGR, which are essential for controlling isoprenoid biosynthesis *via* the MVA pathway in plants and animals (Goldstein and Brown, 1990). These evidences suggest that duplication of genes encoding certain enzymes of the isoprenoid biosynthesis may be a general feature of plants. Gene duplication represents a major driving force for the recruitment involving in a secondary metabolism (Picherskya and Ganga, 2000; Ober, 2005).

2.2 Comparison of the deduced *HbDXRs* amino acid sequences

Using the ClustalW program (Thompson *et al.*, 1994), comparison of the deduced amino acid sequences of both HbDXRs with those of other DXRs showed that these two sequences exhibit relatively high homology to other plant DXRs, with the highest sequence identity to *L. esculentum* (88% identity) (Rodríguez-Concepción *et al.*, 2001) (Table 3.4). These two HbDXRs also share lower sequence identity with those from other prokaryotic DXRs (40-70% identity). Despite these differences, multiple alignments of the deduced amino acid sequences of HbDXRs with those from different organisms reveal a region of striking similarity at the COOH-terminus (Fig. 3.19), suggesting an important role of this region of DXR.

In the NH_2-terminal region of all DXRs, two highly conserved motifs were found (Fig. 3.19). The first motif is the NADPH-binding motif and second motif is the Proline rich motif. The presence of the NADPH-binding motif suggests that DXR proteins require a NADPH as a cofactor, similar to other NADPH-dependent enzyme, such as dehydrogenases (Lesk, 1995). It has been shown that replacement of NADPH by

NADH results in the loss of activity of *E. coli* DXR to 1% of the wild type enzyme with NADPH (Takahashi *et al.*, 1998). Based on the 3D structure of *Z. mobilis* DXR, the main chain amides of N39, R40, and N41 interact with the 2´ phosphate of NADPH by hydrogen bonds. Thus, lack of the 2´ phosphate group in NADH could break hydrogen bonds during the binding of cofactor resulting in a decrease in activity (Ricagno *et al.*, 2004). On the contrary, the Proline rich motif is found only in eukaryotic DXRs, such as plants, green, and red algae. The biological function of this motif in these DXRs is unknown, since it is absent in the bacterial counterparts. Carretero-Paulet *et al.* (2002) proposed that the Proline-rich region of plant DXRs might be involved in specific interactions with regulatory proteins or other enzymes of the MEP pathway. The function of the Proline-rich region in plant DXRs, however, remains to be elucidated. In animals, Proline-rich regions play an important role in signaling events by interacting with Src homology 3 domain of other cellular proteins (Mayer and Baltimore, 1993; Mayer, 1999). The Src homology 3 domain is a small noncatalytic domain of around 50-60 amino acids. This domain mediates protein-protein interactions that are important for coupling of intracellular signaling pathways, regulation of catalytic activity of proteins, recruitment of substrates to enzymes, and localization of proteins to a specific compartment (Mayer and Baltimore, 1993; Mayer, 1999).

In addition to these two conserved motifs, several amino acids were formed the catalytic site, such as D150, E152, E231, E234, H209, and H257 (*E. coli* DXR numbering), and the residues previously shown to be important for the substrate-binding in the *E. coli* DXR (S151, S186, N227, K228, H209, W212, E152, and E231, where numbers are assigned according to the residue positions in *E. coli*), appear to be conserved in both HbDXRs (Fig. 3.19) and all other DXR proteins. These residues are parts of substrate-binding and divalent cation-binding sites, based on the structure of *E. coli* DXR (Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; Yajima *et al.*, 2004; MacSweeney *et al.*, 2005). Mutagenesis studies revealed that mutation of these residues produces enzymes with no activity or produces inactive enzymes in both *E. coli* and *Synechocystis* sp. PCC6803 (Kuzuyama *et al.*, 2000; Fernandes and Proteau, 2006). According to these observations, the cloned *HbDXRs* cDNAs are expected to encode active enzymes and utilize mechanisms similar to that of the *E. coli* DXR, since both HbDXRs contain all important regions characteristic of the DXR. However, their biochemical properties will be further elucidated for better understanding the molecular mechanisms of these proteins.

Since the MEP pathway has been shown to be localized in plastids (Eisenreich *et al.*, 1998; Lichtenthaler, 1999; Eisenreich *et al.*, 2001; Rodríguez-Concepción and Boronat, 2002; Eisenreich *et al.*, 2004), evidence for the presence of the NH_2-terminal leader sequence that encodes the plastid-targeting sequences or cTP in both HbDXRs, as well as other plant DXR, is derived from the multiple alignment and the *in silico* prediction by the ChloroP program (Emanuelsson *et al.*, 1999), as described for the HbDXSs sequences. The results showed that most plant DXRs are predicted to be localized in chloroplast. However, for the two HbDXRs, these predictions did not provide concrete answers (Table 3.5). Scores for potential chloroplast targeting of HbDXR1 and HbDXR2 were 0.485 and 0.500, respectively, where a value that greater than 0.5 is considered as indicative of chloroplast localization (Emanuelsson *et al.*, 1999). Analysis of the HbDXRs sequences revealed that the NH_2-terminal extension sequences of both HbDXRs exhibit common features of the cTP sequence, for example, they contain high numbers of hydroxylated and basic amino acids, as found in the case of HbDXSs. In addition, they also show the same pattern of hypothetical cleavage site for cTP sequence (Xaa↓Cys-Ser-Xaa), similar to other plant DXRs (Fig. 3.19 and Fig. 3.20). Although it does not prove in this study that the HbDXR1 and HbDXR2 proteins are localized in plastids, these findings suggest that these extra amino acids found in both HbDXRs could act as a cTP leader sequence. Thus, both HbDXRs and other plant DXRs could reside within the plastid compartment. The localization of the DXR protein in chloroplast of higher plants was confirmed experimentally by immunoblot analysis and transient expression using the GFP in *A. thaliana* (Carretero-Paulet *et al.*, 2002). The cTP sequences of HbDXRs show more than 50% identity to the cTP sequence of *A. thaliana*.

2.3 Phylogenetic analysis of the HbDXRs sequences

The phylogenetic relationships of HbDXRs (Fig. 3.21) generated by the neighbor-joining method in the MEGA program (version 3.1) (Kumar *et al.*, 2004) as described for the DXS tree, produce a consensus tree similar to the DXR tree previously reported by Yao *et al.* (2007). The DXR sequences from both green and red algae are also included in the analysis. The DXR tree present in this study reveals three distinct lineages corresponding to bacteria, cyanobacteria, and green lineages. The green lineages are further separated into streptophytes and chlorophytes groups. The relationships between streptophytes and chlorophytes are supported by a bootstrap probability of 89% (Fig. 3.21), and they are closely similar to cyanobacteria DXRs, suggesting they are originated from a cyanobacterial progenitor through the primary symbiosis (McFadden, 1999; Martin

and Herrmann, 1998), as found in other genes in the MEP pathway (Lange *et al.*, 2000). Within streptophytes or seed plants, the DXRs form a monophyletic group but they are further resolved into two clades including gymnosperms and angiosperms. The gymnosperm DXRs are resolved as a sister clade to all other angiosperm DXRs. Their relationships are supported by a bootstrapping value of 100% (Fig. 3.21).

In addition, the phylogenetic position of DXR from the rhodophyte *C. merolae* is more closely related to cyanobacterial proteins than to those from green algae and green plants, although its position is not really supported (Fig. 3.21). This result is inconsistent with the recent phylogenetic data suggested that the different algal groups diverged near the base of the eukaryotic tree (Ishida *et al.*, 1997; Archibald *et al.*, 2003; Rodríguez-Ezpeleta *et al.*, 2005). Thus, the divergence of *C. merolae* DXR, and green algae and green plants needs to be further elucidated by using a large amount of sequence data, since limited sequence data and taxon sampling of species may lead to an inaccurate phylogeny. It is worth mentioning that the NH_2-terminal presequence of the *C. merolae* DXR is a little bit longer than those from green algae and green plants (data not shown). This observation suggests that the primary plastids in rhodophytes might have been lost from a common ancestor and thus the secondary plastids might have mediated a *DXR* gene transfer in this lineage. Additional DXR sequences from more species of rhodophytes may help resolve this result.

As mentioned above, plant DXRs can be separated into two groups: gymnosperms and angiosperms. The cloned HbDXRs are grouped within the angiosperms and they are more closely related to *L. esculentum* DXR than to those from other plants (Fig. 3.21). This is consistent with the results obtained from the deduced amino acid sequences comparison. In addition, the two DXRs from *H. brasiliensis* fall into the same phylogenetic branch with each other, with a 100% bootstrap value support (Fig. 3.21), implying that they might result from a recent gene duplication event, as described for the HbDXSs sequences, and they might be under an evolutionary pressure to maintain a functional redundancy of DXR in *H. brasiliensis*. The occurrence of multiple DXRs in *H. brasiliensis* and other respective plants described above may reflect the diversity of DXRs that catalyze reactions for the synthesis of a wide range of primary and secondary metabolites essential for plant growth and development.

2.4 Prediction of secondary and 3D structures of HbDXRs

To date, several crystal structures of DXR from many bacteria have been reported by different research groups. Most of which are from *E. coli* (Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; Yajima *et al.*, 2004; MacSweeney *et al.*, 2005). Therefore, it is interesting to know whether these crystal structures would be useful for the homology modeling of unknown structures. The X-ray data from *E. coli* (PDB code 1K5H) was chosen as a template for the prediction of 3D models of HbDXRs. The models are constructed based on the SWISS-MODEL program (Schwede *et al.*, 2003). Comparison of the deduced amino acids between template and HbDXRs sequences shows 48% identity.

As expected, the overall model structures of HbDXRs are very similar to the structure of template (Fig. 3.24 and Fig. 3.25). The comparison of superimposed structures between the *E. coli* DXR structure and the HbDXRs models revealed a similar topology of many structurally equivalent regions (Fig. 3.26). The NH_2-terminal domain of HbDXRs exhibits a typical dinucleotide-binding motif, composing of βαβ units called a Rossmann fold (Fig. 3.23 and Fig. 3.24). This region participates in the NADPH-binding, consistent with the structural information of DXRs (Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; Yajima *et al.*, 2004; MacSweeney *et al.*, 2005), and it is perfectly maintained throughout the DXRs characterized so far. The Rossmann fold was first identified in the dinucleotide-binding proteins by Rossmann *et al.* (1974). In addition, the active site of DXR is found to be located in the central catalytic domain (Reuter *et al.*, 2002; Yajima *et al.*, 2002; Steinbacher *et al.*, 2003; Yajima *et al.*, 2004; MacSweeney *et al.*, 2005). Several essential residues in this region are considered to be highly conserved in the HbDXRs (Fig. 3.19). The overall folding patterns of both HbDXRs are very similar to the *E. coli* DXR template. Thus, it is possible that the substrate DXP, NADPH, and the metal ion could bind to the HbDXRs in a similar manner. Therefore, these residues are prime targets for site-directed mutagenesis and their role in substrate interactions with this enzyme needs further investigation.

3. Expression analysis of *HbDXSs* and *HbDXRs* mRNAs

3.1 Expression of *HbDXSs* and *HbDXRs* genes in various tissues

The expression of *HbDXSs* and *HbDXRs* genes in various tissues of *H. brasiliensis* was analyzed in both seedlings and mature plants by semi-quantitative RT-PCR. This RT-PCR analysis employed the gene-specific primers for each gene and *18S rRNA*-specific primers as internal control to confirm the equal amount of RNA used as template in each samples. The *18S rRNA* gene is recommended to be used as an internal standard for the mRNA quantification study. Total cellular RNA level cannot be highly modified compared to the *18S rRNA* gene. Generally, the RT-PCR products of the mRNA of interest reflect its expression, which is weak in comparison with the expression of *18S rRNA* gene (Thellin *et al.*, 1999). The *18S rRNA* gene has also been previously used as an internal standard by Sirinupong *et al.* (2005) and Tang *et al.* (2007).

In this study, the transcripts of both the *HbDXSs* and *HbDXRs* genes were found to be present in all tissues analyzed including roots, stems, and leaves of seedling, and flowers, latex, leaves, and leaves of mature plants naturally infected with *O. heveae* (Fig 3.30 and Fig. 3.32). In roots of seedlings (Fig 3.30A) and latex of mature plants (Fig 3.30B), the transcripts of *HbDXS1* are hardly detectable. The different expression patterns of the *HbDXS* and *HbDXR* genes in the different tissues analyzed further supports the idea that both enzymes are indeed encoded by two different genes. In general, these genes show similar expression patterns in the RT-PCR analysis (Fig 3.30 and Fig. 3.32). For example, the transcripts of both *HbDXS1* and *HbDXR1* are highly expressed in photosynthetic tissues such as leaves, while the transcripts of both *HbDXS2* and *HbDXR2* are highly expressed in non-photosynthetic tissues, such as roots, flowers, and latex, with an exception of higher expression of *HbDXR2* transcripts in leaves.

