This is to certify that the
dissertation entitled

Characterization of glucose-6-phosphate dehydrogenase isozymes and their roles in the plant and in seed oil accumulation in *Arabidopsis thaliana*

presented by

Setsuko Wakao

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Genetics

Major Professor's Signature

9/8/05

Date

MSU is an Affirmative Action/Equal Opportunity Institution

DATE DUE	DATE DUE	DATE DUE

CHARACTERIZATION OF GLUCOSE-6-PHOSPHATE DEHYDROGENASE ISOZYMES AND THEIR ROLES IN THE PLANT AND IN SEED OIL ACCUMULATION IN *ARABIDOPSIS THALIANA*

By

Setsuko Wakao

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Graduate Program in Genetics

2005

ABSTRACT

CHARACTERIZATION OF GLUCOSE-6-PHOSPHATE DEHYDROGENASE ISOZYMES AND THEIR ROLES IN THE PLANT AND IN SEED OIL ACCUMULATION OF *ARABIDOPSIS THALIANA*

By

Setsuko Wakao

In green tissues of plants under illumination, photosynthesis is the primary source of NADPH. In non-photosynthetic tissues or under non-photosynthetic conditions, the oxidative pentose phosphate pathway (OPPP) is one of the major sources of NADPH. The first and committed reaction is catalyzed by glucose-6-phosphate dehydrogenase (G6PDH). The six members of the G6PDH gene family in Arabidopsis, *G6PD1*-to-*G6PD6*, were characterized. Transit peptide analysis predicted two cytosolic and four plastidic isoforms. Five of the six genes encode active G6PDHs when expressed in *E. coli*. The recombinant isoforms showed differences in substrate requirements and sensitivities to feed-back inhibition. Plastidic isoforms were inactivated by reduction. One cytosolic isoform was insensitive to redox, while the other was inactivated by oxidation. The respective genes had distinct expression patterns that did not correlate with the activity of the proteins, implying a regulatory mechanism beyond the control of mRNA abundance. Two cytosolic and one presumably plastidic isoform were detected *in vivo* using zymograms, two of which coding genes were identified using T-DNA insertion lines.

The function of the cytosolic G6PDH isoforms were directly addressed using the respective T-DNA insertion mutants and a double mutant. A reciprocal induction in activity was observed for the two cytosolic isoforms in the single mutants, which did not involve increased levels of the respective mRNAs. In contrast, the activity of the isoform presumed to be plastidic was not affected, when one or both of the cytosolic isoforms were lost. The two cytosolic isoforms were found dispensable for growth under normal- and oxidative stress conditions tested. To address the contributions of G6PDH isoforms in seed oil accumulation, the gene expression patterns and the activity of G6PDH isoforms were examined in developing seeds. The transcript of all isoforms was detected at various levels, but the activity of only two, G6PD6 and the presumed plastidic isoform, was found. G6PDH activity was higher at later developmental stages. The oil content of seeds for the single and double mutants was compared to that of the wild-type (WT). Results suggested that the oil content and seed weight of the double mutant were higher than those of WT seeds.

To my parents

ACKNOWLEDGMENTS

First and foremost I would like to thank my advisor and mentor, Christoph Benning for the guidance during these years. I have learned so much from so little that I started with, and it has gone so fast because I truly enjoyed it. On the first day you told me you would train me to be ready to go out into the world. I hope I have reached that stage, thank you for carrying out that difficult task. It was sometimes challenging to convince and excite you with results, but through this I have learned to present and explain data and also to critically analyze them. I have always appreciated your encouragement, fairness, and understanding and I will keep remembering your words. The time in the Benning lab will have a special spot in my heart for the rest of my life.

The members of the Benning lab, I appreciate for all your friendship and scientific interactions. I have learned the importance of sharing ideas, knowledge, and expertise unselfishly. Not only stimulating but the fun atmosphere (especially all those lunch time conversations), has allowed me to work the long hours and in the weekends, and still be the happy person that I am. I wish I could list your names and what each you have given me, but I will just say that I am deeply grateful to all of you for being part of my life at MSU and beyond.

My committee, Douglas Gage, John Ohlrogge, and Steve Triezenberg, thank you for your commitment in guidance through my education. I felt confident to present elsewhere after having discussed my data with you, because it was always at the highest standards. I am also grateful for all the advice and discussions from the many times I just dropped by, it has been very valuable in taking my paths to reach where I am and will be.

I am grateful to the people who make the collaborative atmosphere at MSU, including Biochemistry Dept., the plant community, Genetics program, for teaching me the importance of a happy work life and scientific interaction. I was fortunate to be at the interface of many different "groups" and be exposed to different ideas, which I have greatly enjoyed. In particular, I would like to thank John Wilson for teaching me zymograms that became essential in this work, and also for discussions on characterization of the purified enzymes. Thank you to Dean Dellapenna for advising me in one of my first and primitive writing experience. It was not easy (for either of us), but your comments have helped my writing throughout my graduate studies. The many members of the Ohlrogge lab, who has shared ideas and helped me with experiments, thank you for always welcoming me in your lab.

I would especially like to express my gratitude to Miguel Ballicora, for spending hours teaching me enzyme kinetics, data analyses, among many other topics. Together I would like to mention the rest of Argentines of the Preiss lab, Alejandra Yep and Clarisa Bejar. Thanks for your friendship, good neighbor-ship and the hours we spent chatting about some scientific and more non-scientific matters. Life in Biochemistry would not have been half as fun without you.

Friends, thank you for the great time I had in MSU. I don't think I have enough words to express what good friends I have met, and how much I will miss this time. I do not have the space to list names, but I would like to especially thank Donatella Canella for we are comrades in the battle through the "grad school" (among other common battles we fight). Thanks for all those hours we shared the cool results and the screw-ups, and of course other matters. My latin/South American family and those that are not native to this

area but are included in this group because of various reasons such as mine, I apologize for grouping you together. Thank you for treating me as part of the family, despite that I am only slightly *chancho en misa*. One of the treasures from these years is the warmth of your presence that I have come to take for granted, and I will truly miss my East Lansing family when I leave.

My family, although very very far away, has been the solid support all this time. You are my oasis that I return to and your support is my motor. I am the little black sheep of the family, and to know that you are proud of me makes me very happy of what I do. Thanks to my father who had supported me all my life, has had the earliest influence on my interest in biology. My mother, who is perhaps my closest friend and now equivalent to both my mother and father, and my brothers that I always count on, and their families, thank you for all your love.

Finally, Francisco. Thank you for being my biggest fan, you have always managed to lift me from my lowest points. I will not say that this was not possible without you, but it would have been so difficult I would rather not try. You have indeed made me a better person, and I am truly grateful to have you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abbreviations

ACCase	acetyl-CoA carboxylase
ACP	acyl carrier protein
DAF	days after flowering
DGAT	diacylglycerol acyltransferase
DHAP	dihydroxyacetone phosphate
DTT	dithiothreitol
EST	expressed sequence tag
FA	fatty acid
FAS	fatty acid synthetase
FAT	fatty acid thioesterase
GAP	glyceraldehyde-3-phosphate
GAPDH	3-phosphoglyceraldehyde dehydrogenase
G6PDH	glucose-6-phosphate dehydrogenase
G6P	glucose-6-phosphate
GPAT	glycerol-3-phosphate acyltransferase
GPT	glucose-6-phosphate transporter
HXK	hexokinase
LPAT	lysophosphatidic acid acyltransferase
MDH	malate dehydrogenase
MeV	methylviologen
NADPH	nicotineadenine dinucleotide phosphate (reduced form)
NADH	nicotineadenine dinucleotide (reduced form)
NaTT	sodium tetrathionate
OAA	oxaloacetate
OPPP	oxidative pentose phosphate pathway
PDC	pyruvate dehydrogenase complex
PDAT	phosphocholine:diacylglycerol acyltransferase
PEP	phosphoenolpyruvate
PFK	phosphofructokinase
PET	photosynthetic electron transport
PGA	phosphoglycerate
6PGDH	6-phosphogluconate dehydrogenase
PK	pyruvate kinase
PPP	pentose phosphate pathway
Pyr	pyruvate
Ru5P	ribulose-5-phosphate
Suc	sucrose
TAG	triacylglycerol
TPT	triose-phosphate transporter
VLCFA	very long chain fatty acid
XPT	xylulose-5-phosphate transporter
X5P	xylulose-5-phosphate

Chapter 1

Introduction

1.1 Seed as the transfer of life

The successful development of a competent seed is essential for the plant's transfer of genetic information to the next generation. To accomplish this, a seed must fulfill several requirements: (1) establish the plant body pattern, (2) accumulate storage reserves for germination and (3) acquire dormancy to survive adverse environments (Goldberg *et al.*, 1994).

1.1.1 Seed development

Embryogenesis and seed development have been extensively characterized using Arabidopsis. The development of a seed starts with two fertilization events that lead to the formation of the diploid zygote and the triploid endosperm. The male gametophyte (pollen) contains two sperm cells, one of which fertilizes the haploid egg cell and the other fertilizes the diploid central cell contained in the female gametophyte (embryo sac) (Bewley *et al.*, 2000). During embryogenesis two major processes must be achieved: the development of the zygote into an embryo, and the accumulation of storage compounds as energy reserves for later germination. Storage compounds in the form of starch, lipid, and protein accumulate either in the embryo of plants such as Arabidopsis, Brassica and soy, or in the endosperm of wheat, rice and other cereal crops.

The seeds of Arabidopsis take approximately 20 days to mature under normal greenhouse conditions, from fertilization in which the zygote forms until complete maturation and desiccation of the seed (Koornneef and Karssen, 1994). The development of the seed can be largely separated into two phases, morphogenesis and seed maturation (Goldberg *et al.*, 1994). During morphogenesis, the embryo goes through the globular

and heart stages (for which the names refer to the embryo shape) at the end of which, the complete body pattern of the embryo is established. In the maturation phase, the embryo further develops into torpedo-, bent cotyledon-, and mature stages, during which storage compound deposition occurs. When all biochemical activity has ceased the seed enters dormancy (Goldberg *et al.*, 1994). The endosperm, the other main component of the seed, goes through endoreduplication forming a syncytium that contributes to the large increase in seed size from fertilization to the globular stage. Coincident with the heart stage, cellularization occurs in the endosperm syncytium (Berger, 2003). From this point on the endosperm exhibits little activity of cell division and in turn the embryo goes through rapid cell division followed by storage compound accumulation. The embryo is thought to take up carbon sources from maternal tissues, indicated by the presence of sucrose and peptide transporters involved in seed development (Hirose *et al.*, 1997; Song *et al.*, 1997; Stacey *et al.*, 2002). The role of the endosperm still remains unclear but it seems to have functions other than nutrient supply as its removal has been shown to affect seed germination (Penfield *et al.*, 2004). Finally the embryo fills up most of the mature seed and only the outer cell layer remains of the endosperm.

1.1.2 Storage compound accumulation

A mature Arabidopsis seed is filled with oil bodies containing triacylglycerol (TAG) and protein bodies with storage proteins. The oil bodies can reach up to 60% of the total volume (Mansfield and Briarty, 1992) or fatty acids (FAs) up to 30~40 % of dry-weight (Focks and Benning, 1998; O'Neill *et al.*, 2003). Soon after cell division ceases in the Arabidopsis embryo, large changes in the biochemical activity of the seed are observed.

Although the seed must synthesize large amounts of storage compounds, starch, protein and lipids, these processes occur at distinct phases (Focks and Benning, 1998; White and Benning, 2001; Baud *et al.*, 2002). Between 5-6 days after flowering (DAF) starch is the primary compound deposited and as it is degraded, accumulation of FAs and protein begins (9 DAF). FA accumulation occurs at highest rate 11-14 DAF and ceases at 17 DAF while protein accumulates linearly during this time (Focks and Benning, 1998) (Figure 1.1). This general pattern has been observed also in Brassica (Eastmond and Rawsthorne, 2000). The structural and regulatory genes involved in this complex process have been of great interest to many scientists.

1.2 Oil seed metabolism

Arabidopsis is just one example of an oil seed plant that accumulates lipids as the primary energy reserve. Plant lipids have a world-wide market value of 40 billion US (Gunstone, 2000). The lipids extracted from seeds of plants such as soybean, palm, and Brassica supply substrates for various products from vegetable oil to industrial products such as detergents (Gunstone, 2000). Much research has been carried out with the goal of elucidating the roles and contributions of the metabolic pathways involved in the accumulation of seed oil. Identification of the enzymes involved in FA and lipid biosynthesis and cloning of the respective genes has greatly increased our understanding of FA and lipid metabolism in plants. Based on this knowledge, many attempts have been made to engineer plants into producing commercial products, some resulting in success and others not. But from each case study we have been able to obtain new understanding of the complex regulatory mechanisms involved in lipid metabolism.

In the following sections, an overview of FA and TAG synthesis and the biochemical pathways involved in the carbon and cofactor supply are discussed. Emphasis is placed on the status of reducing equivalents and energy levels in green- and non-green oil seeds. More details of the biochemical, molecular, and genetic aspects of FA and lipid synthesis are provided in several thorough reviews (Ohlrogge *et al.*, 1991; Browse and Somerville, 1991; Slabas and Fawcett, 1992; Ohlrogge and Browse, 1995; Ohlrogge and Jaworski, 1997).

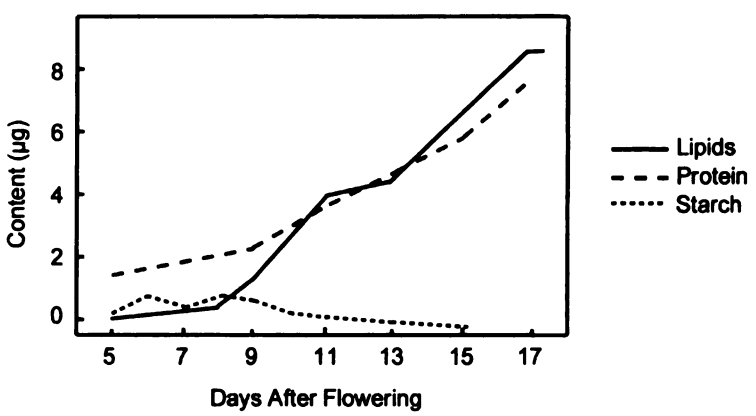


Figure 1.1 Storage compound accumulation in Arabidopsis seeds (Adapted from Figures from Baud *et al.*, 2002.)

1.2.1 Fatty acid synthesis

In the plant cell, all FAs of 16 carbons (C_{16}) and 18 carbons (C_{18}) are synthesized in the plastid (Ohlrogge *et al.*, 1979). FA synthesis is completely light-dependent in leaves (Roughan *et al.*, 1979) and to a lesser extent, in green seeds (Browse *et al.*, 1981; Browse and Slack, 1985; Bao *et al.*, 1998; Eastmond and Rawsthorne, 1998; Goffman *et al.*, 2005). FA synthesis involves the concerted action of many enzymatic reactions to repeatedly add two carbons to the elongating acyl-chain. A summary of FA synthesis is shown in Figure 1.2.

