Lehrstuhl für Biochemie der Technischen Universität München

Characterization and inhibition of the IspC, IspE and IspF proteins involved in the deoxyxylulose phosphate pathway of isoprenoids biosynthesis

Susan Lauw

Vollständiger Abdruck der von der Fakultät für Chemie der Technische Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzende:	Univ.-Prof. Dr. S. Weinkauf
Prüfer der Dissertation:	1. Univ.-Prof. Dr. Dr. A. Bacher, i.R.
	2. Univ.-Prof. Dr. P. Schieberle

Die Dissertation wurde am 14.02.2008 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 11.06.2008 angenommen.

Die experimentellen Arbeiten zur vorliegenden Dissertation wurden von März 2004 bis Februar 2008 am Lehrstuhl für Biochemie der Technischen Universität München durchgeführt.

Acknowledgement

The PhD study in department of Biochemisty and Organic Chemistry, Technische Universität München has been a significant academic challenge for me. This dissertation would not have been written without help and support from the following people. To them, I would like to express my greatest gratitude:

- Prof. Dr. Dr. Adelbert Bacher for the biggest chance in my life to further my study. His great knowledge, wisdom and patience have been always inspiring and motivating me.
- Dr. Felix Rohdich for his knowledge, fruitful discussion and endless support. It has been a pleasure to work with him.
- Dr. Wolfgang Eisenreich for his knowledge and all discussion especially about NMR and chemistry.
- Prof. Dr. Michael Groll for his support and knowledge in crystallography.
- Dr. Johannes Kaiser for his patience in teaching all the knowledge he has during my first work in this department.
- Prof. Dr. Francois Diedrich, Anna Hirsch, Corinne Baumgartner, Christine Crane for the cooperation on publication work.
- Prof. Dr. Avi Golan (Ben Gurion University of the Negev, Israel) and Prof. Dr. Jamal Safi (Environment Protection and Research Institute, Gaza) for the plant materials and the cooperation in working with plant extracts.
- Dr. Boris Illarionov, Dr. Victoria Illarionova, Dr. Ralf Laupitz, Dr. Ferdinand Zepek, Dr. Tobias Gräwert, Dr. Werner Römisch, Katrin Gärtner, Richard Feicht, Christoph Graßberger, Astrid König, Christine Schwarz, Matthias Lee, Dr. Monika Joshi, Elena Ostrozhenkova, Fritz Wendling, Silke Marsch, Eva Eylert and Birgit Keil for all the help and friendship during my study.
- My family, especially my mother for her eternal support.

List of Publication

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

Crane CM, Kaiser J, Ramsden NL, Lauw S, Rohdich F, Eisenreich W, Hunter WN, Bacher A and Diederich F. Fluorescent Inhibitors for IspF, an Enzyme in the Non-Mevalonate Pathway for Isoprenoid Biosynthesis and a Potential Target for Antimalarial Therapy. *Angew. Chem*. Int. Ed. 2006, 45, 1069 –1074.

Kaiser J, Yassinb M, Prakash S, Safib N, Agamid M, Lauw S, Ostrozhenkova E, Bacher E, Rohdich F, Eisenreich W, Safi J and Golan A. Goldhirshc. Anti-malarial drug targets: Screening for inhibitors of 2C-methyl-D-erythritol 4-phosphate synthase (IspC protein) in Mediterranean plants. *Phytomedicine* April 2007, Vol 14, Issue 4, 242-249.

Hirsch AKH, Lauw S,Gersbach P, Schweizer WB, Rohdich F, Eisenreich W, Bacher A, and Diederich F. Non-phosphate inhibitors of IspE protein, a kinase in the non-mevalonate pathway for isoprenoid biosynthesis and a potential target for antimalarial therapy. *ChemMedChem* March 2007, Vol.2, Issue 6, 806-810.

Baumgartner C, Eberle C, DiederichF, Lauw S, Rohdich F, Eisenreich W, Bacher A. Structure-Based Design and Synthesis of the First Weak Non-Phosphate Inhibitors for IspF, an Enzyme in the Non-Mevalonate Pathway of Isoprenoid Biosynthesis. *Helvetica Chimica Acta*, June 2007, Volume 90, Issue 6, Pages 1043-1068.

Crane CM, Hirsch AKH, Alphey MS, Sgraja T, Lauw S, Illarionova V, Rohdich F, Eisenreich W, Hunter WN, Bacher A, and Diederich F. Synthesis and Characterization of Cytidine Derivatives that Inhibit the Kinase IspE of the Non-Mevalonate Pathway for Isoprenoid Biosynthesis. *ChemMedChem* 2008, 3, 91 – 10.

Index of Contents

List of Figures

List of Tabulations

List of Abbreviations

ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CDP	Cytidine 5‘-diphosphate
CDP-ME	4-diphosphocytidyl-2*C*-methyl-D-erythritol
CDP-MEP	4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate
cMEPP	2*C*-metyl-D-erythritol 2,4-cyclodiphosphate
CMP	Cytidine 5‘-monophosphate
CTP	Cytidine 5‘-triphosphate
DMAPP	Dimethylallyl diphosphate
dpmd	Diphosphomevalonate decarboxylase
DTT	Dithiothreitol
DXP	1-Deoxy-D-xylulose 5-phosphate
Dxs	1-Deoxy-D-xylulose 5-phosphate synthase
h	hour
Hmgr	(*S*)-3-Hydroxy-3-methylglutaryl-CoA reductase
Hmg*s*	(*S*)-3-Hydroxy-3-methylglutaryl-CoA synthase
Idi	Isopentenyl diphosphate isomerase
IdiI	Isopentenyldiphosphate isomerase type I
IdiII	Isopentenyldiphosphate isomerase type II
IPP	Isopentenyl pyrophosphate
IPTG	Isopropyl-β-D-thiogalactopyranoside
IspC/ dxr	2*C*-methyl-D-erythritol 4-phosphate synthase
IspD	4-Diphosphocytidyl-2*C*-methyl-D-erythritol synthase

IspE	4-Diphosphocytidyl-2*C*-methyl-D-erythritol kinase
IspF	2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase
IspG	2C-Methyl-D-erythritol 2,4-cyclodiphosphate reductase
IspH	1-Hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate reductase
K_i	Dissociation constant of an inhibitor
M	Molar
MEP	2*C*-Methyl-D-erythritol 4-phosphate
Min	Minute
Mk	Mevalonate kinase
mM	milimolar
NAD^+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
nm	Nanometer
NMR	Nuclear Magnetic Resonance
PEP	Phosphoenol pyruvic acid monopotassium salt
Pmk	Phosphomevalonate kinase
Py-dehydrogenase	Pyruvate dehydrogenase
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TEMED	N,N,N',N'-tetramethylethylenediamine
TLC	Thin Layer Chromatography
UV	ultraviolet

1 Introduction

Terpenoids are large and varied classes of natural products produced by nearly all organisms serving as visual pigments, reproductive hormones, defensive agents, constituents of membranes, components of signal transduction networks, mating pheromones and photoprotective agents. Monoterpenes (C_{10}) and sesquiterpenes (C_{15}) are the elements of the essential oils that are used as flavoring and coloring agents in food, cosmetics, and perfumes as well as therapeutic agents such as antibacterials, fungicides, antiinflammatory compounds, etc. Other terpenoids like taxol (diterpenoid) serve as cytostatic agents while lutein and lycopene (tetraterpenoids) have been registered as oncopreventive agents (Khachik *et al.,* 1995; Demming-Adams and Adams, 2002).

Isoprene (**1**) with the molecular formula C_5H_8 can be considered, in a formal sense, as a basic building block of terpenoids, and certain simple terpenes have the molecular formula $(C_5H_8)_n$. However, only the activated forms of isoprene, i.e. dimethylallyl pyrophosphate (DMAPP)(**3**) and isopentenyl diphosphate (IPP)(**2**) can serve as actual precursors of terpenes (Figure 1). Various assembly and modification of IPP and DMAPP yield more than 30,000 different terpenoids that we know today.

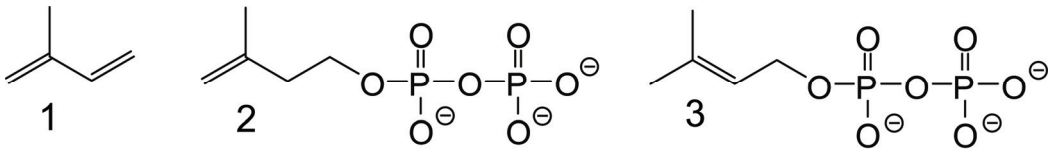


Figure 1: Terpene precursors
Isoprene(**1**) and its activated forms IPP (**2**) and DMAPP (**3**)

Two independent pathways are known for terpenoids biosynthesis, i.e., the mevalonate pathway and the deoxyxylulose phosphate pathway. The occurrence of these pathways among the three kingdoms of life is shown in Table 1. Many pathogenic bacteria have the deoxyxylulose phosphate pathway while mammals use the mevalonate pathway. Plants use the mevalonate pathway in the cytosol and the deoxyxylulose phosphate pathway in their plastids (Eisenreich *et al*., 2004). For this reason, the deoxyxylulose phosphate pathway enzymes are attractive targets for the

design of novel antibiotics and herbicides where toxicity related to selectivity in drug design is resolved.

Table 1: IPP and DMAPP biosynthesis genes in representative examples of completely sequenced organisms

Organisms	Pathway		Isomerase	
	DXP	MEV	IdiI	IdiII
Bacteria				
Aquifales (*Aquifex aeolicus*)	+	—	—	—
Chlamydia group (*Chlamydophila pneumoniae*)	+	—	—	—
Cyanobacteriab (*Synechocystis* sp.)	+	—	—	+
Deinococcus group (*Deinococcus radiodurans*)	+	—	—	+
Firmicutes (*Bacillus subtilis*)	+	—	—	+
(*Mycoplasma genitalium*)	—	—	—	—
(*Staphylococcus aureus*)	—	+	—	+
(*Streptomyces coelicolor*)	+	—	+	—
(*Listeria monozytogenes*)	+	+	—	+
Proteobacteria (*Escherichia coli*)	+	—	+	—
(*Rickettsia prowazeckii*)	—	—	—	+
Spirochaetales (*Treponema pallidum*)	+	—	—	—
(*Borrelia burgdorferi*)	—	+	—	+
Thermotogales (*Thermotoga maritima*)	+	—	—	—
Archaea				
Crenarchaeota (*Aeropyrum pernix*)	—	+	—	+
Euryarchaeota (*Archaeoglobus fulgidus*)	—	+	—	+
Eukaryotes				
Animals (*Homo sapiens*)	—	+	+	—
Plants (*Arabidopsis thaliana*)	+	+	+	—
Plasmodium falciparum	+	—	—	—
Yeasts (*Saccharomyces cerevisiae*)	—	+	+	—

Abbreviations are explained in the List of Abbreviations. This table was published in “Biosynthesis of isoprenoids via the non-mevalonate pathway” by Eisenreich *et al*., 2004

1.1 Mevalonate pathway

The mevalonate pathway (Figure 2) of isoprenoid biosynthesis has been well investigated through the work of Bloch, Lynen and Cornforth (Qureshi and Porter, 1981; Bloch, 1992; Bach, 1995; Bochar *et al*., 1999). Mevalonic acid (**7**) is generated by the condensation of three units of acetyl-CoA (**4**) *via* acetoacetyl-CoA (**2**) and

3-hydroxy-3-methyl-glutaryl-CoA (**5**) followed by a reduction step catalyzed by (*S*)-3-hydroxy-3-methylglutaryl-CoA reductase (*hmgr*). Two consecutive phosphorylation steps catalyzed by mevalonate kinase (*mk*) and phosphomevalonate kinase (*pmk*) lead to the formation of mevalonate phosphate (**8**) and mevalonate diphosphate (**9**), respectively. The final step in this pathway involves the ATP-dependent decarboxylation of mevalonate diphosphate to IPP (**2**) by diphospomevalonate decarboxylase. Its allylic isomer DMAPP (**3**) is then formed by the catalytic action of IPP isomerase.

1.2 Deoxyxylulose phosphate pathway

The discovery of deoxyxylulose phosphate pathway was initiated by the observation of some abnormalities (Rohmer 1999; Eisenreich *et al*, 1998) that could not be explained by means of the mevalonate pathway. In these cases, labeled acetate and mevalonate that were feeded to bacterial as well plant cells were only poorly incorporated into the analyzed terpenes synthesized by those organisms. Additionally, the incorporation of ^{13}C-labeled glucose into ginkgolides of *Ginkgo biloba* showed a labeling pattern which was inconsistent with the mevalonate pathway (Schwarz, 1994). Other experiments (White, 1978; David *et al*., 1981; David *et al*., 1982; Hill *et al*., 1989; Broers, 1994; Rohmer *et al*., 1996; Arigoni and Schwarz, 1999; Eisenreich *et al*., 2001) confirmed these observations leading finally to the elucidation of a new pathway, the non-mevalonate pathway also called deoxyxylulose phosphate pathway (Figure 2).

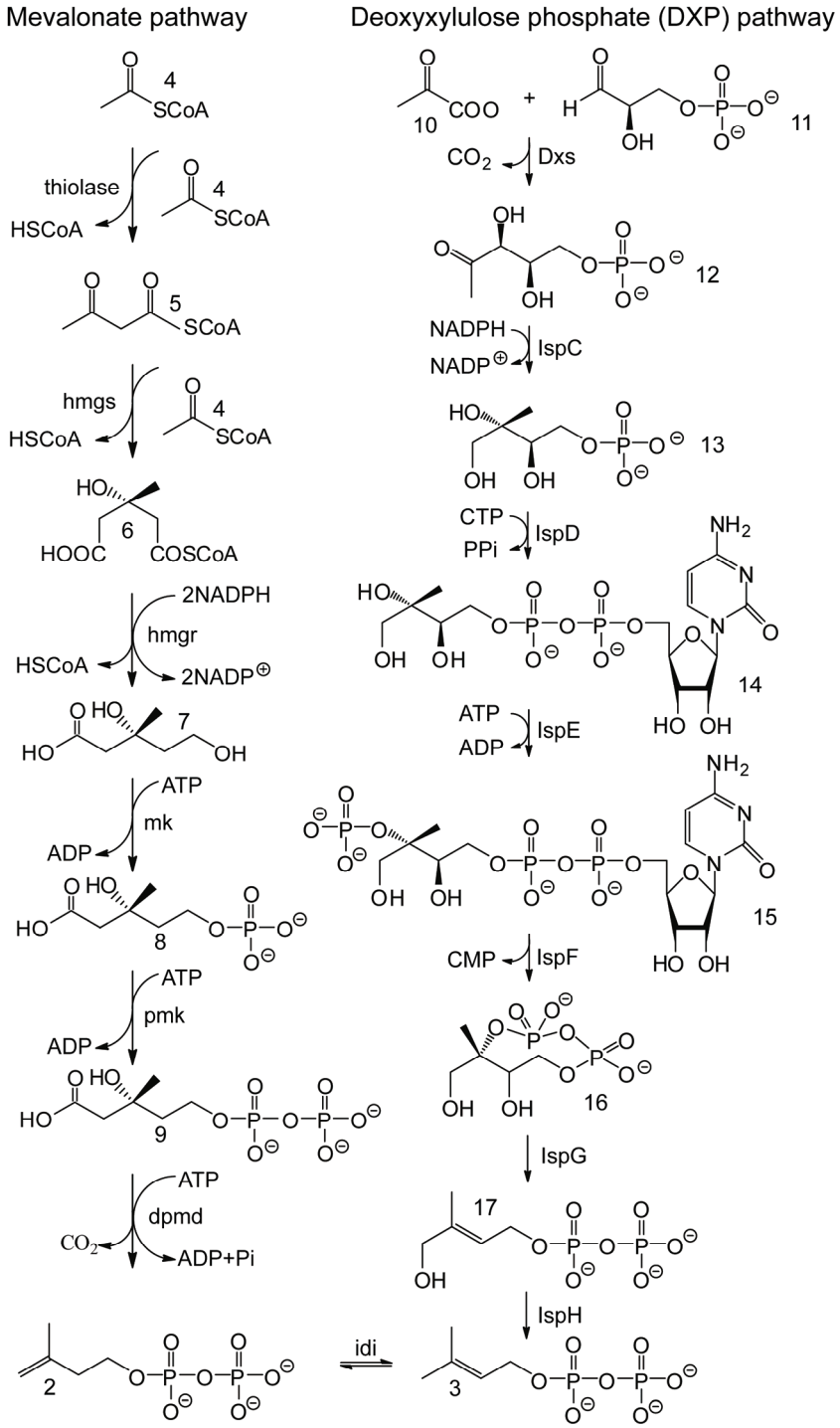


Figure 2: Mevalonate and deoxyxylulose phosphate pathway
Mevalonate pathway : Acetyl-CoA (**4**); acetoacetyl CoA (**5**); 3-hydroxy-3-methyl-glutaryl-CoA (**6**); mevalonic acid (**7**); mevalonate phosphate (**8**); mevalonate diphosphate (**9**).
Deoxyxylulose phosphate pathway: Pyruvate (**10**); D-glyceraldehyde-3-phosphate (**11**); 1-deoxy-D-xylulose-5-phosphate (**12**); 2*C*-methyl-D-erythritol 4-phosphate (**13**); 4-diphosphocytidyl-2*C*-methyl-D-erythritol (**14**); 4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate (**15**); 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate (**16**); 1-hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate (**17**).

On the basis of glycolytic pathways in plants the transfer of ^{13}C glucose to IPP and DMAPP *via* mevalonate or deoxyxylulose pathway can be predicted (Eisenreich *et al*., 2004). The labeling patterns of the ^{13}C enrichment of terpenoid precursors contributed by both terpenoid pathways can be analyzed by ^{13}C nuclear magnetic resonance spectroscopy (see Figure 3).

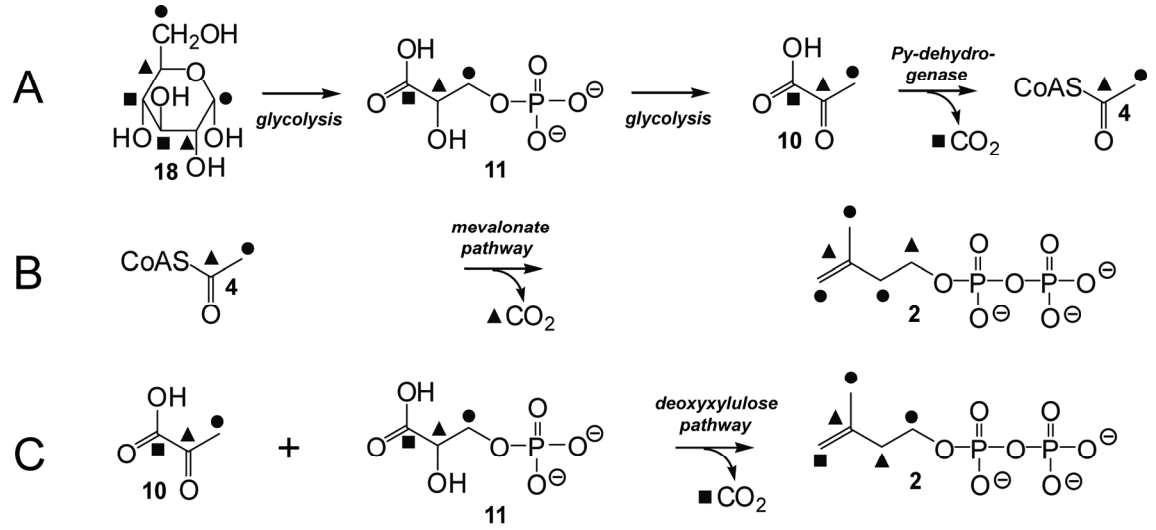


Figure 3: The predicted labeling pattern of IPP synthesized *via* the mevalonate pathway and the deoxyxylulose phosphate pathway.
(**A**), Predicted labeling patterns of glucose (**18**), glyceraldehyde 3-phosphate (**19**), pyruvate (**10**) and acetyl-CoA (**4**) produced from isotope-labeled glucose *via* glycolysis followed by the pyruvate dehydrogenase reaction. (**B**), predicted labeling pattern of IPP (**2**) biosynthesized from acetyl-CoA (**4**) *via* mevalonate. (**C**), predicted labeling pattern of IPP (**2**) biosynthesized from pyruvate (**10**) and D-glyceraldehyde 3-phosphate (**11**) *via* 1-deoxy-D-xylulose 5-phosphate. The symbols indicate biosynthetically equivalent positions. This picture was published by Eisenreich *et al*., 2004.

1.2.1 2*C*-Methyl-D-erythritol 4-phosphate synthase (IspC protein)

2*C*-methyl-D-erythritol 4-phosphate synthase (IspC) catalyzes the first committed step of the non-mevalonate pathway, i.e. the conversion of 1-deoxy-D-xylulose 5-phosphate (**12**) into 2*C*-methyl-D-erythritol 4-phosphate (**13**).

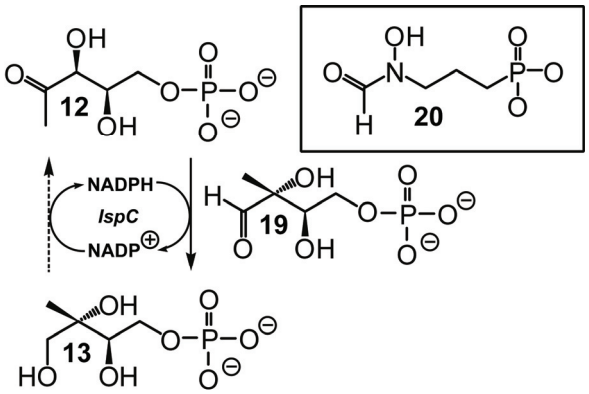


Figure 4: The reaction catalyzed by 2*C*-methyl-D-erythritol 4-phosphate synthase/IspC protein.
1-Deoxy-D-xylulose 5-phosphate (**12**/DXP), 2*C*-methyl-D-erythritol 4-phosphate (**13**/MEP), 2*C*-methyl-D-erythrose 4-phosphate (**19**), inhibitor fosmidomycin (**20**).

The enzyme can also use2*C*-methyl-D-erythritol 4-phosphate (**13**) as substrate and catalyzes the reverse reaction to form 1-deoxy-D-xylulose 5-phosphate (**12**) but with lower catalytic activity as compared to the forward reaction. The enzyme uses NADPH or $NADP^+$ as coenzyme and Mg^{2+} or Mn^{2+} as cofactor.

Fosmidomycin (**20**) has been reported as a strong inhibitor for 2*C*-methyl-D-erythritol 4-phosphate synthases in vitro. Additionally, it inhibits different strains of *Plasmodium falciparum* with IC_{50} in the 300 nM range (Jomaa *et al*., 1999). In the crystal structure of *E. coli* IspC complexed with Mn^{2+} and **20**, **12** can be nicely superimposed onto **20**, indicating that **20** binds in a substrate-like mode (Steinbacher *et al*., 2003).

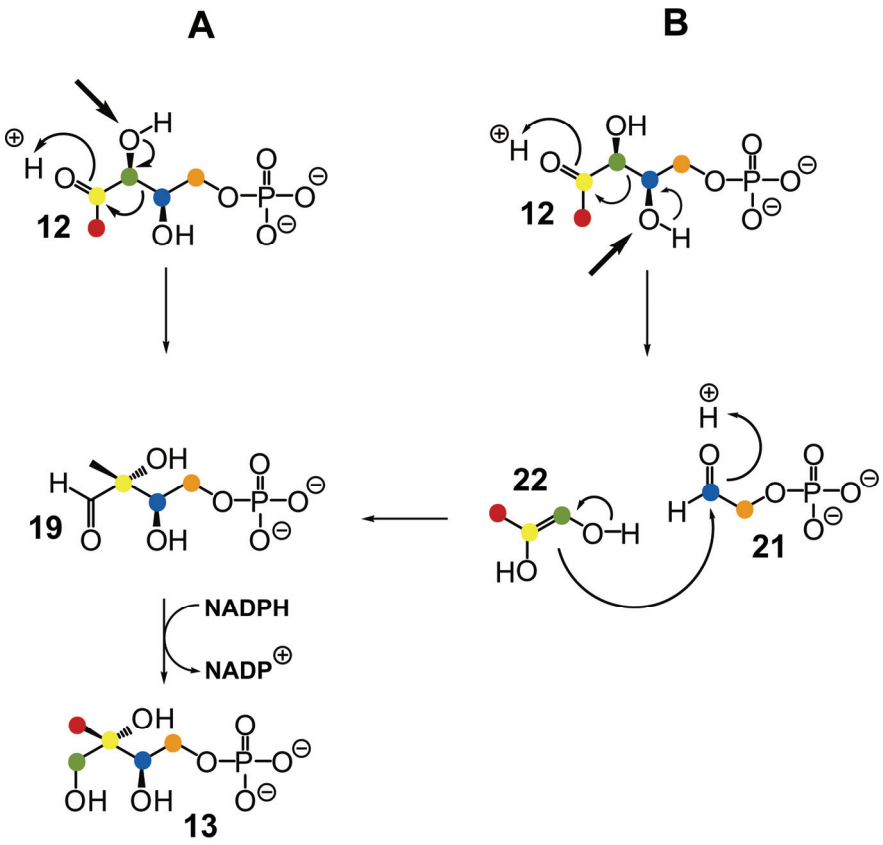


Figure 5: Hypothetical mechanisms of the 2C-methyl-D-erythritol 4-phosphate synthase/IspC reaction.
(**A**) α-ketol/sigmatropic rearrangement mechanism; (**B**) Retroaldol mechanism; 1-deoxy-D-xylulose 5-phosphate/DXP) (**12**); 2*C*-methyl-D-erythritol 4-phosphate/MEP) (**13**); 2*C*-methyl-D-erythrose 4-phosphate (**19**); glycoaldehyde phosphate (**21**); enolate of hydroxyacetone (**22**).The colouring shows the position of the respective carbon labels.

The course of the 2*C*-methyl-D-erythritol 4-phosphate synthase reaction (Figure 5) has been proposed to proceed by a skeletal rearrangement followed by the reduction of the resulting branched-chain aldose (Kuzuyama *et al.*, 1998). An α-ketol/sigmatropic mechanism (see Figure 5A) was suggested based on the analogy with the reaction catalyzed by ketol acid reductoisomerase (Kuzuyama *et al.*, 1998; Reuter *et al.*, 2002; Steinbacher *et al.*, 2003). Both enzymes use NADPH as cofactor and undergo isomerization reactions. However there is no sequence similarity between both enzymes, except the conserved motif LGXTGSIG which was deduced as their NADPH binding site (Takahashi *et al.*, 1998). Since the hypothetical intermediate of the IspC reaction, 2*C*-methyl-D-erythrose 4-phosphate (**19**), could not be trapped, the mechanism was proposed to involve a single intramolecular

rearrangement step whereby the process might be carried out in an enzyme bound stage (Takahashi *et al*., 1998).

As an alternative to a sigmatropic rearrangement, a retroaldol/aldol reaction sequence was introduced (Hoeffler *et al*., 2002). Retroaldol mechanism is well documented for ribulose-5-phosphate 4-epimerase (Johnson and Tanner, 1998; Lee *et al*., 2000). In this mechanism, deprotonation occurs at the C-4 hydroxy group of 1-deoxy-D-xylulose 5-phosphate, followed by cleavage of the C-3—C-4 bond yielding glycoaldehyde phosphate (**21**) and the enolate of hydroxyacetone (**22**).

2*C*-methyl-D-erythrose 4-phosphate (**19**) was shown to be converted into the enzyme product **13** in a reaction containing NADPH as cofactor (Hoffler *et al*., 2002). In a further experiment, this group could also show, that **13** can be converted into the substrate **12** using NADP+ as cofactor, indicating that **19** might be a real intermediate of the IspC reaction. Despite of extensive investigations, it has not been proven until now if sigmatropic or retroaldol is involved in the the mechanism of IspC reaction.

On the other hand, a hydride/methyl shift could be ruled out as on basis of isotope labeling studies, whereby a stereospecific transfer of H_{si} from C-4 of NADPH identifies 2*C*-methyl-D-erythritol 4-phosphate synthase as class B dehydrogenase (Proteau *et al*., 1999; Radykewicz *et al*., 1999; Argyrou and Blanchard, 2004).

The crystal structures of IspC proteins from some organisms have been determined by X-ray crystallography, notably *Escherichia coli* (Reuter *et al*., 2001; Yajima *et al*., 2002, Steinbacher *et al*., 2003; Yajima *et al*., 2004; Mac Sweeney *et al*., 2005; Yajima *et al*., 2007), *Zymomonas mobilis* (Ricagno *et al*., 2003) and *Mycobacterium tuberculosis* (Henriksson *et al*., 2006; Henriksson *et al*., 2007), indicating a similarity of the overall structures. IspC protein forms a homodimer, whereby each monomer can be subdivided into three domains, i.e., the N-terminal domain as an anchor for NADPH, the catalytic residues for the binding site of divalent metal ion and the C-terminal domain that supports the catalytic domain. A flexible loop (206 - 216, *Escherichia coli* IspC) is believed to cover the substrate binding site during the enzymatic reaction (Yajima *et al*., 2002; Mac Sweeney *et al*., 2005). Mn^{2+} was found to bind Asp150, Glu152 and Glu231 of *Escherichia coli* IspC (Steinbacher *et al*., 2003).

1.2.2 4-Diphosphocytidyl-2*C*-methyl-D-erythritol kinase (IspE protein)

IspE protein catalyzes the fourth step in the deoxyxylulose phosphate pathway, converting 4-diphosphocytidyl-2*C*-methyl-D-erythritol (CDP-ME/**14**) in the presence of ATP and a divalent metal ion into 4-diphosphocytidyl-2-*C*-methyl-D-erythritol 2-phosphate (CDP-MEP/**15**) with the release of ADP as shown in Figure 6. Orthologues genes from *Escherichia coli* and *Lycopersicon esculentum* have been cloned and expressed (Rohdich *et al*., 1999; Rohdich *et al*., 2000). The reaction rates for *Escherichia coli* and *Lycopersicon* esculentum were 34 µmol min^{-1} mg^{-1} and 33 µmol min^{-1} mg^{-1}, respectively.

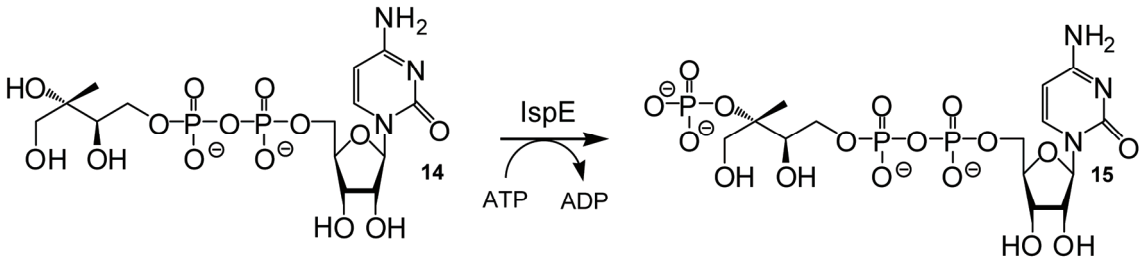


Figure 6: The reaction catalyzed by 4-Diphosphocytidyl-2*C*-methyl-D-erythritol kinase/IspE
14, 4-Diphosphocytidyl 2*C*-methyl-D-erythritol; **15**, 4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate.

The structure of a ternary complex of the homodimeric IspE protein from *Escherichia coli* with the substrate and an ATP analogue has been solved by X-ray crystallography (Miallau *et al*., 2003), affording approximate total mass of 62 kDa. The IspE protein subunits display a close structural relationship to the GHMP kinase superfamily, notably a well conservation of the structure for ATP binding domains and the core of substrate binding domains. The sequence analysis of IspE proteins from some species (Miallau *et al*, 2003), including pathogenic bacteria indicated a high degree of amino acid conservation, about 45 – 60 % over the species tested.

1.2.3 2*C*-Methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF protein)

The enzyme 2*C*-Methyl-D-erythritol 2,4-cyclodiphosphate synthase catalyzes the conversion of 4-diphosphocytidyl-2-*C*-methyl-D-erythritol 2-phosphate (CDP-MEP/**15**) into 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate (**16**) as depicted in Figure 7.

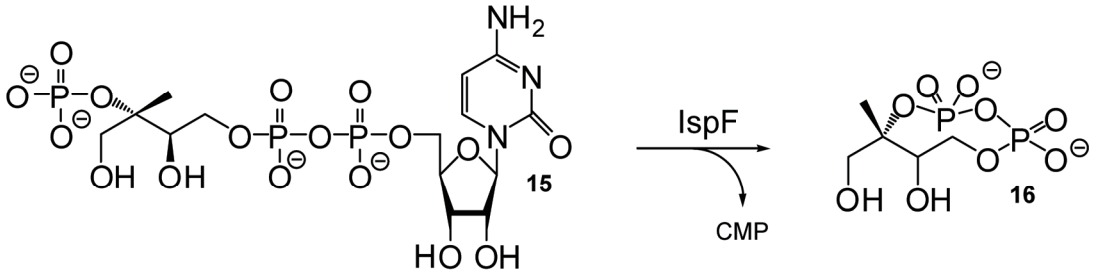


Figure 7: The reaction catalyzed by 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase/IspF
15, 4-Diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate; **16**, 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate.

The IspF protein utilizes divalent metals ions such as Mn^{2+}, Mg^{2+}, Co^{2+} or Zn^{2+} for its catalytic function, and the reaction proceeds only in one direction to the conversion of 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate.

Three dimensional structures for IspF from *Eschericia coli* (Steinbacher *et al*., 2001; Richard *et al*., 2001), *Arabidopsis thaliana* (calisto *et al*., 2007), *Haemophilus influenzae* (Lehmann *et al*., 2002) and *Mycobacterium smegmatis* (Buetow *et al*., 2007) have been determined by X-ray crystallography, showing that the proteins adopt a homotrimeric quaternary structure. The substrate binding site can be subdivided into three pockets, i.e., a central pocket surrounds the ribosyl-5'-diphosphate group of the diphosphocytidyl moiety (pocket I), the bindind site of 2*C*-methyl-D-erythritol 2-phosphate moiety (pocket II) and the pocket for cytidyl moiety (pocket III).

By ^{31}P NMR analysis and electrospray mass spectrometry a mixture of isopentenyl diphosphate (and/or dimethylallyl diphosphate), geranyl diphosphate and farnesyl diphosphate in an approximate ratio of 1:4:2 was identified in the core of the trimeric IspF protein (Richard *et al*., 2002). The modeling of geranyl diphosphate into the

electron density of IspF protein revealed protein-ligand interactions that are important to stabilize the binding of the isoprenoic compound (Kemp *et al*., 2004).

2 Objectives

This study involves the characterization, mechanistic study and inhibition of the enzymes IspC, IspE and IspF involved in the deoxyxylulose phosphate pathway of isoprenoid biosynthesis. Firstly, the recombinant enzymes should be purified by various column chromatographic techniques, and the enzyme substrates should be synthesized by enzyme-assisted procedures. Then, the IspC proteins from *Arabidopsis thaliana* and *Plasmodium falciparum*, the IspE protein from *Aquifex aeolicus* as well as the IspF from *Arabidopsis thaliana* should be enzymatically characterized and their kinetic properties should be determined. Furthermore, inhibition kinetics of chemical compounds against these proteins should be studied in comparison to the respective *E. coli* proteins. IC_{50} values and inhibitions constants (K_i) of the best inhibitors should then be determined.

Additionally, a mechanistic study of IspC proteins from various organisms will be carried out on the bases of NMR analyses.

Both, the finding and evaluating of inhibitors of the deoxyxylulose phosphate pathway as well as the elucidation of enzymatic mechanisms will be beneficial for the rational design of compounds that might be developed in the next future as new antiinfective agents against a broad spectrum of pathogenic microorganisms.

3 Material and Methods

3.1 Materials

3.1.1 Chemicals

Chemical substances are listed n Table 2.

Table 2: List of chemicals

Chemicals	Source
1-Propanol	Merck, Darmstadt, Germany
Acrylamide	Roth, Karlsruhe, Germany
Ammonium formate	Fluka, Neu-Ulm, Germany
Ammonium persulfate	Sigma, Deisenhofen, Germany
Ammonium sulfate	Merck, Darmstadt, Germany
Arginine/HCl	Biomol, Hamburg, Germany
Casein hydrolysate	Gibco-BRL, Eggenstein, Germany
Dimethylsulfoxide	Sigma, Deisenhofen, Germany
Dithioerythritol	Biomol, Hamburg, Germany
Dithiothreitol	Biomol, Hamburg, Germany
EDTA	Merck, Darmstadt, Germany
Hydrochloric Acid	Merck, Darmstadt, Germany
Imidazole	Sigma-Aldrich, Steinheim, Germany
IPTG (Isopropyl-1-thio-ß-D-galactopyranoside)	Eurogentec, Cambridge, England
Magnesium Chloride	Fluka, Neu-Ulm, Germany
Manganese Chloride	Sigma-Aldrich, Steinheim, Germany
Ortho-phosphoric acid 85%	Fluka, Neu-Ulm, Germany
Permablend®	Packard, Zürich, Switzerland
Phenethylamine	Fluka, Neu-Ulm, Germany
Phenylmethanesulfonyl fluoride	Merck, Darmstadt, Germany
Potassium chloride	Fluka, Neu-Ulm, Germany
Propylene glycol	Sigma-Aldrich, Steinheim, Germany
Serva Blue G (Coomassie Brilliant Blue G-250)	Serva, Heidelberg, Germany
Sodium azide	Fluka, Neu-Ulm, Germany
Sodium chloride	Fluka, Neu-Ulm, Germany
Sodium hydroxide	Merck, Darmstadt, Germany
TEMED	Roth, Karlsruhe, Germany
Tetrabutylammonium hydrogen-sulfate	Fluka, Neu-Ulm, Germany
Tris (Hydroxymethyl) aminomethane	Biomol, Hamburg, Germany
Yeast extract	Gibco-BRL, Eggenstein, Germany

3.1.2 Substrates, cofactors and NMR chemicals

The substrates, cofactors and NMR chemicals for the enzymatic assay are listed in Table 3.

Table 3: Substrates, cofactors and NMR chemicals for enzymatic experiments

Compound	Source
[3,4,5-$^{13}C_3$]-DXP	Victoria Illarionova
[1,3,4-$^{13}C_3$]-MEP	Victoria Illarionova
[1-$^{13}C_1$]-MEP	This study
[3-$^{13}C_1$]-MEP	This study
[1,3,4-$^{13}C_3$]-CDP-ME	Victoria Illarionova
[1,3,4-$^{13}C_3$]-CDP-MEP	Victoria Illarionova
[1,3,4-$^{13}C_3$]cMEPP	Victoria Illarionova
[1-$^{13}C_1$]glucose	Omicron (South Bend, Indiana, USA)
[3,4-$^{13}C_2$]glucose	Omicron (South Bend, Indiana, USA)
[2,5-$^{13}C_2$]glucose	Omicron (South Bend, Indiana, USA)
D-Glucose	Merck, Darmstadt
BSA	Biomol
Phosphoenol pyruvate	Merck, Darmstadt
$NADP^+$	Biomol
NADPH	Biomol
NADH	Biomol
NAD^+	Biomol
D-Lactate	Sigma Aldrich
Pyruvate	Sigma Aldrich
ATP	Sigma Aldrich
CMP	Sigma Aldrich
CTP	Sigma Aldrich
Fosmidomycin	Molecular Probes Invitrogen
Thiamin pyrophosphate	Sigma Aldrich
D_2O	Isotec .Inc., Ohio, USA

3.1.3 Enzymes

Recombinant and commercially-available enzymes used in this study are listed in Table 4 (cf. Table 7) and Table 5, respectively.