The detection of both *HbDXSs* and *HbDXRs* transcripts in different parts of *H. brasiliensis* suggests that their encoded proteins have an essential function in plant, but with different physiological roles in controlling the production of different-end products of isoprenoids. In several plants, these *DXS* and *DXR* genes also show differential expression in different tissues of the respective plants, further supporting the notion that the MEP pathway is involved in synthesizing a variety of isoprenoid-end products in different parts of the plants (Estévez *et al.*, 2000; Lois *et al.*, 2000; Carretero-Paulet *et al.*, 2002; Walter *et al.*, 2002; Burlat *et al.*, 2004; Hans *et al.*, 2004; Dudareva *et al.*, 2005; Gong *et al.*, 2005; Gong *et al.*, 2006; Kim *et al.*, 2006; Liao *et al.*, 2007; Phillips *et al.*, 2007; Yao *et al.*, 2007).

The expression of *HbDXSs* and *HbDXR*s transcripts in photosynthetic tissues such as leaves (Fig 3.30 and Fig. 3.32) might indicate the requirement of these products in primary metabolism, i.e. involvement in chlorophyll and carotenoid biosysthesis. Consistent with these roles, previous studies have shown that the expression product of the *CLA1* or *DXS* gene in *A. thaliana* is required for normal chloroplast differentiation and the mutation of this gene gives an albino phenotype (Mandel *et al.*, 1996; Estévez *et al.*, 2000). A mass screening of the T-DNA and transposon insertion lines in *A. thaliana* also allowed Budziszewski *et al.* (2001) to identify an albino mutant disrupted in a *DXR* gene. In addition to *DXS* and *DXR* genes, disruption of other genes in the MEP pathway such as *ispD*, *ispE*, *ispF*, *ispG*, and *ispH*, which encode for other enzymes of the pathway taking place in downstream of *DXS* and *DXR* (Fig. 1.1), also causes a lethal-albino phenotype in plants (Okada *et al.*, 2002; Gutiérrez-Nava *et al.*, 2004; Page *et al.*, 2004; Guevara-García *et al.*, 2005; Hsieh and Goodman 2005; Hsieh and Goodman 2006; Ahn and Pai, 2008; Hsieh *et al.*, 2008). All mutants of the MEP pathway genes deficient in pigment production also died shortly after germination. Loss of photosynthetic pigments may affect the development of the chloroplast and plastid structures of the mutant plants, for example, the accumulation of large vesicles instead of thylakoid membranes was observed inside the mutant chloroplasts (Okada *et al.*, 2002; Gutiérrez-Nava *et al.*, 2004; Page *et al.*, 2004; Guevara-García *et al.*, 2005; Hsieh and Goodman 2005; Hsieh and Goodman 2006; Ahn and Pai, 2008; Hsieh *et al.*, 2008). In contrast to the loss of function of these genes, constitutive overexpression of *DXR* gene in *A. thaliana* results in increased accumulation of chlorophylls and carotenoids (Carretero-Paulet *et al.*, 2006). Moreover, *A. thaliana DXS* overexpresser lines have increased levels of various isoprenoids including chlorophylls, carotenoids, abscisic acids and gibberellins (Estévez *et al.*, 2001). These results suggest the important roles of the genes in the MEP pathway not only for *DXS* and *DXR* genes but also for all genes of this pathway in plastid development as well as in the control of chlorophyll and carotenoid biosynthesis in green tissues and other plastid-derived isoprenoids in plants.

The presence of the *HbDXS2* and *HbDXRs* transcripts in non-photosynthetic tissues such as roots and flowers (Fig 3.30A and Fig. 3.32A) indicates a possible role of the MEP pathway in production of a variety of isoprenoids essential for growth and development in these heterotrophic tissues in *H. brasiliensis*. These results are in good agreement with the expression patterns of both *DXS* and *DXR* genes in *A. thaliana* (Estévez *et al.*, 2000; Carretero-Paulet *et al.*, 2002). The development of roots

and flowers of *H. brasiliensis* might require some special isoprenoids synthesized by the MEP pathway and as yet uncharacterized. The exact function of both genes and their developmental requirement in these tissues is still unclear and further studies are needed to clarify their functions in *H. brasiliensis*. It has been shown in *A. thaliana* that mutant plants disrupted in *DXS* gene do not developed mature flowers, indicating the requirement of specific isoprenoids during developmental processes of flowers (Mandel *et al.*, 1996).

In addition to the development of flowers, another function of floral-derived isoprenoids in plants is the involvement of such compounds in a floral scent to attract insects or other pollinators. Dudareva *et al.* (2005) have shown that in *A. majus*, volatile isoprenoids such as myrcene and nerolidol in a floral scent of *A. majus* involved in the attraction of bee pollinators are synthesized by the MEP pathway. The MEP pathway genes such as *DXS* and *DXR* are preferentially expressed in corolla tissues of flowers, where these volatile isoprenoids are produced and emitted (Dudareva *et al.*, 2005). In addition to scent attractants, flowers of *T. erecta* in dark orange cultivar accumulate carotenoids at levels up to 100 times greater than those found in white or yellow cultivars and the accumulated carotenoids correlate closely with the expression patterns of *DXS* and *DXR* genes (Moehs *et al.*, 2001). Moreover, plant hormones such as gibberellins synthesized using the intermediates derived from the MEP pathway play many roles in plant development. For example, they regulate flowering time, development of flowers, and onset of fruits (Hooley, 1994). Thus, the expression of both *DXS* and *DXR* genes in floral organs of different plants might provide different precursor pools required for the synthesis of various kinds of isoprenoid compounds.

Previous studies reported that the expression of *DXS* and *DXR* genes in roots of some plant species, such as *O. sativa*, *Z. mays*, *T. aestivum*, *H. vulgare*, and *M. truncatula* (only the *MtDXS2* but not the *MtDXS1* isoforms) are up-regulated upon colonization by mycorrhizal fungus, and the increased in levels of these transcripts correlates with the accumulation of apocarotenoids i.e. mycorradicin and glycosylated cyclohexenone derivatives, as well as with the development of mycorrhizal structures (Walter *et al.*, 2000; Walter *et al.*, 2002; Hans *et al.*, 2004). However, very little is known about the mycorrhizal fungus-mediated symbiosis in *H. brasiliensis* roots.

In contrast to fungal symbiosis, in leaves of *H. brasiliensis* invaded naturally with the fungus *O. heveae*, the expression levels of the *HbDXSs* and *HbDXR2* genes seem to be reduced (Fig 3.30B and Fig. 3.32B), suggesting that both genes are elicitor-responsive genes. Fungal infection led to an increase in steady-state transcript

levels of *DXS* class II genes in cell culture of *P. abies* (Phillips *et al.*, 2007). In addition to *DXS* gene, Yamasakia and Akimitsu (2007) have shown that the expression of *ispF* gene in leaves of *Citrus jambhiri* is not induced by either pathogenic or non-pathogenic fungi. In *A. thaliana*, however, the expression of the *ispG* gene has been shown to be repressed by a bacterial elicitor (Gil *et al.*, 2005). Based on these reports, it seems likely that different plants develop different responses to pathogen infection and microbial symbiosis.

In *H. brasiliensis*, latex is the cytoplasmic content of laticifers or latex vessels, a major site of natural rubber biosynthesis (de Faÿ and Jacob, 1989). The chemical composition of natural rubber is *cis*-1,4-polyisoprene, produced from IPP as a common precursor (Fig. 1.10) (Kekwick, 1989). So far, the MVA pathway is considered to be a major pathway for the synthesis of natural rubber (Skilleter and Kekwick, 1971; Hepper and Audley, 1969). Genes in the MVA pathway such as *HMGS* and *HMGR* have been shown to be highly expressed in laticifers (Kush *et al.*, 1990; Chye *et al.*, 1992; Suwanmanee *et al.*, 2002; Suwanmanee *et al.*, 2004; Sirinupong *et al.*, 2005). Since the discovery that plants have two pathways for IPP synthesis: MVA and MEP pathway (Fig. 1.1), it became interesting to investigate whether the novel MEP pathway plays any role in the synthesis of natural rubber. No information about the role of MEP pathway in rubber biosynthesis is available.

In more recent years, the MEP pathway has been considered as a possible alternative route for rubber biosynthesis (Ko *et al*., 2003). In agreement with this notion, both the *HbDXS2* and *HbDXRs* genes are also found to be transcriptionally expressed in the latex of *H. brasiliensis* (Fig 3.30B and Fig. 3.32B), strongly indicating a possible functional role of these genes in metabolism of laticifers, where natural rubber is synthesized. For example, they might involve in rubber biosynthesis by providing IPP, which is further converted to DMAPP by IPP isomerase (Oh *et al.*, 2000) and used as substrate for condensation with other intermediates produced from the MVA pathway to synthesize rubber. Therefore, both pathways could provide substrates to produce rubber in *H. brasiliensis*.

It is worth noticing that the latex of *H. brasiliensis* contains specialized organelles called Frey-Wyssling particles. These particles contain double membranes and accumulate carotenoid compounds, a characteristic of plastids (Gomez and Moir, 1979, de Faÿ *et al*., 1989). Since carotenoids are synthesized from intermediates derived from the MEP pathway, it seems likely that the MEP pathway exists in Frey-Wyssling particles

might have an important function in laticifers by synthesizing various kinds of isoprenoids, such as carotenoids, either instead of or in addition to supplying substrates for rubber biosynthesis. Carotenoids produced in laticifers may function as antioxidants or be involved in plant disease resistance, as reported in other plants (Hundal *et al*., 1995; Maciejewska *et al*., 2002).

Previously, it has been proposed that Frey-Wyssling particles may be involved in rubber biosynthesis, but there is no evidence to support this hypothesis (Dickenson, 1969). Wititsuwannakul *et al.* (2003) have shown that non-rubber particles of latex from the bottom fraction, which contains predominantly Frey-Wyssling particles and lutoids, can synthesize new rubber molecules, indicating that these particles may be another site for rubber initiation. Although it is difficult to put forward a hypothesis concerning the Frey-Wyssling particles in rubber biosynthesis, a possible role of the MEP pathway in laticifers in relation to synthesis of rubber cannot be excluded. IPP generated by Frey-Wyssling particles may be transported out into cytosol and used as substrate for rubber biosynthesis together with the MVA pathway. In plants, cross talk between the two IPP biosynthetic pathways has been documented (Bick and Lange, 2003; Hampel *et al.*, 2005). Thus, the relative involvement of Frey-Wyssling particles to the biosynthesis of natural rubber remains to be resolved.

3.2 Expression of *HbDXS2* and *HbDXR2* genes in different clones

An observation that the transcripts of *HbDXS2* but not the *HbDXS1* are detected in latex and the transcripts of *HbDXR2* are preferentially expressed in latex than the *HbDXR1* led to an investigation of a possible role of the MEP pathway in rubber biosynthesis. The expression pattern of these genes in different *H. brasiliensis* clones is compared with the latex yield. In this study, two clones of *H. brasiliensis* with contrasting latex yield are used: RRIM 600 and wild type. The clone RRIM 600 produces higher rubber yield than the wild type, by about 2.5-folds in latex exported and two-folds in dry rubber content (Fig. 3.33A and Fig. 3.33B), similar with previously reported results (Suwanmanee *et al.*, 2004). These results confirm the earlier report that latex yield is a clonal character (Tupy, 1988; Jacob *et al.*, 1989).

As Fig. 3.33C shows, the expression level of *HbDXS2* transcripts is not significantly different between the two clones, while the expression levels of *HbDXR2* transcripts is higher in clone RRIM 600 than in the wild type, as revealed by RT-PCR. With respect to these results, it seems likely that the level of *HbDXR2* transcripts correlates with isoprenoid biosynthesis in *H. brasiliensis* as judged by the rubber yield.

These results lead me to surmise that the MEP pathway might play some roles in regulating the rubber biosynthesis, and such regulation occurred at least partly at the transcriptional level. In agreement with these results, the participation role of *DXR* gene in the control of isoprenoid accumulation in plants has been investigated. For instance, overexpression of *DXR* gene in transgenic *M. piperita* leads to 50% increased of essential oil monoterpenes (Mahmound and Croteau, 2001). In addition, *A. thaliana*-overexpressing *DXR* lines accumulate the high amount of chlorophylls and carotenoids (Carretero-Paulet *et al.*, 2006). In contrast, down regulation of *DXR* gene leads to reduced accumulation of oil monoterpenes in *M. piperita* (Mahmound and Croteau, 2001), and shows the variegated phenotypes such as reduced in pigmentation and defected in chloroplast development of transgenic *A. thaliana* (Carretero-Paulet *et al.*, 2006). These results support the role of *DXR* important for control of metabolic flux for isoprenoid biosynthesis by the MEP pathway.