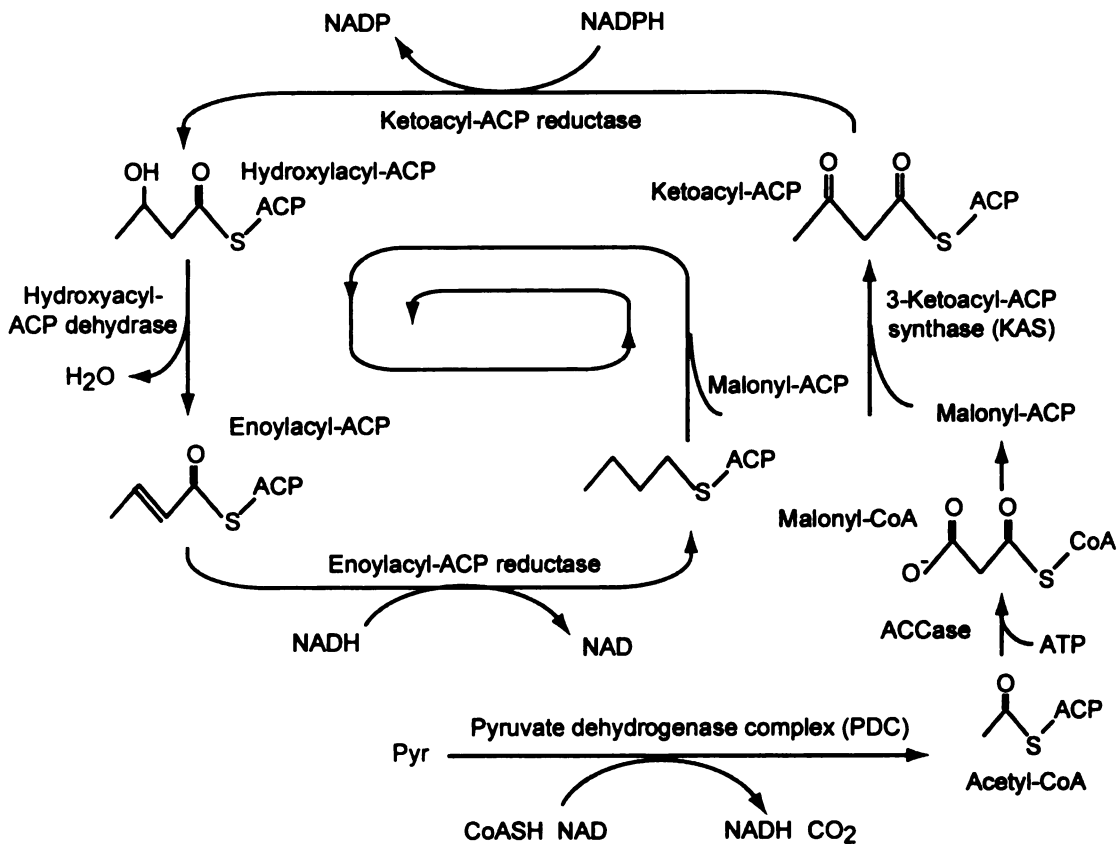


Figure 1.2 Fatty acid synthesis

1.2.1.1 Acetyl-CoA carboxylase (ACCase)

Malonyl-coenzyme A (CoA) is the carrier of the two-carbon units for FA synthesis. The generation of malonyl-CoA from acetyl-CoA is carried out by acetyl-CoA carboxylase (ACCase) and is considered the committed step for FA synthesis. ACCase catalyzes two reactions, carboxylation of the biotin in the enzyme which requires ATP, and the transfer of the carboxyl group to the substrate acetyl-CoA. Two forms have been observed for the enzyme. The multisubunit (heteromeric) ACCase, consists of four subunits: biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP), α-carboxyl-transferase, β-carboxyl-transferase (CT). The multifunctional (homomeric) ACCase, possesses BC, BCCP, CT domains in a single polypeptide. Most plants including Arabidopsis contain

both types, with heteromeric ACCase in the plastid and homomeric ACCase in the cytosol (Konishi *et al.*, 1996). Exceptions are the grass species containing homomeric ACCase in both subcellular localizations which explains the mode of function of some grass herbicides (Konishi and Sasaki, 1994).

The plastidic ACCase plays a major role in the regulation of FA synthesis. Its activity is light-dependent through redox regulation involving thioredoxin and is affected by pH and Mg^{2+} (Sasaki *et al.*, 1997; Hunter and Ohlrogge, 1998). The expression of the genes encoding for the plastidic ACCase subunits correlates with plant tissues showing active FA synthesis, either rapidly generating membrane lipids or accumulating oil (Ke *et al.*, 2000b). Alteration of ACCase activity, either by overexpression or repression of gene expression, has correlated with increased or decreased FA synthesis in several studies (Shintani *et al.*, 1997; Madoka *et al.*, 2002). Elevated ACCase activity could partially redirect carbon to FA in starch storing potato tubers (Klaus *et al.*, 2004).

The cytosolic (homomeric) ACCase is involved in generating malonyl-CoA required for the elongation of FAs beyond C_{18} into very long chain FA (VLCFA) that occurs in the cytosol. The seeds of the Arabidopsis knock-out (KO) plant of one of the two genes coding for cytosolic ACCase was found to accumulate less TAG than wild-type (WT) and to completely lack VLCFAs (Baud *et al.*, 2003).

1.2.1.2 Fatty acid synthetase (FAS)

Large efforts put into cloning and characterizing fatty acid synthetase (FAS) enzymes have revealed many biochemical details of FA synthesis in plants (Figure 1.2) (reviewed in Ohlrogge *et al.*, 1991; Browse and Somerville, 1991; Slabas and Fawcett, 1992;

Ohlrogge and Browse, 1995). Plants have a Type II FAS, which is a multiprotein complex similar to those of bacteria (Type I in animals and fungi is a multifunctional protein complex). The plant FAS resides in the plastid stroma and carries out all the reactions after malonyl-CoA formation by ACCase until the release of acyl-chains from the acyl-carrier protein (ACP).

The cycle of elongation by two-carbon units involves several reactions. First, the malonyl-CoA:ACP transacylase forms malonyl-ACP from malonyl-CoA. From thereon, the acyl-chain is not released from ACP until FA synthesis is complete. The addition of the two carbons onto the acyl-chain is catalyzed by 3-ketoacyl-ACP synthase (KAS), commonly referred to as the condensing enzyme. The initial acyl-chain is formed from malonyl-ACP and acetyl-CoA by KASIII resulting in a 3-ketobutyryl-ACP. This is then reduced by ketoacyl-ACP reductase, forming 3-hydroxybutyryl-ACP, and dehydrated by hydroxyacyl-ACP dehydrase to butenoyl-ACP. Finally butenoyl-ACP is reduced by enoyl-ACP reductase into a fully saturated butyryl-ACP, which serves as the substrate for the second round of the cycle. Three isoforms of KAS are found in many plants, each with unique substrate specificity. KASIII is specific to acetyl-CoA, KASI utilizes acyl-chains between C_4 and C_{16}, and KASII carries out the final addition that yields stearic acid (18:0). The two reductases in FAS have been described in various plants to specifically utilize NADH or NADPH. 3-Ketoacyl-ACP reductase utilizes NADPH and enoyl-ACP reductase utilizes NADH (Caughey and Kekwick, 1982; Shimakata and Stumpf, 1982) (Figure 1.2).

1.2.1.3 FA desaturation

A large portion of FAs occurring in the plant is unsaturated, most of which is a product of oleic acid (18:1). Two types of desaturases are known. One is the soluble Δ-9 acyl-ACP desaturase which most commonly generates oleoyl-ACP (18:1) from stearoyl-ACP (18:0). The family of soluble Δ-9 acyl-ACP desaturases includes those specific to different length acyl-chains that are responsible for some unusual desaturated FAs (Jaworski and Cahoon, 2003). Another type of desaturases is represented by the membrane-bound desaturases in the plastid and the endoplasmic reticulum (ER) that desaturates acyl chains attached to glycerolipids. Such enzymes have been difficult to purify because of their membrane-bound nature. The isolation of Arabidopsis mutants such as *fad2–* to-*fad8* disrupted in the activity of desaturases bound to the ER (FAD2, FAD3) or plastidic (FAD4 through FAD8) membranes has greatly contributed to our understanding (Ohlrogge *et al.*, 1991; Browse and Somerville, 1991). The ER desaturases prefer acyl-chains on phosphatidylcholine (PC) while the plastidic desaturases have various specificities for plastidic lipids.

1.2.1.4 FA thioesterases

FA synthesis is terminated when the acyl-chain is hydrolyzed from the ACP by an acyl-ACP thioesterase (FAT). Two types of FATs (FATA and FATB) have been observed widely in plants with different specificities. FATA hydrolyzes $18:1^{\Delta 9}$ acyl-ACP while FATB targets saturated and shorter chain acyl-ACPs, such as lauroyl-ACP thioesterase from California bay (*Umbellularia californica*). Because the released acyl-chains can follow various fates, the activity and specificity of FATs influence lipid metabolism

quantitatively and qualitatively. It has been shown to be a good target in engineering seed oil composition in several studies (Voelker *et al.*, 1992; Voelker *et al.*, 1996). In Arabidopsis, the FATB KO mutant has a reduction in palmitate, stearate, and in epicuticular wax, which requires saturated FAs for its biosynthesis (Bonaventure *et al.*, 2004).

1.2.1.5 FA elongation

Although the most common FAs have chain lengths of C_{16} and C_{18}, plants also contain longer chain FAs in the form of waxes, sphingolipids and TAGs. Seed storage TAGs in Arabidopsis contain C_{20}, C_{22}, and C_{24} chain FAs. FA elongation occurs in the cytosol, where an ER membrane-bound elongation complex resides. FA elongation follows a similar series of reactions as in FA synthesis except that both malonate and the acyl-chains are thioesterified to CoA and not ACP. The elongation complex consists of 3-ketoacyl-CoA synthase, 3-ketoacyl-CoA reductase, 3-hydroxyacyl-CoA dehydrase and enoyl-CoA reductase. FA elongation in plants requires only NADPH as a reducing equivalent (Barrett and Harwood, 1998; Lardans and Tremolieres, 1992). The cytosolic ACCase at least in part regulates FA elongation as has been shown in Arabidopsis plants with reduced cytosolic ACCase activity, which had reduced very-long-chain FAs (Baud *et al.*, 2003). In contrast to FA synthesis, FA elongation is not light-dependent (Bao *et al.*, 1998) suggesting a different regulatory mechanism of cytosolic homomeric ACCase.

1.2.2 Synthesis of triacylglycerol

TAG in mature seeds is found in oil bodies, which are structures enveloped by a single-layer phospholipid membrane embedded with oleosin, which presumably prevents the oil bodies from fusing to one another and protects them during desiccation. It is generally accepted that oil bodies arise from ER because oleosins are co-translationally inserted into microsomal membranes and enzymes that synthesize TAGs are present in the ER membrane (reviewed by Galili *et al.*, 1998).

Acyl-CoA that entered the ER is incorporated into lipids by the eukaryotic pathway, opposed to the synthesis of glycerolipids in the plastid by the prokaryotic pathway (reviewed by Browse and Somerville, 1991; Ohlrogge and Browse, 1995). In seeds, TAG can be synthesized in the ER through the Kennedy pathway and the relatively recently discovered phosphocholine:diacylglycerol acyltransferase (PDAT) activity (Dahlqvist *et al.*, 2000; Voelker and Kinney, 2001) (summarized in Figure 1.3). This is different from the TAG observed in leaf chloroplast plastoglobuli (Martin and Wilson, 1984), which is likely to be synthesized by a plastidic diacylglycerol acyltransferase (DGAT) (Kaup *et al.*, 2002). In the Kennedy pathway, acyl-chains are esterified to the *sn*-1, *sn*-2, *sn*-3 positions of glycerol-3-phosphate by glycerol-3-phosphate acyltransferase (GPAT), lysophosphatidic acid acyltransferase (LPAT) and DGAT, respectively. The third addition of the acyl-chain by DGAT occurs after the formation of diacylglycerol (DAG) from phosphatidic acid (PA) by phosphatidic acid phosphatase. Alternatively, PC can be formed from the transfer of phosphocholine to the *sn*-3 position of DAG by choline phosphotransferase (CPT) (also called CDP-choline:diacylglycerol cholinephosphotransferase). PDAT can convert PC and DAG into lyso-PC and TAG. The

presence of PDAT indicates that membrane lipids can be directly converted to TAG. Both PDAT and DGAT activity are present in microsomes which raises the question of which enzyme is responsible for TAG biosynthesis. DGAT is thought to be the major activity responsible for the formation of seed TAG because an Arabidopsis mutant, *tag1*, with reduced seed oil is defective in a gene coding for a DGAT (Katavic *et al.*, 1995; Zou *et al.*, 1997; Routaboul *et al.*, 1999).