Table 4: Recombinant proteins used in this study expressed in the *E. coli* strains listed in Table 7

Enzyme	**Source**
1-Deoxy-D-xylulose 5-phosphate synthase from *Bacillus subtilis*	F. Rohdich, TU
2*C*-Methyl-D-erythritol 4-phosphate synthase (IspC) from *Escherichia coli*	F. Rohdich, TU
2*C*-Methyl-D-erythritol 4-phosphate synthase (IspC) from *Arabidopsis thaliana*	F. Rohdich, TU
2*C*-Methyl-D-erythritol 4-phosphate synthase (IspC) from *Mycobacterium tuberculosi*s	F. Rohdich, TU
2*C*-Methyl-D-erythritol 4-phosphate synthase (IspC) from *Plasmodium falciparum*	F. Rohdich, TU
4-Diphosphocytidyl-2*C*-methyl-D-erythritol kinase (IspE) from *Escherichia coli*	F. Rohdich, TU
4-Diphosphocytidyl-2*C*-methyl-D-erythritol kinase (IspE) from *Aquifex aeolicus*	William N. Hunter, University of Dundee, UK
2*C*-Methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF)from *Escherichia coli*	F. Rohdich, TU
2*C*-Methyl-D-erythritol 2,4-cyclodiphosphate synthase from *Arabidopsis thaliana*	J. Kaiser, TU
Fructose 1,6-biphosphate aldolase, Phosphofructokinase	F. Rohdich, TU
Glucose 6-phosphate isomerase from *E. coli*	W. Römisch, TU
6-Phosphogluconate dehydrogenase from *B. subtilis*	W. Römisch, TU

Table 5: Commercially available enzymes

Enzyme	Source
Hexokinase from yeast	Sigma-Aldrich (Germany)
Triosephosphate isomerase from rabbit muscle	Sigma-Aldrich (Germany)
Glutamate dehydrogenase from bovine liver	Sigma-Aldrich (Germany)
Glucose dehydrogenase from *Thermoplasma acidophilum*	Sigma-Aldrich (Germany)
Pyruvate kinase	Sigma-Aldrich (Germany)
Lactose dehydrogenase from rabbit muscle	Sigma-Aldrich (Germany)

3.1.4 Chromatographic materials

FPLC Sepharose Q FF, Phenyl Sepharose 6FF, Source 15 Q, Superdex 200 HR 26/60, Superdex 75 HR 26/60, Resource PHE, Mono Q HR5/5, Red Sepharose CL-6B, Chelating Sepharose FF (Amersham Pharmacia Biotech, Freiburg, Germany.), HIPREP 26/10 column (Amersham)

HPLC Rezex RPM Monosaccharide (Phenomenex, Hösbach, Germany) Nucleosil 10 SB, Nucleosil 5 SB (Macherey Nagel, Düren, Germany),

Multospher 120 RP18-5 (CS: Chromatography Service, Langerwehe, Germany)

TLC POLYGRAM® CEL 300 PEI, POLYGRAM® Sil N-HR/UV$_{254}$ (Macherey Nagel, Düren, Germany) Silica gel 60 F$_{254}$ (Merck, Darmstadt, Germany)

PC Whatman No.3 (Whatman, Springfield Mill, England)

3.1.5 Buffers and solutions

3.1.5.1 Protein determination

Bradford reagent

0.1 g of Serva Blue G was dissolved in 10 ml of ethanol and 25 ml of H_3PO_4 and stirred thoroughly for 10 min. The volume of the mixture was adjusted to 250 ml with distilled water, and the solution was filtered and stored in a dark colored bottle at room temperature.

BSA solution

10 mg of bovine serum albumine (BSA) was dissolved in distilled water. The volume of the solution was adjusted to 100 ml with distilled water. This solution was used as a standard for protein determinations.

3.1.5.2 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Acrylamide Solution 40%	38.9 g Acrylamide
	1.2 g Bis-acrylamide
	The volume was adjusted with distilled water to 100 ml
Coomassie blue staining	2.5 g Coomassie-Blue R-250
	454 ml Methanol
	92 ml Acetic acid
Destaining solution	250 ml Methanol
	100 ml Acetic acid
	650 ml H_2O
Stacking gel buffer	0.25 M Tris/HCl, pH 6.8
	0.2 % (w/v) SDS
SDS-PAGE running buffer	192 mM Glycine
	25 mM Tris/HCl, pH 8.3
	0.1 % (w/v) SDS
SDS-PAGE sample buffer	60 mM Tris/HCl, pH 6.8
	5 % (w/v) SDS
	3% 2-Mercaptoethanol
	30 % (v/v) Glycerol
	0.02 % (w/v) Bromphenol Blue
	10 % (w/v) Sucrose
Separating gel buffer	1.5 M Tris/HCl, pH 8.8
	0.4 % (w/v) SDS

3.1.5.3 Buffer mix for measuring pH optimum of the catalyzed reaction

10x Buffer mix	0.5 M Glycin
	0.2 M Phosphate
	0.5 M Sodium acetate
	0.5 M Tris/HCl
	0.02 % (w/v) NaN_3

3.1.6 Culture media

Complete medium

LB-medium (Luria-Bertani) contained, per liter: 10 g of casein hydrolysate, 5 g of yeast extract and 5 g of NaCl.

LB amp contained per liter: 180 mg ampicillin trihydrate, dissolved in 1 ml of sterilized water and added to 1 l of sterilized LB-medium.

LB amp/kan contained per liter: 180 mg ampicillin trihydrate and 50 mg kanamycin sulphate, dissolved in 1 ml of sterilized water and added to 1 l of sterilized LB-medium.

The components of media were dissolved in deionized water. The medium was sterilized by autoclaving (121 °C, 1 bar, 25 min). Supplements were added as defined.

3.1.7 *Escherichia coli* strains

Recombinant and wild type .*E coli* strains are listed in Table 6 and Table 7

Table 6: Wild type strains of *Escherichia coli*

Strain	Genotype	Reference
XL1-Blue	*RecA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac*, [F´,*proAB, lacIqZDM15*, *Tn10 (tetr*)]	Bullock *et al.*, 1987
M15[pREP4]	*Lac*, *ara, gal*, *mtl*, *recA+, uvr+,* [pREP4, *lacI*, *kanr*]	Zamenhof, P.J. & Villarejo, M. (1972)

Table 7: Recombinant strains of *Escherichia coli* (cf. Table 4)

Strain	Recombinant enzyme	Source
M15p[REP4]pQE30ispC	IspC from *E. coli*	F. Rohdich, TU
XL1-pQEispCara1	IspC from *A. thaliana*	F. Rohdich, TU
XL1-pQEispCara2	IspC from *A. thaliana*	F. Rohdich, TU
XL1-pQEispCara3	IspC from *A. thaliana*	F. Rohdich, TU
XL1-pQEispCara4	IspC from *A. thaliana*	F. Rohdich, TU
XL1-pQEispCMyco	IspC from *M. tuberculosis*	F. Rohdich, TU
XL1-pQEispCplas2	IspC from *P. falciparum*	F. Rohdich, TU
M15pRep4pQE30ispF E.coli2	IspF from *E. coli*	F. Rohdich, TU
M15p[REP4]pQE30ispFara5	IspF from *A. thaliana*	J. Kaiser, TU

3.2 Instruments

FPLC (Amersham Pharmacia Biotech, Freiburg, Germany)
Column for HPL System, Conductivity Monitor, FRAC-100 Fraction Collector, LCC-500 PLUS Controller, MV-7 Motor Valve, P-500 Pump,

	REC100 Recorder, Superloop 10 ml, UV-1 Monitor.
CENTRIFUGE	A14 LabTop Microcentrifuge (jouan,Unterhaching, Germany), Servall Superspeed RC-2B and RC5B Plus with rotors GS3, GSA, and SS34 (Dupont Instruments, Bad Hamburg, Germany), GS-15R Refrigerated centrifuge with rotor S1480 and F2402 (Becman, Fullerton, USA)
Electrophoresis	SE 250 Mighty small II (Hoefer Scientific, San Fransisco, USA)
Power supply	PHERO-stab 200 (Reiskirchen, Ger-many), Biocapt camera software LTF Labortechnik (Wasserburg, Germany)
Heat block	Techne DRI BLOCK DB-2A (Wartheim, Germany)
Incubator	Thermostat 340 (Eppendorf, Hamburg, Germany)
Lyophylizer	CHRIST ALPHA 1-4 (Osterode am Harz, Germany)
NMR	AVANCE DRX500 Spectrometer (Bruker, Karlsruhe, Germany)
PH Meter	E512 Metrohm AG CH-9100 (Herisau, Switzerland)
Protein concentrator	Amicon (Witten, Germany), and membrane: PALL Filtron (Dreieich, Germany)
Shaker	Labshaker and Certomat MO (B.Braun, Melsungen, Germany)
Spectrophotometer	Ultrospec 2000 (Amersham Pharmacia Biotech, Freiburg, Germany) equipped with control temperature holder.
Ultrasonicator	Sonifier B-12 (Branson SONIC Power

	Company, USA)
Autoclave	Sanoclav (Wolf; Geislingen)
Plate reader	Molecular devices, Spectra Max M_2, computer software, Promax, NUNC™ 96 well plate (Nalge Nunc International)
Thermo multiscan spectrum	Thermo electron corporation, SkanIt® software

3.3 Methods

3.3.1 Culture condition

Precultures were made from frozen cultures and incubated in the LB medium with appropriate antibiotics, and incubated at 37 °C (140 rpm) overnight. The overnight culture was subcultured into a 500 ml flask containing the same medium. The cells were grown at 37 °C (140 rpm) for 5 h or 27 °C overnight and harvested by centrifugation at 4°C, 4000 rpm for 30 min, washed in saline (0.9% w/v NaCl) and stored at –20 °C.

3.3.2 Protein Chemical methods

3.3.2.1 Cell extraction

Frozen cell mass of *E. coli* was thawed in buffer containing 100 mM Tris/HCl, pH 8.0 and 150 mM KCl, DNA-ase I and Pefablock (1 mg/100 ml crude extract). The cells were disrupted by ultrasonication (70 % duty cycle, output 4, 10 x 1 min, 1 min pause). The suspension was centrifuged (rotor GSA, 13,000 rpm, 4 °C, 15 min), and the supernatant was used as crude extract.

Note: For individual recombinant protein, the Tris/HCl buffer might be adjusted to a pH which is different from the one indicated above. The detail of the procedures and buffers are explained in the Result part.

3.3.2.2 Protein purification

3.3.2.2.1 Preparation of recombinant protein

The crude extract was applied to a column of Ni-chelating Sepharose FF (Amersham

Pharmacia Biotech; column volume, 34 ml) which had been equilibrated with 50 mM Tris/HCl, pH 8.0, containing 0.5 M NaCl and 20 mM imidazole (flow rate, 3 ml min^{-1}). The column was washed with 50 mM Tris/HCl, pH 8.0, containing 0.5 M NaCl and 20 mM imidazole and was then developed with a gradient of 20 - 500 mM imidazole in 50 mM Tris/HCl, pH 8.0, containing 0.5 M NaCl. Fractions were combined and imidazole was removed by dialysis overnight or by a desalting column HIPREP 26/10 (Amersham) with buffer containing 50 mM Tris/HCl, pH 8.0, 100 mM NaCl, 2 mM DTT and 0.02 % (w/v) NaN_3. The purified protein was stored at -80 °C.

Note: For individual recombinant protein, the Tris/HCl buffer might be adjusted to a pH which is different from the one indicated above. The detail of the procedures and buffers are explained in the Result part.

3.3.2.2.2 Preparation of recombinant protein from *inclusion bodies*

Cells were disrupted according to paragraph 3.4.2.1. After centrifuged (rotor GSA, 13,000 rpm, 4 °C, 15 min), the pellet was dissolved in a buffer containing 50 mM Tris/HCl, pH 8.0, 8 M urea and 0.15 M KCl. The suspension was centrifuged at 15,000 rpm for 30 min. The supernatant was applied to a column of Ni-chelating Sepharose FF (Amersham Pharmacia Biotech) which had been equilibrated with 50 mM Tris/HCl, pH 8.0, containing 8 M urea, 0.15 M KCl and 20 mM imidazole (flow rate, 3 ml min^{-1}). The column was washed with 50 mM Tris/HCl, pH 8.0, containing 8 M urea, 0.15 M KCl and 20 mM imidazole. For renaturation, the column was washed with 10 times column volume of 50 mM Tris/HCl, pH 8.0, containing 0.5 M urea, 0.15 M KCl and 20 mM imidazole and was then developed with a gradient of 20 - 500 mM imidazole in 50 mM Tris/HCl, pH 8.0, containing 0.5 M urea, 0.15 M KCl and 500 mM imidazole. Fractions were combined and dialyzed overnight with dialysis buffer containing 50 mM Tris/HCl, pH 8.0, containing 1 M arginine, 0.05 M KCl, 1 mM dithiothreitol and 0.02 % (w/v) NaN_3 and stored at –80 °C.

3.3.2.3 Determination of protein concentration

The amount of protein was determined by the dye-binding method (Bradford assay) modified by Read and Northcote (1981).

Procedure

950 µl of Bradford reagent was added to a 50 µl protein sample or standard protein solution (0.01 - 0.1 mg ml^{-1}) in a 1 ml cuvette. As a reference, 50 µl of buffer without protein was mixed with 950 µl of Bradford reagent. The absorbance was measured at 578 nm. The amount of protein was estimated from a standard curve using BSA as standard protein.

3.3.2.4 Polyacrylamide gel electrophoresis

The purity of protein fractions was measured by discontinuous sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to the system of Lämmli (1970).

Procedure

The gel was prepared using the SDS-PAGE discontinuous buffer system with vertical slab gels (Table 8). The components of the separating gel solution were mixed together and then loaded onto the slab deposited between two glass plates on the gel caster. The top of the gel was overlayed with isopropanol. The polymerization of acrylamide was completed after 1 h. After removing the isopropanol, the stacking gel solution was prepared and loaded on top of the separating gel in the presence of 10 wells comb. The stacking gel completed its polymerization after 20 min. The slab gel was placed on a SE 250 Mighty Small II electrophoresis system (Hoefer Scientific, San Francisco, USA) which was connected to a cooling system. Protein samples were mixed with SDS-sample buffer at a ratio of 1:1. The mixtures were heated at 95 °C for 5 min. Then the protein samples were loaded into the wells. Standard proteins were used as molecular weight markers. The electrophoresis system was powered from a power supply with 25 mA per gel. After running for 45 min, the gel was removed carefully and stained in staining solution for 30 min. The gel was then destained with destaining solution for 1 h.

Table 8: Mixture for SDS –PAGE electrophoresis

Stock solution	Stacking gel	Separating gel (15 % Acrylamide)
Acrylamide (40 % T, 3 % C)	0.5 ml	3.75 ml
Stacking gel buffer	2.5 ml	-
Separating gel buffer	-	2.5 ml
Water	2.0 ml	3.75 ml
10% (w/v) APS	50 µl	50 µl
TEMED	5 µl	5 µl

3.3.3 Preparation of [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate

3.3.3.1 Synthesis of [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate

A solution containing 18 mM [3,4-$^{13}C_2$]D-glucose or [2,5-$^{13}C_2$]D-glucose, 36 mM PEP, 1 mM TPP, 0.7 mM ATP, 10 mM $MgCl_2$, 5 mM DTT and 150 mM Tris/HCl was adjusted to pH 8.0 by the addition of 5 N NaOH (final volume, 10 ml). Hexokinase (300 µg, 27 units), phosphofructokinase (400 µg, 2 units), fructose 1,6-biphosphate aldolase (540 µg, 2 units), triose phosphate isomerase (4 µg, 21 units), glucose 6-phosphate isomerase (18 µg, 9 units) and Dxs protein from *B. subtilis*. (690 µg, 2 units) were added, and the pH was adjusted to 8.0. The reaction mixture was incubated at 37 °C and monitored by ^{13}C NMR spectroscopy. After 3 h, protein was removed by ultrafiltration (10 kDa cutoff).

$NADP^+$ (8.4 mg, 10 µmol), glucose dehydrogenase (57 µg, 12 units), D-glucose (75 mg) and IspC protein from *E. coli* (800 µg) were added, and the pH was adjusted to pH 8.0 by the addition of 5 N NaOH (final volume, 11 ml). The reaction was controlled by ^{13}C NMR spectroscopy. After 3 h, proteins were removed by ultrafiltration (10 kDa cutoff). The solution was lyophilized.

The residue was dissolved in 3 ml of a solution containing 40 % (v/v) methanol and 20 % (v/v) isopropanol. The solution was applied to a column of microcrystalline cellulose (volume, 40 ml) that had been equilibrated with the same solution. The column was developed with the same solution. Fractions were analyzed by thin layer chromatography and ^{13}C NMR spectroscopy and were then combined and

evaporated to a small volume under reduced pressure. The solution was lyophilized. The residue was dissolved in H_2O and stored at -80 °C.

3.3.4 Synthesis of [U-$^{13}C_5$]ribulose 5-phosphate

[U-$^{13}C_5$]ribulose 5-phosphate was prepared according to published procedures with slight modifications (Volk and Bacher, 1991). A solution containing 18.5 mM [U-$^{13}C_6$]D-glucose, 100 mM Tris/HCl, pH 8.0, 80 mM DTT, 10 mM $MgCl_2$, 30 mM ATP, 81 mM ammonium acetate, 81 mM α-ketoglutarate and 1.6 mM $NADP^+$ was adjusted to pH 8.0 by addition of 1 M NaOH. The reaction was started by addition of hexokinase (8 mg, 300 units), 6-phosphogluconate dehydrogenase (1 mg, 25 units), glucose-6-phosphate dehydrogenase from *B. subtilis* (200 µg, 62 units) and glutamate dehydrogenase from bovine liver (3 mg, 100 units) (final volume, 37 ml). The reaction mixture was incubated at 37 °C, and the reaction was monitored by ^{13}C NMR spectroscopy. After 1 h, the protein was removed by ultrafiltration (3 kDa), and the solution was lyophilized overnight.

3.3.5 Synthesis of [1,2-$^{13}C_2$]glycoaldehyde phosphate

2.5 g of residue from lyophilization (cf. paragraph 3.4.4) was dissolved in 6 ml H_2O containing 0.2 mmol [U-$^{13}C_5$]ribulose 5-phosphate, and the pH was adjusted to 6. The diol cleavage was induced by the addition of 3 mmol of sodium metaperiodate, and the mixture was incubated for 5-10 min. The formation of the product [1,2-$^{13}C_2$]glycoaldehyde phosphate was monitored by ^{13}C NMR spectroscopy. The excess of periodate was quenched with 10 mmol glycerol, and [1,2-$^{13}C_2$]glycoaldehyde phosphate was purified with Dowex Cl^- (anion exchanger).

3.3.6 Enzymatic assays

3.3.6.1 IspC enzymatic assays

3.3.6.1.1 Photometric assay of IspC enzymatic activity

A. Assay for forward reaction: Assay mixtures contained 100 mM Tris/HCl, pH 7.5, 0.4 mM NADPH, 2 – 10 mM Mn^{2+} or Mg^{2+}, 1 mM DXP and IspC protein in a total volume of 200 µl. The assay mixtures were incubated at 37 °C, and the

reaction was monitored spectrophotometrically at 340 nm.

B. Assay for backward reaction: Assay mixtures contained 100 mM Tris/HCl, pH 7.5, 1 mM $NADP^+$, 2 – 10 mM Mn^{2+} or Mg^{2+}, 10 mM MEP and IspC protein in a total volume of 200 µl. The assay mixtures were incubated at 37 °C, and the reaction was monitored spectrophotometrically at 340 nm.

3.3.6.1.2 Determination of the inhibition constant (K_i) of the inhibitor fosmidomycin by initial rate measurements

The assay was prepared as described in paragraph 3.3.6.1.1 using varied substrate concentration and the addition of inhibitor at various concentrations as seen in Table 9.

Table 9: Assay mixtures for K_i measurement of the inhibitors of the reaction catalyzed by IspC protein

Conditions	IspC from *A. thaliana* (Forward reaction)	IspC from *P. falciparum*	
		Forward reaction	Backward reaction
Tris/HCl, pH 7.5	100 mM	100 mM	100 mM
$MgCl_2$	10 mM		
$MnCl_2$		2 mM	2 mM
NADPH	1 mM	1 mM	
$NADP^+$			1.5 mM
IspC protein	As needed	As needed	As needed
DXP	176 - 8000µM	80-911 µM	
MEP			80-1367 µM
Total volume	500 µl	200 µl	
Instrument	Spetrophotometer	Spectrophotometer with plate reader and 96 well plate	
Temperature	37 °C	30 °C	

3.3.6.1.3 ^{13}C NMR spectroscopic assay

3.3.6.1.3.1 ^{13}C NMR spectroscopic assay of the forward reaction of IspC

Assay mixtures contained 100 mM Tris/HCl, pH 7.5, 10 mM metal ion, 3 - 6 mM NADPH, 5 - 10 mM [3,4,5-$^{13}C_3$]-DXP, 10 % (v/v) D_2O and IspC protein in a volume of 500 µl. The mixtures were incubated at 37 °C and terminated by the addition of EDTA to a final concentration of 30 mM. The solutions were analyzed by ^{13}C NMR spectroscopy.

3.3.6.1.3.2 ^{13}C NMR spectroscopic assay of the backward reaction of IspC using a NADPH recycling system

Assay mixtures contained 200 mM Tris/HCl, pH 7.5, 15 mM $MgCl_2$, 15 mM $NADP^+$, 10 mM [1,3,4-$^{13}C_3$]-MEP, lactate dehydrogenase (500 µg, 880 units), 15 mM pyruvate kinase and IspC protein in a volume of 500 µl. The mixture was incubated at 37 °C and was terminated by the addition of EDTA to a final concentration of 30 mM. D_2O was added to a final concentration of 10 % (v/v). The solution was analyzed by ^{13}C NMR spectroscopy.

3.3.6.1.3.3 ^{13}C NMR spectroscopic assay of the backward reaction of IspC using an excess of $NADP^+$

Assay mixtures contained 100 mM Tris/HCl, pH 7.5, 10 mM $MgCl_2$, 215 mM $NADP^+$, 10 mM labeled-MEP, 10 % (v/v) D_2O and IspC protein in a volume of 500 µl. The mixtures were incubated at 37 °C and analyzed by ^{13}C NMR spectroscopy.

3.3.6.1.3.4 ^{13}C NMR study of IspC mechanism

The assay is described in paragraph 3.3.6.1.3.3 but using the substrates and cofactors as seen in Table 10.

Table 10: Substrates used in the ^{13}C NMR assay for the mechanistic study of IspC protein

Assay	Substrates	Cofactor
1	5 mM [1-$^{13}C_1$]-MEP and 5 mM [3-$^{13}C_1$]-MEP	$NADP^+$
2	10 mM [1,3,4-$^{13}C_3$]-MEP and 243 mM unlabeled hydroxyacetone	$NADP^+$
3	2 mM [1,2-$^{13}C_2$]glycoaldehyde phosphate and 243 mM unlabeled hydroxyacetone	NADPH

3.3.6.1.3.5 ^{13}C NMR study of the isomerization from 1-deoxy-D-xylulose 5-phosphate

Assay mixtures contained 100 mM Tris hydrochloride pH 8.0, 10 % (v/v) D_2O, 10 mM [3,4,5-$^{13}C_3$]-DXP, 2.18 mM [1-$^{13}C_1$]D-glucose, IspC enzymes in various concentrations and H_2O in a total volume of 500 µl. The samples were incubated at 37 °C and analyzed by ^{13}C NMR spectroscopy.

3.3.6.2 IspE enzymatic assays

3.3.6.2.1 Photometric assay for IspE enzymatic activity

Assay mixtures contained 100 mM Tris/HCl, pH 8.5, 0.45 mM NADH, 2 mM ATP, 4 mM PEP, 20 mM KCl, 10 mM $MgCl_2$, 1 mM CDP-ME, pyruvate kinase (10 µg, 2 units), lactate dehydrogenase (1 µg, 2 units) and IspE protein in a total volume of 200 µl. The assay mixtures were incubated at 37 °C, and the reaction was monitored spectrophotometrically at 340 nm in a plate reader.

3.3.6.2.2 ^{13}C NMR assay for IspE enzymatic activity

Assay mixtures contained 100 mM Tris/HCl, pH 8.5, 2 mM $MgCl_2$, 2.5 mM [1,3,4-$^{13}C_3$]CDP-ME, 8 mM ATP, IspE protein, 5 mM divalent metal ions, and 10 %

(v/v) D_2O in a total volume of 500 µl. The mixtures were incubated at 37 °C for 30-60 min and terminated by the addition of EDTA to a final concentration of 30 mM. The solution was analyzed by ^{13}C NMR spectroscopy.

3.3.6.2.3 Kinetics of IspE inhibitors

3.3.6.2.3.1 IC_{50} value measurement

Assay mixtures contained 100 mM Tris/HCl, pH 8.5, 0.45 mM NADH, 2 mM ATP, 4 mM PEP, 20 mM KCl, 10 mM $MgCl_2$, 1 mM CDP-ME, pyruvate kinase (10 µg, 2 units), lactase dehydrogenase (1 µg, 2 units), IspE protein and the inhibitor in a total volume of 200 µl. The assay mixtures were incubated at 37 °C, and the reaction was monitored spectrophotometrically at 340 nm in a plate reader.

3.3.6.2.3.2 K_i value measurement

Assay mixtures contained 100 mM Tris/HCl, pH 8.5, 0.45 mM NADH, 2 mM ATP, 4 mM PEP, 20 mM KCl, 10 mM $MgCl_2$, pyruvate kinase (10 µg, 2 units), lactate dehydrogenase (1 µg, 2 units), IspE protein, 53-600 µM CDP-MEP, and inhibitor at various concentrations in a total volume of 200 µl. The assay mixtures were incubated at 24 °C and the reaction was monitored spectrophotometrically at 340 nm in a plate reader. The K_i value and the mode of inhibition were determined with the *Dynafit* program.

3.3.6.3 IspF enzyme assays

3.3.6.3.1 Photometric assay of IspF enzymatic activity

Assay mixtures contained 100 mM Tris/HCl, pH 8.0, 0.45 mM NADH, 2 mM ATP, 4 mM PEP, 20 mM KCl, 10 mM $MgCl_2$, 1 mM CDP-MEP, pyruvate kinase (10 µg, 2 units), lactose dehydrogenase (1 µg, 2 units), CMP kinase (25 µg, 0.5 U) and IspF protein in a total volume of 200 µl. The assay mixtures were incubated at 24 °C, and the reaction was monitored spectrophotometrically at 340 nm in a plate reader.

3.3.6.3.2 ^{13}C NMR spectroscopy assay of IspF enzyme activity

3.3.6.3.2.1 ^{13}C NMR spectroscopy assay of the forward reaction of IspF

Assay mixtures contained 100 mM Tris/HCl, pH 8.0, 10 mM $MgCl_2$, 1 mM [1,3,4-$^{13}C_3$]-CDP-MEP and IspF protein in a volume of 500 µl. The mixture was incubated at 37 °C and was terminated by the addition of EDTA to a final concentration of 30 mM. D_2O was added to a final concentration of 10 % (v/v). The solution was analyzed by ^{13}C NMR spectroscopy.

3.3.6.3.2.2 ^{13}C NMR spectroscopy assay of backward reaction of IspF

Assay mixtures contained 100 mM Tris/HCl, pH 8.0, 10 mM $MgCl_2$, 200 mM CMP, 5 mM [1,3,4-$^{13}C_3$]-cMEPP and IspF protein in a volume of 500 µl. The mixture was incubated at 37 °C and terminated by the addition of EDTA to a final concentration of 30 mM. D_2O was added to a final concentration of 10 % (v/v). The solution was analyzed by ^{13}C NMR spectroscopy.

3.3.6.4 Kinetics of IspF inhibitors

3.3.6.4.1 IC_{50} value measurement

A. Photometric assay: Assay mixtures contained 100 mM Tris/HCl, pH 8.0, 0.45 mM NADH, 2 mM ATP, 4 mM PEP, 20 mM KCl, 10 mM $MgCl_2$, 1 mM CDP-MEP, pyruvate kinase (10 µg, 2 units), lactose dehydrogenase (1 µg, 2 units), CMP kinase (25 µg, 0.5 U), IspF protein and inhibitor at various concentrations in a total volume of 200 µl. The assay mixtures were incubated at 37 °C, and the reaction was monitored spectrophotometrically at 340 nm in a plate reader.

B. ^{13}C NMR assay: Assay mixtures contained 100 mM Tris hydrochloride, pH 8.0, 10 mM $MgCl_2$, 1 mM [1,3,4-$^{13}C_3$]-CDP-MEP, inhibitor in a range of 0 - 250 µM and IspF protein in a total volume of 500 µl. The mixture was incubated at 37 °C and terminated by the addition of EDTA to a final concentration of 30 mM. D_2O was added to a final concentration of 10 % (v/v). The solution was analyzed by ^{13}C NMR spectroscopy.

3.3.6.4.2 *K*i measurement of IspF inhibitors

Assay mixtures contained 100 mM Tris/HCl, pH 8.0, 0.45 mM NADH, 2 mM ATP, 4 mM PEP, 20 mM KCl, 10 mM $MgCl_2$, 53 - 600 µM CDP-MEP, pyruvate kinase (10 µg, 2 units), lactose dehydrogenase (1 µg, 2 units), CMP kinase (25 µg, 0.5 U), IspF protein and inhibitor at various concentrations in a total volume of 200 µl. The assay mixtures were incubated at 37 °C, and the reaction was monitored spectrophotometrically at 340 nm in a plate reader. The K_i and mode of inhibition were analyzed by *Dynafit* program.

4 Result and Discussion

4.1 IspC protein

IspC protein catalyzes the conversion of the substrate 1-deoxy-D-xylulose 5-phosphate (**12**) into the product 2*C*-methyl-D-erythritol 4-phosphate (**13**) with 2*C*-methyl-D-erythrose 4-phosphate (**25**) as intermediate (Figure 8).

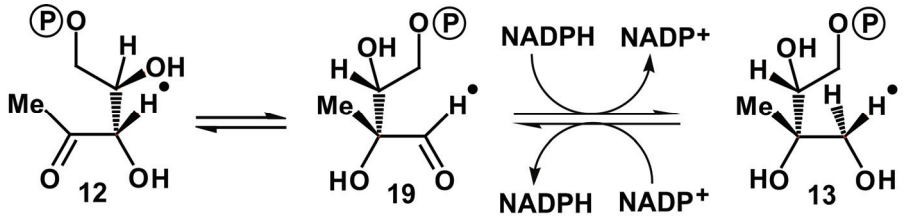


Figure 8: The reaction catalyzed by the IspC protein
12, 1-deoxy-D-xylulose 5-phosphate (DXP); **19**, 2*C*-methyl-D-erythrose 4-phosphate; **13**, 2*C*-methyl-D-erythritol 4-phosphate.

4.1.1 Characterization of the IspC protein from *Arabidopsis thaliana*

4.1.1.1 Cloning and expression of the *ispC* gene from *Arabidopsis thaliana*

The predicted ORF of the *ispC* gene of *A. thaliana* (chromosome V, accession number NC_003076, bp 3 629 999 – 3 633 442; cDNA sequence, accession number AF148852) was amplified by PCR from cDNA and cloned into the plasmid pQE30. DNA sequencing indicated that the resulting plasmid construct pQE30ispCara1 contains an ORF of 1467 codons specifying the vector-derived coding region MRGSHHHHHHGS followed by the complete IspC protein sequence (residues 1 – 477; Figure 9). The recombinant gene could be expressed at a low level in *E. coli* host cells, but the protein was insoluble, and in vitro renaturation experiments were unsuccessful (Rohdich *et al*., 2006).

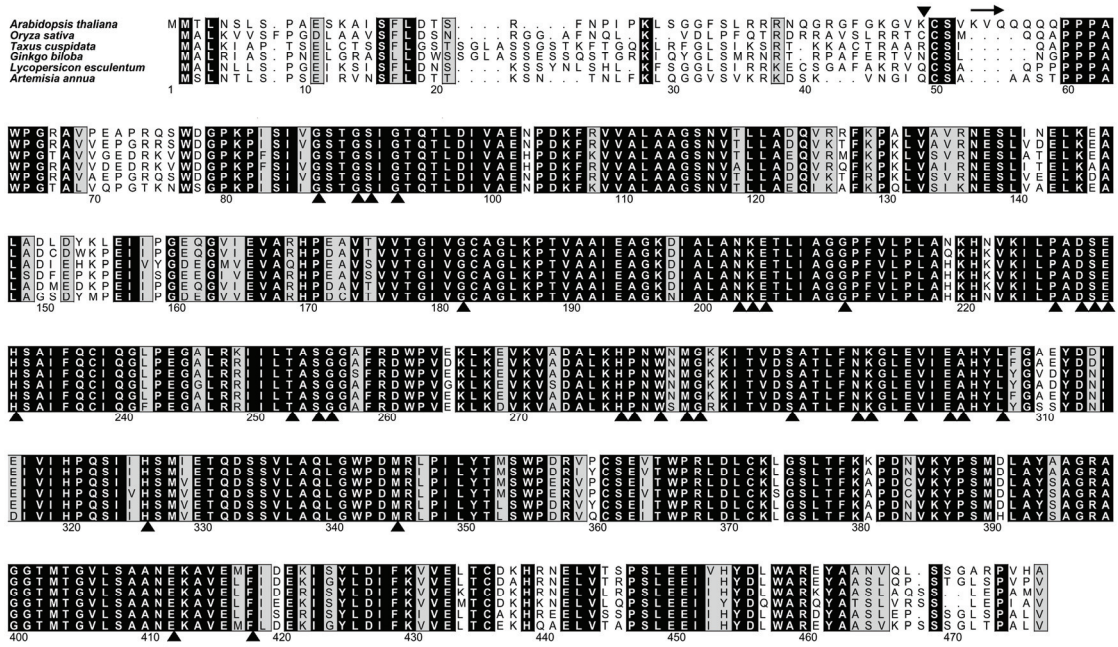


Figure 9: Alignment of IspC amino acid sequences from plants *Arabidopsis thaliana* (Eurosids), *Oryza sativa* (Poaceae), *Taxus cuspidata* (Taxus), *Ginkgo biloba* (Ginkgo), *Lycopersicon esculentum* (Lamiids), *Artemisia annua* (Asteroideae) Inverted triangle, cleavage site predicted for all plant IspC amino acid sequences with the CHLOROP 1.1 server from the Center For Biological Sequence Analysis (http://www.cbs.dtu.dk/services/ChloroP/); arrows, N-terminal region corresponding to the 5´-end of the cloned and successfully expressed *ispC* gene fragment from *A. thaliana*; triangles, absolutely conserved residues in IspC proteins of the bacterial kingdom.

The length of the *A. thaliana* IspC protein exceeds that of bacterial orthologs by approximately 80 amino acid residues. The N-terminal segment of the plant protein has the characteristic features of a plastid-targeting sequence, as shown by computer analysis. This is well in line with the plastid location that has been documented for IspC and other proteins of the nonmevalonate pathway in plants (Carretero-Paulet *et al*., 2002; Guevara-Gàrcia *et al*., 2005; Hsieh *et al*., 2005). It was therefore appropriate to investigate the expression of the plant IspC protein in a pseudomature form. On the basis of the sequence comparison between the *A. thaliana* protein and its bacterial orthologs, several expression constructs were obtained, directing the expression of recombinant proteins starting with amino acid residues 57 – 95 (Figure 10A). In order to enable rapid affinity chromatography purification of native or denatured protein, the recombinant gene constructs were designed to specify N-terminal polyhistidine tags.

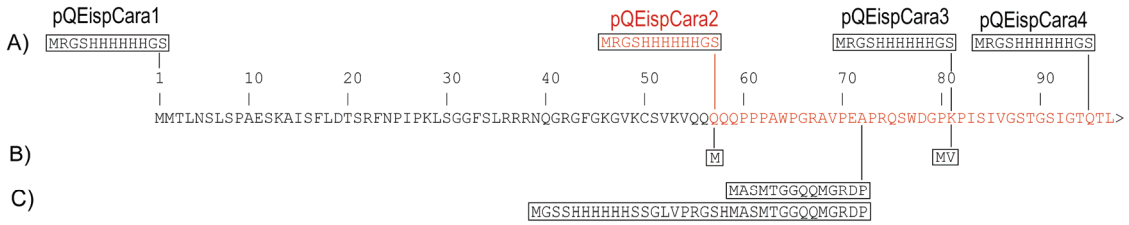


Figure 10: N-terminal region of the IspC protein from *Arabidopsis thaliana*
Starting regions and N-terminal extensions of expressed constructs reported in different papers are indicated: (A) this study; (B) Campos and coworkers (Carretero-Paulet *et al*., 2002); (C) Lichtenthaler and coworkers (Schwender *et al*., 1999; Müller *et al*., 2000). The N-terminal sequence of the expressed construct used for high-yield production of recombinant *A. thaliana* IspC protein is indicated in red.

The constructs pQE30ispCara3 and pQE30ispCara4, starting with amino acid residues 81 and 95, respectively, could be expressed in insoluble form (inclusion bodies) in recombinant *E. coli* cells. The crude cell extracts showed no IspC activity in excess of the background activity provided by the bacterial *ispC* gene. Attempts at in vitro folding of the insoluble proteins under a variety of experimental conditions (3.3.2.2.2) also failed to yield detectable IspC protein activity that could be attributed to the recombinant plant protein. The recombinant proteins could be purified by affinity chromatography under denaturing conditions, but all attempts to renaturate the purified proteins were unsuccessful.

A detailed sequence comparison, including all putative IspC amino acid sequences from plants that are available in public databases (representatives of each family are shown in Figure 9), revealed, among other distinctive features, a proline-rich block of seven contiguous amino acids residues (PPPAWPG) located upstream of the N-terminus of bacterial orthologs that was absolutely conserved in all plant sequences available in the databases. Moreover, a CS motif and a tryptophan residue (residues 50/51 and 77 of the *A. thaliana* protein, respectively) were absolutely conserved within the N-termini of the plant proteins. These findings suggested that the N-terminus of the mature plant enzymes is located significantly upstream in comparison with bacterial enzymes (this part of the sequence had been attributed, hypothetically, to the plastid-targeting sequence in our own initial studies, as well as in studies by other authors). Consequently, an approximate length of only about 50 amino acid residues would be predicted for the plastid-targeting sequences of plant IspC

proteins, and the length of the mature protein would exceed that of the bacterial enzymes by about 30 amino acid residues. Notably, a plastid-targeting sequence with a cutting sequence at the conserved CS motif (residues 50/51 of the *A. thaliana protein*) (Figure 9, down arrow) is predicted by the chlorop 1.1 server (http://www.cbs.dtu.dk/services/ Chloro/) for all currently known plant sequences.