Unlike these notions, the transcripts of *DXR* in *L. esculentum* and *E. guineensis* do not correlate with the accumulation of carotenoids during ripening of fruits, while the levels of *DXS* transcripts correlates (Lois *et al.*, 2000; Rodríguez-Concepción *et al.*, 2001; Khemvong and Suvachittanont, 2005). In flowers of *A. majus* flowers, the level of *DXS* transcripts correlates with the pattern of myrcene and nerolidol emissions (Dudareva *et al.*, 2005). Based on these studies, *DXS*, but not the *DXR*, is indeed control the rate-limiting step of isoprenoid biosynthesis using the MEP pathway in plants. Several investigations also support this conclusion. For example, overexpression of *DXS* gene from *A. thaliana* increases essential oil production in *Lavandula latifolia* in leaf and flower glandular trichomes (Muñoz-Bertomeu *et al.*, 2006), and that from *E. coli* increases carotenoid content in tubers of transgenic *Solanum tuberosum* (Morris *et al.*, 2006). In *A. thaliana*, overexpression of an endogenous *DXS* gene increases various isoprenoids, including total chlorophylls and carotenoids (Estévez *et al.*, 2001). In addition to these observations, both *DXS* and *DXR* genes could have a regulatory role in controlling the production of carotenoids in *O. sativa*, *Z. mays*, *T. aestivum*, and *H. vulgare* (Walter *et al.*, 2000), and terpenoid indole alkaloids in *C. roseus* cell suspension culture (Veau *et al.*, 2000). According to these results, it can be concluded that several enzymes of the MEP pathway in plants may contribute to the control of this metabolic route, but this depends on the species, tissue, and physiological state of plants. In addition, Guevara-Garcia *et al.* (2005) demonstrate that the regulation of the MEP pathway occurred at both transcriptional and posttranscriptional levels in *A. thaliana*. However, there is no other

experimental evidence supporting this aspect in other plants, thus it remains to be further studied.

As the MVA pathway represents the major metabolic route for rubber synthesis in *H. brasiliensis*, previous studies reported that the activity of HMGR and HMGS, and *HMGS* transcripts are higher in high-yielding clones than in low-yielding clones, and these correlate with the latex yield (Wititsuwannakul *et al.*, 1988; Suwanmanee *et al.*, 2004). In addition to these genes, the recent study by Priya *et al.* (2006) reveals that the expression level of rubber elongation factor, considered to be a key gene involved in rubber biosynthesis (Dennis and Light, 1989), is also higher in high-yielding *H. brasiliensis* clones than in the low yielders. These results suggest that the higher expression of genes involved in rubber biosynthesis correlates with the rubber yield. Therefore, the higher expression of *HbDXR2* suggests that its function might contribute in some way to the control of rubber synthesis possibly by producing more IPP precursor, although such precursor is mainly from the MVA pathway. In the future, ectopic expression of these genes in *H. brasiliensis* will provide more information concerning their functions in controlling rubber biosynthesis, which is more complex and not clearly understood to date.

3.3 Expression of *HbDXS2* and *HbDXR2* genes under ethephon treatment

It is widely known that treating *H. brasiliensis* bark with ethephon increases latex yield and regulates expression of genes involved in rubber biosynthesis (Coupé and Chrestin, 1989; Pujade-Renaud *et al*., 1994; Suwanmanee *et al.*, 2004). Ethephon releases ethylene under physiological conditions in plants. In this study, the wild type *H. brasiliensis* treated with 2.5% ethephon increases both exported latex and dry rubber content (Fig. 3.34A and Fig. 3.34B), as compared to the same trees before treatment. In contrast, the levels of both *HbDXS2* and *HbDXR2* transcripts are transiently induced after 24 h of ethephon treatment and then decreased (Fig. 3.34C), while the rubber yield continued to increase. These results suggest that the induced accumulation of rubber production is not preceded by an increase of expression of both *HbDXS2* and *HbDXR2* genes of the MEP pathway, which is different from previously reports on an induction of *HMGS*, *HMGR*, and rubber elongation factor upon ethephon treatment (Chye *et al.*, 1992; Suwanmanee *et al.*, 2004; Priya *et al.*, 2006; Priya *et al.*, 2007). Adiwilaga and Kush (1996) have also shown that the transcripts of *FDS* encoding FPP synthase, a key enzyme in rubber biosynthesis, show no effect after treated with ethephon.

The exact function of ethylene on latex metabolism of *H. brasiliensis* is not clearly understood. The increase of rubber content in the latex may result from a prolonged time of latex flow, causing the increased in volume of exported latex. It has been presumed that ethylene acted on membrane permeability of a latex organelle i.e. lutoids, which are found up to 10-20% in latex (de Faÿ *et al.*, 1989), by preventing or delaying the rupture of lutoid membranes, and thus preventing the coagulation of latex. Since the lutoids are single membrane, it might be easily damaged by physical shear occurring during the flow of latex upon tapping, resulting in an aggregation of latex (Coupé and Chrestin, 1989).

4. Expression of recombinant proteins and Western blot analysis

Attempts are made to overexpress recombinant HbDXS1 and HbDXR1 proteins in *E. coli*, and the results showed that these proteins are successfully expressed using the *E. coli* expression system, but the proteins appear to localize in inclusion bodies. Although it has been established that the expression of foreign proteins in *E. coli* is fast and efficient, the often problem of the most heterologous expression in this system is the low solubility and improper folding of recombinant proteins (Wetzel, 1996). Factors that dictate the formation of insoluble protein has not been fully understood. Many strategies have been reported to prevent the formation of insoluble protein included: co-expression of chaperones, optimization of growth medium, and culturing conditions. These strategies have been proven to be effective in improving the solubility of a number of difficult proteins (Georgiou and Valax, 1996; Nishihara *et al.*, 1998; de Marco *et al.*, 2000; de Marco and De Marco, 2004). These methods have also been applied for the improving expression of both HbDXS1 and HbDXR1 proteins, but they failed to produce more soluble proteins (data not shown). Attempts in *in vitro* refolding of the recombinant HbDXR1 after purification under denaturating conditions also failed to yield an active enzyme, as detected by an activity assay (data not shown).

In addition to *H. brasiliensis*, the recombinant *A. thaliana* DXR was also expressed in insoluble form in inclusion bodies when expressed in *E. coli* (Rohdich *et al.*, 2006). However, one expression-construct containing amino acid residues 57-477 yields an active DXR, and thus allowing authors to characterize its biochemical properties (Rohdich *et al.*, 2006). In the future, it is interesting to construct several types of HbDXR1 or HbDXS1 sequences by shorten more amino acid residues after a predicted cTP cleavage site when expressed in the *E. coli* system to obtain more soluble protein, as

found in the case of *A. thaliana* DXR. Instead of expression in *E. coli*, the expression of both HbDXS1 and HbDXR1 could also be performed using yeast (*S. cerevisiae*) as host in the future, since yeast has some advantages over *E. coli* for expression of eukaryotic genes. For example, yeast can introduce post-translational modifications of proteins i.e. phosphorylation, N- and O-linked glycosylation, disulfide cross-linking, and subcellular targeting, which may be important for the biological functioning and the stability of recombinant proteins (Holz *et al.*, 2002). To date, a number of plant proteins have been expressed successfully in yeast system, for example, a fatty acid desaturase gene from *G. max* (Li *et al.*, 2007), a cytochrome P450 monooxygenase (CYP71AV1) and its partner redox, NADPH: cytochrome P450 oxidoreductase genes from *A. annua* (Ro *et al.*, 2006), 2,3-oxidosqualene triterpenoid cyclase genes from *A. thaliana* (Husselstein-Muller *et al.*, 2001), and a myrosinase gene from *Brassica napus* (Chen and Halkier, 1999).

The purified recombinant HbDXS1 and HbDXR1 proteins were further used to produce antibodies. Both HbDXS1 and HbDXR1 antisera are specific for DXS and DXR proteins expressed in *E. coli*, respectively, as judged by western blot analysis (Fig. 3.37A, lane 1). Further, the HbDXR1 antiserum can detect not only a putative DXR from *H. brasiliensis,* but also from *A. thaliana* (Fig. 3.37A, lane 2 or Fig. 3.38A, lane 2) and *P. trichocarpa* (Fig. 3.38A, lane 1). These cross-reactive reactions may be due to a high sequence similarity of DXR between these plant species i.e. HbDXR1 exhibits 84% and 90% identity with *A.thaliana* DXR and *P. trichocarpa* DXR, respectively. In contrast, the HbDXS1 antiserum failed to recognize the putative DXS protein in any plants event in *H. brasiliensis* (Fig. 3.37A, lane 3), suggesting that the amount of putative DXS protein in *H. brasiliensis* may be low. This is the preliminary characterization of the produced polyclonal antibodies. More experiments need to be performed in the future such as purification of polyclonal antibodies before being used, and reduction of abundant proteins, e.g. Rubisco proteins to increase the resolution of detection of low-abundance proteins.

CHAPTER 5

CONCLUSIONS

1. Molecular cloning and characterization of cDNAs encoding DXS

Two-full length cDNAs encoding for DXS have been isolated from *H. brasiliensis* using the RT-PCR and RACE based methods. The *HbDXS1* cDNA was isolated from leaves, while the *HbDXS2* cDNA was from the latex. The *HbDXS1* cDNA contains an ORF of 2163 bp encoding a polypeptide of 720 amino acids with a predicted molecular mass and pI values of 78 kDa and 7.66. The *HbDXS2* cDNA contains an ORF of 2136 bp encoding a polypeptide of 711 amino acids with predicted molecular mass and pI values of 76 kDa and 7.34. The nucleotide and deduced amino acid sequences of both *HbDXS*s have been deposited to GenBank database under the accession nos. AY502939 (for *HbDXS1*) and DQ473433 (for *HbDXS2*).

The deduced amino acid sequences of both HbDXSs showed 69% identity to each other and they shared high sequence homology with other plant DXSs, with the highest sequence identity to *P. montana* (87%, for HbDXS1) and to *C. roseus* (84%, for HbDXS2). Multiple alignments revealed that the deduced amino acid sequences of both HbDXSs contain all typical characteristics of DXS, such as a TPP-binding motif (GDG-X_{25}-NDN, residues 216 to 246 in the HbDXS1 sequence and residues 209 to 239 in the HbDXS2 sequence), and a transketolase motif (DRAG-X_{28}-P-X-D, residues 498 to 532 in HbDXS1, and residues 491 to 525 in HbDXS2). Several conserved amino acids important for DXS activity (H114, E449, R477, and R557 in the HbDXS1 sequence, and H107, E442, R470, and R550 in the HbDXS2 sequence) are also conserved in both HbDXSs.

The NH_2-terminal extension regions of both HbDXSs show common characteristics of a cTP sequence in plants. The possible cTP cleavage sites were predicted to be recognized between amino acids 57 and 58 in the HbDXS1 sequence and amino acids 34 and 35 in the HbDXS2 sequence. Most plant DXSs (16 out of 24 DXS sequences or about 67%) shares a similar cleavage site motif 'Xaa↓Ala-Ser-Xaa'. After cleavage these sequences, the putative mature protein of HbDXS1 was predicted to

contain 663 amino acids with a molecular weight of 71 kDa and a pI value of 6.4, and HbDXS2was predicted to contain 663 amino acids with a molecular weight of 71 kDa and a pI value of 6.4, and to contain 677 amino acids with a molecular weight of 72 kDa and a pI value of 6.56. A phylogenetic analysis showed that both HbDXSs are grouped within the plant DXS family but branch in different groups. The HbDXS1 is more closely related to other plant DXSs belonging to class I, while the HbDXS2 is close to class II of the plant DXS family. The prediction of 3D structures of both HbDXSs revealed that the predicted 3D models are similar to that of the *D. radiodurans* DXS template. There are three domains: I, II, and III in the modeled monomeric subunit.

2. Molecular cloning and characterization of cDNAs encoding DXR

The RT-PCR and RACE methods were also used to isolate a full length cDNA encoding DXR in *H. brasiliensis*. In this study, two-full length *DXR* cDNAs were obtained. The *HbDXR1* cDNA was isolated from leaves, while the *HbDXR2* cDNA was from the latex. Both *HbDXRs* cDNAs contain an ORF of 1416 bp encoding a polypeptide of 471 amino acids with a predicted molecular mass of about 51 kDa. The predicted pI values of HbDXR1 and HbDXR2 were 5.67 and 6.04, respectively. The nucleotide and deduced amino acid sequences of both *HbDXR*s cDNAs have been submitted to GenBank database under the accession nos. AY502937 (for *HbDXR1*) and DQ473432 (for *HbDXR2*).

The deduced amino acid sequences of both HbDXRs exhibited 94% identity to each other at both nucleotide and amino acid levels. The two HbDXRs showed high sequence homology with other plant DXRs, with the highest identity to *L. esculentum* (88% identity). Multiple alignments revealed that the deduced amino acid sequences of both HbDXRs also contain all typical characteristics of plant DXRs, such as a Proline-rich region $P_{2\text{-}4}$AWPG(R/T) with unknown function in plants (residues 53-61 in both HbDXRs), and a NADPH-binding site GXXGXXG (residues 81-87 in both HbDXRs). In addition, four acidic amino acids: D223, E225, E294, and E297 important for the DXR activity are conserved in both HbDXRs. The invariant amino acids required for a DXP binding, such as K198, S249, H272, W275, S285, N290, and K291, are also conserved in both HbDXRs.

Like HbDXSs proteins, the NH_2-terminal extension regions of both HbDXRs were predicted to have common features of cTP sequences in plants. The cTP recognition sites in both HbDXRs were predicted to be localized between amino acids 48 and 49. All plant DXRs analyzed so far show a similar consensus sequence for cTP cleavage sites. After cleavage of the cTP sequences, the putative mature HbDXRs proteins were predicted to contain 423 amino acids with a molecular mass of 46 kDa. The predicted pI values of HbDXR1 and HbDXR2 were 5.24 and 5.45, respectively. A phylogenetic analysis showed that both HbDXRs form a monophyletic group with those from other plant DXRs. Southern blot analysis suggested that the *DXR* gene in *H. brasiliensis* is most likely to be encoded by a small gene family, with at least two members. The predicted 3D structures of both HbDXRs show the overall folding pattern is consistent with the *E. coli* DXR template. The modeled monomeric subunit contains three domains: an NH_2-terminal NADPH-binding domain, a central catalytic domain, and a COOH-terminal domain.