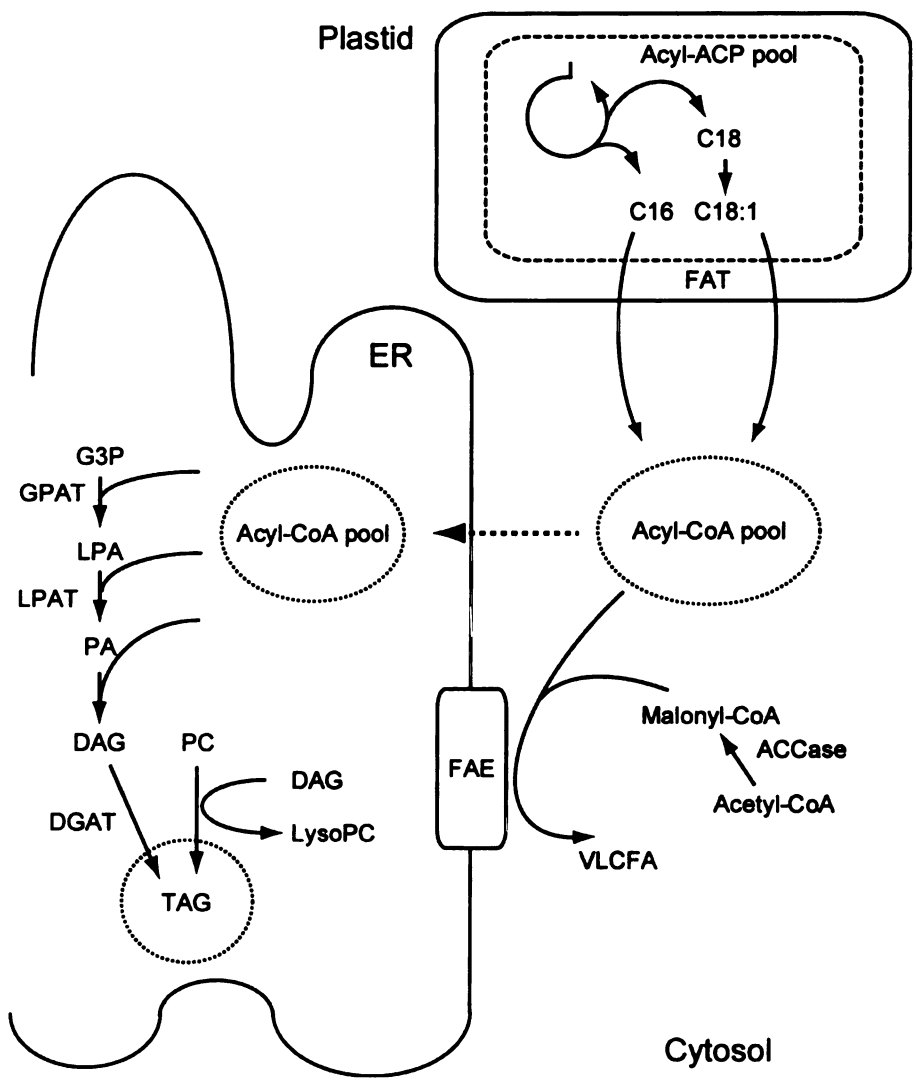


Figure 1.3 Compartmentalization of FA and TAG synthesis

In Arabidopsis, a seed- and flower-specific GPAT has been reported to be targeted to the ER- and mitochondrial membranes. When the gene (*GPAT1*) was disrupted by T-DNA insertion, the plant had severe defect in pollen development and minor changes in seed FA but not in total seed oil content (Zheng *et al.*, 2003). The role of PDAT in seed oil metabolism still cannot be ignored. In Arabidopsis, six genes code for proteins similar to human and yeast PDATs (Stahl *et al.*, 2004). Of the two genes whose encoded proteins had highest similarity, one was expressed ubiquitously while the other was seed-specific. It has been observed in transgenic plants that most novel FAs are predominantly found in the seed TAG and not in membrane lipids (Eccleston and Ohlrogge, 1998; Hooks *et al.*, 1999; Thomaeus *et al.*, 2001). If some FAs are actively removed from the membrane lipids, PDAT could possibly play a role in such process. The mechanism that discriminates certain FAs to be targeted to membrane lipids or TAGs is still under debate. Spatial compartmentalization in subdomains of the ER (Vogel and Browse, 1996; Lacey and Hills, 1996) or substrate specificity of the enzymes involved in membrane lipid and TAG synthesis have been suggested to be involved (Wiberg *et al.*, 1994; Millar *et al.*, 2000).

1.2.3 Supply of substrates and cofactors required for oil accumulation

Arabidopsis seeds are initially transparent and hence heterotrophic and begin to turn green from the globular stage on (Meinke, 1994). Starch is deposited while the seed is mostly non-photosynthetic, and the metabolic transition to accumulate FA and protein occurs as the seed turns green. FA synthesis not only requires large amounts of carbon but also ATP and NAD(P)H. ATP is required at the carboxylation step of acetyl-CoA by

ACCase. The two reductases specifically utilize NADPH (3-ketoacyl-ACP reductase) and NADH (enoyl-ACP reductase) (Caughey and Kekwick, 1982; Shimakata and Stumpf, 1982; Slabas *et al.*, 1986), and therefore eight moles of NADPH and NADH each are required for one mole of stearic acid to be synthesized. The desaturation of stearoyl-ACP by Δ-9 stearoyl-ACP desaturase requires reduced ferredoxin, and ferredoxin accepts electrons from NADPH or photosystem I (Jaworski, 1987). Outside the plastid, synthesis of VLCFAs in the cytosol also demands for NADPH, as the two reductases both utilize NADPH (Lardans and Tremolieres, 1992; Barrett and Harwood, 1998). In the following sections biochemical pathways that may contribute in the supply of substrates and cofactors are discussed. Figure 1.4 gives the overview of all the pathways discussed in their compartments.

1.2.3.1 Photosynthesis

In green tissues, photosynthetic electron transport (PET) generates the bulk of NADPH and ATP, which partly explains the light-dependence of FA synthesis. However most seeds are contained in fruits or siliques, a fact which raises the question of the amount of light the seeds receive. Moreover, many oil accumulating seeds such as sunflower, sesame and castor beans are not green.

It has been reported that the light transmission inside a silique of Brassica is 20-30% of ambient light (Eastmond *et al.*, 1996; King *et al.*, 1998). Plastids in developing Brassica embryos show structural resemblance to those from heterotrophically grown tissues (Asokanthan *et al.*, 1997), and contain 60% photosynthetic capacity of that in a leaf (Eastmond *et al.*, 1996).

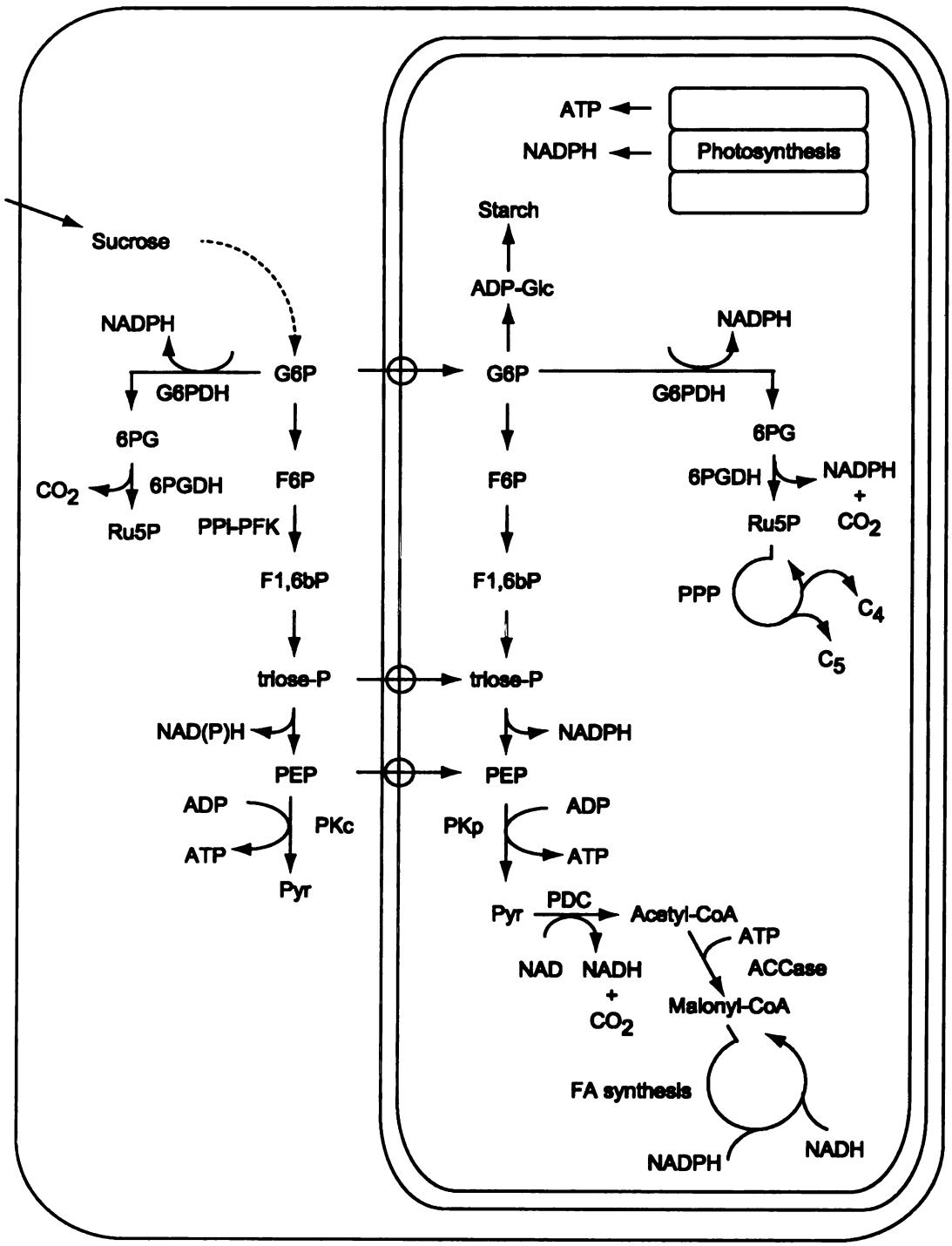


Figure 1.4 Oil seed metabolism

Brassica seeds have high carotenoid content similar to that of shade adapted leaves, indicating altered light harvesting to maximize absorption of light energy (Ruuska *et al.*, 2004). Stimulation of FA synthesis by light is less in embryo plastids compared to that in leaves which is completely dependent on light (Browse and Slack, 1985; Bao *et al.*, 1998; Eastmond and Rawsthorne, 1998). These results suggest that plastids of developing seeds have less photosynthetic activity but other means to supply energy and reductants for FA synthesis.

Despite the structural and metabolic differences between seed plastids and chloroplasts, the presence of active PET in Brassica embryos has been reported by several groups. Isolated seed plastids are able to evolve oxygen in light as the product of PET (Browse and Slack, 1985; Eastmond *et al.*, 1996; Bao *et al.*, 1998; King *et al.*, 1998). Atmospheric carbon incorporation in embryos is low suggesting that seed plastids have PET that could potentially generate NADPH to be utilized in biosynthetic reactions but not in such excess for fixation of exogenous carbon to occur (Eastmond *et al.*, 1996; King *et al.*, 1998).

The contribution of photosynthesis in green seeds to oil accumulation was directly shown by covering half of the silique of Brassica with aluminum foil for ten hours (Ruuska *et al.*, 2004). When the siliques were fed with ^{3}H-labeled water, the seeds in the light had incorporated, on average, 2.5-fold higher label into lipids than those in the dark. This simple experiment displays the contribution by photosynthetic activity in seed oil.

1.2.3.2 Energy status in seeds

The seed needs energy during storage compound accumulation. Starch synthesis requires ATP at the rate-limiting step of ADP-glucose pyrophosphorylase, FA synthesis requires ATP for carboxylation of acetyl-CoA, the synthesis of many amino acids requires ATP (Regierer *et al.*, 2002) and peptide bond formation in protein synthesis consumes ATP equivalents.

The levels of oxygen and ATP have been studied in developing legumes, pea (*Pisum sativum*) and fava beans (*Vicia faba*). Oxygen and ATP levels inside the seed were measured with a microsensor and luminescence, respectively (Borisjuk *et al.*, 2003; Rolletschek *et al.*, 2003). The chlorophyll content, oxygen and ATP levels showed spatial and temporal correlation in the developing seed. Low levels of ATP and oxygen (5-20% of atmospheric oxygen) were detected during early embryo stages, both of which increased as the seed turned green. The authors propose that low ATP levels caused by low respiration rates are limiting factors for metabolic flux during early legume seed development and that photosynthesis may contribute by supplying oxygen. A similar question was addressed in Brassica by measuring oxygen levels in the siliques and seeds (Vigeolas *et al.*, 2003). When atmospheric oxygen was increased beyond ambient concentration, the incorporation of ^{14}C from sucrose (Suc) into protein and lipid was increased, but not to starch. These effects were associated with increased ATP/ADP and UTP/UDP levels. The results of the two studies together indicate that low energy levels due to low oxygen concentration could be limiting in early development and oil accumulation in the seed, although the complex contributions of photosynthesis and respiration in supplying ATP and oxygen are still unclear. Novel biotechnological

approaches to increase oxygen levels in developing seeds were proposed such as engineering oxygen binding proteins (Vigeolas *et al.*, 2003). However, such manipulations should be carefully considered since hypoxia in certain tissues is important for proper development (Porterfield *et al.*, 1999). Recently in contrast to these findings, increase in oxygen levels has been shown to have less effect in increasing oil content than higher light intensities in cultivated Brassica embryos (Goffman *et al.*, 2005).

Arabidopsis plants with altered adenylate pools have been found to be affected in biosynthetic activity. In Arabidopsis, nine genes were found with high similarity to potato genes encoding plastidic adenylate kinases (ADK) (Carrari *et al.*, 2005). The T-DNA KO line of one of the plastidic ADKs showed increases in growth, adenylate pools (not necessarily ATP/ADP ratio), and amino acid content mostly under continuous light conditions. Because of the complex and conditional phenotypes observed in this KO line, the function of a single ADK cannot be specified. However, this study demonstrates that the loss of a single ADK leads to enhanced growth and amino acid biosynthesis most likely caused by increased adenylate levels. Manipulation of adenylate pool through ADK has proven to be effective in potato tubers after failed attempts in engineering enzymes in carbon metabolism, including ectopic expression of an invertase gene and introduction of a bacterial glucokinase (Sonnewald *et al.*, 1997; Trethewey *et al.*, 1998; Trethewey *et al.*, 1999). Starch synthesis and tuber yield was finally enhanced by repression of adenylate kinase (ADK) activity through an increase in the overall adenylate pool and ADP-glucose (Regierer *et al.*, 2002).

The ATP/ADP transporter imports ATP in exchange for ADP to meet the high demands of biosynthetic processes in the plastid, and has been proven to be a good target

to manipulate adenylate levels. Non-photosynthetic plastids are more dependent on the import of ATP, a prominent example being roots and tubers where starch accumulates in the absence of photosynthesis (Neuhaus and Emes, 2000). Consistently, transgenic potato plants with increased ATP/ADP transporter activity resulted in higher starch content (Tjaden *et al.*, 1998). The function of the ATP/ADP transporters has also been investigated using Arabidopsis KO plants (Reiser *et al.*, 2004). Two genes encode ATP/ADP transporters in Arabidopsis (*AtNTT1* and *AtNTT2*). The loss of NTT1 affects the plant slightly, whereas loss of NTT2 results in reduced growth rates only under short-day conditions indicating the importance of ATP import at night. The mature seeds of the Arabidopsis plants lacking both transporters were found to have reduction in both lipid (40%) and proteins (20%) (Reiser *et al.*, 2004).

1.2.3.3 Oxidative pentose phosphate pathway (OPPP)

The oxidative pentose phosphate pathway (OPPP) is the major source of NADPH in non-photosynthetic tissues as well as in photosynthetic tissues under darkness (Neuhaus and Emes, 2000). The two dehydrogenases in the OPPP, glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49) and 6-phosphogluconate dehydrogenase (6PGDH, EC 1.1.1.44), use $NADP^+$ as a cofactor and generate two moles of NADPH during the conversion of one mole glucose-6-phosphate (G6P) to ribulose-5-phosphate (Ru5P) (Figure 1.5). The reaction catalyzed by G6PDH is considered the regulatory step in the OPPP because of its tight regulation by the redox balance, i.e., $NADPH/NADP^+$ (Johnson, 1972; Lendzian, 1980; Williams, 1980).

