The biosynthesis of phylloquinone (vitamin K_1) in higher plants

Dissertation

zur Erlangung des Doktorgrades der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

Jeferson Gross

aus Porto Alegre, Brasilien

2006

1. Gutachter: PD Dr. Jörg Meurer
2. Gutachter: Prof. Dr. Jürgen Soll

Tag der mündlichen Prüfung: 27.10.2006

Abbreviations

µE	microEinstein (1 E = 1 mol of photons)
µg/g	microgram/gram
A_0	Photosystem I primary electron acceptor composed of chlorophyll *a*
A_1	Photosystem I secondary electron acceptor composed of phylloquinone
ADCS	4-amino-4-deoxychorismate synthase
AHAS	acetohydroxy acid synthase
AS	anthranilate synthase
ATP	adenosine 5′-triphosphate
bp	base pairs
cDNA	complementary DNA
CAPS	cleaved amplified polymorphic sequence maker
CDD	conserved domains database
C-terminal	carboxy-terminal part of a protein
N-terminal	amino-terminal part of a protein
dCAPS	derived cleaved amplified polymorphic sequence maker
DNA	desoxyribonucleic acid
dNTPs	desoxynucleoside triphosphates
DTE	1,4-dithioerythritol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol-bis(2-aminoethylether)-*N,N,N′,N′*-tetraacetic acid
EMS	ethyl methanesulfonate
ESTs	expressed sequence tags
ETC	electron transport chain
e-value	expect value
EPR	electron paramagnetic resonance
F_A	[4Fe-4S] cluster of PsaC subunit
F_B	[4Fe-4S] cluster of PsaC subunit
F_m	maximum fluorescence yield
Fs	steady state fluorescence
F_o	minimal fluorescence yield
F_x	interpolypeptide [4Fe-4S] cluster between PsaA and PsaB subunits
g	gram

H-bond	hydrogen-bond
hcf	*high chlorophyll fluorescence*
HPLC	high performance liquid chromatography
ICS	isochorismate synthase
kb	kilobases
kDa	kilodalton
LHCI	chlorophyll-binding Photosystem I light-harvesting complex
LHCII	chlorophyll-binding Photosystem II light-harvesting complex
m	metre
M	molar
MES	2-morpholinoethanesulfonic acid
mg	milligram
ml	millilitre
mM	millimolar
mm	millimetre
MOPS	3-[N-morpholino]propanesulfonic acid
MQ	menaquinone
mRNA	messenger RNA
NA	1,4-dihydroxy-2-naphthoate
NADH	nicotinic adenine dinucleotide, reduced form
NADPH	nicotinic adenine dinucleotide phosphate, reduced form
nm	nanometre
NOX	NADH oxidase
NPQ	non-photochemical chlorophyll *a* fluorescence quenching
OSB	*o*-succinylbenzoate
P700	Photosystem I primary electron donor chlorophyll *a*
PAM	pulse amplitude–modulated fluorometer
PCR	polymerase chain reaction
pha	*phylloquinone absence*
*pha*3c	*pha*3 mutant complemented with the cDNA form 4
*pha*4c	*pha*4 mutant complemented with the cDNA form 4
PhQ	phylloquinone
PM	plasma membrane
PSI	Photosystem I

PSII	Photosystem II
PVDF	polyvinylidene difluoride
qP	photochemical chlorophyll *a* fluorescence quenching
RCF	relative centrifugal force
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcription PCR
SA	salicylic acid
sec	second
SD	standard deviation
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SHCHC	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate
SSLP	simple sequence length polymorphism
T-DNA	transferred DNA
ThDP	thiamine diphosphate
Tris	tris-(hydroxymethyl)-aminomethane
UTR	untranslated region
v/v	volume per volume
w/v	weight per volume

Contents

1. INTRODUCTION 9

1.1. The vitamin K 9

1.1.1. Identification and isolation of vitamin K type compounds: a historical overview 9

1.1.2. The functions of vitamin K in mammalian physiology 10

1.2. Phylloquinone as a cofactor of Photosystem I 11

1.2.1. The identification of phylloquinone as a component of Photosystem I 11

1.2.2. The role of phylloquinone in the electron transport within Photosystem I 12

1.2.3. The Photosystem I structure 13

1.2.4. The Photosystem I function 13

1.3. The biosynthesis of phylloquinone 14

1.3.1. The biosynthesis of menaquinone in eubacteria 14

1.3.2. The biosynthesis of phylloquinone in cyanobacteria 16

1.3.3. The biosynthesis of phylloquinone in plants 18

1.4. Alternative function of phylloquinone in plasma membrane of plants 18

1.5. The use of a chlorophyll fluorescence screening to identify the function of new nuclear-encoded plastid-localized proteins 20

1.5.1. Most of the plastid proteins are nuclear encoded 20

1.5.2. Screening of *high chlorophyll fluorescence* (*hcf*) mutants 21

1.5.3. Nuclear-encoded factors isolated from *hcf* mutants 21

1.6. Goals of the project 22

2.0. MATERIALS AND METHODS 24

2.1. Materials 24

2.1.1. Plant stocks 24

2.1.2. Bacterial strains 24

2.1.3. Cloning vectors 24

2.1.4. Clones 24

2.1.5. Oligonucleotides 25
2.1.6. Chemicals and enzymes 26
2.1.7. Media, solutions, buffers and antibiotics 27
2.1.8. Antibodies 28

2.2. Methods 28
2.2.1. Molecular biology methods 28
2.2.1.1. General methods 28
2.2.1.2. RNA gel blot analysis of the *phyllo* transcript 29
2.2.1.3. Western blot analysis of thylakoid protein complexes 30

2.2.2. *Arabidopsis* methods 31
2.2.2.1. Plant growth, seed sterilization 31
2.2.2.2. Rapid isolation of plant DNA for PCR 31
2.2.2.3. Isolation of total RNA 32

2.2.3. Spectroscopic and fluorimetric methods 32
2.2.3.1. Chlorophyll *a* fluorescence analyses 32
2.2.3.2. Light-induced changes of the P700 redox state 32

2.2.4. Genetic methods 33
2.2.4.1. Mutant selection 33
2.2.4.2. High-resolution genetic mapping of the *pha* mutations 33
2.2.4.3. Complementation analysis 33

2.2.5. High performance liquid chromatography 35
2.2.6. Subcellular localization of the PHYLLO protein by fluorescence imaging 35
2.2.7. Sequence analyses 35

3. RESULTS 37
3. 1. Characterization of the *pha* phenotype 37
3.1.1. General phenotype 37
3.1.2. Phylloquinone absence 37
3.1.3. A *hcf* phenotype associated with Photosystem I lesions 37

3.1.4. Specific impairment of Photosystem I complex accumulation 39
3.1.5. Recovery of the phylloquinone content and Photosystem I activity after 1,4-dihydroxy-2-naphthoate feeding 39

3.2. Localization of the *pha* mutations into the *PHYLLO* locus 41
3.3. Evidences of a single gene related to the *PHYLLO* locus 43
3.4. The *PHYLLO* gene 44
3.5. Complementation of the *pha* mutations 45
3.5.1. The failure to complement the *pha* mutations with the RAFL 09-32-C05 cDNA (form 1) 45
3.5.2. Complementation with the engineered full-length form 4 45
3.5.3. Complementation with the genomic locus 46

3.6. Characterization of the *PHYLLO* product 47
3.6.1. PHYLLO is a plastidial protein 48
3.6.2. The MenD module 48
3.6.3. The MenC module 51
3.6.4. The MenH module 54
3.6.5. The MenF 5'-module 56

3.7. Synteny of the *PHYLLO* locus among different kingdoms 57
3.7.1. Conservation of *PHYLLO* in higher plants 57
3.7.2. Conservation of *PHYLLO* in green algae 58
3.7.3. Partial conservation of the architecture of *PHYLLO* in red algae plastomes and eubacterial *men* operons 59

3.8. Genetic characterization of the MenF enzymatic function in *Arabidopsis* 61
3.8.1. The *ICS*1 and *ICS*2 genes in *Arabidopsis* 61
3.8.2. The ICS proteins of *Arabidopsis* 62

3.9. A gene splitting event of the 3'-part of the *PHYLLO men*F module in higher plants 66
3.10. The phylloquinone content associated to Photosystem I activity in mutant, wild-type and 1,4-dihydroxy-2-naphthoate-fed plants 67

3.11. Presence of other *men* genes in the *Arabidopsis* genome 67

4.0. DISCUSSION 69

4.1. Essential role of phylloquinone in higher plants 69

4.1.1. The function of *PHYLLO*, *ICS*1 and *ICS*2 in the phylloquinone biosynthesis 69

4.1.2. A photosynthetic defect related to the Photosystem I function in the *pha* mutants 69

4.1.3. The bulk of phylloquinone in *Arabidopsis* is not associated with Photosystem I 70

4.2. *PHYLLO*, a plant locus originated from a fusion of four eubacterial genes 72

4.2.1. *PHYLLO* has a composite structure 72

4.2.2. *PHYLLO* was presumably originated from the structure of an operon 73

4.2.3. PHYLLO, a prokaryotic metabolon adapted to eukaryotes 74

4.3. A metabolic link between plant resistance and photosynthesis 76

4.3.1. Gene duplication and splitting events of the *PHYLLO men*F module 76

4.3.2. The *Arabidopsis* ICS1 protein represents a branching point between phylloquinone and salicylic acid biosynthesis 76

4.4. Concluding remark: the phylloquinone biosynthetic pathway in *Arabidopsis thaliana* 78

5. SUMMARY 80

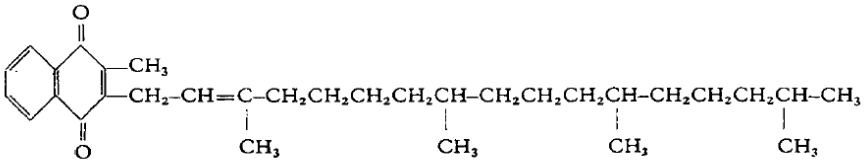
6. LITERATURE 82

1. Introduction

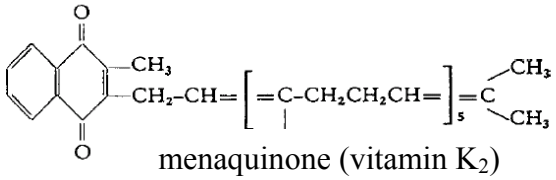
1.1. The vitamin K

1.1.1. Identification and isolation of vitamin K type compounds: a historical overview

In the beginning of the 1930s, blood clot disorders characterized by several hemorrhagic spots under the skin, muscles and other organs were frequently observed in chickens deliberately submitted to a fat-poor diet. Initial attempts to assign the cause of this new experimental hemorrhagic disease to the lack of a known substance in the nutrition, like cholesterol or vitamin C, were unsuccessful until 1934, when the Danish biochemist Henrik Dam demonstrated that the reason was due to the absence of a hitherto unrecognised compound in the diet, of special occurrence in green leaves (Dam, 1946). Furthermore, it became clear with time that the new factor was also important not only for the clot system of other animals, but also for the coagulation process in humans. In 1935 the new compound was characterized as a fat-soluble vitamin and was given the designation vitamin K, for the first letter in the word "Koagulation", according to the Scandinavian and German spelling. This prompted a run to identify the chemical structure of the new component. The decisive step was the isolation, reported in 1939, of a pure sample of vitamin K prepared from green leaves. Subsequently, a different vitamin K type was also isolated from putrefied fish meat (Dam, 1946). Finally, in the same year, the American scientist Edward Doisy not only solved the chemical structure of both substances as different 1,4-naphthoquinone derivatives, but was also able to synthetically prepare them under laboratorial conditions (Dam, 1946; Doisy, 1976). Since then, these two compounds have been known as vitamin K_1 and vitamin K_2. The vitamin K_1 is found in green vegetables and chemically is a 2-methyl-3-phytyl-1,4-naphthoquinone, or phylloquinone (PhQ) (Fig. 1), whereas the vitamin K_2 is of eubacterial origin and structurally is a 2-methyl-3-prenyl-1,4-naphthoquinone, or menaquinone (MQ) (Lamson, and Plaza, 2003). Henrik Dam and Edward Doisy received in 1943 the Nobel Prize of medicine and physiology for their discovery and structural characterization of vitamin K.



phylloquinone (vitamin K_1)



menaquinone (vitamin K_2)

Figure 1. The chemical structures of phylloquinone (vitamin K_1) and menaquinone (vitamin K_2). Structures are depicted according to the original description of Doisy and coworkers (Dam, 1946).

1.1.2. The functions of vitamin K in mammalian physiology

The molecular role of vitamin K in the blood clot physiology as a cofactor of gamma-glutamyl-carboxylases was only established in the mid-1970s (Booth and Suttie, 1998). These enzymes catalyse post-translational carboxylations of glutamic acid residues in the amino-terminal (N-terminal) part (the Gla region) of coagulation factors II (prothrombin), VII, IX, X, as well as the proteins C, S, and Z. The carboxylated form of these proteins corresponds to the active state necessary to interact in the coagulation cascade. In the process of carboxylation, vitamin K epoxides are formed. These are subsequently recycled by sequential action of an epoxide reductase and a quinone reductase that regenerate the vitamin K to the active hydroquinone form (Lamson and Plaza, 2003).

Although vitamin K is usually identified as a critical factor for blood coagulation, recent research in mammals has found that vitamin K is also a cofactor of gamma-glutamyl-carboxylases modifying Gla-containing proteins important for the bone metabolism and arterial calcification (Wallin *et al.*, 2001; Lamson and Plaza, 2003). Furthermore, vitamin K-dependent tyrosine kinases receptors and phosphatases are involved in signal transduction pathways related to cell growth, inhibition of programmed cell death (apoptosis) and in cell transformation (Saxena *et al.*, 2001; Otsuka *et al.*, 2004). The latter properties could account for the anticancer effects observed in experimental trials involving administration of vitamin K type compounds (Lamson and Plaza, 2003).

Now, it is well established that green leafy vegetables are the major source of vitamin K compounds for human nutrition. Its contribution reaches 40-50% of the total diary intake, surpassing other sources like oils, fruits, meat, milk derivatives, eggs and the vitamin K_2 produced in the human gut by the action of bacteria (Booth and Suttie, 1998). Although it was already known in the late 1930s that PhQ is related to chloroplasts and could supposedly exert some role in the photosynthetic process (Dam, 1946), it was only in the 1980s that its function as a cofactor of the photosystem I (PSI) complex in the thylakoid membranes of plastids was definitively established (below).

1.2. Phylloquinone as a cofactor of Photosystem I

1.2.1. The identification of phylloquinone as a component of Photosystem I

By early-1980s it was already known that the redox reaction catalysed by PSI involved an intrinsic electron transport chain (ETC) consisting of a primary electron donor (the chlorophyll *a* P700) and secondary electron acceptors, called A_0, A_1, A_2, which can be distinguished from each other by different electron paramagnetic resonance (EPR) signals and different absorbance contributions to the optical spectra (Bonnerjea and Evans, 1982; Gast *et al.*, 1983). The nature of the centre A_2, as composed by three [4Fe-4S] clusters (named F_X, F_A and F_B), was already solved in the 1970s (Golbeck and Kok, 1978), before that the centre A_0 has been identified in 1982-1983 as another chlorophyll *a* (Bonnerjea and Evans, 1982; Gast *et al.*, 1983). At that time, the remaining centre A_1 was suspected to be a quinone-like compound (Gast *et al.*, 1983). Parallel to this spectroscopic approaches, several biochemical studies in the 1970s and early-1980s indicated an association of PhQ with membrane preparations enriched in PSI polypeptides (Interschick-Niebler and Lichtenthaler, 1981). These trends came together in 1985-1986 when four independent communications reported that PhQ co-purifies with PSI in a ratio of two PhQ per P700 (Takahashi *et al.*, 1985; Schoeder and Lockau, 1986) and that UV and flash absorption spectra related to the component of A_1 centre definitively corresponds to that of PhQ (Brettel *et al.*, 1986; Mansfield and Evans, 1986). Subsequently, experimental data demonstrate that photoreduction activity of PSI is inactivated by hexane/methanol-mediated extraction of PhQ from the A_1 centre, followed by full reconstitution of activity by addition of exogenous PhQ to this centre, as well as partial recovery after replacement with other quinone-like compounds (Biggins and Mathis, 1988; Biggins, 1990; Itoh *et al.*, 2001). This definitively established the

functional role of PhQ in mediating electron transport inside the PSI complex of plants, algae and cyanobacteria.

1.2.2. The role of phylloquinone in the electron transport within Photosystem I

The PSI enzymatic complex catalyses the light-driven electron transfer from a reduced electron donor (cytochrome *c*6 in cyanobacteria or a plastocyanine in plants and algae) localized on the lumenal side of the thylakoid membrane to an acceptor (flavodoxin in cyanobacteria or ferrodoxin in plants and algae) positioned at the stromal side (Fromme *et al.*, 2001). The reaction involves absorption of light energy that leads to formation of an excited state of the primary donor, P700, followed by electron transfer to a primary acceptor, A_0. The electron is then further transferred along the secondary electron acceptors, A_1, F_X, F_A, and F_B, until reduction of the stromal ferrodoxin/flavodoxin (Brettel and Leibl, 2001). The structure of this ETC became apparent with the publications of the solved X-ray structures of the cyanobacteria *Synechococcus elongatus* and the pea PSI (Jordan *et al.*, 2001; Ben-Shem *et al.*, 2003). It consists of 6 chlorophyll *a* (eC-A1 and eC-B1, the P700, eC-A2 and eC-B2, connecting P700 to A_0, eC-A3 and eC-B3, the centre A_0) and the two PhQ, Qk-A and Qk-B (Fig. 2). They are pair wise orientated constituting two branches of the ETC that converge to the iron-sulfur centre F_X (Brettel and Leibl, 2001; Golbeck, 2003).

The two PhQ, Qk-A and Qk-B, present in the PSI are tightly bound, but can be extracted by organic solvents, one being more easily removed than the other. Interestingly, the electron transfer to F_X is not affected by extraction of the first PhQ, but is blocked by depletion of the second (Biggins and Mathis, 1988; Biggins, 1990; Itoh *et al.*, 2001). The binding pocket for PhQ is revealed by the X-ray crystal structure of *Synechococcus elongatus* showing that W697 PsaA and W677 PsaB form the most prominent contacts by π-stacking interactions with the two PhQ (Jordan *et al.*, 2001; Fromme *et al.*, 2001). Also the S692 PsaA and S672 PsaB form hydrogen-bonds (H-bonds) to PhQ, whereas R694 PsaA and R674 PsaB are involved in a H-bonded network of side groups that constitutes an electron transfer pathway between the PhQs and F_X. Both PhQ molecules accept only one hydrogen, which is donated by the NH groups of the backbone of L722 PsaA and L706 PsaB to the carbonyl oxygen at the ortho position to the phytyl chains of PhQ (Xu *et al.*, 2003). The single hydrogen bond to PhQ in the PSI is remarkable, since in all enzymes of known structures containing bound quinones, hydrogens are transferred to both carbonyl groups (Fromme *et al.*, 2001).

1.2.3. The Photosystem I structure

The PSI is a highly conserved protein complex present in the thylakoid membranes of cyanobacteria, algae and plants (Nelson and Ben-Shem, 2004). The X-ray structure of the *Synechococcus elongatus* PSI reveals the presence of 12 protein subunits and 127 cofactors, comprising 96 chlorophyll *a*, 22 carotenoids, 2 PhQ, 3 [4Fe-4S] clusters and 4 lipids (Jordan *et al.*, 2001). The PSI of higher plants is composed of a reaction centre formed by the subunits PsaA–PsaL, PsaN and PsaO (Ben-Shem *et al.*, 2003). The core PsaA–PsaB binds the P700 and all the other cofactors involved in the ETC, including the two PhQ. In addition, this heterodimer contains approximately 80 chlorophylls that function as an intrinsic light-harvesting antenna. The terminal [4Fe-4S] clusters F_A and F_B are bound by the small stromal subunit PsaC. The remaining subunits have roles in the docking of ferredoxin (PsaC, PsaE and PsaD) and plastocyanin (PsaF), the association with light-harvesting complex I (LHCI) (PsaK, PsaG, PsaJ and PsaF), the docking of light-harvesting complex (LHCII) (PsaI, PsaH and PsaL) and the maintenance of the complex integrity (Sheller *et al.*, 2001; Nelson and Ben-Shem, 2004). Furthermore, an extrinsic membrane antenna, composed by the LHC-containing proteins (Lhca1–Lhca5) constituting the LHCI, is peripherical bound to PSI in plants (Scheller *et al.*, 2001; Ben-Shem *et al.*, 2003; Storf *et al.*, 2005).

1.2.4. The Photosystem I function

In oxygenic photosynthesis, PSI operates in line with the photosystem II (PSII) complex. In both photosystems, the sun light is trapped and used as oxidation power in a series of redox reactions that result in translocation of electrons across the thylakoid membrane via a chain of electron carriers (Fromme *et al.*, 2001; Saenger *et al.*, 2002; Nelson and Ben-Shem, 2004). The process initiates with oxidation of two water molecules to O_2 and is enable by the excited state of the primary electron donor of PSII (the P680). The electrons are further carried from PSII to the cytochrome b_6f complex by a mobile plastoquinone pool and then to PSI by the soluble electron carrier proteins plastocyanin (plants and algae) or cytochrome *c*6 (cyanobacteria). Finally, PSI reduces ferredoxin, providing the necessary electrons for the reduction of $NADP^+$ to NADPH by the ferredoxin-$NADP^+$ oxidoreductase. All the electron transfer process is couple to formation of an electrochemical potential across the thylakoid membrane. This proton motive force is further used to drive ATP synthesis.

NADPH and ATP take part in subsequent dark reactions in the stroma that finally reduce CO_2 to carbohydrates (Fromme *et al.*, 2001; Saenger *et al.*, 2002; Nelson and Ben-Shem, 2004).

In addition to the linear electron transfer, PSI also participates in a cyclic electron transport pathway, in which electrons are transferred from reduced ferrodoxin or NADPH back to the plastoquinone. The electrons then return to PSI via the normal route involving the cytochrome *b6f* complex and plastocyanin. This cyclic electron transport around PSI is an important process, serving not only to produce ATP, but also to avoid overreduction of PSI and stroma components (Munekage *et al.*, 2002 and 2004). Another alternative pathway in which PSI participates is the water-water cycle (Ort and Baker, 2002). This involves photoreduction of O_2 by PSI. The superoxide resultant is dismutated to hydrogen peroxide, which is further converted to water by an ascorbate peroxidase. Interestingly, the monodehydroascorbate, which results from the latter reaction, is recycled by electrons also originating from PSI via NADPH, in a reaction catalysed by the monodehydroascorbate reductase. It has been suggested that this pathway could correspond to 30% of the total eletronic flux from photosynthesis under intense light conditions (Ort and Baker, 2002). The PSI also takes part in state transition, a phenomenon that involves relocation of light-harvesting complex II (LHCII) between PSII and PSI, occurring as a rapid response to alterations in light properties and permitting the fast redistribution of absorbed light energy (Lunde *et al.*, 2000; Haldrup *et al.*, 2001). In the case of excess of light captured by PSII, state transition operates by the dissociation from PSII of a pool of LHCII trimmers that associate with the PSI subunits PsaI, PsaL and PsaH. This enables at the same time a decrease in the light absorption by PSII and an enhancement in the light absorption by PSI (Zhang and Scheller, 2004; Haldrup *et al.*, 2001).

1.3. The biosynthesis of phylloquinone

1.3.1. The biosynthesis of menaquinone in eubacteria

MQ is a naphthoquinone related compound (Fig. 1) that serves as sole quinone in ETC of most anaerobic bacteria and many aerobic Gram-positive bacteria. Nonetheless, some facultative aerobic/anaerobic Gram-negative bacteria, like *E. coli*, use both ubiquinone and MQ, the first during aerobiosis and the latter used as an alternative electron carrier in conditions of anaerobiosis (Meganathan, 2001). Additionally, MQ is supposed to be the

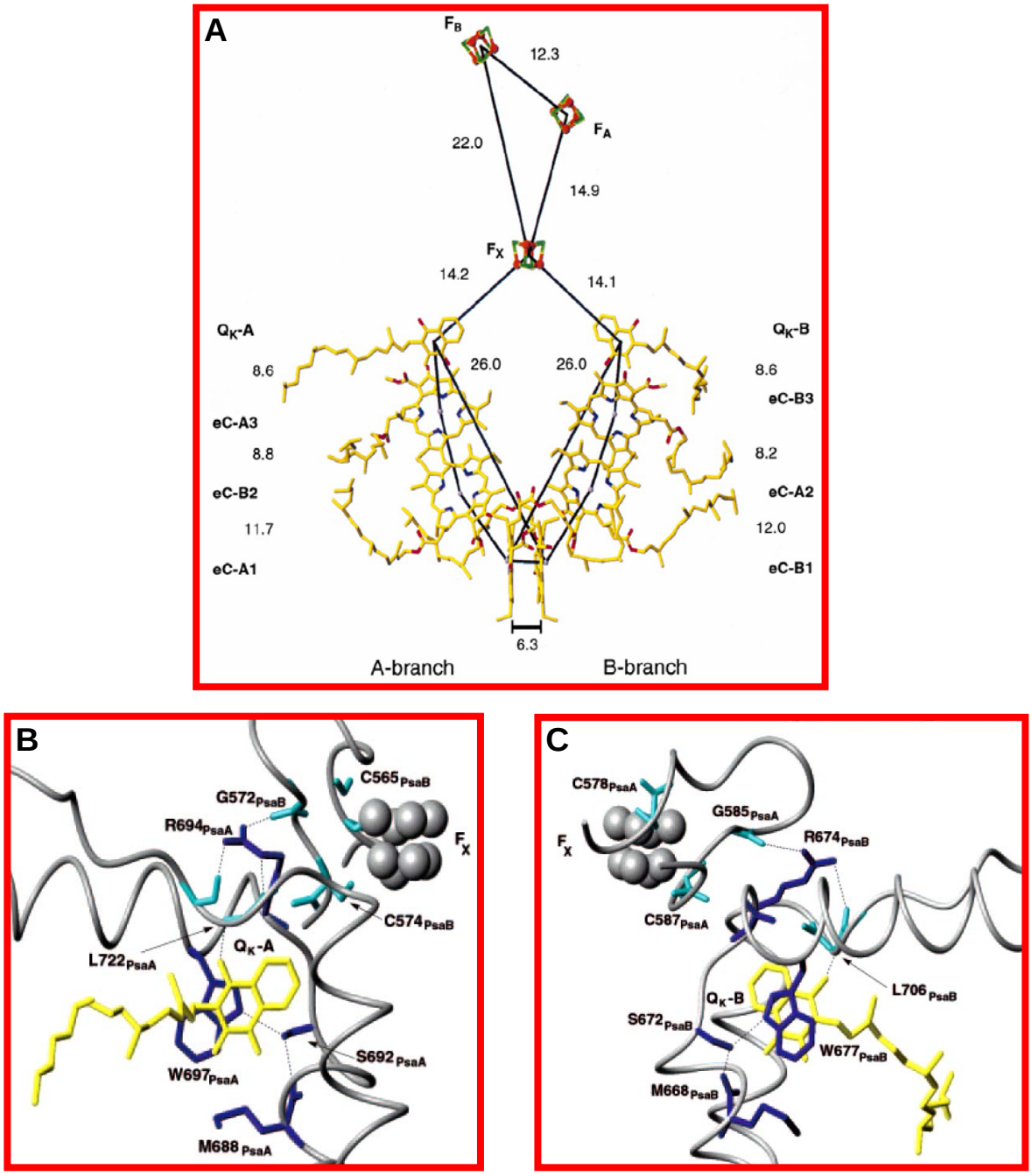


Figure 2. The PhQ in the electron transfer chain of PSI from *Synechococcus elongatus*. (**A**) The redox cofactors chlorophylls and the PhQ are depicted in yellow and are arranged in two branches. The left-hand side corresponds to the A-branch and the right-hand side to the B-branch. The three [4Fe4S] clusters F_X, F_A and F_B are located above these cofactors. The nomenclature of the cofactors is given at the left and right margins, together with the distances (in angstroms) between the corresponding cofactors. Adapted from Fromme *et al.* (2001). (**B** and **C**) The binding pocket for the two PhQ (yellow) close to the iron-sulfur cluster F_X. Some structural amino acids are depicted that constitute a network of contacts between residues extending from the M688 PsaA (M668 PsaB), passing through S692 PsaA (S672 PsaB), W697 PsaA (W677 PsaB), the PhQ, L722 PsaA (L706 PsaB), R694 PsaA (R674 PsaB), until the G572 PsaB (G585 PsaA) in the F_X binding loop. Adapted from Xu *et al.* (2003).

cofactor of the photosynthetic reaction centre type I of green sulfur bacteria (Kjær *et al.*, 1998; Hauska *et al.*, 2001) and, surprisingly, is also present in the PSI of the red alga *Cyanidium caldarium* (Yoshida *et al.*, 2003). From the biosynthetic perspective, it was early demonstrated by experiments involving isotopic traced molecules that MQ is a shikimate-derivative compound with the seven shikimate carbons atoms being incorporated via chorismate into the naphthoquinone nucleus of MQ and the remaining three carbons derived from α-ketoglutarate (Campbell *et al.*, 1967; Bentley and Meganathan, 1982). The MQ biosynthesis has been dissected in *E. coli* and *B. subtilis* during the last three decades by a comprehensive mutational analysis of almost all genes related to the biosynthetic pathway and overexpression and purification of their encoded proteins (Meganathan, 2001). Eight genes (called *men*A, B, C, D, E, F, G, H), encoding Men enzymatic steps necessary for MQ biosynthesis were identified and are often organized into operons in these eubacteria, (Taber *et al.*, 1981; Meganathan, 2001) (Fig. 3).

The MQ biosynthetic pathway is initiated by isomerization of chorismate to isochorismate catalysed by MenF (Daruwala *et al.*, 1996). In the subsequent reaction conducted by MenD, the isochorismate condensates with the thiamine pyrophosphate (TPP) anion of succinic semialdehyde, resulting in the formation of 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) (Palaniappan *et al.*, 1992). This is dehydrated by MenC to the aromatic benzenoid compound *o*-succinylbenzoate (OSB) (Sharma *et al.*, 1993). The OSB is converted by MenE to a CoA thioester (Sharma *et al.*, 1996), followed by cyclization of the naphthalene aromatic ring and thioesterolysis of the CoA by sequential action of MenB and MenH enzymes, generating 1,4-dihydroxy-2-naphthoate (NA) (Sharma *et al.*, 1992; Meganathan, 2001). In the last steps, coordinated by MenA and MenG proteins, NA is prenylated and methylated giving the end product MQ (Suvarna *et al.*, 1992; Koike-Takeshita *et al.*, 1997).

1.3.2. The biosynthesis of phylloquinone in cyanobacteria

PhQ has a C-20 phytol chain in the position C-3, thus differing from the partially unsaturated side chain (usually C-40) observed for the MQ (Fig. 1). This variation is attributed to distinct enzymatic activities verified for the MenA protein, which serves as a phytyl transferase for the PhQ pathway and as a prenyl transferase for the MQ route (Johnson *et al.*, 2000). Despite this difference, the synthesis of the naphthalene ring for both MQ and

PhQ are supposed to follow the same enzymatic steps (Fig.3). In fact, homologs to all *men* genes are present in the genome of *Synechocystis* sp. PCC 6803, which made this organism a model for a systematic reverse genetics approach involving knockout of five genes for PhQ biosynthesis. The disruption of four of them, *men*D, *men*E, *men*B, and *men*A, resulted in a complete absence of PhQ (Johnson *et al.*, 2000 and 2003), which can partially be restored by addition of exogenous vitamin K_1 and several other naphthoquinones (Johnson *et al.*, 2001), confirming that also the biosynthesis of PhQ is operational through the *men* pathway.

The PSI activity in these mutant strains is decreased to 50–60% of wild-type levels, whereas the activity of PS II per cell is comparable to that of the wild type. The phenotype associated to each of these knockout lines is very similar and is a consequence of the imbalance between these two photosystems (Johnson *et al.*, 2000 and 2003). These mutants are not viable under high light conditions (>120 μE m^{-2} s^{-1}) and the knockouts strains can only grow photoautotrophically under low light conditions (20-40 μE m^{-2} s^{-1}), when the observed doubling times are considerably longer than those verified for wild-type cells. This situation is relieved with addition of glucose to the growth medium, but, even so, the mutants are not able to grow under high light conditions, unless atrazine or 10 μM 3-(3,4-dichlorophenyl)-1,1-dimethylurea are present to inhibit PSII activity. These latter results suggest that the phototoxicity verified for the *men* mutants may be a consequence of excess of reductants generated from PSII that are not able to be assimilated by PSI due to the reduced accumulation levels of this complex (Johnson *et al.*, 2000 and 2003).