3. RNA expression analysis by semi-quantitative RT-PCR

Semi-quantitative RT-PCR analysis showed that the transcripts of both *HbDXSs* and *HbDXRs* genes are found to be present in all tissues analyzed; however, the transcripts of *HbDXS1* are hardly detectable in roots of seedlings and latex of mature plants. Both *HbDXSs* and *HbDXRs* genes show similar expression patterns in general, for example, the transcripts of both *HbDXS1* and *HbDXR1* are highly expressed in photosynthetic tissues such as leaves, while the transcripts of both *HbDXS2* and *HbDXR2* are highly expressed in non-photosynthetic tissues such as roots, flowers, and latex, with an exception of higher expression of *HbDXR2* transcripts in leaves.

The expression level of *HbDXR2* transcripts was found to be higher in *H. brasiliensis* clone RRIM 600 (a high-latex yielding clone) than in the wild type plant (a low-latex yielding clone), and seems to be correlated with the latex yield, while the transcripts of *HbDXS2* are not significantly different between these two clones. Treatment of *H. brasiliensis* trees with ethephon, which generates ethylene *in vivo*, increases the rubber yield, while the levels of both *HbDXS2* and *HbDXR2* transcripts are transiently induced.

3. Expression of recombinant proteins and Western blot analysis

The *HbDXS1* and *HbDXR1* cDNAs were used for expression in *E. coli*, and the results showed that they are successfully expressed in *E. coli* but present as inclusion bodies. The recombinant HbDXS1 and HbDXR1 proteins purified under denaturing conditions using the nickel-based affinity chromatography show a major band corresponding to the molecular weight of about 70 and 40 kDa, respectively, consistent with their molecular sizes predicted from their nucleotide sequences, i.e. 71 kDa for HbDXS1 and 46 kDa for HbDXR1. The recovered recombinant HbDXR1 protein after refolding shows no enzymatic activity, as detected by an activity assay.

The recombinant HbDXS1 and HbDXR1 proteins were used for the production of antibodies. As judged by western blot analysis, both HbDXS1 and HbDXR1 antisera can detect their corresponding proteins expressed in *E. coli*. The HbDXS1 antiserum fails to detect the putative DXS proteins from *H. brasiliensis* and other plants, while the HbDXR1 antiserum can detect a putative DXR protein from *H. brasiliensis* and can cross-react with DXR proteins from other plants.

REFERENCES

Adam K.P. and Zapp J. 1998. Biosynthesis of the isoprene units of chamomile sesquiterpenes. *Phytochemistry* 48, 953-959.

Adiwilaga K. and Kush A. 1996. Cloning and characterization of cDNA encoding farnesyl diphosphate synthase from rubber tree (*Hevea brasiliensis*). *Plant Mol. Biol.* 30, 935-946.

Ahn C.S. and Pai H.S. 2008. Physiological function of *IspE*, a plastid MEP pathway gene for isoprenoid biosynthesis, in organelle biogenesis and cell morphogenesis in *Nicotiana benthamiana*. DOI: 10.1007/s11103-007-9286-0.

Altincicek B., Hintz M., Sanderbrand S., Wiesner J., Beck E. and Jomaa H. 2000. Tools for discovery of inhibitors of the 1-deoxy-D-xylulose 5-phosphate (DXP) synthase and DXP reductoisomerase: an approach with enzymes from the pathogenic bacterium *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* 190, 329-333.

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25, 3389-3402.

Archibald J.M., Rogers M.B., Toop M., Ishida K.I. and Keeling P.J. 2003. Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. *Proc. Natl. Acad. Sci. USA* 100, 7678-7683.

Argyrou A. and Blanchard J.S. 2004. Kinetic and chemical mechanism of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate isomeroreductase. *Biochemistry* 43, 4375-4384.

Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A. and Struhl K. 1995. Current Protocols in Molecular Microbiology. New York: Wiley.

Bailey A.M., Mahapatra S., Brennan P.J. and Crick D.C. 2002. Identification, cloning, purification, and enzymatic characterization of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate synthase. *Glycobiology* 12, 813-820.

Balasubramaniam S., Goldstein J.L. and Brown M.S. 1977. Regulation of cholesterol synthesis in rat adrenal gland through coordinate control of 3-hydroxy-3-methylglutaryl-coenzyme A synthase and reductase activity. *Proc. Natl. Acad. Sci. USA* 74, 1421-1425.

Barlow A.J., Becker H. and Adam K.P. 2001. Biosynthesis of the hemi- and monoterpene moieties of isoprenyl phenyl ethers from the liverwort *Trichocolea tomentella*. *Phytochemistry* 57, 7-14.

Bick J.A and Lange B.M. 2003. Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.* 415, 46-54.

Boucher Y. and Doolittle W. F. 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Mol. Microbiol.* 37, 703-716.

Bouvier F., d'Harlingue A., Suire C., Backhaus R.A. and Camara B. 1998. Dedicated roles of plastid transketolases during the early onset of isoprenoid biogenesis in pepper fruits. *Plant Physiol.* 117, 1423-1431.

Budziszewski G.J., Lewis S.P., Glover L.W., Reineke J., Jones G., Ziemnik L.S., Lonowski J., Nyfeler B., Aux G., Zhou Q., McElver J., Patton D.A., Martienssen R., Grossniklaus U., Ma H., Law M. and Levin J.Z. 2001. Arabidopsis genes essential for seedling viability: isolation of insertional mutants and molecular cloning. *Genetics* 159, 1765-1778.

Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L., Geoghagen N.S. and Venter J.C. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273, 1058-1073.

Burlat V., Oudin A., Courtois M., Rideau M. and St-Pierre B. 2004. Co-expression of three MEP pathway genes and geraniol 10-hydroxylase in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of intermediates during the biosynthesis of monoterpene indole alkaloids and isoprenoid-derived primary metabolites. *Plant J.* 38, 131-141.

Cane D.E., Chow C., Lillo A. and Kang I. 2001. Molecular cloning, expression, and characterization of the first three genes in the mevalonate-independent isoprenoid pathway in *Streptomyces coelicolor*. *Bioorg. Med. Chem.* 9, 1467-1477.

Carretero-Paulet L., Ahumada I., Cunillera N., Rodríguez-Concepción M., Ferrer A., Boronat A. and Campos N. 2002. Expression and molecular analysis of the Arabidopsis DXR gene encoding 1-deoxy-D-xylulose 5-phosphate

reductoisomerase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. *Plant Physiol.* 129, 1581-1591.

Carretero-Paulet L., Cairó A., Botella-Pavía P., Besumbes O., Campos N., Boronat A. and Rodríguez-Concepción M. 2006. Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. *Plant Mol. Biol.* 62, 683-695.

Celniker S.E. and Rubin G.M. 2003. The *Drosophila melanogaster* genome. *Annu. Rev. Genomics Hum. Genet.* 4, 89-117.

Chahed K., Oudin A., Guivarc'h N., Hamdi S., Cheniéux J.C., Rideau M. and Clastre M. 2000. 1-Deoxy-D-xylulose 5-phosphate synthase from periwinkle: cDNA identification and induced gene expression in terpenoid indole alkaloid-producing cells. *Plant Physiol. Biochem.* 38, 559-566.

Charlwood B.V. and Banthorpe D.V. 1978. The biosynthesis of monoterpenes. *Prog. Phytochem.* 5, 65-125.

Charon L., Hoeffer J.F., Pale-Grosdemange C., Lois L.M., Campos N., Boronat A., and Rohmer M. 2000. Deuterium labeled isotopomers of 2-*C*-methyl-D-erythritol as tools for the elucidation of the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway for isoprenoid synthesis. *Biochem. J.* 346, 737-742.

Chen S. and Halkier B.A. 1999. Functional expression and characterization of the myrosinase MYR1 from *Brassica napus* in *Saccharomyces cerevisiae*. *Protein Expr. Purif.* 17, 414-420.

Chow K.S., Wan K.L., Noor Mat Isa M., Bahari A., Tan S.H., Harikrishna K. and Yeang H.Y. 2007. Insights into rubber biosynthesis from transcriptome analysis of *Hevea brasiliensis* latex. *J. Exp. Bot.* 58, 2429-2440.

Chye M.L., Kush A., Tan C.T. and Chua N.H. 1991. Characterization of cDNA and genomic clones encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase from *Hevea brasiliensis*. *Plant Mol. Biol.* 16, 567-577.

Chye M.L., Tan C.T. and Chua N.H. 1992. Three genes encode 3-hydroxy-3-methylglutaryl-CoA reductase in *Hevea brasiliensis*: *hmg1* and *hmg3* are differentially expressed. *Plant Mol. Biol.* 19, 473-484.

Cornish K. and Scott D.J. 2005. Biochemical regulation of rubber biosynthesis in guayule (*Parthenium argentatum* Gray). *Ind. Crops Prod.* 22, 49-58.

Coupé M. and Chrestin H. 1989. Physiochemical and biochemical mechanisms of hormonal (ethylene) stimulation. In: d'Auzac J, Jacob J.L. and Chrestin H., Editors, *Physiology of Rubber Tree Latex*, CRC Press, Boca Raton, pp. 295-319.

d'Auzac J. and Jacob J.L. 1989. The composition of latex from *Hevea brasiliensis* as a laticiferous cytoplasm. In: d'Auzac J, Jacob J.L. and Chrestin H., Editors, *Physiology of Rubber Tree Latex*, CRC Press, Boca Raton, pp. 59-96.

De-Eknamkul W. and Potduang B. 2003. Biosynthesis of β-sitosterol and stigmasterol in Croton sublyratus proceeds *via* a mixed origin of isoprene units. *Phytochemistry* 62, 389-398.

de Faÿ E. and Jacob J.L. 1989. Anatomical organization of the laticiferous system in the bark. In: d'Auzac J, Jacob J.L. and Chrestin H., Editors, *Physiology of Rubber Tree Latex*, CRC Press, Boca Raton, pp. 3-14.

de Faÿ E., Hébant C.H. and Jacob J.L. 1989. Cytology and cytochemistry of the laticiferous system. In: d'Auzac J, Jacob J.L. and Chrestin H., Editors, *Physiology of Rubber Tree Latex*, CRC Press, Boca Raton, pp. 15-30.

de Marco A., Volrath S., Bruyere T., Law M. and Fonnè-Pfister R. 2000. Recombinant maize protoporphyrinogen IX oxidase expressed in *Escherichia coli* forms complexes with GroEL and DnaK chaperones. *Prot. Exp. Purif.* 20, 81-86.

de Marco A. and De Marco V. 2004 A. Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *J. Biotechnol.* 109, 45-52.

Dennis M.S. and Light D.R. 1989. Rubber elongation factor from *Hevea brasiliensis*: identification, characterization and role in rubber biosynthesis. *J. Biol. Chem.* 264, 18608-18617.

Dhiman R.K., Schaeffer M.L., Bailey A.M., Testa C.A., Scherman H. and Crick D.C. 2005. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase (IspC) from *Mycobacterium tuberculosis*: Towards understanding mycobacterial resistance to fosmidomycin. *J. Bacteriol.* 187, 8395-8402.

Dickenson PB. 1969. Electron microscopical studies of latex vessel system of *Hevea brasiliensis*. *J. Rubber Res. Inst. Malaya* 21, 543-551.

Disch A. and Rohmer M. 1998. On the absence of the glyceraldehyde 3-phosphate/pyruvate pathway for isoprenoid biosynthesis in fungi and yeasts. *FEMS Microbiol. Lett.* 168, 201-208.

Disch A., Schwender J., Mueller C., Lichtenthaler H.K. and Rohmer, M. 1998. Distribution of the mevalonate and glyceraldehyde phosphate/pyruvate pathways for isoprenoid biosynthesis in unicellular algae and the cyanobacterium *Synechocystis* PCC 6714. *Biochem. J.* 333, 381-388.

Doolittle W.F., Boucher Y., Nesbo C.L., Douady C.J., Andersson J.O. and Roger A.J. 2003. How big is the iceberg of which organellar genes in nuclear genomes are but the tip?. *Phil. Trans. Roy. Soc. ser. B Biol. Sci.* 358, 39-58.

Dubey V.K., Bhalla R. and Luthra R. 2003. An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *J. Biosci.* 28, 637-646.

Dudareva N., Andersson S., Orlova I., Gatto N., Reichelt M., Rhodes D. Boland W. and Gershenzon J. 2005. The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. *Proc. Natl. Acad. Sci. USA* 102, 933-938.

Duvold T., Cali P., Bravo J.M. and Rohmer M. 1997. Incorporation of 2-*C*-methyl-D-erythritol, a putative isoprenoid precursor in the mevalonate-independent pathway, into ubiquinone and menaquinone of *Escherichia coli*. *Tetrahedron Lett.* 38, 6181-6184.

Eisenreich W., Menhard B., Hylands P.J., Zenk M.H. and Bacher A. 1996. Studies on the biosynthesis of taxol: the taxane carbon skeleton is not of mevalonoid origin (NMR spectroscopy/*Taxus chinensis*/plant cell culture/terpene). *Biochemistry* 93, 6431-6436.