OPPP resides in the cytosol and in the plastid (for summary, Figure 1.4). In non-photosynthetic tissues or under non-photosynthetic conditions, OPPP is the major sources of NADPH, used for various reductive reactions such as nitrate assimilation and FA synthesis (Neuhaus and Emes, 2000). Plastidic G6PDHs are subject two modes of regulation by photosynthesis, feed-back inhibition by NADPH and reductive inactivation through the thioredoxin-ferredoxin system (see later section 1.4 on G6PDH and OPPP).

The importance of OPPP in non-photosynthetic plastids can be observed by the prevalence of G6P transporters in these plastids, as G6P is a substrate for starch synthesis and OPPP (Emes and Neuhaus, 1997). The plastidic G6P transporters (GPT) from heterotrophic tissues have broad substrate specificity for G6P, triose phosphate, and PGA (Kammerer *et al.*, 1998). The gene coding for GPT is expressed mostly in heterotrophic tissues where it presumably imports G6P for starch synthesis and OPPP. The Arabidopsis genome contains two genes coding for GPT (*GPT1* and *GPT2*). Gene disruption by T-DNA insertion in *GPT2* does not affect the plant, while disruption of *GPT1* is lethal at the stages of pollen and embryo sac development. It is speculated that in heterotrophic pollen, where large amounts of FAs are synthesized, import of G6P into the plastid is strictly required for the generation of NADPH by OPPP (Niewiadomski *et al.*, 2005). In Arabidopsis seeds, the transcript for GPT is more abundant in early stages when starch accumulation occurs (Ruuska *et al.*, 2002).

In animal cells where no photosynthesis exists as source of NADPH, the relationship between FA synthesis and OPPP has long been suggested. In Drosophila, genotypic variation exists resulting in different activities of OPPP. Flies with higher OPPP activity had higher rates of FA synthesis resulting in a fatter fly (Cavener and Clegg, 1981). In rat

adipose tissue, the increase in the activity of OPPP coincides with the increased rate of FA synthesis (Ayala *et al.*, 1991). In mouse liver, the splicing of G6PDH mRNA to the mature form is inhibited by polyunsaturated FAs (Tao *et al.*, 2002), indicating a strong connection between OPPP and FA synthesis presumably through the supply of NADPH.

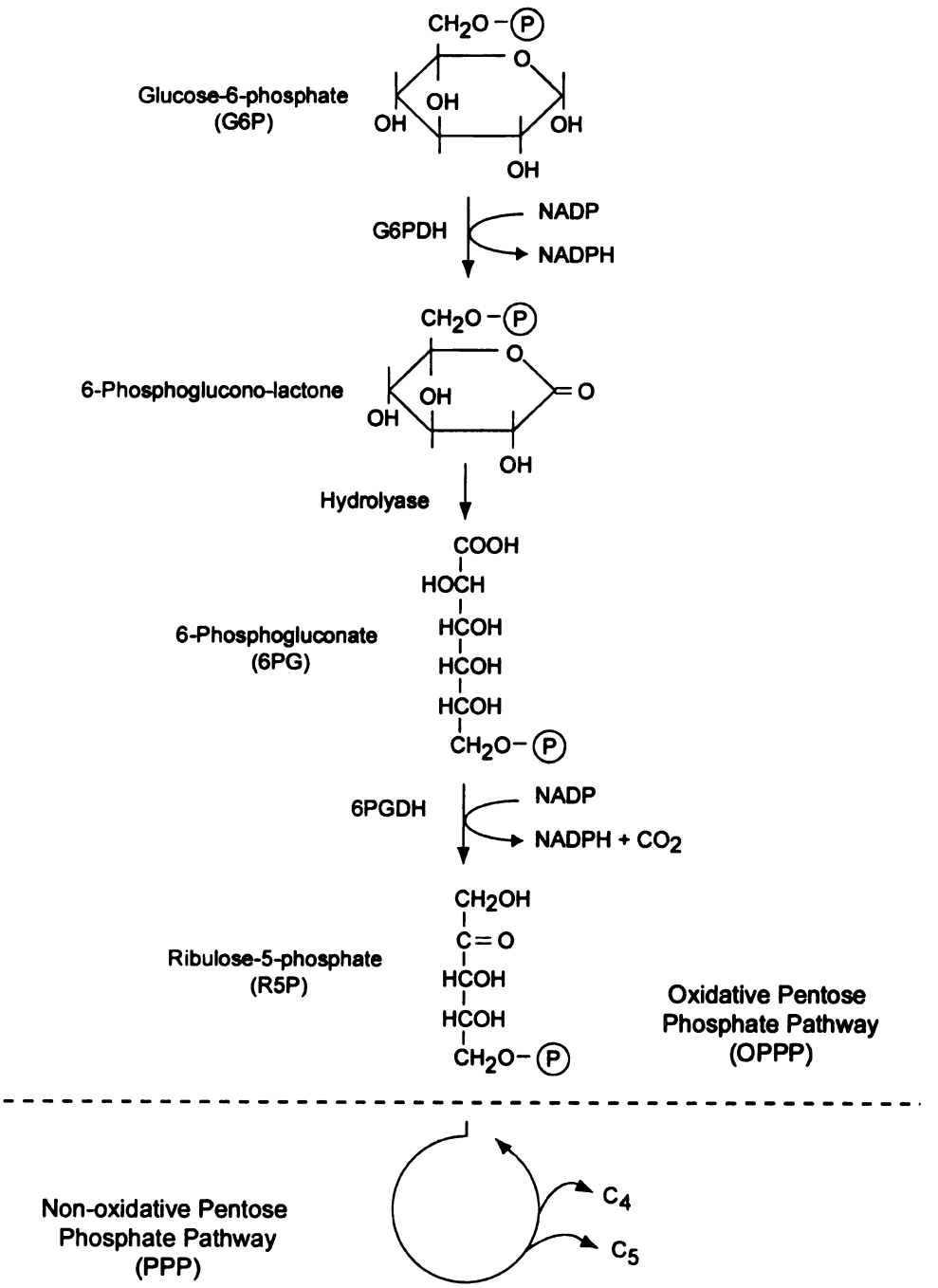


Figure 1.5 Oxidative pentose phosphate pathway

1.2.3.4 Glycolysis

Early seeds are heterotrophic and rely on the import of carbon from the maternal tissues in the form of Suc (Hirose *et al.*, 1997). The primary role of glycolysis is to break down sugar and generate pyruvate (Pyr), and to produce reducing equivalents and ATP for anabolic reactions (reviewed by Plaxton, 1996). Plant glycolysis is distinct from that of animals in several aspects, some of which with relevance to seed oil metabolism are discussed here. First, it occurs in plastids as well as in the cytosol. In heterotrophic tissues, glycolytic intermediates are transported across the plastid envelope membranes to supply precursors for biosynthetic processes such as FA, starch and amino acid synthesis. Second, cytosolic glycolysis possesses a pyrophosphate-dependent phosphofructokinase (PPi-dependent PFK), which conserves ATP in contrast with the reaction carried out by ATP-dependent PFK. The activity of PPi-dependent PFK has been found to be reduced in *wri1*, a mutant with 80% reduction in seed oil content associated with the down-regulation of several glycolytic enzyme activities. Third, Pyr can be formed in many ways other than by pyruvate kinase (PK), such as malic enzyme that produces pyruvate and NADPH from malate found in castor bean. This is an example of the metabolic flexibility to respond to specific demands of certain tissues. Nonetheless, PK, that irreversibly converts phosphoenolpyruvate (PEP) and ADP into Pyr and ATP, is thought to be one of the primary regulatory sites of plant glycolysis, because PEP is an inhibitor of several ATP-dependent PFKs and PK activity is affected by many downstream cell-specific metabolites (Plaxton, 1996, and references therein).

The reduction in seed oil of the *wri1* mutant is accompanied by reduced activity of several glycolytic enzymes including hexokinase (HXK), PPi-dependent PFK, and

pyruvate kinase (PK) (Focks and Benning, 1998). The effect on these adenylate metabolism-related enzymes in *wri1* may suggest that the involvement of glycolysis in seed oil accumulation is to supply not only precursors but also energy. Plastidic PK (PKp) activity has been indicated to have an important role during seed oil accumulation in an analysis of the temporal gene expression profiles using microarrays (Ruuska *et al.*, 2002) (for positions of PKc and PKp in the metabolic map, see Figure 1.5). The cytosolic PK (PKc) gene was expressed high during early seed development and declined while the PKp gene showed an opposite pattern with a peak at later stages, similar to that of many genes related to FA synthesis. The transition of the PK genes was associated with the peak in the gene expression of the plastidic PEP transporter. Presumably PKp provides Pyr and ATP for FA synthesis from PEP, a process which is reassured by the presence of a PEP transporter. An Arabidopsis mutant defective in one of the two PEP transporters (PPT1 and PPT2), *ppt1* (*cue1*) has a reduced content of aromatic amino acids and phenolic compounds (of which PEP is a precursor), and defective leaf development (Streatfield *et al.*, 1999; Knappe *et al.*, 2003). The effect on seed oil accumulation in *cue1* has not been described. Considering the severe defects of the *cue1* plant, it will be interesting to see the effect on seed oil in a plant with seed-specific reduction of PPT activity.

1.2.3.5 Pyruvate dehydrogenase complex (PDC)

The source of acetyl-CoA for FA synthesis in green oil seeds is accepted to be the plastidic pyruvate dehydrogenase complex (PDC) (Williams and Randall, 1979) (Figure 1.5). This is supported by the fact that its gene expression correlates with FA synthesis

during seed development in Arabidopsis (Ke *et al.*, 2000a) and because pyruvate has been found to be a preferred substrate of FA synthesis in plastids of Brassica embryos (Kang and Rawsthorne, 1994). PDC simultaneously provides acetyl-CoA and NADH for FA synthesis resulting in a stoichiometric generation of NADH for every two carbons added to the acyl-chain suggesting that the supply of NADH for FA synthesis is not limited. PDC also resides in the mitochondria. The activity of PDC is not easily manipulated because it is a multiprotein complex but plants with altered mitochondrial PDC (mtPDC) activity have been described (Marillia *et al.*, 2003). Activation of mtPDC was achieved in Arabidopsis plants with repressed expression of a gene coding for a pyruvate dehydrogenase kinase (PDHK) that inactivates mtPDC through phosphorylation. The constitutive and seed-specific repression of PDHK in transgenic plants led to elevated respiration, increased seed oil content, and higher rate of radiolabeled Pyr incorporation into seed lipids. Changes in flux or the acetyl-CoA pool in mitochondria or plastids were not examined, but the authors speculated that increased generation of acetyl-CoA by mtPDC led to the increased export from mitochondria and import into the plastid of acetate, and resulted in larger carbon flux into FA synthesis.

1.2.3.6 Green and non-green model seeds

Isolated plastids from oil seeds have been used extensively in radiolabel feeding experiments to address the significance of individual biochemical pathways in seed oil accumulation. Not only green seeds such as those of Arabidopsis and Brassica accumulate oil. Some non-green seeds such as castor beans, sesame, and sunflower are

important sources of commercial oil. FA synthesis in the absence of photosynthesis has been addressed using these systems.

1.2.3.6.1 Sunflower (*Helianthus annuus*)

Sunflower seeds are non-photosynthetic and accumulate oil in the embryo. Radiolabel feeding experiments showed that isolated plastids of sunflower embryos at the oil accumulating stage prefer malate and to a lesser extent pyruvate as substrates of FA synthesis (Pleite *et al.*, 2005). The flux through the plastidic OPPP measured by the release of labeled CO_2 from [1-^{14}C] G6P was reduced by the addition of malate while it was greatly increased by Pyr. G6P did not serve as a substrate of FA synthesis, but when added together with Pyr, greatly enhanced the incorporation of Pyr into FAs. This study suggests that two sources of reducing equivalents exist for FA synthesis in sunflower seed embryo, the OPPP and the malic enzyme, similar to that in castor bean endosperm (Figure 1.6a).

1.2.3.6.2 Castor bean (*Ricinus communis*)

Castor beans are not photosynthetic and accumulate large amounts of TAGs in the endosperm, with the majority of FAs constituting an unusual hydroxy FA, ricinoleic acid. Isolated leucoplasts of developing castor bean endosperm prefer malate as the substrate for FA synthesis (Smith *et al.*, 1992). The incorporation of malate and Pyr into FAs bypasses the requirement of reducing equivalents in contrast to that of acetate. It is speculated that a malic enzyme converts malate to Pyr, simultaneously generating NADPH and CO_2, and subsequently PDC generates acetyl-CoA and NADH.

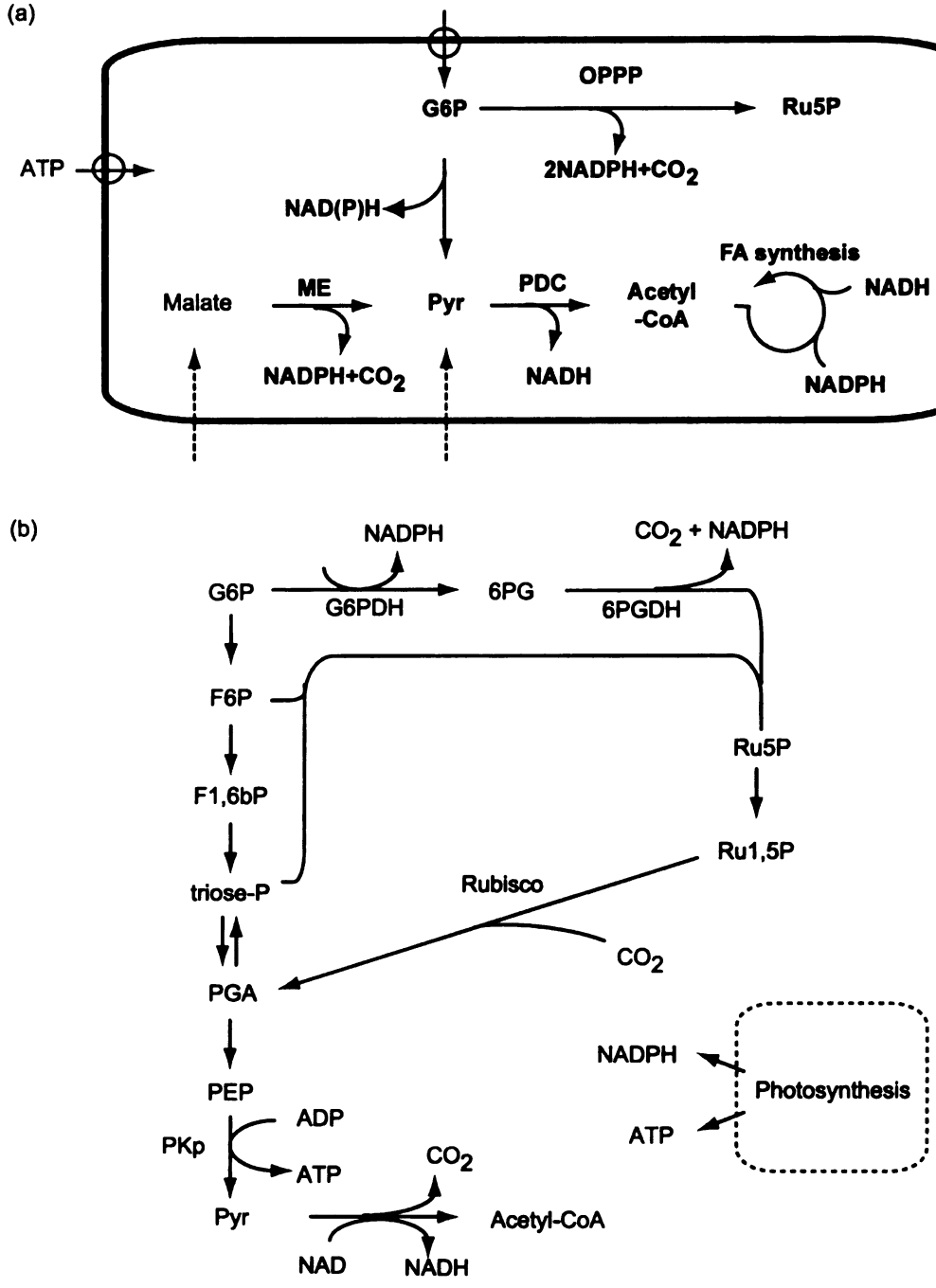


Figure 1.6 Models of cofactor supply in oil seed metabolism
(a) Alternative sources of NADPH in plastids of various oil seeds.
(b) Refixation of CO_2 by Rubisco increases carbon economy but requires energy and NADPH (Adapted from Schwender et al., 2004, Figure 1).