Interestingly, solvent extracts from PSI trimmers of *men* mutants were analysed by high performance liquid chromatography (HPLC) and demonstrated the presence of plastoquinone-9 (Johnson *et al.*, 2000). This observation was confirmed by EPR analysis of PSI complexes isolated from the mutants, presenting signals that were compatible with the presence of a plastoquinone in the centre A_1 (Zybailov *et al.*, 2000), which may sustain the remaining PSI activity still present in the *men* knockouts despite the absence of PhQ (Semenov *et al.*, 2000). Such a phenomenon was not observed for the *men*G knockout mutant, in which EPR studies indicated the presence of a demethylphylloquinone in the site A_1, confirming the impairment of the methyltransferase step of the pathway. Unless high light intensities of 300 μE m^{-2} s^{-1} were used, the growth rates and photosynthetic activity observed for the *men*G mutant strain resembled those verified for wild-type cells (Sakuragi *et al.*, 2002).

1.3.3. The biosynthesis of phylloquinone in plants

In plants, relatively little is known about PhQ biosynthesis, although it has been suggested that the principle reactions occur at the chloroplast envelope membrane in the presence of light (Schultz *et al.*, 1981; Schütte, 1993). Enzymatic studies involving cell-free systems and addition of traced compounds revealed that chorismate is a direct precursor in formation of NA and that this synthesis proceeds via the intermediate OSB (Schütte, 1993). It was also demonstrated that synthesis of OSB is catalysed from the isochorismate and α-ketoglutarate precursors in presence of thiamine diphosphate (ThDP) and Mn^{+2} and that SHCHC is an intermediate in this reaction (Simantiras and Leistner, 1991). Furthermore, enzymatic activities corresponding to MenA and MenG were identified in subfractions of chloroplasts with the phytylation step associated with envelope membrane and the methylation related to the thylakoid fraction (Schultz *et al.*, 1981; Kaiping *et al.*, 1984). Taken together, these results support the idea that PhQ biosynthesis in plant chloroplasts also proceeds via the *men* pathway.

Recently, the *men*A homolog in *A. thaliana*, *Atmen*A, was isolated by T-DNA tagging and was demonstrated to encode a product localized in the chloroplast (Shimada *et al.*, 2005). The loss of function of this gene abolished PhQ biosynthesis and extremely affected the accumulation of PSI polypeptides. In this mutant the P700 activity was virtually inexistent, in sharp contrast to the results obtained for the *men* genes in *Synechocystis*, which preserve 50-60% of PSI activity. Furthermore, the *Atmen*A knockout plant was also affected in the accumulation of PSII polypeptides, a result interpreted as an indirect effect related to the strong decrease in plastoquinone amounts observed in this mutant to 3% of the wild-type level (Shimada *et al.*, 2005).

1.4. Alternative function of phylloquinone in plasma membrane of plants

Besides the very well established role of PhQ as PSI cofactor in chloroplasts, several reports have directly or indirectly demonstrated the presence of PhQ also in the plasma membrane (PM) of plant cells. Direct evidences for this occurrence came from PM isolations

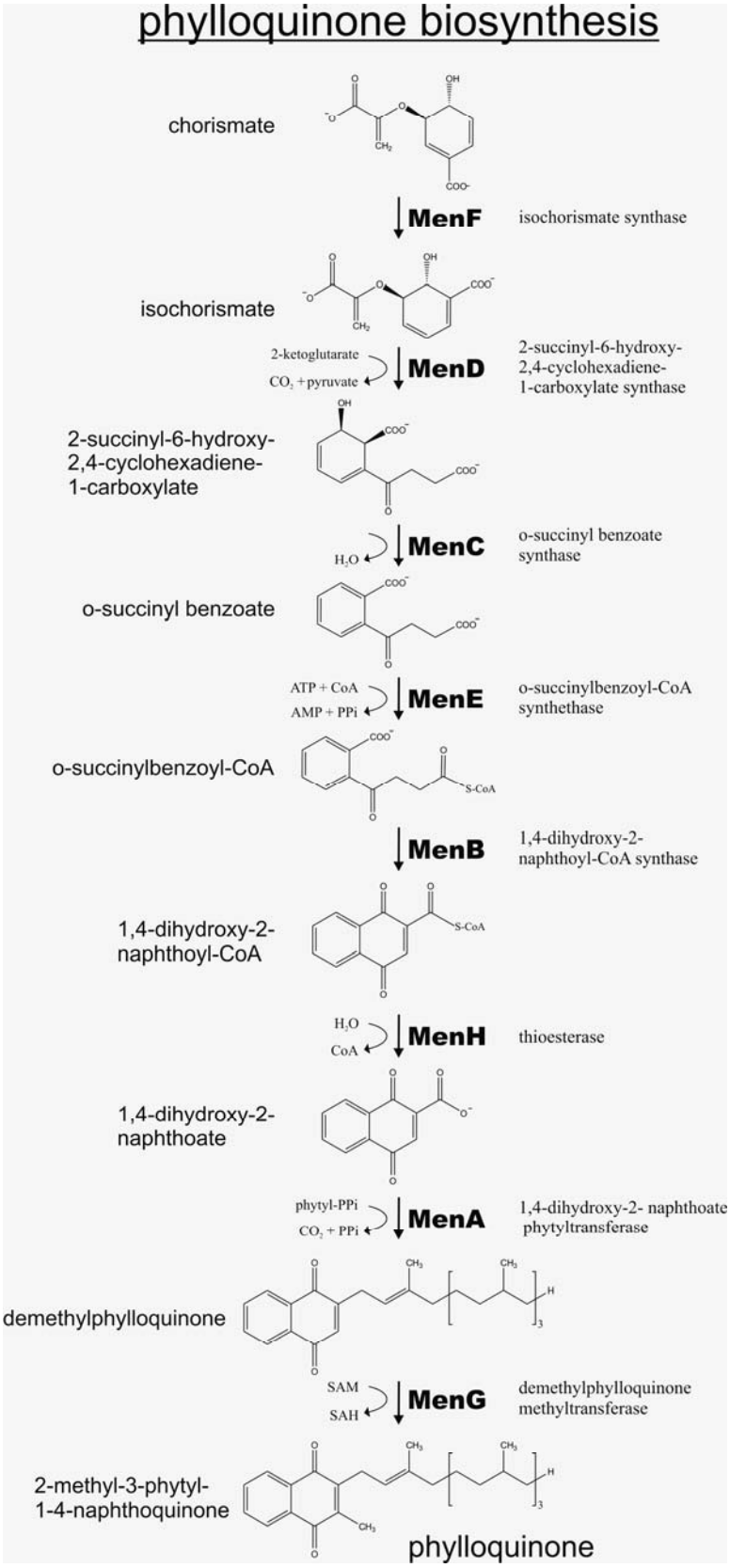


Figure 3. The biosynthesis of PhQ occurs in eight steps. The enzymatic route for MQ synthesis is essentially the same of that for PhQ, contrasting only in the MenA-mediated step that transfers a prenyl tail for the MQ biosynthesis and a phytyl chain for the PhQ pathway. PPi, pyrophosphate; SAM, S-adenosylmethionine; SAH, S-adenosylhomocystein.

obtained from soybean hypocotyls and maize roots (Lüthje *et al.*, 1997). In this latter case, the presence of PhQ was detected in *n*-heptane extracts from PM by HPLC and further confirmed by gas chromatography and mass spectroscopy analysis (Lüthje *et al.*, 1997, Lochner *et al.*, 2003). These positive conclusions are paralleled by negative results in attempts to isolate ubiquinone from PM, even when similar extraction methods were used (Lochner *et al.*, 2003).

It has been suggested that PhQ could serve as a mobile electron carrier in redox systems known to exist in the PM of plants (Lüthje *et al.*, 1997, Lochner *et al.*, 2003). Accordingly, there are a number of publications demonstrating the abolishment of PM redox activity by depletion or inhibition of PhQ in PM, caused by solvent extraction, UV destructive irradiation, or administration of the vitamin K antagonists dicumarol and warfarin (Döring *et al.,* 1992; Barr *et al.*, 1992; Lüthje *et al.*, 1997). In some of these cases, exogenous addition of vitamin K_1 to the PhQ-depleted PM restored the lost redox activity (Barr *et al.*, 1992). Furthermore, NAD(P)H-dependent quinone reductase activities have been isolated from several tissues and plant species, two of them, of 27 kDa and 31 kDa, respectively, have been purified from the PM and are supposed to be localized at the cytoplasmic surface (Luster and Buckout, 1989; Serrano *et al.*, 1994 and 1995; Cordoba *et al.*, 1996). Remarkably, one NADH oxidase (NOX) was partially purified from PM and demonstrated to exhibit PhQ hydroxyquinone oxidase activity coupled to reduction of O_2 and protein disulfides (Bridge *et al.*, 2000). These findings suggest the existence of an electron transfer system in which a PhQ pool serves as a transmembrane intermediate to shuttle electrons originated from the cytoplasmic donor NAD(P)H to acceptors localized at the outer surface of the PM (Lüthje *et al.*, 1997; Bridge *et al.*, 2000).

1.5. The use of a chlorophyll fluorescence screening to identify the function of new nuclear-encoded plastid-localized proteins

1.5.1. Most of the plastid proteins are nuclear encoded

Chloroplasts are genetic compartments of the plant cell that arose in the evolution from a cyanobacteria-like organism incorporated within a eukaryotic host (Raven and Allen, 2003; Timmis *et al.*, 2004). In the course of a long endosymbiotic process the eubacterial-derived organelle has progressively lost autonomy by massive transfer of its genes to the nuclear genome of the eukaryote. As a result, contemporary plastid genomes encode only between 60-

200 proteins. On the other hand, it is estimated that about 2.100-3.100 nuclear encoded proteins are present in the chloroplast proteome of higher plants (Leister, 2003). These comprise not only cyanobacterial-derived proteins, but also proteins already existing in the eukaryotic host, as well as products of novel genes (Martin *et al.*, 2002; Raven and Allen, 2003). Most of them present a characteristic N-terminal extension, the so called transit peptide, that ensures a functional post-translational relocation of proteins from the cytoplasm into the plastid compartment, where they are active in several functional and structural aspects of the plastid physiology and biogenesis (Soll 2002; Raven and Allen, 2003; Timmis *et al.*, 2004).

1.5.2. Screening of *high chlorophyll fluorescence* (*hcf*) mutants

Several nuclear-encoded and plastid-localized proteins are directly or indirectly involved in the photosynthetic process within the chloroplast (Martin *et al.*, 2002). Lesions of these genes affect the photosynthetic electron transport and also the ability of plants to efficiently quench the light energy trapped by the chlorophyll antenna systems in form of excited states. As a consequence, a fraction of the absorbed energy is re-emitted in form of red fluorescence light (Maxwell and Johnson, 2000). This enables an easy procedure to identify mutants presenting photosynthetic lesions, the so called *hcf* mutants, simply by examining them by use of optical devices that can detect the altered pattern of re-emitted fluorescence (Meurer *et al.*, 1996; Pesaresi *et al.*, 2001). Furthermore, analysis of the chlorophyll fluorescence characteristics and the absorbance changes in the reaction centre of PSI (P700) of the mutants allows not only a quantitative description of the effect of the mutation, but also an estimation of the sector of the photosynthetic electron transport or carbon assimilation which might be affected by the mutation. These studies can be complemented by additional molecular methods, like immunoblot and Northern blot analysis of thylakoid protein complexes and transcripts encoding photosynthetic proteins, respectively (Meurer *et al.*, 1996).

1.5.3. Nuclear-encoded factors isolated from *hcf* mutants

Several mutant plants generated by ethyl methane sulfonate (EMS) or T-DNA tagging affected in various aspects of the chloroplast biogenesis have been isolated by the fluorescence screening approach. Two examples demonstrating well this diversity in function are the *hcf*136 mutant defective in a protein crucial for the stability/assembly of PSII

complexes (Meurer *et al.*, 1998a) and the mutant affected in the γ subunit of the plastidial ATP synthase (Dal Bosco *et al.*, 2004), both inactivated by T-DNA insertions. Other mutants have been characterized which are defective in the fine control of postranscriptional processes affecting the stability and functional accumulation of plastidial mRNAs. Some of the corresponding genes have already been described or are in process of isolation. They correspond to the mutants *hcf*107, *hcf*109, *hcf*135, *hcf*145 and *pac* (Felder *et al.*, 2001; Meurer *et al.*, 1998b and 2002; Lezhneva and Meurer, 2004). Two new protein factors involved in the [4Fe-4S] cluster metabolism were identified by analysing mutants affected in the PSI assembly/stability. The HCF101 is an ancient member of the P-loop ATPase superfamily and is distributed among the three domains of life (Lezhneva *et al.*, 2004). The APO protein belongs to a novel gene family restricted to plants (Amann *et al.*, 2004). Both proteins exert a crucial role in assembly or stability processes involving the cofactors [4Fe-4S] of PSI.

1.6. Goals of the project

The two basic objectives underlying this work were: (i) to improve the understanding of aspects involving biogenesis and function of PSI; (ii) to investigate PhQ biosynthesis in higher plants. To achieve these goals, a forward genetics approach was chosen using the model plant *Arabidopsis thaliana*. The project involved several steps:

1) Functional and molecular characterization of two *phylloquinone absence* (*pha*) mutant lines, *pha*1 and *pha*2, involved in PSI function and PhQ biosynthesis;
2) Molecular mapping of the defective alleles related to *pha*1 and *pha*2 and the identification of the respective lesions at the DNA level;
3) Characterization of two further mutated alleles, the T-DNA insertion lines *pha*3 and *pha*4, shown to belong to the same complementation group of *pha*1 and *pha*2;
4) Identification of the *PHYLLO* gene as a monocistronic transcriptional unit and characterization of its coding sequence;
5) Complementation of the *pha* mutations with a functional cDNA and the corresponding genomic locus using *Agrobacterium*-mediated transformation;
6) Characterization of the composite *PHYLLO* product and its role in the biosynthesis of PhQ;

7) Identification of the isochorismate synthase (ICS) function related to the PhQ biosynthesis by the genetic characterization of two *ICS* genes in the *Arabidopsis* genome;
8) Understanding the evolutionary dynamics of the *PHYLLO* and *ICS* genes in higher plants.

2.0. Materials and methods

2.1. Materials

2.1.1. Plant stocks

The *pha*1 mutant, formerly called *hcf*113 (Amann *et al.*, 2004), of the accession Wassilewskija, was obtained from a T-DNA collection (Feldmann, 1991). The *pha*2 plants, formerly *hcf*104, of the accession Columbia, was obtained from an EMS-induced mutant collection (Meurer *et al.*, 1996). *Pha*3 , *pha*4 , *ics*1 and *ics*2 mutants were obtained from the T-DNA insertion lines Salk_137597, Salk_039309, Salk_042603 and Salk_073287, respectively (Salk Institute, La Jolla, Ca). The insertion sites and the genotype of these lines were confirmed by PCR, nucleotide sequence and segregation analyses. Wild-type seeds of *Arabidopsis* ecotypes Columbia, Landsberg *erecta* and Wassilewskija, were obtained from plants propagated in the greenhouse.

2.1.2. Bacterial strains

The bacterial strains used were: *E. coli* DH5α (Bethesda Res. Lab., 1986) and *Agrobacterium tumefaciens* GV3101 (pMP90RK) (Koncz *et al*., 1994).

2.1.3. Cloning vectors

The cloning vectors used in this work were: pPCR-Script (Stratagene, Heidelberg, Germany); the plant binary expression vector pSEX001-VS (Reiss *et al*., 1996) and the pOL-DsRed vector (Mollier *et al.*, 2002).

2.1.4. Clones

The BAC clone T6L1 was purchased from the DNA Stock center of the *Arabidopsis* Biological Resource Center from the Ohio State University. It consists of an *Arabidopsis* genomic fragment cloned into the vector pBeloBAC11 (Choi *et al.*, 1995). The RAFL 09-32-

CO5 clone is a full-length cDNA of the splicing form 1 (see section 3.4.) obtained from the RIKEN company (Seki *et al.*, 2003).

2.1.5. Oligonucleotides

Table 1. Oligonucleotides used for PCR, RT-PCR, sequencing, hybridisation probes and PCR-based cloning. When not specified, the position of the annealing site of the oligonucleotides is relative to the *PHYLLO* locus. The numbers indicate the location in base pairs of the 5' to 3' interval of oligonucleotides relative to start codon of genes or cDNAs. In some primers used for PCR-based cloning, nucleotides not existing in the annealing site were introduced into the sequence of the primer and are indicated in lower case.

Name	Sequence (5' to 3')	Annealing site (bp)
MenD1_for	CCTTCTCACCAGCCATTGAT	(-52)-(-33)
MenD1_rev	TTGGGCAGAAATGAGAAGAAA	1759- 1739
MenD2	ACCATAGGCCGTCACCATAG	860- 841
MenD3	CTCATGGGGTTTTTGGGATT	645- 664
MenD4_for	AGCAGTTACATTGGCATGGA	1074- 1093
MenD4_rev	CAATAGCGTGAAACGCAAGA	3089- 3070
MenD5	GAGCATATGGGCTTTTTCACC	2148- 2168
MenD6	TTGACGTCAGCCTGATCTCTC	2197- 2177
MenD7_for	TGGATCAAGGTCTTCCCATC	2989- 3008
MenD7_rev	AACCCATCAATGCCACTAGC	4992- 4973
MenD8	GTTGTTGAAGCCAGTGAGGA	3288- 3307
MenD9	CAACATCAAACTCTATCAAATTCCTG	4044- 4019
MenD10_for	AAAGAGCTGATGTGGCCTGT	3885- 3904
MenD10_rev	AGCAAAGTGGGAAAGGAAAA	5559- 5540
MenD11	CCAATGCCTGTAGATGACCA	4586- 4567
MenD12	TGAGTGCAGAACTACCATGTCA	4923- 4944
MenD13	CAAAGCTTTAGCCCTTGGAG	541- 522
MenD14	ACCTGACGTTGTTTTCTCAAA	2381- 2401
MenD15	TGTCTGGACTTAGTTTCCTTTTCA	2094- 2117
MenD16	TCAAAATTCATCACAGCGAAG	15- 35 of *phyllo*
MenD17	TGCACTCGTCTTGGTTTGAC	2879- 2898
MenD18	GGTATGCCCTTTTGGAGAGA	2532- 2551
MenD19	AACTCTGACCGCTGGAAGAA	3175- 3156
MenD20	TCTGTGCTTGTGGTTTGGAA	1483- 1502
MenD21	ATCCTTGCATCCTGTTCAGC	1900- 1881
MenD23	AAGCGTTCTCTGCTACGTCTG	98- 78
MenD24	CCATTTGCATACATTTTGGTTG	4424- 4445
MenD25	CGGTGTTTGACCAATCAGG	5150- 5132
MenD26_for	CATCACAGCGAAGAAGATGC	23- 42 of *phyllo*
MenD26_rev	TTGAACTCCTTGAGCGACCT	212- 193 of *phyllo*
MenD27_for	AGCAGAGGATTGCAAGGAAA	2623- 2642 of *phyllo*
MenD27_rev	TTCGTCCCAACGTGAACATA	2815- 2796 of *phyllo*
MenD28_for	AACCGTCATTGTGGGAAGAG	4888- 4907 of *phyllo*
MenD28_rev	TCGAGATGTACAGCGTGACC	5074- 5055 of *phyllo*
MenD30	TGCCCAAATGTAGAGGATCA	1753- 1772
MenD31_for	GGCTTTTTCACCTTTTCTGG	2157- 2176
MenD31_rev	AGCTTGCAATGGAACTGACA	2377- 2358
MenD32	CCAGGGGCTACACAAAAGTA	2991- 2972
MenD34	CTGATGCCCATACAGCATTG	2864- 2845
MenD35	AGGTGAAAGCTTGCAATGGA	2384- 2365
MenD36	AAAGCATGCAAGACACGTTG	3058- 3039
MenD37	TGGTATCTTCTTTACCCTG	2203- 2221
MenD38	ACGTCAAGGAACTGAAAAG	991- 1009
MenD39_for	CATTGCAAGCTTTCACCTGA	2367- 2386
MenD39_rev	TGCCCATACAGCATTGATGT	2860- 2841
MenD40_for	CAATGCTGTATGGGCATCAG	2845- 2864
MenD40_rev	GCGATAGCAAGATGGGAAGA	3018- 2999
MenD41_for	CAGGAACTGCCGTTTCAAAT	3135- 3154
MenD41_rev	AGTTCAGGAGGACGATCTGC	3352- 3333
MenD42	ACCCTGCTTTTGTTGCAG	1016- 1034 of form 2
MenD43	TTGTATACTTTTCTGCAAC	2351- 2333
MenD100	TAAAGGCGTTGTGGGTTTTG	(-127)-(-108)
MenD100_rev	TGTTCAAATGGGAGAGAAGGA	188- 168

MenD101_for	TGATATTGTCCCCTTTGAAAAA	9247- 9268
MenD101	TGCGAGCTTGGAACTTGTTA	9766- 9785
MenD102	TCAATACAATTGAAATTTGAAACAA	9661- 9637
MenD103	ACTGTCGGTGTGATTATGGTTG	9766- 9745
MenD104_for	TCAAATGGGATAATGTCTGATGA	9218- 9240
MenD104_rev	GTCCTAGGGAATGCGAGCTT	9796- 9777
MenC1_for	TGGAATCTTCCGACTTCTTCC	5399- 5419
MenC1_rev	TCACACAGTTTGACTCTGCAA	6200- 6180
MenC2_for	CTTCTTTGCCAAGTCTCTGGA	6069- 6089
MenC2_rev	TCAGAATCAAGAAGGGCACA	6885- 6866
MenC3_for	TGGTCTGGAAATGGCTCTTC	6748- 6767
MenC3_rev	TGGTAAGCATACGCAGAGGA	7583- 7564
MenC4_for	GGTTTATTGGTGAACAGCTGTAAT	7318- 7341
MenC4_rev	AAATCCTTCAACAAATCCACTG	8166- 8145
MenC5_for	CATCCACAGAGCAGAAGCAA	8030- 8049
MenC5_rev	CACTTCGAATTTTCCGTGCT	8875- 8856
MenC6_for	CATGGCTTTGAGGTTTAGCA	8708- 8727
MenC6_rev	GCCTGAAACAGGTCCAATCA	8162- 8140
MenC7	CCTTCAACAAATCCACTGTAAGG	8162- 8140
MenC8	CTCCCTGATACCACGACAGC	8834- 8815
MenC9	TTCTCAGAAGCTCTTGTTAGCACT	9530- 9553
MenC10	TTCAACTGGAATTGGGTCAA	10445-10426
AtRT1-for	ATGGTGACGGCCTATGGTT	843- 861
AtRT1_rev	ATGACTGCTGGTTTGAGGGA	3123- 3104
AtRT2_for	TACTTTTTCGTTCCTCAG	918- 935
RFPmen_for	CCTCGTTACAGTTCTAGAAAATCCA	52- 76
Del2_for	AAGAAACATTGAGGTCGCTCA	182- 202
Del2_rev	TCTCTCGATTCAGCATGTTGTT	1086- 1065
113dCAP_for	CACTTTTGGTTCATTTTCTTGCAGGTgTCT	1714- 1742
113dCAP_rev	TGCAGGAAACAGTTTGATGAAcAGACAAAT	1940- 1911
104CAPS_for	TTTGGAAGCTATTCTATAGATCGAA	8787- 8811
104CAPS_rev	CAGTTTAAGTCCGGGGAACC	9031- 9012
104mut	ATATAATGGAATTTGGAAGCTATTCT	8776- 8801
Spemen_for	GGACCTGTTTCAGGCCTTTA	9609- 9628
Spemen_rev	CGGTGGCGGCCGCTaCTAGtCCC	Specific for pPCR-Script vector
Ecomen_for	AGCCATTGATaTCACAATCTCAA	(-42)- (-21)
Ecomen_rev	TGCTACGTCTGATGGATTTTCTA	66- 88
locEco	CTTCgTCgACCAGCCATTGATaTCACAATC	(-51)-(-32)
Cor2	TCAATTTTTCAGGTTTCTTCTCA	834- 856 of form 1
MenCor	GATTGTATACTTTTCcAGGGTAAAGAAG	1035- 1009 of form 1
MenGFP1_for	TCAAAATTCAgTCgACAGCGAAGAAGAT	15- 40 of form 1
MenRFP3_rev	CATATGgTCgAcTCTCAAGTGGATGTATCT	956- 928 of form 1
TcDNAmen1	GGGTCTTGCGAAGGATAGTG	pSEX, 35S promoter
TcDNAmen1_rev	AAACCTCCTCGGATTCCATT	pSEX, 35S promoter
T6L1-2_for	TACCCAACAGCAGCAAGATG	86239-86258 of BAC T6L1
T6L1-2_rev	CCTGGTTCAAACCACGACTT	87111-87092 of BAC T6L1
F14K14-1_for	TGGTGAAAGAAGAGGTCATTTTG	3026- 3004 of BAC T6L1
F14K14-1_rev	TGTGCGCTAAGAAAGAAGAAGA	2930- 2951 of BAC T6L1
ICS1C_for	TGCAACTATTGCTTGGGATG	1648- 1667
ICS1B_rev	CCTTCACGCTGTCACACAAA	2130- 2111
ICS2B_for	TTAAGCCACGTGGAGCTCTT	259- 278
ICS2B_rev	TTTCTTCCAGATAAACCCATGAA	621- 599
LBb1	GCGTGGACCGCTTGCTGCAACT	left border of the T-DNA
Ac_for	TCCTAGTATTGTGGGTCGTCCTCG	*Arabidopsis* actine3 gene
Ac_rev	GCTCATTCTGTCGGCGATTCCAGG	*Arabidopsis* actine3 gene
A/B_for	CTCCTGCTACTCAGCCTAGAGCCTTGAGCA	*psa*A/B operon
A/B_rev	TCATCATGACTCTCGAAGTCATGTGCGG	*psa*A/B operon

2.1.6. Chemicals and enzymes

Chemicals and materials used in this work were purchased from the following companies: Applichem (Darmstadt, Germany), Biozym (Oldendorf, Germany), Fluka (Steinhein, Germany), ICN Biomedicals GmbH (Eschwege, Germany), Merck (Darmstadt, Germany), Pharmacia (Uppsala, Sweden), Roth (Karlsruhe, Germany), Serva (Heidelberg, Germany), Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany), and USB (Cleveland,

USA). Other materials were obtained from Biomol (Hamburg, Germany), Eppendorf (Hamburg, Germany), Greiner Bio-One GmbH (Frickenhausen, Germany), Millipore (Eschborn, Germany), Pall Bio Support Division (Dreieich, Germany), Qiagen (Hilden, Germany), and Schleicher and Schüll (Dassel, Germany).

Enzymes were obtained from Invitrogen (Karlsruhe, Germany), MBI Fermentas (St. Leon-Rot, Germany), New England Biolabs (Frankfurt/Main, Germany), Promega (Mannheim, Germany), Qiagen (Hilden; Germany), Roche Diagnostics (Mannheim, Germany), and Stratagene (Heidelberg, Germany).

Radioactive nucleotides were purchased from Amersham Biosciences Europe GmbH (Freiburg, Germany).

2.1.7. Media, solutions, buffers and antibiotics

Media utilized for plant cultivation and *Agrobacterium*-mediated plant transformation are described below:

MS-medium:
1x MS-salts (Murashige and Skoog, 1962)
1,5% sucrose
2,5 mM MES-NaOH, pH 5,7
0,3% gelrite

Infiltration medium:
5% sucrose
0,05% Silvet L-77 (Clough and Bent, 1998)

YEB medium:
5 g/l beef extract
5 g/l bacteriological peptone
1 g/l yeast extract
5 g/l sucrose
2 mM $MgCl_2$
supplemented with 20 g/l agar for solid medium

Further media and antibiotics used for bacterial growth and solutions, reagents and buffers utilized for general molecular biology methods and other techniques have been described in Sambrook *et al*., (1989).

2.1.8. Antibodies

Table 2. Antibodies used for Western analysis.

Protein or protein complex	Subunit	Source of the antibody
Photosystem I	PsaA/B	R. Nechushtai (Hebrew University, Jerusalem, Israel)
	PsaE	R. Herrmann
Photosystem II	PsbD (D2)	J. Mullet
	PsbE	R. Herrmann
Cytochrome b_6f complex	Cytochrome b_6 (PetB)	R. Berzborn
	Rieske (PetC)	R. Herrmann
ATP synthase	α subunit (CF_1 α, AtpA)	R. Berzborn
NAD(P)H-dehydrogenase	NdhH	R. Herrmann

2.2. Methods

2.2.1. Molecular biology methods

2.2.1.1. General methods

The *E.coli* strain DH5α, used in this work for transformation with recombinant plamids, was grown in LB medium or on LB-agar plates at 37°C with addition of 70 μg/ml ampicillin. Preparation of plasmidial DNA for sequencing or cloning was performed with QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) according to instructions of the manufacturer. Preparation of DNA from the BAC clone T6L1 was performed with QIAfilter Plasmid Maxi Kit (Qiagen, Hilden, Germany). Standard procedures involving manipulation of nucleic acids

like precipitation, gel electrophoresis, staining and quantification of nucleic acids are described in Sambrook *et al.* (1989).

The DNA recombinant methods involving preparation of vectors and fragments, ligation reaction, preparation of competent cells, heat-shock transformation of *E. coli* and final selection of recombinants are described in Sambrook *et al*. (1989). The enzymes for modification of nucleic acids like DNA ligases, DNA phosphatases, restriction enzymes and RNases, were used during these procedures with buffers and conditions specified by the manufacturers. DNA fragments were purified from agarose gels using the QIAEX®-II Gel extraction kit and PCR products were purified with the QIAquick® PCR purification Kit (Qiagen, Hilden). For PCR-based cloning the *pfuTurbo* DNA polymerase from Stratagene (Heidelberg, Germany) was used according to manufacturers instructions.

For PCR amplifications of DNA templates 0,5 mM oligonucleotide primers, 0,2 mM dNTPs, 5 mM $MgCl_2$ and 1U Taq polymerase were used in a 20 µl final reaction volume. An initial denaturation step was performed at 95°C for 5 minutes, and afterwards 30 cycles of denaturation (95°C for 20 sec), annealing (54 - 62°C for 20 sec) and extension (72°C for 1 minute per 1 kb DNA) were performed followed by a final extension step at 72°C for 5 minutes. Reverse transcription was performed with 0,1 µg mRNA using SuperScript II RNase H^- Reverse Transcriptase (Invitrogen, Karlsruhe, Germany) and hexanucleotides (Roche Molecular Biochemicals, Mannheim, Germany) as random primers according to the manufacturers instructions. Quantitative two-step real-time RT-PCR for *pha*3 and wild-type plants was carried out essentially as described (Lezhneva and Meurer, 2004) using the primer combinations MenD27_for-MenD27_rev and MenD28_for-MenD28_rev, annealing in the middle and 3' parts of the *phyllo* transcript, respectively.

2.2.1.2. RNA gel blot analysis of the *phyllo* transcript

The RNA samples from *Arabidopsis* leaves were first denaturated through incubation with 30% glyoxal (McMaster and Carmichael, 1977). For hybridisation with the *phyllo* probes 6,3 µg of $polyA^+$ mRNA was loaded on 1,2% agarose gel. For hybridisation with a probe specific for the plastid *psa*A/B operon 8,0 µg total RNA were used. The RNA was then electrophoretically separated and transferred by capillarity onto a Biodyne A nitrocellulose membrane (0,45 µm; Schleicher & Schuell BioScience GmbH, Dassel, Germany) in 20x SSC

buffer. RNA was fixed to the membrane by exposure to a temperature of 80°C (Thomas, 1980).

Three probes were amplified from the RAFL 09-32-CO5 cDNA (form 1) with the specific primer combinations D26_for-D26_rev, D27_for-D27_rev, and D28_for-D28_rev that anneal in the 5', middle and 3'-parts of the *phyllo* transcript, respectively. One further probe, specific for the *psa*A/B operon, was amplified with the primer combinations A/B_for-A/B_rev. These probes were labelled using the Random Primed DNA Labelling Kit (Roche Molecular Biochemicals, Mannheim, Germany) according to the method of Feinberg and Vogelstein (1983) and used for overnight hybridisations with the specific filters. All hybridisations were performed at 62°C as described (Thomas, 1980), preceded by prehybridisations carried out in the same buffer for at least two hours. After hybridisation filters were repeatedly washed for 30 minutes, wrapped with plastic foils and analysed by phosphorimaging (BASIII Fuji Bio Imaging plates and BAS2000 software package and the AIDA software package v3.25 beta; Raytest, Straubenhardt, Germany).

2.2.1.3. Western blot analysis of thylakoid protein complexes

Material from leaves of three-week-old plants was homogenized after grounding to fine powder by applying two volumes of homogenisation buffer (10 mM EDTA, 2 mM EGTA, 50 mM Tris.HCl; pH 8,0, 10 mM DTE). The homogenate was filtered through two Miracloth layers (100 µm, Calbiochem, La Jolla, USA) and centrifuged for 10 minutes at 10.000 rpm at 4°C (SS 34-rotor, Sorvall). The pellet corresponding to the membrane protein fraction was resuspended in one volume of sample buffer (100 mM Na_2CO_3, 10% w/v saccharose, 50 mM DTE) (Meurer *et al.*, 1996).