On the basis of these findings, an ORF specifying a polyhistidine tag followed by amino acid residues 57 – 477 of the *A. thaliana ispC* gene was constructed (clone pQE30ispCara2; see Figure 10A, sequence labeled in red).

4.1.1.2 Purification of IspC protein from *Arabidopsis thaliana*

The expression pattern of cell extracts from the recombinant strain (Figure 11, lane C) differed significantly from that of the *E. coli* host strain, but showed no visible expression of the recombinant protein as judged by SDS-PAGE (Figure 11, lane B). However, these cell extracts were shown to catalyze the conversion of 1-deoxy-D-xylulose 5-phosphate (**12**) into 2*C*-methyl-D-erythritol 4-phosphate (**13**) at a rate of 90 nmol min^{-1} mg^{-1} as compared to a rate of about 10 nmol min^{-1} mg^{-1} in cell extracts of the host strain. Affinity chromatography on nickel-chelating Sepharose followed by desalting (see Methods 3.3.2.2.1) afforded a protein fraction that appeared pure as judged by SDS-PAGE (Figure 11, lane D).

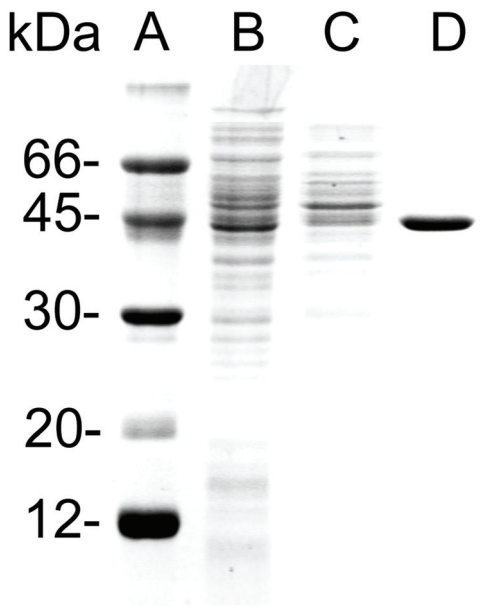


Figure 11: Purification of recombinant IspC protein from *Arabidopsis thaliana*
SDS-PAGE: lane A, molecular weight markers; lane B, cell extract of *E. coli* wild-type; lane C, cell extract of a recombinant *E. coli* strain expressing the *ispC* gene from *A. thaliana* (XL1- pQE30ispCara2); lane D, recombinant IspC protein after affinity chromatography on Ni-chelating Sepharose FF and desalting chromatography on an HIPREP 26/10.

Using a photometric assay to monitor the consumption of NADPH (see Methods 3.3.6.1.1A), the protein was shown to catalyze the formation of compound **12** from compound **13** at a rate of 5.6 µmol min^{-1} mg^{-1} (Table 11). The apparent purification factor of about 60 as judged from the specific activities of the crude extract and the purified protein suggests that the recombinant protein had been expressed at a level of about 2% on the basis of total cell protein (Table 11).

Table 11: Purification of recombinant IspC protein from *Arabidopsis thaliana*

Procedure	Total protein (mg)	Specific activity (nmol min^{-1} mg^{-1})	Total activity (nmol min^{-1})	Yield (%)	Purification factor
Cell extract	1523	90	137,070	100	1
Ni^{2+}-Sepharose FF/ HIPREP 26/10	9	5600	50,400	37	62

Partial Edman degradation afforded the N-terminal sequence MRGSHHHHHH GSQQQPPPAW, in agreement with the DNA sequence of the plasmid construct pQE30ispCara2 (Figure 10A).

4.1.1.3 Metal dependence of IspC protein from *Arabidopsis thaliana*

The recombinant enzyme requires divalent cations for catalytic activity. Mn^{2+} and Mg^{2+} are most efficient as activators, but Fe^{2+}, Cu^{2+} and Ni^{2+} also afford considerable stimulation (Table 12). On the other hand, no activation was observed with Zn^{2+}. Without addition of any divalent cations, the activity is below the detection limit (< 0.001% as compared to the maximum activity observed in the presence of Mn^{2+} or Mg^{2+}, respectively).

Table 12: Activation of recombinant IspC protein from *Arabidopsis thaliana* by divalent metal ions

Metal ion	Relative activity (%)
Mn^{2+}	100
Mg^{2+}	43
Fe^{2+}	28
Cu^{2+}	16
Ni^{2+}	12
Co^{2+}	6
Ca^{2+}	2
Zn^{2+}	2
None	< 0.001

The reaction mixtures were prepared and analyzed with the photometric assay as described under Methods 3.3.6.1.1A. The final concentration of each metal ion was 5 mM.

Half-maximal activation was observed at a Mg^{2+} concentration of 1 mM and Mn^{2+} concentration of 60 µM, respectively (Figure 12). Considering the typical ion concentrations in the plastid, Mg^{2+} and Mn^{2+} are probably the important ionic species for activation of the IspC protein in vivo.

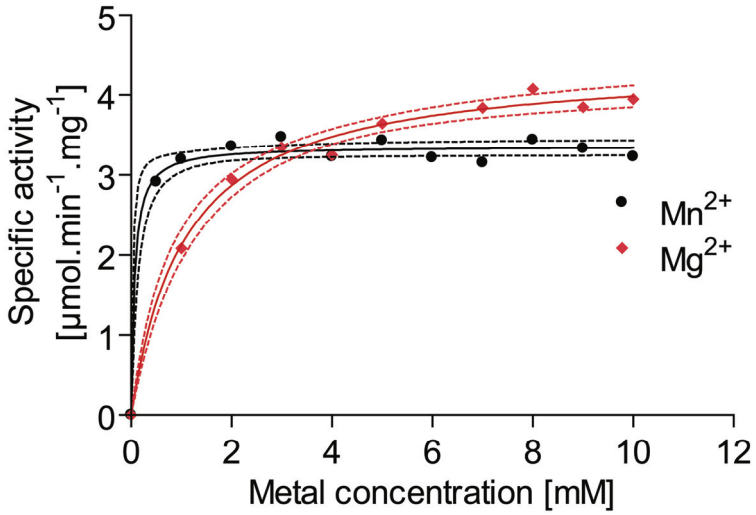


Figure 12: Influence of Mn^{2+} and Mg^{2+} on the activity of IspC protein from *Arabidopsis thaliana*
Assays were performed spectrophotometrically as described in 3.3.6.1.1A. Dotted lines, range of 95% confidentiality (calculated with the GraphPad PRISM 4.03 software).

4.1.1.4 PH dependence and the stability of IspC protein from *Arabidopsis thaliana*

In the presence of 10 mM Mg^{2+} (i.e. saturating conditions), the pH optimum was 8.0 (Figure 13). Although the highest activities were found at pH 8.0, it became obvious that the enzyme is more stable at a pH of 7.5. Therefore, the assays were performed at pH 7.5.

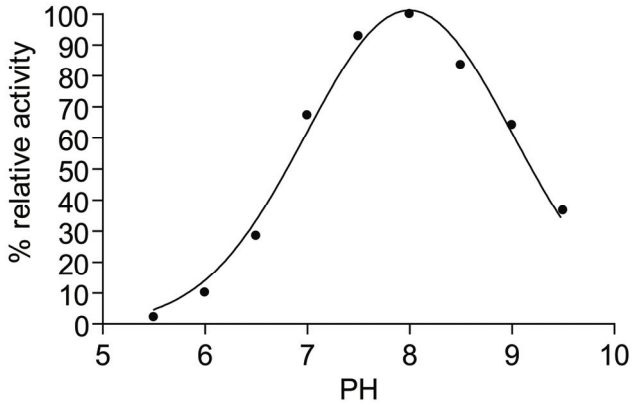


Figure 13: Relative activities versus pH values of IspC protein from *Arabidopsis thaliana*
Enzymatic activities were assayed (see Methods 3.3.6.1.1A) over a pH value range of 5.5 - 9.5 in buffer containing 50 mM Tris/HCl 50 mM glycine, 20 mM sodium phosphate and 50 mM sodium acetate.

After storage at 4 and 20 °C for 6 months, the enzyme retained 4 and 56% of its activity, respectively. The enzyme was stored in 100 mM Tris/HCl, pH 7.5 containing 50 mM potassium chloride, 10% (v/v) glycerol and 10 mM DTT at 80 °C. Under these conditions, the enzyme did not loose activity after storage for 6 months.

4.1.1.5 Temperature dependence of IspC protein from *Arabidopsis thaliana*

To test the temperature dependence of *A. thaliana* IspC, the assay was prepared as described under Methods 3.3.6.1.3.1, containing 100 mM Tris pH 7.5, 10 mM $MgCl_2$, 5 mM NADPH, 4 mM DXP, 10 µg protein and 10 % (v/v) D_2O to a total volume of 600 µl and incubated at different temperatures.

The enzyme appeared to have a narrow temperature optimum, with more than 90% relative activity between 35 and 42 °C (Figure 14). The maximum activity was observed at 37 °C, and the activation energy was calculated to be 48 kJ mol^{-1}. The enzyme retained 78 and 30% of enzymatic activity for 24 h at 20 and 37 °C, respectively.

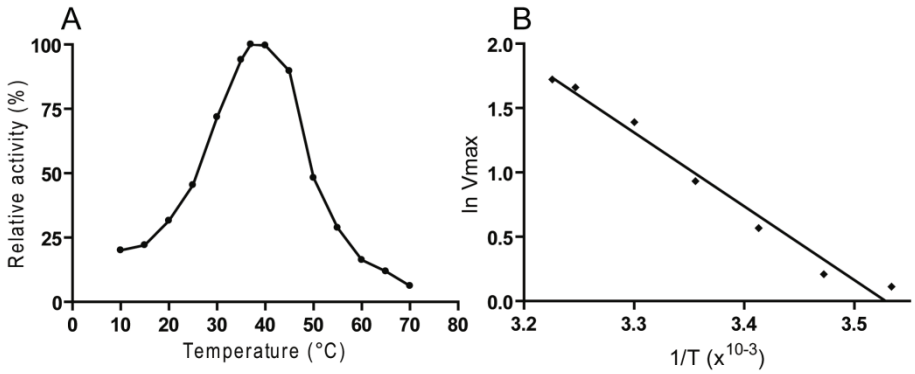


Figure 14: Temperature dependence of IspC protein from *Arabidopsis thaliana*
(A) Relative activities versus temperature. (B) Arrhenius plot used for the calculation of the activation energy.

4.1.1.6 Substrates dependence of IspC protein from *Arabidopsis thaliana*

Initial rate kinetic analysis using 10 mM Mg^{2+} (see Methods 3.3.6.1.1A) as activator showed typical Michaelis-Menten characteristics with regard to the NADPH concentration, with a K_M value of 30 µM (Figure 15A). On the other hand, initial rate kinetic analysis (Figure 15B) also showed substantial deviation from Michaelis-

Menten behavior with respect to the concentration of the substrate, 1-deoxy-D-xylulose 5-phosphate (**12**) (apparent K_M 132 µm).

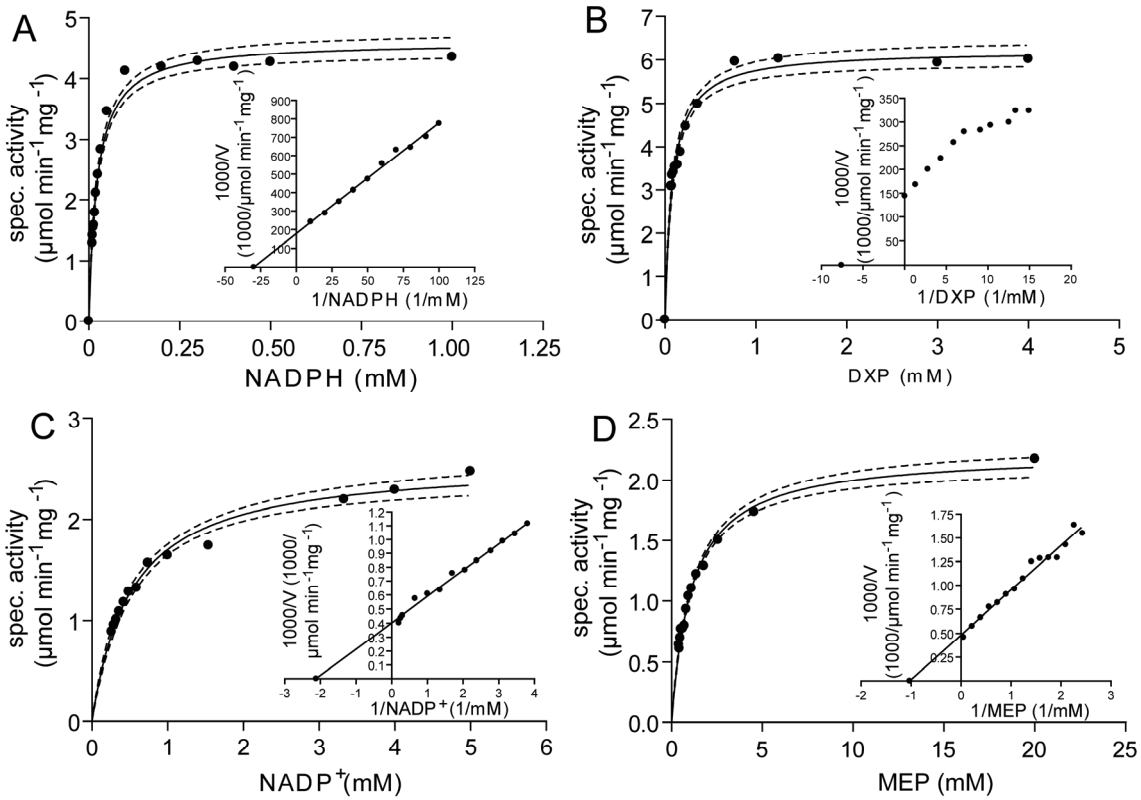


Figure 15: Michaelis-Menten kinetics of IspC protein from *Arabidopsis thaliana*
(A) Initial rates versus NADPH concentration. (B) Initial rates versus 1-deoxy- D-xylulose 5-phosphate concentration. (C) Initial rates versus $NADP^+$ concentration. (B) Initial rates versus 2*C*-methyl-D-erythritol 4-phosphate concentration. Insets, respective Lineweaver–Burk plots; dotted lines, ranges of 95% confidentiality.

NADH can substitute for NADPH, albeit at low rates of about 0.8 µmol min^{-1} mg^{-1} (14% as compared to the rates with NADPH; Table 13).

Table 13: Substrate specificity of *Arabidopsis thaliana* IspC

Substrate	Specific activity (µmol min^{-1} mg^{-1})	Relative activity (%)
NADPH	5.6	100
$NADP^+$	0.8	14.28

The reaction was also monitored by ^{13}C NMR spectroscopy using [3,4,5-$^{13}C_3$]-**12** as substrate (Figure 16). Owing to $^{13}C^{13}C$ coupling, the ^{13}C signals of the labeled carbon

atoms of the substrate [3,4,5-$^{13}C_3$]-**12** and the product [1,3,4-$^{13}C_3$]-**13** appear as complex multiplets that provide an unmistakable fingerprint (cf. Table 14). The apparent rate as determined from NMR observation was approximately 6 µmol min^{-1} mg^{-1}, in good agreement with the result of the photometric assay.

Table 14: ^{13}C NMR data of [3,4,5-$^{13}C_3$]1-deoxy-D-xylulose 5-phosphate and [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate.

Position	Chemical shift, ppm	Coupling constant, Hz	
	$^{13}C^a$	J_{CC}	J_{PC}
1-Deoxy-D-xylulose 5-phosphate (**12**)			
3	77.2	39.8 (d)	
4	70.9	41.1 (dd), 42.0 (dd)	6.9
5	64.5	42.9 (dd)	4.8
2*C*-Methyl-D-erythritol 4-phosphate (**13**)			
1	66.7		
3	74.1	42.0 (dd)	6.4
4	65	42.7 (dd)	4.2

[a]referenced to external trimethylsilylpropane sulfonate, the ^{13}C signals of DXP and MEP have been assigned in the thesis of Stefan Hecht, 2002.

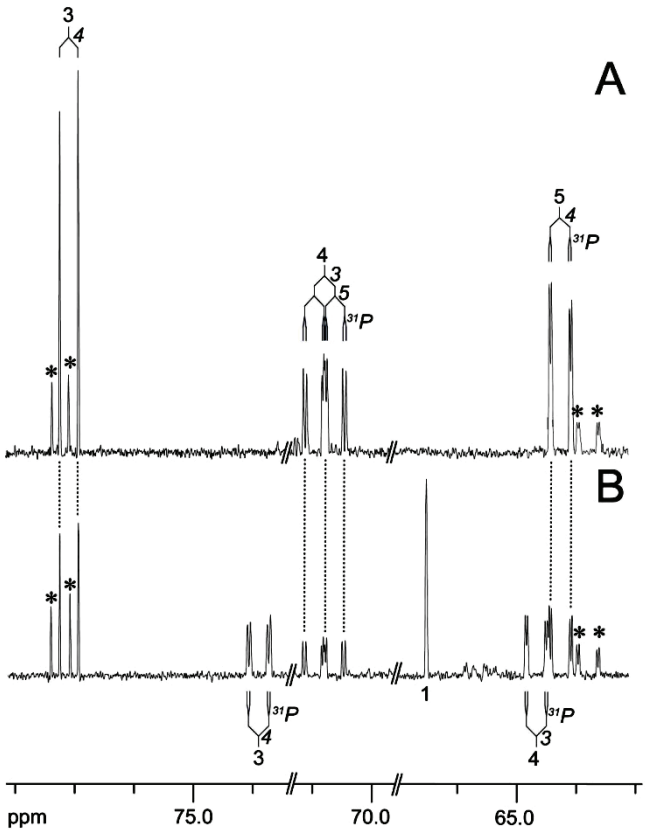


Figure 16: ^{13}C NMR signals detected in a reaction mixture converting [3,4,5-$^{13}C_3$]1-deoxy-D-xylulose 5-phosphate into [1,3,4-$^{13}C_3$]2*C*methyl- D-erythritol 4-phosphate Assay mixtures were prepared as described under Methods 3.3.6.1.3.1 and contained 100 mM Tris/HCl, pH 7.5, 10 mM $MgCl_2$, 6 mM NADPH, 6 mM [3,4,5-$_{13}C_3$]-**12**, 4mM DTT, and 15 µg of IspC protein in a volume of 500 µl. ^{13}C NMR signals detected in the assay mixture before (A) and after (B) incubation with the IspC protein from *Arabidopsis thaliana*; ^{13}C coupling patterns are indicated; asterisks designate inpurities.

Kinetic analysis of the reverse reaction was performed with 2*C*-methyl-D-erythritol 4-phosphate (**13)** as substrate and $NADP^+$ as cosubstrate (see Figure 15C and D, the assay is described under Methods 3.3.6.1.1B). The formation of NADPH was monitored photometrically at 340 nm. The rate constant k_{cat} had a value of 1.6 s^{-1} (V_{max} 2.1 µmol min^{-1} mg^{-1}). The K_M values for **12** and $NADP^+$ were 972 and 471 µM, respectively. Control assays without substrate and/or IspC protein were conducted to confirm that the NADPH formation is essentially linked to the formation of compound **12** catalyzed by the IspC protein.

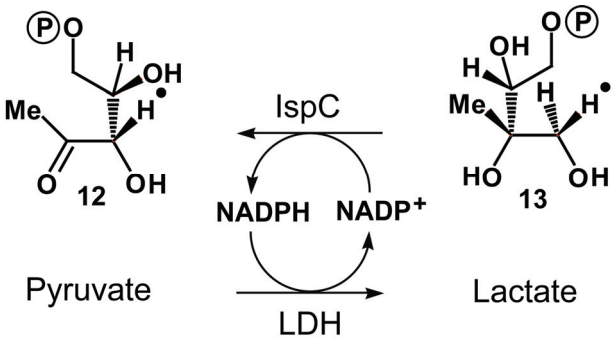


Figure 17: Coupling of the IspC reaction with a recycling system for NADPH
The reaction involves the conversion of 2*C*-methyl-D-erythritol 4-phosphate (**13**) to 1-deoxy-D-xylulose 5-phosphate (**12**) in the IspC reaction (cf. Figure 18) and could proceed to the complete formation of **12**. The assay was prepared as described under Methods 3.3.6.1.3.2 and analyzed by ^{13}C NMR spectroscopy.

The reaction was also monitored by ^{13}C NMR spectroscopy using [1,3,4-$^{13}C_3$]-**13** as substrate and a recycling system consisting of lactate dehydrogenase and pyruvate for regeneration of $NADP^+$ (Figure 17 and Figure 18). Closely in line with the result of the photometric assay, a specific catalytic rate of 2.3 µmol min^{-1} mg^{-1} was determined.

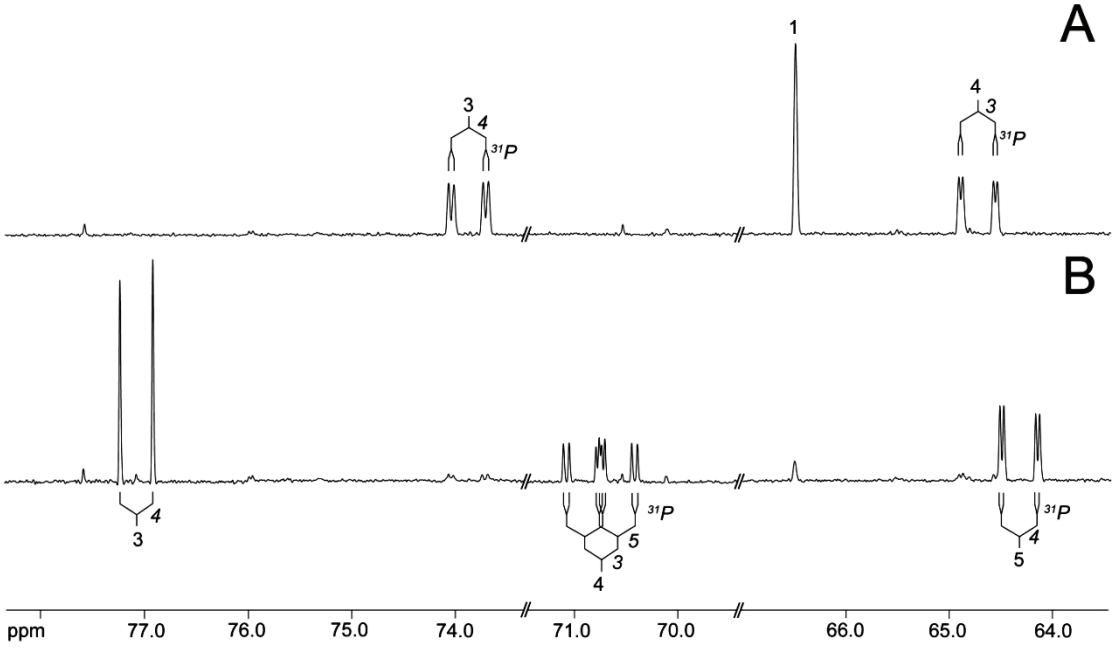


Figure 18: ^{13}C NMR signals of a conversion of [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate into [3,4,5-$^{13}C_3$]1-deoxy-D-xylulose 5-phosphate
The assay mixture was prepared as described under Methods 3.3.6.1.3.2. ^{13}C NMR signals detected in the assay mixture before (A) and after (B) incubation with the

IspC protein from *Arabidopsis thaliana*; ^{13}C coupling patterns are indicated (cf. Table 14).

4.1.1.7 Inhibition of IspC protein from *Arabidopsis thaliana* by fosmidomycin

The antibiotic fosmidomycin is a potent inhibitor of IspC enzymes from bacteria (Kuzuyama *et al*., 1998) and from *Plasmodium falciparum* (Jomaa *et al*., 1999). The compound, which was found to act as a slow binding inhibitor (Koppisch *et al*., 2002; Dhiman *et al*., 2005; Kuntz *et al*., 2005), inhibits the plant enzyme competitively with IC_{50} and K_i of 1.2 µM and 85 nm, respectively (Figure 19), in good agreement with the reported herbicidal activity of the antibiotic (Rodríguez-Concepción *et al*., 2001; Hans *et al*., 2004; Veau *et al*., 2000).

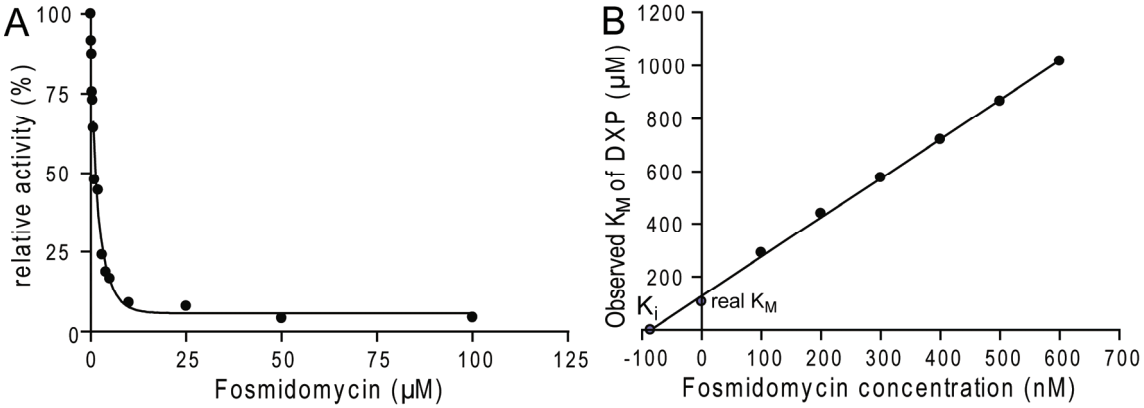


Figure 19: Inhibition of IspC protein from *Arabidopsis thaliana* by fosmidomycin
A, IC_{50} curve; B, Observed K_M of 1-deoxy-D-xylulose 5-phosphate (**12**) versus fosmidomycin concentration (the assay was prepared as described in paragraph 3.3.6.1.2); X-axis intercept denotes minus K_i; Y axis intercept denotes the real K_M.

4.1.1.8 Kinetic parameters of IspC protein from *Arabidopsis thaliana*

Kinetic parameters of *Arabidopsis thaliana* IspC are summarized in Table 15 below.

Table 15: Kinetic parameters of recombinant *Arabidopsis thaliana* IspC

Parameter	Value
k_{cat} (DXP → MEP)$_{NADPH}$	4.4 s^{-1}
k_{cat} (DXP → MEP)$_{NADH}$	0.63 s^{-1}
k_{cat} (MEP → DXP)$_{NADP^+}$	1.63 s^{-1}
K_M (NADPH)	30 µM
K_M (DXP)	132 µM
K_M ($NADP^+$)	471 µM
K_M (MEP)	972 µM
pH optimum	7.5
T_{max}	37 °C
K_i (fosmidomycin)	85 nM
Metal ion preference	Mn^{2+}, Mg^{2+}

4.1.1.9 Discussion

Several early reports were published on the heterologous expression of *ispC* genes from plants. However, as pointed out in a recent review by Proteau (Proteau, 2004), none of these reports included a description of the enzymes' properties to any significant depth. More specifically, Lichtenthaler and his coworkers reported an expression plasmid specifying amino acid residues 72–477 of the *A. thaliana* IspC protein (see Figure 10C, upper lane) (Schwender *et al.*, 1999); this coding sequence did not specify amino acid residues 55–71, which are crucial for expression in an enzymatically active form according to the experiments in this study (see 4.1.1.2). For enzymatic activity tests, the authors of that study used the radiolabeled radioactive enzyme substrate [2-^{14}C]-**12** that had been prepared enzymatically and was used without purification. After incubation, the reaction mixtures were treated with alkaline phosphatase and were subjected to thin-layer chromatography. The radioactivity of the dephosphorylated compounds was then detected by phosphoimaging. Using this method, the authors reported an enzymatic activity in crude cell extracts of the recombinant strain that was not specified in quantitative terms (Schwender *et al*., 1999). A follow- up publication of the same group (Müller *et al*., 2000) reported on an expression clone specifying the same polypeptide (see Figure 10C, upper lane) preceded by an N-terminal His tag (Figure 10C, lower lane). The authors reported on recombinant expression and purification of the cognate protein. An IC_{50} value of 0.28 µm was reported for fosmidomycin (Müller *et al*., 2000), and some kinetic parameters (Table 16) were given by Müller, 2003.

Campos and coworkers (Carretero-Paulet *et al*., 2002) reported residues 57–95 (Figure 10B) to be essential for the enzymatic activity of the IspC protein from *A. thaliana*, as diagnosed by in vivo complementation of an *ispC*-deficient mutant *E. coli* by plasmids carrying different *ispC* gene constructs. Although the authors did not perform in vitro enzyme studies, their findings are in agreement with our enzymatic data. As described under Results (4.1.1.2), among a set of expression clones specifying polypeptides of different lengths (see Figure 10A), only the clone specifying the protein extending from amino acid residues 57 to 477 afforded enzymatically active protein. Thus, it is concluded that the N-terminal segment extending from residues 57 to 71 is required in order to enable the expression of the *ispC* gene of *A. thaliana* in recombinant *E. coli* host strains. As that part of the gene

had not been reported in public databases by the time of the study by Lichtenthaler *et al*., it is not surprising that they failed to include that crucial sequence segment in their expression constructs.

Table 16: Kinetic parameters[a] of recombinant IspC proteins from different sources

Source	K_M (µM)		k_{cat} (s^{-1})	K_i (nM) **20**	Reference
	12	NADPH			
A. thaliana					
His-tagged	132	30	4.4	85	This study
His-tagged	108	7.2	-[b]	280[c]	Müller *et al*., 2000, Müller, 2003
E. coli					
Native	115-720	0.5	22-116		Proteau, 2004; Koppisch *et al*., 2002
His-tagged	97-99	18	8	38	Hoeffler *et al*., 2002; Kuzuyama *et al*., 1998; Koppisch *et al*., 2002; Hecht *et al*., 2001, Kuzuyama *et al*., 2000
M. tuberculosis					
Native	42	5.0	2.1	310[c]	Argyrou and Blanchard, 2004; Dhiman *et al*., 2005
S. coelicolor					
Native	190	190	19.2	-	Cane *et al*., 2001
***Synechocystis* PCC 6803**					
Native	134	5.0	5	-	Yin and Proteau, 2003; Woo *et al*., 2005
His-tagged	170	4.6	7	57	Yin and Proteau, 2003
Z. mobilis					
Native	300	5	14	600	Grolle *et al*., 2000

[a]All Values for kinetic parameters were those determined with Mg^{2+}, with the exception of the kinetic parameter of the Zymomonas mobilis enzyme, which were determined with Mn^{2+}. [b]Two different V_{max} values were reported, due to apparent instability of the protein (Müller, 2003).[c]IC_{50} value determined as described in Dhiman *et al*., 2005 and Müller, 2003, respectively; **12**, 1-deoxy-D-xylulose 5-phosphate; **20**, fosmidomycin.

The biochemical properties of the recombinant plant protein as reported in the present study are similar to those reported for the bacterial orthologs (Proteau,

2004). The K_M values for $NADP^+$ and compound **13** are 30-fold and 3-fold higher than the respective values for NADPH and compound **12**. The K_i value for the inhibitor fosmidomycin is similar to the value obtained with the *E. coli* ortholog (Table 16).

Sequence comparison showed a high degree of similarity among IspC protein members from the plant kingdom (more than 75% sequence identity, and at least 87% conservation) (Figure 9). As expected, homology to the bacterial IspC proteins is significantly lower. In total, 33 amino acid residues are conserved within all IspC amino acid sequences currently available in the database. A phylogenetic analysis of all currently available IspC sequences of plants and cyanobacteria is shown in Figure 20. Not surprisingly, all plant sequences (total number, 17) group together, and the cluster of cyanobacterial sequences (total number, 11) is relatively close to that of the plants. It should be noted that one green sulfur bacterium (*Chlorobium tepidum*) and one eubacterium (*Exiguobacterium* sp. 255-15) cluster with the cyanobacterial proteins, whereas the distances between most bacterial proteins and the plant/cyanobacterial cluster are much larger. It is also notable that the sequences of the two gymnosperms *Ginkgo biloba* and *Taxus cuspidata* form a lineage that is separated from the other plant species. The sequences from the grass species serving as the major staples for human nutrition and animal husbandry (i.e. rice, wheat, maize and barley) are all very similar.

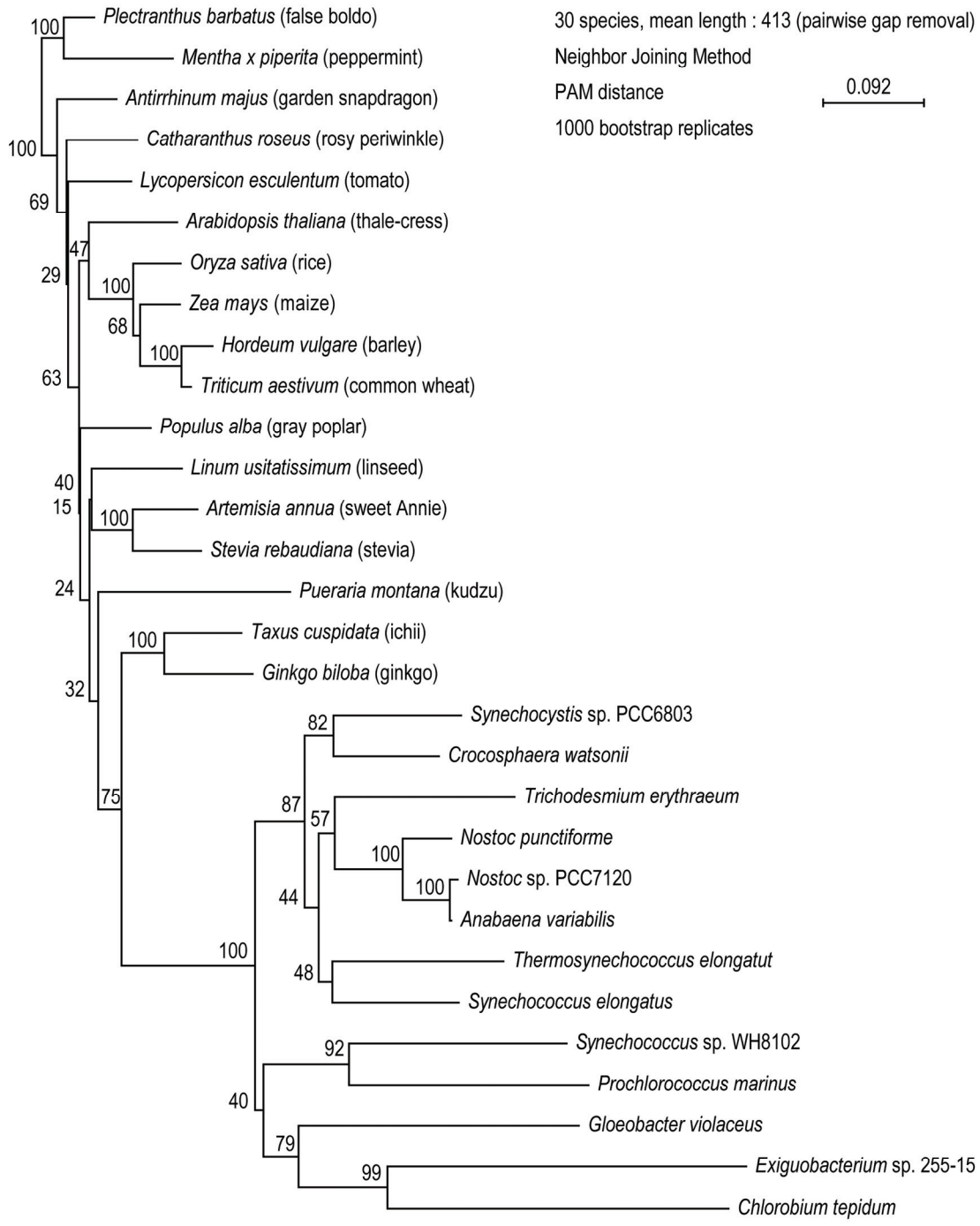


Figure 20: Consensus cladogram of IspC proteins
The simplified tree was constructed by neighbor-joining analysis from an alignment of IspC amino acid sequences from 17 plant species, 11 cyanobacterial species, one green sulfur bacterium (*Chlorobium tepidum*) and one eubacterium (*Exiguobacterium* sp. 255 – 15). Gaps were removed from the alignment, and the total number of positions taken into account was 413. The numbers at the nodes are the statistical confidence estimates computed by the bootstrap procedure. The bar represents 0.092 PAM distance.

Whereas the deoxyxylulose phosphate pathway has been proposed earlier as a putative target for the development of novel herbicides (Eisenreich *et al*., 2004; Zeidler *et al*., 1998; Rohdich *et al*., 2004; Rodríguez-concepción, 2004), in this study it has been shown for the first time that recombinant plant IspC protein can be produced in bulk amounts sufficient for in vitro high-throughput screening.

4.1.2 Characterization of IspC protein from *Plasmodium falciparum*

4.1.2.1 The cloning of IspC protein from *Plasmodium falciparum*

The *P. falciparum* open reading frame AE014826 (GenBank) specifies a 2*C*-methyl-D-erythritol 4-phosphate synthase (IspC protein) ortholog comprising 488 amino acid residues (Jomaa *et al*., 1999). Based on computer analysis, the N-terminal sequence residues 1 - 70 are believed to represent a bipartite leader sequence consisting of a signal peptide (residues 1 - 21) followed by a transit peptide (residues 22 - 70) typical for proteins that are targeted to the apicoplast (Figure 21) (Waller *et al*., 2000; Foth *et al*., 2003). In contrast to plants, protein import into the malarial plastid equivalent is a two-step process, whereby the protein is directed into the endomembrane system by the signal peptide and, after removal of the signal peptide, is diverted away from the default secretory pathway into the apicoplast (Waller *et al*., 2000; Dooren *et al*., 2001). Since sequence similarity to bacterial orthologs begins at amino acid position 74 (data not shown), we decided to express amino acid residues 74–488. For that purpose, a synthetic gene was designed that was optimized for expression in *E. coli* (Figure 21) (Lauw *et al*., 2008; publication in preparation).