Recently the presence of a leucoplastic malic enzyme activity and its correlation with the onset of FA synthesis has been demonstrated (Shearer *et al.*, 2004). OPPP has been suggested to contribute 20 to 27% of total reducing equivalents required for FA synthesis in castor beans using 3-^{3}H-labeled glucose (Agrawal and Canvin, 1970; Agrawal and Canvin, 1971).

1.2.3.6.3 Brassica

Not only because it is one of the major seed oil crops, but also because its seeds are large enough to manipulate, Brassica has been the preferred plant in many experiments addressing seed oil accumulation. As such, the light transmission of silique walls and the structural analysis of seed plastids as well as their photosynthetic activity have been investigated using Brassica (Eastmond *et al.*, 1996; Asokanthan *et al.*, 1997; King *et al.*, 1998).

The presence of a complete glycolytic pathway and the two dehydrogenases of the OPPP, G6PDH and 6PGDH, has been shown in both the cytosol and plastids (Kang and Rawsthorne, 1994). In isolated plastids from Brassica embryos, Pyr is the most preferred substrate for FA synthesis, whose incorporation into FA is greatly enhanced by the presence of G6P (Kang and Rawsthorne, 1996). This was the first indication that G6P contributes to FA synthesis by providing NAPDH through OPPP. This effect increased as seeds developed from the starch- to the oil accumulating stage (Eastmond *et al.*, 1996; Kang and Rawsthorne, 1996). Isolated leaf chloroplasts and seed plastids of Brassica were compared in terms of energy and light requirement for storage compound synthesis (Eastmond and Rawsthorne, 1998). Interestingly, FA and starch syntheses of leaf plastids

are completely dependent on light whereas those from seeds are dependent on the addition of ATP and less dependent on light. These results have led to the conclusion that seed plastids depend on other means than photosynthesis, most likely OPPP for the supply of reductant and import of ATP.

Recently, OPPP activity in isolated plastids of developing embryos was measured in the presence of various compounds to mimick different electron sink sizes. The mimicked conditions and the compounds added were as follows: (1) glutamate synthesis (glutamine and 2-oxoglutarate), (2) FA synthesis (carbonate, ATP CoA and thiamine pyrophosphate), (3) oxidation of PET (an electron accepting herbicide methylviologen, MeV). The flux through OPPP increased under all the conditions in the order of 1, 2, 3 (Hutchings *et al.*, 2005). The authors conclude that OPPP is not limited in FA synthesis because the flux could be increased further by addition of MeV. Although these findings offer insight into the demand of reductant for the three biochemical processes, the results of this experiment rely on the efficiency of import of the cofactor and the reconstitution of *in vivo* rates of each electron consuming pathway. Thus, further experiments are required to corroborate these conclusions.

1.2.3.6.4 Flux analysis in Brassica

Feeding radiolabeled substrates to isolated plastid tells us the biochemical capacities of the plastid under substrate saturated conditions, but may not reflect the true flux *in vivo.* Metabolic flux analysis was carried out in Brassica embryos cultivated in media mimicking endosperm liquid. Two main questions were addressed: (1) what are the contributions of photosynthesis and OPPP in supplying NADPH for FA synthesis, and

(2) where does the CO_2 released in various reactions go. The second question arose from the observation that for every acetyl-CoA formed by PDC CO_2 is released leading to one-third of carbon loss from glycolysis. The loss of carbon is even greater when considering that CO_2 is also released during the conversion of 6-phosphogluconate to Ru5P by 6PGDH in OPPP. The flux through a biochemical pathway is measured by comparing substrate labeled at a specific carbon with the labeled product, and deducing from the patterns of label which pathway it derived from. Flux analysis requires the establishment of a metabolic network through assumptions based on the knowledge in the biochemical activities in the system under investigation. This was made possible by the abundant studies reported on oil seed metabolism.

Initial experiments focused on the first question by measuring flux of ^{13}C-labeled sugar into various products such as lipids, starch, amino acids. Of the glucose that entered the embryo, 10% was found to have been metabolized by the OPPP, which accounts for 44% of the NADPH and 22% of the total reducing equivalents required for FA synthesis (Schwender *et al.*, 2003). In the following experiments the second question was addressed by adding ^{13}C-labeled alanine in combination with glucose (Schwender *et al.*, 2004). Alanine entering glycolysis in the form of Pyr allowed the measurement of CO_2 release by PDC independent of the upstream glycolysis (see Schwender *et al.*, 2004, Figure 1b).

The data suggested that the carbon, which was lost during acetyl-CoA formation was refixed by ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) resulting in the

labeling of PGA (Schwender *et al.*, 2004) (Figure 1.6b). PGA then re-enters glycolysis and FA synthesis bypassing the conventional regeneration of pentose-phosphate through the Calvin cycle. This model improves the carbon economy but also causes a dilemma for the seeds of increased demand for ATP and NADPH required for carbon fixation, which presumably comes from photosynthesis and OPPP. Photosynthesis and OPPP are not exclusive in their contributions and the seeds possibly possess a dynamic metabolic network that responds to the environmental and diurnal light fluctuations.

Measurement of metabolic flux is a powerful method in developing the flux map but it relies on how well the experimental conditions reproduce that of the living organism. As of now in many cases, it is difficult to distinguish metabolic pathways in different compartments. In the future, the use of mutants disrupted in a single step of a pathway may partly contribute by targeting a specific step in a subcellular compartment.

1.2.3.7 Alternative means of transferring reducing power

1.2.3.7.1 Indirect transport of reducing equivalents by transporters

Interestingly, the mitochondrial membrane has the capacity to directly shuttle reducing equivalents (Neuburger *et al.*, 1984), but plastidic membranes have not been shown to allow NAD(P)H to cross either by diffusion or a transporter (Heineke *et al.*, 1991) (reviewed in Flugge and Heldt, 1984; Neuhaus and Wagner, 2000). However, plastidic transporters such as the triose-phosphate transporter (TPT) and the malate valve can shuttle reducing power across the plastidic membrane by coupling it to the oxidation of triose-phosphate (triose-P) or malate, respectively (Heineke *et al.*, 1991; Scheibe, 1991; Scheibe, 2004).

TPT transports dihydroxyacetone phosphate (DHAP) or glyceraldehyde-3-phosphate (GAP) out of the plastid in exchange for an incoming 3-phosphoglycerate (PGA), an intermediate required for the Calvin cycle. This serves as an indirect export of reducing equivalents and ATP because the two reactions that convert triose-phosphate to 3-PGA by 3-phosphoglyceraldehyde dehydrogenase (GAPDH) and 3-phosphoglycerate kinase generate NADH and ATP (Flugge and Heldt, 1984) (Figure 1.7a). These two reactions can be replaced by an NADP-dependent non-phosphorylating GAPDH that produces NADPH. In Arabidopsis, TPT is involved in the export of photosynthates in the form of triose-P (and the indirect export of energy and reducing power) and the export of carbon mobilized from starch (Schneider *et al.*, 2002). The transcript for TPT decreases during development in Arabidopsis seeds temporally correlating with the synthesis and mobilization of starch (White *et al.*, 2000; Ruuska *et al.*, 2002).

Another possible mechanism of indirect shuttling of reducing power is the malate valve. In chloroplasts, the malate valve exports excess reducing power to avoid over-reduction by photosynthesis (Scheibe, 2004). The chloroplastic malate dehydrogenase (MDH) utilizes NADPH to reduce malate to oxaloacetic acid (OAA). OAA is exported to the cytosol and malate is formed in the reverse reaction by a cytosolic MDH utilizing NAD(P) (Scheibe, 2004) (Figure 1.7b). In non-green plastids or chloroplasts in the dark, triose-P is imported and oxidized to generate ATP by GAPDH. NADH that is concomitantly produced is used by MDH, and a similar shuttle as that of photosynthetic chloroplasts functions to allow continuous generation of ATP (Figure 1.7c).

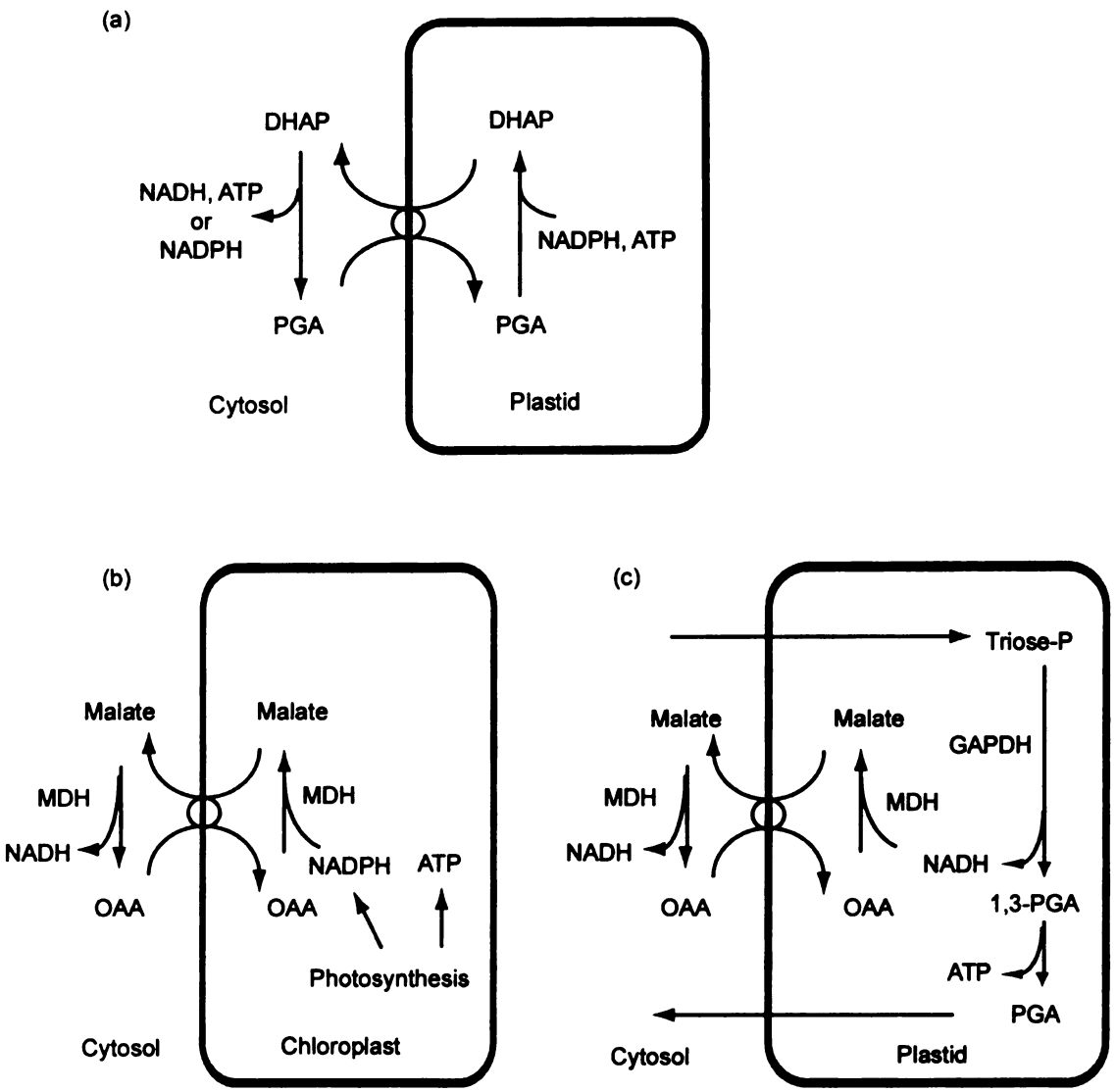


Figure 1.7 Indirect transport of reducing equivalents
(a) Triose-phosphate transporter, (b) Malate valve in the chloroplasts, (c) Malate valve in non-green plastids or chloroplasts in the dark (adapted from Scheibe, 2004, Figure 1).

A transporter that shuttles OAA has been found in maize and spinach chloroplasts (Eastmond *et al.*, 1997), and more recently in Arabidopsis (Taniguchi *et al.*, 2002). The kinetic parameters of the latter suggest its involvement in the malate valve.

Transport mechanisms that indirectly import cofactors (energy and reducing equivalents) into the plastids have been implied to support FA synthesis in seeds. Import of malate into castor bean endosperm plastids supplies NADPH (Smith *et al.*, 1992) and the PEP transporter in Arabidopsis and Brassica supplies ATP for FA synthesis (Kubis and Rawsthorne, 2000; Ruuska *et al.*, 2002). It is possible that other transport mechanisms similar to the TPT and the malate valve also contribute in FA synthesis in oil seeds.

1.2.3.7.2 Transhydrogenase, NAD(H) kinase, NADP(H) phosphatase

Many studies on NAD(P)H metabolism in mitochondria have been published, some of which are discussed in this section to provide insights into alternative mechanisms of NADPH generation. But so far these mechanisms are associated only with mitochondria and have not been reported in relation with the plastids.