Eisenreich W., Sagner S., Zenk M.H. and Bacher A. 1997. Monoterpenoid essential oils are not of mevalonoid origin. *Tetrahedron Lett.* 38, 3889-3892.

Eisenreich W., Schwarz M, Cartayrade A., Arigoni D., Zenk M.H. and Bacher A. 1998. The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chem. Biol.* 5, 221-233.

Eisenreich W., Rohdich F. and Bacher A. 2001. Deoxyxylulose phosphate pathway to terpenoids. *Trends Plant Sci.* 6, 78-84.

Eisenreich W., Bacher A., Arigoni D. and Rohdich F. 2004. Biosynthesis of isoprenoids *via* the non-mevalonate pathway. *Cell. Mol. Life Sci.* 61, 1401-1426.

Emanuelsson O., Nielsen H. and von Heijne G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978-984.

Engprasert S., Taura F. and Shoyama Y. 2005. Molecular cloning, expression and characterization of recombinant 1-deoxy-D-xylulose-5-phosphate reducto-isomerase from *Coleus forskohlii* Briq. *Plant Sci.* 169, 287-294.

Estévez J.M., Cantero A., Romero C., Kawaide H., Jiménez L.F., Kuzuyama T. Seto H. Kamiya Y. and León P. 2000. Analysis of the expression of *CLA1*, a gene that encodes the 1-deoxy-D-xylulose 5-phosphate synthase of the 2-*C*-methyl-D-erythritol-4-phosphate pathway in *Arabidopsis*. *Plant Physiol.* 124, 95-103.

Estévez J.M., Cantero A., Reindl A., Reichler S. and León P. 2001. 1-Deoxy-D-xylulose 5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J. Biol. Chem.* 22, 22901-22909.

Eubanks L.M. and Poulter C.D. 2003. *Rhodobacter capsulatus* 1 deoxy-D-xylulose 5-phosphate synthase: steady-state kinetics and substrate binding. *Biochemistry* 42, 1140-1149.

Farrell B.D., Dessourd D.E. and Mitter C. 1991. Escalation of plant defenses: do latex and resin canals spur plant diversification? *Am. Nat.* 138, 881-900.

Fernandes R.P.M., Phaosiri C. and Proteau P.J. 2005. Mutation in the flexible loop of 1-deoxy-D-xylulose 5-phosphate reductoisomerase broadens substrate utili-zation. *Arch. Biochem. Biophys.* 444, 159-164.

Fernandes R.P.M. and Proteau P.J. 2006. Kinetic characterization of *Synechocystis* sp. PCC6803 1-deoxy-D-xylulose 5-phosphate reductoisomerase mutants. *Biochim. Biophys. Acta* 1764, 223-229.

Flesch G. and Rohmer M., 1988. Prokaryotic hopanoids: the biosynthesis of the bacteriohopan skeleton. *Eur. J. Biochem.* 175, 405-411.

Flournoy D.S. and Frey P.A. 1989. Inactivation of the pyruvate dehydrogenase complex of *Escherichia coli* by fluoropyruvate. *Biochemistry* 28, 9594-9602.

Flügge U.I. and Gao W. 2005. Transport of isoprenoid intermediates across chloroplast envelope membranes. *Plant. Biol.* 7, 91-97.

Gaberc-Porekar V. and Menart V. 2001. Perspectives of immobilized-metal affinity chromatography. *J. Biochem. Biophys. Methods* 49, 335-360.

Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D. and Bairoch A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucl. Acids Res.* 31, 3784-3788.

Gavel Y. and von Heijne G. 1990. A conserved cleavage-site motif in chloroplast transit peptides. *FEBS Lett.* 261, 455-458.

Georgiou G. and Valax P. 1996. Expression of correctly folded proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* 7, 190-197.

Giaever G., Chu A.M., Ni L., Connelly C., Riles L., Veronneau S., et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387-391.

Gil M.J., Coego A., Mauch-Mani B., Jorda L. and Vera P. 2005. The Arabidopsis *csb3* mutant reveals a regulatory link between salicylic acid-mediated disease resistance and the methyl-erythritol 4-phosphate pathway. *Plant J.* 44, 155-166.

Goldstein J.L. and Brown M.S. 1990. Regulation of the mevalonate pathway. *Science* 343, 425-430.

Gong Y.F., Liao Z.H., Chen M., Zuo K.J., Guo L., Tan Q.M., Huang Z.S., Kai G.Y., Sun X.F. Tan F. and Tang K.X. 2005. Molecular cloning and characterization of a 1-deoxy-D-xylulose 5-phosphate reductoisomerase gene from *Ginkgo biloba*. *DNA Seq.* 16, 111-120.

Gong Y.F., Liao Z.H., Guo B.H., Sun X.F. and Tang K.X. 2006. Molecular cloning and expression profile analysis of *Ginkgo biloba DXS* gene encoding 1-deoxy-D-xylulose-5-phosphate synthase, the first committed enzyme of the 2-*C*-methyl-D-erythritol 4-phosphate pathway. *Planta Med.* 72, 329-335.

Grolle S., Bringer-Meyer S. and Sahm H. 2000. Isolation of the *dxr* gene of *Zymomonas mobilis* and characterization of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *FEMS Microbiol. Lett.* 191, 131-137.

Gu Z., Steinmetz L.M., Gu X., Scharfe C., Davis R.W. and Li W.H. 2003. Role of duplicate genes in genetic robustness against null mutations. *Nature* 421, 63-66.

Guevara-García A., San Román C., Arroyo A., Cortés M.E., Gutiérrez-Nava M.L. and León P. 2005. Characterization of the Arabidopsis *clb6* mutant illustrates the importance of posttranscriptional regulation of the methyl-d-erythritol 4-phosphate pathway. *Plant Cell* 17, 628-643.

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

Gutiérrez-Nava M.L., Gillmor C.S., Jiménez L.F., Guevara-García A. and León P. 2004. CHLOROPLAST BIOGENESIS genes act cell and noncell autonomously in early chloroplast development. *Plant Physiol.* 135, 471-482.

Hahn F.M., Eubanks L.M., Testa C.A., Blagg B.S.J., Baker J.A. and Poulter C.D. 2001. 1-Deoxy-D-xylulose 5-phosphate synthase, the gene product of open reading frame (ORF) 2816 and ORF 2895 in *Rhodobacter capsulatus*. *J. Bacteriol.* 183, 1-11.

Hampel D., Mosandl A. and Wust M. 2005. Biosynthesis of mono- and sesquiterpenes in carrot roots and leaves (*Daucus carota* L.): metabolic cross talk of cytosolic mevalonate and plastidial methylerythritol phosphate pathways. *Phytochemistry* 66, 305-311.

Han Y.S., Heijden R.V.D., Lefeber A.W.M., Erkelens C. and Verpoorte R. 2002. Biosynthesis of anthraquinones in cell cultures of Cinchona 'Robusta' proceeds via the methylerythritol 4-phosphate pathway. *Phytochemistry* 59, 45-55.

Han Y.S., Roytrakul S., Verberne M.C., Heijden R.V.D., Linthorst H.J.M. and Verpoorte R. 2003. Cloning of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate synthase from *Morinda citrifolia* and analysis of its expression in relation to anthraquinone accumulation. *Plant Sci.* 164, 911-917.

Hans J., Hause B., Stack D. and Walter M.H. 2004. Cloning, characterization, and immunolocalization of a mycorrhiza-inducible 1-deoxy-D-xylulose 5-phosphate reductoisomerase in arbuscule-containing cells of maize. *Plant Physiol.* 134, 614-624.

Hawkins C.F., Borges A. and Perham R.N. 1989. A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS Lett.* 255, 77-82.

Hemmerlin A., Hoeffler J.-F., Meyer O., Tritsch D., Kagan I.A., Grosdemange-Billiard C., Rohmer M. and Bach T. J. 2003. Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells. *J. Biol. Chem.* 278, 26666-26676.

Henriksson L.M.,Bjorkelid C., Mowbray S.L. and Unge T. 2006. The 1.9 Å resolution structure of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate reductoisomerase, a potential drug target. *Acta Cryst.* 62, 807-813.

Hepper C.M. and Audley B.G. 1969. The biosynthesis of rubber from hydroxyl-methyl-glutaryl coenzyme A in *Hevea brasiliensis* latex. *Biochem. J.* 114, 379-386.

Hertewich U. Zapp J., Becker H. and Adam K.P. 2001. Biosynthesis of a hopane triterpene and three diterpenes in the liverwort *Fossombronia alaskana*. *Phytochemistry* 58, 1049-1054.

Himmeldirk K., Kennedy I.A., Hill R.E., Sayer B.G. and Spenser I.D. 1996. Biosynthesis of vitamins B-1 and B-6 in *Escherichia coli*: concurrent incorporation of 1-deoxy-D-xylulose into thiamin (B-1) and pyridoxol (B-6). *J. Chem. Soc. Chem. Commun.* 10, 1187-1188.

Hoeffler J.F., Tritsch D., Grosdemange-Billiard C. and Rohmer M. 2002. Isoprenoid biosynthesis *via* the methylerythritol phosphate pathway: mechanistic investigations of the 1-deoxy-d-xylulose 5-phosphate reductoisomerase. *Eur. J. Biochem.* 269, 446-457.

Holz C., Hesse O., Bolotina N., Stahl U. and Lang C. 2002. A micro-scale process for high-throughput expression of cDNAs in the yeast *Saccharomyces cerevisiae*. *Protein Expr. Purif.* 25, 372-378.

Hooley R. 1994. Gibberellins: perception, transduction and responses. *Plant Mol. Biol.* 26, 1529-1555.

Hsieh M.H. and Goodman H.M. 2005. The Arabidopsis *IspH* homolog is involved in the plastid nonmevalonate pathway of isoprenoid biosynthesis. *Plant Physiol.* 138, 641-653.

Hsieh M.H. and Goodman H.M. 2006. Functional evidence for the involvement of Arabidopsis *IspF* homolog in the nonmevalonate pathway of plastid isoprenoid biosynthesis. *Planta* 223, 779-784.

Hsieh M.H., Chang C.Y., Hsu S.J. and Chen J.J. 2008. Chloroplast localization of methylerythritol 4-phosphate pathway enzymes and regulation of mitochondrial genes in *ispD* and *ispE* albino mutants in Arabidopsis. *Plant Mol. Biol.* DOI: 10.1007/s11103-007-9286-0.

Huang X. and Madan A. 1999. CAP3: a DNA sequence assembly program. *Genome Res.* 9, 868-877.

Husselstein-Muller T., Schaller H. and Benveniste P. 2001. Molecular cloning and expression in yeast of 2,3-oxidosqualene-triterpenoid cyclases from *Arabidopsis thaliana*. *Plant Mol. Biol.* 45, 75-92.

Itoh D. Karunagoda R.P. Fushie T. Katoh K. and Nabeta K. 2000. Nonequivalent labeling of the phytyl side chain of chlorophyll a in callus of the hornwort *Anthoceros punctatus*. *J. Nat. Prod.* 63, 1090-1093.

Ishida, K. I., Cao Y., Hasegawa M., Okada N. and Hara Y. 1997. The origin of chlorarachinophyte plastids, as inferred from phylogenetic comparisons of amino acid sequences of EF-Tu. *J. Mol. Evol.* 45, 682-687.

Jacob J.L., Prévôt J.C. and Kekwick R.G.O. 1989. General metabolism of *Hevea brasiliensis*. In: d'Auzac J., Jacob J.L. and Chrestin H., Editors, *Physiology of Rubber Tree Latex*, CRC Press, Boca Raton, pp. 101-144.

Jennewein S., Wildung M.R., Chau M., Walker K. and Croteau R. 2004. Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc. Natl. Acad. Sci. USA* 101, 9149-9154.

Jomaa H., Wiesner J., Sanderbrand S., Altincicek B., Weidemeyer C., Hintz M., Turbachova I., Eberl M., Zeidler J., Lichtenthaler H.K., Soldati D. and Beck E. 1999. Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* 285, 1573-1576.

Kekwick R.G.O. 1989. The formation of isoprenoids in *Hevea* latex. In: d'Auzac J, Jacob J.L. and Chrestin H., Editors, *Physiology of Rubber Tree Latex*, CRC Press, Boca Raton, pp. 145-164.

Khemvong S. and Suvachittanont W. 2005. Molecular cloning and expression of a cDNA encoding 1-deoxy-D-xylulose-5phosphate synthase from oil palm *Elaeis guineensis Jacq*. *Plant Sci*. 169, 571-578.

Kim B.R., Kim S.U. and Chang Y.J. 2005. Differential expression of three 1-deoxy-D-xylulose-5-phosphate synthase genes in rice. *Biotechnol. Lett.* 27, 997-1001.

Kim S.M., Kuzuyama T., Chang Y.J., Song K.S. and Kim S.U. 2006. Identification of class 2 1-deoxy-D-xylulose 5-phosphate synthase and 1-deoxy-D-xylulose 5-phosphate reductoisomerase genes from *Ginkgo biloba* and their transcription in embryo culture with respect to ginkgolide biosynthesis. *Plant Med.* 72, 234-240.

Kishimoto S. and Ohmiya A. 2006. Regulation of carotenoid biosynthesis in petals and leaves of chrysanthemum (*Chrysanthemum morifolium*). *Physiol. Plantarum* 128, 436-447.