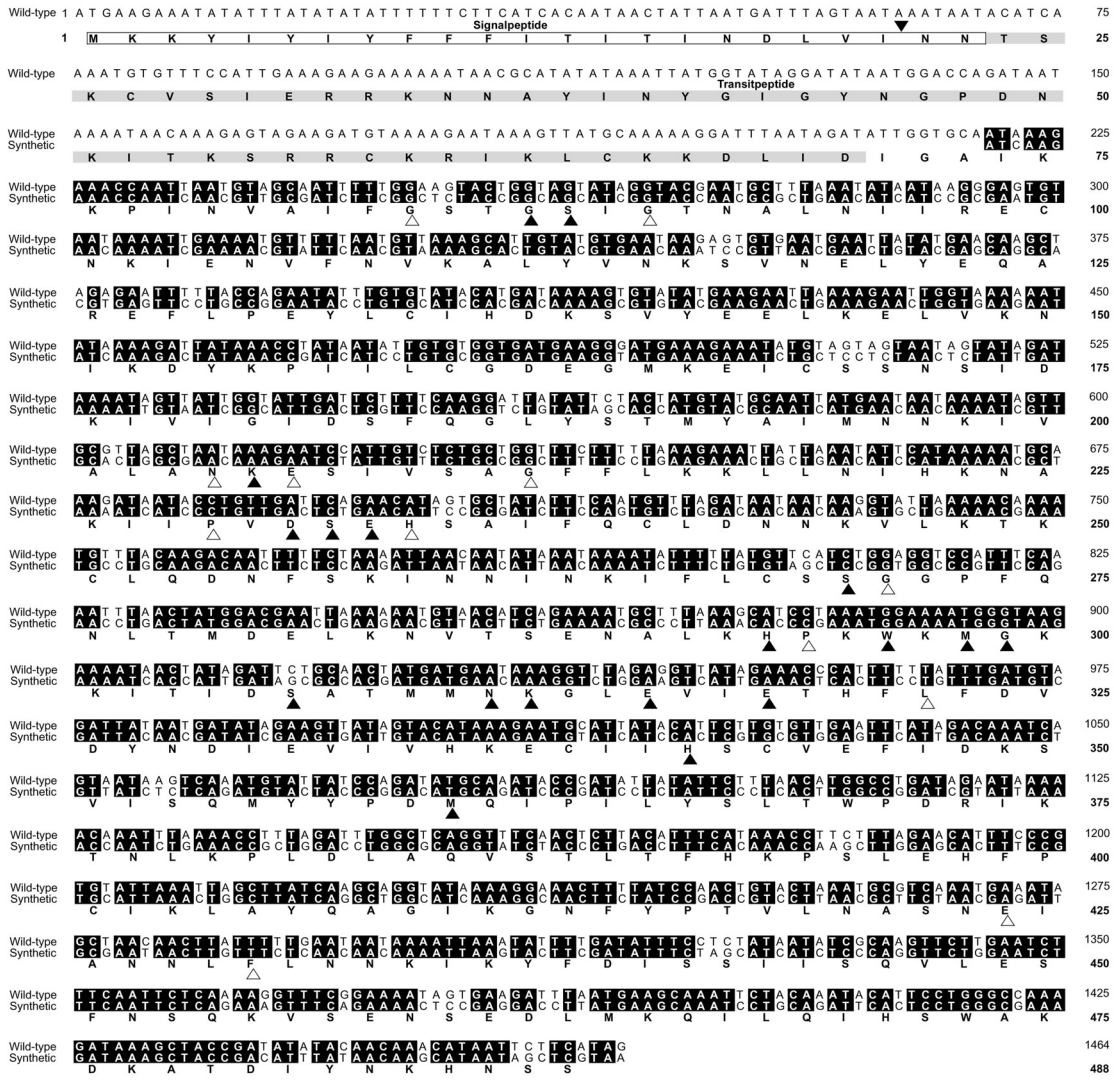


Figure 21: Nucleic and amino acid sequences of IspC protein from *Plasmodium falciparum*.
The synthetic construct starts at nucleic acid position 220 of the IspC open reading frame from the wild-type sequence; identical bases are shown in inverse contrast. Signal (boxed), transit peptide (grey background) and cleavage site (▼) as determined with SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP), PATS (http://gecco.org.chemie.uni-frankfurt.de/pats-index.php) and PlasmoAP (http://v4-4.plasmodb.org/restricted/PlasmoAPcgi.shtml) are indicated. ▲, absolutely conserved residues; filled up triangles, residues involved in enzyme catalysis.

The synthetic gene was obtained by custom synthesis and cloned into the vector pQE30 resulting in an expression construct preceded by an N-terminal MRGSHHHHHHGS sequence.

4.1.2.2 Purification of IspC protein from *Plasmodium falciparum*

The extract from recombinant cells showed no visible expression as judged by SDS polyacrylamide gel electrophoresis (Figure 22, lane C) but could be shown to catalyze the conversion of 1-deoxy-D-xylulose 5-phosphate (**12**) into 2*C*-methyl-D-erythritol 4-phosphate (**13**) with a rate of 50 nmol min^{-1} mg^{-1} beyond the background of wild type activity of 10 nmol min^{-1} mg^{-1}.

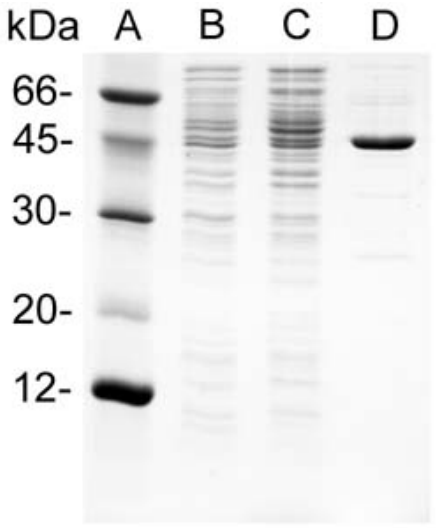


Figure 22: Purification of recombinant IspC protein from *Plasmodium falciparum*
SDS-PAGE: lane A, molecular weight markers; lane B, cell extract of *E. coli* wild-type; lane C, cell extract of a recombinant *E. coli* strain expressing the *ispC* gene from *P. falciparum* (XL1-pQEispCplas2); lane D, recombinant IspC protein after affinity chromatography on Ni-chelating Sepharose FF and desalting chromatography on a HIPREP 26/10.

The recombinant protein was isolated by nickel chelation chromatography followed by desalting (see Methods 3.3.2.2.1, buffer containing 100 mM Tris/HCl, pH 7.5 and 10 % (v/v)). The resulting protein appeared to approximately 90 % pure as judged by SDS polyacrylamide gel electrophoresis with an apparent mass of 49 kDa as judged by SDS polyacrylamide gel electrophoresis (Figure 22, lane D).

A photometric assay based on NADPH consumption was performed as described under Methods 3.3.6.1.1A, showing the formation of **13** at a rate of 2 µmol min^{-1} mg^{-1} in the presence of 10 mM Mg^{2+}. The apparent purification factor (Table 17) indicates that the protein has been expressed at a level about 2 % from the total cell protein.

Table 17: Purification of recombinant IspC protein from *Plasmodium falciparum*

Procedure	Total protein (mg)	Specific activity (µmol $min^{-1}mg^{-1}$)	Total activity (µmol min^{-1})	Yield (%)	Purification factor
Cell extract	3391	0.05	159	100	1
Ni^{2+}-Sepharose, HIPREP 26/10	20	2	40	25	43

N-terminal Edman sequencing confirmed the expected MRGSHHHHHHGSIKKPINVAIFGST motif (Figure 21). Mass spectrometry (MALDI TOF) afforded a mass of 48,580 Da in good agreement with the expected value of 48,587 Da.

4.1.2.3 Metal dependence of IspC protein from *Plasmodium falciparum*

Assays to determine the metal dependence were prepared as described under Methods 3.3.6.1.1A with the addition of 1 mM EDTA. The metal ion was added to initiate the reaction.

Table 18: Activation of recombinant IspC protein from *Plasmodium falciparum* by divalent metal ions

Metal ion	Relative activity (%)
Mn^{2+}	100
Mg^{2+}	41
Co^{2+}	22
Fe^{2+}	18
Ca^{2+}	5
Cu^{2+}	4
Zn^{2+}	1
Ni^{2+}	< 0.001
None	< 0.001

The addition of 1 mM EDTA reduced the activity to an undetectable level. Enzyme that had been inactivated by treatment with EDTA could be reactivated by a variety of divalent cations (Table 18. Notably, the highest activity was obtained with Mn^{2+}, but

substantial reactivation was also observed with Mg^{2+}, Co^{2+} and Fe^{2+}, whereas Ni^{2+} was without effect. Generally, the metal preference profile of the *Plasmodium* enzyme resembles that of the ortholog from *Arabidopsis thaliana* that has been reported earlier in this study (see 4.1.1.3).

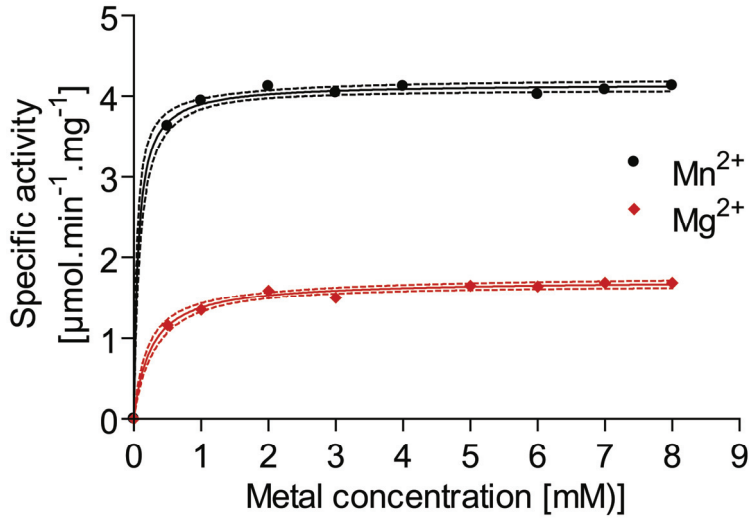


Figure 23: Mn^{2+} and Mg^{2+} dependence of IspC protein from *Plasmodium falciparum*

The optimum concentration of both Mn^{2+} and Mg^{2+} was 2 mM as depicted in Figure 23. Similar to *Arabidopsis thaliana* IspC, protein, assays containing Mn^{2+} and Mg^{2+} afforded Michaelis-Menten kinetics with half activation at concentrations of 65 and 251 µM, respectively. Mn^{2+} induced higher enzyme activity than Mg^{2+} (4.1 as compared to 1.7 µmol min^{-1} mg^{-1}). Based on this finding, Mn^{2+} was used as metal ion of favoured choice in the assays.

4.1.2.4 PH dependence and stability of IspC protein from *Plasmodium falciparum*

The pH optimum of the enzyme is in the range of 7.5 – 8.0 (Figure 24). Upon storage at 4 °C in Tris/HCl, pH 7.5, the activity loss was in the range of 20 % per day. At room temperature, the decline of activity per day was in the range of 10 % per day. Under the strorage condition in buffer containing 50 mM Tris/HCl, pH 7.5 and 10 % (v/v) glycerol, the enzyme did not loose activity after 1 year.

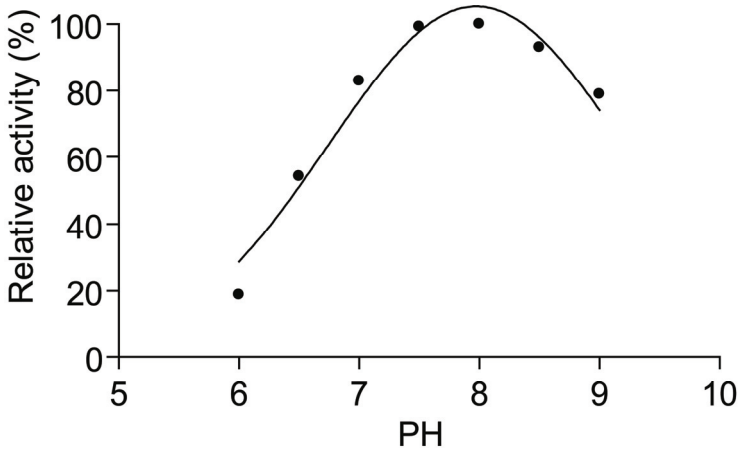


Figure 24: Relative activity versus pH values of IspC protein from *Plasmodium falciparum*
Enzymatic activities were assayed (see paragraph 3.3.6.1.1A) over a pH range of 6 - 9 in buffer containing 50 mM Tris/HCl, 50 mM glycine, 20 mM sodium phosphate and 50 mM sodium acetate.

4.1.2.5 Temperature dependence of IspC protein from *Plasmodium falciparum*

To test the temperature dependence of *P. Falciparum* IspC, ^{13}C NMR assay mixtures were prepared as described under Methods 3.3.6.1.3.1. These mixtures contain 100 mM Tris/HCl, pH 7.5, 2 mM $MnCl_2$, 6 mM NADPH, 4 mM [3,4,5-$^{13}C_3$]1-deoxy-D-xylulose 5-phosphate (**12**).

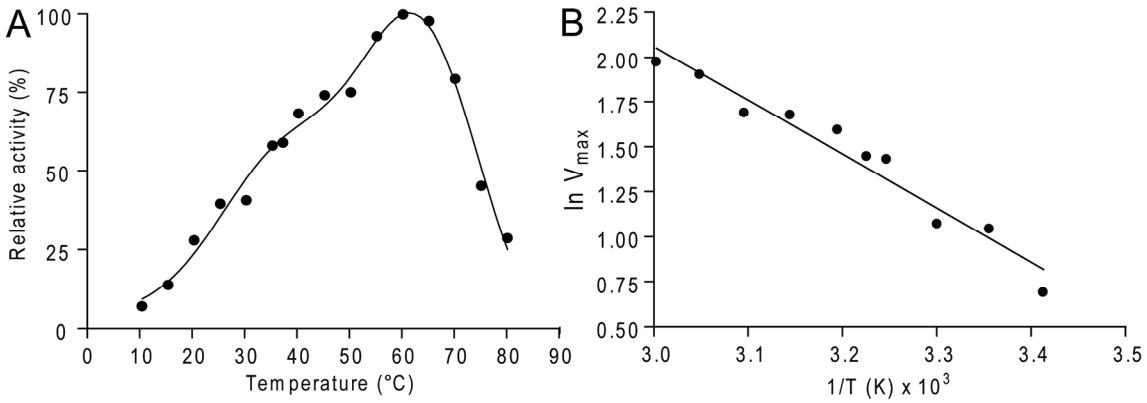


Figure 25: Temperature dependence of IspC protein from *Plasmodium falciparum*
(A) Relative activity versus temperature; (B) Arrhenius plot used for the calculation of the activation energy.

In the enzyme assays conducted over a period of 30 min at different temperatures, the apparent activity reached a plateau value approximately at 60 °C. The Arrhenius plot indicates activation energy of 25 kJ mol^{-1} for the forward reaction (conversion of **12** into **13**) (Figure 25).

4.1.2.6 Substrate dependence of IspC protein from *Plasmodium falciparum*

The photometric assay of IspC activity in the forward direction was performed as described under Methods 3.3.6.1.1A using Mn^{2+} as the metal ion in a concentration of 2 mM. An Initial rate kinetic analysis showed typical Michaelis-Menten characteristics with regard to NADPH and 1-deoxy-D-xylulose 5-phosphate (DXP/**12**). The K_M values were determined to 124 and 281 µM, respectively (Figure 26A and B), while the k_{cat} value was 3.5 s^{-1} (V_{max} = 4.4 µmol min^{-1} mg^{-1}).

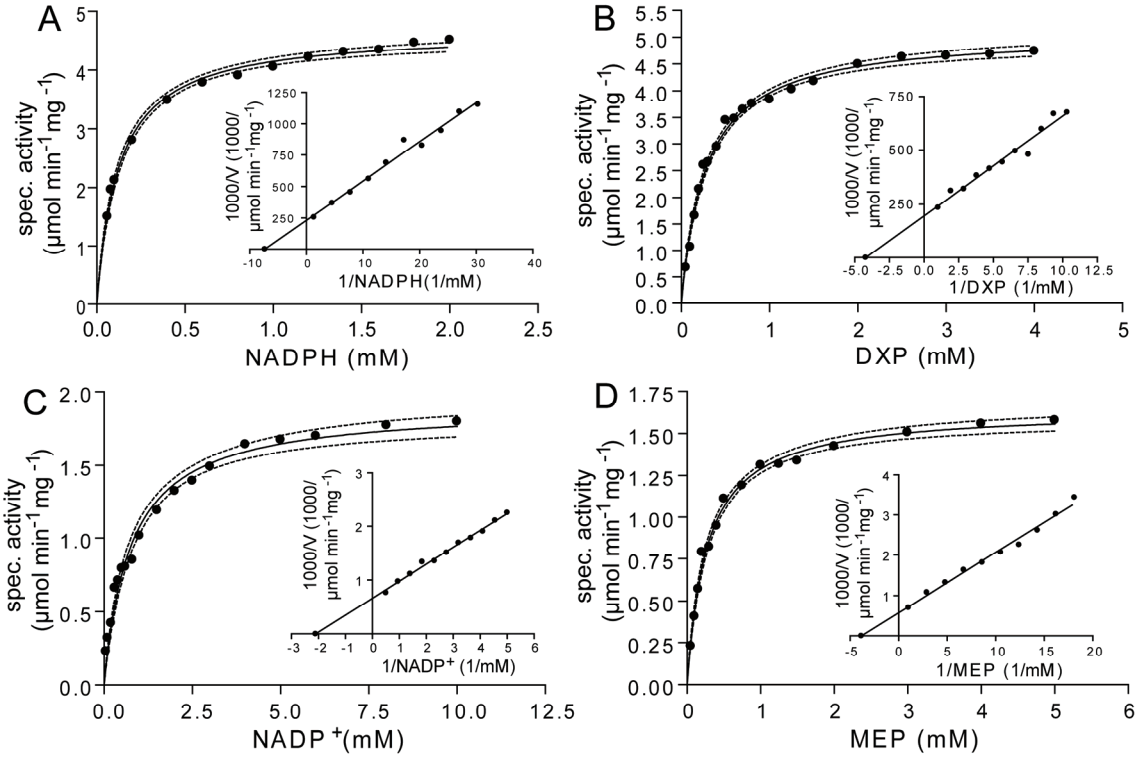


Figure 26: Michaelis-Menten kinetics of IspC protein from *Plasmodium falciparum*
(A) Initial rates versus NADPH concentration. (B) Initial rates versus 1-deoxy-D-xylulose 5-phosphate (**12**) concentration. (C) Initial rates versus $NADP^+$ concentration. (D) Initial rates versus 2*C*-methyl-D-erythritol 4-phosphate (**13**) concentration. Insets, respective Lineweaver-Burk plots; dotted lines ranges of 95 % confidentiality.

The reverse reaction was assessed as described under Methods 3.3.6.1.1B using **13** and $NADP^+$ as substrates in the presence of 2 mM Mn^{2+}. The rate constant k_{cat} had a value of 1.3 s^{-1} (V_{max} = 1.6 µmol min^{-1} mg^{-1}).

NADH can substitute for NADPH, but the maximum reaction rate is lower, about 1.1 µmol min^{-1} mg^{-1} (28 % as compared to the rates with NADPH, Table 19).

Table 19 Substrate specificity of *P. falciparum* IspC

Substrate	Specific activity (µmol min^{-1} mg^{-1})	Relative activity (%)
NADPH	3.9	100
NADH	1.1	28

4.1.2.7 Inhibition of *Plasmodium falciparum* IspC by fosmidomycin

The assay of the *P. falciparum* enzyme was implemented on an automated plate reader using the 96-well format (see Methods 3.3.6.1.2, Table 9). The automated assay was monitored photometrically at 340 nm. Using equation 1, a Z factor of ≥ 0.8 was determined indicating a high level of robustness.

$$Z = 1 - \frac{(3\sigma_s + 3\sigma_c)}{|\mu s - \mu c|}$$

Equation 1: Z factor
σ, the standard deviation; μ, the mean value; s, the signal, and c, the control.

Using this automated assay, the inhibition of the enzyme by fosmidomycin was studied. Fosmidomycin has been shown earlier to inhibit the *Plasmodium* enzyme (Jomaa *et al*., 1999). The forward reaction showed a mixed type inhibition with K_i = 5.4 ± 1.8 and K_{is} = 30 ± 17 nM. The reverse reaction showed competitive inhibition with K_i = 6.0 ± 0.6 nM (Table 20).

Table 20: K_i values and mode of inhibition of the *Plasmodium falciparum* IspC protein by fosmidomycin

IspC reaction	Mode of Inhibition	K_i (nM)
Forward	Mixed	K_i = 5.4 ± 1.8 K_{is} = 30 ± 17
Backward	Competitive	K_i = 6 ± 0.6

4.1.2.8 Kinetic parameters of Ispc protein from*Plasmodium falciparum*

Kinetic parameters of *Plasmodium falciparum* IspC are summarized in Table 21 below.

Table 21: Kinetic parameters of recombinant IspC protein from *Plasmodium falciparum*

Parameter	Value
k_{cat} (DXP → MEP)$_{NADPH}$	3.55 s^{-1}
k_{cat} (DXP → MEP)$_{NADH}$	0.99 s^{-1}
k_{cat} (MEP → DXP)$_{NADP^+}$	1.29 s^{-1}
K_M (NADPH)	136 µM
K_M (DXP)	281 µM
K_M ($NADP^+$)	475 µM
K_M (MEP)	260 µM
pH optimum	7.5
T_{max}	60 °C
K_i (fosmidomycin)	6 nM
Metal ion preference	Mn^{2+}, Mg^{2+}

4.1.2.9 Discussion

The recombinant expression of 2*C*-methyl-D-erythritol 4-phosphate synthase has been reported earlier by Jomaa and coworkers (*Jomaa et al*., 1999). Using various strains of *P. falciparum*, the authors performed an in vitro test to determine the inhibition efficacy of fosmidomycin and its derivative FR-900098 in comparison to currently used antimalarial drugs such as chloroquine and pyrimethamine. These assays showed that fosmidomycin and FR-900098 inhibit the growth of the respective *P. falciparum* strains with comparable IC_{50} values as compared to that of chloroquine and pyrimethamine (Table 22, Jomaa *et al*., 1999). These in vitro tests were supported by in vivo drug tests where mice infected with a rodent malaria causative were treated with 30 mg kg^{-1} of fosmidomycin and were cured over a period of 8 days. Furthermore, fosmidomycin showed a low toxicity in the animal model (Jomaa *et al*., 1999) which might be due to the absence of the deoxyxylulose phosphate pathway in the animal host. Earlier studies had already suggested that fosmidomycin is well tolerated by volunteers and patients who suffered from bacterial infections (Kuemmerle *et al*., 1987). These findings indicate that IspC protein is a valid target and that inhibitors of this enzyme could serve as potential leads for the discovery of antimalarial agents.

Table 22: In vitro inhibition of various *P. falciparum* strains by fosmidomycin, FR-900098, chloroquine and pyrimethamine

Plasmodium falciparum strains	IC_{50} (nM)			
	Fosmidomycin	FR-900098	Chloroquine	Pyrimethamine
HB3	350 ± 170	170 ± 100	20 ± 5	60 ± 42
A2	370 ± 45	170 ± 45	37 ± 7	4 ± 2
Dd2	290 ± 130	90 ± 20	200 ± 30	2500 ± 1000

This table was published by Jomaa *et al*., 1999. IC_{50}, the concentration that causes half of maximum inhibition of growth rates. Mean values and standard deviations from 3-4 independent experiments are shown.

The efforts to discover antimalarial agents are carried out intensively and continuously. One approach for this aim is the rational design of inhibitors. With this

respect, the structure of IspC protein from *P. falciparum* was modeled on basis of its homology to the respective proteins of *E. coli* (Singh *et al*., 2005) and *Zymomonas mobilis* from which three-dimensional structures have been reported (Ricagno *et al*., 2003) in order to molecularly design new antimalarial drugs. An alternative is the high-throughput screening of 2*C*-methyl-D-erythritol 4-phosphate synthase (IspC protein) using a rapid photometric assay method that has been previously developed in our laboratory (Illarionova *et al.*, 2006). In order to accomplish such a screening relatively large amounts of the recombinant target enzyme as well as high Z factors indicating a robust assay system are neccessary. Both these prerequisites have been compiled within the studies of the present PhD thesis (see 4.1.2.2 and 4.1.2.7).

4.1.3 Mechanistic study of IspC protein

As already mentioned (4.1), 2*C*-methyl-D-erythritol 4-phosphate synthase (IspC protein) catalyzes the first committed step in the non-mevalonate isoprenoid biosynthetic pathway. The reaction involves the isomerization of 1-deoxy-D-xylulose 5-phosphate (DXP/**12**) affording the branched chain aldose-derivative (**19**) which is subsequently reduced under formation of 2*C*-methyl-D-erythritol 4-phosphate (MEP/**13**). The isomerization step has been proposed to proceed as an intramolecular rearrangement or a retroaldol/aldol sequence (Kuzuyama *et al*., 2008; Hoeffler *et al*., 2002) (Figure 27).

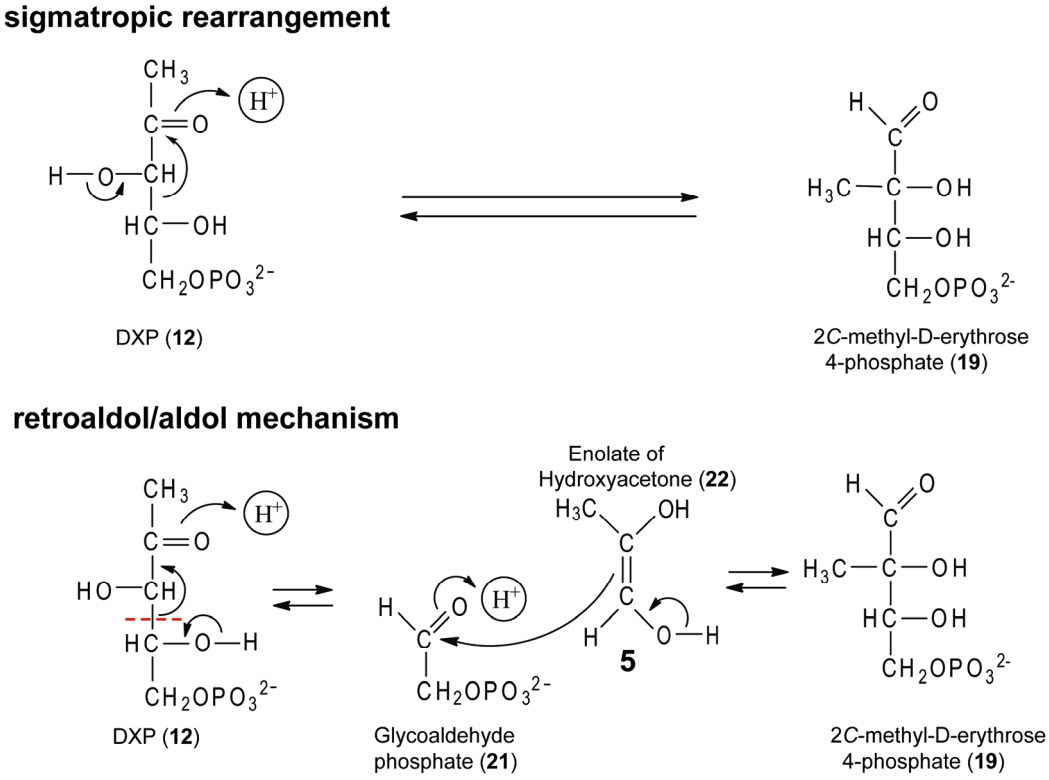


Figure 27: Hypothetical mechanisms of the IspC reaction.

In the sigmatropic rearrangement, deprotonation occurs at C-3 of DXP followed by a displacement of a C-2 unit affording 2*C*-methyl-D-erythrose 4-phosphate. As opposed to sigmatropic rearrangement, the hypothetical retroaldol mechanism is initiated by deprotonation at C-4 of DXP followed by cleavage between C-3 and C-4 affording glycolaldehyde phosphate and the enolate of hydroxyacetone. A consecutive aldol addition could then afford 2*C*-methyl-D-erythrose 4-phosphate (Figure 27).

We report the preparation of ^{13}C-labeled substrate isotopologues which were designed to optimize the detection of an exchange of putative cleavage products that might occur in the hypothetical retroaldol/aldol reaction sequence. 2*C*-Methyl-D-erythritol 4-phosphate synthases of *Escherichia coli*, *Mycobacterium tuberculosis* and *Arabidopsis thaliana* were selected for parallel enzyme studies. The dendogram (Figure 28) shows a relatively large genetic distance between these three species. The degree of sequence identity of the enzyme from *E. coli* with those from *M. tuberculosis* and *A. thaliana* is 40 and 43 %, respectively. Notably, the three enzymes are located in three different sequence clusters comprising pro- and eukaryotic orthologues, respectively (shown in Figure 28).

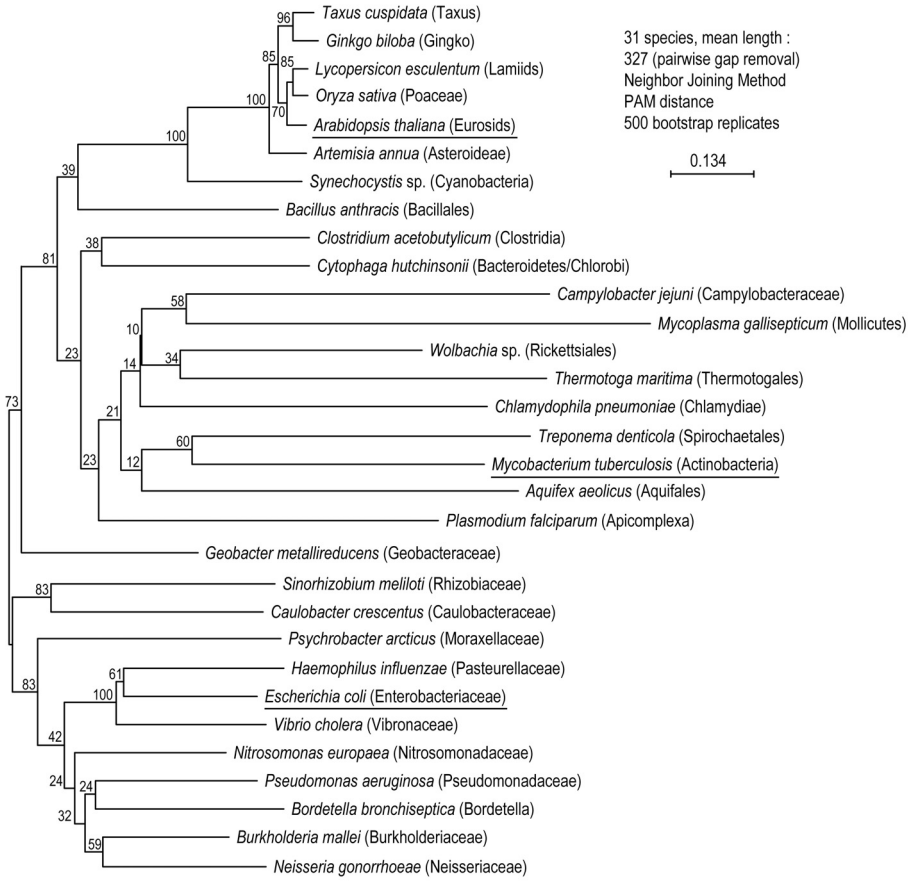


Figure 28: Phylogenetic tree of IspC proteins from various organisms
The consensus cladogram was constructed by Neighbor-joining analysis from an alignment of IspC amino acid sequences from 6 plant species, one cyanobacterium (*Synechocystis* sp.), one protist (*P. falciparum*) and 24 eubacteria representing different families. Gaps were removed from the alignment, and the total number of positions taken into account was 327. The numbers at the nodes are the statistical confidence estimates computed by the bootstrap procedure. The bar represents 0.134 PAM distance.

4.1.3.1 Synthesis of [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate

Two isotopologues of glucose were used in the synthesis as described under Methods 3.3.3.1. Specifically, [2,5-$^{13}C_2$]D-glucose and [3,4-$^{13}C_2$]D-glucose and were used to produce [1-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate; and [3-$^{13}C_1$] 2*C*-methyl-D-erythritol 4-phosphate. Each mol of double-labeled D-glucose produced 2 mol of single-labeled substrate 2*C*-methyl-D-erythritol 4-phosphate in the synthesis.

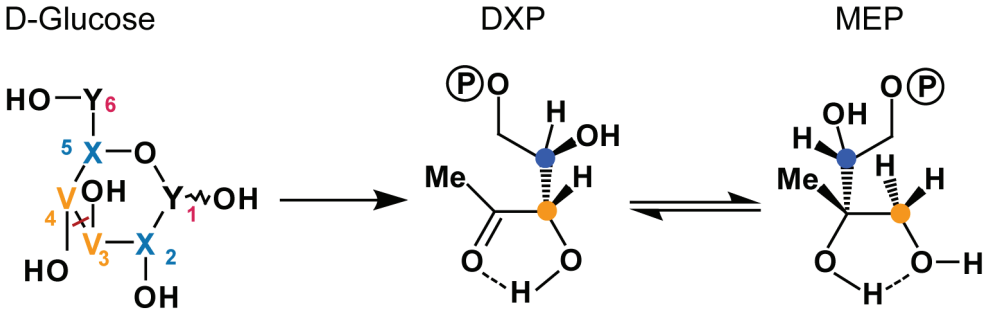


Figure 29: Labeling pattern of MEP synthesized from D-glucose
The [3,4-$^{13}C_2$]D-glucose (yellow) and [2,5-$^{13}C_2$]D-glucose (blue) and their respective product during the synthesis of [1-$^{13}C_1$]- and [3-$^{13}C_1$]2C-methyl-D-erythritol 4-phosphate. The labeling position was signed with yellow dots for [3-$^{13}C_1$]-DXP and [1-$^{13}C_1$]-MEP and blue dots for [4-$^{13}C_1$]-DXP and [3-$^{13}C_1$]-MEP.

The products were analyzed by ^{13}C NMR measurement affording the expected signals of [1-$^{13}C_1$]- MEP as a singlet at 66.5 ppm and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate as a doublet due to the phosphorus coupling at 73.7 ppm as shown in Figure 30.

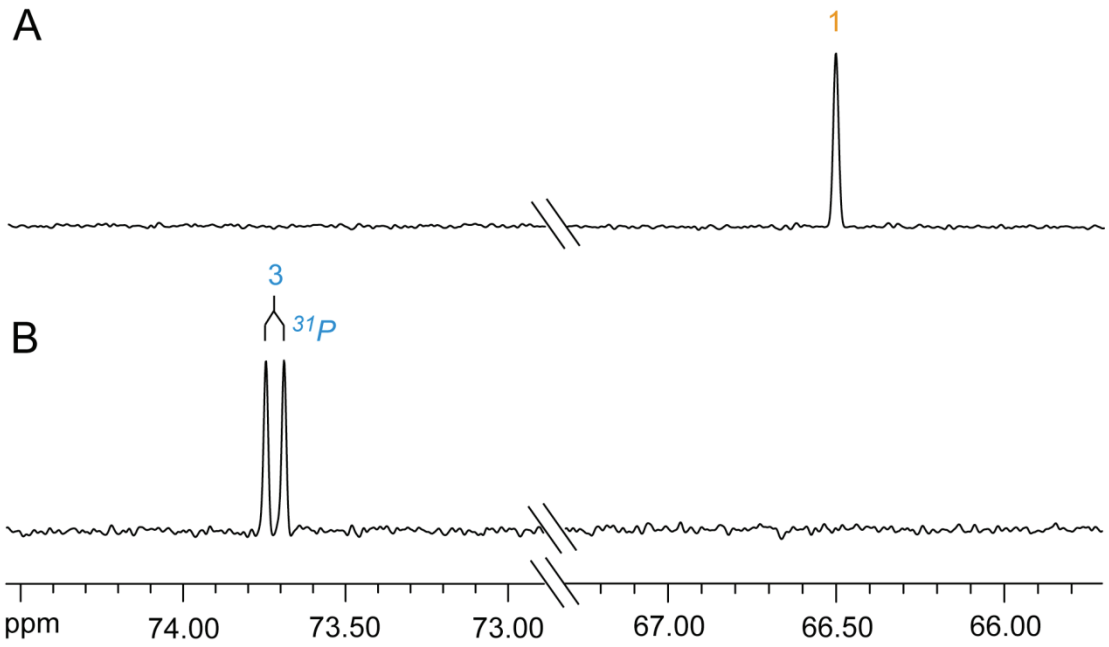


Figure 30: ^{13}C NMR signals of [1-^{13}C]-MEP (A) and [3-^{13}C]-MEP (B)

The recovery of product 2*C*-methyl-D-erythritol 4-phosphate after purification through cellulose column was about 40 %.

4.1.3.2 Synthesis of [1,2-$^{13}C_2$]glycolaldehyde phosphate

The synthesis of [U$^{13}C_5$]ribulose 5-phosphate is described under Methods 3.3.4. The formation of [U$^{13}C_5$]ribulose 5-phosphate was monitored by ^{13}C NMR spectroscopy and the signals fit excellently with the ^{13}C signals assigned by Volk (Volk, TU dissertation, 1989). The yield after lyophilization was 90 % (mol/mol) when compared to the initial [U$^{13}C_6$]D-glucose.

The synthesis of [1,2-$^{13}C_2$]glycoaldehyde phosphate was carried out according to the Methods 3.3.5 from the crude [U$^{13}C_5$]ribulose 5-phosphate. The formation of [1,2-$^{13}C_2$]glycoaldehyde phosphate was monitored by ^{13}C NMR spectroscopy based on the published ^{13}C signals by Hoeffler (Hoeffler *et al*., 2002). Specifically, the signals of [1,2-$^{13}C_2$]glycoaldehyde phosphate at 125 MHz and 50 mM Tris/HCl, pH 8.0 were C-1 (J^{13}C, 88.6 ppm; J^{31}P, 8.9 ppm, dd) and C-2 (J^{13}C, 67.5 ppm; J^{31}P, 5.1 ppm, dd) as seen in Figure 31.

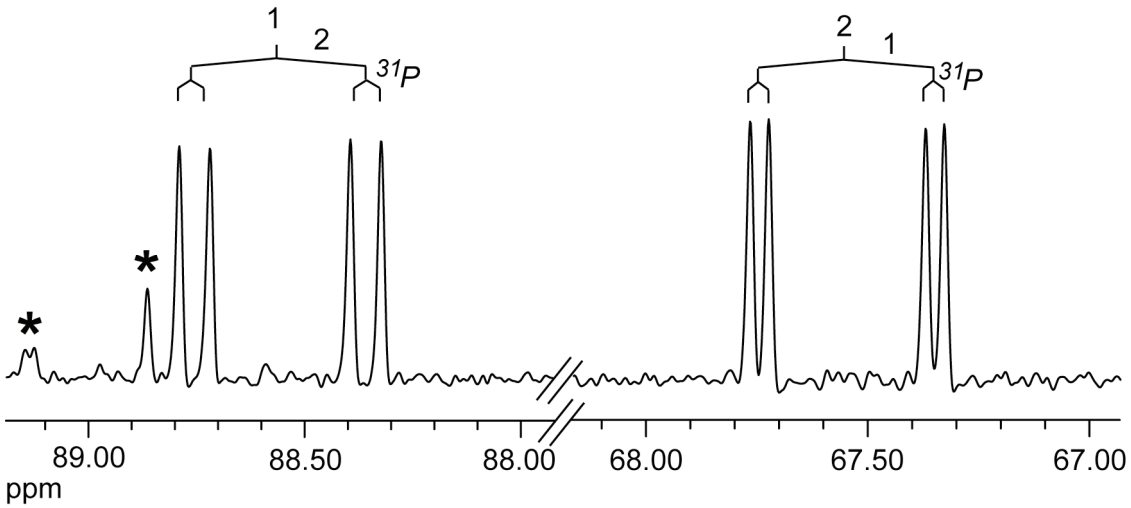


Figure 31: ^{13}NMR spectra of [1,2-$^{13}C_2$]glycolaldehyde phosphate.
Asterisks denote impurities.