Mammalian mitochondria possess a transmembrane transhydrogenase activity that transfers the hydrogen from NADH in the intermembrane space to NADP in the matrix. It is driven by the proton motive force and is responsible for the high NADPH/NADP ratio (ten-fold) in the matrix (Hoek and Rydstrom, 1988; Lenartowicz, 1990; Moller and Rasmusson, 1998; Moller, 2001). The presence of a transhydrogenase activity in plant mitochondria has been suggested (Carlenor *et al.*, 1998; Bykova *et al.*, 1999). Alternatively, an activity capable of direct conversion between NADPH and NADH,

NAD(H) kinase and NADP(H) phosphatase, has been detected in oat seeds (Gallais *et al.*, 2000a; Gallais *et al.*, 2000b). Such mechanisms have not been reported in plastids.

1.3 The genetics and molecular biology of seed oil accumulation

1.3.1 Arabidopsis mutants affected in seed oil

The use of forward genetics in Arabidopsis has allowed the discovery of factors that are important in oil accumulation independently of the characterization of their biochemical activity. The isolation of a genetic mutant is the proof of a certain activity playing a role in the process in question. Such mutants involved in seed oil accumulation include *wrinkled1* (*wri1*), *triacylglycerol biosynthesis defect1* (*tag1*), and *shrunken seed1* (*sse1*) discussed below.

In a screen for Arabidopsis mutants affected in very long-chain fatty acids (VLCFAs), an ethyl methanesulfonate-induced mutant was identified that had altered seed oil FA composition. The mutant plant had reduced DGAT activity, reduced total seed oil content and altered FA composition which indicated that the biosynthesis of TAG affects the FA composition in TAG (Katavic *et al.*, 1995; Routaboul *et al.*, 1999). This *TAG1* locus codes for a DGAT whose activity is associated with microsomes and although *TAG1* is expressed throughout the plant, the mutant phenotype was restricted to TAG in the seeds (Zou *et al.*, 1999). Seed-specific over-expression of *TAG1* complemented the FA composition phenotype and also caused an increase in oil content and seed size (Jako *et al.*, 2001).

The mutant *wri1* was found in a screen for mutants with decreased seed oil content (Focks and Benning, 1998). In the mutant seed TAG is reduced to 20% of WT level

along with a global down-regulation of glycolytic enzyme activity, most severely hexokinase (HXK) and PPi-dependent PFK with 80% reduction, and fructokinase, aldolase, phosphoglycerate mutase, enolase and PK, with 40% reduction. A higher rate of Pyr incorporation into TAG was observed in the mutant, indicating FA synthesis was not impaired. *WRI1* codes for a putative protein classified as an AP2 domain transcription factor (Cernac and Benning, 2004). Because glycolytic enzymes such as HXK, enolase and PK are down-regulated in the mutant seeds, WRI1 is speculated to be a direct transcriptional regulator of the genes coding for these enzymes. The expression of *WRI1* complements the seed oil phenotype of the mutant. Strikingly, over-expression of *WRI1*, only in the presence of sugar, produces seedlings that show embryo-like identity that accumulate seed oil and storage protein (Cernac and Benning, 2004).

The *sse1* mutant was identified by the shrunken appearance of the seeds and showed reduced oil content with concomitant accumulation of starch. *SSE1* encodes a protein similar and functionally equivalent to the yeast Pex16p, a peroxisomal protein involved in peroxisome biogenesis (Lin *et al.*, 1999). Radiolabeled substrate feeding to *sse1* show that FA synthesis is impaired but not the incorporation of FA into TAG (Lin *et al.*, 2004). The seed oil phenotype has not been complemented by a functional SSE1, and hence the direct cause of reduced accumulation of TAG in the mutant and how that involves abnormal peroxisomes remain unknown. Nonetheless, the fact that peroxisomes are present in WT developing seeds in close proximity to growing oil bodies and findings of this study indicate that some biochemical and/or developmental process takes place that involves peroxisomes during seed development.

1.3.2 Genetic engineering of seeds to increase oil production

In contrast to forward genetics for the study of seed oil accumulation, targeted approaches to increase oil content using transgenic plants have been possible because of the abundant information in FA and lipid synthesis. Some case studies are discussed in this section.

An LPAT from yeast with specificity to long-chain FAs was introduced into a high erucic acid line of Brassica (Zou *et al.*, 1997). This led to an increase in erucic acid content in TAG and a 50% increase in total seed TAG compared to WT. Some descriptions of the transgenic plant such as the subcellular targeting of the product of the protein and the effect of the constitutive transgene expression on the whole plant were not included in this report. Nonetheless, this result together with the increased seed oil achieved by enhanced DGAT activity (Jako *et al.*, 2001), demonstrates that increasing the sink size at later steps of TAG synthesis such as the activity of acyltransferases, can increase the flux from FA to TAG.

Removal of FAs from the acyl-ACP pool by increasing the activity of acyl-ACP thioesterase (FAT) has been shown to influence the flux into TAG. The coding gene of lauroyl-ACP thioesterase from California bay was constitutively overexpressed in Arabidopsis (Voelker *et al.*, 1992) and Brassica (Voelker *et al.*, 1996; Eccleston and Ohlrogge, 1998) and in both cases, this approach led to increase in laurate content in seed TAG. Interestingly, a simultaneous induction of β-oxidation and FA synthesis was observed in Brassica (Eccleston and Ohlrogge, 1998), but not in Arabidopsis (Hooks *et al.*, 1999), which instead showed an increase in the medium chain acyl-CoA pool.

Two additional steps have been demonstrated to affect the flux of FA into seed oil. The plant *Cuphea wrightii* accumulates caprylate (8:0) and caprate (10:0) in the seed TAG. The genes coding for the condensing enzyme (KASIV) in FAS and the acyl-ACP thioesterase (FAT), both specific to medium-chain FAs (MCFAs), were cloned and introduced into Arabidopsis (Dehesh *et al.*, 1998) and Brassica (Leonard *et al.*, 1998). In both cases, the accumulation of MCFA in seed TAG was observed. The transgenic Brassica plants contained increased medium chain acyl-CoA pool (Larson *et al.*, 2002) indicating that the incorporation of the novel FA into TAG was still limited by other steps. These studies suggest that the upper limit of seed oil content is determined by the sink size, which then leads to rapid turnover and retention of FA in the acyl-CoA pool. Whether the introduction of LPAT or DGAT to such transgenic lines will further increase the flux of novel FAs into seed TAG has not been tested.

These cases with transgenic plants tell us that the qualitative and quantitative control of FA and TAG synthesis is a concerted process that is specific to each plant. An important observation is that when lipid metabolism was altered either by a loss of an endogenous or introduction of a heterologous enzyme, the plant responded with compensatory processes to maintain homeostasis. In several cases, novel FAs accumulate predominantly in seed TAG and are excluded from membrane lipids despite constitutive expression of the transgene (Zou *et al.*, 1997; Thomaeus *et al.*, 2001). In some transgenic plants β-oxidation is increased to remove novel FAs leading to a futile cycling (Eccleston and Ohlrogge, 1998; Poirier *et al.*, 1999; Moire *et al.*, 2004). Such a process can be presumed to affect the whole plant carbon economy, which is not beneficial in agronomic productivity. On the other hand, not all novel FAs can be excluded from membrane

lipids. Transgenic Arabidopsis plants expressing a gene coding for the very-long-chain FA (VLCFA) condensing enzyme (*FAE1*) accumulated VLCFAs in membrane lipids and exhibited severe morphological phenotypes (Millar *et al.*, 1998). This problem was neatly solved by seed-specific expression of a single *FAE* gene from Garden nasturtium (*Tropaeolum majus*) in Arabidopsis and tobacco (*Nicotiana tabacum*) (Mietkiewska *et al.*, 2004). The *FAE* gene was isolated from the seeds of Garden nasturtium that accumulate a large amount of erucic acid in the seed TAG, and increased its incorporation into TAG. Apparently, the key to successful seed oil engineering lies not only in establishing a strong source by selecting the right enzymes that produce the target FA, but also in sink compatibility, ensuring that the acceptor plant has the capacity to process it.

1.3.3 Natural variation in seed quality and oil

Natural variation in seed oil has been observed among Arabidopsis accessions that are now abundantly available (O'Neill *et al.*, 2003). In an attempt to find the genetic factors affecting seed oil content and composition, quantitative trait loci (QTL) analysis has been carried out using recombinant inbred lines (RILs) of accessions Cape Verdi Islands (Cvi) and Landsberg *erecta* (L*er*) (Hobbs *et al.*, 2004). QTL affecting oil content were mapped to chromosomal regions, and the polymorphic alleles remain to be identified. Interestingly, this work showed the lack of correlation between seed mass and oil content, and that up to 30% of the variation in seed oil content was due to environmental factors.

1.3.4 Transcriptome analysis of developing seeds

Transcript profiling has been reported in many different types of seeds including castor bean (Van De Loo *et al.*, 1995), Brassica (Soeda *et al.*, 2005), sesame (Suh *et al.*, 2003) and barley (Sreenivasulu *et al.*, 2002) and proteomic analysis has been reported in *Medicago truncatula* as a model legume (Gallardo *et al.*, 2003). Such studies are useful by themselves, but also comparative analyses between different species have given rise to questions and hypotheses and have led to new discoveries. Arabidopsis is by far the most advantageous plant for bioinformatic studies because of the size of the available database. This section will focus on what we have learned from the Arabidopsis developing seed transcriptome.

As a starting point in studying oil seed metabolism, expressed sequence tag (EST) analysis of Arabidopsis developing seeds (5 to 13 DAF) was carried out (White *et al.*, 2000). More than 10,000 clones were sequenced after the pre-screening removal of storage protein and other abundant transcripts. This and the enrichment of seed mRNA allowed identification of approximately ~4,200 novel sequences representing 3,049 genes (March 2000), 40% of which belonged to the category whose closest matches in the database had unidentified function or non-significant homology. To identify the seed-specific genes, the abundance of genes in a cDNA library from seeds (8-11 DAF) was compared with that from leaves and roots (Girke *et al.*, 2000). A total of 2,715 putative unique clones, some of which may represent the same gene, were spotted on a cDNA array. In total, 25% of the genes were expressed more than 2-fold higher in seeds compared to leaves and roots. The categories of those found to be expressed seed-specifically were in order of their abundance, storage protein, unidentified function and

non-significant homology, development, and lipid metabolism. The analysis of the promoter regions of such genes has allowed identification of novel seed-specific promoters.

Finally the dynamics of gene expression was studied using an array of ~6,000 spots representing ~3,500 unique genes, more than 100 of which are genes related to lipid metabolism (Ruuska *et al.*, 2002). The array was hybridized with cDNA from seeds of different stages at five time points between 5 to 13 DAF, during which storage compound metabolism was predicted to be the highest. Surprisingly, only 35% of the genes reliably detected over background were found to change more than 2-fold during development. Changes of more than 2-fold were detected in 70% of the seed-specific genes, and 40% of the lipid genes. Three major temporal patterns were found for the lipid genes: bell-shaped curve that peaked between 8 and 11 DAF, late peak with increase starting at 8 DAF, and no large change during development. Surprisingly, different patterns were observed even for genes whose proteins belong to the same complex, such as KASI and KASIII within FAS. The most striking finding in this study was the contrapuntal pattern of genes belonging to different metabolic groups. Genes related to starch metabolism were expressed in early seeds, whereas lipid genes showed a later peak and declined. Storage protein genes showed a later onset than lipids and maintained high levels. These patterns strongly resembled that of storage compound accumulation in developing seed, though temporally shifted. Although the patterns did not apply to all genes involved in a pathway which could be due to various reasons, the patterns suggest the involvement of a specific gene product in a biochemical pathway and reveal the highly concerted regulation of the metabolic network.

1.4 OPPP and G6PDH

OPPP and G6PDH have been studied in various organisms involved in different physiological processes. G6PDH was first reported in the 1930's and in the earliest studies its involvement in yeast fermentation was examined (Green and Keilin, 1936). In humans, genetic variants for G6PDH have been recognized for decades to associate with hypersensitive response to certain nutritional components, oxidative stress, and more recently, resistance to malaria infections (Vulliamy *et al.*, 1992; Kletzien *et al.*, 1994; Martini and Ursini, 1996). In plants, the study of OPPP and G6PDH has begun more recently and many questions on the roles of G6PDH *in vivo* remain to be answered.

1.4.1 OPPP and G6PDH in plants

Plants are unique in that OPPP resides in at least two subcellular compartments: the cytosol and the plastid (Figure 1.4). The two key enzymes in the OPPP, G6PDH and 6PGDH, have been studied in various plants (Figure 1.5). In all plants examined, the two enzyme activities are present in both cytosol and plastids (Schnarrenberger *et al.*, 1973; Debnam and Emes, 1999). To date, only one study reports a peroxisomal G6PDH (protein and activity) in pea leaves (*Pisum sativum*), but neither nucleotide nor amino acid sequences are available (Corpas *et al.*, 1998). It is inactivated by reduction and its structure is predicted to resemble that typical to G6PDHs because it is recognized by immunolabeling. The authors speculate that G6PDH as well as 6PGDH, the second enzyme in OPPP, are involved in the regeneration of NADPH required for the ascorbate-glutathione cycle occurring in the leaf peroxisome.

The abundance of information on G6PDH and 6PGDH from various organisms has allowed their phylogenetic analyses (Wendt *et al.*, 1999; Krepinsky *et al.*, 2001). It has been suggested that plant G6PDHs have an eubacterial origin (both cytosolic and plastidic isoforms are more similar to those of animals and fungi than to bacterial isoforms), and that the plastidic isoform is a product of gene duplication in the ancestor that precedes the endosymbiotic event that acquired the plastid (Wendt *et al.*, 1999; Krepinsky *et al.*, 2001). In contrast to G6PDHs, cytosolic and plastidic 6PGDH are most similar to each other and to cyanobacteria than to other non-photosynthetic eukaryotes, indicating that the 6PGDH genes were acquired in the nucleus from the endosymbiont and the pre-existing enzyme was eventually lost (Krepinsky *et al.*, 2001).

Of the two key enzymes of the OPPP, G6PDH has been far more studied than 6PGDH. Because it is the first enzyme of the pathway and has many more regulatory mechanisms identified than 6PGDH, G6PDH is believed to be the rate-limiting step of the pathway. The most widespread regulation of eukaryotic G6PDH is feed-back inhibition by NADPH, hence it is presumed to act as the cellular redox sensor. In cyanobacteria and plant plastids of green tissues, G6PDH co-exists with photosynthesis, the primary source of NADPH. This has given rise to a redox regulatory mechanism of G6PDHs in cyanobacteria (Cossar *et al.*, 1984) and plant plastids that involves its inactivation in the light (Lendzian, 1980; Scheibe and Anderson, 1981; Fickenscher and Scheibe, 1986; Graeve *et al.*, 1994; Wenderoth *et al.*, 1997; Wendt *et al.*, 2000). Three conserved cysteine residues have been identified in algae and plant plastidic G6PDH, two of which are responsible for the redox regulation through thioredoxin (Wenderoth *et al.*, 1997). Site directed mutagenesis of the two residues resulted in the loss of redox

regulation of the enzyme activity. In cyanobacteria, two conserved cysteine residues that are involved in thioredoxin-mediated redox regulation are present and found at different positions of the protein from those of plastidic G6PDHs (Wendt *et al.*, 1999). This indicates that the redox regulatory mechanisms arose independently in plants and cyanobacteria to prevent carbon oxidation when NADPH generation by photosynthesis is sufficient.