The proteins were quantified according to Bradford (1976). Prior to the SDS-PAGE, samples were prepared by heating for 5 minutes at 80°C in 2% SDS mixed with 1/10 volume glycerol/dye solution. The electrophoretic separation was carried out on 10% SDS/polyacrylamide gels at 30 mA for 12-15 hours at RT. For immunodetection, proteins were transferred to polyvinylidene difluoride (PVDF) membranes (Amersham Buchler, Braunschweig, Germany) using a semi-dry electroblotting device (Peqlab, Erlangen, Germany), incubated with specific antisera and visualized using the enhanced

chemioluminescence technique (Amersham Buchler, Braunschweig, Germany) (Meurer *et al.*, 1996).

2.2.2. *Arabidopsis* methods

2.2.2.1. Plant growth, seed sterilization

Propagation of plants containing the lethal *pha* mutation occurred via heterozygous offspring grown on soil. For analysis of homozygous mutants the plants were grown on sterile MS-medium. For this purpose, seeds were surface-sterilized in a solution composed of 32% (v/v) "bleach" and 0,8% N-laurylsarcosinate for 3 minutes and washed three times with sterilized water. Seeds were then distributed on Petri dishes (94 x 16 mm) containing solid MS-media supplemented with 1,4% sucrose and kept for 2 days at 4°C to synchronize germination. Seedlings were grown under continuous light at a photon flux density of 20-40 μE m^{-2} sec^{-1} and at a constant temperature of 21°C. For feeding experiments with NA hypocotyls of three-week-old plants were cut off at their base and plants were transferred and kept for five days either on MS medium or on medium containing 1,0 mM NA (ABCR, Karlsruhe, Germany).

2.2.2.2. Rapid isolation of plant DNA for PCR

DNA was isolated from *Arabidopsis* leaves homogenized in 400 μl extraction buffer (0,2 M Tris/HCl, pH 7,5; 0,25 M NaCl; 0,025 M EDTA and 0,5% (w/v) SDS) in racks containing 96 well-microtubes of 1,2 ml (Qiagen, Hilden, Germany) by disruption for 3 minutes with 3 mm tungsten carbide beads in the Mixer Mill MM 300 (Retsch, Haan, Germany) according to the manufacturer's instructions. The extract was centrifuged for 30 minutes at 3.000 units of relative centrifugal force (RCF) at RT (Centrifuge 4-15C, Sigma, Osterode am Harz, Germany). 300 μl supernatant were transferred to a new rack of 96 well-microtubes and 300 μl isopropanol was added. The mixture was incubated at RT for 2 minutes and centrifuged for 15 minutes at 3.000 RCF at RT. The supernatant was discarded, and the pellet was washed with 70% ethanol and centrifuged for 15 minutes. Finally, the pellet was air-dried and resuspended in 200 μl H_2O. For PCR amplification 6 μl DNA were used for a 20 μl final reaction volume.

2.2.2.3. Isolation of total RNA

Isolation of total RNA was carried out from leaf material in 8 ml homogenisation solution (0,33 M sorbitol; 0,2 Tris-NaOH, pH 9,0; 0,3 M NaCl; 10 mM EDTA; 10 mM EGTA; 2% SDS) mixed with 4 ml phenol and 4 ml chloroform at 40°C. Selective precipitation of RNA was achieved with 2 M LiCl (Barlow *et al.*, 1963). Extraction of polyA+ mRNA mRNA was performed according to the Dynabeads method (Dynal, Oslo, Norway).

2.2.3. Spectroscopic and fluorimetric methods

2.2.3.1. Chlorophyll *a* fluorescence analyses

Chlorophyll *a* fluorescence was measured by a pulse amplitude–modulated fluorometer (PAM101; Walz, Effeltrich, Germany). This was equipped with a data acquisition system (PDA-100; Walz) used in combination with a personal computer running a Wincontrol version 1,72 software (Walz, Effeltrich, Germany) for data collection. The following settings were used for the PAM101 unit: light intensity, 4; gain, 6; damping, 9. After induction, saturating pulses of 4.000 μE m^{-2} sec^{-1} light intensity and 1 sec duration were applied in 30 sec intervals to estimate quenching parameters (Meurer *et al*., 1996). For analysis of photochemical (qP) and non-photochemical (NPQ) chlorophyll *a* fluorescence quenching at RT, the actinic light intensity was 20 μE m^{-2} sec^{-1} and the intensity of the saturating light pulses (1 sec, 20 sec intervals) used for detection of the quenching parameters during induction was 4.000 μE m^{-2} sec^{-1}. The analyses were performed using plants grown under the same conditions.

2.2.3.2. Light-induced changes of the P700 redox state

Light-induced changes of the P700 redox state were recorded by absorbance changes at 830 nm with the same PAM system described above equipped with a dual wavelength emitter-detector unit (Meurer *et al*., 1996). Multiple turn-over flashes of 80 μsec were induced by a Xenon lamp (Walz, Effeltrich, Germany) and saturating light pulses of 1 sec were applied by halogen lamps.

2.2.4. Genetic methods

2.2.4.1. Mutant selection

Two to three-week-old plants grown on MS medium were analysed by a chlorophyll fluorescence video imaging system (FluorCam690M, Photon Systems Instruments, Brno, Czech Republic). Mutant plants were clearly distinguishable from wild-type plants by the presence of a typical *hcf* phenotype characterized by re-emission of the absorbed light in form of red fluorescence (Meurer *et al.*, 1996).

2.2.4.2. High-resolution genetic mapping of the *pha* mutations

A mapping population segregating the *pha*1 mutation was generated by pollinating emasculated flowers of the accession Landsberg *erecta* with Wassilewskija plants heterozygous for the *pha*1 mutation. Back-crossed plants segregating the *pha*2 mutation were generated by pollinating emasculated flowers of the accession Landsberg *erecta* with Columbia plants heterozygous for the *pha*2 mutation. F_2 recombinant plants segregating the mutation were grown on soil and their progeny was screened for the presence of the *hcf* phenotype. DNA was extracted from each of the F_3 plants showing a mutant phenotype and analysed with Simple Sequence Length Polymorphism (SSLP) markers (Bell and Ecker, 1994) covering the entire genome. After this phase of rough mapping, the SSLP marker F14K14-1 (primers F14K14-1_for and F14K14-1_rev) and a cleaved amplified polymorphic sequence (CAPS) marker (Konieczny and Ausubel, 1993), called T6L1-2 (primers T6L1-2_for and T6L1-2_rev, product cleavage with the enzyme *Pst*I), were used in the close vicinity of the defective locus for fine localization of the mutated gene.

2.2.4.3. Complementation analysis

For complementation analysis, a fragment containing the whole cDNA corresponding to the form 1 of *phyllo* was prepared from the RAFL 09-32-C05 clone by sequential digestion with *Bgl*I and fill-in with the Klenow subunit of DNA polymerase and subsequent subcloning into the vector pPCR-Script (Stratagene, Heidelberg, Germany) within the *Sma*I restriction site, generating the pB clone. Restriction sites for *EcoR*V and *Spe*I were introduced into pB in

positions flanking the 5' and 3' regions of the form 1 cDNA by PCR-based cloning of fragments generated with the primer combinations Ecomen_for-Ecomen_rev and Spemen_for-Spemen_rev, respectively. The resulting clone, pBSpe, was digested with *Eco*RV and *Spe*I and the fragment was cloned into the plant binary transformation vector pSEX001-VS (Reiss *et al*., 1996) within compatible *Sma*I and *Xba*I restriction sites generating the pbinC05 clone.

To generate a full-length cDNA for complementation of *pha* mutants, the *Xba*I and *Bst*1107I cleaved fragment of pbinC05 was replaced by a PCR product generated with the primer combinations RFPmen_for and MenCor. A guanine was introduced in this latter primer within the sequence corresponding to the stop codon of form 1 (Table 1). The resulting pCor plasmid contains a full-length cDNA encoding a single reading frame of the *PHYLLO* locus (form 4).

A fragment containing the *PHYLLO* locus was generated by cleavage of the BAC clone T6L1 with *Hpa*I and *Spe*I restriction enzymes. This fragment was cloned within the *Eco*RV and *Spe*I restriction sites of the clone pBSpe, previously prepared with the same restriction enzymes by releasing the fragment correspondent to the cDNA form 1 used for the cloning of pbinC05. The resulted clone pBB, was subcloned by introduction by PCR-based cloning (primers locEco and Ecomen_rev) of a restriction site for the enzyme *Eco*RV in the position flanking the 5' coding region of the *PHYLLO* locus. The resulting clone, pBBEco, was then digested with *Eco*RV and *Spe*I and the fragment was cloned into the plant binary transformation vector pSEX001-VS, within compatible *Sma*I and *Xba*I restriction sites, generating the pG10 clone.

The constructs obtained were introduced into *Agrobacterium tumefaciens* GV3101 (pMP90RK) (Koncz *et al*., 1994) and transformed into progenies of mutant segregants using the floral dip method (Clough and Bent, 1998). The transformed offspring was selected with 10 mg/L sulfadiazine and genotyped with specific CAPS markers for the *pha*1 and *pha*2 mutations (section 3.2.) or with combinations of primers annealing with the T-DNA and flanking genomic regions for the *pha*3 and *pha*4 mutants (section 3.5.). The presence of the transformation construct in resistant lines was confirmed by primers annealing either in the internal part of the vector or in the inserted fragment. Homozygosity of complemented lines

was also checked in the progeny and segregation of the wild-type phenotype with sulfadiazine resistance was observed to confirm the success of the transformation procedures.

2.2.5. High performance liquid chromatography

HPLC analysis was performed by Dr. Jon Falk in the Botanisches Institut of Christian-Albrechts-Universität zu Kiel. Material of two to three-weeks-old seedlings grown on MS medium or leaves from five to six-weeks-old plants grown on soil were analysed. 100-500 mg of ground material from leaves were extracted with 2 volumes n-heptane at 20°C overnight and re-extracted with 100 µl n-heptane. The combined extract was analysed using a LiChrosphere Si 100 column (10 x 250 mm, 5µm) with n-heptane/2-propanol (99,05+0,05) as eluant at a flow rate of 1,0 mL/minute. PhQ was detected by its absorbance at 270 nm using a UV detector (484 Tunable Absorbance Detector, Waters Instruments, Eschborn, Germany). Individual peaks were verified by their absorbance spectra recorded with a Photodiode Array Detector (SPD-M10Avp, Shimadzu, Duisburg, Germany).

2.2.6. Subcellular localization of the PHYLLO protein by fluorescence imaging

A PCR product was amplified from the RAFL 09-32-CO5 using the primers MenGFP1_for and MenRFP3_rev and cloned into a *Sal*I restriction site of the RFP expression vector pOL-DsRed (Mollier *et al.*, 2002). This produced a traslational fusion of the the MenF module of PHYLLO with the *Discosoma* sp. (DsRed) protein (Jach *et al.*, 2001). The resulting construct and a control corresponding to the empty vector pOL-Ds were expressed in tobacco (*Nicotiana tabacum*) protoplasts according to the polyethylene glycol protocol (Lyznik *et al.*, 1991). Images were captured using a fluorescence microscope equipped with a digital camera (Axioplan; Zeiss, Jena, Germany).

2.2.7. Sequence analyses

Nucleotide sequences were determined using the ABI377 system (Applied Biosystems, Foster City, CA) and chromatograms were analysed using the Sequencher 3.0 software (Gene Codes Corp., Ann Arbor, MI). General procedures for computational analysis of DNA and protein sequences are described in Gross *et al.* (2003). Database searches for homologous genes and proteins were performed using the BLAST algorithm (Altschul *et al.,* 1997) at the

server of the NCBI (http://www.ncbi.nlm.nih.gov/Blast). Global multiple alignments of sequences were performed using the ClustalW (Thompson *et al.*, 1994) and DIALIGN (Morgenstern *et al.*, 1998) algorithms and then edited with the GeneDoc program (Nicholas and Nicholas, 1997). For alignments of proteins the substitution matrix BLOSUM62 (Henikoff and Henikoff, 1992) was used as a parameter for conservation of amino acids between two sequences. For local alignments of two sequences the BLAST 2 version of the original BLAST algorithm was used (Tatusova and Madden, 1999). Analysis of functional patterns of proteins was performed by searches against the Conserved Domain Database (CDD) (Marchler-Bauer and Bryant, 2004). The algorithms PCLR Chloroplast Localization Prediction (Schein *et al.*, 2001), ChloroP (Emanuelsson *et al.*, 2000), iPSORT (Bannai *et al.*, 2002), and Predotar (Small *et al.*, 2004) were used to predict putative transit peptides in the Men homologs of *Arabidopsis*. Primers were designed with the assistance of the Primer3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). The tools of the BCM Search Launcher: Sequence Utilities (http://searchlauncher.bcm.tmc.edu/seq-util/seq-util.html) were used for general manipulation of sequences.

3. Results

3. 1. Characterization of the *pha* phenotype

The *pha* phenotype is characteristic of all mutant plants analysed in this work. It is related to the *pha*1, *pha*2, *pha*3, and *pha*4 mutants associated to the *PHYLLO* locus, as well as to the *ics*1 and *ics*2 double mutant plants. The characteristics of the *pha* phenotype are described below.

3.1.1. General phenotype

The mutant plants present a recessive lesion that is lethal to the plant in the homozygous state, indicating impairment in the function of an essential gene. Accordingly, these mutants can not grow photoautotrophically and are only able to reach the seedling state when maintained on sucrose supplemented MS medium, typical of mutants impaired in photosynthesis (Meurer *et al.*, 1996). When growing under sterile heterotrophic conditions, the seedlings are clearly distinct from the wild-type plants by their pale and debilitated appearance.

3.1.2. Phylloquinone absence

All the *pha*1, *pha*2, *pha*3, and *pha*4 mutants, the *ICS*1 and *ICS*2 double mutant plants were systematically analysed by HPLC during this work and always revealed a complete lack PhQ (Fig. 4A, Table 3). The absence of vitamin K_1 is not a secondary effect related to the failure of plants to accumulate PSI, since other *Arabidopsis* mutants, like *hcf*101 and *hcf*145, with PSI amounts below 5% (Lezhneva et al., 2004; Lezhneva and Meurer, 2004), possess normal levels of PhQ (Table 3). This differentiates the *pha* plants from the other *hcf* and PSI mutants.

3.1.3. A *hcf* phenotype associated with Photosystem I lesions

Upon illumination the mutants are not able to efficiently quench the light energy, consequently the energy trapped in the chlorophyll *a* is re-emitted in form of fluorescence. This *hcf* phenotype is typically observed in photosynthetic mutants (Meurer *et al.*, 1996). PSI

activity, measured by the redox kinetics of the P700 chlorophyll during absorbance changes at 820 nm, indicated a basal activity of about 5-15% of wild-type levels (Fig. 4B). Chlorophyll fluorescence analysis revealed that the potential yield of PSII, showed by the ratio of variable to maximum fluorescence, is reduced to about 75% of wild-type (Fig. 4C), demonstrating that this photosystem is only moderately affected in the *pha* mutants. This is also observed for mutants not affected in PSII (Lezhneva *et al.*, 2004; Lezhneva and Meurer, 2004; Amann *et al.*, 2004), indicating that the limiting step of the linear electron transport in the *pha* mutants lies behind this photosystem, in accordance with specific photosynthetic lesions related to PSI function.

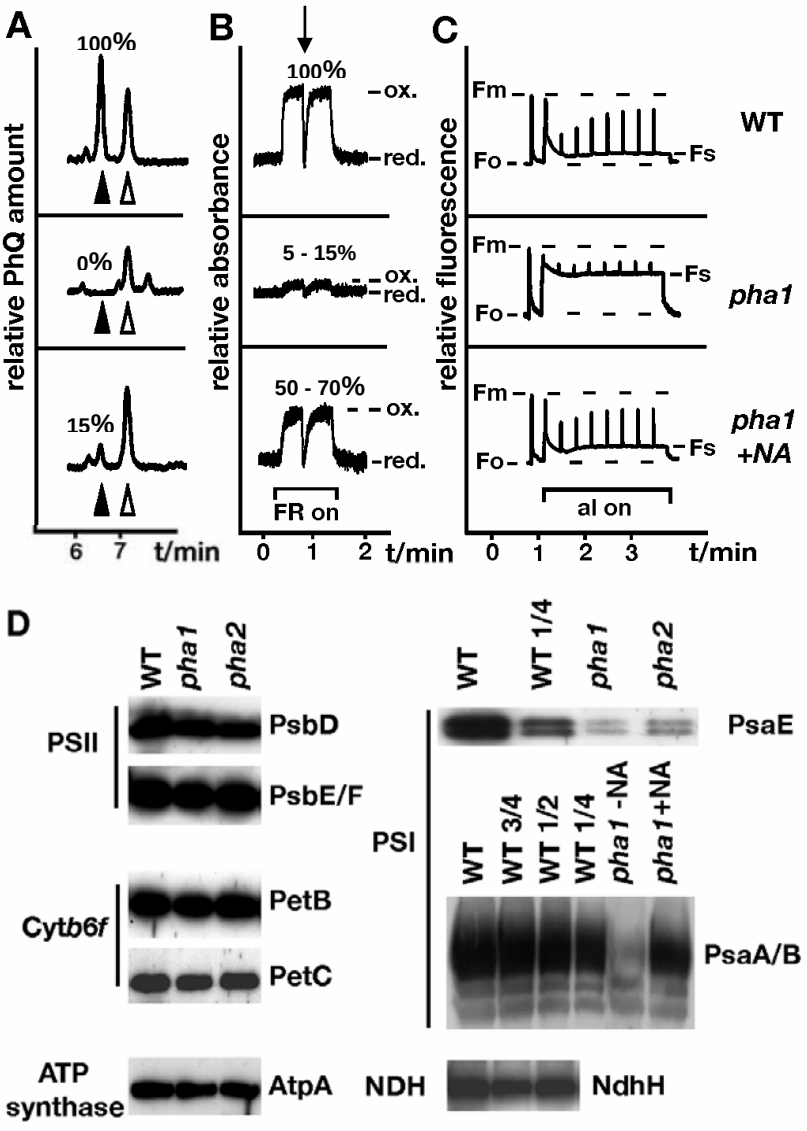


Figure 4. Characterization of the *pha* mutants. (**A**) HPLC analysis of wild-type, *pha*1, and NA-supplemented *pha*1 mutant plants. Black triangles indicate the PhQ retention peak, the white triangles the plastoquinone peak. The PhQ amounts are expressed as percentage relative to wild-type (**B**) PSI activity induced by 720 nm far-red light (FR)- and actinic light (al) pulse-dependent absorbance changes of the reaction centre chlorophyll P700 at 830 nm. First, P700 was completely reduced (red.) in the dark. FR light was applied to fully oxidize (ox.) P700. A saturating pulse (arrow) of white light was given to follow the PSII-dependent re-reduction of P700. (**C**) Chlorophyll *a* fluorescence at RT reflects PSII activity. The minimal fluorescence yield (F_o) of dark-adapted *pha*1 mutant and wild-type plants was excited by a pulsed red measuring beam of low intensity. Leaves were exposed to 800 ms saturating light pulses to determine the maximum fluorescence yield (F_m). The fluorescence was then measured upon switching on continuous actinic light and additional superimposed saturating light pulses. NA-fed mutant plants showed fluorescence characteristics similar to those of wild-type indicating recovery of photosynthesis. Fs, steady state fluorescence. Equivalent results shown for *pha*1 in panels **A**, **B** and **C** were obtained for all *pha* mutants, *pha* crosses, and the *ics*1/*ics*2 double-knockout plants. (**D**) Representative immunoblot analysis from five thylakoid membrane complexes: PSII, cytochrome b_6f complex, ATP synthase, NAD(P)H-dependent dehydrogenase (NDH), and PSI. Eight µg of thylakoid membrane proteins from three-week-old wild-type and *pha* mutant plants were loaded on the gels for analysis. Thylakoid complexes are indicated at the left and corresponding proteins are labeled at the right. Serial dilutions of wild-type membranes were used to quantify PSI polypeptides. +NA, plants supplemented with NA; –NA, without NA.

3.1.4. Specific impairment of Photosystem I complex accumulation

Immunoblot analysis of mutant proteins showed a reduced accumulation of the PSI core complex reaching approximately 5-15% of wild-type levels. The other thylakoid membrane complexes, i.e. ATP synthase, cytochrome b_6f complex, NAD(P)H-dependent dehydrogenase and PSII, however, accumulated at almost normal levels in these mutants (Fig. 4D). The present data combined with the spectroscopic profile described above, unequivocally indicate a specific reduction of PSI stability in the *pha* plants.

3.1.5. Recovery of the phylloquinone content and Photosystem I activity after 1,4-dihydroxy-2-naphthoate feeding

When grown in the presence of the metabolic precursor of vitamin K_1, 1,4-dihydroxy-2-naphthoate (NA) (Fig. 3), all mutants analysed in this work typically show a partial recover of the PhQ content to about 0,5 µg/g fresh weight (Table 3), corresponding to 15% wild-type levels (Fig. 4A). The mutants greened upon the treatment (Fig. 5) and reestablished the PSI activity and accumulation to 50-70% of wild-type levels (Fig. 4B-D). The partial recovery of PhQ content after NA feeding indicates that the *pha* mutants are directly affected in the first

steps of the PhQ biosynthetic pathway, prior to the MenA enzymatic activity (Fig. 3).

Table 3. PhQ content of wild type and *Arabidopsis* PSI mutant plants. Values are expressed in µg PhQ/g fresh weight of plant material. -NA, without 1,4-dihydroxy-2-naphthoate; +NA, plants grown on 1 mM 1,4-dihydroxy-2-naphthoate. *Pha*3c and *pha*4c, complemented lines of the *pha*3 and *pha*4 mutants with the form 4, respectively. *pha*3G10-15,-55,-69,-87, complemented lines of *pha*3 mutant with the genomic locus. ND, not determined. All experiments are means ± standard deviations (SD) of at least three independent measurements.

Lines	-NA	+NA
Wild type	3,04 ± 0,52	5,15 ± 0,96
***hcf*145**	3,21 ± 0,42	5,36 ± 0,36
***hcf*101**	3,22 ± 0,35	5,63 ± 0,06
***pha*1**	0,0	0,41 ± 0,10
***pha*2**	0,0	0,59 ± 0,08
***pha*3**	0,0	0,48 ± 0,13
***pha*4**	0,0	ND
***pha*3c**	0,55 ± 0,11	ND
***pha*4c**	0,72 ± 0,16	ND
***pha*3G10-15**	0,61± 0,28	ND
***pha*3G10-55**	0,53± 0,27	ND
***pha*3G10-69**	0,66± 0,26	ND
***pha*3G10-87**	0,515± 0,28	ND
***ics*1/*ics*1; *ics*2/*ics*2**	0,0	ND
***ics*1/*ics*1; *ICS*2/*ics*2**	0,60 ± 0,11	ND

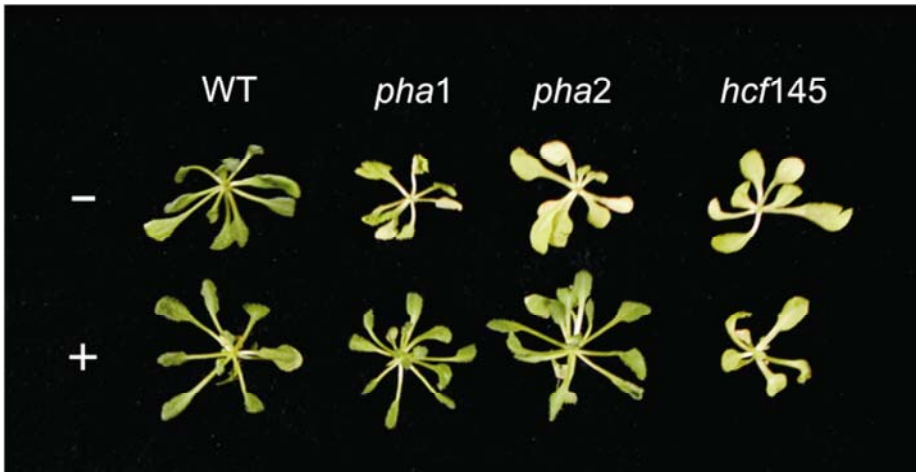


Figure 5. 1,4-dihydroxy-2-naphthoate feeding of *Arabidopsis* plants. The wild-type (WT) plant and the *pha1*, *pha*2, *hcf*145 mutants were grown for approximately 2 weeks and then transferred to a 1,0 mM 1,4-dihydroxy-2-naphthoate supplemented medium (+) or kept in normal MS medium without NA (-). After five days photos were taken. The *pha*1 and *pha*2 mutants greened upon the treatment and partially recover the photosynthetic parameters. The control *hcf*145 mutants remained pale.

3.2. Localization of the *pha* mutations into the *PHYLLO* locus

Initial efforts were concentrated on the localization of *pha*1 and *pha*2 mutations mapped to the lower part of the chromosome I. Two specific molecular markers (F14K14-1 and T6L1-2), related to the BAC clone T6L1, detected no recombinants segregating with the *pha* phenotype, therefore, indicating a localization of the *pha*1 mutation somewhere in the middle of T6L1. Fine sequence analysis revealed that the intermediate region of this genomic clone harbours a 9.788 bp interval (*PHYLLO* locus, starting on the position 25.900.613 of the chromosome I), surprisingly, showing homology to four eubacterial *men* genes, *men*F, *men*D, *men*C, and, *men*H (Fig. 6; Fig. 7A) related to the first steps of PhQ biosynthesis. Nucleotide sequence analysis of the *pha*1 and *pha*2 mutants and wild-type uncovered the genetic defects precisely assigned to the *PHYLLO* locus (Fig. 7A). An insertion of two nucleotides (CT) in the *pha*1 mutant in the region homologous with *men*F causes a change in the reading frame and a premature stop codon (Fig.7C). The resulting mutation was confirmed by a designed derived cleaved amplified polymorphic sequence (dCAPS) marker (*pha*1dCAPS, primers 113dCAP_for and 113dCAP_rev), which generates a *BsmA*I restriction fragment of 177 bp for the mutant, polymorphic to that of 212 bp verified for the wild-type (Fig. 7D). In the *pha*2 mutant, an exchange of a guanine to an adenine (transition) in the 3'-part of the region related to *men*H produced an in-frame stop codon (Fig. 7C). Again, the point mutation could be confirmed by a designed CAPS marker (*pha*2CAPS, primers 104CAPS_for and 104CAPS_rev), which generates a *Rsa*I restriction fragment of 270 bp for the mutant, polymorphic to that of 234 bp verified for the wild-type (Fig. 7E).

Furthermore, two additional T-DNA mutant alleles, Salk_137597, called *pha*3, inserted at the 5' region of the *PHYLLO* locus, and Salk_039309, called *pha*4, inserted at the 3' region of this locus were also included in this study (Fig. 7A). In the homozygous form, the latter two mutants exhibited all the characteristics related to the *pha* phenotype, reinforcing the essential function of the *PHYLLO* locus in vitamin K_1 biosynthesis.

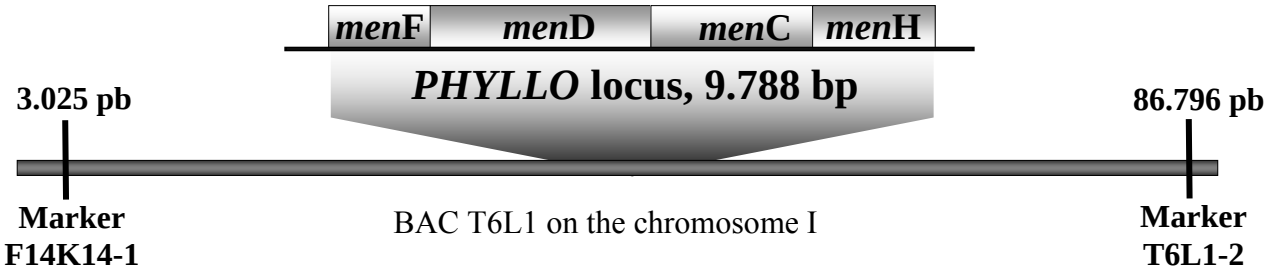


Figure 6. Fine mapping of the *pha*1 mutant inside the BAC clone T6L1. The position of the molecular markers F14K14-1 and T6L1-2 used in the final part of the mapping procedures are indicated in base pairs. The results uncovered the *PHYLLO* locus presenting homology to four Men proteins that are shown in the sequence of appearance in projection above the locus.

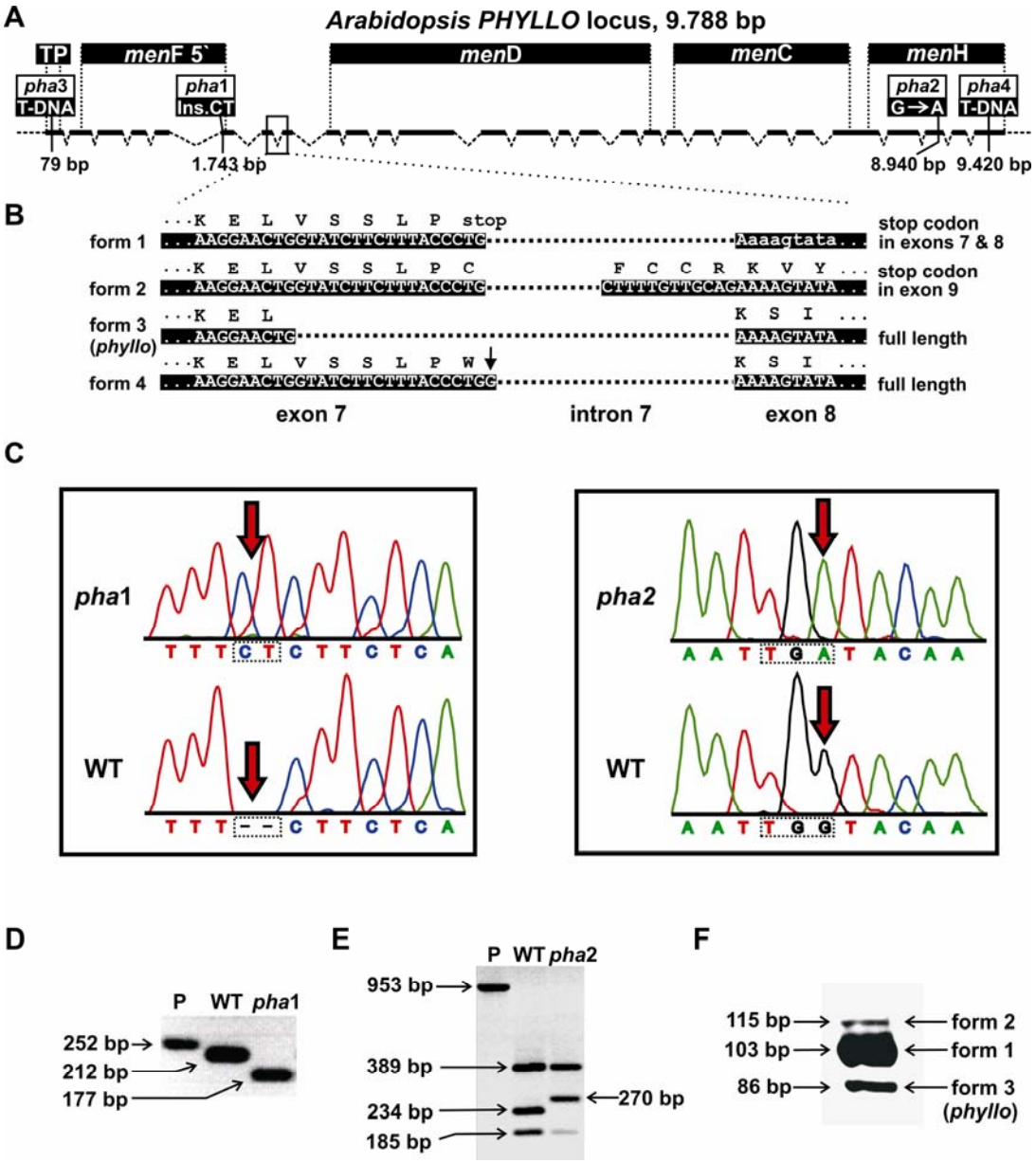


Figure 7. Physical characterization of the *PHYLLO* locus (**A**) Physical map of *PHYLLO* depicting exons (solid lines), introns (traced angular lines), the transit peptide (TP) for chloroplast targeting, *men* homologous regions (black bars), and the localization of *pha* mutations relative to the start codon. *Pha*1, insertion (Ins.) of a cytosine/thiamine (CT); *pha*2, transition of guanine (G) to adenine (A); *pha*3 and *pha*4, T-DNA insertions. The squared box delimits the region of alternative splicing. (**B**) Alternatively spliced forms 1-3 between exons 7 and 8 are presented in black bars containing the nucleotide sequences. Deduced amino acid residues are placed above. The arrow indicates the guanine added to form 1 to generate form 4. (**C**) Representative sequence chromatograms showing the *pha*1 and *pha*2 mutations (indicated by arrows) comparing to the wild-type

sequence. (**D**) The *pha*1dCAPS polymorphic marker distinguishes the *pha*1 mutation from the wild-type sequence. The DNA of the 252 bp PCR product (P, lane 1) and the *BsmA*I restriction fragments of 212 bp (WT) and 177 bp (*pha*1) were separated by electrophoresis on a 2% agarose gel. (**E**) The *pha*2CAPS polymorphic marker distinguishes the *pha*2 mutation from the wild-type sequence. The DNA of the 953 bp PCR product (P, lane 1) and the *Rsa*I restriction fragments of 389, 234 and 185 bp in the wild-type and 389, 270 and 185 bp in the *pha*2 mutant were separated by electrophoresis on a 2% agarose gel. (**F**) Eletrophoretic separation on a 3% agarose gel of RT-PCR products with the primers MenD31_for and MenD31_rev annealing in the 7th and 8th exons, respectively. Three products of 103 bp, 115 bp and 86 bp were observed and correspond to the alternative splicing variants form 1, form 2 and form 3 (*phyllo*), respectively.