After purification by anion exchange, the yield of [1,2-$^{13}C_2$] glycoaldehyde phosphate was 40 % (mol/mol) when compared to the initial [U$^{13}C_5$]ribulose 5-phosphate. However, [1,2-$^{13}C_2$] glycoaldehyde was unstable and rapidly degraded, therefore it should be always freshly prepared prior the IspC assay and used within 3 - 4 h.

4.1.3.3 Purification of IspC proteins from *E. coli*, *M. tuberculosis* and *A. thaliana*

The purification was performed according to the Methods 3.3.2.2.1 with the modifications as seen in Table 23, affording the total amount of recombinant proteins as shown in Table 23.

Table 23 Purification of IspC proteins from *E. coli*, *M. tuberculosis* and *A. thaliana*

IspC from various organisms	Purification method	Cell (g)	Pure protein (mg)
Escherichia coli	Nickel-chelated followed by dialysis	2	240
Mycobacterium tuberculosis	Nickel-chelated followed by HiPrep desalting column	40	40
Arabidopsis thaliana	Nickel-chelated followed by HiPrep desalting column	40	20

The SDS-PAGE (Figure 32) shows the bands of the purified recombinant IspC proteins from *E. coli*, *M. tuberculosis* and *A. thaliana*.

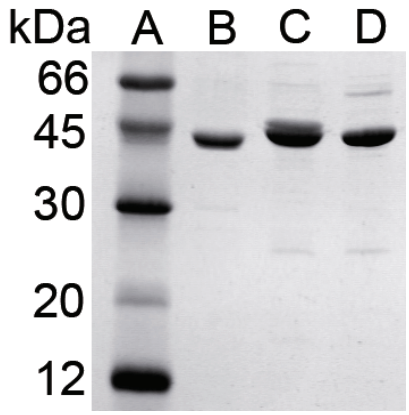


Figure 32: SDS-PAGE of purified IspC proteins from different organisms.
Protein marker (A) and IspC proteins from *E. coli* (B), *M. tuberculosis* (C) and *A. thaliana (D)*.

4.1.3.4 Mechanistic study of IspC reaction using [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate in the enzyme assay

In order to measure the frequency of any potential intermediate exchange, two substrate isotopomers specifically designed to maximize sensitivity for the diagnosis of fragment exchange by ^{13}C NMR spectroscopy were prepared. Specifically, [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate (**13a** and **13b,** respectively, Figure 33) were prepared from [3,4-$^{13}C_2$]- and [2,5-$^{13}C_2$]glucose, respectively, by the enzyme-assisted one-pot reaction strategy described earlier (Hecht *et al*., 2001). An enzyme-mediated recombination of fragments **21**, **22a**, **21a** and **22** generated from a mixture of [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythrose 4-phosphate (**19a** and **19b**, respectively) *via* the proposed retroaldol/aldol mechanism should result in the formation of four isotopologue species of 1-deoxy-D-xylulose 5-phosphate (**12a** - **12d**, Figure 33). Notably, the enzyme-mediated recombination of [1-$^{13}C_1$]glycolaldehyde (**21a**) and the enolate of [1-$^{13}C_1$]hydroxyacetone (**22a**) could then afford [3,4-$^{13}C_2$]1-deoxy-D-xylulose 5-phosphate (**12c**). This double-labeled species would be detected by satellite lines in the ^{13}C NMR spectrum due to $^{13}C^{13}C$ coupling as simulated in Figure 34C.

Alternatively, a sigmatropic rearrangement would develop an intermolecular rearrangement, affording **12a** from **13a** and **12b** from **13b** (Figure 33). The ^{13}C NMR signals of these species are simulated in Figure 34B.

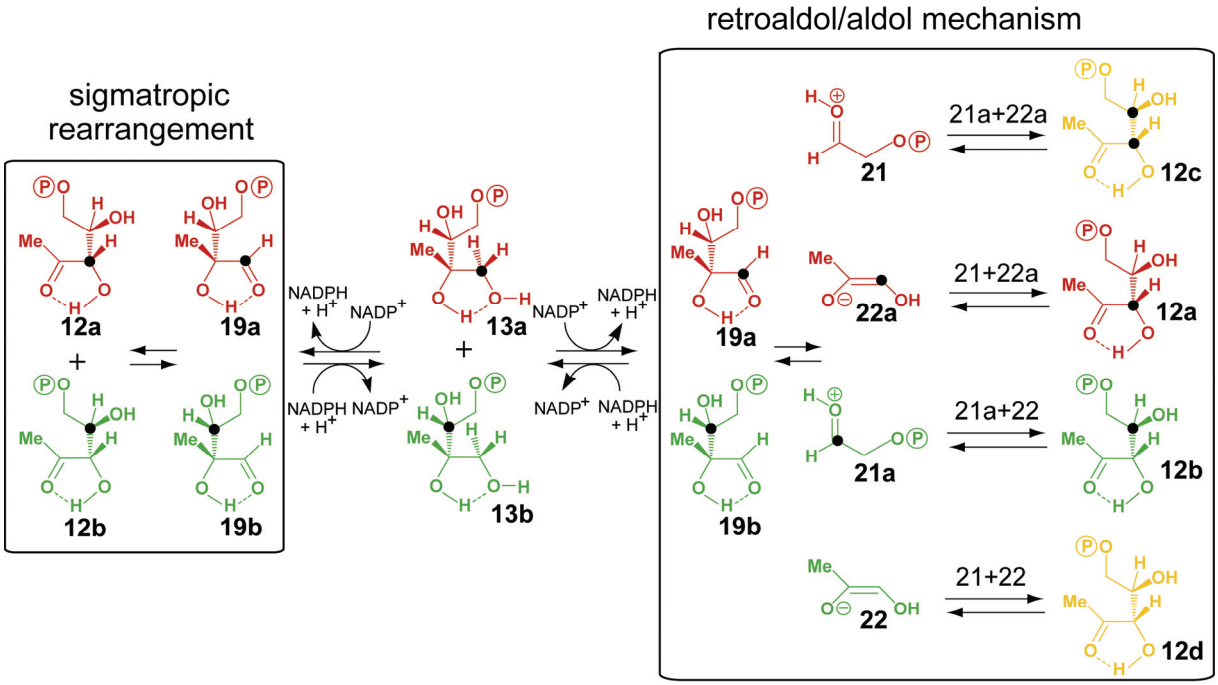


Figure 33: Hypothetical course of the IspC reaction starting from [1-$^{13}C_1$]- (**13a**) and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate (MEP/**13b**) as substrates
12a, [3-$^{13}C_1$]-; **12b**, [4-$^{13}C_1$]-, **12c**, [3,4-$^{13}C_2$]-; **12d**, unlabeled 1-deoxy-D-xylulose 5-phosphate (DXP), **19a**, [1-$^{13}C_1$]-; **19b**, [3-$^{13}C_1$] 2*C*-methyl-D-erythrose 4-phosphate, **21**, unlabeled-; **21a**, [1-$^{13}C_1$]glycolaldehyde phosphate, **22**, unlabeled-; **22a**, [1-$^{13}C_1$]enolate of hydroxyacetone.

In order to improve the diagnostic sensitivity, it was decided to conduct the experiments under steady state conditions where the reactants **12** and **13** are present in similar amounts at thermodynamic equilibrium. For that purpose, reaction mixtures containing 5 mM **13a**, 5 mM **13b**, 215 mM $NADP^+$ and 0.15-0.25 mM IspC protein from *E. coli*, *M. tuberculosis* or *A. thaliana* were incubated at pH 8 and 37 °C for 24 h and were monitored by ^{13}C NMR spectroscopy. The partial reduction of $NADP^+$ by the enzyme rapidly resulted in steady state conditions where the concentrations of **12** and **13** were approximately equal (cf. Figure 33). Consequently, the forward and the reverse reaction rate under equilibrium condition are also bound to be approximately equal. Notably, the IspC enzymes were present in very high

(near-stoichiometric) concentrations. Under these conditions, the substrate molecules should be engaged by enzyme molecules on a near-permanent basis.

The residual enzyme activity after 24 h incubation was measured after massive dilution of an aliquot of the reaction mixture using **1** as substrate. The decrease in activity during the 24 h incubation period was in the range of 27 – 37 % for the three different enzymes under study (Table 24).

Table 24: Calculated conversion and equilibrium constants for the IspC reaction

Organism	Residual activities after 24 h (%)	Conversion (%)	Cycles after 24 h	Apparent equilibrium constant
Escherichia coli	72	52	>8800	2.9×10^{-10} M
Mycobacterium tuberculosis	73	53	>12100	2.9×10^{-10} M
Arabidopsis thaliana	63	50	>2400	2.5×10^{-10} M

The ratio of conversion and the equilibrium constants were calculated after 1.5 h incubation (cf. Figure 34). For details, see under Methods 3.3.6.1.3.4.

From the starting conditions and the enzyme stability measurements under the reaction conditions in this study, it follows that an average substrate molecule should have passed through approximately 8,800, 12,100 and 2,400 forward/reverse cycles in the experiments with enzymes from *E. coli*, *M. tuberculosis* and *A. thaliana*, respectively (Table 24).

For the equilibrium constant of the reaction catalyzed by 2*C*-methyl-D-erythritol 4-phosphate synthase as defined by Equation 2, we obtained a value of $2.8 \pm 0.2 \times 10^{-10}$ M at pH 8.0. This is well in line with a value of $4.6 \pm 0.5 \times 10^{-10}$ M that had been reported earlier by Rohmer and coworkers under closely similar conditions (Hoeffler *et al*., 2002).

$$K = \frac{[NADPH]_{eq}.[\mathbf{12}]_{eq}.[H^+]}{[NADP^+]_{eq}.[\mathbf{13}]_{eq}}$$

Equation 2: Equilibrium constant
K, Equlibrium constant; **12**, 1-deoxy-D-xylulose 5-phosphate; **13**, 2*C*-methyl-D-erythritol 4-phosphate.

Figure 34 shows ^{13}C NMR signals of the reaction mixtures prior to the addition of enzyme (Figure 34D) and after incubation with enzymes from *E. coli* (Figure 34E), *M. tuberculosis* (Figure 34F) and *A. thaliana* (Figure 34G), respectively. Reaction mixtures treated with enzymes from the three different organisms studied showed closely similar results.

The crucial result is the absence of any detectable excess of the $^{13}C^{13}C$ coupling satellites beyond the natural abundance level for the signals of C-3 and C-4 of a hypothetical product **12c.** The hypothetical positions of the $^{13}C^{13}C$ coupling satellites expected in the spectrum of **12c** (as simulated in Figure 34C) are marked by arrows in Figure 34E-G. In each case, the integrals of the satellite signals are in the range of 1 % as compared to the central signal. Signals of that size would be expected in the complete absence of fragment exchange where they reflect the presence of about 1.1 % ^{13}C (i.e. natural ^{13}C abundance) in those carbon atoms of the reactant that were not labeled.

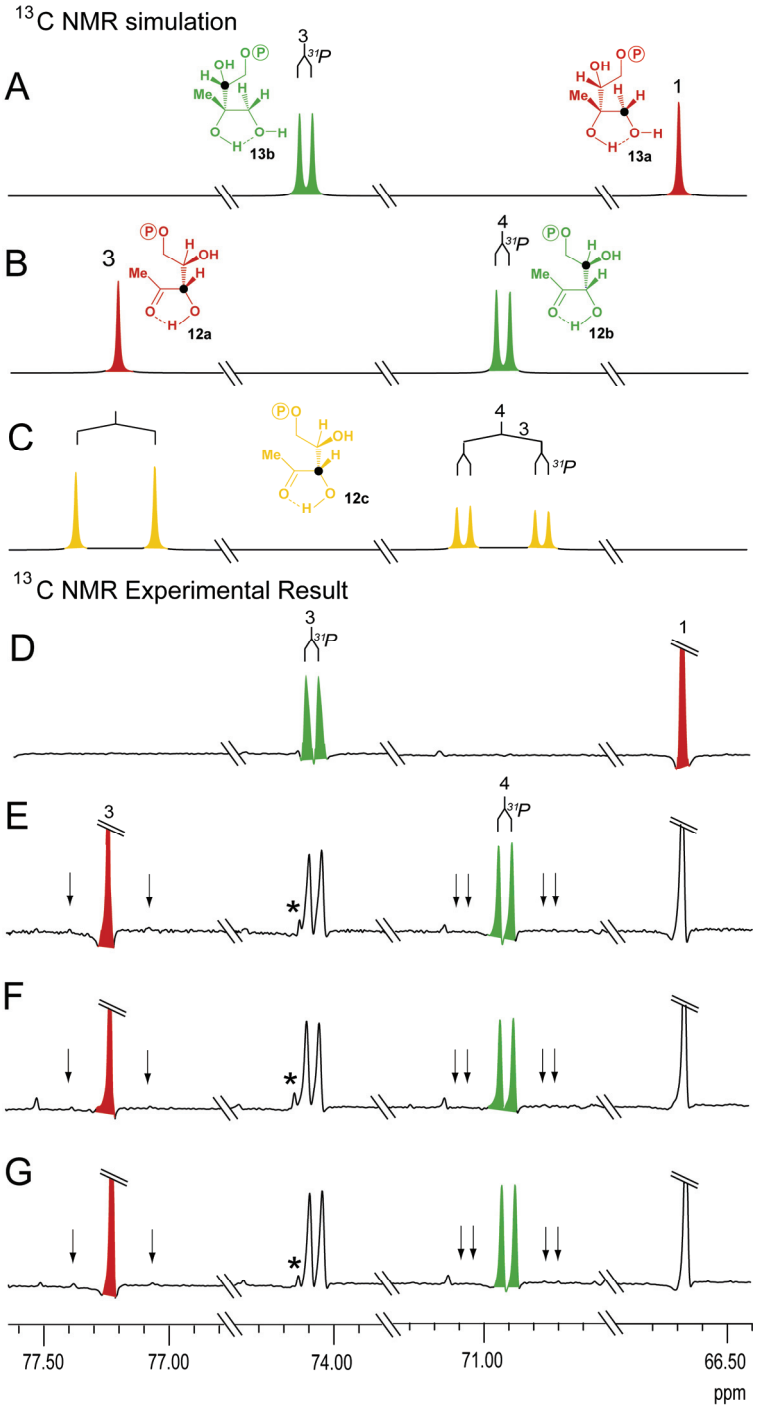


Figure 34: ^{13}C NMR simulation and experimental result of the IspC reaction using [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate as subtrates.
A-C are^{13}C NMR simulation of IspC reaction using [1-$^{13}C_1$]- and [3-$^{13}C_1$]2*C*-methyl-D-erythritol 4-phosphate as substrates with (A), NMR signals from initial substrates, (B), sigmatropic rearrangement; and (C), retroaldol/aldol mechanism (cf. Figure 33). The colors and structures shown are corresponding to the reaction in Figure 33D-F are the resulted NMR signals from ^{13}C-NMR spectroscopy with (D), initial substrates and incubation with IspC protein from (E), *E coli*; F, *M.tuberculosis* and (G), *A. thaliana*. The arrows show the chemical shift of signals should arise from retroaldol mechanism. Asterisks denote impurity.

On the basis of the quantitative evaluation of the ^{13}C NMR signal intensities and coupling satellites in experiments with ^{13}C-labeled substrates, it can be estimated that less than one fragment exchange has occurred during more than 100,000 reaction cycles. Whereas these data are not sufficient to rule out a retroaldol/aldol reaction sequence, they do show that a hypothetical retroaldol/aldol sequence would require extremely tight confinement of the intermediary molecular fragments at the active site of the enzyme. The limit for escape and reutilization of a retroaldol fragment would be less than once in 100,000 forward/reverse cycles. In this context, it is also worth noting that the branched intermediate 2*C*-methyl-D-erythrose 4-phosphate (**19**) can be used as substrate by the enzyme at a rate that is comparable with the conversion rate of the substrate **12** (Hoeffler *et al*., 2002); thus, strict confinement seems at least not to apply to that intermediate.

4.1.3.5 Mechanistic study of IspC reaction using [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate and unlabeled hydroxyacetone in the enzyme assay

To check whether exogenous hydroxyacetone, whose enolate is the predicted intermediate of the hypothetical retroaldol/aldol mechanism, can be incorporated into the reactants **12** and 1**3** by fragment exchange, [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate (**13c**) and unlabeled hydroxyacetone (**23**) were used in the enzyme assays (Figure 35). Preliminary experiments had shown that hydroxyacetone does not significantly change the apparent reaction rate of 2*C*-methyl-D-erythritol 4-phosphate synthase when present in concentrations up to 2 % (v/v). The reaction mixtures contained 10 mM [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate (**13c**, Figure 35), 215 mM NADP+, 100 mM Tris hydrochloride, pH 8.0, and 0.23 – 0.25 mM IspC protein from *E. coli*, *M. tuberculosis* and *A. thaliana*, respectively, and were incubated for 8 h and analyzed by NMR spectroscopy.

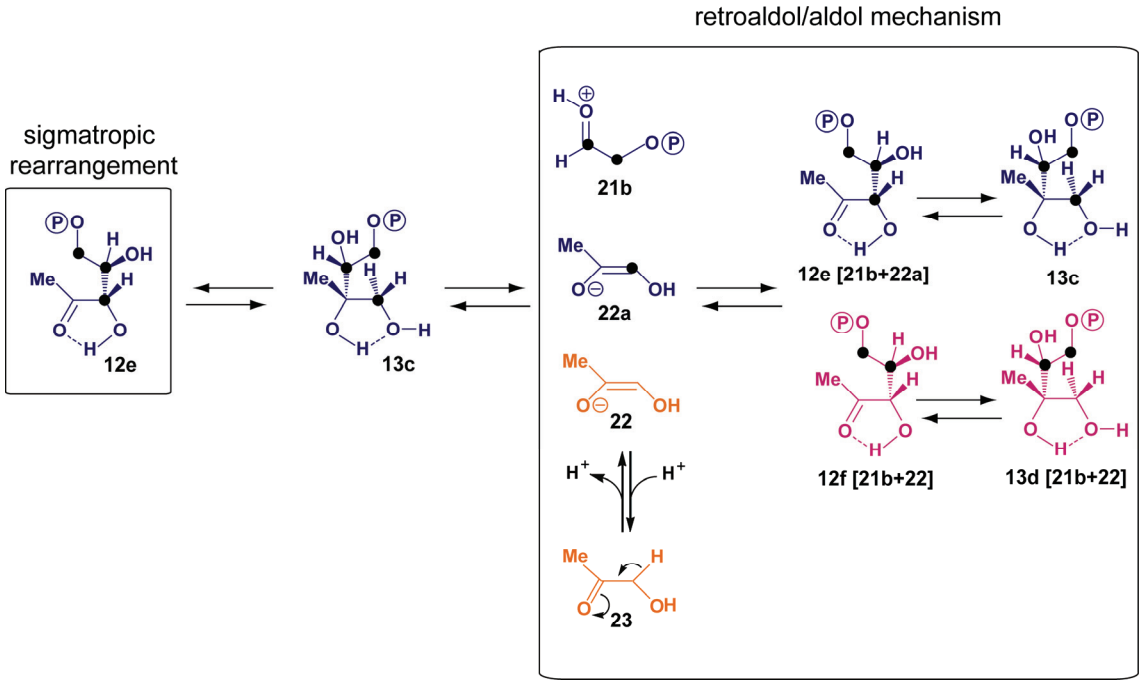


Figure 35: Hypothetical course of the IspC reaction starting from [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate (**13c**) as substrate in the presence of an excess of unlabeled hydroxyacetone (**23**)
12e, [3,4,5-$^{13}C_3$]-;**12f**, [4,5-$^{13}C_2$]1-deoxy-D-xylulose 5-phosphate; **13d**, [3,4-$^{13}C_2$]2*C*-methyl-D-erythritol 4-phosphate; **21b**, [1,2-$^{13}C_2$]glycolaldehyde phosphate; **22**, unlabeled-; **22a**, [1-$^{13}C_1$]enolate of hydroxyacetone.

As described before (see 4.1.3.4), these initial conditions were rapidly conducive to steady conditions where **12** and **13** were present in closely similar concentrations, and the rates of the forward reaction (conversion of **12** to **13**) and the backward reaction (conversion of **13** into **12**) were also essentially the same. The diversion of unlabeled hydroxyacetone (**23**) to the reactants **12** and **13**, by occasional exchange with the enolate of [1-$^{13}C_1$]hydroxyacetone (**22a**) formed by retroaldol cleavage of **12e** or **13c** (Figure 35), should result in **12f** which is characterized by only two ^{13}C atoms. This isotopologue would become apparent in the ^{13}C NMR spectra by a distinctive double doublet signature of C-4 of **12f** as simulated in Figure 36C (indicated by arrows in Figure 36E-G). Notably, the formation of double labeled **12f** would argue against a sigmatropic reaction mechanism.

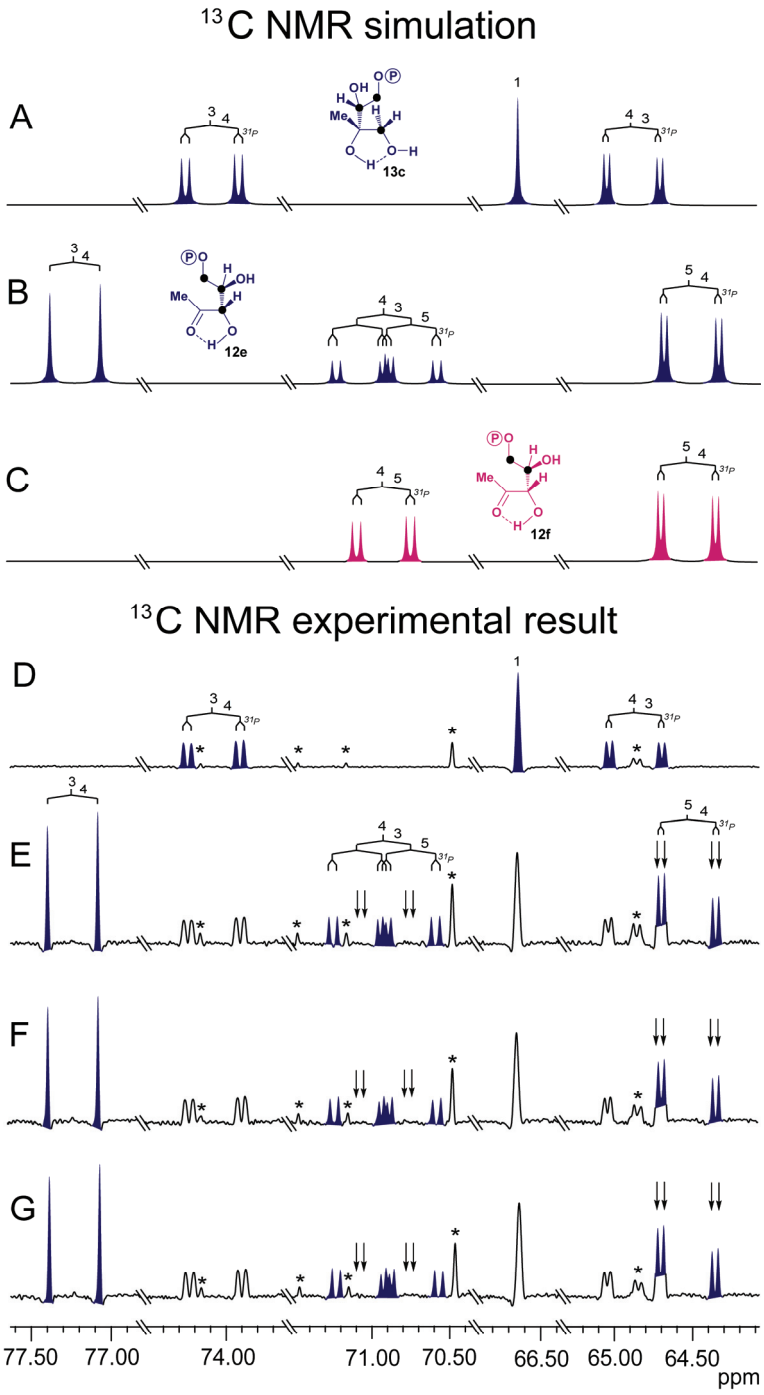


Figure 36: ^{13}C NMR simulation and experimental result of the IspC reaction using [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate (**13c**)as substrates in the presence of an excess of unlabeled hydroxyacetone (cf. Figure 35)
A-C are ^{13}C NMR simulation of the arising signals of (cf. Figure 35): (A), initial substrate; (B), sigmatropic mechanism; (C), retroaldol mechanism. D-G are ^{13}C NMR result depicting the signals from (D), initial substrate; and E-G, IspC assay from *E. coli*, *M. tuberculosis* and *A. thaliana*, respectively. The arrows show the chemical shift where the signals from retroaldol mechanism should arise. Asterisks denote impurity.

Figure 36E-G shows ^{13}C NMR signals detected in the exchange experiment with unlabeled hydroxyacetone. Signal intensities showed that 59 % of the **13c** had been converted into **12e** (Table 25). The crucial double doublets as expected for a retroaldol mechanism (Figure 36E-G) were absent. Again, no indication for a retroaldol/aldol mechanism could be obtained.

Table 25: Calculated conversion and cycles for the IspC reaction containing [3,4,5-$^{13}C_3$]-DXP and hydroxyacetone

Organism	Residual activities after 8 h (%)	Conversion (%)	Cycles after 8 h
Escherichia coli	71	59	>4100
Mycobacterium tuberculosis	86	59	>4000
Arabidopsis thaliana	59	59	>1000

The ratio of conversion was calculated after 2 h incubations. For details see 3.3.6.1.3.4.

4.1.3.6 Mechanistic study of IspC reaction using [1,2-$^{13}C_2$]glycolaldehyde phosphate and unlabeled hydroxyacetone in the enzyme assay

In the following set of experiments, it was checked whether glycolaldehyde phosphate and hydroxacetone can serve as substrates for IspC protein from different organisms to form 2*C*-methyl-D-erythritol 4-phosphate as shown in the hypothetical reaction sequence as illustrated in Figure 37. Specifically, the reaction mixtures contained 2 mM [1,2-$^{13}C_2$]glycolaldehyde phosphate (**21b**), 243 mM hydroxyacetone (**22**), Tris hydrochloride, pH 8.0, 3 mM NADPH and 0.21 – 0.34 mM IspC protein from *E. coli*, *M. tuberculosis* and *A. thaliana*, respectively.

As shown in Figure 37, a retroaldol mechanism in the presence of NADPH as the cofactor in the reaction would convert the double labeled [1,2-$^{13}C_2$]glycolaldehyde phosphate (**21b**) and unlabeled hydroxyacetone into their respective double labeled product [3,4-$^{13}C_2$]2*C*-methyl-D-erythritol 4-phosphate (**13d**). The ^{13}C NMR signals of the product **13d** was determined by the addition of [3,4,5-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate (**13c**) as a standard. The crucial signals of **13d** would arise as double doublet at 73.8 ppm as shown in Figure 37E. Alternatively, in a sigmatropic

rearrangement; both fragments would not be recognized by the IspC protein and therefore, any reaction would not occur.

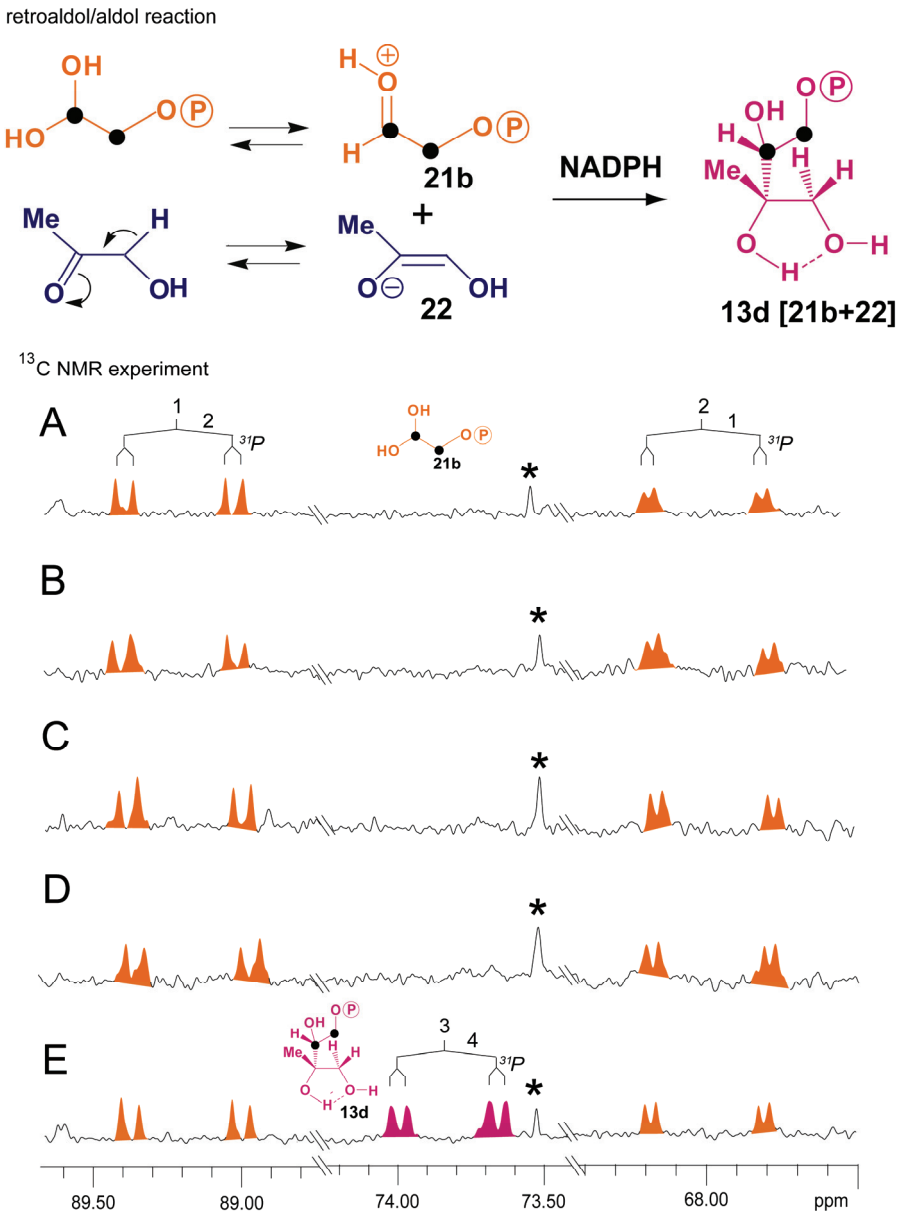


Figure 37: Hypothetical isotopologue species formed by a retroaldol mechanism of the IspC reaction with protonated [1,2-$^{13}C_2$]glycolaldehyde phosphate (**21b**) and the enolate of hydroxacetone (**22**) as initial substrates
13d, [3,4-$^{13}C_2$]2*C*-methyl-D-erythritol 4-phosphate. A-E are ^{13}C NMR spectra obtained from IspC reactions using [1,2-$^{13}C_2$]glycolaldehyde phosphate (**21b**) and hydroxacetone (**23**) as substrates. (A), without IspC enzyme; (B), (C) and (D) with IspC from *E. coli*, *M. tuberculosis* and *A. thaliana*, respectively. (E), with [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate after 3 h incubation of the reaction mixture, (B). Signals arising from C-3 of **13d**.

The ^{13}C NMR spectra obtained after an incubation period of 1.5 and 3 h, respectively, showed only double doublet signals at 89.2 ppm due to the presence of the hydrate of **21b** (Figure 37B-D**)**. As shown in Figure 37, no evidence for the presence of [3,4-$^{13}C_2$]2*C*-methyl-D-erythritol 4-phosphate (**13d**) could be obtained. Notably, it would have been possible to detect any **13d** by the specific double doublet signature of C-3, as confirmed by a titration experiment with [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 4-phosphate (**13c**) (Figure 37E). As a result, the ^{13}C NMR spectrum showed no signals of **13d** eventually emerging in the reaction (Figure 37B-D), hence, no indication for the existence of retroaldol mechanism.

4.1.3.7 Investigation of the role of 1-deoxy-L-ribulose 5-phosphate in the IspC catalyzed reaction

Wong and Cox reported the formation of the 4-epimer 1-deoxy-L-ribulose 5-phosphate of **12** (**24b**, Figure 38) in an IspC reaction mixture without NADPH and a divalent metal ion (Wong and Cox, 2007). Specifically, they observed a new ^{13}C NMR signal at 71.6 ppm, which was assigned to C-4 of **24**. In contrast to this finding, we observed a corresponding signal (71.49 ppm) together with the signals of C-3 and C-5 in our preparations of [3,4,5-$^{13}C_3$]-**12** before any incubation.

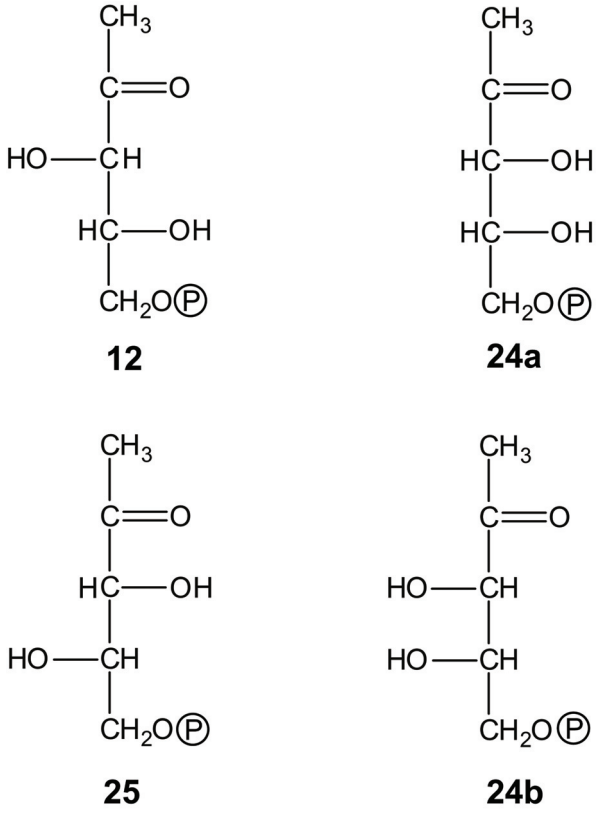


Figure 38: The enantiomer and diastereomers of 1-deoxy-D-xylulose 5-phosphate
12, 1-deoxy-D-xylulose 5-phosphate; **24a**, 1-deoxy-D-ribulose 5-phosphate; **24b**, 1-deoxy-L-ribulose 5-phosphate; **25**, 1-deoxy-L-xylulose 5-phosphate.

As shown in the ^{13}C NMR spectrum (Figure 39A), [3,4,5-$^{13}C_3$]-**24** is present in a ratio of 2.5 % as compared to [3,4,5-$^{13}C_3$]-**12**. After incubation with 0.2 mg and 8 mg IspC protein for 20 h in the mixture containing 4 mM EDTA, 100 mM Tris/HCl, pH 8.0 and 10 % (v/v) D_2O in a total volume of 500 µl, the ratio of [3,4,5-$^{13}C_3$]-**24** increased up to

5 % similarly as a control assay without any enzyme (Figure 39). Any new signals could not be observed.

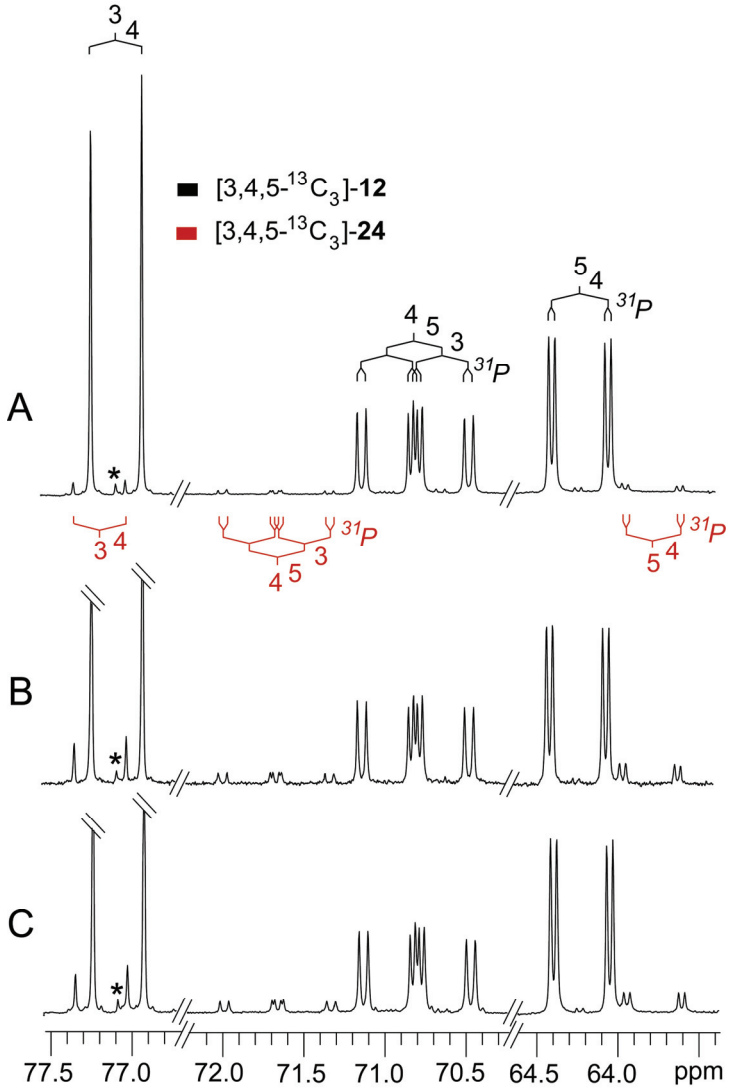


Figure 39: ^{13}C NMR measurement of the reaction catalyzed by IspC protein in the absence of NADPH and metal ions
The reaction using [3,4,5-$^{13}C_3$]- 1-deoxy-D-xylulose 5-phosphate as substrate was observed; A, before incubation; B, without enzyme and C, with IspC enzyme from *E. coli*; **24**, potential diastereomer of **12**; asterisk denotes impurity.

Consequently, **12** was incubated in buffer containing Tris/HCl, pH 8.0 and 10 % (v/v) D_2O at 37 °C for a period of 8 days and analyzed the sample by ^{13}C NMR spectroscopy in time intervals of 1 day. The ^{13}C NMR spectra (Figure 40) showed the increase of the formation of **24** followed by simultaneous decrease of **12**. The rate constant for the formation of **24** was calculated to 3 x 10^{-7} s^{-1}. The ratios for

consumption of **1** and the formation of **24**, respectively (Figure 40, in box), indicate that the equilibrium is reached after 8 days of incubation. The equilibrium constant of [**12**]/[**24**] was calculated to 3.45.