In contrast to the dual localization of OPPP, the non-oxidative portion of the pentose phosphate pathway (PPP) in the cytosol is incomplete, judged by the lack of PPP enzymes such as transketolase and transaldoase (Schnarrenberger *et al.*, 1995; Debnam and Emes, 1999; Henkes *et al.*, 2001; Kruger and von Schaewen, 2003; Hutchings *et al.*, 2005). Thus the fate of the cytosolic Ru5P (product of OPPP) was unclear, although there were indications that a pentose phosphate transporter was present because pentose-P could support OPPP for nitrite assimilation similarly to G6P (Bowsher *et al.*, 1992). A pentose-P transporter with specificity to xylulose-5-phosphate (X5P) and triose-P was found in Arabidopsis (XPT, xylulose-5-phosphate transporter) (Eicks *et al.*, 2002). It is speculated that XPT connects the cytosolic and plastidic OPPPs at the level of pentose phosphates by shuttling X5P formed from Ru5P by Ru5P epimerase. The expression of the gene was detected in green vegetative parts, root tips, and specific floral tissues, but the precise physiological role of XPT is unknown.

1.4.2 Current questions

1.4.2.1 Reductant supply for nitrogen assimilation

In plants, G6PDH has been most frequently described for its role in nitrogen assimilation. The induction of its activity or transcript has been described in various systems including pea roots (Bowsher *et al.*, 1992), Chlamydomonas (Huppe *et al.*, 1994), barley roots (Wright *et al.*, 1997; Esposito *et al.*, 2003), maize roots (Redinbaugh and Campbell, 1998), tobacco roots and leaves (Debnam *et al.*, 2004) and Arabidopsis (Wang *et al.*, 2003). Several lines of evidence indicate the direct role of G6PDH in nitrogen metabolism by supplying reductants. The initial observation was made in pea root plastids, in which G6P (and R5P) supported glutamate synthesis in correlation with the release of CO_2 from [1-^{14}C] G6P (indicator of OPPP flux), which was abolished when glutamate synthase was inhibited by azaserine (Bowsher *et al.*, 1992). Electron transport from G6PDH to nitrite reductase could be reconstituted *in vitro* (Wright *et al.*, 1997; Jin *et al.*, 1998), and simultaneous supply of G6P with nitrate increased NADPH production (Jin *et al.*, 1998). The increase in G6PDH activity was associated with the appearance of a novel plastidic isoform in Chlamydomonas (Huppe and Turpin, 1996) and barley roots (Esposito *et al.*, 2001). In barley roots G6PDH has been shown to correlate with the formation of glutamate from glutamine, suggesting that glutamine synthetase (GS) and G6PDH activity are required (Esposito *et al.*, 2003). The increase in activity and the appearance of G6PDH was associated with the induction of NADH-dependent glutamine:oxoglutarate amino transferase (GOGAT) in roots and ferredoxin-dependent GOGAT in leaves. The reducing power of NADPH is presumably transferred to

ferredoxin by ferredoxin:NADP reductase, but how reducing power is converted from NADPH to NADH to support NADH-dependent GOGAT is still unknown.

1.4.2.2 Response to oxidative stress

A correlation of G6PDH with various oxidative stresses has been observed at the level of mRNA abundance or enzyme activity. Inhibitors of PET, such as methylviologen (also called paraquat), activate a plastidic isoform which is associated with the loss of phosphorylation of the protein (Hauschild and von Schaewen, 2003). Transgenic tobacco plants with decreased levels of a plastidic isoform caused by antisense expression showed an unexpected increase in resistance to methylviologen (Debnam *et al.*, 2004). Although this study showed a seemingly opposite role of G6PDH in oxidative stress response, it demonstrated that the alteration of metabolite levels due to change in G6PDH activity changed the plant's response to oxidative stress. The correlation of G6PDH with pathogenesis has been reported with fungal elicitors (Batz *et al.*, 1998) and viral infection (Sindelar and Sindelarova, 2002). In the latter study, G6PDH activity correlated with viral replication in leaves infected with different virus strains. Interestingly, higher G6PDH activity was observed in local leaves with less virus content. This may suggest G6PDH involvement in multiple processes during viral infection, such as signaling or short- and long-term oxidative stress response, as well as supplying viral replication with a nucleotide precursor, ribose-5-phosphate, an intermediate of the non-oxidative PPP. The differential increase in G6PDH transcripts under salt stress has also been observed (Nemoto and Sasakuma, 2000) suggesting common pathways in stress signaling.

1.4.2.3 Differential roles of plastidic isoforms P1 and P2

The finding that G6PDH activity is modulated by redox through thioredoxin (Scheibe and Anderson, 1981) led to the initial assumption that chloroplastic G6PDH would be inactive during the day or under photosynthetic conditions. Studies in potato (*S. tuberosum*) found two plastidic isoforms, P1 and P2, with distinct sequences (Graeve *et al.*, 1994; von Schaewen *et al.*, 1995; Wendt *et al.*, 2000). The two groups have been found to exist broadly in the plant kingdom, leading to the assumption that they may have distinct roles in plants (Wendt *et al.*, 2000).

In potato, the gene for P1 is expressed predominantly in green tissues, while that of P2 is ubiquitously expressed. Isoform P2 is less sensitive to inactivation by reduction by DTT and less sensitive to feed-back inhibition by NADPH. From these findings it was proposed that the two isoforms evolved to have different sensitivity to regulation by photosynthesis. Chloroplastic (P1) isoforms, because the transcript is predominant in green tissues, are suggested to be tightly regulated by photosynthetic generation of NADPH. Plastidic (P2) isoforms on the contrary are less tightly regulated by photosynthesis and ubiquitous. It is proposed that *in vivo* under high NADPH/NADP ratios, while P1 isoforms would be inactivated in green tissues under light conditions, P2 isoforms are able to function. After the initial exposure to nitrogen, high NADPH/NADP ratio has been observed during nitrogen assimilation in Chlamydomonas and barley roots (Wright *et al.*, 1997; Jin *et al.*, 1998). Indeed in barley root plastids, a P2 isoform with low sensitivity to feed-back inhibition and reductive inactivation was found (Esposito *et al.*, 2001). Such a G6PDH isoform less subject to these regulatory mechanisms is postulated to play a role in providing NADPH in non-photosynthetic tissues or in green

tissues independently of photosynthesis to support sudden increases in the demand for NADPH, e.g., under oxidative stress and FA synthesis (Wendt *et al.*, 2000; Esposito *et al.*, 2003).

A side-by-side comparison of the biochemical activities present in autotrophic (chloroplast) and heterotrophic plastids (chromoplast) in ripening pepper fruits has been reported (Thom *et al.*, 1998). The activity of the two OPPP enzymes was not markedly different in contrast to some glycolytic enzymes. A higher stimulation of OPPP was observed when substrates of nitrogen assimilation were added to chromoplasts than to chloroplasts, indicating that the chromoplasts were more dependent on OPPP than chloroplasts. Liquid assays for G6PDH activity cannot distinguish the change in the present isoforms, and it could be a reason for not detecting changes in OPPP. Comparative analysis between different plastids from a single plant addressing the presence of the different G6PDH isoforms (cytosolic, P1, P2) and their biochemical characteristics will greatly contribute in elucidating their roles *in vivo*.

1.4.3 New questions: The regulation of G6PDH and its role in whole cell, whole plant physiology

Little is known of the regulatory mechanisms of plant G6PDHs. Recently, a sugar-induced cytosolic G6PDH activity was found to occur at the mRNA level, which was associated with the presence of sugar response DNA *cis*-elements in the promoter region (Hauschild and von Schaewen, 2003). In mammalian cells, cytosolic G6PDH is also regulated by nutritional cues and hormones through alternative splicing (Tao *et al.*, 2002); G6PDH activity is increased by sugar and inhibited by lipids (Salati *et al.*, 2004).

The regulation of plastidic isoform (P1) was found to occur at the protein level through changes in redox, possibly involving a change in phosphorylation of the enzyme (Hauschild and von Schaewen, 2003).

Plant OPPP is unique in its dual localization and compartmentalization, which raises the complexity when considering its role in whole cell physiology. Few studies address such questions. Maize mutants with the cytosolic 6PGDHs disrupted were affected in the capacity of nitrogen assimilation (Averill *et al.*, 1998), indicating the presence of redox communication between the cytosol and the plastid. The discovery of the XPT (pentose-phosphate transporter) demonstrated the physical connection between plastidic and cytosolic non-oxidative pentose phosphate pathway (PPP) but how the two pathways coordinate in supplying NADPH to the cell still remains unknown. The close connection between biochemical pathways in central metabolism makes interpretation of an effect observed in plants with altered biochemical activity difficult. It is not always clear whether the effect is due to decreased supply of NADPH or precursors generated from PPP because of the cyclic nature of PPP and OPPP. Overlap of intermediates with glycolysis and the Calvin cycle also makes it difficult to discern whether the effect is primarily due to OPPP or is secondary. For example, transgenic tobacco plants decreased in plastidic transketolase activity had severe effects that extended to phenylpropanoid metabolism and photosynthesis (Henkes *et al.*, 2001), and demonstrated the complex network of metabolic flux and regulation.

To date, there is an accumulating knowledge on the specific characteristics of the different plant G6PDH isoforms. This together with the increasing number of genome sequences which allows bioinformatic identification of multiple G6PDH isoforms in an

organism will enable comparative analyses that will contribute to defining the roles of G6PDHs. Furthermore, many new questions await to be addressed such as, how G6PDH contributes in various subcellular compartments to short- and long-term response during oxidative stress, how the redox state in different organelles is communicated through OPPP, and where and how OPPP and G6PDH are placed in the complex metabolic network. The function of G6PDHs in the whole cell and in whole plant physiology still remains to be elucidated.

Reference

Agrawal, P.K., Canvin, D.T. (1970) Contribution of pentose phosphate pathway in developing castor bean endosperm. *Can. J. Bot.* **49,** 267-272.

Agrawal, P.K., Canvin, D.T. (1971) The pentose phosphate pathway in relation to fat synthesis in developing castor oil seed. *Plant Physiol.* **47,** 672-675.

Asokanthan, P.S., Johnson, R.W., Griffith, M., Krol, M. (1997) The photosynthetic potential of canola embryos. *Physiol. Plant.* **101,** 353-360.

Averill, R.H., Bailey-Serres, J., Kruger, N.J. (1998) Co-operation between cytosolic and plastidic oxidative pentose phosphate pathways revealed by 6-phosphogluconate dehydrogenase-deficient genotypes of maize. *Plant J.* **14,** 449-457.

Ayala, A., Fabregat, I., Machado, A. (1991) The role of NADPH in the regulation of glucose-6-phosphate and 6-phosphogluconate dehydrogenases in rat adipose tissue. *Mol. Cell Biochem.* **105,** 1-5.

Bao, X., Pollard, M., Ohlrogge, J. (1998) The biosynthesis of erucic acid in developing embryos of brassica rapa. *Plant Physiol* **118,** 183-190.

Barrett, P.B., Harwood, J.L. (1998) Characterization of fatty acid elongase enzymes from germinating pea seeds. *Phytochemistry* **48,** 1295-1304.

Batz, O., Logemann, E., Reinold, S., Hahlbrock, K. (1998) Extensive reprogramming of primary and secondary metabolism by fungal elicitor or infection in parsley cells. *Biol. Chem.* **379,** 1127-1135.

Baud, S., Boutin, J.P., Miquel, M., Lepiniec, L., Rochat, C. (2002) An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiol. Biochem.* **40,** 151-160.

Baud, S., Guyon, V., Kronenberger, J., Wuilleme, S., Miquel, M., Caboche, M., Lepiniec, L., Rochat, C. (2003) Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in Arabidopsis. *Plant J.* **33,** 75-86.

Berger, F. (2003) Endosperm: the crossroad of seed development. *Curr. Opin. Plant Biol.* **6,** 42-50.

Bewley JD, Hempel FD, McCormick S, Zambryski P (2000) Reproductive Development. In *Biochemistry & Molecular Biology of Plants* (Buchanan, B.B., Gruissem, W., Jones, R.J., eds). Rockville, MD: American Society of Plant Biologists, pp 988-1043.

Bonaventure, G., Bao, X., Ohlrogge, J., Pollard, M. (2004) Metabolic responses to the reduction in palmitate caused by disruption of the *FATB* gene in Arabidopsis. *Plant Physiol* **135,** 1269-1279.

Borisjuk, L., Rolletschek, H., Walenta, S., Panitz, R., Wobus, U., Weber, H. (2003) Energy status and its control on embryogenesis of legumes: ATP distribution within *Vicia faba* embryos is developmentally regulated and correlated with photosynthetic capacity. *Plant J.* **36,** 318-329.

Bowsher, C.G., Boulton, E.L., Rose, J., Nayagam, S., Emes, M.J. (1992) Reductant for glutamate synthase is generated by the oxidative pentose-phosphate pathway in nonphotosynthetic root plastids. *Plant J.* **2,** 893-898.

Browse, J., Roughan, P.G., Slack, C.R. (1981) Light control of fatty acid synthesis and diurnal fluctuations of fatty acid composition in leaves. *Biochem. J.* **196,** 347-354.

Browse, J., Slack, C.R. (1985) Fatty-acid synthesis in plastids from maturing safflower and linseed cotyledons. *Planta* **166,** 74-80.

Browse, J., Somerville, C.R. (1991) Glycerolipid synthesis - biochemistry and regulation. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **42,** 467-506.

Bykova, N.V., Rasmusson, A.G., Igamberdiev, A.U., Gardestrom, P., Moller, I.M. (1999) Two separate transhydrogenase activities are present in plant mitochondria. *Biochem. Biophys Res. Commun.* **265,** 106-111.

Carlenor, E., Persson, B., Glaser, E., Andersson, B., Rydstrom, J. (1998) On the presence of a nicotinamide nucleotide transhydrogenase in mitochondria from potato tuber. *Plant Physiol.* **88,** 303-308.

Carrari, F., Coll-Garcia, D., Schauer, N., Lytovchenko, A., Palacios-Rojas, N., Balbo, I., Rosso, M., Fernie, A.R. (2005) Deficiency of a plastidial adenylate kinase in Arabidopsis results in elevated photosynthetic amino acid biosynthesis and enhanced growth. *Plant Physiol* **137,** 70-82.

Caughey, I., Kekwick, R.G. (1982) The characteristics of some components of the fatty acid synthetase system in the plastids from the mesocarp of avocado (*Persea americana*) fruit. *Eur. J. Biochem.* **123,** 553-561.