3.3. Evidences of a single gene related to the *PHYLLO* locus

Four lines of evidence demonstrate unequivocally that the observed regions of the *PHYLLO* locus homologous with four *men* genes represent a single, monocistronic gene. (i) Crossing the mutants with lesions located in the 5' region of the locus (*pha*1 and *pha*3) with those defective in the 3' region (*pha*2 and *pha*4) (Fig. 8A) resulted in offspring showing a clear *pha* phenotype. In all cases the results of the crosses were confirmed by PCR and nucleotide sequence, indicating unambiguously that all mutants belong to the same complementation group. (ii) In the *pha*3 background (T-DNA insertion in the *men*F 5'-region, Fig. 7A) no accumulation of transcripts related to *men*D and *men*H regions could be detected

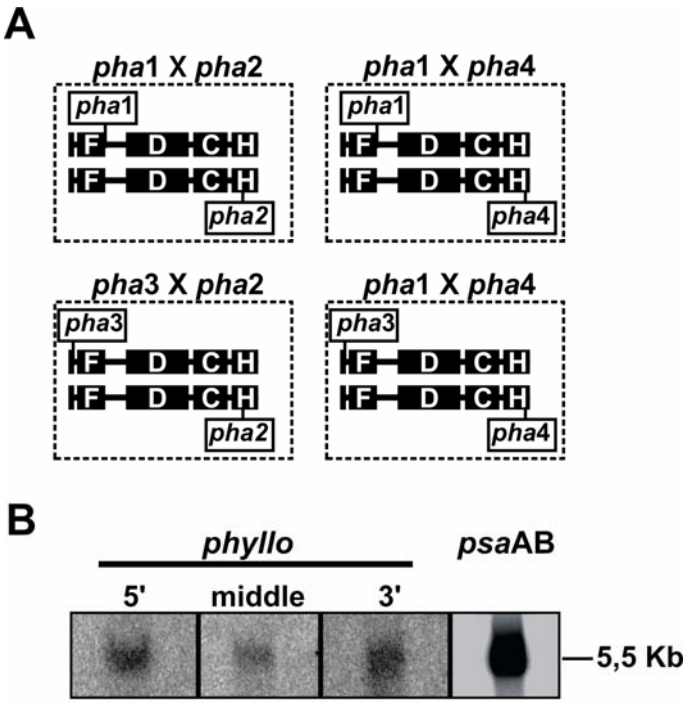


Figure 8. Evidences that the *PHYLLO* locus encodes a single gene. (**A**) Schematic presentation of allelism test crosses. A clear *pha* phenotype was observed in all cases (Fig. 4A-C). Localisation of the mutations and the designation of the *men* genes (abbreviated by capital letters) are depicted. (**B**) RNA gel blot analysis using probes amplified from the 5´, middle and 3´ parts of the *phyllo* locus. For hybridisation with the *phyllo* probes 6,3 µg of polyA$^+$ mRNA were loaded resulting in the detection of very weak signal of 5,5 kb. For hybridization with a probe specific for the plastid *psa*A/B operon 8,0 µg total RNA was loaded and a strong signal of the same size could be detected.

by real time RT-PCR analysis (data not shown). This suggests that only one promoter drives the expression of the entire *PHYLLO* region. (iii) RNA gel blot analysis revealed a single band of 5,5 kb when probing 5'-, central, and 3'-parts of the *PHYLLO* locus (Fig. 8B), indicating that an unique transcript covers the entire *PHYLLO* interval. (iv) The most convincing evidence is the existence of a functional full-length coding frame of the complete *PHYLLO* gene (section 3.4., below).

3.4. The *PHYLLO* gene

Comprehensive nucleotide sequence analyses of overlapping RT-PCR products revealed that *PHYLLO* is encoded by 28 exons containing all the four *men*-homologous regions and covering a 9.788 bp interval of the T6L1 BAC clone (Fig. 7A). Additionally, three alternative splicing events between exons 7 and 8, named forms 1-3, were detected (Fig. 7B and F). One splicing variant (form 1) corresponds to the RAFL 09-32-C05 cDNA, which carries a premature termination codon at position 1.005 bp overlapping exons 7 and 8 and leads to expression only of the region homologous with *men*F (Fig.7B). The downstream sequence links the *men*D, *men*C, and *men*H regions in a single reading frame. Form 2 extends the 5' part of exon 8 and harbours a stop codon at nucleotide position 1092 in exon 9, also resulting in the expression of a truncated MenF product. Differently from the previous splicing variants, form 3 contains shorter boundaries of exons 7 and 8, which bypasses the latter stop codons and generates a reading frame encoding a fused product of the *men*F, *men*D, *men*C, and *men*H modules preceded by transit peptide. Taken together, since this third variant solely has the coding capability for the entire locus spanning all four *men*-homologous regions, it's to be concluded that this form encodes and corresponds to the functional full-length *PHYLLO* product.

The determined nucleotide sequences of the 5.364 bp RAFL 09-32-C05 cDNA clone and the 5.347 nucleotides *phyllo* transcript were deposited in the GenBank and are available under the accession numbers DQ084385 and DQ084386, respectively. The 5.148 nucleotides reading frame of the *phyllo* transcript is preceded by a 5'-UTR of 38 bases and is followed by a 3'-UTR of 161 nucleotides.

3.5. Complementation of the *pha* mutations

3.5.1. The failure to complement the *pha* mutations with the RAFL 09-32-C05 cDNA (form 1)

Both splicing forms 1 and 2 present a premature stop codon and have a bicistronic structure, encoding two ORFs that correspond to a MenF and a downstream MenDCH products, respectively. It could be possible that these forms might be decoded by rare mechanisms involving internal initiation of translation or translational frame-shifting, sometimes verified in polycistronic-like mRNA in eukaryotes (Blumenthal, 1998; Ivanov *et al.*, 2000). If this were the case then the RAFL 09-32-C05 cDNA containing a stop codon (form 1) should complement all *pha* mutations. To analyse this possibility, the pbinC05 construct containing the form 1 was transformed into the background of all *pha* mutants. In a total of 170 combined transformants analysed, 57 plants were heterozygous for the *pha* mutations and no one was homozygous. From these heterozygous plants, 20 transformants (five from each of the *pha* mutations) were selected and self-crossed. Their offspring were grown on MS medium in the presence of the selecting antibiotic. In all cases a clear *pha* phenotype co-segregated with the complementation vector, indicating unambiguously that the stop codon-containing cDNA was not able to rescue the *pha* mutations and that the truncated splicing forms 1 and 2 might not be functional.

3.5.2. Complementation with the engineered full-length form 4

To confirm the previous results indicating that the *PHYLLO* locus encodes a single composite gene, a full-length cDNA containing one tetra-modular reading frame of the four *men*-homologous regions was generated by inserting a guanine nucleotide into the context of the stop codon of form 1 in the pbinC05 clone (Fig. 7B). The resulting form 4 complemented four independent lines of the *pha*4 mutant (Fig. 9), allowing photoautotrophic growth, partial

PhQ accumulation to approximately 23,5% wild-type levels (Table 3) and recovery of PSI activity. The form 4 was introduced into the background of the *pha*3 mutation by crossing. Accordingly, homozygous *pha*3 plants in the F2 progeny were also complemented by the form 4 (Table 3; Fig. 9), leading to partial PhQ accumulation to approximately 18% of wild-type levels, confirming that the *PHYLLO* locus encodes a tetra-modular fused gene.

3.5.3. Complementation with the genomic locus

Furthermore, a genomic fragment containing the *PHYLLO* coding region, subcloned from the BAC clone T6L1, was introduced into the background of the *pha*3 mutation and complemented several independent transformed lines, also recovering photoautotrophic growth. Offspring of the complemented lines *pha*3G10-15, *pha*3G10-55, *pha*3G10-69 and *pha*3G10-87 were analysed showing different levels of PhQ accumulation (Table3) and correspondent photosynthetic activity (section 3.10).

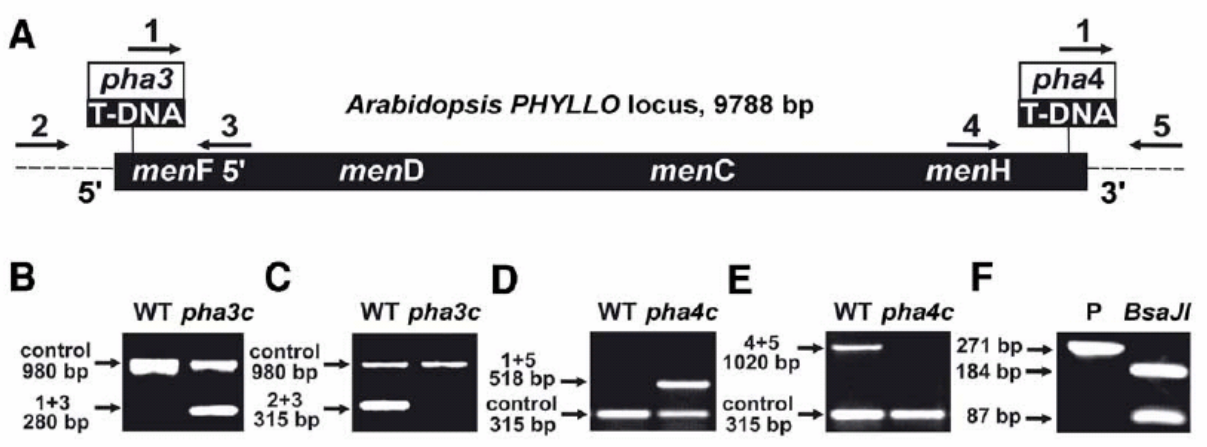


Figure 9. Complementation of the *pha*3 and *pha*4 mutations with the form 4. (**A**) Schematic representation of the *PHYLLO* locus indicating the locations of *pha*3 and *pha*4 mutations and the positions of primers (arrows) used for screening of the transformed plants. (**B**) PCR product of 280 bp using primers 1 (LBb1) and 3 (MenD100_rev) indicates the presence of the T-DNA in *pha*3 complemented line (*pha*3c). (**C**) PCR product of 315 bp using primers 2 (MenD100) and 3 is only present in the wild-type (WT) plant, indicating homozygozity of the T-DNA insertion in the *pha*3c. (**D**) PCR product of 518 bp using primers 1 and 5 (MenD104_rev) indicates the presence of the T-DNA in *pha*4 complemented line (*pha*4c). (**E**) PCR product of 1.020 bp using primers 4 (104mut) and 5 is only present in the wild-type, indicating homozygozity of the T-DNA insertion in the *pha*4c. In all the PCR reactions positive controls generating products of 980 bp (primers Ac_for and Ac_rev) (panels **B** and **C**) and 315 bp (primers MenD100 and MenD100_rev) (panels **D** and **E**) were used. (**F**) The presence of form 4 in the *pha*3c and *pha*4c plants was confirmed by cleavage of a specific PCR product (p) of 271 bp (primers Cor2 and Del2_rev) using the enzyme *BsaJ*I, generating fragments of 184 and 87 bp.

3.6. Characterization of the *PHYLLO* product

The *PHYLLO* product consists of 1.715 deduced amino acid residues comprising the four Men-homologous modules preceded by a predicted transit peptide for chloroplast targeting (Fig. 10). The boundaries of the Men-homologous modules were delimitated by alignments with homologous sequences of eubacteria and searches with PHYLLO against the Conserved Domain Database. All modules are separated by small stretches showing no significant similarity to Men proteins. These spacer sequences are of 83 amino acid residues between MenF and MenD, 50 amino acids separating MenD from MenC, and 46 amino acids between MenC and MenH.

```
1    MRSSFLVSNPPFLPSLIPRYSSRKSIRRSRERFSFPESLRVSLLHGIQRN
51   IEVAQGVQFDGPIMDRDVNLDDDLVVQVCVTRTLPPALTLELGLESLKEA
101  IDELKTNPPKSSSGVLRFQVAVPPRAKALFWFCSQPTTSDVFPVFFLSKD
151  TVEPSYKSLYVKEPHGVFGIGNAFAFVHSSSVDSNGHSMIKTFLSDESAM
201  VTAYGFPDIEFNKYSTVNSKDGSSYFFVPQIELDEHEEVSILAVTLAWNE
251  SLSYTVEQTISSYEKSIFQVSSHFCPNVEDHWFKHLKSSLAKLSVEEIHP
301  LEMEHMGFFTFSGRDQADVKELKSIQSSCQFHCKLSPDVVFSNNMLNRET
351  EVSNFLRDEANINAVWASAIIEECTRLGLTYFCVAPGSRSSHLAIAAANH
401  PLTTCLACFDERSLAFHAIGYAKGSLKPAVIITSSGTAVSNLLPAVVEAS
451  EDFLPLLLLTADRPPELQGVGANQAINQINHFGSFVRFFFNLPPPTDLIP
501  VRMVLTTVDSALHWATGSACGPVHLNCPFRDPLDGSPTNWSSNCLNGLDM
551  WMSNAEPFTKYFQVQSHKSDGVTTGQITEILQVIKEAKKGLLLIGAIHTE
601  DEIWASLLLAKELMWPVVADVLSGVRLRKLFKPFVEKLTHVFVDHLDHAL
651  FSDSVRNLIEFDVVIQVGSRITSKRVSQMLEKCFPFAYILVDKHPCRHDP
701  SHLVTHRVQSNIVQFANCVLKSRFPWRRSKLHGHLQALDGAIAREMSFQI
751  SAESSLTEPYVAHMLSKALTSKSALFIGNSMPIRDVDMYGCSSENSSHVV
801  DMMLSAELPCQWIQVTGNRGASGIDGLLSSATGFAVGCKKRVVCVVGDIS
851  FLHDTNGLAILKQRIARKPMTILVINNRGGGIFRLLPIAKKTEPSVLNQY
901  FYTAHDISIENLCLAHGVRYVHVGTKSELEDALFVPSVEEMDCIVEVESS
951  INANAIVHSTLERFARQAAENSLGIVSASSFLHPMIKNVLLCQVSGIQYS
1001 QYRVKLCDRPTICSDEFSQFHREGFILSLTLEDGSIGYGEVAPLNSNVEN
1051 LMDVEGQLQLVLHLMNEAKFSYMLPLLNGSISSWIWSELGITASSIFPSV
1101 RCGLEMALLNAMAVRHDSSLLGILHYQKEENGSAQPHSVQICALLDSEGT
1151 PLEVAYVARKLVQEGFSAIKLKVGRRVSSVQDALVMQEVRRAVGVQIELR
1201 ADANCRWTFEEAREFGLLVNSCNLKYIEEPVQNKDDLIRFHEETGLPVAL
1251 DETLDDFEECPLRMLTKYTHPGIVAVVIKPSVVGGFENAALIARWAQQHG
1301 KMAVISAAYESGLGLSAYILFASYLEMENVKASTEQKQGTPPSVAHGLGT
1351 YRWLSEDVMMNTLGIFRSPYSGFVEGFIADASRNLKDVKINNDVIVRTSK
1401 GIPVRRYELRVDVDGFSHFIRVHDVGENAEGSVALFLHGFLGTGEEWIPI
1451 MTGISGSARCISVDIPGHGRSRVQSHASETQTSPTFSMEMIAEALYKLIE
1501 QITPGKVTIVGYSMGARIALYMALRFSNKIEGAVVVSGSPGLKDPVARKI
1551 RSATDDSKARMMVDNGLYIFIENWYNGGLWKSLRNHPHFSKIAASRLLHG
1601 DVPSVAKLLSDLSSGRQPSLWEELEDCDTNISLVFGEKDVKYKQIATRMY
1651 REMSKSKKSVNNIIEIVEIPEAGHAVHLESPLRVILALRKFLTRVHNSST
1701 ETELSQKLLLALKEM
```

Figure 10. Amino acid sequence of the deduced *PHYLLO* product. The predicted transit peptide is indicated in red; the MenF 5'-module in blue; the MenD module in lavender; The MenC module in green; and the MenH module in orange. Amino acid residues not coloured correspond to spacer regions.

3.6.1. PHYLLO is a plastidial protein

PHYLLO has a putative N-terminal transit peptide for chloroplast targeting. The ChloroP tool predicted a proteolytic cleavage site after the amino acid residue at position 19. A fusion of the MenF 5'-module of PHYLLO with the DsRed fluorescence reporter was constructed for transient expression analysis in tobacco protoplasts. The results confirmed the localization of PHYLLO in the chloroplast, into dense spatial structures verified inside this organelle (Fig. 11).

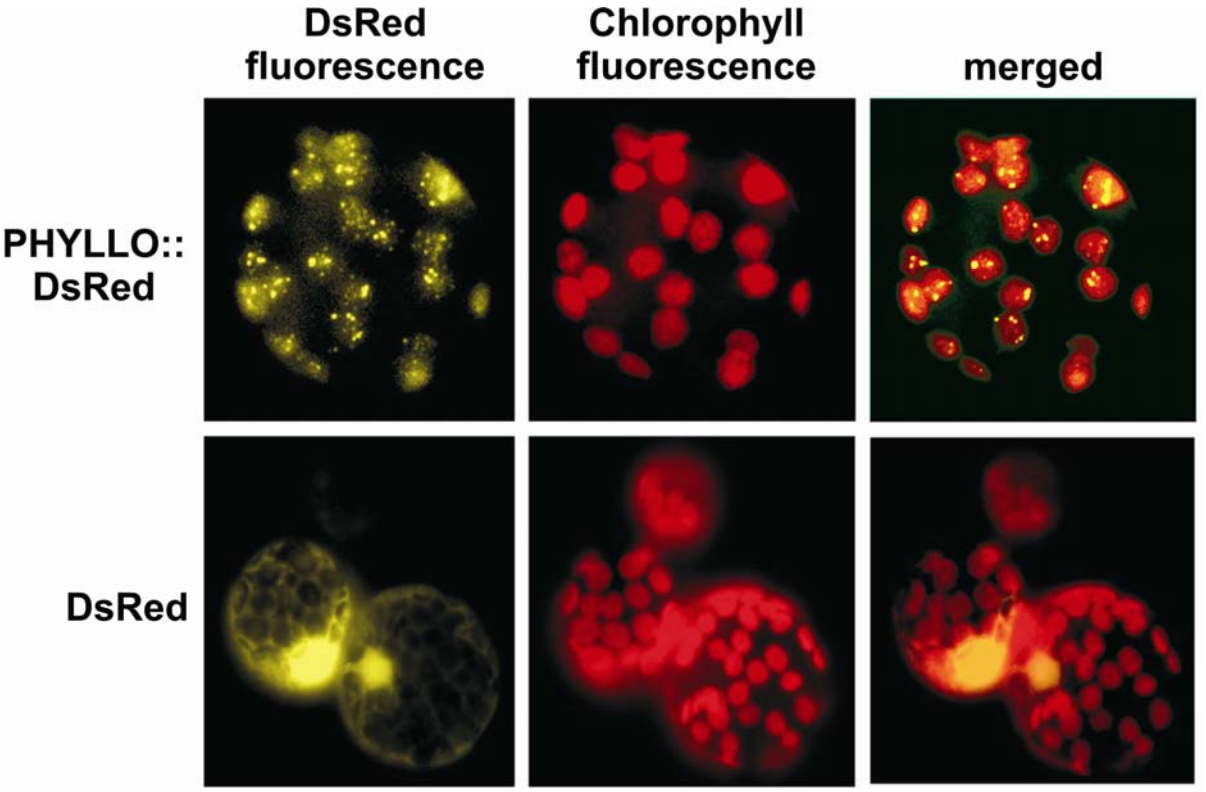


Figure 11. Localization of PHYLLO in chloroplasts is revealed by merging the fluorescence induced by the PHYLLO::DsRed fusion and the chlorophyll. The DsRed control is expectedly localized in the nucleus.

3.6.2. The MenD module

The PHYLLO MenD module comprises 615 amino acids ranging from the residue position 356 to the position 970 (Fig. 10). Searches with this module against the CDD retrieved a hit with expect value (e-value) of 4e-122 with the deposited MenD domain, as well as a hit of 1e-08 with the thiamine diphosphate (ThDP) domain, characteristic of this protein. The MenD enzyme, or 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC)

synthase, catalyses the ThDP-dependent decarboxylation of 2-ketoglutarate, the subsequent addition of the resulting succinyl-ThDP moiety to isochorismate, and the elimination of the pyruvate to yield SHCHC, pyruvate and carbon dioxide (Bashin *et al.*, 2003). These enzymatic catalysis, corresponding to the second step of the PhQ and MQ pathway (Fig. 3), are entirely dependent on ThDP and a divalent ion (Mg^{+2} or Mn^{+2}) (Palaniappan *et al.*, 1992; Bashin *et al.*, 2003). Pair wise alignments showed identities for the MenD module of PHYLLO ranging from 20% to 30% with eubacterial MenD homologs (Table 4), in accordance with the characteristic low conservation observed between the members of the MenD family (Bhasin *et al.*, 2003).

Table 4. Conservation levels of the MenD module of PHYLLO. The identity (ID) and similarity (Sim) levels of the MenD module of PHYLLO with representative eubacterial MenD proteins are shown. GenBank accession numbers (accession) of proteins are indicated.

MenD of	accession	ID	Sim
Chlorobium tepidum	AAM73060	30%	48%
Rubrobacter xylanophilus	ZP_00199577	30%	45%
Cyanidium caldarium	NP_045109	28%	48%
Crocosphaera watsonii	ZP_00514771	28%	47%
Escherichia coli	NP_416767	28%	43%
Bacillus subtilis	CAB15060	28%	46%
Trichodesmium erythraeum	ZP_00071903	26%	46%
Synechococcus elongatus	ZP_00164512	27%	44%
Cyanidioschyzon merolae	NP_849026	25%	42%

The MenD proteins belong to a superfamily of ThDP-dependent 2-oxo acid decarboxylases. Structural and functional data has already been provided for the *E. coli* MenD and for the close related member of the superfamily, acetohydroxy acid synthase (AHAS) of yeast (Pang *et al.*, 2002; Bashin *et al.*, 2003). The MenD structure can be divided in three distinct domains (Fig. 12). The N-terminal part corresponds to the ThDP/pyrimidine binding domain spanning from residues 356 to 539 in the PHYLLO protein. The residues G387, E411 and S434 in PHYLLO are highly conserved to the G31, E55 and T78 in *E. coli* MenD (Fig. 12 and Fig 13) and were demonstrated in the AHAS homolog to directly contact the aminopyrimidine portion of the ThDP (Pang *et al.*, 2002). Especially the E411 is part of a Asp-Glu-Arg strictly conserved in the whole MenD family (Bashin *et al.*, 2003). The central domain, from a.a 540 to 723, is characteristically low conserved in all the superfamily members, probably reflecting the absence of functional elements. The carboxy-terminal (C-terminal) part (amino acids 724 to 970) corresponds to the ThDP/pyrophosphate binding

domain, which carries the characteristic ThDP-binding motif that stretches from residues 848 to 877 in PHYLLO. Also this domain in the PHYLLO MenD module contains a highly conserved S780 and R784 that may be involved in the binding of the isochorismate substrate during catalysis (Bashin *et al.*, 2003). Taken together, these data confirm the overall conservation of the MenD module PHYLLO with related eubacterial proteins involved in MQ biosynthesis (Fig. 12 and Fig. 13).

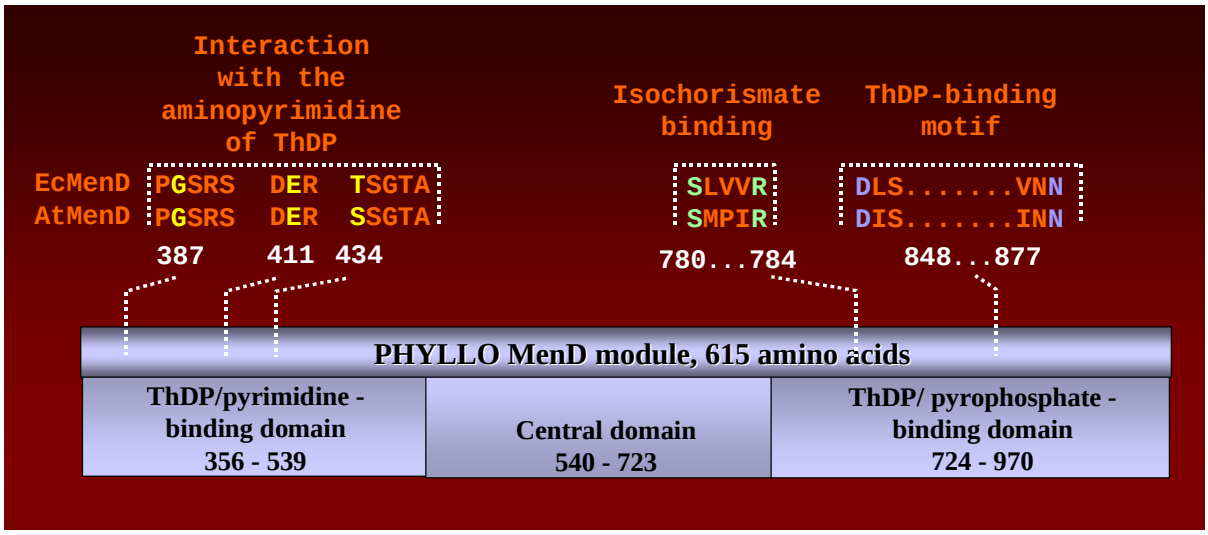


Figure 12. Schematic representation of the 615 amino acid residues of the PHYLLO MenD module. Three characteristic domains are indicated below the structure of the MenD module (central axis) with numbers corresponding to the amino acid positions of the full-length PHYLLO protein. The presence of conserved motifs in the *Arabidopsis thaliana* MenD module of PHYLLO (AtMenD) with the *E. coli* MenD (EcMenD) are indicated above by local alignments between the two proteins. The yellow amino acid residues (positions 387, 411 and 434 of PHYLLO) are supposed to directly contact the aminopyrimidine part of ThDP. The green amino acid residues (position 780 and 784 of PHYLLO) possibly bind the isochorismate precursor. The blue amino acid residues are highly conserved residues that take part of the ThDP-binding motif (positions 848 to 877 of PHYLLO).

```
                    *        20         *        40         *        60         *
E. coli     : -------------------MSVSAFNRRWAAVILEALTRHGVRHICIAPGSRSTPLTLAAAENSAFIHHTHFDER :  56
Syn. WH     : ------------------MQLDFRNINTLWSSVIVETLSRLGLTTAVICPGSRSTPLTLAFARHPHIETTPILDER :  58
Cyanidium   : ----------------------MMDNIWSNAIVEELVRQKITYFFIAPGSRSTLLALAVANNKQSKSIVHFDER :  52
Chlorobium  : MGRGRSENRGHSCHNAADLMNSKQITTLWCAVIVEELIRQEAGFFCISPGSRSTPLTLAVASNPKARFRMFPDER :  75
AtMenD      : -------------------LRDEANINAVWASAIIEECTRLGLTYFCVAPGSRSSHLAIAAANHPLTTCLACFDER :  57

                 80         *        100        *        120        *        140        *
E. coli     : GLGHLALGLAKVSKQPVAVIVTSGTAVANLYPALIEAGLTGEKLILLTADRPPELIDCGANQAIRQPGMFASHPT : 131
Syn. WH     : SAAFFALGIAKRIHLPTALICTSGTAGANFYPAVIEAKESQVPLLIFTGDRPPELRNCHAGQTIDQVKLYSNYCR : 133
Cyanidium   : SIGYAAVGFSKSFNTAAVIIVTSGTAVGNLLPCIIEASMSCIPLLVLSADRPYELKFCGANQSIDQNKIFCNFTR : 127
Chlorobium  : SAGFYALGYARATGMPAVLVCTSGTAVANYFPAVVEASADAQPMLVLSADRPFELLECGANQAIRQQNIFGSYTR : 150
AtMenD      : SLAFHAIGYAKGSLKPAVIITSSGTAVSNLLPAVVEASEDFLPLLLLTADRPPELQGVGANQAINQINHFGSFVR : 132

                160        *        180        *        200        *        220
E. coli     : HSISLPRPTQDIAARWLVSTIDHALGTLHAG----GVHINCPFAEPLYGEMDDTGLSWQQRLGDWWQDDKE---- : 198
Syn. WH     : WYTELSLPSLDIKMLSYLRQTIITAWDHTLVPYRGVVHLNCPFREPLAPIVT-PEINTLSSTINFDVFFTN---- : 203
Cyanidium   : WFFEIPPPEEKLKIKAIISAILYACIKMKGDTP-GPVHINCLFREPFFHDNKNHNDTGLIYFREWIHSHEIFN--- : 199
Chlorobium  : WSFELPEPGIATPLASLLSTVDHAVRKSLSLP-AGPVHLNLPFREPLEPEAPDPGHPWAAPLETWQASGEE---- : 220
AtMenD      : FFFNLPPPTDLIPVRMVLTTVDSALHWATGS-ACGPVHLNCPFRDPLDGSPTNWSSNCLNGLDMWMSNAEPFTKY : 206
```

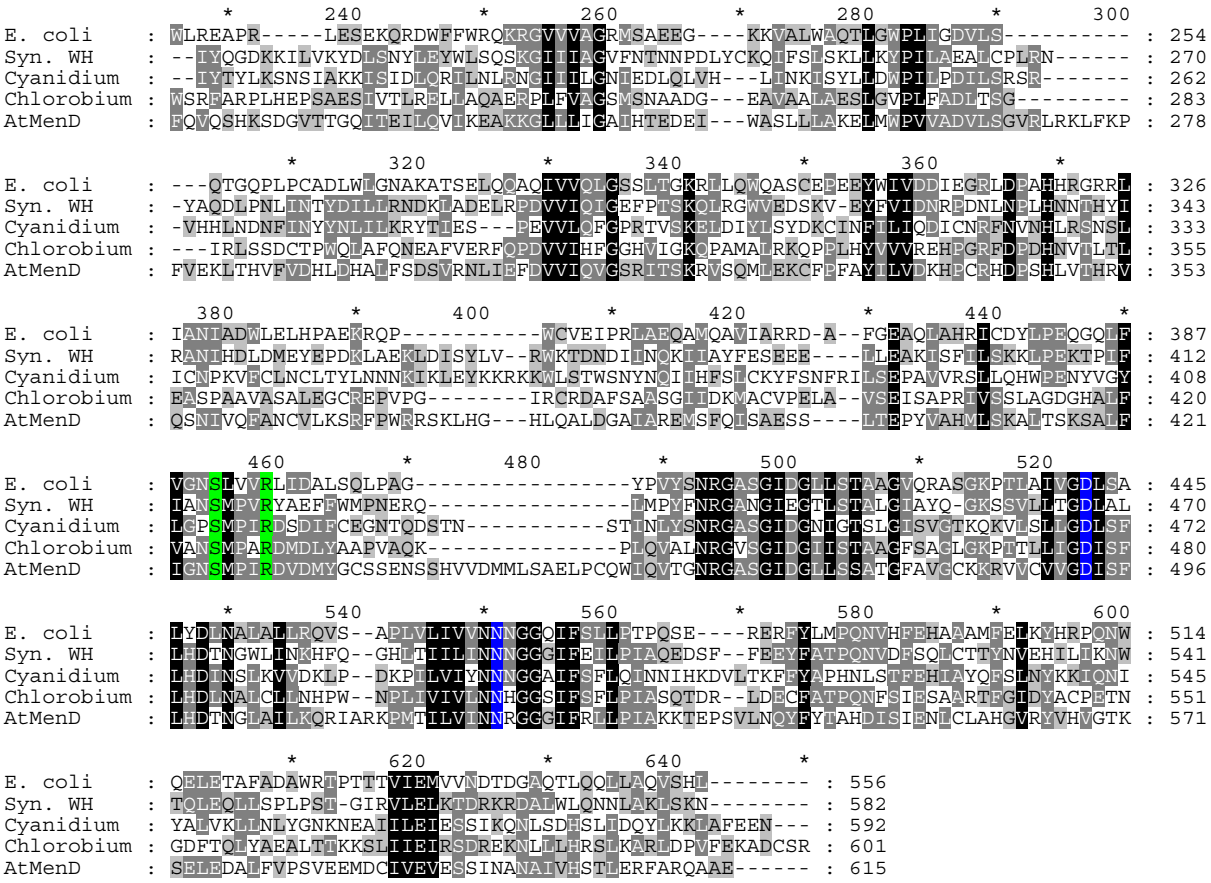


Figure 13. Alignment of the *Arabidopsis thaliana* MenD module of PHYLLO (AtMenD) with MenD proteins of the representative eubacteria *Escherichia coli* K12 (E. coli, accession number NP_416767), *Synechocystis* sp. WH 8501 (Syn. WH, ZP_00179049), *Cyanidium caldarium* (Cyanidium, NP_045109), and *Chlorobium tepidum* (Chlorobium, AAM73060). Residues with black background (X) -100% conservation, grey background (X) -60% conservation, grey background with black letter (X) -30% conservation. Residues coloured in yellow, green and blue correspond to the highly conserved motifs discussed in the text and depicted in the Fig. 12.