Klenk H.P., Clayton R.A., Tomb J.F., White O., Nelson K.E., Ketchum K.A., et al. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* 390, 364-370.

Ko J.H., Chow K.S. and Han K.H. 2003. Transcriptome analysis reveals novel features of the molecular events occurring in the laticifers of *Hevea brasiliensis* (para rubber tree). *Plant Mol. Biol.* 53, 479-492.

Koonin E.V. 2000. How many genes can make a cell: the minimal-gene-set concept. *Annu. Rev. Genomics Hum. Genet.* 1, 99-116.

Koppisch A.T., Fox D.T., Blagg B.S.J. and Poulter C.D. 2002. *E. coli* MEP synthase: steady-state kinetic analysis and substrate binding. *Biochemistry* 41, 236-243.

Krushkal J, Pistilli M., Ferrell K.M., Souret F.F. and Weathers P.J. 2003. Computational analysis of the evolution of the structure and function of 1-deoxy-D-xylulose-5-

phosphate synthase, a key regulator of the mevalonate-independent pathway in plants. *Gene* 313, 127-138.

Kumar S., Tamura K. and Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform.* 5, 150-163.

Kush A., Goyvaerts E., Chye M.L. and Chua N.H. 1990. Laticifer specific gene expression in *Hevea brasiliensis* (rubber tree). *Proc. Natl. Acad. Sci. USA* 87, 1787-1790.

Kuzuyama T., Shimizu T., Takahashi S. and H. Seto. 1998. Fosmidomycin, a specific inhibitor of 1-deoxy-D-xylulose 5-phosphate reductoisomerase in the nonmevalonate pathway for terpenoid biosynthesis. *Tetrahedron Lett.* 39, 7913-7916.

Kuzuyama T., Takahashi S., Takagi M. and Seto H. 2000a. Characterization of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, an enzyme involved in isopentenyl diphosphate biosynthesis, and identification of its catalytic amino acid residues. *J. Biol. Chem.* 275, 19928-19932.

Kuzuyama T., Takagi M., Takahashi S. and Seto H. 2000b. Cloning and characterization of 1-deoxy-D-xylulose 5-phosphate synthase from *Streptomyces* sp. strain CL190, which uses both the mevalonate and nonmevalonate pathways for isopentenyl diphosphate biosynthesis. *J. Bacteriol.* 182, 891-897.

Laemmli U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.

Lange B.M., Wildung M.R., McCaskill D. and Croteau R. 1998. A family of transketolases that directs isoprenoid biosynthesis *via* a mevalonate-independent pathway. *Proc. Natl. Acad. Sci. USA* 95, 2100-2104.

Lange B.M. and Croteau R. 1999. Isoprenoid biosynthesis *via* a mevalonate-independent pathway in plants: cloning and heterologous expression of 1-deoxy-D-xylulose-5-phosphate reductoisomerase from peppermint. *Arch. Biochem. Biophys.* 365, 170-174.

Lange B.M., Rujan T., Martin W. and Croteau R. 2000. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. USA* 97, 13172-13177.

Laule O., Fuerholz A., Chang H. S., Zhu T., Wang X., Heifetz P.B., Gruissem W. and Lange M. 2003. Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 100, 6866-6871.

Lee J.K., Oh D.K. and Kim S.Y. 2007. Cloning and characterization of the *dxs* gene, encoding 1-deoxy-D-xylulose 5-phosphate synthase from *Agrobacterium tumefaciens*, and its overexpression in *Agrobacterium tumefaciens*. *J. Biotechnol.* 128, 555-566.

Lell B., Ruangweerayut R., Wiesner J., Missinou M.A., Schindler A., Baranek T., Hintz M., Hutchinson D., Jomaa H. and Kremsner P.G. 2003. Fosmidomycin, a novel chemotherapeutic agent for malaria. *Antimicrob. Agents Chemother.* 47, 735-738.

Lesk, A.M. 1995. NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* 5, 775-783.

Lewis L.A. and McCourt R.M. 2004. Green algae and the origin of land plants. *Am. J. Bot.* 91, 1535-1556.

Li S.M., Wemakor E. and Heide L. 2001. PCR amplification of 1-deoxy-D-xylulose 5-phosphate synthase (*dxs*) genes from different *Streptomyces* species: evidence for the existence of two *dxs* gene families. *Mol. Biol. Today* 2, 61-66.

Li L., Wang X., Gai J. and Yu D. 2007. Molecular cloning and characterization of a novel microsomal oleate desaturase gene from soybean. *J. Plant Physiol.* 164, 1516-1526.

Liao Z., Chen R., Chen M., Yang C., Wang Q. and Gong Y. 2007. A new 1-deoxy-D-xylulose 5-phosphate reductoisomerase gene encoding the committed-step enzyme in the MEP pathway from *Rauvolfia verticillata*. *Z. Naturforsch. [C]*. 62, 296-304.

Lichtenthaler H.K., Schwender J., Disch A. and Rohmer M. 1997. Biosynthesis of isoprenoids in higher plant chloroplasts proceeds *via* a mevalonate-independent pathway. *FEBS Lett.* 400, 271-274.

Lichtenthaler HK. 1999. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu. Rev. Plant Mol. Biol.* 50, 47-65.

Lindqvist Y., Schneider G., Ermler U. and Sundstrom M. 1992. Three-dimensional structure of transketolase, a thiamine diphosphate dependent enzyme, at 2.5 Å resolution. *EMBO J.* 11, 2373-2379.

Lois L.M., Campos N., Putra S.R., Danielsen K., Rohmer M. and Boronat A. 1998. Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxylylulose 5-

phosphate, a common precursor for isoprenoid, thiamine, and pyridoxol biosynthesis. *Proc. Natl. Acad. Sci. USA* 95, 2105-2110.

Lois L.M., Rodríguez-Concepcíón M., Gallego F., Campos N. and Boronat A. 2000. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant J.* 22, 503-513.

Lund O., Nielsen M., Lundegaard C. and Worning P. 2002. CPHmodels 2.0: X3M a computer program to extract 3D Models. In Abstract at the CASP5 conference, p. A102.

Lynen F. 1969. Biochemical problems of rubber synthesis. *J. Rubb. Res. Inst. Malaya* 21, 389-406.

MacSweeney A., Lange R., Fernandes R.P., Schulz H., Dale G.E., Douangamath A., Proteau J.P. and Oefner C. 2005. The crystal structure of *E. coli* 1-deoxy-D-xylulose-5-phosphate reductoisomerase in a ternary complex with the antimalarial compound fosmidomycin and NADPH reveals a tight-binding closed enzyme conformation. *J. Mol. Biol.* 345, 115-127.

Mahmoud S.S. and Croteau R.B. 2001. Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. *Proc. Natl. Acad. Sci. USA* 98, 8915-8920.

Mandel M.A., Feldmann K.A., Herrera-Estrella L., Rocha-Sosa M. and León P. 1996. CLA1, a novel gene required for chloroplast development, is highly conserved in evolution. *Plant J.* 9, 649-658.

Martin W. and Herrmann R.G. 1998. Gene transfer from organelles to the nucleus: How much, what happens and why? *Plant Physiol.* 118, 9-17.

Martin W. and Mueller M. 1998. The hydrogen hypothesis for the first eukaryote. *Nature* 392, 37-41.

Matsuzaki M., Misumi O., Shin-I T., Maruyama S., Takahara M., Miyagishima S.Y., Mori T., Nishida K., Yagisawa F., Nishida K., Yoshida Y., Nishimura Y., Nakao S., Kobayashi T., Momoyama Y., Higashiyama T., Minoda A., Sano M., Nomoto H., Oishi K., Hayashi H., Ohta F., Nishizaka S., Haga S., Miura S., Morishita T., Kabeya Y., Terasawa K., Suzuki Y., Ishii Y., Asakawa S., Takano H., Ohta N., Kuroiwa H., Tanaka K., Shimizu N., Sugano S., Sato N., Nozaki H., Ogasawara N., Kohara Y. and Kuroiwa T. 2004. Genome sequence of the ultra small unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428, 653-657.

Mayer B.J. and Baltimore D. 1993. Signalling through SH2 and SH3 domains. *Trends Cell. Biol.* 3, 8-13.

Mayer B.J. 1999. Protein-protein interactions in signaling cascades. *Mol. Biotechnol.* 13, 201-213.

McCaskill D. and Croteau R. 1995. Monoterpene and sesquiterpene biosynthesis in glandular trichomes of peppermint (*Mentha x piperita*) rely exclusively on plastid-derived isopentenyl diphosphate. *Planta* 197, 49-56.

McFadden G.I. 1999. Endosymbiosis and evolution of the plant cell. *Curr. Opin. Plant. Biol.* 2, 513-519.

Meyer O., Grosdemange-Billiard C., Tritsch D. and Rohmer M. 2003. Isoprenoid biosynthesis via the MEP pathway. Synthesis of (3*R*,4*S*)-3,4-dihydroxy-5-oxohexylphosphonic acid, an isosteric analogue of 1-deoxy-D-xylulose 5-phosphate, the substrate of the 1-deoxy-D-xylulose 5-phosphate reducto-isomerase. *Org. Biomol. Chem.* 1, 4367-4372.

Miller B., Heuser T. and Zimmer W. 1999. A *Synechococcus leopoliensis* SAUG 1402-1 operon harboring the 1-deoxyxylulose 5-phosphate synthase gene and two additional open reading frames is functionally involved in the dimethylallyl diphosphate synthesis. *FEBS Lett.* 460, 485-490.

Miller B., Heuser T. and Zimmer W. 2000. Functional involvement of a deoxy-D-xylulose 5-phosphate reductoisomerasegene harboring locus of *Synechococcus leopoliensis* in isoprenoid biosynthesis. *FEBS Lett.* 481, 221-226.

Moehs C.P., Tian L., Osteryoung K.W. and Dellapenna D. 2001. Analysis of carotenoid biosynthetic gene expression during marigold petal development. *Plant Mol. Biol.* 45, 281-293.

Mori H. and Cline K. 2001. Post-translational protein translocation into thylakoids by the Sec and pH-dependent pathways. *Biochim. Biophys. Acta* 1541, 80-90.

Morris W.L., Ducreux L.J., Hedden P., Millam S. and Taylor M.A. 2006. Overexpression of a bacterial 1-deoxy-D-xylulose 5-phosphate synthase gene in potato tubers perturbs the isoprenoid metabolic network: implications for the control of the tuber life cycle. *J. Exp. Bot.* 57, 3007-3018.

Müller C., Schwender J., Zeidler J. and Lichtenthaler H.K. 2000. Properties and inhibition of the first two enzymes of the non-mevalonate pathway of isoprenoid biosynthesis. *Biochem. Soc. Trans.* 28, 792-793.

Muñoz-Bertomeu J, Arrillaga I, Ros R. and Segura J. 2006. Up-Regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. *Plant Physiol.* 142, 890-900.

Nabeta K. Kawae T. Kikuchi T. Saitoh T. and Okuyama H. 1995. Biosynthesis of chlorophyll a from ^{13}C-labelled mevalonate and glycine in liverwort: nonequivalent labeling of phytyl side chain. *J. Chem. Soc. Chem. Commun.* 2529-2530.

Nicholas K.B., Nicholas H.B. Jr. and Deerfield D.W. II. 1997. GeneDoc: analysis and visualization of genetic variation. *Embnet News* 4, 1-4.

Nilsson U., Meshalkina L., Lindqvist Y. and Schneider G. 1997. Examination of substrate binding in thiamin diphosphate-dependent transketolase by protein crystallography and site-directed mutagenesis. *J. Biol. Chem.* 272, 1864-1869.

Nishihara K., Kanemori M., Kitagawa M., Yanagi H., Yura T. 1998. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64, 1694-1699.

Nishiyama T., Fujita T., Shin-I T., Seki M., Nishide H., Uchiyama I., Kamiya A., Carninci P., Hayashizaki Y., Shinozaki K., Kohara Y. and Hasebe M. 2003. Comparative genomics of *Physcomitrella patens* gemetophytic transcriptome and *Arabidopsis thaliana*: Implication for land plant evolution. *Proc. Natl. Acad. Sci. USA* 100, 8007-8012.

Ober D. 2005. Seeing double: gene duplication and diversification in plant secondary metabolism. *Trends Plant Sci.* 10, 444-449.

Office of the rubber replanting aid funds. 2006. URL: http://www.thailandrubber.thai-gov.net/news-/number5.pdf

Oh S.K., Kang H., Shin D.H., Yang J. and Han K.H. 2000. Molecular cloning and characterization of a functional cDNA clone encoding isopentenyl diphosphate isomerase from *Hevea brasiliensis*. *J. Plant Physiol.* 157, 549-557.

Okada K., Kawaide H., Kuzuyama T., Seto H., Curtis I.S. and Kamiya Y. 2002. Antisense and chemical suppression of the nonmevalonate pathway affects ent-kaurene biosynthesis in Arabidopsis. *Planta* 215, 339-344.

Okuhara M., Kuroda Y., Goto T., Okamoto M., Terano H., Kohsaka M., Aoki H. and Imanaka H. 1980a. Studies on new phosphonic acid antibiotics: III. Isolation and characterization of FR-31564, FR-32863 and FR-33289. *J. Antibiot.* 33, 24-28.