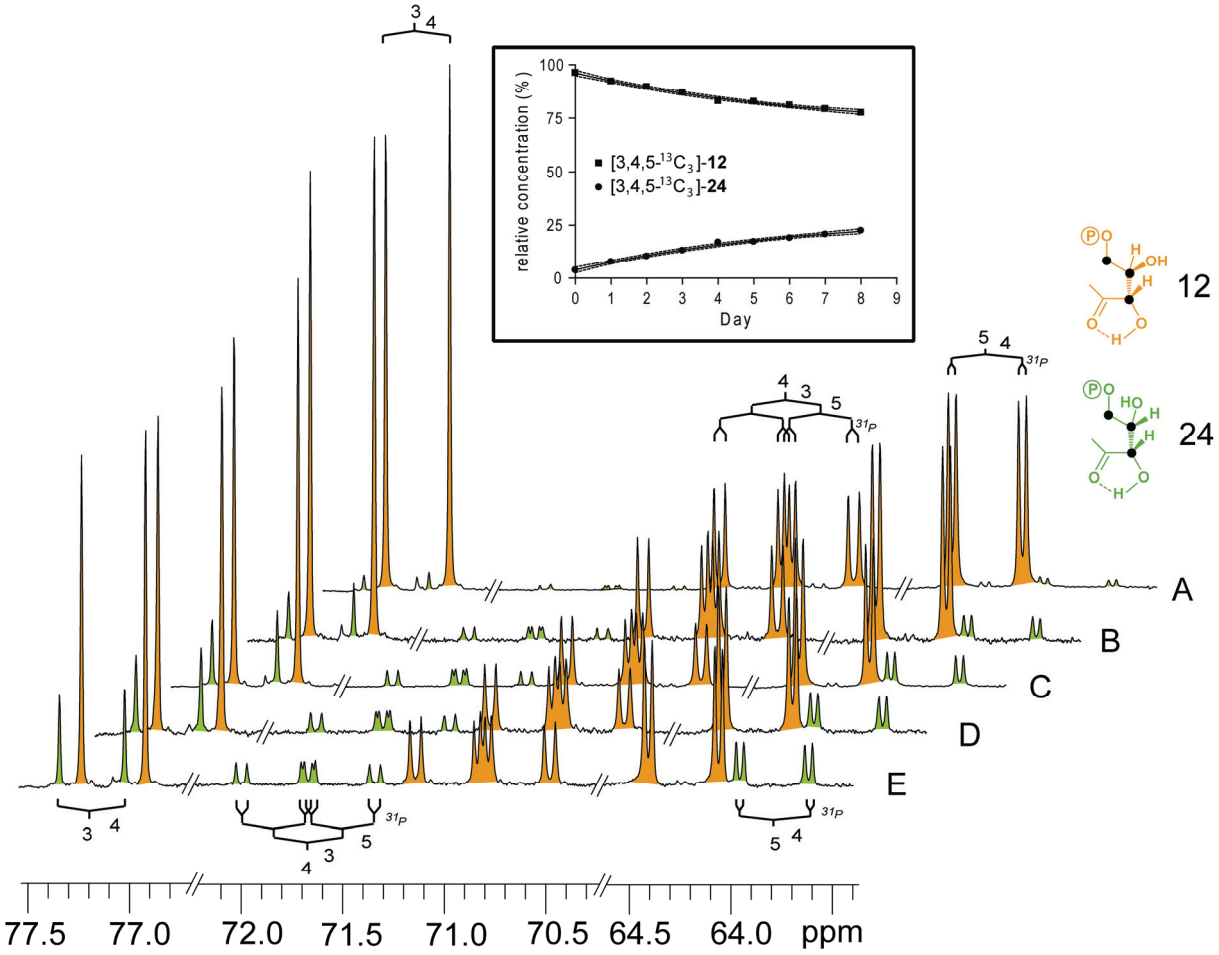


Figure 40: ^{13}C NMR spectra of [3,4,5-$^{13}C_3$]1-deoxy-D-xylulose 5-phosphate and its diastereomer (**24**)
The inserted curve in box shows the relative concentration of [3,4,5-$^{13}C_3$]1-deoxy-D-xylulose 5-phosphate (**12**) and its potential diastereomer (**24**) during 8 days incubation.

The next experiment was performed to investigate the structure of compound **24** *via* NMR analysis. For this purpose, [U-^{13}C]-**12** was prepared according to the published procedure (Hecht *et al*., 2001; Illarionova *et al*., 2006). The ^{13}C NMR spectra (Figure 41) shows the presence of **24** after the preparation of **12**, specifically the mixture contains 82 % of [U-^{13}C]-**12** and 18 % [U-^{13}C]-**24**.

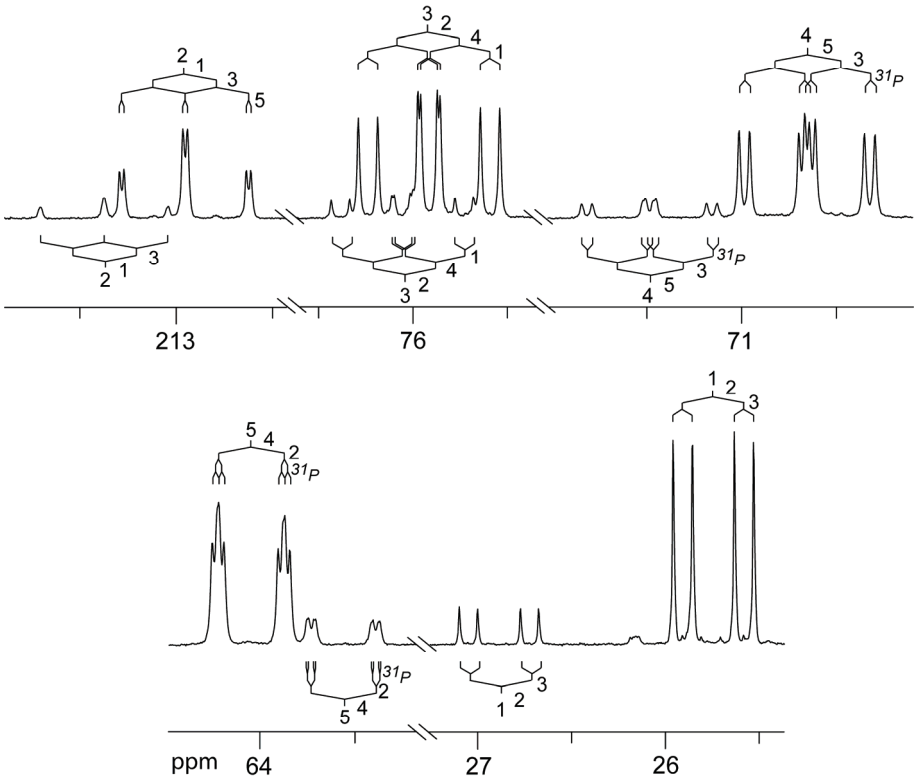


Figure 41: ^{13}C NMR signals of [U^{13}-C]1-deoxy-D-xylulose 5-phosphate (**12**) and its diastereomer (**24**) (cf. Figure 38)
The upper ^{13}C coupling patterns are originated from U^{13}C 1-deoxy-D-xylulose 5-phosphate (**12**) and the lower ^{13}C coupling patterns, to 1-deoxyribulose 5-phosphate (**24)**.

All ^{13}C signals of **24** (resulted from a preparation of uniformly labelled **12**) were assigned on the basis of the coupling constants for $^{13}C^{13}C$ and $^{13}C^{31}P$ interactions. The ^{1}H NMR signals were then assigned by HMQC spectroscopy (Figure 41 and Table 26). The NMR data were in close correspondence with those of 1-deoxy-L-ribulose-5-phosphate (**24b**, Figure 38) chemically synthesized earlier (Phaosiri and Proteau, 2004).

As seen in Figure 41 and Table 26, the ^{13}C chemical shifts as well as the $^{13}C^{13}C$ and $^{13}C^{31}P$ coupling constants of DXP (**12**) and DRP (**24**) are similar. This suggested that both compounds are structurally related and might display isomers, such as diastereomer. Notably there exist two diastereomers of **12**, namely 1-deoxy-D-ribulose 5-phosphate (**24a**) and 1-deoxy-L-ribulose 5-phosphate (**24b**) (seeFigure 38) which show identical NMR data. On the basis of NMR, the configuration of **24** can not be distinguished. Thus, **24** could be **24a**, **24b** or a racemate of **24a** and **24b**.

Table 26 Chemical shifts and coupling constants of [U-^{13}C]1-deoxy-D-xylulose 5-phosphate (DXP/**12**) and 1-deoxyribulose 5-phosphate (DRP/**24**) obtained by ^{13}C NMR spectroscopy and HMQC.

^{13}C Position (ppm)	Chemical shift (ppm)				Coupling Constant (Hz)			
	^{13}C		^{1}H		^{13}C^{31}P		^{13}C^{13}C	
	DXP	DRP	DXP	DRP	DXP	DRP	DXP	DRP
1	25.75	26.90	2.17	2.20	-	-	41.1 (2) 13.0 (3)	40.7 (2) 12.6 (3)
2	212.96	213.38	-	-	-	-	41.1 (1) 41.1 (3) 2.8 (5)	41.7 (1) 41.7 (3) -
3	76.92	77.06	4.40	4.25	-	-	41.6 (2) 39.4 (4) 12.8 (1)	42.5(2) 40.5(4) 12.6(1)
4	70.66	71.49	4.21	3.97	7.1	7.1	43.6 (5) 39.6 (3)	42.5 (5) 40.2 (3)
5	64.06	63.58	3.75	3.72	4.6	4.8	43.6 (4) 3.3 (2)	42.3 (4) 1.3(2 **?**)

4.1.3.8 Mechanistic discussion of IspC protein from the structural point of view

The importance of the hydroxy groups at C-3 and C-4 of 1-deoxy-D-xylulose 5-phosphate (**12)** has been shown by analogues of **12** comprising hydroxy functions at the respective positions. Two substrate analogues, namely 3-deoxy-**12** and 4-deoxy-**12** served as weak inhibitors (Hoeffler *et al*., 2002). On the other hand, the other substrate analogues 1-fluoro,1-deoxy-D-xylulose 5-phosphate (Fox and Poulter, 2004) and 1-deoxy-D-xylulose 5-phosphonate (Phaosiri and Proteau, 2004) have been demonstrated to replace 1-deoxy-D-xylulose 5-phosphate (**12**) as substrates in the IspC reaction. Notably, both alternative substrates possess hydroxy groups at C-3 and C-4 position of **12**. Regarding the important role of hydroxy groups at C-3 and C-4 of **12**, the reaction can be considered to follow a retroaldol mechanism. Alternatively, the respective hydroxy groups might act as metal binding sites, indicating the essential role of the respective hydroxy groups in both sigmatropic and retroaldol mechanisms.

The experiments performed in this study did not provide any evidence for a retroaldol mechanism. In the experimental setting using [1-^{13}C]- and [3-^{13}C]2*C*-methyl-D-

erythritol 4-phosphate (**13a** and **13b**) as substrates, the crucial double labeled isotopologues [3,4-$^{13}C_2$]-**12 (12c**), were not detected *via* ^{13}C NMR analysis (Figure 33 and Figure 34). Additionaly, the unlabeled hydroxyacetone was not incorporated into [1,3,4-$^{13}C_3$]-**13 (13c**) in the IspC reaction, therefore did not produce [4,5-$^{13}C_2$]-**12** (**12e**) (Figure 35 and Figure 36). Moreover, [1,2-$^{13}C_2$]glycoaldehyde phosphate (**21b**) and unlabeled hydroxyacetone (**23**) were not recognized as substrates in the IspC reaction, notably ^{13}C NMR spectra did not show the formation of **13d** (Figure 37). Instead, the above-mentioned experiments afforded the products that can be explained by a sigmatropic mechanism.

The next experiment was performed to detect the enzymatic generation of **24b** (Figure 38) which was claimed as a product of retroaldol mechanism (Wong and Cox, 2007). As a surprise, it was observed that ^{13}C NMR signals due to **24b** or its enantiomeric form were already present in the assay mixture containing [1,3,4-$^{13}C_3$]-**12** prior the incubation. The signal intensities of **24b** were 2.5 % as compared to the signals due to [1,3,4-$^{13}C_3$]-**12**. The incubation of the substrate [1,3,4-$^{13}C_3$]-**12** at 37°C for 20 h resulted in an increase of the concentration of **24b** by a factor of 2 in the absence of metal ion. Notably, this increase was not dependent on the presence of IspC protein. On the basis of this result, it can be concluded that **24b** is a spontaneous isomerization product of **12** and not an intermediate in the IspC reaction *via* a retroaldol mechanism in the absence of metal ion as claimed earlier (Wong and Cox, 2007). Moreover, the crystal structure of *E. coli* IspC protein in complex with **24b** illustrated that the C-3 hydroxy group of **24b** is not involved in any hydrogen bonding (Mac Sweeney *et al*., 2005) explaining why **24b** is not recognized as a substrate. Instead, it was found that it functions as a weak inhibitor with a K_i value of 180 µM (Phaosiri and Proteau, 2004).

The crystal structure of IspC protein from *E. coli* in complex with the inhibitor fosmidomycin and Mn^{2+} (Steinbacher *et al*., 2003) showed that the metal ion is ligated to the carboxylic oxygen atoms of Asp-150, Glu-152, Glu-231, and with the oxygen atoms at C-1 and N-hydroxy of fosmidomycin. The crystal structure of IspC from *E. coli* in complex with fosmidomycin, Mg^{2+} and NADPH (Yajima *et al*., 2007) confirmed this binding topology and displayed the same conformation of fosmidomycin (Figure 42A, green). When the substrate, 1-deoxy-D-

xylulose 5-phosphate (**12**), was modeled into the structure, an interaction between the metal ion and the carbonyl group of C-2**,** as well as with the hydroxy group at C-3 of **12** was suggested (Steinbacher *et al*., 2003). This topology which is similar to the first step of the reaction catalyzed by acetohydroxy acid isomeroreductase, illustrates that the carbonyl group of C-2 is polarized by the metal ion, thus facilitating the rearrangement (sigmatropic) reaction in a scheme where the hydroxy group at C-3 is fixed by the coordination to the metal ion (Steinbacher *et al*., 2003).

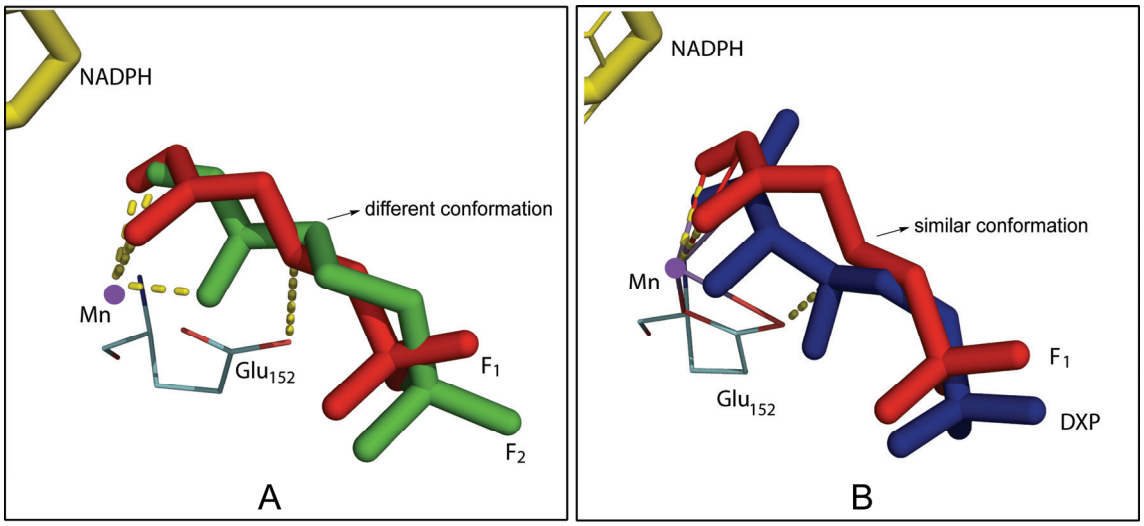


Figure 42: Comparison of DXP modeled into the IspC crystal structures at different conformations
A, comparison of DXP conformation from pdb files: 1Q0Q (Mac Sweeney *et al*.,2005, red) and 1ONP (Steinbacher *et al*.,2003, green); B, comparison of DXP conformation from pdb files:1Q0Q (Mac Sweeney 2005, red) and 2JCZ (Henriksson *et al*.,2007, blue).

On the other hand, a different conformation of fosmidomycin (Figure 42B, blue) was reported on the basis of a complex structure of IspC from *M. tuberculosis*, with the inhibitor, Mn^{2+} and NADPH. When the substrate was modeled into this scaffold, the hydroxy group at C-4 of **12** was found close to Mn^{2+} and the neighbouring carboxylate group of Glu153 (corresponding to Glu152 in the *E. coli* protein) could serve as an acceptor for H^{+} from the C-4 hydroxy group. Additionally, Glu222 (corresponding to Glu231 in the *E. coli* protein) was found to be perfectly positioned for deprotonation at C-3 (Henriksson *et al*., 2007). Indeed, a complex structure with the substrate, 1-deoxy-D-xylulose 5-phosphate (Figure 42A and B, red), and with NADPH showed the predicted binding (Mac Sweeney *et al*.,

2005). This topology supports a hypothetical aldol/retroaldol-type mechanism where deprotonation at the C-4 hydroxy group is followed by the cleavage between C-3 - C-4 of **12** and deprotonation at C-3 hydroxy yielding glycoaldehyde phosphate and hydroxyacetone as intermediates.

The crystallographic data of *E. coli* IspC protein also revealed the existence of a flexible loop at the active site (amino acid residues 206-216) (Reuter *et al*., 2001; Steinbacher *et al*., 2003; Mac Sweeney *et al*., 2005; Yajima *et al*., 2007) which is capable to fold into at least three different conformations depending on the ligands in complex with the protein. Specifically, the structure of the apoenzyme displays an unordered loop comprising amino acids 206 – 216 (Figure 43A), whereas the NADPH and especially the complex with bound NADPH and 1-deoxy-D-xylulose 5-phosphate showed a well ordered loop closing the active site region of the enzyme (Figure 43B). This can be taken as evidence that upon substrate binding, this variable, initially disordered lid interacts with NADPH and closes the active site, protecting the reactants completely from the solvent environment. The finding could also explain why the intermediate, 2*C*-methyl-D-erythrose 4-phosphate (**19**) as well as other putative intermediates can not be trapped.

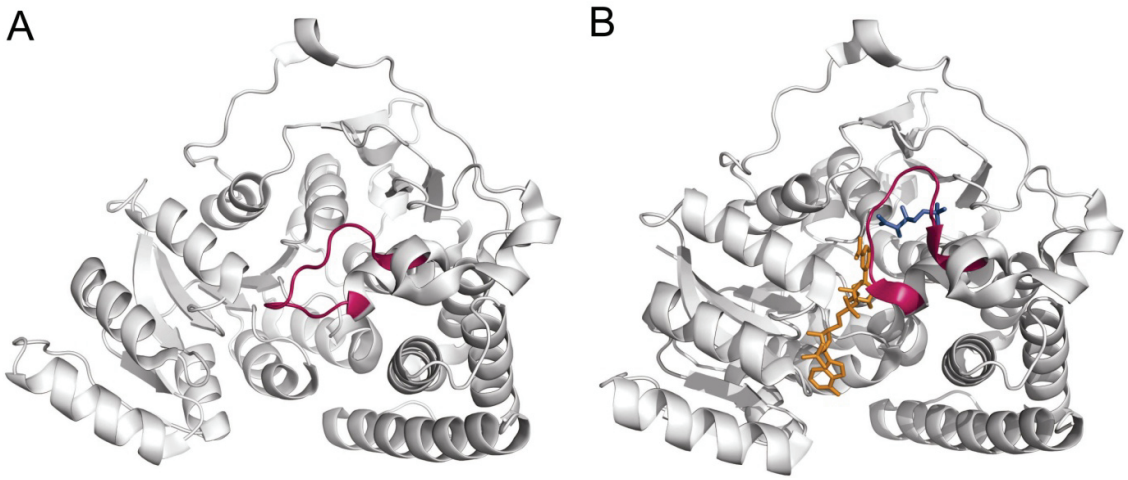


Figure 43: Crystal structures of monomeric IspC protein from *E. coli*.
(A), apoenzyme (pdb file 1K5H(Reuter *et al*.,2002)) and (B), enzyme in complex with NADPH (orange) and 1-deoxy-D-xylulose 5-phosphate (blue) (pdb file 1Q0Q(Mac Sweeney *et al*., 2005)). The flexible loop (residues 206-216) in both structures is shown in magenta.

Peculiarly, the loop is capable to open the cavity of the active site in the presence of compounds like 1-deoxy-1-fluoro-D-xylulose 5-phosphate (Fox and Poulter, 2004) and 1-deoxy-D-xylulose 5-phosphonate (Phaosiri and Proteau, 2004) as analogues of the substrate as well as 2*C*-methyl-D-eryhtrose 4-phosphate (**19**) as analogue of the product (Hoeffler *et al*., 2002). These analogues were accepted as substrates by IspC protein from *E. coli* with K_M of 227 µM, 690 µM and 158 µM respectively. This concept could explain why no IspC reaction precedes when hydroxyacetone and glycoaldehyde phosphate were used instead of substrate. Both substances do not act as substrate analogues but rather as possible intermediates in the retroaldol mechanism. They are not able to induce the opening of the lid of the active site, therefore are not recognized by the protein.

4.2 IspE protein

4.2.1 Characterization of IspE protein from *Aquifex aeolicus*

4-Diphosphocytidyl 2*C*-methyl-D-erythritol kinase (IspE protein) catalyzes the conversion of 4-diphosphocytidyl 2*C*-methyl-D-erythritol into 4-diphosphocytidyl

2*C*-methyl-D-erythritol 2-phoshate. The catalytic activity of the recombinant IspE protein was measured by 13C NMR spectroscopy and photometry.

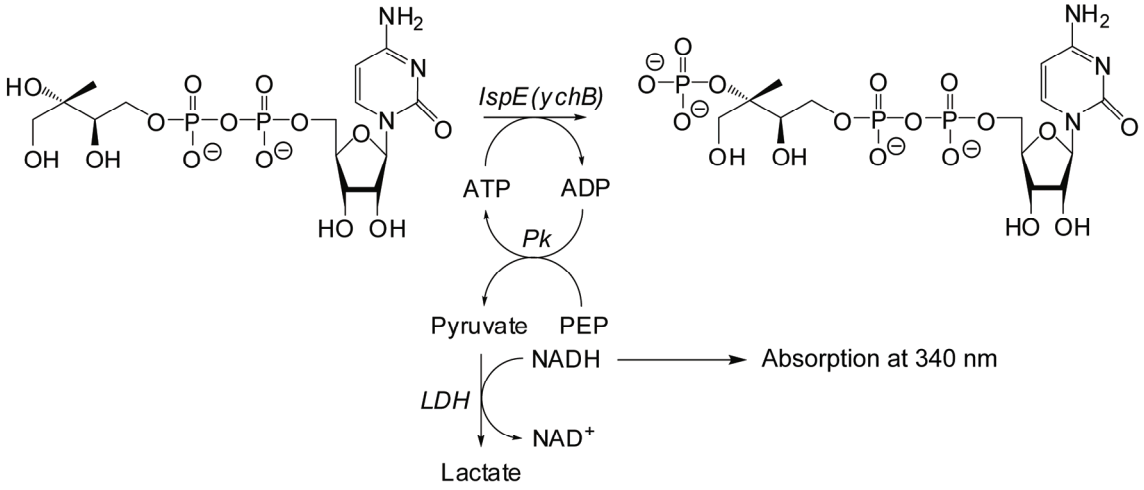


Figure 44: Photometric assay of the IspE reaction

The photometric assay (see Methods 3.3.6.2.1) was performed according to published procedure (Illarionova *et al*., 2006) using auxiliary enzymes in order to observe the decrease of NADH absorption at 340 nm as shown in Figure 44.

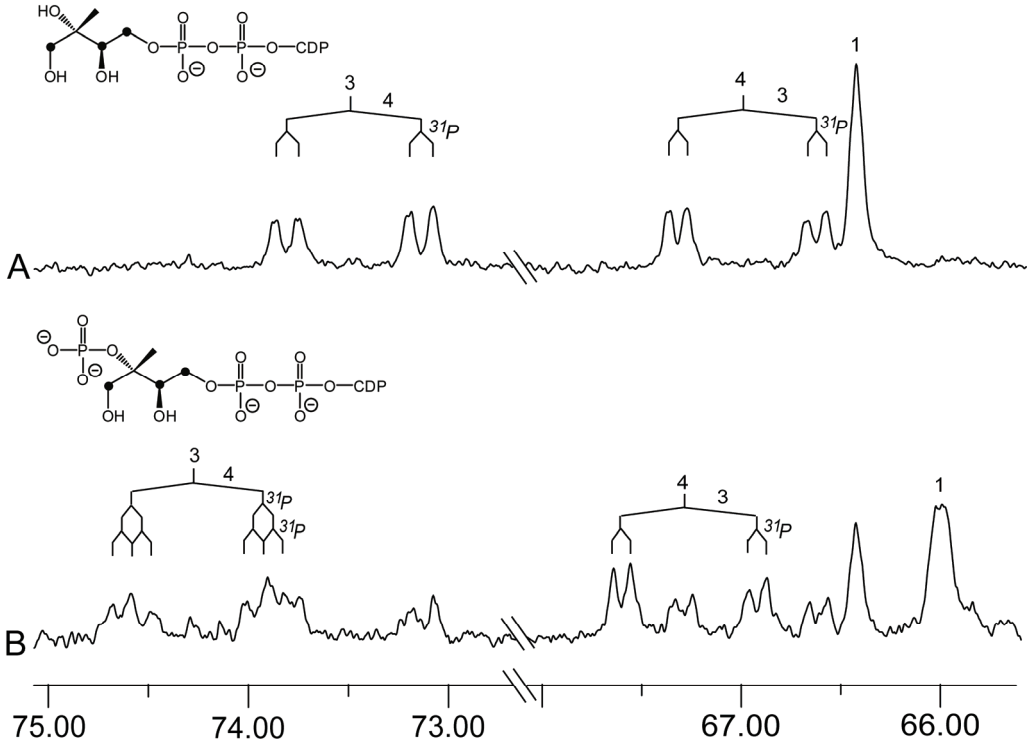


Figure 45: ^{13}C NMR spectra of [1,3,4-$^{13}C_3$]4-diphosphocytidyl-2*C*-methyl-D-erythritol and [1,3,4-$^{13}C_3$]4-diphosphocytidyl-2*C*-methyl erythritol-2-phosphate

The ^{13}C NMR measurement was performed using multiply ^{13}C labeled substrate [1,3,4-$^{13}C_3$]4-diphosphocytidyl 2*C*-methyl-D-erythritol to enhance the sensitivity and selectivity of the observation (see Methods 3.3.6.2.2). ^{13}C NMR signals detected in a typical NMR assay are shown in Figure 45. The chemical shifts as well as the $^{13}C^{13}C$ coupling constants of 4-diphosphocytidyl 2*C*-methyl-D-erythritol and 4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate have been published earlier (Rohdich *et al*., 1999; Lüttgen *et al*., 2000).

4.2.1.1 Cloning and expression of the *ispE* gene from *Aquifex aeolicus*

The cloning and expression of *ispE* gene from *A. aeolicus* was carried out in the laboratory of 'Division of Biological Chemistry and Drug discovery, College of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK'. The information on the recombinant protein is shown in Table 27.

Table 27 Cloning of the *ispE* gene from *Aquifex aeolicus*

Accesion number	Uniprot (entry O67060)
Vector	pET15b vector (Novagen)
Host	*E. coli*, BL21 (DE3)

4.2.1.2 Purification of recombinant IspE protein from *Aquifex aeolicus*

The purification of IspE protein from *Aquifex aeolicus* was completed in the laboratory of 'Division of Biological Chemistry and Drug discovery, College of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK'. The gene was preceded by a His_6-tag, to enable the purification of the recombinant protein *via* metal chelating affinity chromatography. The polyhistidine tag was removed by thrombine-mediated proteolysis, followed by anion exchange chromatography affording the purified protein in a buffer containing 10 mM Tris/HCl, pH 8.5, 20 mM NaCl and 1 mM DTT. The purified protein had an apparent mass of 30 kDa as judged by gel-filtration chromatography, in a good agreement with sedimentation velocity experiment affording a calculated mass of 29.8 kDa.

4.2.1.3 Metal dependence of IspE protein from *Aquifex aeolicus*

The activation of IspE protein by metal ions was analyzed using ^{13}C NMR spectroscopy as described under Methods 3.3.6.2.2. The reaction was started by the addition of 5 mM of various metal ions as shown in Table 28.

Table 28 Metal dependence of *Aquifex aeolicus* IspE

Metal	Relative activity (%)
None	0
Mg^{2+}	100
Mn^{2+}	94
Co^{2+}	52
Cu^{2+}	40
Fe^{2+}	16
Zn^{2+}	13.5
Ni^{2+}	10.7
Ca^{2+}	8.1

Aquifex aeolicus IspE showed maximum activity in the presence of Mn^{2+} or Mg^{2+}. The optimum concentration for both Mn^{2+} and Mg^{2+} was 2 mM (Figure 46).

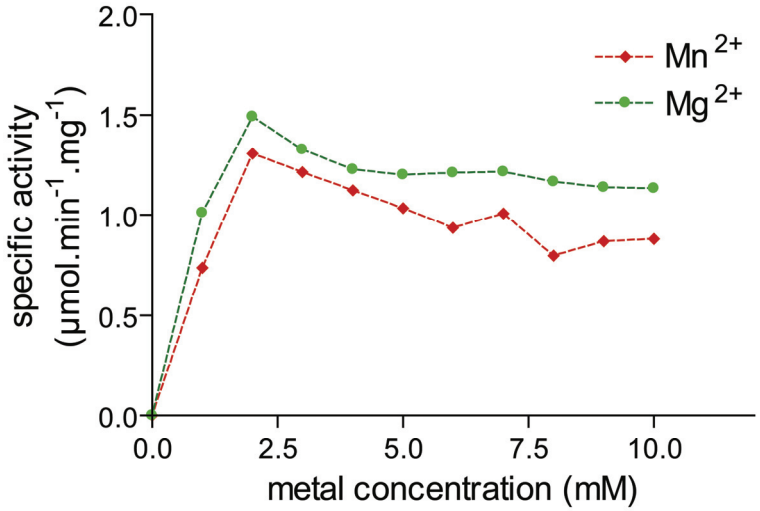


Figure 46: Mg^{2+} and Mn^{2+} dependence of IspE protein from *Aquifex aeolicus*

4.2.1.4 PH dependence of IspE protein from *Aquifex aeolicus*

The catalytic activity of *A. aeolicus* IspE was determined using the photometric assay in buffer containing 50 mM citrate, 50 mM HEPES, 50 mM Tris/HCl and 50 mM boric acid in a total volume of 200 µl. The pH was adjusted to values of 5 - 10 with HCl or NaOH. The assay mixture was prepared and measured according to the Methods 3.3.6.2.1.

The result of this study is shown in Figure 47 as a curve of relative activity of IspE protein at different pH values. The maximum activity is found at pH 8.5. Based on this study, the buffer containing 100 mM Tris/HCl, pH 8.5 was then always used for the assay of *A. aeolicus* IspE.

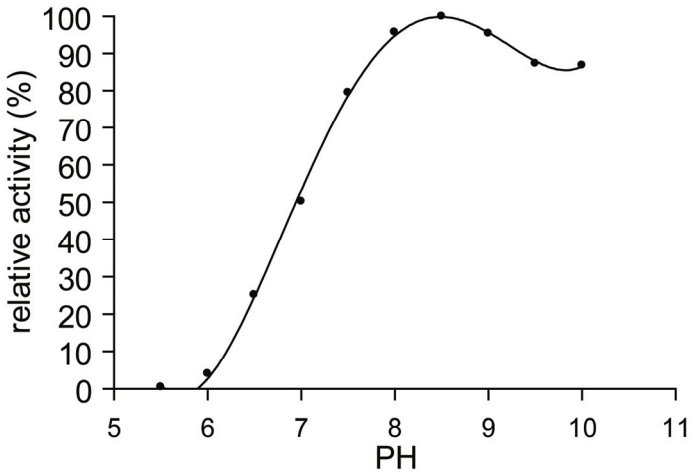


Figure 47: PH dependence of *Aquifex aeolicus* IspE

4.2.1.5 Temperature dependence of IspE protein from *Aquifex aeolicus*

The specific activities of *Aquifex aeolicus* IspE at various temperatures were calculated using ^{13}C NMR spectroscopy. Assay mixtures were prepared as described under Methods 3.3.6.2.2 and incubated at temperature ranging from 20 – 80 °C. The reaction was terminated by the addition of EDTA to a final concentration of 30 mM before ^{13}C NMR measurement.

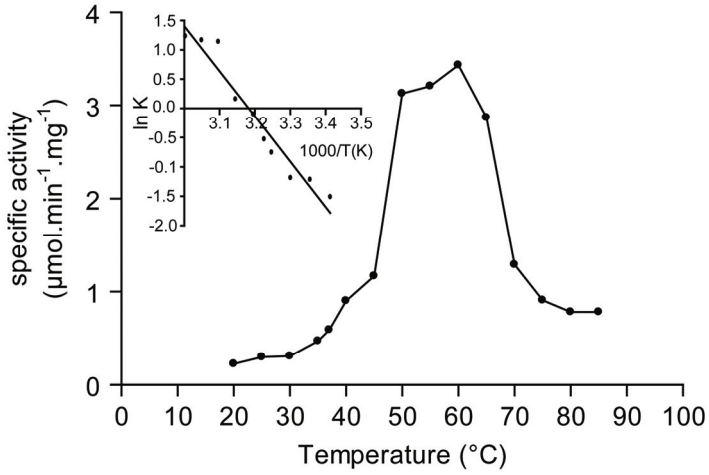


Figure 48: Temperature dependence of *Aquifex aeolicus* IspE with inserted Arrhenius plot

The result is depicted in Figure 48 as a curve of IspE activity at different temperature. As can be seen, the *A. aeolicus* IspE reached maximum activity at 60 °C. The Arrhenius plot was depicted accordingly, affording an activation energy 64.2 kJ.mol^{-1}.

4.2.1.6 CDP-ME and ATP dependence of IspE protein from *Aquifex aeolicus*

The measurements were performed using the photometric assay as described under Methods 3.3.6.2.1. For the measurement of ATP dependence, the concentration of ATP was varied from 10 to 1300 µM, while 4-diphosphocytidyl 2*C*-methyl-D-erythritol (CDP-ME) was fixed to a saturated concentration of 1 mM (see Figure 49A). On the other hand, for the measurement of CDP-ME dependence, the concentration of CDP-ME was varied from 10 – 1300 µM and measured with a fixed 1 mM concentration of ATP (see Figure 49B). The catalytic activities were measured at 37 °C.

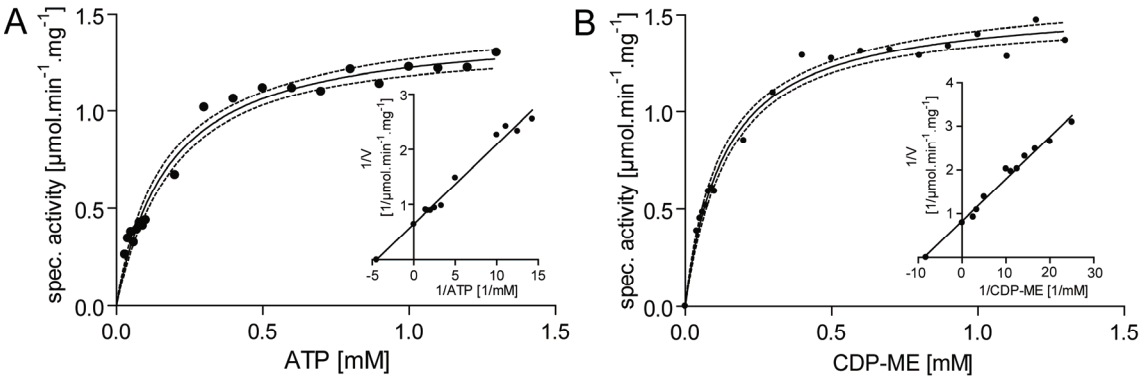


Figure 49: CDP-ME and ATP dependence of IspE protein from *Aquifex aeolicus* (A) Michaelis-Menten kinetics showing initial rate versus ATP concentration; (B) Michaelis-Menten kinetics showing initial rate versus 4-diphosphocytidyl-2*C*-methyl-D-erythritol concentration; insets, Liveweaver-Burk plots; dotted lines show 95 % confidentiality.

The results of this study are shown in Figure 49A and B. The curves show that the substrates dependence of *A. aeolicus* protein follows Michaelis-Menten kinetics the K_M for CDP-ME and ATP of 121 and 222 µM, respectively. The maximum catalytic activity (V_{max}) at 37 °C was 1.5 µmol min^{-1} mg^{-1}.

4.2.1.7 Kinetic parameters of IspE protein from *Aquifex aeolicus*

As a summary, the kinetic parameters of IspE *Aquifex aeolicus* are listed in Table 29 below.

Table 29 Kinetic parameters of IspE protein from *Aquifex aeolicus*

Kinetic parameters	Values
k_{cat} (CDP-ME → CDP-MEP)$_{37°C}$	0.75 s^{-1}
K_M (CDP-ME)	121 µM
K_M (ATP)	222 µM
Metal ion preference	Mn^{2+}, Mg^{2+}
PH optimum	8.5
T_{max}	60 °C
Activation energy	64.2 kJ.mol^{-1}

4.2.1.8 Discussion

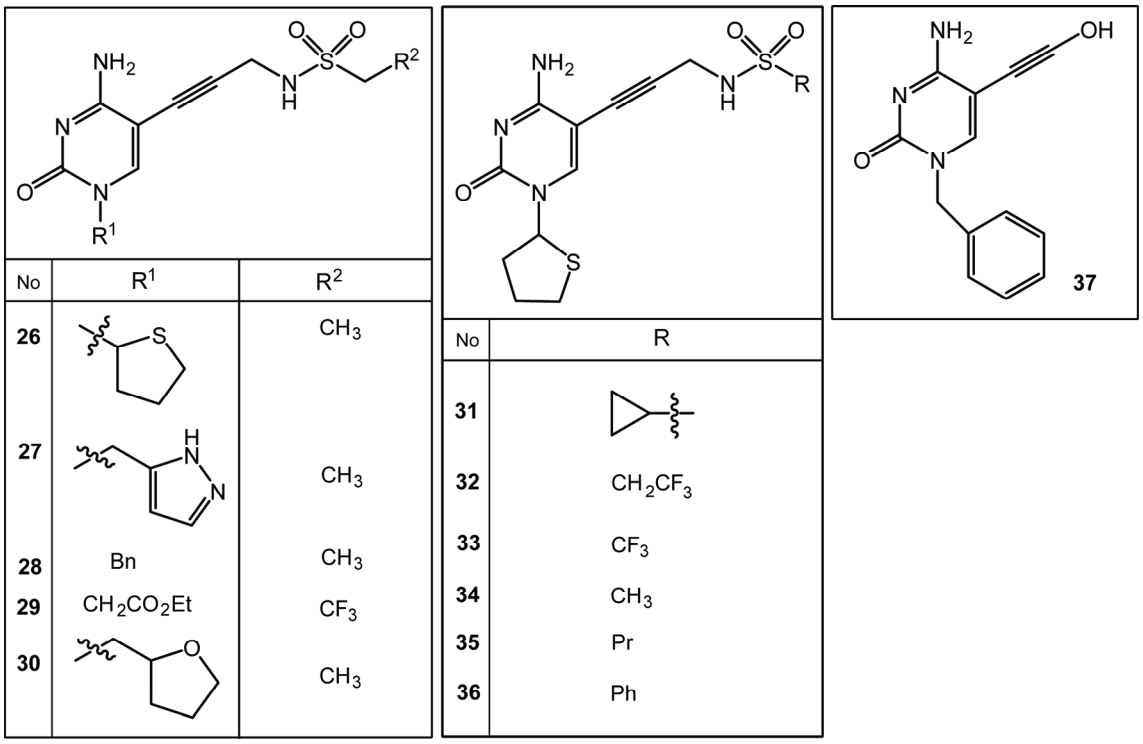
The characterization of IspE proteins from some organisms have been reported (Rohdich *et al*., 2000; Lüttgen *et al*., 1999; Bernal *et al*., 2004). The kinetic parameters from some published data are shown in Table 30. The sequence identities of IspE from *Aquifex aeolicus* are 30 % compared to *E. coli* and 13 % when compared to *Lycopersicon esculentum* show the high sequence conservation among the bacterial enzymes.