3.6.3. The MenC module

The PHYLLO MenC module comprises 347 amino acid residues ranging from the residue position 1.021 to the position 1.367 (Fig. 10). Searches with this module against the CDD retrieved a hit with an e-value of 2e-22 with the deposited MenC domain. The MenC enzyme, or *o*-succinyl-benzoate synthase, catalyses a dehydration reaction, in which SHCHC is converted to OSB. This reaction is strictly dependent on the cofactor Mg^{+2} (Palmer *et al.*, 1999; Klenchin *et al.*, 2003). The MenC is a member of the muconate lactonizing subgroup of the enolase superfamily. The sequence similarity relating different members of the family typically is lower than 25% (Thompson *et al.*, 2000). In accordance, the PHYLLO MenC

module has low overall conservation with the closest related eubacterial proteins (Table 5). Nonetheless, family members can be distinguished by highly conserved structural and functional residues.

Table 5. Conservation levels of the MenC module of PHYLLO. The identity (ID) and similarity (Sim) levels of the MenC module of PHYLLO with representative eubacterial MenC proteins are shown. GenBank accession numbers (accession) of proteins are indicated.

MenC of	accession	ID	Sim
Chlorobium tepidum	AAM73067	21%	32%
Pasteurella multocida	AAK03178	19%	33%
Photobacterium profundum	CAG21001	19%	32%
Escherichia coli	AAC75321	18%	32%
Bacillus subtilis	CAB13155	14%	26%
Trichodesmium erythraeum	ZP_00071895	12%	29%
Cyanidium caldarium	AAF12990	11%	26%
Nostoc sp. PCC 7120	NP_484078	11%	26%
Cyanidioschyzon merolae	NP_849028	9%	23%

Crystallographic data are available for the MenC of *E. coli* in complex with OSB (Thompson *et al.*, 2000), demonstrating with high resolution the amino acids related to the structure of the reaction centre (Fig. 14). The binding pocket for OSB can be divided in three regions (Thompson *et al.*, 2000). (i) The divalent metal interaction composed by amino acids D161, E190 and D213 in *E. coli* that are identical in the PHYLLO MenC in positions D1201, E1228 and D1251 (Fig. 14; Fig. 15). (ii) The region of interaction with the carboxyl and carboxyl groups, composed by residues K131, N163 and S262 that directly interact with OSB, as well as by R159 and D290 that indirectly interact with the product by mean of water bridges. In *Arabidopsis,* these amino acid residues are identical (Fig. 14, Fig. 15) with exception of the D290 that is absent or corresponds to an amino acid residue not obviously evident from the alignment position. Two additional lysines residues, strictly conserved in all the superfamyly members, make direct contact with the OSB being responsible by the abstraction of a α-proton of the carboxylate anion during the catalysis (Palmer *et al.*, 1999; Klenchin *et al.*, 2003). Both are present in PHYLLO at positions K1172 and K1279. (iii) The third region corresponds to a.a residues implicated in hydrophobic contacts with the benzene ring of OSB. Hydrophobic residues are also present in corresponding positions in the MenC module of PHYLLO (Fig. 14; Fig. 15). Taken together, these data support the function of the MenC module of PHYLLO as an OSB synthase.

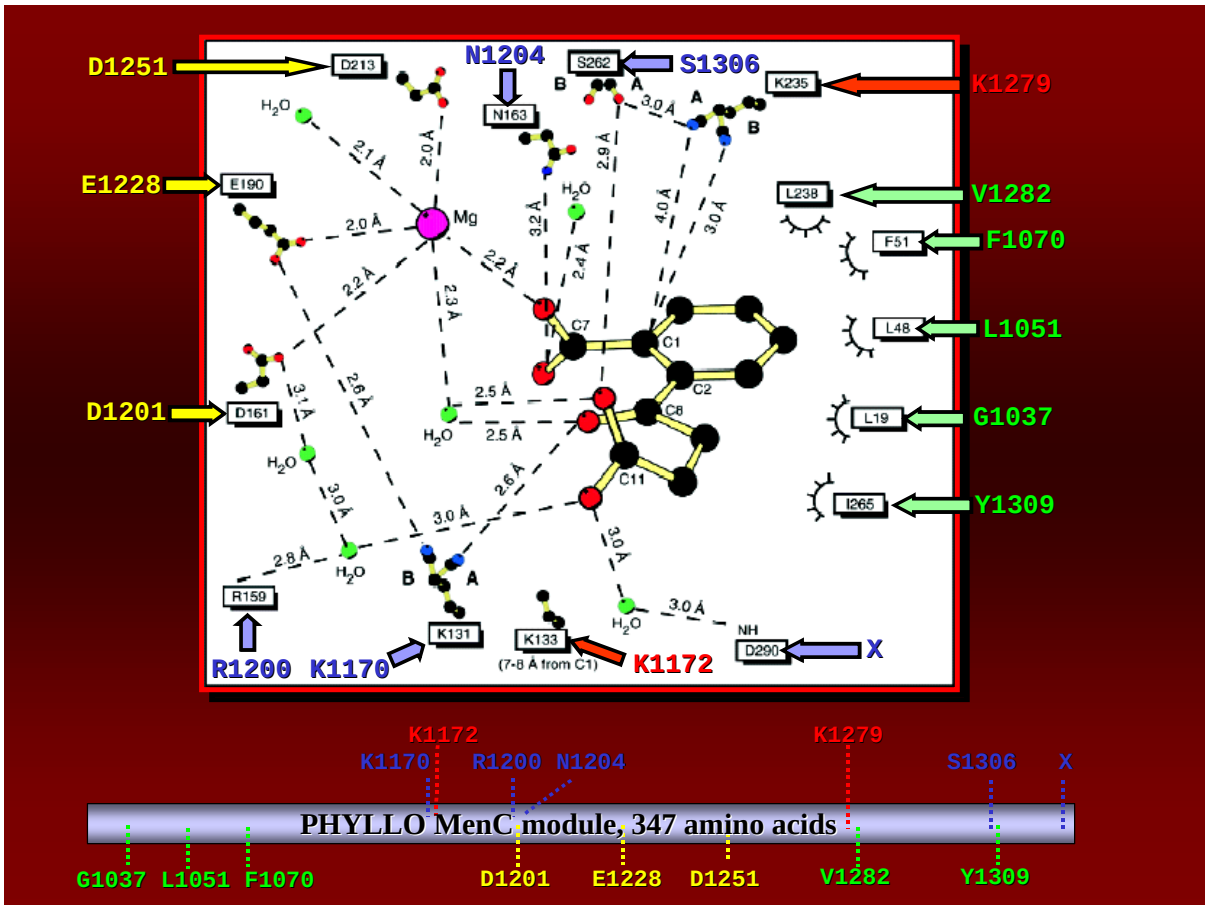


Figure 14. Schematic representation of conserved a.a in the PHYLLO MenC module. The structure of the OSB binding pocket of *E. coli* MenC protein (according to Thompson *et al.*, 2000) is showed above indicating the involved amino acids. The conserved amino acids present in the MenC module of PHYLLO are indicated by arrows pointing to the corresponding amino acid residues of *E. coli* MenC. The position of these amino acids related to the structure of PHYLLO is also depicted below. The yellow residues bind the divalent cation cofactor. The green amino acids make hydrophobic contacts to the benzene ring of OSB. The red coloured lysines constitute highly conserved residues directly involved in catalysis. The blue residues interact directly or indirectly with the carbonyl/carboxyl groups of OSB. The X represents an amino acid corresponding to the D290 of *E. coli* MenC that is either absent from PHYLLO or represented by an amino acid not proximally positioned in the alignment.

```
                    *       20        *       40        *       60        *
E. coli     : ----------MRSAQVYRWQIPMDAGVVLRDRRLKTRDGLYVCLREGEREGWGEISPLPG---FSQETWE :  57
Photobacte  : MELIMAKTVVLTEEYIRSAKLYQYQLPMDSGVILRTQRLATRECWIVELRDGDKVGFGEIAPLPE---FSHETVE :  72
Pasteurell  : ------------MTMIRKFKLYQYSIPVDSQLILRNRFLKKREGLLVQVCCDDAQGWGEIAPLPE---FSQETLE :  60
Chlorobium  : -------------MKPLHADICRYEMDFTAPVTVRGVLLARRQGLLLRLKSEGVTAYGEVAPLIG---LHTESLD :  59
AtMenC      : ----------------HREGFILSITLEDGSIGYGEVAPLNSNVENLMDVEG--QLQLVLHLMNEAKFSYMLPL :  56

                  80        *       100       *       120       *       140       *
E. coli     : EAQSVLLAWVNNMLAGDCELP----------QMPSVAFGVSCALAELTDTLPQ--------------------A : 101
Photobacte  : QAGEQAKQLLESWIQGETIDLSQ---------ALPSVAFGLSMAMLEQASELPM--------------------K : 118
Pasteurell  : QARFQALAWLEEWSASDGSAGKLP-----LEHLFPSVAFGLSCALAEMKGLLHE--------------------E : 110
Chlorobium  : EALQALATFIPELSRLDWNASDGRQRLLDEAALPPSVTTGIEMALINLEATERSSLPS---------FTDEFPPA : 125
AtMenC      : LNGSISSWIWSELGITASS------------IFPSVRCGLEMALLNAMAVRHDSSLLGILHYQKEENGSAQPHS : 118
```

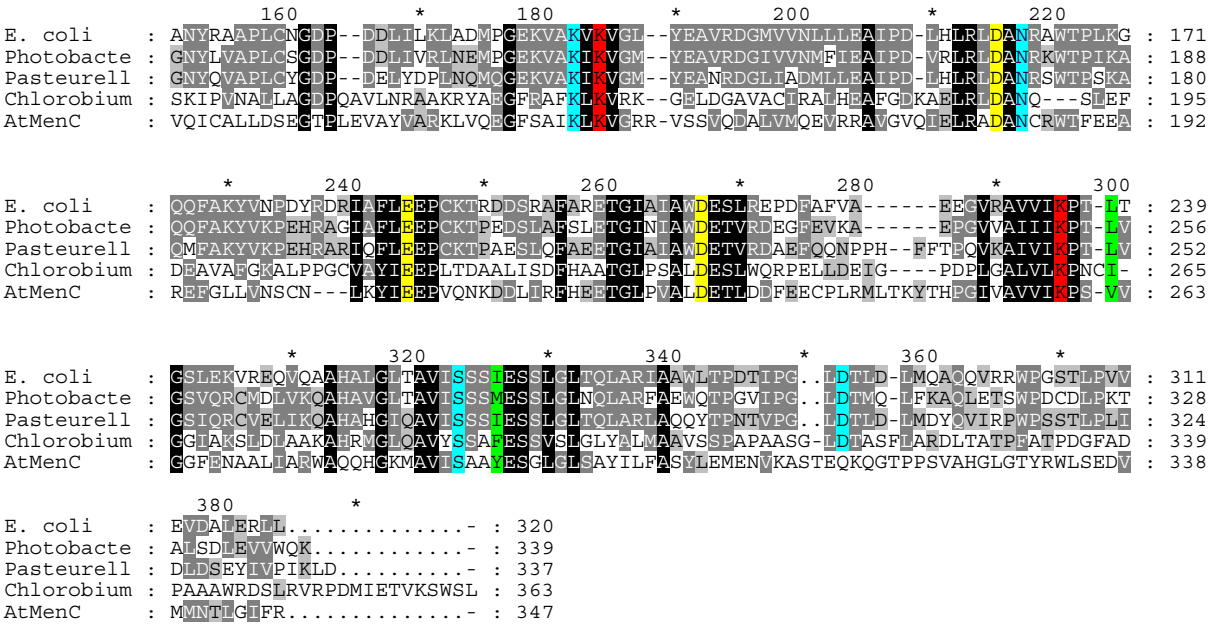


Figure 15. Alignment of the *Arabidopsis thaliana* MenC module of PHYLLO (AtMenC) with MenC proteins of the representative eubacteria *Escherichia coli* K12 (E. coli, accession number AAC75321), *Photobacterium profundum* (Photobacte, CAG21001), *Pasteurella multocida* (Pasteurell, AAK03178), and *Chlorobium tepidum* (Chlorobium, AAM73067). Residues with black background (X) -100% conservation, grey background (X) -60% conservation, grey background with black letter (X) -30% conservation. Residues coloured in yellow, green, blue and red correspond to highly conserved amino acids discussed in the text and depicted in the Fig. 14.

3.6.4. The MenH module

The PHYLLO MenH module comprises 302 amino acid residues ranging from the residue position 1.414 to the position 1.715 (Fig. 10). Blast searches with this module returned hits with minimum e-value of 4e-35 with the MenH proteins of eubacteria. In general, these proteins display low levels of sequence identity (Table 6) with PHYLLO, as also observed for the MenD and MenC eubacterial homologs. In addition to MenH proteins, BLAST searches also retrieved hits with several other proteins having the so-called α/β hydrolase fold. This was confirmed in a search against the CDD returning hits with an e-value 4e-18 with deposited hydrolases and esterases domains, as well as with many subgroups of the α/β hydrolase fold. This superfamily comprises as many different members as proteases, lipases, dehalogenases, peroxidases, epoxide hydrolases and esterases (Nardini and Dijkstra, 1999). This latter function is consistent with the thioesterase activity of MenH proteins.

Table 6. Conservation levels of the MenH module of PHYLLO. The identity (ID) and similarity (Sim) levels of the MenH module of PHYLLO with representative eubacterial MenH proteins are shown. GenBank accession numbers (accession) of proteins are indicated.

MenH of	accession	ID	Sim
Trichodesmium erythraeum	ZP_00071896	27%	48%
Listeria innocua	CAC97013	27%	43%
Nostoc punctiforme	ZP_00112005	26%	47%
Chlorobium tepidum	AAM73065	26%	44%
Nostoc sp. PCC 7120	BAB74452	25%	44%
Bacillus subtilis	CAB15059	23%	39%
Escherichia coli	AAC75323	20%	35%

The α/β hydrolase fold has been described as consisting of parallel, eight-stranded β sheet surrounded on both sides by α helices (Fig. 16), providing a stable scaffold for the active sites of a wide variety of enzymes (Ollis *et al.*, 1992; Nardini and Dijkstra, 1999). The catalytic residues always constitute a highly conserved triad: a nucleophile (serine, cysteine or aspartate) positioned after strand β5, an acidic residue almost always positioned after strand

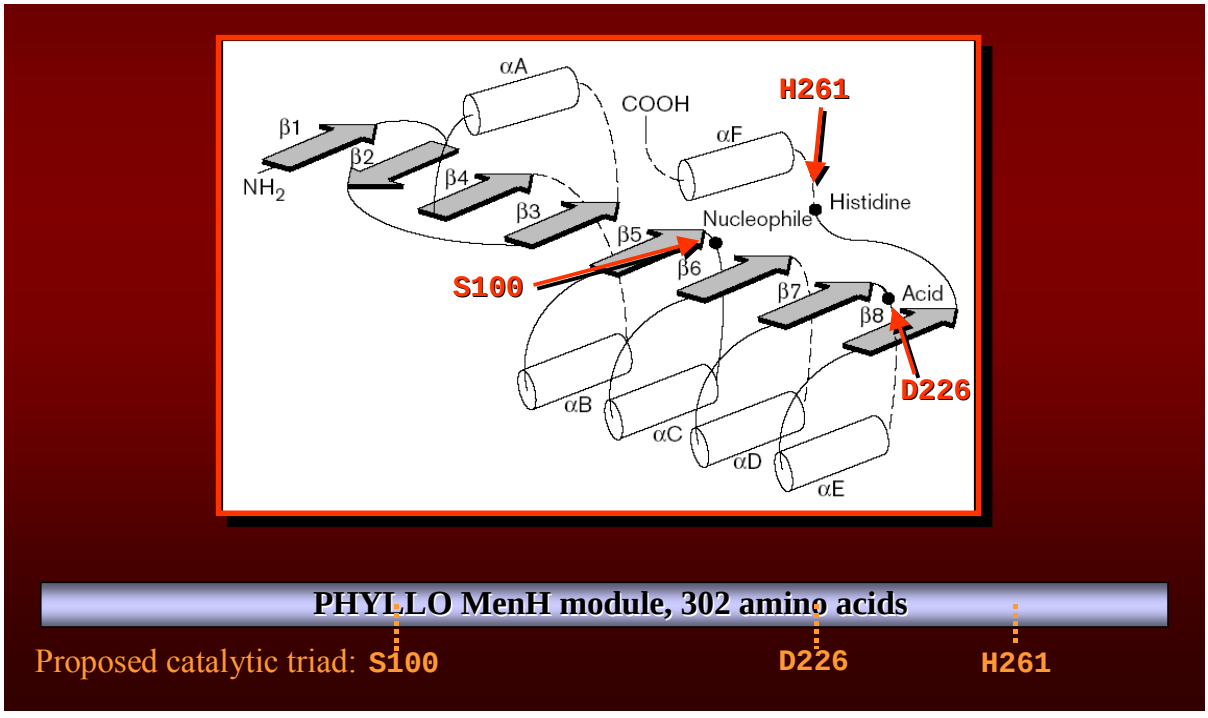


Figure 16. The proposed catalytic triad for the *Arabidopsis* MenH module of PHYLLO. The secondary structure of the canonical α/β hydrolase fold (according to Ollis *et al.*, 1992) is showed above. The α helices and β strands are represented by white cylinders and grey arrows, respectively. The location of the proposed catalytic triad S100, H261 and D226 for the PHYLLO MenH is indicated by arrows and black dots in the upper cartoon and by doted lines in the schematic structure of the MenH module displayed at the lower part of the figure.

β7 and an absolutely conserved histidine following the last β strand (Ollis *et al.*, 1992; Nardini and Dijkstra, 1999). In the *Arabidopsis* MenH module of PHYLLO such highly conserved residues are observed in the relative position of the protein structure. We propose that the nucleophile S1455, the acidic D1639 and the invariant H1674 are the three residues involved in catalysis in the *Arabidopsis* MenH module of PHYLLO, as well as in the other MenH proteins (Fig. 16; Fig. 17).

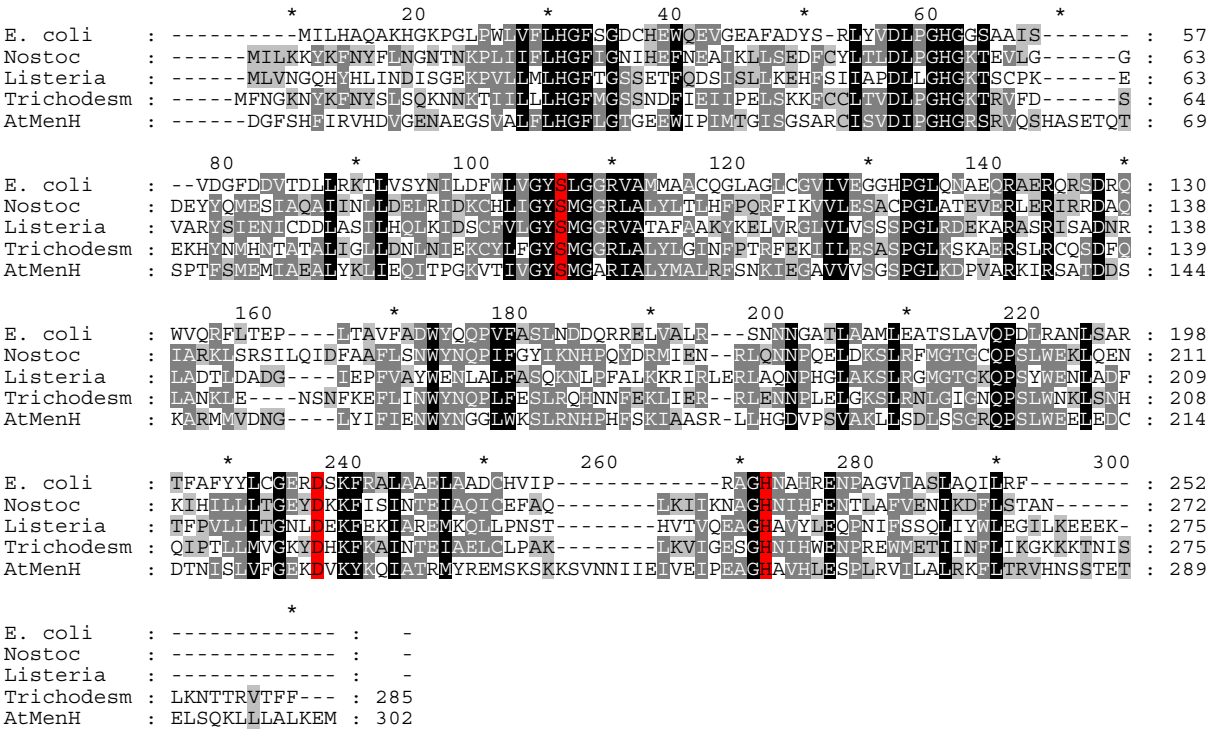


Figure 17. Alignment of the *Arabidopsis thaliana* MenH module of PHYLLO (AtMenH) with MenH proteins of the representative eubacteria *Escherichia coli* K12 (E. coli, accession number AAC75323), *Nostoc punctiforme* (Nostoc, ZP_00112005), *Listeria innocua* (Listeria, CAC97013), and *Trichodesmium erythraeum* IMS101 (Trichodesm, ZP_00071896.1). Residues with black background (X) -100% conservation, grey background (X) -60% conservation, grey background with black letter (X) -30% conservation. Residues coloured in red correspond to proposed catalytic triad discussed in the text and depicted in the Fig. 16.

3.6.5. The MenF 5'-module

The PHYLLO MenF module comprises 198 amino acids ranging from the residue at position 75 to the position 272 (Fig. 10). BLAST searches with the MenF module returned

hits of minimum e-value of 2e-17 and similarities ranging from 45 to 50% with another plant isochorismate synthases experimentally demonstrated to have this enzymatic activity (van Tegelen *et al.*, 1999; Wildermuth *et al.*, 2001). This result unambiguously identifies the relationship of the first module of PHYLLO with MenF proteins that catalyse the first step of the MQ and PhQ pathways (Fig. 3). But, surprisingly, the full extension of 198 amino acid residues of the MenF module aligns only with the N-terminal part of isochorismate synthases (see sections 3.8 and 3.9), clearly indicating that the first module of *PHYLLO* encodes a truncated MenF protein. The absence of the C-terminal part strongly suggests that the MenF module of PHYLLO (hereafter called MenF 5'-module) is not functional. This topic will be further discussed in the sections 3.8 and 3.9.

3.7. Synteny of the *PHYLLO* locus among different kingdoms

3.7.1. Conservation of *PHYLLO* in higher plants

A TBLASTN search with PHYLLO as query sequence against the GenBank non redundant database retrieved hits of minimum e-value 1e-73 with the rice genome, revealing a cluster of *men*F, *men*D, *men*C and *men*H homologous regions present in the positions 22.434.685 bp to 22.447.732 bp on the chromosome 2 of this plant (Fig. 18). An associated cDNA (accession number AK120415) clearly covers the *men*F, *men*D and *men*C homologues regions, implying that also in rice there is a transcriptional link between *men*-similar regions. A conserved *PHYLLO* locus is also present in the genome of the tree *Populus trichocarpa* with the four *men*-homologous regions grouped in the position 13.272.139 to 13.292.047 of the scaffold LG_X (http://genome.jgi-psf.org/euk_cur1.html). The corresponding gene model fgenesh1_pg.C_LG_X001248 predicts the fusion of *men*F, *men*D, *men*C and *men*H modules (Fig. 18). Interestingly, like the *Arabidopsis PHYLLO*, both rice and *Populus* have a truncated *men*F 5'-module (Fig. 18) preceded by a predicted transit peptide for chloroplast targeting, indicating common structural characteristics for *PHYLLO* homologs in plants. In addition to these genomic sequences, several expressed sequence tags (ESTs) from diverse plant organisms presenting significant similarities to the modules of *PHYLLO* were found in databases. Taken together, these results strongly indicate the conservation of *PHYLLO* among higher plants, both in monocots and dicots.

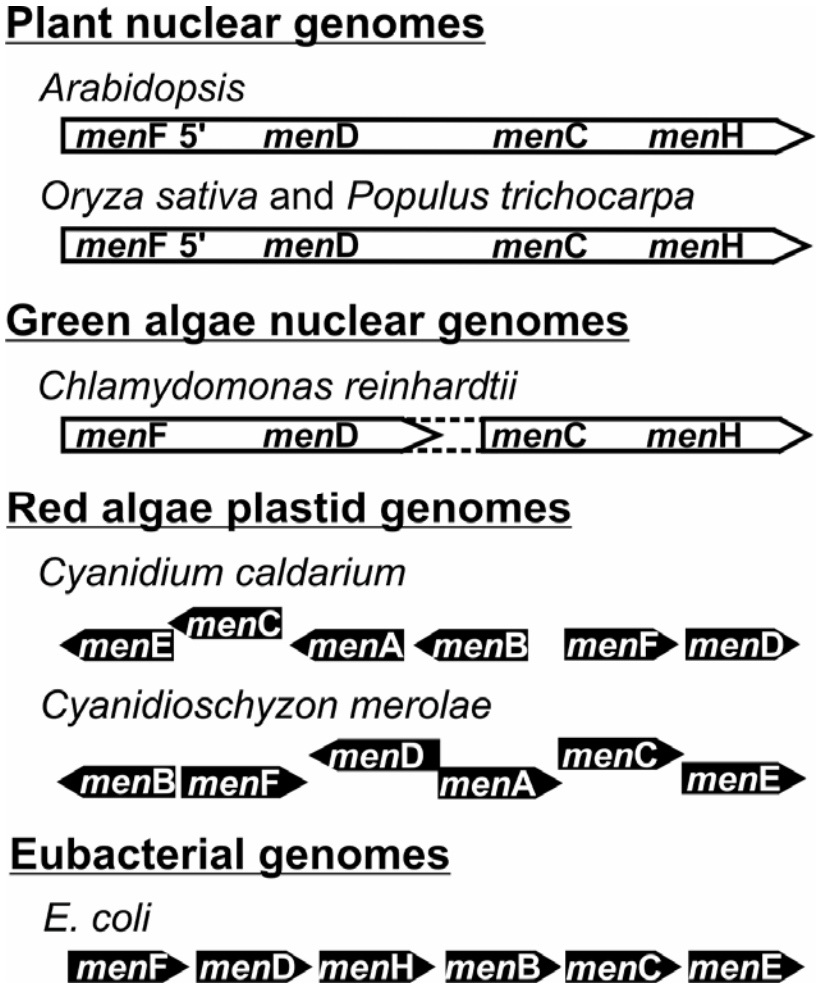


Figure 18. Conservation of the PHYLLO architecture. The *PHYLLO* homologs found in plant and green algae nuclear genomes are indicated by open boxes. Black boxes represent *men* genes found in red algae plastomes and eubacterial genomes. The accession numbers of the genes are: *Arabidopsis* (DQ084386), rice (AP008208), *Populus* (gene prediction fgenesh1_pg.C_LG_X001248), *Chlamydomonas* (gene models C_60079 and C_2490001), *C. caldarium* (AF022186), *C. merolae* (AB002583), and the representative eubacterial genome of *E. coli* (U00096). A fusion of the two gene models (traced lines) in the green alga is still a matter of dispute due to the incomplete sequence status of the *Chlamydomonas* genome project. The tips of the black bars indicate the transcriptional orientation. The vertical displacements of black boxes in the red algae operons indicate overlapping of genes.

3.7.2. Conservation of *PHYLLO* in green algae

The updated draft genome sequence of the green algae *Chlamydomonas reinhardtii* reveals that the four *men*-homologous regions are grouped in two different scaffolds (Fig. 18). The scaffold_6 encodes linked *men*F and *men*D regions from the position 649.286 bp to

657.243 bp (http://genome.jgi-psf.org/euk_cur1.html). The *men*C and *men*H homologs are grouped in-between 403 bp and 4.347 bp of the scaffold_247. For both clusters there are gene models (C_60079 and C_2490001, respectively) predicting a transcriptional fusion between the contiguous regions. The deduced *men*FD and *men*CH products present homology to the related Men proteins and show hits in searches against the Conserved Domain Database of e-value of 8e-43, 2e-66, 6e-05 and 1e-07 with the MenF, MenD, MenC and α/β hydrolase fold domains, respectively. Although no tetramodular fusion can be predicted from the unfinished sequence status of the *Chlamydomonas* genome project, the observed clusters provide evidence for an architectural conservation of the *PHYLLO* locus in green algae. Interestingly, in a TBLASTN search to the *Chlamydomonas* draft genome in January 2003 the four *men*-homologous regions were grouped in the interval 3.913 bp to 17.012 bp of the single scaffold_794 (data not shown). Presently, this scaffold contains a different sequence. Therefore, it remains possible that the *PHYLLO* structure, fusing *men*FDCH, is conserved in *Chlamydomonas* and that the *men*FD and *men*CH regions have been only recently false assembled into two different scaffolds.

3.7.3. Partial conservation of the architecture of *PHYLLO* in red algae plastomes and eubacterial *men* operons

Remarkably, *men* genes are organized into operons in the plastid genomes of the red algae *Cyanidium caldarium* and *Cyanidioschyzon merolae* (Fig 18). Taking into account that the chloroplast is a remnant of an ancestral cyanobacterium (Raven and Allen, 2003; Timmis *et al.*, 2004) it is conceivable that *PHYLLO* was originated from the structure of an ancestral operon in the chloroplast and was transferred into the nuclear genome during endosymbiosis. Such an inter-kingdom transfer is supported by the structural similarity of *PHYLLO* with *men* operons present in red algal plastomes (Fig. 18) and in many eubacterial genomes (Fig. 18, Table 7) that have the same co-localization and transcriptional orientation of *men*F and *men*D genes. Surprisingly, also the archaebacterium *Halobacterium* species NRC-1 has a *men* cluster containing adjacent *men*F and *men*D, presumably as the consequence of a lateral gene transfer (Kennedy *et al.*, 2001), which emphasizes the tendency for this operon to be mobilized as a block. This model is also consistent with previous studies demonstrating that juxtaposed coding frames in operons are at the root of gene fusions, both in prokaryotes and eukaryotes (Yanai *et al.*, 2002; Andersson & Roger, 2002). A prokaryotic origin for *PHYLLO* is further supported by the chloroplast localization of its product (Fig. 11).

Table 7. Representative bacterial genomes with co-localized *men*F and *men*D genes. Co-localization of *men*F and *men*D genes on chromosomes of 35 representative eubacteria and of the archaebacterium *Halobacterium* species NRC-1. The nucleotide positions are given in bp according to BLAST alignments against sequences deposited under the indicated accession numbers.