Okuhara M., Kuroda Y., Goto T., Okamoto M., Terano H., Kohsaka M., Aoki H. and Imanaka H. 1980b. Studies on new phosphonic acid antibiotics. I. FR-900098, isolation and characterization. *J. Antibiot.* 33, 13-17.

Page J.E., Hause G., Raschke M., Gao W., Schmidt J., Zenk M.H. and Kutchan T.M. 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, 1401-1413.

Philippe H. and Douady C.J. 2003. Horizontal gene transfer and phylogenetics. *Curr. Opin. Microbiol.* 6, 498-505.

Phillips M.A., Walter M.H., Ralph S.G., Dabrowska P., Luck K., Uros E.M., Boland W., Strack D., Rodríguez-Concepción M., Bohlmann J. and Gershenzon J. 2007. Functional identification and differential expression of 1-deoxy-D-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). *Plant Mol. Biol.* 65, 243-257.

Picherskya E. and Ganga D.R. 2000. Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci.* 5, 439-445.

Piel J., Donath J., Bandemer K. and Boland W. 1998. Mevalonate-independent biosynthesis of terpenoid volatiles in plants: induced and constitutive emission of volatiles. *Angew. Chem. Int. Ed. Engl.* 37, 2478-2481.

Priya P., Venkatachalam P. and Thulaseedharan A. 2006. Molecular cloning and characterization of the rubber elongation factor gene and its promoter sequence from rubber tree (*Hevea brasiliensis*): a gene involved in rubber biosynthesis. *Plant Sci.* 171, 470-480.

Priya P., Venkatachalam P. and Thulaseedharan A. 2007. Differential expression pattern of rubber elongation factor (REF) mRNA transcripts from high and low yielding clones of rubber tree (*Hevea brasiliensis* Muell. Arg.). *Plant Cell Rep.* DOI: 10.1007/s00299-007-0402-z.

Pujade-Renaud V., Clement A., Perrot-Rechenman C., Prevot J.C., Chrestin H., Jacob J.L. and Guern J. 1994. Ethylene-induced increase in glutamine synthetase activity and mRNA levels in *Hevea brasiliensis* latex cells. *Plant Physiol.* 105, 127-132.

Quackenbush J., Liang F., Holt I., Pertea G. and Upton J. 2000. The TIGR gene indices: reconstruction and representation of expressed gene sequences. *Nucl. Acids Res.* 28, 141-145.

Querol J., Rodríguez-Concepción M., Boronat A. and Imperial S. 2001. Essential role of residue H49 for activity of *Escherichia coli* 1-deoxy-D-xylulose 5-phosphate synthase, the enzyme catalyzing the first step of the 2-*C*-methyl-D-erythritol-4-phosphate pathway for isoprenoid synthesis. *Biochem. Biophys. Res. Commun.* 289, 155-160.

Querol J., Grosdemange-Billiard C., Rohmer M., Boronat A. and Imperial S. 2002. Enzymatic synthesis of 1-deoxysugar-phosphates using *E. coli* 1-deoxy-D-xylulose 5-phosphate synthase. *Tetrahedron Lett.* 43, 8265-8268.

Reuter K., Sanderbrand S., Jomaa S., Wiesner J., Steinbrecher I., Beck E., Hintz M., Klebe G. and Stubbs M.T. 2002. Crystal structure of 1-deoxy-D-xylulose-5-phosphate reductoisomerase, a crucial enzyme in the non-mevalonate pathway of isoprenoid biosynthesis. *J. Biol. Chem.* 277, 5378-5384.

Ricagno S., Grolle S., Bringer-Meyer S., Sahm H., Lindqvist Y. and Schneider G. 2004. Crystal structure of 1-deoxy-D-xylulose-5-phosphate reductoisomerase from *Zymomonas mobilis* at 1.9 Å resolution. *Biochim. Biophys. Acta* 1698, 37-44.

Rieder C., Strauss G., Fuchs G., Arigoni D., Bacher A. and Eisenreich W. 1998. Biosynthesis of the diterpene verrucosan-2β-ol in the phototrophic eubacterium *Chloroflexus aurantiacus*. A retrobiosynthetic NMR study *J. Biol. Chem.* 273, 18099-18108.

Ro D.-K., Paradise E.M., Ouellet M., Fisher K.J., Newman KL., Ndungu J.M., Ho K.A., Eachus R.A., Ham T.S., Kirby J., Chang M.C.Y., Withers S.T., Shiba Y., Sarpong R. and Keasling J.D. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440, 940-943.

Rodríguez-Concepción M., Ahumada I., Diez-Juez E., Sauret-Güeto S., Lois L.M., Gallego F., Carretero-Paulet L., Campos N. and Boronat A. 2001. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase and plastid isoprenoid biosynthesis during tomato fruit ripening. *Plant J.* 27, 213-222.

Rodríguez-Concepción M. and Boronat A. 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids: a metabolic milestone achieved through genomics. *Plant Physiol.* 130, 1079-1089.

Rodríguez-Ezpeleta N., Brinkmann H., Burey S.C., Roure B., Burger G., Löffelhardt W., Bohnert H.J., Philippe H. and Lang B.F. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15, 1325-1330.

Rohdich F., Lauw S., Kaiser J., Feicht R., Köhler P., Bacher A. and Eisenreich W. 2006. Isoprenoid biosynthesis in plants-2C-methyl-D-erythritol-4-phosphate synthase (IspC protein) of *Arabidopsis thaliana*. *FEBS J.* 273, 4446-4458.

Rohmer M., Knani M., Simonin P., Sutter B. and Sahm H. 1993. Isoprenoid biosynthesis in bacteria: a novel pathway for early steps leading to isopentenyl diphosphate. *Biochem. J.* 295, 517-524.

Rohmer M. 1998. Isoprenoid biosynthesis via the mevalonate-independent route, a novel target for antibacterial drugs? *Prog. Drug. Res.* 50, 135-154.

Rosa Putra S., Disch A., Bravo J.M. and Rohmer M. 1998. Distribution of mevalonate and glyceraldehyde 3-phosphate/pyruvate routes for isoprenoid biosynthesis in some gram-negative bacteria and mycobacteria. *FEMS Microbiol. Lett.* 164, 169-175.

Rossmann, M.G., Moras, D. and Olsen, K.W. 1974. Chemical and biological evolution of a nucleotide-binding protein. *Nature* 250, 194-199.

Sacchettini J.C. and Poulter C.D. 1997. Creating isoprenoid diversity. *Science* 277, 1788-1789.

Sambrook J., Fritsch E.F. and Maniatis T. 1989. Molecular cloning. A laboratory manual, 2nd edn. Cold Spring Harbour Laboratory, Cold Spring Harbour.

Sanger F., Nicklen S. and Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.

Schenk G., Layfield R., Candy J.M., Duggleby R.G. and Nixon P.F. 1997. Molecular evolutionary analysis of the thiamine-diphosphate-dependent enzyme, transketolase. *J. Mol. Evol.* 44, 552-572.

Schneider G. and Linvsqvist Y. 1998. Crystallography and mutagenesis of transketolase: mechanistic implications for enzymatic thiamin catalysis. *Biochim. Biophys. Acta* 1385, 387-398.

Schuermann M., Schuermann M. and Sprenger G. A. 2002. Fructose 6-phosphate aldolase and 1-deoxy-D-xylulose 5-phosphate synthase from *Escherichia coli* as tools in enzymatic synthesis of 1-deoxysugars. *J. Mol. Catal.* B, 247-252.

Schwede T., Kopp J., Guex N. and Peitsch M.C. 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucl. Acids Res.* 31, 3381-3385.

Schwender J., Seemann M., Lichtenthaler H.K. and Rohmer, M. 1996. Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) *via* a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga *Scenedesmus obliquus*. *Biochem. J.* 316, 73-80.

Schwender J., Müller C., Zeidler J. and Lichtenthaler H.K. 1999. Cloning and heterologous expression of a cDNA encoding 1-deoxy-D-xylulose-5-phos-phate reductoisomerase of *Arabidopsis thaliana*. *FEBS Lett.* 455, 140-144.

Seto H., Watanabe H. and Furihata K. 1996. Simultaneous operation of the mevalonate and non-mevalonate pathways in the biosynthesis of isopentenyl diphosphate in *Streptomyces aeriouvifer*. *Tetrahedron Lett.* 37, 7979-7982.

Seto H., Orihara N. and Furihata K. 1998. Studies on the biosynthesis of terpenoids produced by *Actinomycetes*. Part 4. Formation of BE-40644 by the mevalonate and nonmevalonate pathways. *Tetrahedron Lett*. 39, 9497-9500.

Sharkey T.D., Yeh S., Wiberley A.E., Falbel T.G., Gong D. and Fernandez D.E. 2005. Evolution of the isoprene biosynthetic pathway in kudzu. *Plant Physiol.* 137, 700-712.

Sirinupong N., Suwanmanee P., Doolittle R.F. and Suvachitanont W. 2005. Molecular cloning of a new cDNA and expression of 3-hydroxy-3-methylglutaryl-CoA synthase gene from *Hevea brasiliensis*. *Planta* 221, 502-512.

Skilleter D.N. and Kekwick R.G.O. 1971. The enzymes forming isopentenyl pyrophosphate from 5-phosphomevalonate (mevalonate-5-phosphate) in the latex of *Hevea brasiliensis*. *Biochem. J.* 124, 407-415.

Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H., Dubois J., Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K., Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D., Spadafora R., Vicaire R., Wang Y., Wierzbowski J., Gibson R., Jiwani N., Caruso A., Bush D. and Reeve J.N. 1997. Complete genome sequence of *Methanobacterium thermo-autotrophicum* deltaH: functional analysis and comparative genomics. *J. Bacteriol.*179, 7135-7155.

Soler E., Clastre M., Bantignies B., Marigo G. and Ambid C. 1993. Uptake of isopentenyl diphosphate by plastids isolated from *Vitis vinifera* L. cell suspensions. *Planta* 191, 324-329.

Souret F.F., Weathers P.J. and Wobbe K.K. 2002. The mevalonate-independent pathway is expressed in transformed roots of *Artemisia annua* and regulated by light and culture age. *In Vitro Cell. Dev. Biol.-Plant.* 38, 581-588.

Sprenger G.A., Schörken U., Wiegert T., Grolle S., Graaf A.A., Taylar S.V., Begley T.P., Bringer-Meyer S. and Sahm H.. 1997. Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose

5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc. Natl. Acad. Sci. USA* 94, 12857-12862.

Stancu C. and Sima A. 2001. Statins: mechanism of action and effects. *J. Cell Mol. Med.* 5, 378-387.

Steinbacher S., Kaiser J., Eisenreich W., Huber R., Bacher A. and Rohdich F. 2003. Structural basis of fosmidomycin action revealed by the complex with 2-C-methyl-D-erythritol 4-phosphate synthase (IspC): implications for the catalytic mechanism and anti-malaria drug development. *J. Biol. Chem.* 278, 18401-18407.

Steinbüchel A. 2003. Production of rubber-like polymers by microorganisms. *Curr. Opin. Microbiol.* 6, 261-270.

Sterky F., Bhalerao R.R., Unneberg P., Segerman B., Nilsson P., Brunner A.M., Charbonnel-Campaa L., Lindvll J.J., Tandre K., Strauss S.H., Sundberg B., Gustafsson P., Uhlen M., Bhalerao R.P., Nilsson O., Sandberg G., Karlsson J., Lundeberg J. and Jansson S. 2004. A *Populus* EST resource for plant functional genomics. *Proc. Natl. Acad. Sci. USA* 101, 13951-13956.

Suvachittanont W. and Wititsuwannakul R. 1995. 3-Hydroxy-3-methylglutaryl coenzyme A synthase in *Hevea brasiliensis*. *Phytochemistry* 40, 757-761.

Suwanmanee P., Suvachittanont W. and Fincher G.B. 2002. Molecular cloning of 3-hydroxy-3-methylglutaryl coenzyme A synthase in *Hevea brasiliensis* Muell Arg. *Sci. Asia* 28, 29-36.

Suwanmanee P., Sirinupong N. and Suvachittanont W. 2004. Regulation of the expression of 3-hydroxy-3-methylglutaryl-CoA synthase gene in *Hevea brasiliensis* (B.H.K) Mull Arg. *Plant Sci.* 166, 531-537.

Tang C., Qi J., Li H., Zhang C. and Wang Y. 2007. A convenient and efficient protocol for isolating high-quality RNA from latex of *Hevea brasiliensis* (para rubber tree). *J. Biochem. Biophys. Methods* 70, 749-754.

Takahashi S., Kuzuyama T., Watanabe H. and Seto H. 1998. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalysing the formation of *2*-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc. Natl. Acad. Sci.USA* 95, 9879-9884.

The international rubber research and development board. 2005. URL: http:// www.-irrdb.com.

Thellina O., Zorzi W., Lakayec B., De Bormanc B., Coumansa B., Hennenb G., Grisarc T., Igoutb A. and Heinena E. 1999. Housekeeping genes as internal standards: use and limits. *J. Biotechnol.* 75, 291-295.

Thompson J.D., Higgins D.G. and Gibson T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22, 4673-4680.

Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F. and Higgins D.G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* 24, 4876-4882.