Table 30 Kinetic parameters of recombinant IspE proteins from different sources

Organism	Paper	Metal ion	V_{max} µmol min^{-1} mg^{-1}	K_M (ATP)	K_M(CDP-ME)
E. coli	Lüttgen *et al*., 2000	Mg^{2+}	33		
	Bernal et al., 2004	Mg^{2+}	1.6	420	150
Lycopersicon esculentum	Rohdich *et al*., 2000	Mg^{2+}	34		
Aquifex aeolicus	This study	Mn^{2+} and Mg^{2+}	1.5	222	121

4.2.2 Inhibition kinetics of the inhibitors of the reaction catalyzed by IspE protein

The X-ray crystal structure of the ternary complex (Miallau *et al*, 2003) of *Escherichia coli* protein shows IspE as a homodimer and the presence of three pockets at the active site: the adenosine, the cytidine, and the methyl erythritol phosphate pockets. A molecular modeling analysis was performed by the group of ETH Zürich using the program MOLOC that revealed the existence of an additional small hydrophobic pocket at the cytidine binding site. However, the structure-activity-relationship suggests that this hydrophobic pocket is not used by the substrate.



No	R^1	R^2
26		CH_3
27		CH_3
28	Bn	CH_3
29	CH_2CO_2Et	CF_3
30		CH_3

No	R
31	
32	CH_2CF_3
33	CF_3
34	CH_3
35	Pr
36	Ph

37

Figure 50: Potential inhibitors of the IspE protein
The inhibitors were synthesized by the group of ETH Zürich.

The potential inhibitors (Figure 50) were designed to occupy both the cytidine and hydrophobic subpocket as shown by the binding of representative inhibitor **31** in Figure 51.

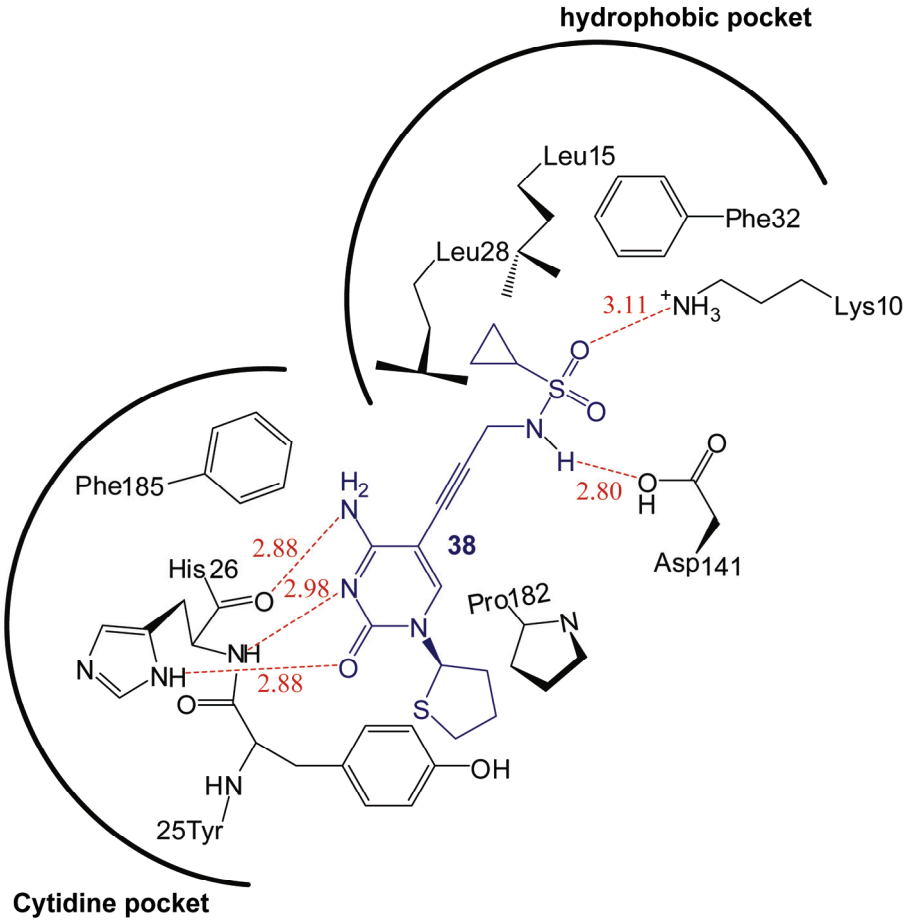


Figure 51: Schematic representation of the binding mode of inhibitor **31**
Potential H-bond is depicted in red dash; distance in Å.

To confirm the binding of the ligand in the IspF protein from *Escherichia coli*, the photometric assay was performed as described under Methods 3.3.6.2.1 in the presence of inhibitor. The IC_{50} and K_i were determined photometrically as described under Methods 3.3.6.4.1 and 3.3.6.4.2, respectively.

The result of this study is shown in Table 31. The ligands were shown to be potent inhibitors of *E. coli* IspE, yielding IC_{50} and dissociation constants of inhibitors to nanomolar range. The three best ligands **26**, **31** and **32** inhibit *E. coli* IspE with inhibition constants (K_i) of 640, 290 and 360 µM, respectively. The ^{13}C NMR measurements (see Methods 3.3.6.2.2) in the presence of the inhibitors were performed and the result (not shown) confirmed the inhibition of IspE protein.

Table 31: IC_{50}, mode of inhibition and K_i of *Escherichia coli* IspE inhibitors

Ligand	Mode	IC_{50} (µM)	K_i (µM)	K_{is} (µM)
26	Competitive	6 ± 0.1	0.64 ± 0.1	—
27	Competitive	19 ±0.1	1.6 ± 0.1	—
28	mixed	79 ± 5.2	3.7 ± 0.5	23.5± 7.1
29	mixed	398 ± 5.1	4.2 ± 0.6	21.6± 6.2
30	Competitive	71 ± 0.1	32.3 ± 2.8	—
31	Competitive	8 ± 0.1	0.29 ± 0.1	—
32	Competitive	6 ± 0.1	0.36 ± 0.1	—
33	Competitive	8 ± 0.2	1.2 ± 0.3	—
34	Competitive	22 ± 1	2.6 ± 0.1	—
35	mixed	48 ± 17	8.2 ± 1.7	—
36	Competitive	102 ± 15	16.3 ± 1.0	—
37	—	± 500 µM	—	—

K_i, competitive inhibition constant; K_{is}, Uncompetitive inhibition constant; mode inhibition and inhibition constant were estimated with program *dynafit* (Kuzmic, P, 1996).

4.3 IspF protein

4.3.1 Characterization of IspF protein from *Arabidopsis thaliana*

IspF protein, known also as 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase, catalyzes the conversion of 4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate into 2*C*-methyl-D-erythritol-2,4-cyclodiphosphate. The enzyme utilizes divalent metal ions such as Mn^{2+}, Mg^{2+}, Zn^{2+}, and Co^{2+}.

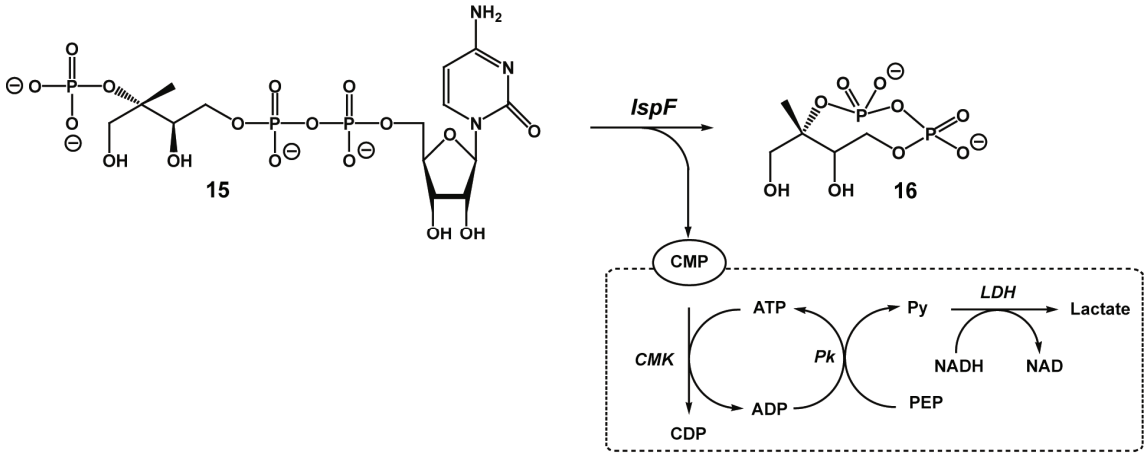


Figure 52: Photometric assay of the IspF reaction coupled with auxiliary enzymes (dashed box)
The enzyme converts 4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate (**15**) into 2*C*-methyl-D-erythritol-2,4-cyclodiphosphate (**16**).

The reaction can be observed *via* ^{13}C NMR spectroscopy using [1,3,4-$^{13}C_3$]4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate (see Herz *et al*., 1999). Alternatively, an enzyme-coupled assay with auxiliary proteins (see Illarionova *et al*., 2006) allows the reaction to be monitored spectrophotometrically at 340 nm (Figure 52). The spectrophotometric measurement has the advantage that the detection of the catalytical reaction is more sensitive, allowing to measure the initial activity of the protein with a smaller concentration of enzyme and substrate whereby the ^{13}C NMR spectroscopy provides a straight forward detection of substrate and product. Both assays are adapted from papers published by Herz *et al*., 2001 and Illarionova *et al*., 2006.

4.3.1.1 Cloning and expression of the *ispF* gene from *Arabidopsis thaliana*

Arabidopsis thaliana ispF gene was cloned by Johannes Kaiser (TUM, 2004). Information about the recombinant protein is summarized in Table 32.

Table 32: Cloning of the *ispF* gene from *Arabidopsis thaliana*

Accession number	NP-564819
Vector	pQE30 vector (N-terminal His-tag)
Host	*E. coli*, M15[pREP4]

The recombinant protein used in this study was produced as a synthetic construct adapted to *E. coli* codon usage. The construct was initiated by MRGSHHHHHHGS as N-terminal His-tag followed by amino acids number 76 - 223. The amino acid arginine/R 76 is the starting point of the sequence homology (cf. Figure 53) and was therefore used as the starting amino acid residue in the construct of this study.

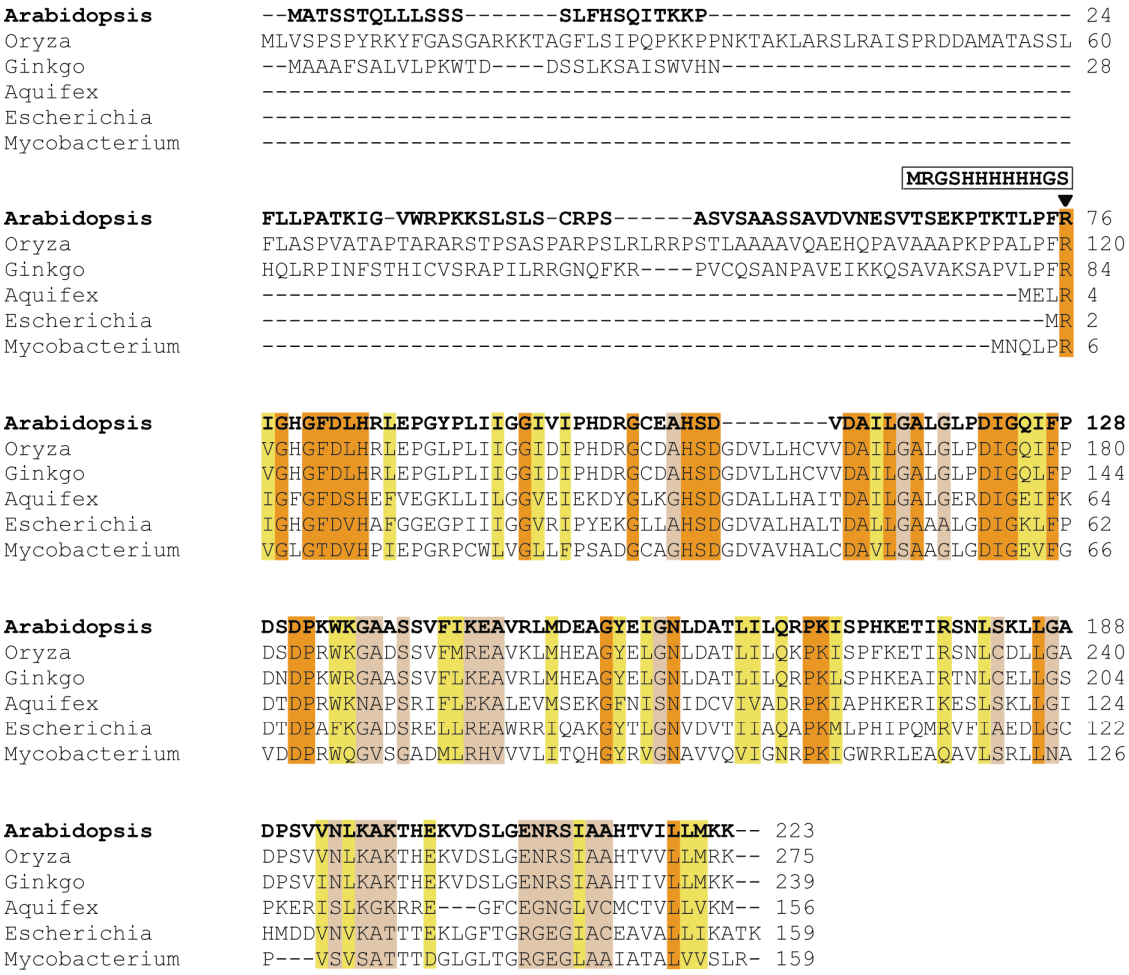


Figure 53: Sequence aligment of IspF proteins from some plants and bacterias. The amino acid sequence of *Arabidopsis thaliana* IspF was shown in bold. ▼ shows the initial amino acid in the *Arabidopsis thaliana* IspF clone used in this study, and the letters in the box refer to the additional amino acid due to the N-terminal Histag attachment. Orange color refers to identical residues in all organisms, yellow color refers to conserved substitutions, brown color refers to semi-conserved substitutions.

Sequence comparison showed high degree of homology among the plant species (55 % identity), whereas the homology to the bacterial orthologs is significantly lower (18 % identity). The enzyme activity of the crude extract was determined to 951 nmol min^{-1} mg^{-1} and the expression test showed a high expression level of IspF protein.(see Figure 54C).

4.3.1.2 Purification of recombinant IspF protein from *Arabidopsis thaliana*

Recombinant IspF protein from *A. thaliana* was purified by Nickel-chelating column chromatography as described under Methods (3.3.2.2.1) followed by dialysis against buffer containing 50 mM Tris hydrochloride, pH 8.0, 0.05 M potassium chloride, 1 mM DTT and 0.02 % NaN_3. This procedure afforded highly purified protein (Figure 54D) with a specific enzyme activity of 3.6 µmol min^{-1} mg^{-1}. The yield of the 3.8 fold enriched protein was 79 % (Table 33).

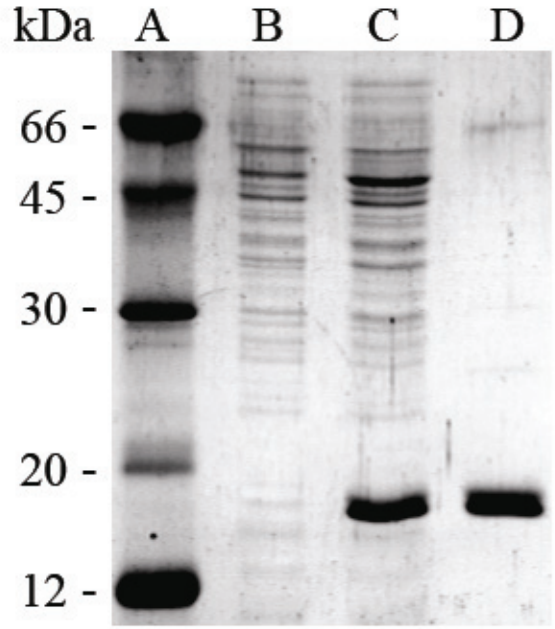


Figure 54: SDS-PAGE of the purification of recombinant IspF protein from *Arabidopsis thaliana*
Note: A, marker; B, wild-type crude extract; C, recombinant crude extract; D, purified IspF protein.

Table 33: Purification of IspF protein from *Arabidopsis thaliana*.

Procedure	Total protein (mg)	Specific activity (µmol min^{-1} mg^{-1})	Total activity (µmol min^{-1})	Yield	Purification factor
Cell extract	1910	0.951	1817	100	1
Ni^{2+} sepharose FF	399	3.59	1432.41	79	3.8

Partial Edman degradation afforded the N-terminal sequence **MRGSHHHHHH GSRIGHGFDLHRLEP** in agreement with the construct as shown in Figure 53. MALDI TOF analysis of the recombinant protein gave a value of 18,263 Da in good agreement with the theoretical value of 18,261.47 Da.

4.3.1.3 Metal dependence of IspF protein from *Arabidopsis thaliana*

The dependence of IspF enzyme activity on divalent metal ions was tested using the NMR assay described in 3.3.6.3.2.1. The reaction was started by the addition of the metal ion in a final concentration of 5 mM. Relative catalytic rates of 100 %, 38.5 % and 29.7 % were found with Co^{2+}, Mg^{2+} and Mn^{2+} ions, respectively. The addition of 1 mM EDTA yielded a significant decrease of enzyme activity that could not be restored when Mg^{2+} and Mn^{2+}, respectively were used in the assay. However, when Co^{2+} was used in the assay no effect was measured when EDTA was added to the assay mixture. Interestingly, the addition of Zn^{2+} can significantly increase enzyme activity when Mg^{2+} together with EDTA were used in the assays, whereas the enzyme activity was quiet low when Zn^{2+} was tested alone.

Table 34: Relative activity of *Arabidopsis thaliana* IspF protein in the presence of metal ions and 1 mM EDTA

Metal ion	Relative activity (%)
Co^{2+}	100
Mg^{2+} [a]	33.8
Fe^{2+}	11
Mn^{2+}	4.6
Ni^{2+}	4.2
Zn^{2+}	2.2
Ca^{2+}	0.8
Cu^{2+}	0
Mg^{2+}	0

The concentration of metal ions and EDTA was 5 mM and 1 mM, respectively.

[a]Addition of 1 mM Zn^{2+}

A metal dependence measurement in the absence of EDTA showed the highest activation of IspF protein by the addition of Co^{2+}, followed by Mn^{2+} and Mg^{2+} with maximum activities of 13.6, 5.7 and 3.5 µmol min^{-1} mg^{-1}, respectively. The optimum concentrations for Co^{2+}, Mg^{2+} and Mn^{2+} were 3, 2 and 0.2 mM, respectively (Figure 55).

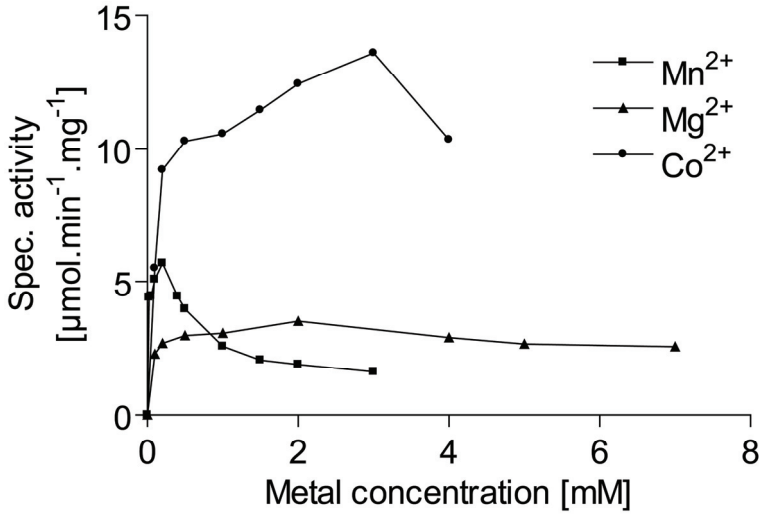


Figure 55: Co^{2+}, Mn^{2+} and Mg^{2+} dependence of IspF protein from *Arabidopsis thaliana*

4.3.1.4 PH dependence of IspF protein from *Arabidopsis thaliana*

The catalytic activity of *A. thaliana* IspF protein was measured at various pH values in the range of 5 – 10, in a buffer containing 50 mM glycine, 20 mM potassium phosphate, 50 mM sodium acetate, 50 mM Tris/HCl, 0.2 mM Mn^{2+}, 10 % (v/v) D_2O, 4 diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate and *A. thaliana* IspF as described under Methods 3.3.6.3.2.1. The assays were analyzed by ^{13}C NMR spectroscopy.

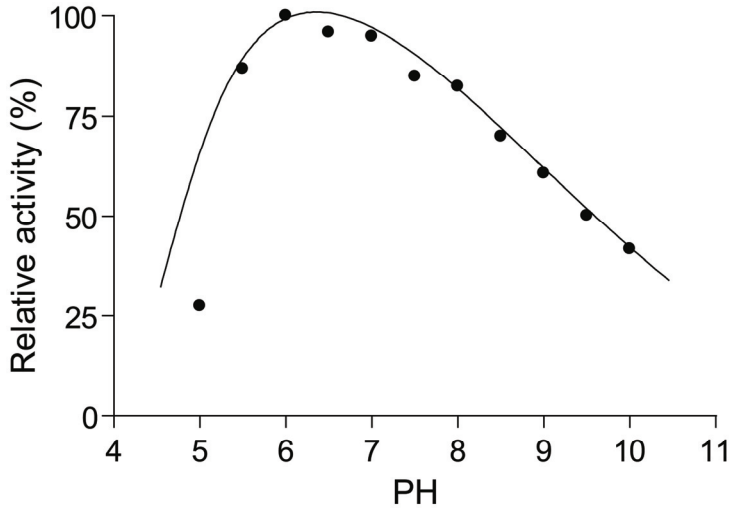


Figure 56: PH dependence of the catalytic activity of IspF protein from *Arabidopsis thaliana*

The result of this study is depicted in Figure 56 as a curve of relative activity of IspF protein at different pH values. As can be seen, maximal activity is found at pH 6.0.

4.3.1.5 Temperature dependence of IspF protein from *Arabidopsis thaliana*

The assay contained 100 mM MES pH 6.0, 0.2 mM Mn^{2+}, 10 % (v/v) D_2O, 4 diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate and IspF protein *from Arabidopsis thaliana.* Measurements were performed by ^{13}C NMR spectroscopy after incubation at different temperatures of 10 - 80 °C as described under Methods (3.3.6.3.2.1). Figure 57 shows the specific enzyme activity of the *Arabidopsis thaliana* IspF protein at different temperatures. It was found that the temperature optimum is at 65 °C. The activation energy was calculated to 44.5 kJ mol^{-1}.

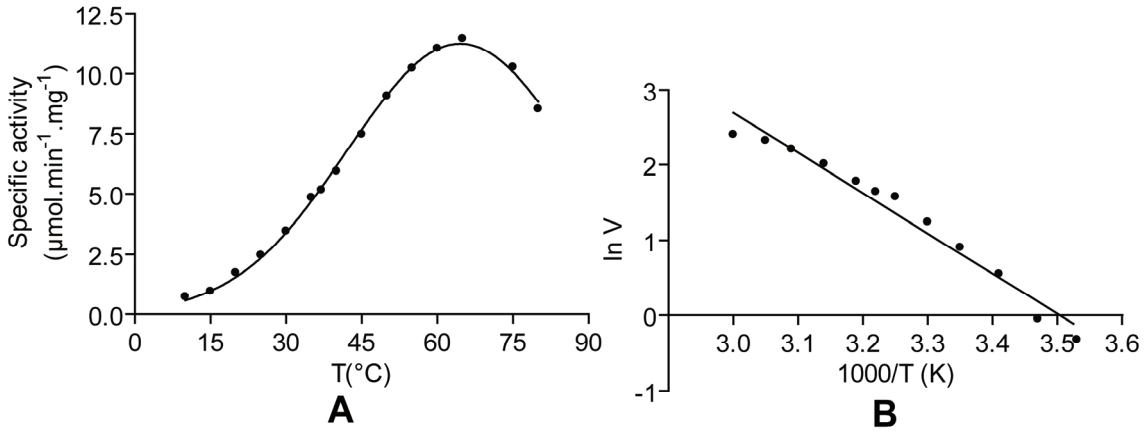


Figure 57: Temperature dependence of *Arabidopsis thaliana* IspF (A) and the Arrhenius plot (B)

4.3.1.6 Substrate dependence of the catalytic activity of the IspF protein from *Arabidopsis thaliana*

Measurements were performed by the photometric coupled enzyme assay as described under Methods 3.3.6.3.1 using 2 mM Mn^{2+}. The reaction was monitored with a spectrophotometer at 340 nm.

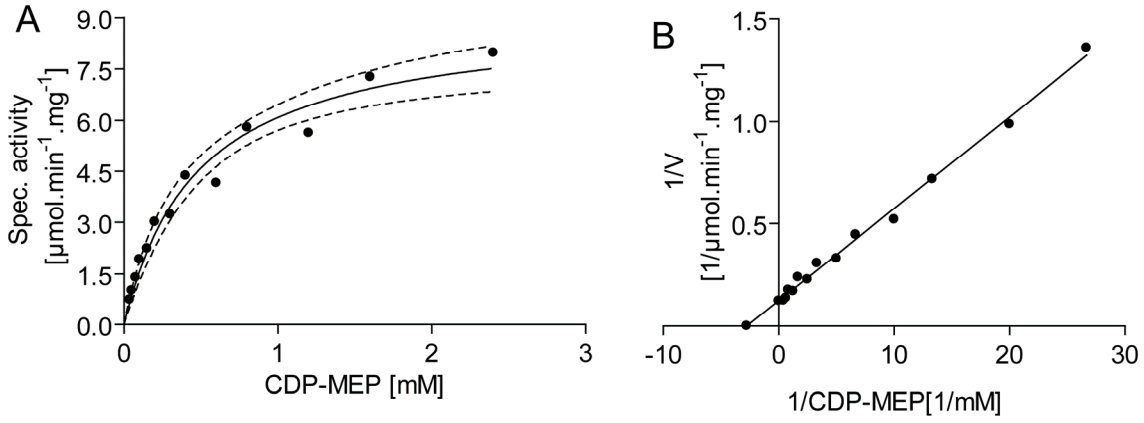


Figure 58: CDP-MEP dependence of IspF protein from *A. thaliana*
A, Michaelis-Menten curve and B, Lineweaver-Burk plot were measured at varied concentration of CDP-MEP.

The Michaelis-Menten and Lineweaver-Burk plot, respectively are shown in Figure 58. The K_M and V_{max} values for the substrate CDP-ME were 483 µM and 8.1 µmol min^{-1} mg^{-1}, respectively.

4.3.1.7 NMR forward and reverse reaction

The assay mixture of IspF forward reaction contained 100 mM Tris hydrochloride, pH 8.0, 10 % (v/v) D_2O, 2 mM $MnCl_2$, 1 mM [1,3,4-$^{13}C_3$]4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate and IspF protein in a total volume of 500 µl (see Methods 3.3.6.1.3.1). After incubation at 37°C, followed by ^{13}C NMR analysis, the assay showed about 60 % conversion from [1,3,4-$^{13}C_3$]4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate (Figure 59A) to [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol-2,4-cyclodiphosphate as shown in Figure 59C.

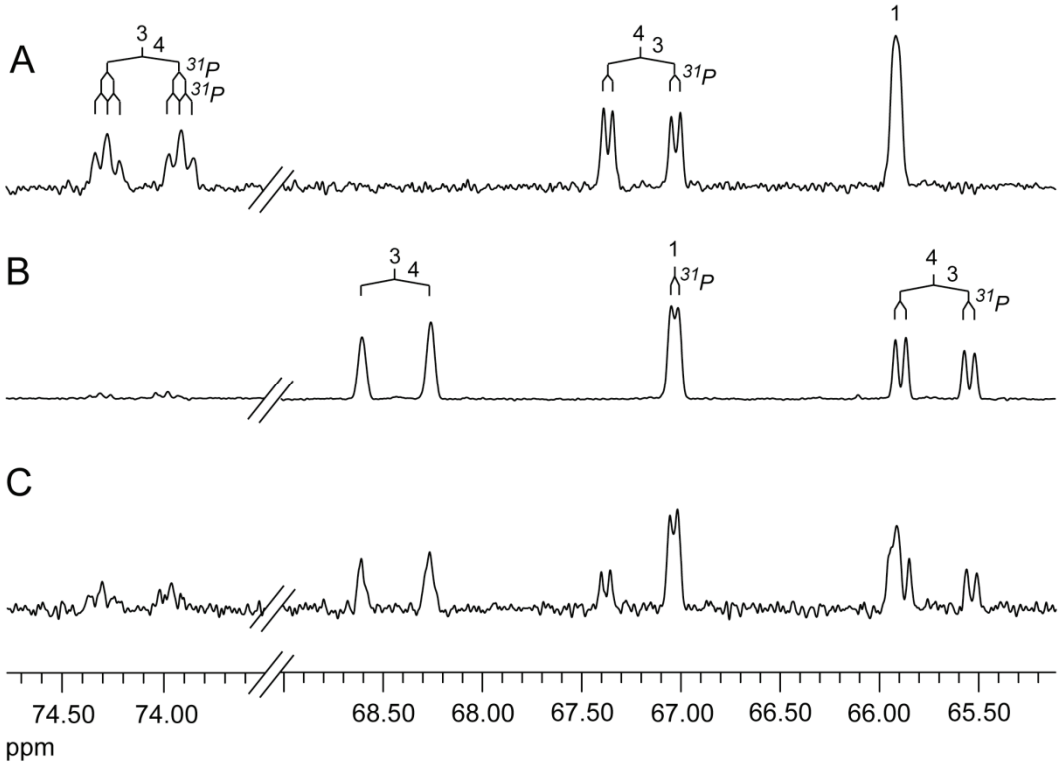


Figure 59: ^{13}C NMR spectra of the reaction catalyzed by *Arabidopsis thaliana* IspF [1,3,4-$^{13}C_3$]4-Diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate (A), [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol-2,4-cyclodiphosphate (B) and the mixture of [1,3,4-$^{13}C_3$]4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate and [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol-2,4-cyclodiphosphat (C).

To test the possible reverse reaction of IspF protein, the ^{13}C NMR analysis was implemented. The assay mixture was prepared according to the Methods 3.3.6.3.2.2. After 1 h incubation and ^{13}C NMR measurement, no signals from [1,3,4-$^{13}C_3$]4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate were generated from [1,3,4-

$^{13}C_3$]2*C*-methyl-D-erythritol-2,4-cyclodiphosphate. According to this finding, the IspF reaction appeared to catalyze an irreversible reaction.

4.3.1.8 Kinetic parameters of IspF protein from *Arabidopsis thaliana*

The overall of kinetic parameters of IspF *Arabidopsis thaliana* is shown in Table 35.

Table 35: Kinetic parameters of recombinant IspF protein from *Arabidopsis thaliana*

Parameter	Value
k_{cat} (CDP-MEP → CYCLO)$_{Mn,\ 37\ °C}$	2.47 s^{-1}
K_M *(* CDP-MEP*)*	483 µM
pH optimum	6.0
T_{max}	65°C
Activation energy	44.5 kJ mol^{-1}
Metal ion preference	Co^{2+}, Mn^{2+}, Mg^{2+}
Reaction	Irreversible

4.3.1.9 Discussion

The heterologous expression of *ispF* genes from *A. thaliana* has been reported earlier (Calisto *et al.*, 2007) in order to obtain the three dimensional structure. However, the kinetics study of *A. thaliana* IspF was not done before. Here we report the characterization of IspF protein from *A. thaliana*, affording K_M of substrate 4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate and V_{max} of 483 µM and 8.1 µmol min^{-1} mg^{-1} that are similar to the K_M and V_{max} reported earlier for *P. falciparum* IspF of 252 µM and 4.3 µmol min^{-1} mg^{-1}, respectively. The metal ions are necessary for the activity of both recombinant proteins. Specifically, Mn^{2+} and Mg^{2+} are the metal of choice for *P. falciparum* IspF while for *A. thaliana* IspF, Co^{2+}, Mn^{2+} and Mg^{2+} are the metal of choice. The three dimensional structure of *A. thaliana* IspF has been revealed before (Calisto *et al*., 2007), confirming the relevance of the trimer as the biological assembly which is similar to *E. coli* protein (Richard *et al*.,2002; Steinbacher *et al*., 2001; Kemp *et al*., 2005). The importance of Zn^{2+} for the catalysis was remarked in the paper published earlier about three dimensional structure of *E. coli* protein (Steinbacher *et al*., 2001), illustrating that Zn^{2+} ion is bound at the active

site, helping to position the substrate for a direct attack of the 2-phosphate group on the β-phosphate. It was therefore considered as metal of choice for IspF protein. This finding was supported by a paper published recently about tetrahedral coordination of Zn^{2+} at the base of the active site pocket of *A. thaliana* IspF (Calisto *et al*., 2007). Based on this finding, although our experiment showed modest activity with zinc ion, it is still considered to be the most important metal for IspF protein under physiological condition.

Despite of the similarity in the coordination of zinc ion, the crystal structure of IspF proteins from *A. thaliana* and *E. coli* show a major difference in the large molecular cavity that forms between subunits and involve the residues that are higly conserved among the plants. In some bacterial enzymes, especially from *E. coli* (Kemp *et al*., 2004) and *M. smegmatis* (Buetow *et al*., 2007), the corresponding cavity has been shown to be the binding pocket of isoprenoid-like ligands such as isopentenyl diphosphate (and or dimethylallyl diphosphate), geranyl diphosphate and farnesyl diphosphate that might be important in the feedback regulatory role (Kemp *et al*., 2004). This cavity is absent in the thermophilic IspF protein from *Thermus thermophilus* where packing of the subunit is tighter and more stable (Kishida *et al*., 2003).On the other hand, the modeled interaction of the structure of *A. thaliana* IspF showed that the electron density of the cavity is suited for the binding of monophosphate ion and discard the possibility for the binding of geranyl diphosphate and farnesyl diphosphate (Calisto *et al*., 2007).

However, the active site architecture is shown to be very well conserved among bacteria and plants. For this reason, the IspF protein becomes an attractive target for the discovery of herbicides and antibiotics. Notably, the depletion of IspF protein has been demonstrated to severely decelerate the growth of bacteria (Campbell and Brown, 2002).

4.3.2 Fluorescent Inhibitors of *Escherichia coli* IspF

The crystal structures from *E. coli* IspF have been reported earlier (Steinbacher *et al*., 2001; Richard *et al*, 2002; Kemp *et al*., 2002) as a C_3-symmetric homotrimer. The substrate binding sites are divided into three distinctive pockets: the first pocket that surrounds the ribosyl-5´-diphosphate group of the diphosphocytidyl moiety, pocket II

where the 2*C*-methyl-D-erythritol 2-phosphate moiety of substrate binds and pocket III that binds the cytidyl moiety. Pocket II and III are represented in Figure 60 by the interaction between the products CMP and 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate and IspF protein (Steinbacher *et al.*,2001). Additionaly, the central pocket (I) does not make a direct contact to the substrate.

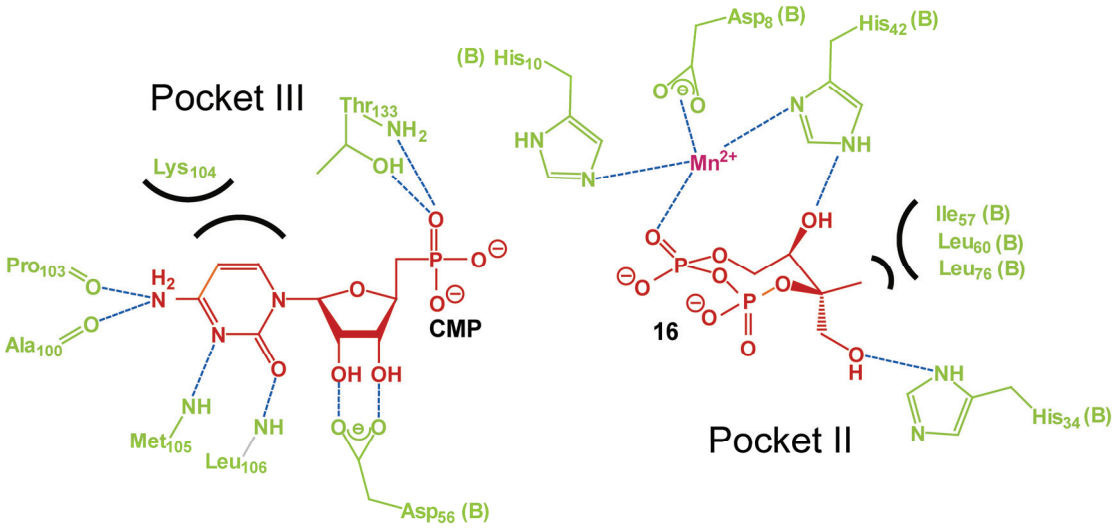


Figure 60: Substrate binding sites of the IspF protein from *Escherichia coli*

Based on the binding sites of IspF protein, three inhibitors that occupy pocket I and II were designed by molecular modeling software (MOLOC). These putative inhibitors were synthesized at the Laboratorium für Organische Chemie, ETH Hönggerberg, Zurich by the work from C. M. Crane and Prof. Diedrich.

4.3.2.1 The structure of inhibitors

The inhibitors **38**, **39** and **40** consist of CDP moieties of the natural substrate 4-diphosphocytidyl-2*C*-methyl-D-erythritol 2-phosphate (**15**) to occupy pocket III, diphosphate as the linker and aromatic residues to occupy pocket II as shown in Figure 61. To develop a fluorescence-based enzyme inhibition assay, the fluorescent anthranilate (2-amino-benzoate) and dansyl (5,5-dimethylaminonaphthalenesulfamoyl) residues were chosen as aromatic moieties.