Bacteria	**Accession Number**	***Men*F**	***men*D**
Escherichia coli	U00096	2.378.440-2.377.373	2.377.281-2.375.614
Bacillus subtilis	BSUB0016	138.699-137.935	137.836-136.196
Shigella flexneri	AE005674	2.393.148-2.392.081	2.391.989-2.390.322
Salmonella enterica	AE017220	2.432.731-2.431.667	2.431.575-2.429.908
Photorhabdus luminescens	BX571869	114.368-113.286	113.178-111.487
Yersinia pseudotuberculosis	BX936398	3.030.337-3.029.216	3.028.976-3.027.288
Yersinia pestis	AE017135	279.352-278.231	277.991-276.303
Erwinia carotovora	BX950851	1.377.972-1.379.066	1.379.344-1.381.026
Mannheimia succiniciproducens	AE016827	1.791.189-1.790.140	1.790.122-1.788.434
Pasteurella multocida	AE006039	2.955-3.933	3.959-5.648
Haemophilus influenzae	U32714	9.631-8.848	8.828-7.139
Photobacterium profundum	CR378671	254.435-253.398	253.320-251.644
Vibrio parahaemolyticus	BA000031	965.550-966.635	966.665-968.272
Vibrio vulnificus	BA000037	1.122.835-1.123.857	1.123.887-1.125.455
Chlorobium tepidum	AE006470	1.740.409-1.741.224	1.741.260-1.742.891
Vibrio fischeri	CP000020	1.885.507-1.884.461	1.884.437-1.882.845
Vibrio cholerae	AE004272	13.701-12.616	12.625-10.877
Geobacillus kaustophilus	BA000043	2.902.661-2.901.873	2.901.838-2.900.153
Bacillus licheniformis	AE017333	3.104.850-3.104.086	3.104.008-3.102.332
Oceanobacillus iheyensis	BA000028	2.377.802-2.377.019	2.377.014-2.375.328
Listeria monocytogenes	AL591980	81.141-80.446	80.365-78.677
Listeria innocua	AL596169	230.736-230.029	229.960-228.272
Staphylococcus epidermidis	AE016746	126.670-127.455	127.514-129.115
Enterococcus faecalis	AE016948	89.427-90.419	90.512-92.161
Staphylococcus aureus	BX571856	1.058.076-1.058.891	1.058.920-1.060.527
Bacillus thuringiensis	AE017355	4.637.740-4.636.940	4.636.871-4.635.192
Bacillus cereus	AE017280	11.571-10.771	10.702-9.065
Bacillus anthracis	AE017225	4.635.402-4.634.602	4.634.533-4.632.896
Listonella anguillarum	AY738107	768-1.853	1.882-2.212 (partial)
Porphyromonas gingivalis	AE017177	92.616-91.839	91.798-90.091

Lactococcus lactis	E006306	11.229-10.294	10.258-8.573
Bacteroides thetaiotaomicron	AE016946	126.546-127.274	127.354-128.979
Bacteroides fragilis	CR626927	1.558.622-1.559.383	1.559.443-1.561.077
Desulfotalea psychrophila	CR522870	275.724-274.912	274.802-273.432
Parachlamydia-related symbiont	BX908798	1.276.351-1.275.692	1.275.701-1.274.181
Halobacterium species NRC-1	AE005040	9.828- 8.833	8.788- 7.079

3.8. Genetic characterization of the MenF enzymatic function in *Arabidopsis*

3.8.1. The *ICS*1 and *ICS*2 genes in *Arabidopsis*

The absence of the 3'-region from the *men*F 5'-module of *PHYLLO* in *Arabidopsis*, rice and *Populus* suggests that this module may not encode a functional ICS in higher plants. There are two additional *ICS* genes present in the *Arabidopsis* genome that could provide a functional product for the first step of the PhQ biosynthesis (Fig. 3). The *ICS*1 gene (At1g74710), localized at the bottom of chromosome I on position 28.074.052 bp to 28.077.556 bp, is a 3.448 bp long gene containing 13 exons that corresponds to a coding sequence of 1.710 bp (accession number AY056055). The other paralog, *ICS*2 (At1g18870), is localized at the top of chromosome I spanning from position 6.515.737 bp to 6.519.164 bp and is a 3.431 bp long gene containing 14 exons that corresponds to a coding sequence of 1.539 bp (NM_101744). T-DNA insertion lines related to these genes were analysed (Fig. 19). The Salk_042603 line has a T-DNA integrated in the intron 4 of the *ICS*1 gene at the position 1.888 bp with reference to the start codon. The Salk_073287 line has a T-DNA integrated in the first intron of the *ICS*2 gene at the position 493 bp relative to the start codon (Fig. 19). Single knockouts of these genes showed no photosynthetic defects, whereas double *ICS* knockout-plants, generated by crossing the single knockouts, demonstrated a clear *pha* phenotype with complete lack of PhQ (Fig. 4 and Table 3), indicating unambiguously that, under normal conditions, ICS1 and ICS2 overlap in their function to synthesize the isochorismate precursor of PhQ and that the product of the *men*F 5'-module of *PHYLLO* is unable to complement this deficiency.

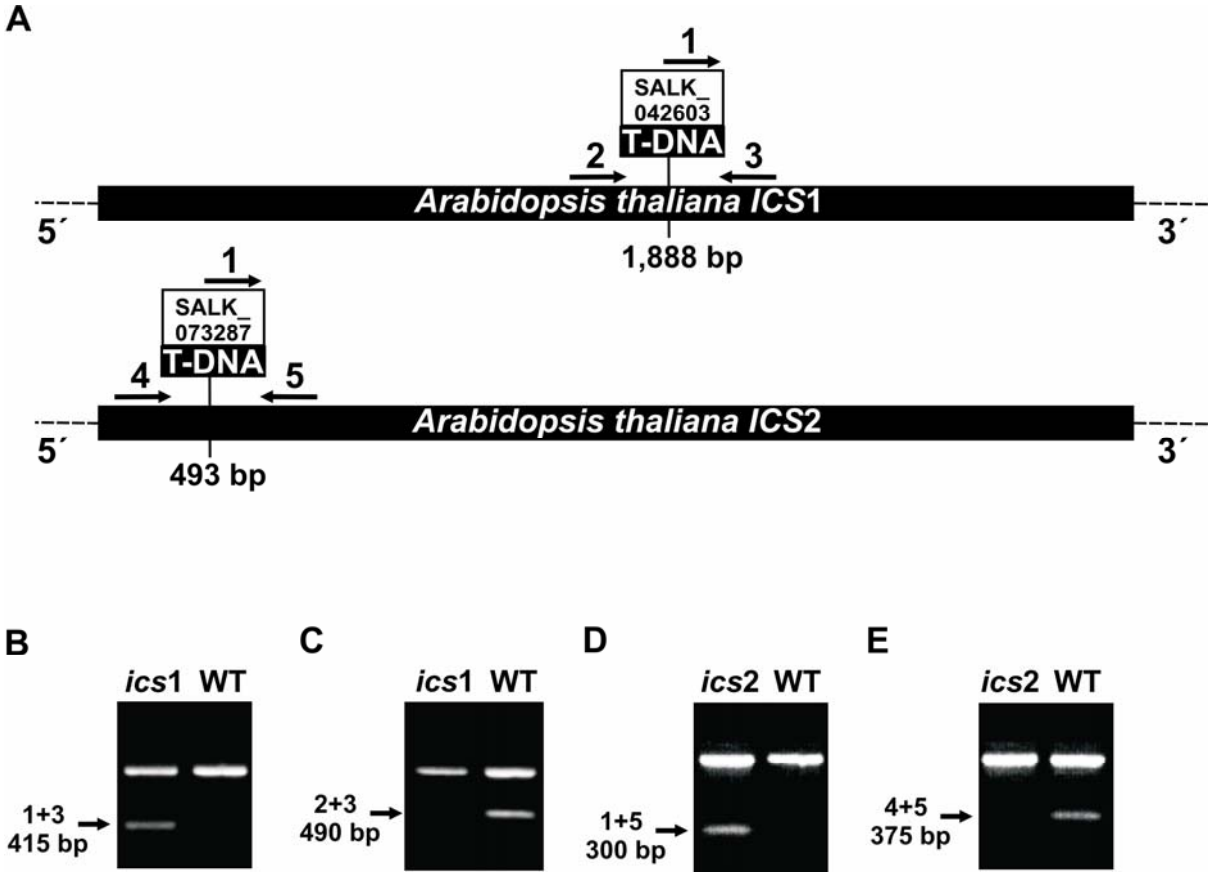


Figure 19. The *ICS*1 and *ICS*2 knockouts (**A**) Schematic representation of the *ICS*1 and *ICS*2 locus indicating the locations (relative to the start codon) of T-DNAs Salk_042603 and Salk_073287 insertions in the *ICS*1 and ICS2 genes, respectively. The positions of primers (arrows) used for screening of the knockout plants are also represented schematically. (**B**) PCR product of 415 bp using primers 1 (LBb1) and 3 (ICS1B_rev) indicates the presence of the T-DNA in *ics*1 mutant. (**C**) PCR product of 490 bp using primers 2 (ICS1C_for) and 3 is only present in the wild-type (WT) plant, indicating homozygozity of the T-DNA insertion in the *ics*1. (**D**) PCR product of 300 bp using primers 1 and 5 (ICS2B_rev) indicates the presence of the T-DNA in *ics*2 mutant. (**E**) A PCR product of 375 bp using primers 4 (ICS2B_for) and 5 is only present in the wild-type (WT) plant, indicating homozygozity of the T-DNA insertion in the *ics*2. In all PCR reactions a positive control amplified from primers Ac_for and Ac_rev were used, generating a product of 980 bp (upper bands in panels B-E).

3.8.2. The ICS proteins of *Arabidopsis*

The ICS1 protein (accession number AAL17715) is a 569 amino acids long polypeptide deduced from the cloned cDNA AY056055. This protein is preceded by a predicted transit peptide to the chloroplast with a proteolytic cleavage site at position 45. The ICS2 (NP_173321) consists of 512 amino acid residues containing no predicted transit peptide. Since the sequence NP_173321 is a computational prediction from the *Arabidopsis* genome, it remains possible that the 5'-part of the putative mRNA is missanotated. Biochemical evidence for an ICS activity has been provided for the close ICS homolog of *Catharanthus roseus* (van

Tegelen *et al.*, 1999). Also, mutational analysis of *ICS*1 in *Arabidopsis* has shown that the gene product is essential for an intermediate step of the salicylic acid (SA) synthesis required for plant defence (Wildermuth *et al.*, 2001).

The ICS proteins catalyse the replacement of the hydroxyl group from the position C4 to C2 of the chorimate ring with an oxygen atom of a water molecule as displacement agent in the reaction (He *et al.*, 2004). The end product is isochorismate. Five common chorismate utilizing enzymes have been characterized so far. These include ICS, anthranilate synthase (AS) 4-amino-4-deoxychorismate synthase (ADCS), chorismate mutase and chorismate lyase. A common feature of these enzymes is the presence of a conserved C-terminal domain required for binding the chorismate precursor and for the catalytic process (He *et al.*, 2004). For AS of *Serratia marcescens* and ADCS of *E. coli* there are X-ray structures available (Spraggon *et al.*, 2001; Parsons *et al.*, 2002) enabling a model for the chorismate binding pocket (Fig. 20) (He *et al.*, 2004). The amino acids implicated in contacts with chorismate and the cofactor Mg^{+2} are all conserved in the two ICS proteins of *Arabidopsis* (Fig. 20; Fig. 21).

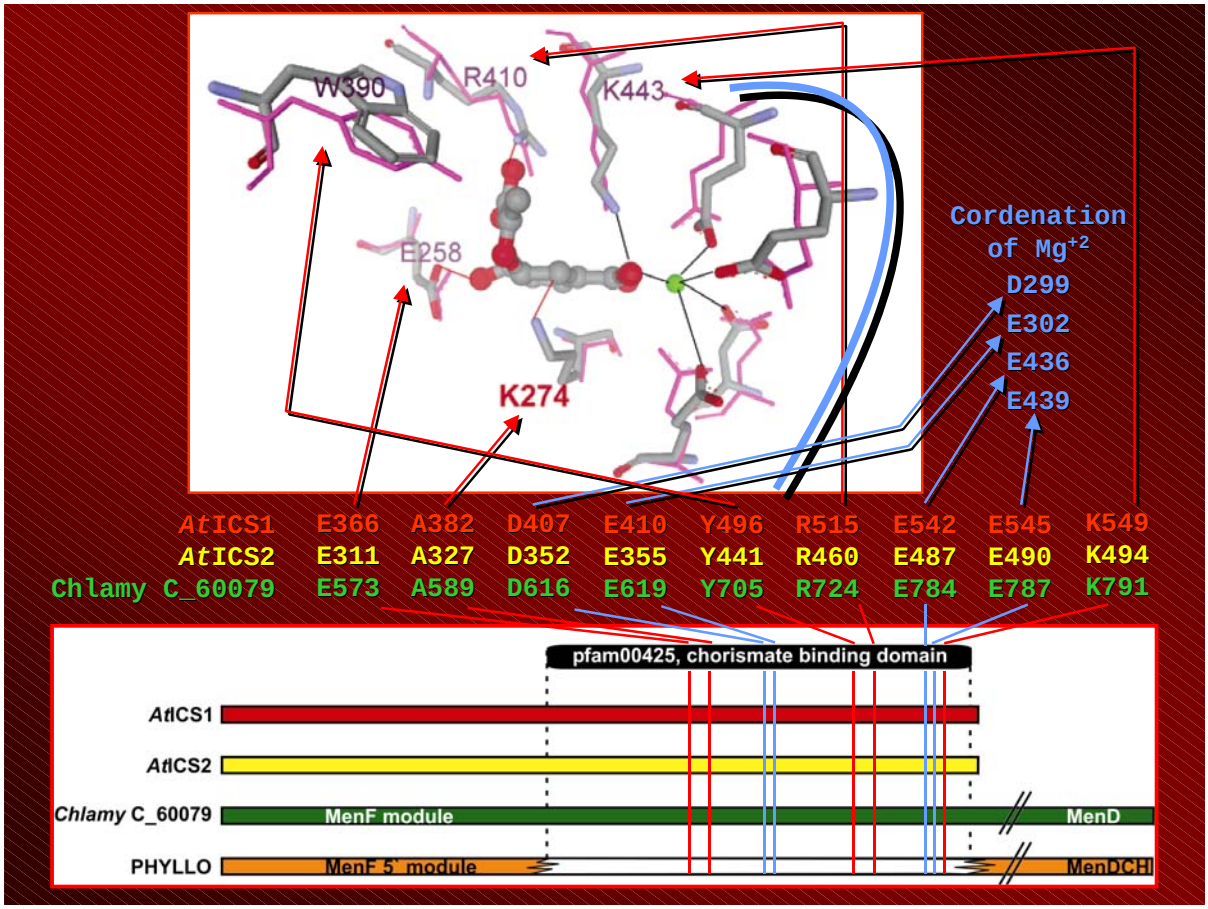


Figure 20. Schematic representation of the conservation of the chorismate binding domain. The protein structure present above represents the binding pocket for chorismate (docked in the centre) by a superposition of the X-ray structures of *E. coli* ADCS (CPK colors) and *Serratia marcescens* AS (purple) modelled according to He *et al.* (2004). The specific amino acids that make contact with the chorismate are labeled corresponding to the ADCS structure. The green atom is Mg^{+2} and the four amino acids that coordinate this ion are shown in the right part of the cartoon (adjacent to the blue arch) and listed in blue at the right part outside the cartoon. The middle part of the figure represents amino acids present in the primary structure of the *Arabidopsis* ICS1 (*At*ICS1, in red), ICS2 (*At*ICS2, in yellow) and the *Chlamydomonas* gene model C_60079 product (Chlamy C_60079, in green) that are conserved to the respective amino acids present in ADCS and AS (indicated by arrows) that are in contact with chorismate (red arrows) or to the Mg^{+2} cofactor (blue arrows). The position of these amino acids in the respective protein structures is indicated by line projections from these residues to the schematic alignment of AtICS1, AtICS2, the product of *Chlamydomonas* gene model C_60079 and PHYLLO from plants. All of them tracing to the C-terminal part of their counterpart proteins, the chorismate binding domain (pfam00425, according to the CDD). The "fissures" in PHYLLO MenF module of plants (orange) remark the gene splitting event resulting in the loss of this functional chorismate biding domain from these proteins. This domain is present in the MenF module of the *Chlamydomonas* PHYLLO homolog.

The absence of these functional amino acids from the MenF 5'-module of PHYLLO in plants, as well as the whole C-terminal part, is evident from the scheme of figure 20 and the alignment of figure 21, demonstrating unequivocally that this module is not a functional chorismate-utilizing enzyme.

```
                        *        20         *        40         *        60         *
AtMenF5'   : -----------MRSSFLVSNPPFLPSLIPRYSSRKSIRRSRERFSFPESLRVSLLHGIQRNIEVAQGVQFDGPIM :  64
PtMenF5'   : -----------MKPHILLLTNNLPLPSPNLPFGTRKSLPLAFLSRNTSIHFPRFQSPKFKVVEAVRFDCPVTDV :  63
OsMenF5'   : MLAVVFSSGHRLLPLPVLPGTFTTPPPPPPPPLLSPRRPLLAPRRRRCLCGGGGGGLLLLRAVAARRAGIVIDVD :  75
AtICS1     : ----------------------------MASLQFSSQFLGSNTKTHSSIISISRSYSPTPFTRFSRKKYESCSMS :  47
AtICS2     : ------------------------------------------------------------------------------ :   -
CrICS      : ---------------------MASITGHCVAHFTDLSTRKSSFFSNSNNNSSLFRRKSTNIVTRKKYIFCSTSLS :  54

              80         *       100         *       120         *       140         *
AtMenF5'   : DRDVNLDDDLVVQVCVTRTLPPALTLELGLESLKEAIDELKTNPPK-SSSGVLRFQVAVPPRAKALFWFCSQPTT : 138
PtMenF5'   : TELEGEDCELVLETCITRTLPPALTLERGIESIKAAVDDLKSNPPC-SLHGVFRFQVAVPPSPKALNWFCSLPES : 137
OsMenF5'   : EVGEVGDRDLPVDVSFTRRLPPVLTLGDGLAALRRAGEEVKACPPAAAASGVIRFEVLVPPSTKALKWLCTQFKR : 150
AtICS1     : MNGCDGDFKTPLGTVETRTMTAVLSPAAATERLISAVSELKSQPPS-FSSGVVRLQVPIDQQIGAIDWLQAQNE- : 120
AtICS2     : MNGCEADHKAPLGTVETRTLSTVPSPAAATERLITAVSDLKSQPPP-FSSGIVRLQVPIEQKIGAIDWLHAQNE- :  73
CrICS      : MNGCNGDPRAPVGTIETRTLPAVSIPALAMERLSSAVANLKSTLPS-AQSGIIRLEVPIEEHIEALDWLHSQDQK : 128

                160         *       180         *       200         *       220
AtMenF5'   : SDVFPVFF--LSKDTV-EPSYKSLYVKEPH-------------------GVFGICNAFAFVHSSS-VDSNGHSMI : 190
PtMenF5'   : DGVFPRFF--LSKETE-DASCKKLYLHRTR-------------------GVFGLGSAICEEASSYRAPEKLK-RI : 189
OsMenF5'   : SSLFPQFY--LSRKQTTDSSIQLEIS-------------------------GAGSAICF-HGSSRVD-NGFDLI : 195
AtICS1     : -IQPRCFFSRRSDVGRPDLLLDLANENGNGNGNG------TVSSDRNLVSVAGIGSAVFFRDLDPFSHDDWR-SI : 187
AtICS2     : -ILPRSFFSRRSDSGRPDLLQDFSSDNG-------------SSDHNPVSVAGIGSAVFFRDLDPFSHDDWR-SI : 132
CrICS      : NLLPRCYFSGRSQVTFSDFTSNDLTNRNGSAANGHLQRISTSSDDKNLVSVAGVGSAVLFRSPNPFSFDDWL-SI : 202

              *       240         *       260         *       280         *       300
AtMenF5'   : KTFLSDESAMVTAYGFPDIEFNKYSTVNSKD-GSSYFFVPQIELDEHEEVSILAVTLAWNESLSYTVEQTISSYE : 264
PtMenF5'   : RRYLSSDSTHIMTYGFMDINFNKESSSIKHEAGSFYFLIPEVELDEQEEASILVITLAWDENSCWTFEQAIQSFE : 264
OsMenF5'   : SRYLSFNSHLIRAYGSVGVKYDKELLSLEERIGSFYFFIPQVELSEFDGYSMLSSTIVWDDSVSHTFEDSVCLFE : 270
AtICS1     : RRFLSSTSPLIRAYGGMRFDPNGKIAVEWEPFGAFYFSVPQVEFNEFGGSSMLAATIAWDDELSWTLENAIEALQ : 262
AtICS2     : RRFLSSKSPLIRAYGGLRFDPTGKIAVEWEHFGSFYFTVPQVEFDEFGGSSMLAATVAWDNELSWTLENAIEALQ : 207
CrICS      : KRFLSKNCPLIRAYGAIRFDARPHIAPEWKAFGSFYFVVPQVEFDELHGSSMIAATVAWDNALSLTYQQAIVALQ : 277
```

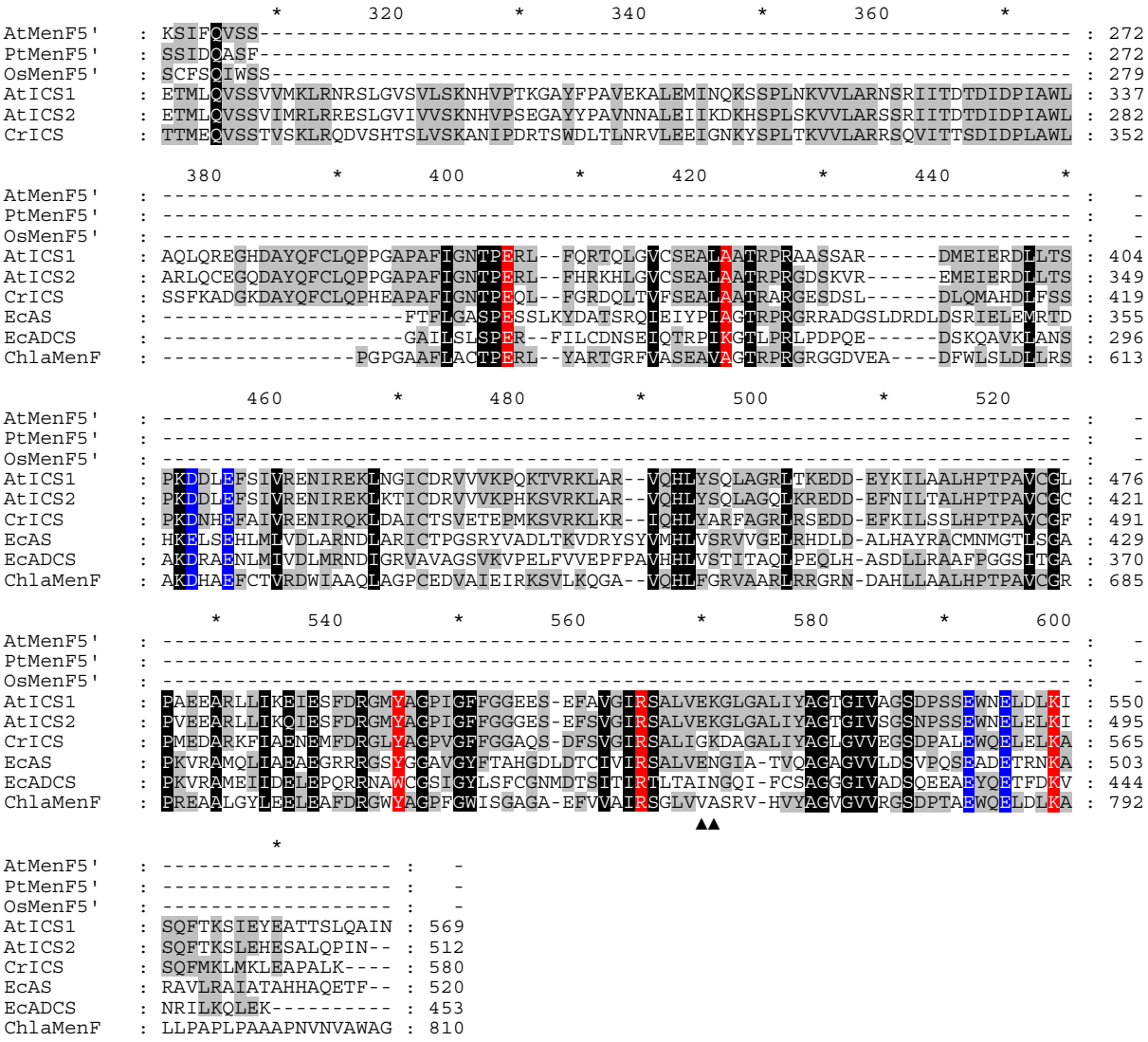


Figure 21. Alignment of the MenF 5´-module of PHYLLO from *Arabidopsis thaliana* (AtMenF 5'), *Populus trichocarpa* (PtMen 5', gene model fgenesh1_pg.C_LG_X001248) *Oryza sativa* (OsMenF 5', from cDNA accession number AK120415) with *Arabidopsis thaliana* ICS1 (AtICS1, AAL17715) and ICS2 (AtICS2, NP_173321), *Catharanthus roseus* ICS (CrICS, CAA06837), *Chlamydomonas* MenF (ChlaMenF, gene model C_60079) and the *E. coli* AS (EcAS, AAC74346) and ADCS (EcADCS, AAC74882). The N-terminal part of EcAS, EcADCS and ChlaMenF proteins were omitted from the alignment due to low similarity to other ICS at the primary sequence level. Also a stretch from the ChlaMenF protein was omitted from the alignment. This is flanked by amino acid positions 729 and 764 of the ChlaMenF protein (▲▲, below the alignment, indicate the flanking amino acids) and exhibits no similarities to the other sequences. The amino acid residues in red represent the ligants to chorismate. The amino acids in blue indicate those involved in coordination of the Mg^{+2} ion. Residues with black background (X) -70% conservation, grey background with black letter (X) -30% conservation.

3.9. A gene splitting event of the 3'-part of the *PHYLLO men*F module in higher plants

Surprisingly, a TBLASTN alignment of the *At*ICS1 protein with the scaffold_6 clearly demonstrates that the similarities between these two proteins spread through the whole ICS1, suggesting that the MenF module in *Chlamydomonas* is full-length. Although the central part is divergent, the presence of the functional C-terminal part is evident from the BLAST graphic output of figure 22, from the scheme of figure 20 and the alignment of figure 21. Furthermore, the amino acid residues required for docking of chorismate and catalysis are all observed in the deduced product of *Chlamydomonas* gene model C_60079 (Fig. 20; Fig. 21) that also presents a hit of e-value of 1e-34 to the chorismate binding domain pfam00425 deposited in the CDD (Fig. 20). Taken together, these results strongly suggest that the *men*F module of the *PHYLLO* homolog in *Chlamydomonas* encodes a functional isochorismate synthase. The fact that the *PHYLLO men*F module is truncated in *Arabidopsis*, rice and *Populus*, strongly suggest that the 3' part of the *men*F module was recently splitted apart from *PHYLLO* in the evolution of higher plants. This resulted in the inactivation of the *men*F product as an isochorismate synthase enzyme, with the corresponding activity taken over by additional copies of *ICS* present in the genome of plants, as verified for the *ICS*1 and *ICS*2 of *Arabidopsis*.

The length of 810 amino acid residues observed for the *Chlamydomonas* MenF module is disproportionate to that of 512 and 569 amino acid residues observed for the two *Arabidopsis* ICS (Fig 20; Fig 21). This could presumably be due to an incorrect prediction of exons assigned to the coding sequence of the gene model C_60079.

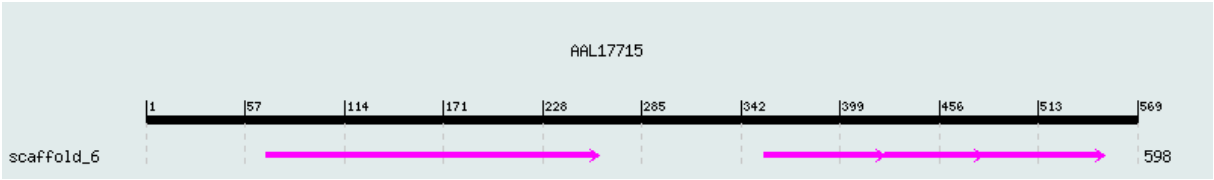


Figure 22. Graphic output from a TBLASTN alignment of the *At*ICS with the scaffold_6 of the *Chlamydomonas* genome. The 569 amino acids from the query sequence AtICS1 (AAL17715) are schematically displayed above. The four pink arrows indicate the extension of the four hits verified against the scaffold_6. Note that no hits were observed for the central part of the ICS1. This can be due to weak similarities or low quality of the sequence reads assign to this part of the scaffold_6. The alignment was done with the BLAST interface of the *Chlamydomonas* genome database (http://genome.jgi-psf.org/euk_cur1.html).

3.10. The phylloquinone content associated to Photosystem I activity in mutant, wild-type and 1,4-dihydroxy-2-naphthoate-fed plants

Interestingly, offspring homozygous for *ICS*1 (*ics*1/*ics*1) and heterozygous for *ICS*2 (*ics*2/*ICS*2) contains only about 18% of the wild-type PhQ content (Table 3) but 50-70% of wild-type PSI activity (data not shown). Similarly, approximately 15% PhQ accumulation in the *pha* mutants after NA supply (Table 3, Fig. 4A) was sufficient to recover up to 70% of PSI amount and activity (Fig. 4B-D). Furthermore, the transformed lines *pha*3G10-15 and *pha*3G10-69 contain above 20% wild-type vitamin K_1 content (Table 3) and normal PSI activity (data not shown), whereas *pha*3G10-55 and *pha*3G10-87 posses below 20% PhQ accumulation (Table 3) and about 80% PSI activity (data not shown). Three major conclusions can be drawn from this non-proportional numerical correlation: (i) PSI is a preferential metabolic sink for newly synthesized PhQ. (ii) An amount of approximately 20-25% PhQ is sufficient for 100% PSI activity. (iii) Consequently, at least 80% of the vitamin K_1 is not associated to PSI and may be located elsewhere in plant membranes. The PhQ pool can be expanded even further from about 3,1 to 5,4 µg/g fresh weight when wild-type, *hcf*145 and *hcf*101 plants are fed with NA (Table 3), notably, without affecting PSI activity (data not shown). The fact that large amounts of vitamin K_1 in *Arabidopsis* were not associated with photosynthesis reinforces previous ideas suggesting that PhQ could also be present in plant plasma membranes (Bridge *et al.*, 2000; Lochner *et al.*, 2003).

3.11. Presence of other *men* genes in the *Arabidopsis* genome

The presence of genes in the *Arabidopsis* genome encoding steps of the PhQ biosynthesis additional to those corresponding to the here described ICS1, ICS2 and PHYLLO and the previously reported At1g60600 (MenA) (Shimada *et al.*, 2005), were accessed by BLAST searches with eubacterial Men proteins as query sequences. The At1g60550 and At1g23360 (AAF87008) annotated proteins present the best BLAST hits to the MenB and MenG homologs of *E. coli* and cyanobacteria, respectively, making them suitable candidates for the respective enzymes of the pathway in *Arabidopsis*. Furthermore, these proteins have N-terminal extensions not observed in the eubacterial homologs that might correspond to transit peptides to the chloroplast. In the case of MenE, the presence of several homologs in the *Arabidopsis* genome makes it difficult to assign a possible candidate.

Interestingly, At1g60550, At1g60600, corresponding to MenB and MenA, are close together encoded in the genome separated by only 16,5 Kb. This close proximity suggests that this gene pair may encode interaction partners. The positions of the proposed genes for the PhQ pathway in *Arabidopsis* are indicated in the map of chromosome I of figure 23.

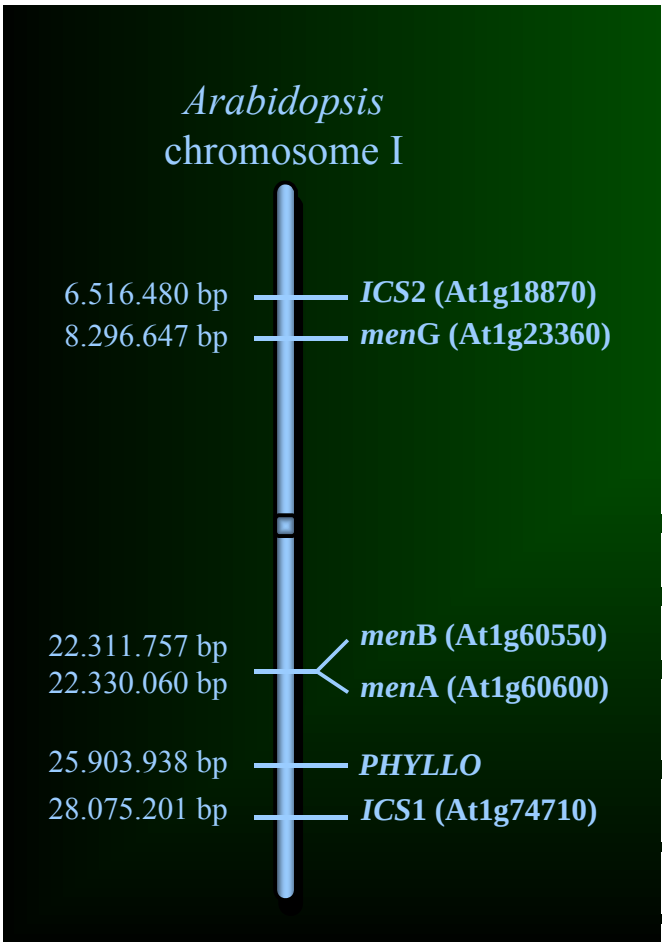


Figure 23. The genomic landscape of the PhQ pathway in *Arabidopsis*. All the known and the putative genes involved in the PhQ biosynthesis are positioned in the chromosome I of *Arabidopsis*. Numbers in bp indicate the star codon of the genes.

4.0. Discussion

4.1. Essential role of phylloquinone in higher plants

4.1.1. The function of *PHYLLO*, *ICS*1 and *ICS*2 in the phylloquinone biosynthesis

This work demonstrated the function of the essential gene *PHYLLO* in *Arabidopsis*. *PHYLLO* has a composite structure whose predicted product displays homology to four known enzymes related to the first steps of the PhQ and MQ biosynthetic pathways, MenF, MenD, MenC and MenH. Additional evidences for the involvement of *PHYLLO* in the PhQ pathway are: (i) the analysis of four allelic *pha* mutants, presenting loss of function of *PHYLLO*, that completely lack PhQ; (ii) the recovery of PhQ in these *pha* mutants after feeding with NA; (iii) the recovery of photoautotrophic growth, PSI function and PhQ accumulation in the *pha*3 and *pha*4 lines complemented with the cDNA form 4 related to *PHYLLO* and with the genomic locus in the case of *pha*3. Furthermore, this work characterized the function of the genes *ICS*1 and *ICS*2 associated to the first step of the PhQ pathway. The homology of the products of these genes to known MenF proteins and the lack of PhQ in the *ics*1/*ics*2 double-knockouts unambiguously demonstrate the requirement of *ICS*1 and *ICS*2 for PhQ biosynthesis.