Totté N., Van den Ende W., Van Damme E.J.M, Compernolle F., Baboeuf I. and Geuns J.M.C. 2003. Cloning and heterologous expression of early genes in gibberellin and steviol biosynthesis *via* the methylerythritol phosphate pathway in *Stevia rebaudiana* Bertoni. *Can. J. Bot.* 81, 517-522.

Tupy J. 1988. Ribosomal and polyadenylated RNA content of rubber tree latex, association with sucrose level and latex pH. *Plant Sci.* 55, 137-144.

Veau B., Courtois M., Oudin A., Chenieux J.C., Rideau M. and Clastre M. 2000. Cloning and expression of cDNAs encoding two enzymes of the MEP pathway in *Catharanthus roseus*. *Biochim. Biophys. Acta* 1517, 159-163.

Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., et al. 2001. The sequence of the human genome. *Science* 291, 1304-1351.

von Heijne G. and Nishikawa K. 1991. Chloroplast transit peptide. *FEBS Lett.* 278, 1-3.

Walter M.H., Fester T. and Strack D. 2000. Arbuscular mycorrhizal fungi induce the non-mevalonate methylerythritol phosphate pathway of isoprenoid bio-synthesis correlated with accumulation of the 'yellow pigment' and other apocarotenoids. *Plant J.* 21, 571-578.

Walter M.H., Hans J. and Strack D. 2002. Two distantly related genes encoding 1-deoxy-D-xylulose 5-phosphate synthases: differential regulation in shoots and apocarotenoid-accumulating mycorrhizal roots. *Plant J.* 31, 243-254.

Wetzel R. 1996. For protein misassembly, it's the ''I'' decade. *Cell.* 86, 699-702.

Wikner C, Meshalkina L, Nilsson U, Nikkola M, Lindqvist Y, Sundström M, Schneider G. 1994. Analysis of an invariant cofactor-protein interaction in thiamin diphosphate-dependent enzymes by site-directed mutagenesis-glutamic acid 418 in transketolase is essential for catalysis. *J. Biol. Chem.* 269, 32144-32150.

Wikner C., Nilsson U., Meshalkina L., Udekwu C., Lindqvist Y. and Schneider G. 1997. Identification of catalytically important residues in yeast transketolase. *Biochemistry* 36, 15643-15649.

Wititsuwannakul R., Wititsuwannakul D., Sothibanhhu R., Suvachittanont W. and Sukonrat W. 1988. Correlation studies on 3-hydroxy-3-methylglutaryl coenzyme A reductase activity and dry rubber yield in *Hevea brasiliensis*. In: C.R. Coll. Expl. Physiol. Amel. Hevea, IRCA-CIRAD, France, 2-7 Nov., pp. 161-172.

Wititsuwannakul D. and Wititsuwannakul R. 2001. Biochemistry of natural rubber and structure of latex. In: Koyama T. and Steinbuechel A., Editors, *Biopolymers*, Weinheim, Wiley-VCH, pp. 151-201.

Wititsuwannakul D., Rattanapittayaporn A. and Wititsuwannakul R. 2003. Rubber biosynthesis by a *Hevea* latex bottom-fraction membrane. *J. Appl. Polym. Sci.* 87, 90-96.

Wititsuwannakul D., Rattanapittayaporn A. Koyama T. and Wititsuwannakul R. 2004. Involvement of *Hevea* latex organelle membrane proteins in the rubber bio-synthesis activity and regulatory function. *Macromol. Biosci.* 4, 314-323.

Woo Y.H., Fernandes R.P.M. and Proteau P.J. 2006. Evaluation of fosmidomycin analogs as inhibitors of the *Synechocystis* sp. PCC6803 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *Bioorg. Med. Chem.*14, 2375-2385.

Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., et al. 2002. The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415, 871-880.

Xiang S., Usunow G., Lange G., Busch M. and Tong L. 2007. Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J. Biol. Chem.* 282, 2676-2682.

Yajima S., Nonaka T., Kuzuyama T., Seto H. and Ohsawa K. 2002. Crystal structure of 1-deoxy-D-xylulose 5-phosphate reductoisomerase complexed with cofactors: implications of a flexible loop movement upon substrate binding. *J. Biochem.* 131, 313-317.

Yajima S., Hara K., Sanders J.M., Yin F., Ohsawa K. and Wiesner J. 2004. Crystallographic structures of two bisphosphonate:1-deoxyxylulose-5-phos-phate reductoisomerase complexes. *J. Am. Chem. Soc.* 126, 10824-10825.

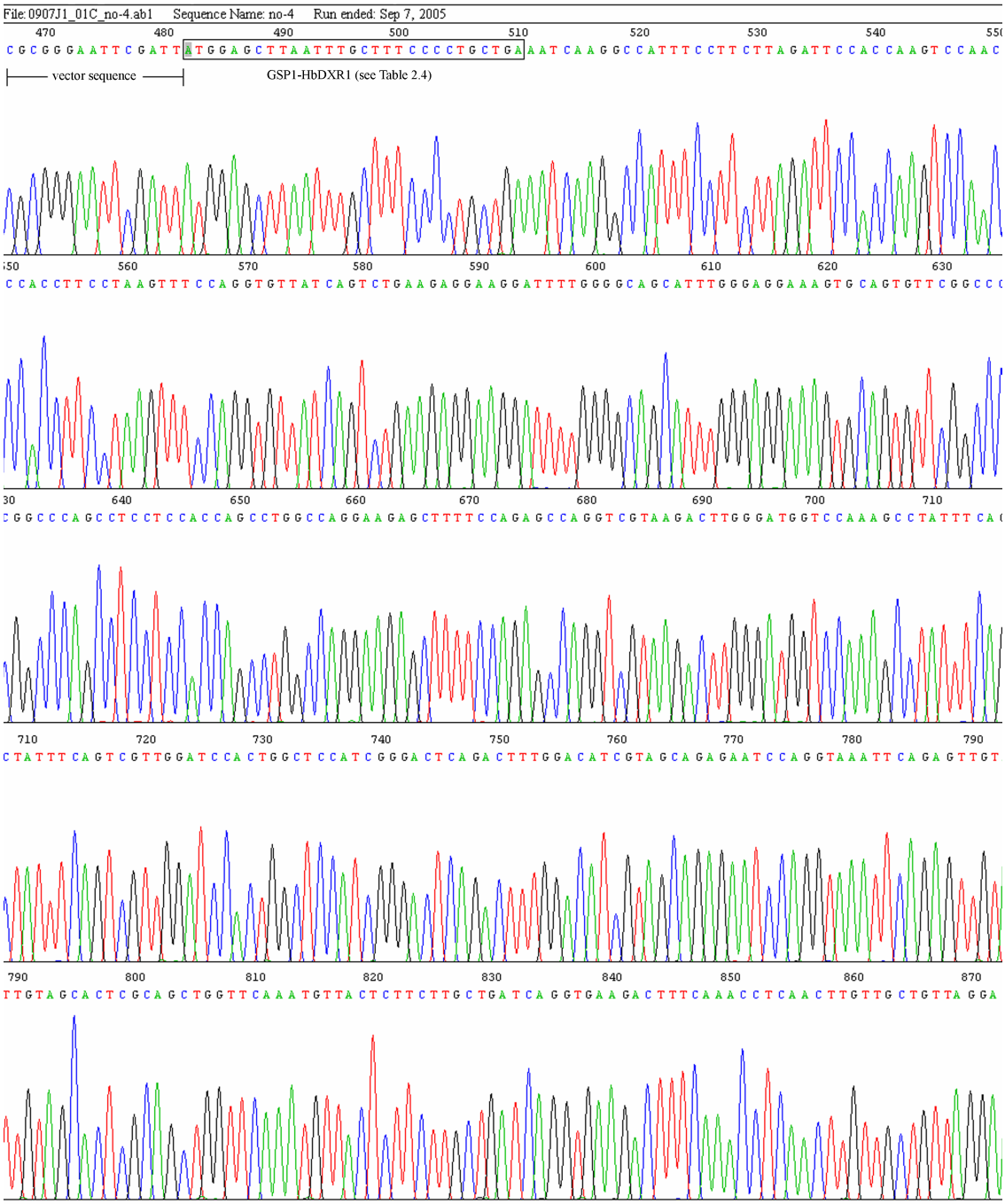
Yamasakia Y. and Akimitsu K. 2007. *In situ* localization of gene transcriptions for monoterpene synthesis in irregular parenchymic cells surrounding the secretory cavities in rough lemon (*Citrus jambhiri*). *J. Plant Physiol.* 164, 1436-1448.

Yao H., Gong Y., Zuo K., Ling H., Qiu C., Zhang F., Wang Y., Pi Y., Liu X., Sun X. and Tang K. 2007. Molecular cloning, expression profiling and functional analysis of a *DXR* gene encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Camptotheca acuminata*. *J. Plant Physiol.* doi:10.1016/-j.jplph.2006.12.001.

Yin X. and Proteau P.J. 2003. Characterization of native and histidine-tagged deoxyxylulose 5-phosphate reductoisomerase from the cyanobacterium *Synechocystis* sp. PCC6803. *Biochim. Biophys. Acta* 1652, 75-81.

1. The chromatogram profile of nucleotide sequencing

The nucleotide sequence was analyzed by automate DNA sequencing using the ABI PRISM System (model 377, PE Applied Biosystems).



2. Amino acids and their abbreviations

Amino acid	Three letter code	Single letter code
Glycine	Gly	G
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tryptophan	Trp	W
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Cysteine	Cys	C
Tyrosine	Tyr	Y
Asparagine	Asn	N
Glutamine	Gln	Q
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Lysine	Lys	K
Arginine	Arg	R
Histidine	His	H

3. DNA codons and their corresponding amino acids

<table>
<tr><th></th><th>T</th><th>C</th><th>A</th><th>G</th><th></th></tr>
<tr><td rowspan="4">T</td><td rowspan="2">Phe [F]</td><td rowspan="4">Ser [S]</td><td>Tyr [Y]</td><td>Cys [C]</td><td>T</td></tr>
<tr><td rowspan="3">Ter [end]</td><td>C</td></tr>
<tr><td rowspan="2">Leu [L]</td><td>Ter [end]</td><td>A</td></tr>
<tr><td>Trp [W]</td><td>G</td></tr>
<tr><td rowspan="4">C</td><td rowspan="4">Leu [L]</td><td rowspan="4">Pro [P]</td><td rowspan="2">His [H]</td><td rowspan="4">Arg [R]</td><td>T</td></tr>
<tr><td>C</td></tr>
<tr><td rowspan="2">Gln [Q]</td><td>A</td></tr>
<tr><td>G</td></tr>
<tr><td rowspan="4">A</td><td rowspan="3">Ile [I]</td><td rowspan="4">Thr [T]</td><td rowspan="2">Asn [N]</td><td rowspan="2">Ser [S]</td><td>T</td></tr>
<tr><td>C</td></tr>
<tr><td rowspan="2">Lys [K]</td><td rowspan="2">Arg [R]</td><td>A</td></tr>
<tr><td>Met [M]</td><td>G</td></tr>
<tr><td rowspan="4">G</td><td rowspan="4">Val [V]</td><td rowspan="4">Ala [A]</td><td rowspan="2">Asp [D]</td><td rowspan="4">Gly [G]</td><td>T</td></tr>
<tr><td>C</td></tr>
<tr><td rowspan="2">Glu [E]</td><td>A</td></tr>
<tr><td>G</td></tr>
</table>

VITAE

Name Mr. Yortyot Seetang-Nun

Student ID 4323001

Educational Attainment

Degree	Name of Institution	Year of Graduation
Bachelor of Science (Biology)	Prince of Songkla University	2000

Scholarship Awards during Enrolment

The Royal Golden Jubilee Graduate Program (PHD/0115/2543) from Thailand Research Fund (TRF), 2000-2005.

List of Publication and Proceeding

1. **Seetang-Nun, Y.**, Sharkey T.D. and Suvachittanont W. 2008. Isolation and characterization of two distinct classes of *DXS* genes in *Hevea brasiliensis*. *DNA Seq.* 19, 291-300.
2. **Seetang-Nun, Y.**, Sharkey T.D. and Suvachittanont W. 2008. Molecular cloning and characterization of two cDNAs encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Hevea brasiliensis*. *J. Plant Physiol.* 165, 991-1002.
3. **Seetang-nun, Y.** and Suvachittanont W. 2004. Cloning and sequencing of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Hevea brasiliensis* (Willd. ex A. Jusss.) Muell. Arg. (Poster presentation), RGJ-Ph.D. Congress V, Apr 23-25, Pattaya, Thailand.
4. **Seetang-nun, Y.** and Suvachittanont W. 2004. A *DXR* gene and its expression in rubber tree (Oral presentation), RGJ Seminar Series XXX: Biosciences and Biotechnology for the Development of Southern Thailand, Aug 13, Faculty of Agro-Industry, Prince of Songkla University, Songkhla, Thailand.
5. **Seetang-nun, Y.** and Suvachittanont W. 2004. Nucleotide sequence of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate (DXP) synthase from *Hevea brasiliensis* (Poster presentation), The 17th FAOBMB Symposium/2nd IUBMB special Meeting/7th A-IMBN Conference on “Genomics and Health in the 21st century”, Nov 22-26, Bangkok, Thailand.