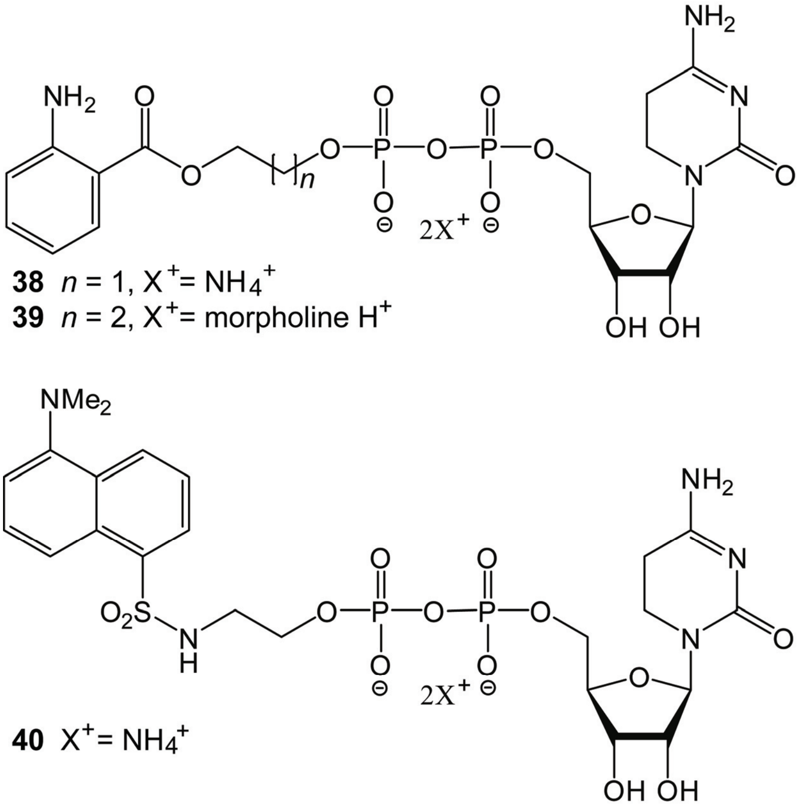


Figure 61: Structures of substrate-derived fluorescent inhibitors of the IspF protein

The dissociation constants of the complex IspF from *E. coli* and ligands **38**, **39** and **40** at 25°C were determined as 36, 23 and 15 µM.

4.3.2.2 Purification of *Escherichia coli* IspF from inclusion body

IspF protein was renaturated and then purified from the inclusion bodies (pellet) according to the aforementioned Methods 3.3.2.2.2. The pellet was obtained from the disruption of 1.9 g cell of *E. coli* strain M15pRep4PQE30ispF E.coli2 and stirred overnight with 15 ml of 8 M urea. The mixture was then centrifuged and the supernatant was taken as crude extract for purification and loaded into a 6 ml volume Nickel-chelated column. After renaturation followed by purification and dialysis, 3 mg of pure IspF protein was obtained with the specific activity of 11.4 µmol min^{-1} mg^{-1}.

4.3.2.3 Inhibition test

The preliminary assay was analyzed by ^{13}C NMR measurement from a mixture containing 50 mM Tris/HCl pH 8.0, 4 mM [1,2-$^{13}C_2$]acetate (as standard), 25 mM [1,3,4-$^{13}C_3$]4-Diphosphocytidyl-2*C*-methyl-D-erythritol-2-phosphate, 10 mM $MgCl_2$,

2.79 µg *Escherichia coli* IspF (renaturated from Inclusion body) and 60 µl D_2O in total volume of 600 µl at 37 ° C.

The signal of the product [1,3,4-$^{13}C_3$]2*C*-methyl-D-erythritol 2,4-cyclodiphosphate was integrated on the basis of acetate standard. Product formation was calculated every 13 min and depicted as curves in Figure 62.

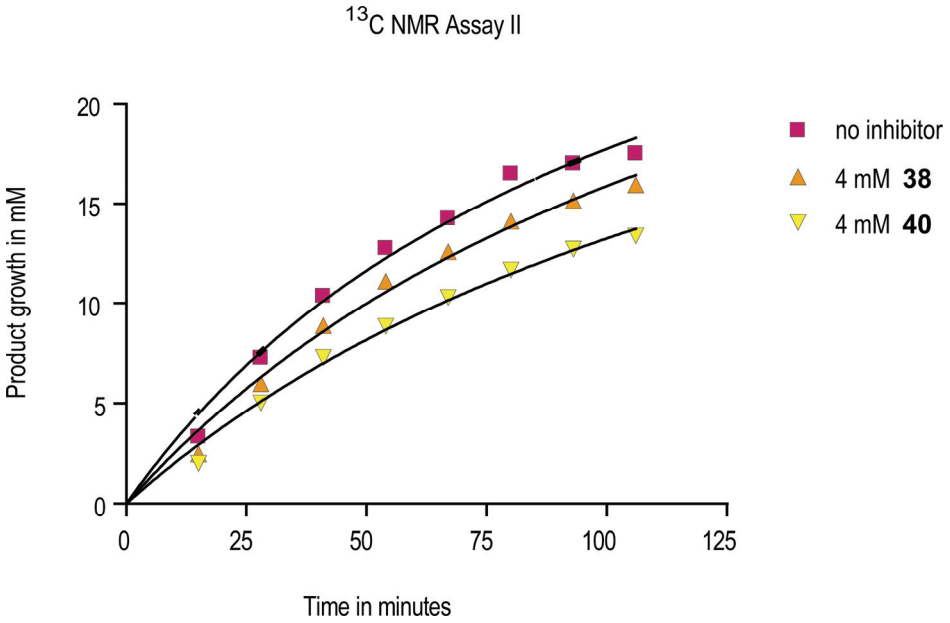


Figure 62: ^{13}C NMR assay of *Escherichia coli* IspF protein in the absence and in the presence of 4 mM of **38** and **40**, respectively

As a result, the calculated residual activities obtained from ^{13}C NMR assay containing inhibitor **38** and inhibitor **40** were 91.2 % and 76.5 % respectively, when compared to the assay without inhibitor.

The IC_{50} analysis of inhibitor **40** was performed in comparison with CDP and CMP using ^{13}C NMR spectroscopy. The curves of relative activity of IspF protein in the presence of inhibitors are depicted in Figure 63. The IC_{50}s obtained from the curves were 3 mM, 7.3 mM and 15 mM for inhibitor **40**, CDP and CMP respectively. The crystal structure of *E. coli* IspF in complex with inhibitor **40** confirmed that the cytidine moiety of **40** binds tightly into the pocket III (Crane *et al*., 2006).

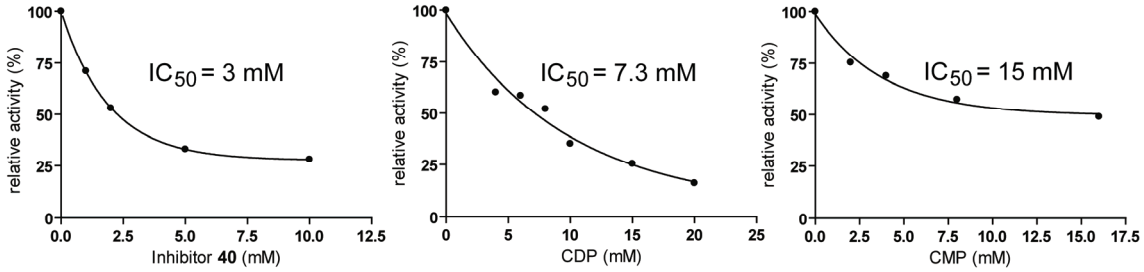


Figure 63: Inhibition of the reaction catalyzed by *Escherichia coli* IspF protein by inhibitor **40**, CDP and CMP, respectively

4.3.3 Non-fluorescent inhibitors of IspF protein

The screening of a compound library showed the clusters of inhibition effects originated from some chemical building blocks (the complete result not shown here). Two strongest inhibitors **S41** and **S42** (structures not shown) originated from different building blocks were further investigated for IC_{50} and inhibition constant (K_i) in the reaction containing IspF protein from various organisms.

Spectrophotometric assay (as described in the Methods 3.3.6.3.2) was carried out in the presence of 100 mM inhibitors and IspF proteins from *A. thaliana*, *P. falciparum* or *Mycobacterium tuberculosis*, respectively. The IC_{50} measurements were performed using both ^{13}C NMR spectroscopy and spectrophotometry, affording a slightly different result as shown in Table 36. This discrepancy is probably caused by different protein concentration and the different nature of both assays.

Photometric assay measures the initial velocity of protein while NMR assay only measures the average velocity. Despite of its less sensitive measurement, the ^{13}C NMR analysis has the advantage of direct detection of substrate and product through their ^{13}C signals, hence it is always recommended to confirm of inhibitory activity that has been previously shown *via* photometric assay.

Table 36: Inhibition kinetics of IspF protein

IspF from different organisms	Inh code	Inhibition assay				
		IC_{50} (µM)		K_i measurement		
		Photometric	NMR	Mode	K_i	Error (%)
Arabidopsis thaliana	**S41**	63	50	Noncompetitive	K_{is}=2.9±0.55	19
	S42	16	42	Mixed	K_i=3.6±1.8 K_{is}=20.4±6.5	50 32
Plasmodium falciparum	**S41**	49	36	Mixed	K_i=2.2±0.59 K_{is}=4.2±1.3	27 31
	S42	24	50	Noncompetitive	K_{is}=38.7±6	15
Mycobacterium tuberculosis	**S41**	47		Noncompetitive	K_{is}=4±2.2	55
	S42	9		Noncompetitive	K_{is}=14±1.5	11
Escherichia coli	**S41**	undetectable		No inhibition		
	S42	82		Noncompetitive	K_{is}=80.2±19	24

The result of this study (Table 36) shows that inhibitors **S41** and **S42** inhibit IspF proteins from *A. thaliana*, *P. falciparum* and *M. tuberculosis* relatively stronger than from *E. coli* IspF. Specifically inhibitor **S41** did not show any inhibition and inhibitor **S42** showed significantly lower K_i when compared to *A. thaliana* IspF while the inhibition of recombinant IspF from *P. falciparum* and *E. coli* showed relatively similar result when compared to *A. thaliana* IspF. The dissimilarity in the inhibition potencies can not be explained by the functional homologues of the amino acid sequence, notably the sequence identities of IspF proteins from *A. thaliana – P. falciparum*, *A-thaliana – M. tuberculosis* and *A. thaliana – E. coli* were calculated as 34 %, 37 % and 50 %, respectively.

From the data (Table 36) it can be seen that the modes of inhibition of both inhibitors in different organisms are either noncompetitive or mixed type, indicating that the inhibitors might not bind at the active site of the protein but rather somewhere else at the protein structure that can induce the conformational change and lower the binding affinity of the substrate. Moreover, the structure of the inhibitors **S41** and **S42** do not show any similarity with 4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate (**15**) (see Figure 52). The mode of inhibition of the inhibitors might explain

why the inhibition potencies of both inhibitors are significantly lower for *E. coli* IspF when compared to *A. thaliana* IspF, despite of the high sequence similarity. The sequence identity can be related to the active site architecture of the protein and the binding sites of substrate and cofactor. The fact that the inhibitors might not bind at the active site of the protein makes the inhibition potencies not related to the degree of sequence conservation.

5 Summary

Terpenoids are one of the largest groups of natural products with a wide diversity in biological function and application (e.g. as food additives, cosmetics and therapeutic agents). Terpenoid biosynthesis is known to proceed *via* two independent pathways, i.e. the mevalonate pathway and the deoxyxylulose phosphate pathway. Many pathogenic bacteria use the deoxyxylulose phosphate pathway whereas animals use the mevalonate pathway. Plants use the mevalonate pathway in the cytosol and the deoxyxylulose phosphate pathway in their plastids.

This study covers the characterization of 2*C*-methyl-D-erythritol 4-phosphate synthase (IspC), 4-diphosphocytidyl-2*C*-methyl-D-erythritol kinase (IspE) and 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF) involved in the deoxyxylulose phosphate pathway. The recombinant proteins served as basis for the development of inhibition assays designed to discover inhibitors *via* High Throughput Screening and to investigate the mode of inhibition. Additionally, the reaction mechanism of IspC protein was studied with the aim to elucidate the binding of substrates, intermediates and inhibitors.

Arabidopsis thaliana IspC utilizes NADPH and 1-deoxy-D-xylulose 5-phosphate as substrates in the forward reaction and converts them into $NADP^+$ and 2*C*-methyl-D-erythritol 4-phosphate with K_M values of 30 and 132 µM, respectively; the k_{cat} value is 4.4 s^{-1}. NADH could replace NADPH with a k_{cat} value of 0.63 s^{-1}. In the reverse reaction, *Arabidopsis thaliana* IspC uses $NADP^+$ and 2*C*-methyl-D-erythritol 4-phosphate as substrates with K_M values of 471 and 972 µM, respectively; the k_{cat} value is 1.6 s^{-1}. *Arabidopsis thaliana* IspC showed maximum catalytic activity at 37 °C and at a pH value of 7.5 in the presence of 3 mM Mn^{2+} or Mg^{2+}.

Plasmodium falciparum IspC uses NADPH and 1-deoxy-D-xylulose 5-phosphate as substrates in the forward reaction with K_M values of 136 and 281 µM, respectively; the k_{cat} value is 3.5 s^{-1}. NADH could replace NADPH with a k_{cat} value of 0.99 s^{-1}. In the reverse reaction, *Plasmodium falciparum* IspC uses $NADP^+$ and 2*C*-methyl-D-erythritol 4-phosphate as substrates with K_M values of 475 and 260 µM,

respectively; the k_{cat} is 1.3 s^{-1}. The reaction catalyzed by *Plasmodium falciparum* IspC showed maximum catalytic activity at 60 °C and at a pH value of 7.5 in the presence of 2 mM Mn^{2+} or Mg^{2+}.

A mechanistic study of the IspC reaction was performed using IspC proteins from *Escherichia coli*, *Mycobacterium tuberculosis* and *Arabidopsis thaliana* in order to rule out one of two potential reaction mechanisms, i.e., a sigmatropic and a retroaldol mechanism (see Figure 27). NMR analysis of the IspC reaction using various isotopologues of 1-deoxy-D-xylulose 5-phosphate and 2*C*-methyl-D-erythritol 4-phosphate did not support the retroaldol mechanism that has been proposed in the literature. However, the products arising in the reaction assays could be explained by a sigmatropic rearrangement.

The crystallographic data of IspC proteins from *Escherichia coli* and *Mycobacterium tuberculosis* indicated the existence of a flexible loop (see Figure 43) that closes the active site during the catalysis and therefore prevents solvent access (Reuter *et al*., 2001; Steinbacher *et al*., 2003; Mac Sweeney *et al*., 2005; Yajima *et al*., Henriksson *et al*., 2007). This loop seems to open the active site only for its substrates and analogues thereof, i.e., analogues of 1-deoxy-D-xylulose 5-phosphate and 2*C*-methyl-D-erythritol 4-phosphate. These analogues have been demonstrated to act either as substrate alternates or inhibitors of the reaction. Notably, the substrate alternates must possess hydroxy groups at both the C-3 and C-4 positions.

4-Diphosphocytidyl 2*C*-methyl-D-erythritol kinase (IspE) catalyzes the conversion of 4-diphosphocytidyl 2*C*-methyl-D-erythritol (CDP-ME) into 4-diphosphocytidyl 2*C*-methyl-D-erythritol 2-phosphate (CDP-MEP). The *Aquifex aeolicus* IspE was studied *via* ^{13}C NMR spectroscopic and photometric assays. These studies afford a k_{cat} value of 0.75 s^{-1} and K_M values of 121 and 222 µM for CDP-ME and ATP, respectively. Maximum activity was observed at 60 °C, pH 8.5 and 2 mM of either Mn^{2+} or Mg^{2+}. Using these assays, it was shown that compounds containing sulfonamide moieties inhibit IspE enzyme from *E. coli* with K_i values in the nanomolar range (Figure 50 and Table 31).

IspF protein from *Arabidopsis thaliana* catalyzes the conversion of 4-diphosphocytidyl 2*C*-methyl-D-erythitol 2-phosphate (CDP-MEP) into 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate (cMEPP) with a k_{cat} value of 2.5 s^{-1} and a K_M value of 483 µM.

The maximum activity was observed at 65 °C, at a pH value of 6.0 and 3 mM Co^{2+}. Fluorescent derivatives of the substrates were tested according to their inhibition activity against IspF protein of *E. coli*. The IC_{50} measurements were performed for the most potent inhibitor (**40**) in comparison with CDP and CMP, affording IC_{50} values of 3.0, 7.3 and 15.0 mM, respectively (Figure 61 and Figure 63). High throughput screening was performed for potential inhibitors of *Arabidopsis thaliana* IspF, affording two inhibitors (**S41** and **S42**) with K_i values of 2.9 and 3.6 µM (Table 36).

References

Altincicek B, Hintz M, Sanderbrand S, Wiesner J, Beck E & Jomaa H. 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* 2000, 190, 329–333.

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

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

Arigoni D, Sagner S, Latzel C, Eisenreich W, Bacher A and Zenk MH.Terpenoid biosynthesis from 1-deoxy-D-xylulose in higher plants by intramolecular skeletal rearrangement. *Proc Natl Acad Sci* USA 1997, 94, 10600–10605.

Arigoni D. and Schwarz MK. Gingkolide biosynthesis. In: *Comprehensive Natural Product Chemistry* 1999, vol 2, pp. 45–68, Cane D. (ed.), Pergamon, Oxford.

Bach T.J. Some new aspects of isoprenoids biosynthesis in plants. *Lipids*, 1995, 30, 191–202.

Barton GJ. ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng* 1993, 6, 37-40.

Biou V, Dumas R, Cohen-Addad C, Douce R, Job D and Pebay-Peyroula E. The crystal structure of plant acetohydroxy acid isomeroreductase complexed with NADPH, two magnesium ions and a herbicidal transition state analog determined at 1.65 A resolution. *Embo J* 1997, 16, 3405-3415.

Bloch K. Sterol molecule: structure, biosynthesis, and function. *Steroids* 1992, 57, 378-383.

Bochar DA, Friesen, J. A., Stauffacher, C. V. , Rodwell, V. W. Biosynthesis of mevalonic acid from acetyl-CoA. In *Comprehensive natural product chemistry* (Cane, DE, ed) 1999, pp. 15-44. Pergamon, Oxford.

Borrmann S, Issifou S, Esser G, Adegnika AA, Ramharter M, Matsiegui PB, Oyakhirome S, Mawili-Mboumba DP, Missinou MA, Kun JF, Jomaa H, Kremsner PG. Fosmidomycin-clindamycin for the treatment of Plasmodium falciparum malaria. *J. Infect. Dis*. 2004, 190, 1534-1540

Borrmann S, Lundgren I, Oyakhirome S, Impouma B, Matsiegui PB, Adegnika AA, Issifou S, Kun JF, Hutchinson D, Wiesner J, Jomaa H & Kremsner PG. Fosmidomycin plus clindamycin for treatment of pediatric patients aged 1 to 14 years with Plasmodium falciparum malaria. *Antimicrob. Agents Chemother.*2006, 50, 2713-2718.

Broers STJ. Über die frühen Vorstufen der Biosynthese von Isoprenoiden in Escherichia coli, 1994, ETH Zürich.

Bullock WO, Fernandez JM and Short JM. XL1-Blue: a high efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. *Biotechniques* 1987, 5, 376-379.

Calisto B.M, Perez-Gil J, Bergua M, Querol-Audi J, Fita I, and Imperial S. Biosynthesis of isoprenoids in plants: Structure of the 2C-methyl-D-erithrytol 2,4-cyclodiphosphate synthase from *Arabidopsis thaliana*. Comparison with the bacterial enzymes. *Protein Science*2007, 16, 2082-2088.

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

Carretero-Paulet L, Ahumada I, Cunillera N, Rodrıíguez-Concepción M, Ferrer A, Boronat A and Campos N. 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* 2002, 129, 1581–1591.

Crane CM, Hirsch AKH, Alphey MS,Sgraja T, Lauw S, Illarionova V, Rohdich F, Eisenreich W, Hunter WN, Bacher A, and Diederich F. Synthesis and Characterization of Cytidine Derivatives that Inhibit the Kinase IspE of the Non-Mevalonate Pathway for Isoprenoid Biosynthesis. *ChemMedChem* 2008, 3, 91 – 10.

Crane CM, Kaiser J, Ramsden NL, Lauw S, Rohdich F, Eisenreich W, Hunter WN, Bacher A,and Diederich F. Fluorescent Inhibitors for IspF, an Enzyme in the Non-Mevalonate Pathway for Isoprenoid Biosynthesis and a Potential Target for Antimalarial Therapy. *Angew. Chem*. Int. Ed. 2006, 45, 1069 –1074.

David S, Estramareix B., Fischer JC. and Thérisod M. 1-Deoxy-D-threo-2-pentulose: the precursor of the five-carbon chain of the thiazole of thiamine. J. Am. Chem. Soc 1981, 103, 7341–7342.

David S, Estramareix B., Fischer JL. and Thérisod M. The biosynthesis of thiamine. Syntheses of [1,1,1,5- 2H4]-1-deoxy-D-threo-2-pentulose and incorporation of this sugar in biosynthesis of thiazole by *Escherichia coli* cells. *J. Chem. Soc* 1982, Perkin Trans. I, 2131–2137.

Dayhoff MO, Schwartz RM and Orcutt BC. A model of evolutionary change in proteins. In *Atlas of Protein Sequence and Structure* (Dayhoff MO, ed.) 1978, vol. 5, pp. 345–352.

Demming-Adams B. and Adams W. W. Antioxidants in photosynthesis and human nutrition. *Science* 2002, 298**,** 2149–2153 1420.

Dhiman RK, Schaeffer ML, Bailey AM, Testa CA, Scherman H and Crick DC. 1-Deoxy-D-xylulose 5- phosphate reductoisomerase (IspC) from Mycobacterium tuberculosis: towards understanding mycobacterial resistance to fosmidomycin. *J Bacteriol* 2005, 187, 8395–8402.

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

Eisenreich W, Menhard B, Hylands PJ, Zenk MH and Bacher A. Studies on the biosynthesis of taxol: the taxane carbon skeleton is not of mevalonoid origin. *Proc. Natl. Acad. Sci. U.S.A.* 1996, 93, 6431-6436.

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

Eisenreich W, Schwarz M, Cartayrade A, Arigoni D, Zenk MH, and Bacher A. The deoxyxylulose pathway of terpenoid biosynthesis in plants and microorganism. *Chemi.Biol* 1998, 5, R221-R233.

Felsenstein J. PHYLIP (Phylogeny Interference Package) 1995, Version 3.57c. Department of Genetics, University of Washington, Seattle.

Foth BJ and McFadden GI. The apicoplast: a plastid in Plasmodium falciparum and other Apicomplexan parasites. *Int. Rev. Cytol* 2003, 224 57-110.

Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF, McFadden GI. Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. *Science* 2003, 299, 705-708.

Fox DT and Poulter CD. Mechanistic studies with 2-C-methyl-d-erythritol 4-phosphate synthase from *Escherichia coli*. Biochemistry 2005, 44, 8360-8368.

Fox DT and Poulter CD. Synthesis and evaluation of 1-deoxy-D-xylulose 5-phosphoric acid analogues as alternate substrates for methylerythritol phosphate synthase. *J Org Chem* 2005, 70, 1978–1985.

Galtier N, Gouy M & Gautier C. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* 1996, 12, 543–548.

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

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

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

Hecht S, Kis K, Eisenreich W, Amslinger S, Wungsintaweekul J, Herz S, Rohdich F & Bacher A. Enzyme-assisted preparation of isotope-labeled 1-deoxyd- xylulose 5-phosphate. *J Org Chem* 2001, 66, 3948–3952.

Hecht S, Wungsintaweekul J, Rohdich F, Kis K, Radykewicz T, Schuhr CA, Eisenreich W, Richter G, Bacher A. Biosynthesis of terpenoids: efficient multistep biotransformation procedures affording isotope-labeled 2C-methyl-D-erythritol 4-phosphate using recombinant 2C-methyl-D-erythritol 4-phosphate synthase. *J. Org. Chem* 2001, 66, 7770-7775.

Hecht, S. Intermediate und Enzyme des alternativen Terpenbiosyntheseweges. 2002. Dissertation TU-München.

Henriksson LM, Unge T, Carlsson J, Aqvist J, Mowbray SL and Jones TA. Structures of Mycobacterium tuberculosis 1-deoxy-D-xylulose-5-phosphate reductoisomerase provide new insights into catalysis. *J. Biol. Chem* 2007, 282, 19905-19916.

Hill RE, Sayer BG and Spenser ID. Biosynthesis of vitamin B6: incorporation of D-1-deoxyxylulose. *J Am Chem Soc* 1989, 117, 1661–1662.

Hirsch AKH, Lauw S, Gersbach P, Schweizer WB, Rohdich F, Eisenreich W, Bacher A, and Diederich F. Non-phosphate inhibitors of IspE protein, a kinase in the non-mevalonate pathway for isoprenoid biosynthesis and a potential target for antimalarial therapy. *ChemMedChem* March 2007, Vol.2, Issue 6, 806-810.

Hoeffler JF, Tritsch D, Grosdemange-Billiard C and Rohmer M. Isoprenoid biosynthesis via the methylerythritol phosphate pathway. Mechanistic investigations of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *Eur J Biochem*, 2002, 269, 4446–4457.

Hoffman SL, Subramanian GM, Collins FH and Venter JC. Plasmodium, human and Anopheles genomics and malaria. *Nature* 2002, 415, 702-709.

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

Illarionova V, Kaiser J, Ostrojenkova E, Bacher A, Eisenreich W and Rohdich F. Non-mevalonate terpene biosynthesis enzymes as antiinfective drug targets. Substrate synthesis and high throughput screening methods. *J. Org. Chem* 2006, 71, 8824-8834.

Johnson AE. and Tanner ME. Epimerization via carbon–carbon bond cleavage. L-Ribulose-5-phosphate 4-epimerase as amasked class II aldolase. *Biochemistry* 1998, 37, 5746–5754.

Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, Hintz M, Turbachova I, Eberl M, Zeidler J, Lichtenthaler HK et al. Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* 1999, 285, 1573–1576.

Kaiser J , Yassinb M, Prakash S, Safib N, Agamid M, Lauw S, Ostrozhenkova E, Bacher A, Rohdich F, Eisenreich W, Safi J, Golan A Goldhirshc. Anti-malarial drug targets: Screening for inhibitors of 2C-methyl-D-erythritol 4-phosphate synthase (IspC protein) in Mediterranean plants. *Phytomedicine* April 2007, Vol 14, Issue 4, 242-249.

Kaiser, J. Untersuchungen an Enzymen der Terpenbiosynthese, 2005. Dissertation TU-München.

Kemp LE, Bond CS and Hunter WN. Structure of 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase: An essential enzyme for isoprenoid biosynthesis and target for antimicrobial drug development. *PNAS* 2002, 10, 6591-6596.

Khachik F., Beecher G.R. and Smith JC Jr. Lutein, lycopene and their oxidative metabolites in chemoprevention of cancer. J. Cell Biochem 1995, 22, 236–246.

Koppisch AT, Fox DT, Blagg BS and Poulter CD. E. coli MEP synthase: steady-state kinetic analysis and substrate binding. *Biochemistry* 2002, 41, 236–243.

Kuemmerle HP, Murakawa T and De Santis F. Pharmacokinetic evaluation of fosmidomycin, a new phosphonic acid antibiotic. *Chemioterapia* 1987, 6, 113-119.

Kuemmerle HP, Murakawa T, Soneoka K, Konishi T. Fosmidomycin: a new phosphonic acid antibiotic. Part I: Phase I tolerance studies. Int*. J. Clin. Pharmacol. Ther. Toxicol.* 1985 23: 515-520.

Kuntz L, Tritsch D, Grosdemange-Billiard C, Hemmerlin A, Willem A, Bach TJ and Rohmer M. Isoprenoid biosynthesis as a target for antibacterial and antiparasitic

drugs: phosphonohydroxamic acids as inhibitors of deoxyxylulose phosphate reducto-isomerase. *Biochem J* 2005, 386, 127–135.

Kuzmic P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem*. 1996, 237: 260-273.

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

Kuzuyama T, Takahashi S, Takagi M and Seto H. 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* 2000, 275, 19928–19932.

Lange BM and Croteau R. 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* 1999, 365, 170–174.

Laule O, Fürholz A, Chang HS, Zhu T, Wang X, Heifetz PB, Gruissem W & Lange M. Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in Arabidopsis thaliana. *Proc Natl Acad Sci USA* 2003, 100, 6866–6871.

Lee LV., Vu MV and Cleland WW. ^{13}C and deuterium isotope effects suggest an aldol cleavagemechanism for L-ribulose- 5-phosphate 4-epimerase. *Biochemistry* 2000, 39, 4808–4820.

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

Logemann J, Schell J and Willmitzer L. Improved method for the isolation of RNA from plant tissues. *Anal Biochem* 1987, 163, 16–20.

Mac Sweeney A, Lange R, Fernandes RP, Schulz H, Dale GE, Douangamath A, Proteau PJ and Oefner C. 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.* 2005, 345, 115-127.

Miallau L, Alphey MS, Kemp LE, Leonard GA, McSweeney SM, Hecht S, Bacher S, Eisenreich W, Rohdich F, and Hunter WN. Biosynthesis of isoprenoids: Crystal structure of 4 diphosphocytidyl-2C-methyl-D-erythritol kinase. *Proc. Natl Acad. Sci. USA* 2003, 100, 9173 – 9178.

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

Missinou MA, Borrmann S, Schindler A, Issifou S, Adegnika AA, Matsiegui PB, Binder R, Lell B, Wiesner J, Baranek T, Jomaa H, Kremsner PG. Fosmidomycin for malaria. *Lancet* 2002, 360: 1941-1942.

Morlais I, Mori A, Schneider JR and Severson DW. A targeted approach to the identification of candidate genes determining susceptibility to Plasmodium gallinaceum in Aedes aegypti. *Mol. Genet. Genomics* 2003, 269: 753-764.

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

Müller C. Untersuchungen u¨ ber zwei Enzyme der plastida¨ ren Isoprenoidbiosynthese: DOXP-Synthase und DOXP-Reduktoisomerase. *Karls Beitr Pflanzenphysiol* 2003, 39, 1–129.

National Biomedical Research Foundation, Washington, DC. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985, 39, 783–791.

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng*. 1997, 10: 1-6.

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

Phaosiri C and Proteau PJ. Substrate analogs for the investigation of deoxyxylulose 5-phosphate reductoisomerase inhibition: synthesis and evaluation. *Bioorg. Med. Chem. Lett.* 2004, 14, 5309-5312.

Proteau PJ, Woo YH, Williamson RT and Phaosiri C. Stereochemistry of the reduction step mediated by recombinant 1-deoxy-D-xylulose 5-phosphate isomeroreductase. *Org. Lett.* 1999, 1, 921-923.

Proteau PJ. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase: an overview. *Bioorg Chem* 2004, 32, 483–493.

Putra SR, Lois LM, Campos N, Boronat A and Rohmer M. Incorporation of [2,3-$^{13}C_2$]- and [2,4-$^{13}C_2$]-D-1-deoxyxylulose into ubiquinone of *Escherichia coli* via the mevalonate-independent pathway for isoprenoid biosynthesis. *Tetrahedron Lett.* 1998, 39, 23-26.

Qureshi N and Porter JW. Conversion of acetyl-coenzyme A to isopentenyl pyrophosphate. In *Biosynthesis of isoprenoid compounds* (Porter, JW and Spurgeon, SL, eds). 1981. John Wiley, New York.

Radykewicz T, Rohdich F, Wungsintaweekul J, Herz S, Kis K, Eisenreich W, Bacher A, Zenk MH and Arigoni D. Biosynthesis of terpenoids: 1-deoxy-D-xylulose-5-phosphate reductoisomerase from Escherichia coli is a class B dehydrogenase. *FEBS Lett.* 2000, 465, 157-160.

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

Richard SB, Ferrer JL, Bowman ME, Lillo AM, Tetzlaff, CN Cane DE and Noel JP. Structure and Mechanism of 2-*C*-Methyl-D-erythritol 2,4-Cyclodiphosphate Synthase. An enzyme in the mevalonate-independent isoprenoid biosynthetic pathway. *J. Biol. Chem*. 2002, 277, 8667 – 8672.

Rieder C, Eisenreich W, O'Brien J, Richter G, Gotze E, Boyle P, Blanchard S, Bacher A and Simon H. Rearrangement reactions in the biosynthesis of molybdopterin--an NMR study with multiply $^{13}C/^{15}N$ labelled precursors. *Eur. J. Biochem* 1998, 255, 24-36.

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

Rodríguez-Concepción M. The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr Pharm Des* 2004, 10, 2391–2400.

Rohdich F, Bacher A & Eisenreich W. Perspectives in anti-infective drug design. The late steps in the biosynthesis of the universal terpenoid precursors, isopentenyl diphosphate and dimethylallyl diphosphate. *Bioorg. Chem.* 2004, 32, 292-308.

Rohdich F, Bacher A and Eisenreich W. Isoprenoid biosynthetic pathways as anti-infective drug targets. *Biochem. Soc. Trans.* 2005, 33, 785-791.

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

Rohmer M, Grosdemange-Billiard C, Seemann M and Tritsch D. Isoprenoid biosynthesis as a novel target for antibacterial and antiparasitic drugs. *Curr. Opin. Investig. Drugs* 2004, 5, 154-162.

Rohmer M. A mevalonate-independent rout to isopentenyl diphosphate. In *Comprehensive natural product chemistry* (Cane, DE, ed) 1999, pp. 45-68. Pergamon, Oxford.

Rohmer M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat.prod.Rep* 1999, 16, 565-574.

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

Rohmer M., Seemann M., Horbach S., Bringer-Meyer S. and Sahm H. Glyceraldehyde 3-phosphate and pyruvate as precursors of isoprenic units in an alternative pathway for terpenoid biosynthesis. J. Am. Chem. Soc. 1996, 118, 2564–2566.

Römisch W. Coenzym-Isotopologe asl spektroskopische Sonden, 2005, TU München.

Römisch W. Synthese isotopmarkierter RNA, 2000. TU München.

Sacchettini JC. and Poulter CD. Creating isoprenoid diversity. Science 1997, 277, 1788–1789.

Schwarz MK and Arigoni D. Ginkgolide biosynthesis. *In Comprehensive Natural Product Chemistry* (Cane DE, ed.) 1999, pp. 367–399. Pergamon, Oxford.

Schwarz MK. Terpen-Biosynthese in *Gingko biloba*: Eine überraschende Geschichte, PhD Thesis, 1994, ETH Zürich, Switzerland.

Schwender J, Müller C, Zeidler J and Lichtenthaler HK. Cloning and heterologous expression of a cDNA encoding 1-deoxy-D-xylulose-5-phosphatereductoisomerase of Arabidopsis thaliana. *FEBS Lett* 1999, 455, 140–144.

Steinbacher S, Kaiser J, Eisenreich W, Huber R, Bacher A and Rohdich F. 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*. 2003, 278: 18401-18407.

Steinbacher S, Kaiser J, Wungsintaweekul J, Hecht S, Eisenreich W, Gerhardt S, Bacher A and Rohdich F. Structure of 2C-Methyl-D-erythritol-2,4- cyclodiphosphate Synthase Involved in Mevalonateindependent Biosynthesis of Isoprenoids. *J. Mol. Biol*. 2002, 316, 79-88.

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

Van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI. Translocation of proteins across the multiple membranes of complex plastids. *Biochim. Biophys. Acta* 2001, 1541, 34-53.

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

Vennerstrom JL, Arbe-Barnes S, Brun R, Charman SA, Chiu FC, Chollet J, Dong Y, Dorn A, Hunziker D, Matile H, McIntosh K, Padmanilayam M, Santo Tomas J, Scheurer C, Scorneaux B, Tang Y, Urwyler H, Wittlin S and Charman WN. Identification of an antimalarial synthetic trioxolane drug development candidate. Nature 2004, 430, 900-904

Volk R. and Bacher A. Biosynthesis of riboflavin. Studies on the mechanism of L-3,4-dihydroxy-2-butanone 4-phosphate synthase. *J. Biol. Chem.* 1991, 266, 20610-20618.

Volk R. Biosynthese der Riboflavinvorstufe 3,4-Dihydroxybutanon-4-phosphat, 1989. Dissertation TU-München.

Waller RF, Reed MB, Cowman AF, McFadden GI. Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *Embo J*. 2000, 19: 1794-1802.

White RH. Stable isotope studies on the biosynthesis of the thiazole moiety of thiamin in *Escherichia coli*. Biochemistry 1978, 17, 3833–3840.

Wong A, Munos JW, Devasthali V, Johnson KA and Liu HW. Study of 1-deoxy-D-xylulose-5-phosphate reductoisomerase: synthesis and evaluation of fluorinated substrate analogues. *Org. Lett.* 2004, 6, 3625-3628.

Wong U and Cox RJ. The Chemical Mechanism of D-1-Deoxyxylulose-5-phosphate Reductoisomerase from Escherichia coli. *Angew. Chem. Int. Ed. Engl* 2007, .46, 4926-4929.

Woo YH, Fernandes RP and Proteau PJ. Evaluationof fosmidomycin analogs as inhibitors of the Synechocystis sp. PCC6803 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *Bioorg Med Chem* 2005, 14, 2375–2385.

Wungsintaweekul J. Enzymes of the alternativ terpenoid pathway in *Escherichia coli*, 2001, Dissertation TU-München.

Yajima S, Hara K, Sanders JM, Yin F, Ohsawa K, Wiesner J,| Jomaa H,| and Oldfield E. Crystallographic Structures of Two Bisphosphonate:1-Deoxyxylulose-5-Phosphate Reductoisomerase Complexes. *J. AM. CHEM. SOC. 2004*, 126, 10824-10825.

Yajima S, Nonaka T, Kuzuyama T, Seto H, and Ohsawa K. Crystal Structure of 1-Deoxy-D-xylulose 5-phosphate Reductoisomerase Complexed with Cofactors:-

Implications of a Flexible Loop Movement upon Substrate Binding1 *J. Biochem* 2002. 131, 313-317.

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

Zeidler J, Schwender J, Müller C, Wiesner J, Weidemeyer C, Beck E, Jomaa H and Lichtenthaler HK. Inhibition of the non-mevalonate 1-deoxy-Dxylulose phosphate pathway of plant isoprenoid biosynthesis by fosmidomycin. *Z Naturforsch [C]* 1998, 53, 980–986.

Zhou D. and White RH. Early steps of isoprenoid biosynthesis in Escherichia coli, *Biochem. J*. 1991, 273, 627-634.

Zuegge J, Ralph S, Schmuker M, McFadden GI and Schneider G. Deciphering apicoplast targeting signals--feature extraction from nuclear-encoded precursors of Plasmodium falciparum apicoplast proteins. *Gene* 2001, 280, 19-26.