4.1.2. A photosynthetic defect related to the Photosystem I function in the *pha* mutants

The analysis of *pha* and *ics* mutants showed that PhQ is essentially required for photoautotrophic growth in *Arabidopsis*. This is due to a specific defect in PSI activity and accumulation, as revealed by P700 redox kinetics and western blot analysis. The deficiency at the level of PSI accumulation, which could be partially overcome by supplementation of the PhQ metabolic precursor NA, indicates that the lack of the cofactor was the primary cause for instability of PSI proteins. This is in accordance with the effects observed for the absence of other cofactors of PSI, like chlorophyll and the [4Fe-4S] clusters (Kim *et al.*, 1994; Lezhneva *et al.*, 2004; Amann *et al.*, 2004), and also with the reduction of PSI activity observed in S*ynechocystis* strains in which the PhQ biosynthetic pathway was disrupted (Johnson *et al.*, 2000 and 2003). Nevertheless, 5-15% of the PSI core subunits and a corresponding activity of this complex could be detected in *pha* mutants and *ics* double-knockouts. One possible explanation for this residual activity could be a functional replacement of the missing PhQ by

plastoquinone, as has similarly been reported *in vivo* for S*ynechocystis men*-mutants (Johnson *et al.*, 2000). This idea is also in accordance with several *in vitro* studies showing that a variety of quinone-like compounds bound at the A_1 site after the extraction of the intrinsic PhQs by organic solvents treatment are able to sustain forward electron flow to the [4Fe-4S] clusters (Biggins, 1990; Itoh *et al.*, 2001). However, as observed in the S*ynechocystis men*-mutants, plastoquinone is able to support 50–60% of wild-type levels of PSI activity, contrasting with the basal P700 oxidation observed for the *pha* and *ics* double-knockout mutants. Such a difference indicates that if there is a quinone replacement in PhQ-mutants in *Arabidopsis* it is not as efficient as in cyanobacteria and it is not sufficient to allow photoautotrophic growth of the mutated plants.

In contrast to the *pha* mutants and *ics* double knockouts analysed in this work, the recently described *Arabidopsis Atmen*A mutant (Shimada *et al.*, 2005), also impaired in the PhQ synthesis in a more advanced step of the pathway than PHYLLO, completely missed the basal PSI activity and displayed a significant reduction of PSII polypeptides. This latter characteristic was interpreted as an indirect consequence of the reduced accumulation of plastoquinone to 3% wild-type levels observed in this mutant (Shimada *et al.*, 2005). Furthermore, the *Atmen*A mutant has a marked phenotype with young leaves having a pale-green appearance and old leaves presenting an albinotic aspect. All these characteristics are in sharp contrast with the *pha* and *ics* double-mutants that not only present a specific photosynthetic defect restricted to PSI activity, but always have a pale-green phenotype. The differences observed between phenotypes of mutants involved in the same metabolic pathway could be due to different light intensities used during growth of the plants analysed in this work and for the *Atmen*A knockout plants. It is possible that a high light intensity used during growth of *Atmen*A could account for the severity of this mutant phenotype. Alternatively, the accumulation of the NA precursor in the impaired *Atmen*A mutant could have a possible impact on the metabolomic profile of the mutant, indirectly influencing the plastoquinone content and altering the photosynthetic performance of the *Atmen*A knockout plant. Despite these differences, both MenA and PHYLLO are targeted to plastids when fused to fluorescence reporter proteins, confirming previous results indicating that biosynthesis of PhQ takes place in chloroplasts (Schultz *et al.*, 1981; Kaiping *et al.*, 1984).

4.1.3. The bulk of phylloquinone in *Arabidopsis* is not associated with Photosystem I

Previous reports suggested the presence of PhQ in the plasma membrane of plant cells (Lüthje *et al.*, 1997; Bridge *et al.*, 2000; Lochner *et al.*, 2003), implying that a substantial pool of vitamin K_1 is not associated with PSI. The data present in this work are in accordance with this assumption. Only 15% accumulation of PhQ in NA-fed *pha* mutants (Table 3) allowed a relatively fast re-establishment of 50–70% of PSI accumulation and activity, suggesting a metabolic sink for newly synthesized vitamin K_1 that is preferentially directed to PSI. Similar numbers were also obtained for plants homozygous for *ICS*1 (*ics*1/*ics*1) and heterozygous for *ICS*2 (*ics*2/*ICS*2) that have about 18% wild-type levels of PhQ and 50-70% PSI activity. Furthermore, analysis of the transformed lines *pha*3G10-15 and *pha*3G10-69 possessing about 20% wild-type vitamin K_1 content (Table 3) indicated that these plants have normal PSI activity. Taken together, these data revealed that about 80% of PhQ are probably not part of PSI and may be associated with other cellular compartments. NA supply to wild-type and other PSI control mutant plants significantly increased the PhQ content without any effect on PSI activity. Apparently, this increase to approximately 170% of vitamin K_1 corresponds to a free and flexible pool outside PSI as well.

A transmenbrane redox system (Fig. 24A) has been hypothesised in plants where PhQ plays a key role in shuttling electrons from cytoplasmic-orientated NAD(P)H-dependent quinone reductases to soluble acceptors at the outer surface of the PM (Lüthje *et al.*, 1997; Bridge *et al.*, 2000; Lochner *et al.*, 2003). In accordance with this idea, a NADH oxidase (NOX) was demonstrated to accept electrons from hydroxyquinone and transfer them to O_2 and protein disulfides (Bridge *et al.*, 2000). Although the physiological role of this transmembrane electron chain is still elusive, it has been suggested that it might function in the scavenging of active oxygen forms at the plasma membrane inhibiting lipid peroxidation or proteins carbonylation (Bridge *et al.*, 2000; Lochner *et al.*, 2003). Contrarily, this PhQ-intermediated redox system might be responsible to induce the accumulation of reactive oxygen species as a result of pathogen attack or action of some stressors (Fig. 24B), as has been demonstrated for other NAD(P)H-dependent oxidases in plants, the family of *Rboh* respiratory burst oxidases (Torres and Dangl, 2005). In line with this idea is the observation that the activity of the 31 KDa NAD(P)H-dependent quinone reductase is stimulated by infiltration of tobacco plants with a solution containing bacterial protein-lipopolysaccharide complexes, with a parallel increase in the formation of the product H_2O_2. This has been considering as an evidence for the possible involvement of this PM electron system in the plant defence against phytopathogens (Valenti *et al.*, 1989; Guerrini *et al.*, 1994; Lochner *et*

al., 2003). Moreover, a second NAD(P)H-dependent quinone reductase of 27 KDa was demonstrated to exhibit superoxide formation at low pH (Vianello and Macri, 1989).

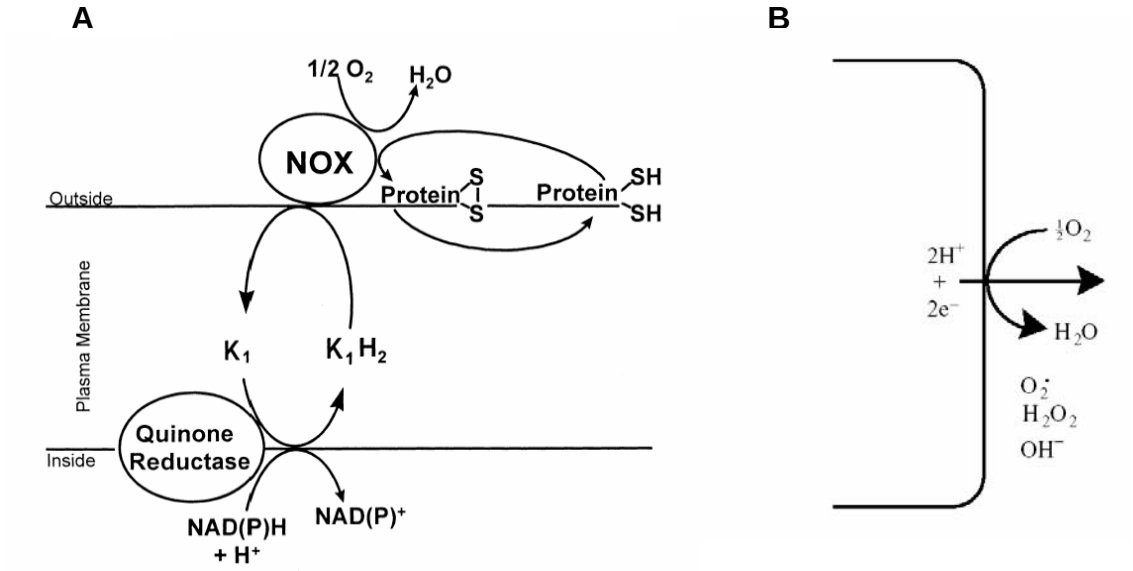


Figure 24. The proposed redox system across the PM of plants. (**A**) Electrons are directed from the cytoplasmic side of the PM to the apoplastic surface by the sequential action of a NAD(P)H quinone reductase and a NADH quinone oxidase (NOX). A pool of vitamin K_1, alternating between the reduced form (K_1H_2) and the oxidased form (K_1), constitutes a mobile pool of PM electron carriers that shuttle electrons across the membrane. The NOX enzyme donates electrons either to O_2 or to protein disulfides. Adapted from Bridge *et al.* (2000). (**B**) Scheme representing the hypothesis that the trans-membrane electron chain depicted in panel **A**, could potentially generate reactive oxygen species. Adapted from Morré *et al.* (2000).

4.2. *PHYLLO*, a plant locus originated from a fusion of four eubacterial genes

4.2.1. *PHYLLO* has a composite structure

This work characterized the *PHYLLO* locus in the genome of *Arabidopsis* as a cluster of encoded-regions homologous with the *men*F, *men*D, *men*C and *men*H genes of eubacteria. Allelism test-crosses and transcription analysis of the *PHYLLO* locus by Northern blot and real time RT-PCR strongly suggest the existence of a single gene for the whole region. Furthermore, the existence of a full-length cDNA (form3, called *phyllo*), containing the four *men*-homologous regions, together with the failure to complement the *PHYLLO* mutations with the truncated form 1 and the successful complementation of two mutants, *pha*3 and *pha*4, with a cDNA form 4 (resembling the form 3), are compelling evidences that *PHYLLO*

has a composite structure fusing the four *men*-modules. Nonetheless, *pha*3c and *pha*4c plants only accumulate approximately 18-24% of the wild-type content of PhQ. This partial complementation can be explained by several reasons: (i) The form 4 has an artificial nature with six additional amino acids not existent in the natural form of PHYLLO (product of form 3). (ii) These results could be interpreted as a consequence of a gene-silencing phenomenon resulting from overexpression or multiple integrations of the complementation construct (Meyer and Saedler, 1996). (iii) The overexpression of PHYLLO causes a dominant-negative phenotype with titration of other subunits of the multienzyme complex for PhQ biosynthesis (see section 4.2.2.). In line with this idea is the fact that none of the *pha*3 lines complemented with the genomic locus could restore 100% PhQ content of the plant (Table 3).

4.2.2. *PHYLLO* was presumably originated from the structure of an operon

Clustering of functionally related genes has been described in the genomes of eukaryotes (Lee and Sonnhammer, 2003; Hurst *et al.*, 2004; Williams and Bowles, 2004). They can vary from sparse groups of genes, like a cluster of genes for avenacin biosynthesis genetically linked in 3,6 centimorgans of the oat genome (Qi *et al.*, 2004), to tight juxtaposed transcriptional units located in the same chromosomal territory, like the *DAL* cluster for allantoin catabolism comprising six adjacent genes in the yeast genome (Wong and Wolf, 2005). Whatever the degree of gene proximity, it has been suggested that the major force driving the formation of gene clusters is an aggregation process directed by selection for co-regulation of gene expression, ensured by the presence of the genes in the same chromatin ambience (Hurst *et al.*, 2004; Wong and Wolf, 2005). This has been interpreted by the notion of dosage balance, postulating that clustering genes whose products physically interact ensures better co-regulation and maintenance of the right stoichiometry of gene products facilitating assembly of functional protein complexes (Teichmann and Veitia, 2004). According to this idea, genes with similar expression profiles are more likely to encode interaction partners. Under this perspective one can say that composite genes, sometimes verified in genomes of eukaryotes, like the pentafunctional *aro*M in *Aspergillus nidulans* (Charles *et al.*, 1986) and the here described *PHYLLO*, may represent examples of higher degree of gene clustering, co-expression and consequent association of its encoded protein domains.

Clustering of functionally related genes in operons is common in the genomic organization of prokaryotes (Lawrence, 2003). Therefore, an alternative mechanism for the formation of gene clusters might be the transfer and maintenance of already existing blocks of genes from eubacterial-derived organelle genomes to the nuclear genome of eukaryotes. In this case, operon-clustered genes, which have been transferred from the chloroplasts and mitochondria to the nucleus, can not be maintained as polycistronic units (Lawrence, 2003), but selective driving forces may exist that preserve operon-based clusters either in the form of sparse groups of genes or as encoded composite proteins. The former explanation has been suggested for the origin of a cluster of genes in the *Arabidopsis* nuclear genome encoding mitochondrial proteins for DNA and RNA metabolism (Elo *et al.*, 2003). The latter possibility is here proposed for the origin of *PHYLLO* (below).

The presence of *men* genes in polycistronic units of eubacterial chromosomes and in plastomes of red algae indicates that *PHYLLO* could have been originated from a plastidial operon in a fusion event predating or postdating the transference of the gene block into the nuclear genome during the course of endosymbiosis. The possibility of a gene cluster transfer is in accordance with the experimental verification in tobacco that relocation of genes from the chloroplast genome to the nucleus is an ongoing process intermediated by DNA molecules that are not restricted to the size of a single gene, but often involves multiples and large fragments between 6,0 and 22,3 kb (Huang *et al.*, 2004). Also the presence of large stretches of integrated organelle DNA in nuclear genomes of eukaryotes that can even comprise major parts of a genome, like a 620 Kb segment in the nucleus of *Arabidopsis* representing the mitochondria genome (Stupar *et al.*, 2001), demonstrates that organelle-to-nucleus transfer of large multigene blocks is a feasible process (Leister, 2005).

4.2.3. PHYLLO, a prokaryotic metabolon adapted to eukaryotes

The principle architecture of *PHYLLO* has been maintained in the genomes of plants and the green alga *Chlamydomonas*, indicating the existence of a strong selective driven force for the preservation of this eubacterial gene cluster. It is possible that this structural conservation, fusing consecutive enzymatic steps of this pathway, reflects a need of a multienzyme system required to channel metabolic intermediates of PhQ biosynthesis (Dandekar *et al.*, 1998; Marcotte *et al.*, 1999; Tsoka and Ouzounis, 2000; Yanai *et al.*, 2001 and 2002; Winkel, 2004; Jorgensen, K. *et al.*, 2005). This has been similarly suggested for the menaquinone pathway in

E. coli and *B. subtilis*, in which co-localization of *men*F and *men*D genes in operons is essential for the proper channelling of isochorismate, the product of MenF, to the menaquinone pathway via MenD (Rowland and Taber, 1996; Buss *et al.*, 2001). Although gene orders inside operons are commonly lost in long term evolution (Itoh *et al.*, 1999), the tendency to co-localize the *men*F and *men*D genes is extended to many eubacterial genomes (Table 7), in line with the idea that conserved gene orders in operons of different bacteria often correspond to physical interactions of their gene products (Dandekar *et al.*, 1998; Lawrence, 2003). These results suggest that the transcriptional linkage of *men* genes in an operon ensures a co-translational folding and association of Men proteins promoting efficient assembly of a multienzymatic complex (Dandekar *et al.*, 1998; Buss *et al.*, 2001; Lawrence, 2003).

PHYLLO also has co-localized *men*FD modules which are linked to the *men*C module encoding the successive enzymatic step of the pathway (Fig. 25). Therefore, it is conceivable that *PHYLLO* arose from the framework of a transferred *men* operon to functionally adapt an analogous operon-directed association of enzymes to the eukaryotic context of the host genome, that can not decipher polycistrons (Marcotte *et al.*, 1999; Enright *et al.*, 1999; Tsoka and Ouzounis, 2000; Yanai *et al.*, 2001 and 2002; Winkel, 2004; Jorgensen, K. *et al.*, 2005). The single gene fusion ensures expression and multifunctional association of the previously four enzymatic activities in one polypeptide, which is post-translationally imported into the chloroplast. Furthermore, the additional presence of the MenH module in PHYLLO, corresponding to an enzymatic step non-successive to that of MenDC in the pathway, could imply that PHYLLO takes part of a macromolecular enzymatic complex in association with the intermediate enzymes MenE and MenB, which could channel the MenDC products to MenH, as suggested similarly for a bifuctional protein that does not catalyse consecutive reactions in the aspartate pathway (James and Viola, 2002). Moreover, the NA-dependent increase of the PhQ pool was up to five times lower in *pha* mutants than in wild-type, *hcf*101, and *hcf*145 plants (Table 3), uncovering deficiencies of *pha* plants for subsequent enzymatic steps after NA entry, catalyzed by MenA or MenG, presumably due to destabilization of the multienzyme complex in absence of the interacting *PHYLLO* product. This strongly suggests that MenA and/or MenG contribute to the metabolon as well. Also the tight cluster of the genes encoding MenA and MenB separated by 16,5 Kb is in accordance with the dosage balance notion supporting that genes encoding subunits of stable complexes tend to be clustered in the genome (Hurst *et al.*, 2004; Teichmann and Veitia, 2004). In fact, at least 7

out of 8 genes for the PhQ pathway in *Arabidopsis* are encoded on the chromosome I of this plant (Fig. 23).

4.3. A metabolic link between plant resistance and photosynthesis

4.3.1. Gene duplication and splitting events of the *PHYLLO men*F module

The fact that the *PHYLLO men*F module is full-length in *Chlamydomonas*, but is truncated in *Arabidopsis*, rice and *Populus*, strongly suggests that the 3' part of the *men*F module was recently splitted apart from *PHYLLO* in the evolution of higher plants. Although the reason for preservation of the *PHYLLO men*F 5'-region is still elusive and suggest a potential function for this truncated module, it is obvious that the split of the 3'-region resulted in inactivation of the *men*F product as an isochorismate synthase enzyme, with the corresponding activity taken over by additional copies of ICS encoded in the genome of plants, as verified in *Arabidopsis*. Such a gene fission event should have been preceded by establishment of a second copy of an individual *men*F gene, either by direct duplication of the *men*F module of *PHYLLO* or by independent transfer from the plastome. BLAST 2 comparisons of the N-terminal region of ICS1 and ICS2 of *Arabidopsis* against eubacterial and red algae MenF proteins, the translated gene model C_60079 of *Chlamydomonas*, and the product of the *PHYLLO men*F 5'-module, returned hits of minimum e-value of about 1e-18 with PHYLLO MenF 5'-module and 1e-09 with the C_60079 product, whereas no similarity was observed with plastidial and eubacterial MenF proteins. Interestingly, BLAST 2 alignments of the PHYLLO MenF 5'-module to the C_60079 product of *Chlamydomonas* revealed absence of similarity between the two proteins, unambiguously demonstrating that the N-terminal of the ICS proteins of *Arabidopsis* has closer relationship to the product of the *men*F 5'-module of *PHYLLO*. Furthermore, a search against the *Chlamydomonas* draft genome revealed no additional copy of *ICS* besides the one related to the gene model C_60079, encoding the fused MenF and MenD (data not shown). In summary, these data support the idea of a direct duplication of the *PHYLLO men*F module before the splitting of the 3'-region of this module.

4.3.2. The *Arabidopsis* ICS1 protein represents a branching point between phylloquinone and salicylic acid biosynthesis

It is possible that the gene duplication and fission events, establishing an alternative ICS copy and splitting apart this activity from the MenF module of *PHYLLO* product in the course of evolution of higher plants, represent a consequence of a need for at least one separate *ICS* gene outside of the *PHYLLO* context that is especially important under conditions of phytopathogenic attack to induce the metabolic flow from chorismate towards the synthesis of SA required for plant resistance (Wildermuth *et al.*, 2001; Durrant and Dong, 2004; Brodersen *et al.*, 2005). The SA acts as a central signal in the process of systemic and local acquired resistance and is associated with accumulation of pathogenesis-related proteins, which are though to contribute to plant defence. It has been demonstrated in *Arabidopsis* that the isochorismate pathway, initiated by the induction of the *ICS*1 gene, is the major source of SA during systemic acquired resistance (Wildermuth *et al.*, 2001; Durrant and Dong, 2004). The present work provides the genetic evidence that *ICS1* is also required for the synthesis of PhQ and exerts overlapping functions with *ICS2*. Therefore, duplication and inactivation of the *men*F module of *PHYLLO* by a gene split event generated a branching point between PhQ and SA biosynthesis (Fig. 25).

This metabolic link has two possible consequent implications for the partitioning of the metabolic flow between the two routes when *ICS*1 is induced upon infection by pathogens. (i) The product of ICS1 is preferentially channelled to the synthesis of SA at the expense of channelling isochorismate for PhQ biosynthesis. Under this condition, it is conceivable that the isochorismate essential for the PhQ pathway could be favoured by the *ICS2* gene product. Another intriguing possibility is that the switch of the metabolic flow between PhQ and SA biosynthesis could be fine-tuned by regulation of alternative splicing of exons 7 and 8 of *PHYLLO*, dictating the output rate of truncated MenF 5'-module/active *PHYLLO* full-length product. Ironically, alternative splicing takes place in the splitted region of *PHYLLO* in-between the *men*F 5' and the *men*D modules. (ii) Conversely, the biosynthesis of both, PhQ and SA, could simultaneously increase following strong induction of *ICS*1 by phytopathogenic attack. This is in line with the dosis effect of ICS mRNA and the impact of NA feeding on PhQ levels (Table 3) and with a recent report suggesting that PhQ or its intermediates could trigger cell death responses associated with plant defence against phytopathogens (Brodersen *et al.*, 2005). Interestingly, as described in section 4.1.3., the production of H_2O_2 by a NAD(P)H-dependent quinone reductase in PM of tobacco was stimulated by infiltration of the plants with a solution containing bacterial protein-lipopolysaccharide complexes (Valenti *et al.*, 1989; Guerrini *et al.*, 1994). It is possible that

the surplus of PhQ produced by induction of *ICS*1 and not associated with PSI could be directed to PM to sustain the increase in activity of a NAD(P)H-dependent quinone reductase operational during plant defence. In any case, a metabolic link between photosynthesis and plant resistance has now been established and provokes further investigations.

4.4. Concluding remark: the phylloquinone biosynthetic pathway in *Arabidopsis thaliana*

The results present in this work permitted to construct a clear scenario for the biosynthesis of PhQ in *Arabidopsis* that is, presumably, applicable to all higher plants (Fig. 25). The here described function of the new nuclear locus *PHYLLO* and the *ICS*1 and *ICS*2 genes definitively established that the biosynthesis of PhQ in plants is operational in

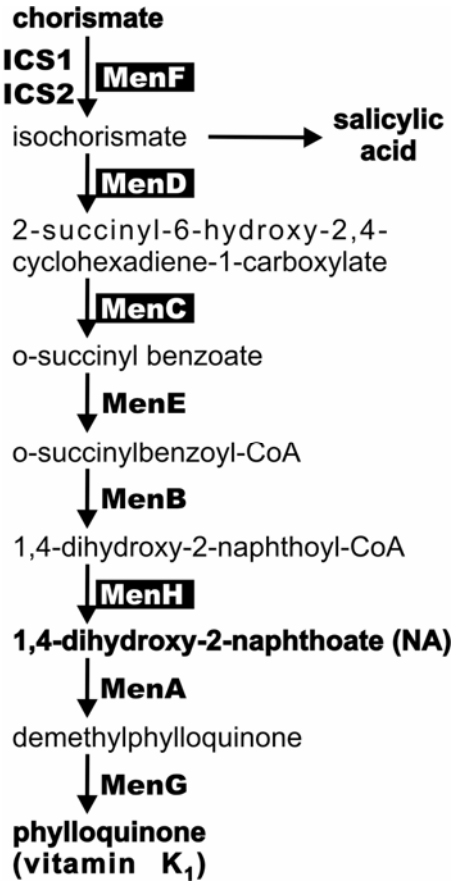


Figure 25. The PhQ pathway in *Arabidopsis* branches off for SA biosynthesis. Bacterial MenF, MenD, MenC, and, MenH proteins (labelled in black boxes) reveal homology to the *PHYLLO* composite. In *Arabidopsis*, ICS1 and ICS2 replaced the corresponding function of the MenF module.

chloroplasts through the *men* pathway. The synthesis started from chorismate that is converted to isochorismate by two redundant enzymes, ICS1 and ICS2. The ICS1 constitutes a branch point for the synthesis of the SA, which is stimulated by pathogen attack. The pathway proceeds via PHYLLO that, by mean of the MenD and MenC modules, converts the isochorismate to *o*-succinylbenzoate. The MenE and MenB (At1g60550) enzymes close the naphthalene ring and presumably establish a metabolic bridging between the MenDC and the MenH modules, this latter constitutes the thiosterase activity of the pathway. The synthesis of PhQ is completed by the phytylation and methylation of the naphthalene ring by the MenA (At1g60600) and MenG (At1g23360) enzymes, respectively. As discussed in section 4.2.2., possibly most of these enzymes are associated in a multienzyme complex necessary for the PhQ biosynthesis.

5. Summary

Phylloquinone is a compound present in all plants serving as cofactor for photosystem I mediated electron transport during photosynthesis. This work reports on the identification and analysis of several *Arabidopsis thaliana phylloquinone absence* (*pha*) and *isochorismate synthase* (*ics*) mutants impaired in the biosynthesis of PhQ (vitamin K_1). Besides the complete lack of PhQ, these plants show a typical phenotype characterized by seedling lethality, photosynthetic defects specifically related to impaired photosystem I accumulation/activity to 5-15% of wild-type levels and partial recovery of 15% PhQ content and 50-70% PSI accumulation/activity after feeding with the metabolic precursor of vitamin K_1, 1,4-dihydroxy-2-naphthoate.

Map-based localization of the mutated allele in the *pha* plants identified a new gene, called *PHYLLO*. It consists of a fusion of four previously individual eubacterial genes, *men*F, *men*D, *men*C, and *men*H, required for the biosynthesis of the photosynthetic phylloquinone in cyanobacteria and the respiratory menaquinone in eubacteria. The fact that homologous *men* genes still reside as polycistronic units in plastomes of red algae and in eubacterial chromosomes strongly suggests that *PHYLLO* derived from an operon present in the proto-organelle precursor of all plastids. The principle architecture of the *PHYLLO* locus is conserved in the nuclear genomes of plants and the green alga *Chlamydomonas reinhardtii*, indicating that selective forces have been acting to maintain the cluster structure in the form of a gene fusion, presumably as an adaptation of an multifunctional association of four enzymatic activities already pre-existing in the chloroplast. In line with this finding, the data present in this work suggest that the *PHYLLO* composite product is part of a metabolon for the biosynthesis of phylloquinone.

The *men*F module of *PHYLLO* in *Chlamydomonas*, encoding the isochorismate synthase activity, is full-length, whereas in higher plants this module surprisingly lacks the functional 3' part, uncovering a recent gene splitting event during evolution. Such a gene fission event, which resulted in inactivation of the encoded ICS enzymatic activity from *PHYLLO*, must have been preceded by establishment of a second functional copy of the *men*F gene. Accordingly, double-knockouts of the *ICS*1 *and ICS*2 genes in *Arabidopsis* analysed during this work, were unable to synthesize PhQ, demonstrating that the activity of the *men*F module of *PHYLLO* has been replaced after the splitting of the 3'-region by at least one more *ICS*

gene present in genomes of higher plants. The fact that *ICS*1 is also required for salicylic acid biosynthesis in *Arabidopsis*, establishes a metabolic link between photosynthesis and systemic acquired resistance. Therefore, gene fusion, duplication and fission events adapted a eubacterial multienzymatic system to the metabolic requirements of plants.

Despite the essential function of PhQ for PSI stability and plant viability, analyses of *ics* heterozygous knockout plants, as well as complementation of the *pha* mutants by NA feeding and transgenic forms of *PHYLLO* demonstrate that the bulk of cellular phylloquinone is not associated with photosystem I, opening the possibility for additional functions of vitamin K_1 in plant cell membranes.

6. Literature

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

Amann, K., Lezhneva, L., Wanner, G., Herrmann, R.G. and Meurer J. (2004) ACCUMULATION OF PHOTOSYSTEM ONE1, a member of a novel gene family, is required for accumulation of [4Fe-4S] cluster-containing chloroplast complexes and antenna proteins. Plant Cell 16, 3084-3097.

Andersson, J.O. and Roger A.J. (2002) Evolutionary analyses of the small subunit of glutamate synthase: gene order conservation, gene fusions, and prokaryote-to-eukaryote lateral gene transfers. Eukaryot. Cell 1, 304-310.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K. and Miyano, S. (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18, 298-305.

Barlow, J.J., Mathias, A.P., Williamson, R. and Gammack, D.B., (1963). A simple method for the quantitative isolation of undegraded high molecular weight ribonucleic acid. Biochem. Biophys. Res. Commun. 13, 61-66.

Barr, R., Pan, R.S., Crane, F.L., Brightman, A.O. and Morré, D.J. (1992) Destruction of vitamin K_1 of cultured carrot cells by ultraviolet radiation and its effect on plasma membrane electron transport reactions. Biochem. Int. 27, 449-456.

Bashin, M., Billinsky, J.L. and Palmer, D.R.J. (2003) Steady-state kinetics and molecular evolution of the *Escherichia coli* MenD [(1*R*,6*R*)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase], an anomalous thiamine diphosphate-dependent decarboxylase-carboligase. Biochemistry, 42, 13496-13504.

Bell, C.J. and Ecker, J.R. (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19, 137-144.

Ben-Shem, A., Frolow, F. and Nelson, N. (2003) Crystal structure of plant photosystem I. Nature 426, 630-635.

Bentley, R. and Meganathan, R. (1982) Biosynthesis of vitamin K (menaquinone) in bacteria. Microbiol. Rev. 46, 241-280.

Biggins, J. and Mathis, P. (1988) Functional role of vitamin K in photosystem I of the cyanobacterium Synechocystis sp. 6803. Biochemistry, 8, 1494 -1500.

Biggins, J. 1990. Evaluation of selected benzoquinones, naphthoquinones, and anthraquinones as replacements for phylloquinone in the A1 acceptor site of the photosystem I reaction center. Biochemistry, 29, 7259-7264.

Blumenthal, T. (1998) Gene clusters and polycistronic transcription in eukaryotes. BioEssays 20, 480-487.

Bonnerjea, J. and Evans, M.C.W. (1982) dentification of multiple components in the intermediary electron carrier complex of photosystem I. FEBS Lett., 148, 313-316.

Booth, S.L. and Suttie, J.W. (1998) Dietary intake and adequacy of vitamin K_1. J. Nutr., 128, 785-788.

Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-254.

Brettel K. and Leibl, W. (2001) Electron transfer in photosystem I. Biochim. Biophys. Acta 1507, 100-114.

Brettel, K., Setif, P. and Mathis P. (1986) Flash-induced absorption changes in photosystem I at low temperature: evidence that the electron acceptor A_1 is vitamin K_1. FEBS lett., 203, 220-224.

Bridge, A., Barr, R. and Morré, D.J. (2000) The plasma membrane NADH oxidase of soybean has vitamin K(1) hydroquinone oxidase activity. Biochim. Biophys. Acta 1463, 448-458.

Brodersen P., Malinovsky, F.G., Hematy, K., Newman, M.A. and Mundy, J. (2005) The Role of Salicylic Acid in the Induction of Cell Death in *Arabidopsis* acd11. Plant Physiol. 138, 1037-1045.

Buss, K., Muller, R., Dahm, C., Gaitatzis, N., Skrzypczak-Pietraszek, E., Lohmann, S., Gassen, M. and Leistner, E. (2001) Clustering of isochorismate synthase genes *menF* and *entC* and channeling of isochorismate in *Escherichia coli*. Biochim. Biophys. Acta. 1522, 151-157.

Campbell, I.M., Coscia, C.J., Kelsey, M. and Bentley, R. (1967) Origin of the aromatic nucleus in bacterial menaquinones. Biochem. Biophys. Res. Commun. 28, 25-29.

Charles, I.G., Keyte, J.W., Brammar, W.J., Smith, M. and Hawkings, A.R. (1986) The isolation and nucleotide sequence of the complex *AROM* locus of *Aspergillus nidulans*. Nucl. Acids Res. 14, 2201-2213.

Choi, S.D., Creelman, R., Mullet, J. and Wing, R.A. (1995) Construction and characterization of bacterial artificial chromosome library for *Arabidopsis thaliana*. Weeds World 2, 17-20.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16, 735-743.

Cordoba, M.C., Serrano, A., Cordoba, F., Gonzales-Reyes, J.A., Navas, P. and Villalba, J.M. (1994) Topography of the 27- and 31-kDa electron transport proteins in the onion root plasma membrane. Biochem. Biophys. Res. Commun. 216, 1054-1059.

Dal Bosco, C., Lezhneva, L., Biehl, A., Leister, D., Strotmann, H., Wanner, G. and Meurer, J. (2004) Inactivation of the chloroplast ATP synthase γ subunit results in high non-photochemical fluorescence quenching and altered nuclear gene expression in *Arabidopsis thaliana*. J. Biol. Chem. 279, 1060-1069.

Dam, H. (1946) The discovery of vitamin K, its biological functions and therapeutical application. Nobel lecture, available online: http://nobelprize.org/medicine/laureates/1943/dam-lecture.pdf.

Dandekar, T., Snel, B., Huynen, M. and Bork, P. (1998) Conservation of gene order: a fingerprint of proteins that physically interact. Trends Biochem. Sci. 23, 324-328.

Daruwala, R., Kwon, O., Meganathan, R. and Hudspeth, M.E.S. (1996) A new isochorismate synthase specifically involved in menaquinone (vitamin K_2) biosynthesis encoded by the *men*F gene.

Doisy, E.A. (1976) An autobiography. Annu. Rev. Biochem. 45, 1-9.

Döring, O., Lüthje, S. and Böttger, M. (1992) Inhibitors of the plasma membrane redox system of Zea mays L. roots. The vitamin K antagonists dicumarol and warfarin. Biochim. Biophys. Acta 1110, 235-238.

Durrant, W.E. and Dong, X. (2004) Systemic Acquired Resistance. Annu. Rev. Phytopathol. 42, 185-209.

Elo, A., Lyznik, A., Gonzalez, D.O., Kachman, S.D. and Mackenzie, S.A. (2003) Nuclear genes that encode motochondrial proteins for DNA and RNA metabolism are clustered in the *Arabidopsis* genome. Plant Cell 15, 1619-1631.

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300, 1005–1016.

Enright, A.J., Iliopoulos, I., Kyrpides, N.C. and Ouzounis, C.A. (1999) Protein interaction maps for complete genomes based on gene fusion events. *Nature* 402, 86–90.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

Felder, S., Meierhoff, K., Sane, A.P., Meurer, J., Driemel, C., Plücken, H., Klaff, P., Stein, B., Bechtold, N. and Westhoff, P. (2001) The nucleus-encoded HCF107 gene of *Arabidopsis* provides a link between intercistronic RNA processing and the accumulation of translation-competent *psbH* transcripts in chloroplasts. Plant Cell 13, 2127-2141.

Feldmann, K.A. (1991) T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. Plant J. 1, 71-82.

Fromme, P., Jordan, P. and Krauß, N. (2001) Structure of photosystem I. Biochim. Biophys. Acta 1507, 5-31.

Gast, P., Swarthoff, T., Ebskamp, F.C.R., Hoff, A.J. (1983) Evidence for a new early acceptor in photosystem I of plants. An ESR investigation of reaction center triplet yield and of the reduced intermediary acceptors. Biochim. Biophys. Acta, 722, 163-175.

Golbeck, J.H. and Kok, B. (1978) Further Studies of the Membrane-Bound Iron-Sulfur Proteins and P700 in a Photosystem I Subchloroplast Particle. Arch. Biochem. Biophys., 188, 233-242.

Golbeck J.H. (2003) The binding of cofactors to photosystem I analyzed by spectroscopic and mutagenic methods. Annu. Rev. Biophys. Biomol. Struct.. 32, 237–256.

Gross, J., Stein, R.J., Fett-Neto, A.G. and Fett, J.P. (2003) Iron homeostasis related genes in rice. Genet. Mol. Biol., 26, 477-497.

Guerrini, F., Lombini, A., Bizarri, M. and Pupillo, P. (1994) The effect of calcium chelators on microssomal pyridine nucleotide-linked dehydrogenases of sugarbeet cells. J. Exp. Bot. 45, 1227-1233.

Haldrup A., Jensen P.E., Lunde, C. and Scheller, H.V. (2001) Balance of power: a view of the mechanism of photosynthetic state transitions. Trends Plant Sci. 6, 301-305.

Hauska, G., Schoedl, T., Remigy, H. and Tsiotis G. (2001) The reaction center of green sulfur bacteria. Biochim. Biophys. Acta 1507, 260-277.

He, Z., Lavoie, K.D.S., Bartlett, P.A. and Toney, M.D. (2004) Conservation of Mechanism in Three Chorismate-Utilizing Enzymes. J. Am. Chem. Soc., 126, 2378-2385.

Henikoff, S. and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89, 10915-10919.

Huang, C.Y., Ayliffe, M.A. and Timmis, J.N. (2004) Simple and complex nuclear loci created by newly transferred chloroplast DNA in tobacco. Proc. Natl. Acad. Sci. USA 101, 9710–9715.

Hurst, L.D., Pal, C. and Lercher, M.J. (2004) The evolutionary dynamics of eukaryotic gene order. Nat. Rev. Genet. 5, 299.

Interschick-Niebler, E. and Lichtenthaler, H.K. (1981) Partition of phylloquinone K_1 between digitonin particles and chlorophyll-proteins of chloroplast membranes from *Nicotiana tabacum*. Z. Naturforsch. 36c, 276-283.

Itoh, T., Takemoto, K., Mori, H. and Gojobori, T. (1999) Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes. Mol. Biol. Evol. 16, 332-346.

Itoh S., Iwaki, M. and Ikegami I. (2001) Modification of photosystem I reaction center by the extraction and exchange of chlorophylls and quinones. Biochim. Biophys. Acta., 1507, 115-138.

Ivanov, I.P., Matsufuji, S., Murakami, Y., Gesteland R.F. and Atkins, J.F. (2000) Conservation of polyamine regulation by translational frameshifting from yeast to mammals. EMBO J. 19, 1907-1917.

Jach, G., Binot, E., Frings, S., Luxa, K. and Schell, J. (2001) Use of red fluorescence protein from *Discosoma* sp. (dsRED) as a reporter for plant gene expression. Plant J., 28, 483-491.

James, C.L. and Viola, R.E. (2002) Production and characterization of bifunctional enzymes. Substrate channeling in the aspartate pathway. Biochemistry 41, 3726-3731.

Johnson, T.W., Shen, G., Zybailov, B., Kolling, D., Reategui, R., Beauparlant, S., Vassiliev, I.R., Bryant, D.A., Jonesi, A.D., Golbeck, J.H. and Chitnis P.R. (2000) Recruitment of a foreign quinone into the A(1) site of photosystem I. I. Genetic and physiological characterization of phylloquinone biosynthetic pathway mutants in *Synechocystis sp*. pcc 6803. J. Biol. Chem. 275, 8523-8530.

Johnson, T.W., Zybailov, B., Jones, A.D., Bittl, R., Zech, S., Stehlik, D., Golbeck, J.H. and Parag R. Chitnis, P.R. (2001) Recruitment of a foreign quinone into the A1 site of photosystem I. *In vivo* replacement of plastoquinone-9 by media-supplemented naphthoquinones in phylloquinone biosynthetic pathway mutants of *synechocystis* sp. PCC 6803. J. Biol. Chem. 276, 39512–39521.

Johnson, T.W., Naithani, S., Stewart Jr., C., Zybailov, B., Jones, A.D., Golbeck, J.H. and Chitnis, P.R. (2003) The *menD* and *menE* homologs code for 2-succinyl-6-hydroxyl-2,4-cyclohexadiene-1-carboxylate synthase and O-succinylbenzoic acid-CoA synthase in the phylloquinone biosynthetic pathway of *Synechocystis sp*. PCC 6803. Biochim. Biophys. Acta 1557, 67-76.

Jordan, P., Fromme, P., Witt, H.T., Klukas, O., Saenger, W. and Krauss, N. (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. Nature 411, 909-917.

Jorgensen, K., Rasmussen, A.V., Morant, M., Nielsen, A.H., Bjarnholt, N., Zagrobelny, M., Bak, S. and Moler, B.L. (2005) Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. Curr. Opin. Plant Biol. 8, 280-291.

Kaiping, S., Soll, J. and Schultz, G. (1984) Site of methylation of 2-phytyl-1,4-naphthoquinol in phylloquinone (vitamin K_1) synthesis in spinach chloroplasts. Phytochemistry 23, 89-91.

Kennedy, S.P., Ng, W.V., Salzberg, S.L., Hood, L. and DasSarma, S. (2001) Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence. Genome Res. 11, 1641-1650.

Kim, J., Eichacker, L.A., Rüdiger, W. and Mullet, J.E. (1994) Chlorophyll regulates accumulation of the plastid-encoded proteins P700 and D1 by increasing apoprotein stability. Plant Physiol. 104, 907-916.

Kjær, B., Frigaard, N.U., Yang, F., Zybailov, B., Miller, M., Golbeck, J.H. and Scheller, H.V. (1998) Menaquinone-7 in the Reaction Center Complex of the Green Sulfur Bacterium *Chlorobium* V*ibrioforme* Functions as the Electron Acceptor A_1. Biochemistry 37, 3237-3242.

Klenchin, V.A., Ringia, E.A.T., Gerlt, J.A. and Rayment, I. (2003) Evolution of enzymatic activity in the enolase superfamily: structural and mutagenic studies of the mechanism of the reaction catalized by *o*-succinylbenzoate synthase from *Escherichia coli*. Biochemistry 42, 14427-14433.

Koike-Takeshita, A., Koyama, T. and Ogura, K. (1997) Identification of a novel gene cluster participating in menaquinone (vitamin K2) biosynthesis. J. Biol. Chem. 272, 12380-12383.

Koncz, C., Martini, N., Szabados, L., Hrouda, M., Bachmair, A. and Schell, J. (1994) Specialized vectors for gene tagging and expression studies. In: Plant Molecular Biology Manual. Vol. B2 (Gelvin, S.B. and Schilperoort, R.A., eds). Dordrecht: Kluwer Academic Publishers, pp. 1-22.

Konieczny, A. and Ausubel, F.M. (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific markers. Plant J. 4, 403-410.

Lamson, D.W. and Plaza, S.M. (2003) The anticancer effects of vitamin K. Altern. Med. Rev., 8, 303-318.

Lawrence, J.G. (2003) Gene organization: selection, selfishness, and serendipity. Annu. Rev. Microbiol. 57, 419.

Lee, J.M. and Sonnhammer, E.L. (2003) Genomic gene clustering analysis of pathways in eukaryotes. Genome Res. 13, 875-882.

Leister, D. (2003) Chloroplast research in the genomic age. Trends in Genet. 19, 47-56.

Leister, D. (2005) Origin, evolution and genetic effects of nuclear insertions of organelle DNA. Trends in Genet. 21, 655-663.

Lezhneva, L. and Meurer, J. (2004) The nuclear factor HCF145 affects chloroplast *psaA-psaB-rps14* transcript abundance in *Arabidopsis thaliana*. Plant J. 38, 740-753.

Lezhneva, L., Amann, K. and Meurer, J. (2004) The universally conserved HCF101 protein is involved in assembly of [4Fe-4S]-cluster-containing complexes in *Arabidopsis thaliana* chloroplasts. Plant J. 37, 174-185.

Lochner, K; Döring, O. and Böttger, M. (2003) Phylloquinone, what can we learn from plants? BioFactors 18, 73-78.

Lunde, C.P., Jensen, P.E., Haldrup, A., Knoetzel, J. and Scheller, H.V. (2000) The PSI-H subunit of photosystem I is essential for state transitions in plant photosynthesis. Nature 408, 613–615.

Lüthje, S., Döring, O., Heuer, S., Lüthen, H. and Böttger, M. (1997) Oxidoreductases in plant plasma membranes. Biochim. Biophys. Acta 1331, 81-102.

Luster, D.G. and Buckhout, T.J. (1989) Purification and identification of a plasma membrane associated electron transfer protein from maize (Zea mays L.) roots. Plant Physiol. 91, 1014-1019.

Lyznik, L.A., Peng, J.Y. and Hodges T.K. (1991) Simplified procedure for transient transformation of plant protoplasts using polyethylene glycol treatment. Biotechniques 10, 294-300.

Mansfield, R.W., Evans, M.C.W. (1986) UV optical difference spectrum associated with the reduction of electron acceptor A_1 in photosystem I of higher plants. FEBS lett., 203, 225-229.

Marchler-Bauer, A. and Bryant, S.H. (2004) CD-Search: protein domain annotations on the fly. Nucleic. Acids. Res. 32, 327-331.

Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., Eisenberg, D. (1999) Detecting protein function and protein-protein interactions from genome sequences. Science 285, 751-753.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M. and Penny, D. (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA 99,12246-12251.

Maxwell K. and Johnson G.N. (2000) Chlorophyll fluorescence-a practical guide. J. Exp. Bot. 345, 659-668.

McMaster, G.K. and Carmichael, G.G. (1977) Analysis of single and double-stranded nucleic acids on polyacrilamid and agarose gels by using glyoxal and acridine orange. Proc. Natl Acad. Sci. USA 74, 4835-4838.

Meganathan, R. (2001) Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): a perspective on enzymatic mechanisms. Vitam. Horm. 61, 173-218.

Meurer, J., Meierhoff, K. and Westhoff, P. (1996) Isolation of high-chlorophyll-fluorescence mutants of *Arabidopsis thaliana* and their characterization by spectroscopy, immunoblotting and RNA gel hybridization. Planta 198, 385–396.

Meurer, J., Plücken, H., Kowallik, K.V. and Westhoff, P. (1998a) A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in *Arabidopsis thaliana*. EMBO J. 17, 5286–5297.

Meurer J., Grevelding, C., Westhoff, P. and Reiss, B. (1998b) The PAC protein affects the maturation of specific chloroplasts mRNA in *Arabidopsis thaliana*. Mol. Gen. Genet., 258, 342-351.

Meurer, J., Lezhneva, L., Amann, K., Gödel, M., Bezhani, S., Sherameti, I. and Oelmüller, R. (2002) A peptide chain release factor 2 affects the stability of UGA-containing transcripts in *Arabidopsis* chloroplasts. Plant Cell 14, 3255-3269.

Meyer, P. and Saedler, H. (1996) Homology-dependent gene silencing in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 23-48.

Mollier, P. Hoffmann, B. Debast, C. and Small, I. (2002) The gene encoding *Arabidopsis thaliana* mitochondrial ribosomal protein S13 is a recent duplication of the gene encoding plastid S13. Curr. Genet. 40, 405-409.

Morgenstern, B., Frech, K., Dress, A. and Werner, T. (1998) DIALIGN: finding local similarities by multiple sequence alignment. Bioinformatics 14, 290-294.

Morré, D.M., Lenaz, G. and Morré, D.J. (2000) Surface oxidase and oxidative stress propagation in aging. J. Exp. Biol. 203, 1513-1521.

Munekage, Y., Hojo, M., Meurer, J., Endo, T., Tasaka, M. and Shikanai, T. (2002) PGR5 is involved in cyclic electron flow around photosystem I and is essential for photoprotection in *Arabidopsis*. Cell 110, 361-371.

Munekage, Y., Hashimoto, M., Miyake, C., Tomizawa, K., Endo, T., Tasaka, M. and Shikanai, T. (2004) Cyclic electron flow around photosystem I is essential for photosynthesis. Nature 429, 579-582.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. 15, 473-497.

Nardini, M. and Dijkstra, B.W. (1999) α/β hydrolase fold enzymes: the family keeps growing. Curr. Opin. Struct. Biol. 9, 732-737.

Nelson, N. and Ben-Shem, A. (2004) The complex architecture of oxygenic photosynthesis. Nat. Rev. Mol. Cell Biol., 5, 971-982.

Nicholas, K.B. and Nicholas, H.B. (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author. http://www.psc.edu/biomed/genedoc/.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M. Remington, S.J., Silman, I., Schrag, J. *et al.* (1992) The α/β hydrolase fold. Protein Eng. 5, 197-211.

Ort, D.R. and Baker, N.R. A photoprotective role for O_2 as an alternative electron sink in photosynthesis? (2002) Curr. Opin. Plant Biol. 5, 193-198.

Otsuka, M., Kato, N., Shao, R-X., Hoshida, Y., Ijichi, H., Koike, Y., Taniguchi, H., Moriyama, M., Shiratori, Y., Kawabe, T. and Omata, M. (2004).Vitamin K_2 inhibits the growth and invasiveness of hepatocellular carcinoma cells via protein kinase A activation. Hepatology 40, 243–251.

Palaniappan, C., Sharma, V., Hudspeth, M.E.S. and Meganathan, R. (1992) Menaquinone (vitamin K2) biosynthesis: evidence that the *Escherichia coli menD* gene encodes both 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid synthase and a-ketoglutarate decarboxylase activities. J. Bacteriol. 174, 8111–8118.

Palmer, D.R., Garrett, J.B., Sharma, V., Meganathan, R., Babbitt, P.C. and Gerlt, J.A. (1999) unexpected divergence of enzyme function and sequence: "N-acylamino acid racemase" is *o*-succinylbenzoate synthase. Biochemistry 38, 4252-4258.

Pang, S.S., Duggleby, R.G. and Guddat LW (2002) Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors. J. Mol. Biol. 317, 249-262.

Parsons, J.F., Jensen, P.Y., Pachikara, A.S., Howard, A.J., Eisenstein, E. and Ladner, J.E. (2002) Structure of *Escherichia coli* aminodeoxychorismate synthase: architectural conservation and diversity in chorismate-utilizing enzymes. Biochemistry 41, 2198-2208.

Pesaresi, P., Varotto, C., Richly, E., Kurth, J., Salamini, F. and Leister, D. (2001) Functional genomics of *Arabidopsis* photosynthesis. Plant Physiol. Biochem. 39, 285-294.

Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R. and Osbourn, A. (2004) A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. Proc. Natl. Acad. Sci. USA 101, 8233-8238.

Raven, J.A. and Allen, J.F. (2003) Genomics and chloroplast evolution: what did cyanobacteria do for plants? Genome Biol. 4, 209.1-209.5.

Reiss, B., Klemm, M., Kosak, H. and Schell, J. (1996) RecA protein stimulates homologous recombination in plants. Proc. Natl. Acad. Sci. USA 93, 3094-3098.

Rowland, B.M. and Taber, H.W. (1996) Duplicate isochorismate synthase genes of *Bacillus subtilis*: regulation and involvement in the biosyntheses of menaquinone and 2,3-dihydroxybenzoate. J. Bacteriol. 178, 854-61.

Saenger, W., Jordan, P. and Krauss, N. (2002) The assembly of protein subunits and cofactors in photosystem I. Curr. Opin. Struct. Biol. 12, 244-254.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning – a laboratory manual. 2 ed. Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press.

Sakuragi, Y., Sakuragi, Y., Zybailov, B., Shen, G., Jones, A.D., Chitnis, P.R., van der Est, A., Bittl, R., Zech, S., Stehlik, D., Golbeck, J.H. and Bryant, D.A. (2002) Insertional inactivation of the *menG* gene, encoding 2-phytyl-1,4-naphthoquinone methyltransferase of *Synechocystis sp*. PCC 6803, results in the incorporation of 2-phytyl-1,4-naphthoquinone into the A(1) site and alteration of the equilibrium constant between A(1) and F(X) in photosystem I. Biochemistry 41, 394-405.

Saxena, S.P., Israels, E.D. and Israels, L.G. (2001) Novel vitamin K-dependent pathways regulating cell survival. Apoptosis 6, 57-68.

Scheller, H.V., Jensen, P.E., Haldrup, A., Lunde, C. and Knoetzel, J. (2001) Role of subunits in eukaryotic Photosystem I. Biochim. Biophys. Acta 1507, 41-60.

Schoeder H.U. and Lockau W. (1986). Phylloquinone copurified with the large subunit of photosystem I. FEBS Lett., 199, 23–27.

Seki, M. *et al.* (2003) Functional annotation of a full-length *Arabidopsis* cDNA collection. Science 302, 842-846.

Semenov, A.Y., Vassiliev, I.R., van der Est, A., Mamedov, M.D., Zybailov, B., Shen, G., Stehlik, D., Diner, B.A., Chitnis, P.R. and Golbeck, J.H. (2000) Recruitment of a foreign quinone into the A1 site of photosystem I. Altered kinetics of electron transfer in phylloquinone biosynthetic pathway mutants studied by time-resolved optical, EPR, and electrometric techniques. J. Biol. Chem. 275, 23429–23438.

Serrano, A., Cordoba, F., Gonzales-Reyes, J.A., Santos, C., Navas, P. and Villalba, J.M. (1994) Purification and Characterization of Two Distinct NAD(P)H Dehydrogenases from Onion (Allium cepa L.) Root Plasma Membrane. Plant Physiol. 106, 87-96.

Serrano, A., Cordoba, F., Gonzales-Reyes, J.A., Santos, C., Navas, P. and Villalba, J.M. (1995) NADH-specific dehydrogenase from onion root plasma membrane: purification and characterization. Protoplasma 184, 133-139.

Sharma, V., Suvarna, K., Meganathan, R. and Hudspeth M.E.S. (1992) Menaquinone (vitamin K_2) biosynthesis: nucleotide sequence and expression of the *men*B gene from *Escherichia coli*. J. Bacteriol. 174, 5057-5062.

Sharma, V., Meganathan, R. and Hudspeth, M.E.S. (1993) Menaquinone (vitamin K_2) biosynthesis: nucleotide sequence, and expression of the *men*C gene from *Escherichia coli*. J.Bacteriol. 175, 4917-21.

Sharma, V., Hudspeth, M.E.S. and Meganathan, R. (1993) Menaquinone (vitamin K_2) biosynthesis: localization and characterization of the *men*E gene from *Escherichia coli*. Gene 168, 43-48.

Schein, A.I., Kissinger, J.C. and Ungar, L.H. (2001) Chloroplast transit peptide prediction: a peek inside the black box. Nucleic Acids Res. 29, E82.

Shimada, H., Ohno, R., Shibata, M., Ikegami, I., Onai, K., Ohto, M.A. and Takamiya, K. (2005) Inactivation and deficiency of core proteins of photosystems I and II caused by genetical phylloquinone and plastoquinone deficiency but retained lamellar structure in a T-DNA mutant of *Arabidopsis*. Plant J. 41, 627-637.

Schultz, G., Ellerbrock, B.H. and Soll, J. (1981) Site of prenylation reaction in synthesis of phylloquinone (vitamin K_1) by spinach chloroplasts. Eur. J. Biochem. 117, 329-332.

Schütte, H.-R. (1993) in *Progress in Botany*, eds. Behnke, H.-D., Lüttge, U., Esser, K., Kadereit, J.K. & Runge, M. (Springer, Berlin, Heidelberg), pp. 218-236.

Simantiras, M. and Leistner, E. (1991) Cell free synthesis of o-succinylbenzoic acid in protein extracts from anthraquinone and phylloquinone (vitamin K_1) producing plant cell suspension cultures. Occurrence of intermediates between isochorismate and o-succinylbenzoic acid. 46, 364-370.

Small, I., Peeters, N., Legeai, F., Lurin, C. (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics, 4, 1581-1590.

Spraggon, G., Kim, C., Nguyen-Huu, X., Yee, M.-C., Yanofsky, C. and Mills, S.E. (2001) The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan. Proc. Natl. Acad. Sci. USA, 98, 6021–6026.

Soll, J. (2002) Protein import into chloroplasts. Curr. Opin. Plant Biol. 5, 529-535.

Storf, S., Jansson, S. and Schmid, V.H.R. (2005) Pigment binding, fluorescence properties, and oligomerization behavior of Lhca5, a novel light-harvesting protein. J. Biol. Chem. 280, 5163-5168.

Stupar, R.M., Lilly, J.W., Town, C.D., Cheng, Z., Kaul, S., Buell, C.R. and Jiang, J. (2001) Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proc. Natl. Acad. Sci. U. S. A. 98, 5099–5103.

Suvarna, K., Stevenson, D., Meganathan,, R. and Hudspeth M.E.S. (1998) Menaquinone (vitamin K_2) biosynthesis: localization and characterization of the *men*A gene from *Escherichia coli*. J. Bacteriol. 180, 2782–2787.

Tatusova, T.A. and Madden, T.L. (1999) BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. 174, 247-250.

Taber, H.W., Dellers, E.A. and Lombardo, L.R. (1981) Menaquinone biosynthesis in *Bacillus subtilis*: isolation of *men* mutants and evidence for clustering of *men* genes. J. Bacteriol. 145, 321-327.

Takahashi, Y., Hirota, K. and Katoh, S. (1985) Multiple forms of P700–chlorophyll a–protein complexes from Synechococcus sp.: The iron, quinone and carotenoid contents. Photosynth. Res. 6, 183–192.

Teichmann, S.A. and Veitia, R.A. (2004) Genes encoding subunits of stable complexes are clustered on the yeast chromosomes: and interpretation from a dosage balance perspective. Genetics 167, 2121.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y. and Martin, W. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat. Rev. Genet. 5, 123-135.

Thomas, P.S. (1980) Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77, 5201-5205.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 11, 4673-4680.

Thompson, T.B., Garrett, J.B., Taylor, E.A., Meganathan, R., Gerlt, J.A. and Rayment, I. (2000) Evolution of enzymatic activity in the enolase superfamily: structure of *o*-succinylbenzoate synthase from *Escherichia coli* in complex with Mg^{2+} and *o*-succinylbenzoate. Biochemistry 39, 10662-10676.

Torres, M.A. and Dangl, J.L. (2005) Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. Curr. Opin. Plant Biol. 8, 387-403.

Tsoka, S. and Ouzounis, C.A. (2000) Prediction of protein interactions: metabolic enzymes are frequently involved in gene fusion. Nat. Genet. 26, 141-142.

Valenti, V., Minardi, P., Guerrini, F., Mazzucchi, U. and Pupillo, P. (1989) Increase of plasma membrane NADH-duroquinone reductase in tobacco leaves treated with protein-lipopolysaccharide complexes. Plant Physiol. Biochem. 27, 569-576.

van Tegelen, L.J., Moreno, P.R., Croes, A.F., Verpoorte, R. and Wullems, G.J. (1999) Purification and cDNA cloning of isochorismate synthase from elicited cell cultures of *Catharanthus roseus*. Plant Physiol. 119, 705-712.

Vianello, A. and Macri, F. (1989) NAD(P)H oxidation elicits anion superoxide formation in radish plasmalemma vesicles. Biochim. Biophys. Acta 980, 202-208.

Wallin, R., Wajih, N., Greenwood, G.T. and Sane D.C. (2001) Arterial calcification: a review of mechanisms, animal models, and the prospects for therapy. Med. Res. Rev., 21, 274-301.

Wildermuth, M.C., Dewdney, J., Wu, G. and Ausubel, F.M. (2001) Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature 414, 562-565.

Williams, E.J.B. and Bowles, D.J. (2004) Coexpression of neighbouring genes in the genome of *Arabidopsis thaliana*. Genome Res. 14, 1060-1067.

Winkel, B.S.J. (2004) Metabolic channeling in plants. Annu. Rev. Plant Biol. 55, 84-107.

Wong, S. and Wolfe K.H. (2005) Birth of a metabolic gene cluster in yeast by adaptive gene relocation. Nat. Genet. 37, 777-782.

Xu, W., Chitnis, P., Valieva, A., van der Est, A., Pushkar, Y.N., Krzystyniak, M., Teutloff, C., Zech, S.G., Bittl, R., Stehlik, D., Zybailov, B, Shen, G. and Golbeck, J.H. (2003) Electron Transfer in Cyanobacterial Photosystem I. I physiological and spectroscopic characterization of site-directed mutants in a putative electron transfer pathway from A_0 through A_1 to F_X. J. Biol. Chem. 278, 27864–27875.

Yanai, I., Derti, A. and DeLisi, C. (2001) Genes linked by fusion events are generally of the same functional category: a systematic analysis of 30 microbial genomes. Proc. Natl. Acad. Sci. USA 98, 7940-7945.

Yanai, I., Wolf, Y.I. and Koonin, E.V. (2002) Evolution of gene fusions: horizontal transfer versus independent events. Genome Biol. 3, 1-13.

Yoshida, E., Nakamura, A. and Watanabe, T. (2003) Reversed-phase HPLC determination of chlorophyll *a* and naphthoquinones in photosystem I of red algae: existence of two menaquinone-4 molecules in photosystem I of *Cyanidium caldarium*. Anal Sci. 19, 1001-1005.

Zhang, S. and Scheller, H.V. (2004) Light-harvesting complex II binds to several small subunits of photosystem I. J. Biol. Chem. 279, 3180–3187.

Zybailov, B., van der Est, A., Zech, S., Teutloff, C., Johnson, T.W., Shen, G., Bittl, R., Stehlik, D., Chitnis, P.R. and Golbeck, J.H. (2000) Recruitment of a foreign quinone into the A1 site of photosystem I. II. Structural and

functional characterization of phylloquinone biosynthetic pathway mutants by EPR and electron-nuclear double resonance spectroscopy , J. Biol. Chem., 275, 8531-8539.

Acknowledgements

I would like to thank PD Dr. Jörg Meurer for providing me the opportunity to perform my Ph.D. under his supervision and for the exciting project that he gave me. I am also very grateful to him for the introduction into the *Arabidopsis* research and plant molecular genetics as well as for his permanent support and all the freedom that I had during my work. The continuous adventurous atmosphere in his lab, the creative discussions, his technical contributions and his enthusiastic view of science were crucial for the successful outcome of the *PHYLLO* project.

This work was performed at the Department Biologie I, Bereich Botanik, Ludwig-Maximilians-Universität München in the laboratory of Prof. Dr. R.G. Herrmann. I would like to thank him for the all his support. I am also grateful for the excellent technical assistance of Elli Gerick, Ingrid Duschaneck, Uli Wißnet and Claudia Nickel.

I would like to acknowledge PD Dr. Jon Falk, Botanisches Institut der Christian-Albrechts-Universität zu Kiel. Without his contribution to the PhQ measurements by HPLC this work would not have been possible.

I would like to express my thanks to all my colleagues of the lab Cristina, Pavan, Won, Katrin, Lina, Stephan, Rhea, Serena, Agata, Lada, Jarda, Uwe, Martha, Brigitte, Manfred, Rainer, Michi, Peter, Anna, Lena, Irina P., Giusy, Irina G., Jutta, Pat, Petra, Cordelia, for all the nice moments and the collaborative atmosphere in the lab. I am especially grateful to Won and Lina for the collaboration on the *PHYLLO* project.

Thanks to Christiane for sharing all the good and difficult moments during my Ph.D. and for her continuous patience, stimuli and support.

Thanks to my parents Alfredo and Delsi for providing me the basis of my education and for all the support during my study. I am also grateful to my aunt Gerda for providing me support during my stay in German.

I would like to express my acknowledgments to the CAPES institution, related to the Brazilian Ministry of Education, for providing me a scholarship for my Ph.D. study.

Curriculum vitae

Personal data

Family name:	Gross
First name:	Jeferson
Home address:	Schwanthalerstr 111/ App 459, PLZ-80339, Munich
Date of birth:	03.06.1970
Place of birth:	Porto Alegre, Brazil
Marital status:	single

Education

1º CYCLE: "Escola Estadual de 1º Grau Três de Outubro". Porto Alegre, Brazil (1977-1984).
2º CYCLE: "Escola Técnica Parobé" and "Escola de 2º Grau Mauá". Porto Alegre, Brazil (1985-1986).

Undergraduate studies

Bachelor Art, Music: Federal University of Rio Grande do Sul (UFRGS). Porto Alegre, Brazil (1988-1993).
Bachelor Science, Pharmaceutical Sciences: UFRGS. Porto Alegre, Brazil (1994- 1999).

Graduate studies

Master Science: Program in Cell and Molecular Biology- Biosciences Institute, Center of Biotechnology, UFRGS. Porto Alegre, Brazil (March 2000 - April 2002).
Ph.D.: Department Biologie I, Bereich Botanik; Ludwig-Maximilians-Universität München, Germany (since April 2003).

Research experience

Undergraduate Research Assistant (August 1997- December 1999): Adviser: Prof. Dr. Jeverson Frazzon. Biological Nitrogen Fixation Laboratory, Center for Biotechnology, UFRGS.
Master Science (March 2000-April 2002): Advisers: Dr. Arthur G. Fett-Neto and Dr. Janette P. Fett. Plant Physiology Laboratory, Center for Biotechnology, UFRGS.
Ph.D. (since April 2003): Advisers: PD Dr. Jörg Meurer and Prof. Reinhold G. Herrmann. Department Biologie I, Bereich Botanik; Ludwig-Maximilians-Universität München, Germany.

Publications

Gross J., Stein J.R., Fett-Neto A.G. and Fett J.P. (2003) Iron homeostasis related genes in rice. Genet. Mol. Biol. 26, 477-497.

Sperotto R.A., Gross J., Vedoy C., Passaglia L.M.P. and Schrank I.S. (2004) The electron transfer flavoprotein *fix*ABCX gene products from *Azospirillum brasilense* show a NifA-dependent promoter regulation. Curr. Microbiol., 49, 267-273.

Gross J, Cho W.K., Lezhneva L., Falk J., Krupinska K., Shinozaki K., Seki M., Herrmann R.G. and Meurer J. (2006) A plant locus for phylloquinone (vitamin K_1) biosynthesis originated from a fusion of four eubacterial genes, (submitted).

Eidesstattliche Erklärung:

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt habe.
Zudem erkläre ich, dass ich keine früheren Promotionsversuche unternommen habe.

München, Januar 2006

Jeferson Gross