Cavener, D.R., Clegg, M.T. (1981) Evidence for biochemical and physiological differences between enzyme genotypes in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A* **78,** 4444-4447.

Cernac, A., Benning, C. (2004) *WRINKLED1* encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant J.* **40,** 575-585.

Corpas, F.J., Barroso, J.B., Sandalio, L.M., Distefano, S., Palma, J.M., Lupianez, J.A., del Rio, L.A. (1998) A dehydrogenase-mediated recycling system of NADPH in plant peroxisomes. *Biochem. J.* **330,** 777-784.

Cossar, J.D., Rowell, P., Stewart, W.D.P. (1984) Thioredoxin as a modulator of glucose-6-phosphate dehydrogenase in a N_2-fixing cyanobacterium. *J. Gen. Microbiol.* **130,** 991-998.

Dahlqvist, A., Stahl, U., Lenman, M., Banas, A., Lee, M., Sandager, L., Ronne, H., Stymne, S. (2000) Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. *Proc. Natl. Acad. Sci. U. S. A* **97,** 6487-6492.

Debnam, P.M., Emes, M.J. (1999) Subcellular distribution of enzymes of the oxidative pentose phosphate pathway in root and leaf tissues. *J. Exp. Bot.* **50,** 1653-1661.

Debnam, P.M., Fernie, A.R., Leisse, A., Golding, A., Bowsher, C.G., Grimshaw, C., Knight, J.S., Emes, M.J. (2004) Altered activity of the P2 isoform of plastidic glucose 6-phosphate dehydrogenase in tobacco (*Nicotiana tabacum cv. Samsun*) causes changes in carbohydrate metabolism and response to oxidative stress in leaves. *Plant J.* **38,** 49-59.

Dehesh, K., Edwards, P., Fillatti, J., Slabaugh, M., Byrne, J. (1998) KAS IV: a 3-ketoacyl-ACP synthase from *Cuphea sp.* is a medium chain specific condensing enzyme. *Plant J.* **15,** 383-390.

Eastmond, P., Kolacna, L., Rawsthorne, S. (1996) Photosynthesis by developing embryos of oilseed rape (*Brassica napus* L.). *J. Exp. Biol.* **47,** 1763-1769.

Eastmond, P.J., Dennis, D.T., Rawsthorne, S. (1997) Evidence that a malate/inorganic phosphate exchange translocator imports carbon across the leucoplast envelope for fatty acid synthesis in developing castor seed endosperm. *Plant Physiol* **114,** 851-856.

Eastmond, P.J., Rawsthorne, S. (1998) Comparison of the metabolic properties of plastids isolated from developing leaves or embryos of *Brassica napus* L. *J. Exp. Biol.* **49,** 1105-1111.

Eastmond, P.J., Rawsthorne, S. (2000) Coordinate changes in carbon partitioning and plastidial metabolism during the development of oilseed rape embryos. *Plant Physiol* **122,** 767-774.

Eccleston, V.S., Ohlrogge, J.B. (1998) Expression of lauroyl-acyl carrier protein thioesterase in brassica napus seeds induces pathways for both fatty acid oxidation and biosynthesis and implies a set point for triacylglycerol accumulation. *Plant Cell* **10,** 613-622.

Eicks, M., Maurino, V., Knappe, S., Flugge, U.I., Fischer, K. (2002) The plastidic pentose phosphate translocator represents a link between the cytosolic and the plastidic pentose phosphate pathways in plants. *Plant Physiol* **128,** 512-522.

Emes, M.J., Neuhaus, H.E. (1997) Metabolism and transport in non-photosynthetic plastids. *J. Exp. Bot.* **48,** 1995-2005.

Esposito, S., Carfagna, S., Massaro, G., Vona, V., Rigano, V.D. (2001) Glucose-6-phosphate dehydrogenase in barley roots: kinetic properties and localisation of the isoforms. *Planta* **212,** 627-634.

Esposito, S., Massaro, G., Vona, V., Rigano, V.D., Carfagna, S. (2003) Glutamate synthesis in barley roots: the role of the plastidic glucose-6-phosphate dehydrogenase. *Planta* **216,** 639-647.

Fickenscher, K., Scheibe, R. (1986) Purification and properties of the cytoplasmic glucose-6-phosphate-dehydrogenase from pea leaves. *Arch. Biochem. Biophys.* **247,** 393-402.

Flugge, U.I., Heldt, H.W. (1984) The phosphate-triose phosphate-phosphoglycerate translocator of the chloroplast. *Trends Biochem. Sci.* **9,** 530-533.

Focks, N., Benning, C. (1998) *wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* **118,** 91-101.

Galili, G., Sengupta-Gopalan, C., Ceriotti, A. (1998) The endoplasmic reticulum of plant cells and its role in protein maturation and biogenesis of oil bodies. *Plant Mol. Biol.* **38,** 1-29.

Gallais, S., de Crescenzo, M.A., Laval-Martin, D.L. (2000a) Evidence of active $NADP^+$ phosphatase in dormant seeds of *Avena sativa* L. *J. Exp. Bot.* **51,** 1389-1394.

Gallais, S., Pou de Crescenzo, M.A., Laval-Martin, D.L. (2000b) Changes in soluble and membrane-bound isoforms of calcium-calmodulin-dependent and -independent NAD^+ kinase, during the culture of after-ripened and dormant seeds of *Avena sativa*. *Aust. J. Plant Physiol.* **27,** 649-658.

Gallardo, K., Le Signor, C., Vandekerckhove, J., Thompson, R.D., Burstin, J. (2003) Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol* **133,** 664-682.

Girke, T., Todd, J., Ruuska, S., White, J., Benning, C., Ohlrogge, J. (2000) Microarray analysis of developing Arabidopsis seeds. *Plant Physiol* **124,** 1570-1581.

Goldberg, R.B., Depaiva, G., Yadegari, R. (1994) Plant embryogenesis-zygote to seed. *Science* **266,** 605-614.

Goffman, F.D., Alonso, A.P., Schwender, J., Shachar-Hill Y., Ohlrogge, J.B. (2005) Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. *Plant Physiol* **138,** 2269-2279.

Graeve, K., von Schaewen, A., Scheibe, R. (1994) Purification, characterization, and cDNA sequence of glucose-6-phosphate dehydrogenase from potato (*Solanum tuberosum* L.). *Plant J.* **5,** 353-361.

Green, D.E., Keilin, D. (1936) Biological oxidations and reductions. *Annu. Rev. Biochem.* **5,** 1-24.

Gunstone, F. (2000) Soybeans pace boost in oilseed production. *Inform* **11,** 1287-1289.

Hauschild, R., von Schaewen, A. (2003) Differential regulation of glucose-6-phosphate dehydrogenase isoenzyme activities in potato. *Plant Physiol.* **133,** 47-62.

Heineke, D., Riens, B., Grosse, H., Hoferichter, P., Peter, U., Flugge, U.I., Heldt, H.W. (1991) Redox transfer across the inner chloroplast envelope membrane. *Plant Physiol.* **95,** 1131-1137.

Henkes, S., Sonnewald, U., Badur, R., Flachmann, R., Stitt, M. (2001) A small decrease of plastid transketolase activity in antisense tobacco transformants has dramatic effects on photosynthesis and phenylpropanoid metabolism. *Plant Cell* **13,** 535-551.

Hirose, T., Imaizumi, N., Scofield, G.N., Furbank, R.T., Ohsugi, R. (1997) cDNA cloning and tissue specific expression of a gene for sucrose transporter from rice (*Oryza sativa* L.). *Plant Cell Physiol* **38,** 1389-1396.

Hobbs, D.H., Flintham, J.E., Hills, M.J. (2004) Genetic control of storage oil synthesis in seeds of Arabidopsis. *Plant Physiol* **136,** 3341-3349.

Hoek, J.B., Rydstrom, J. (1988) Physiological roles of nicotinamide nucleotide transhydrogenase. *Biochem. J.* **254,** 1-10.

Hooks, M.A., Fleming, Y., Larson, T.R., Graham, I.A. (1999) No induction of β-oxidation in leaves of Arabidopsis that over-produce lauric acid. *Planta* **207,** 385-392.

Hunter, S.C., Ohlrogge, J.B. (1998) Regulation of spinach chloroplast acetyl-CoA carboxylase. *Arch. Biochem. Biophys* **359,** 170-178.

Huppe, H.C., Farr, T.J., Turpin, D.H. (1994) Coordination of chloroplastic metabolism in N-limited *Chlamydomonas reinhardtii* by redox modulation (II. Redox modulation activates the oxidative pentose phosphate pathway during photosynthetic nitrate assimilation). *Plant Physiol* **105,** 1043-1048.

Huppe, H.C., Turpin, D.H. (1996) Appearance of novel glucose-6-phosphate dehydrogenase isoforms in *Chlamydomonas reinhardtii* during growth on nitrate. *Plant Physiol* **110,** 1431-1433.

Hutchings, D., Rawsthorne, S., Emes, M.J. (2005) Fatty acid synthesis and the oxidative pentose phosphate pathway in developing embryos of oilseed rape (*Brassica napus* L.). *J. Exp. Bot.* **56,** 577-585.

Jako, C., Kumar, A., Wei, Y., Zou, J., Barton, D.L., Giblin, E.M., Covello, P.S., Taylor, D.C. (2001) Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. *Plant Physiol* **126,** 861-874.

Jaworski J (1987) Biosynthesis of monoenoic and polyenoic fatty acids. In *The biochemistry of plants Volume 9 Lipids: Structure and function* (Stumpf, P.K., ed). Orlando: Academic press, pp 159-174.

Jaworski, J., Cahoon, E.B. (2003) Industrial oils from transgenic plants. *Curr. Opin. Plant Biol.* **6,** 178-184.

Jin, T., Huppe, H.C., Turpin, D.H. (1998) In vitro reconstitution of electron transport from glucose-6-phosphate and NADPH to nitrite. *Plant Physiol* **117,** 303-309.

Johnson, H.S. (1972) Dithiothreitol - inhibitor of glucose-6-phosphate-dehydrogenase activity in leaf extracts and isolated chloroplasts. *Planta* **106,** 273-277.

Kammerer, B., Fischer, K., Hilpert, B., Schubert, S., Gutensohn, M., Weber, A., Flugge, U.I. (1998) Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. *Plant Cell* **10,** 105-117.

Kang, F., Rawsthorne, S. (1994) Starch and fatty acid synthesis in plastids from developing embryos of oilseed rape (*Brassica napus* L.). *Plant J.* **6,** 795-805.

Kang, F., Rawsthorne, S. (1996) Metabolism of glucose-6-phosphate and utilization of multiple metabolites for fatty acid synthesis by plastids from developing oilseed rape embryos. *Planta* **199,** 321-327.

Katavic, V., Reed, D.W., Taylor, D.C., Giblin, E.M., Barton, D.L., Zou, J., Mackenzie, S.L., Covello, P.S., Kunst, L. (1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in Arabidopsis thaliana affecting diacylglycerol acyltransferase activity. *Plant Physiol* **108,** 399-409.

Kaup, M.T., Froese, C.D., Thompson, J.E. (2002) A role for diacylglycerol acyltransferase during leaf senescence. *Plant Physiol* **129,** 1616-1626.

Ke, J., Behal, R.H., Back, S.L., Nikolau, B.J., Wurtele, E.S., Oliver, D.J. (2000a) The role of pyruvate dehydrogenase and acetyl-coenzyme A synthetase in fatty acid synthesis in developing Arabidopsis seeds. *Plant Physiol* **123,** 497-508.

Ke, J., Wen, T.N., Nikolau, B.J., Wurtele, E.S. (2000b) Coordinate regulation of the nuclear and plastidic genes coding for the subunits of the heteromeric acetyl-coenzyme A carboxylase. *Plant Physiol* **122,** 1057-1071.

King, S.P., Badger, M.R., Furbank, R.T. (1998) CO_2 refixation characteristics of developing canola seeds and silique wall. *Aust. J. Plant Physiol.* **25,** 377-386.

Klaus, D., Ohlrogge, J.B., Neuhaus, H.E., Dormann, P. (2004) Increased fatty acid production in potato by engineering of acetyl-CoA carboxylase. *Planta* **219,** 389-396.

Kletzien, R.F., Harris, P.K., Foellmi, L.A. (1994) Glucose-6-phosphate dehydrogenase: a "housekeeping" enzyme subject to tissue-specific regulation by hormones, nutrients, and oxidant stress. *FASEB J.* **8,** 174-181.

Knappe, S., Lottgert, T., Schneider, A., Voll, L., Flugge, U.I., Fischer, K. (2003) Characterization of two functional *phosphoenolpyruvate/phosphate translocator* (*PPT*) genes in *Arabidopsis--AtPPT1* may be involved in the provision of signals for correct mesophyll development. *Plant J.* **36,** 411-420.

Konishi, T., Sasaki, Y. (1994) Compartmentalization of two forms of acetyl-CoA carboxylase in plants and the origin of their tolerance toward herbicides. *Proc. Natl. Acad. Sci. U. S. A* **91,** 3598-3601.

Konishi, T., Shinohara, K., Yamada, K., Sasaki, Y. (1996) Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant Cell Physiol* **37,** 117-122.

Koornneef M, Karssen CM (1994) Seed Dormancy and Germination. In *Arabidopsis* (Meyerowitz, E.M., Somerville, C.R., eds). New York: Cold Spring Harbor Laboratory Press, pp 313-334.

Krepinsky, K., Plaumann, M., Martin, W., Schnarrenberger, C. (2001) Purification and cloning of chloroplast 6-phosphogluconate dehydrogenase from spinach - Cyanobacterial genes for chloroplast and cytosolic isoenzymes encoded in eukaryotic chromosomes. *Eur. J. Biochem.* **268,** 2678-2686.

Kruger, N.J., von Schaewen, A. (2003) The oxidative pentose phosphate pathway: structure and organisation. *Curr. Opin. Plant Biol.* **6,** 236-246.

Kubis, S.E., Rawsthorne, S. (2000) The role of plastidial transporters in developing embryos of oilseed rape (*Brassica napus* L.) for fatty acid synthesis. *Biochem. Soc. Trans.* **28,** 665-666.

Lacey, D.J., Hills, M.J. (1996) Heterogeneity of the endoplasmic reticulum with respect to lipid synthesis in developing seeds of *Brassica napus* L. *Planta* **199,** 545-551.

Lardans, A., Tremolieres, A. (1992) Fatty-acid elongation activities in subcellular-fractions of developing seeds of *Limnanthes alba*. *Phytochemistry* **31,** 121-127.

Larson, T.R., Edgell, T., Byrne, J., Dehesh, K., Graham, I.A. (2002) Acyl CoA profiles of transgenic plants that accumulate medium-chain fatty acids indicate inefficient storage lipid synthesis in developing oilseeds. *Plant J.* **32,** 519-527.

Lenartowicz, E. (1990) A complex effect of arsenite on the formation of alpha-ketoglutarate in rat liver mitochondria. *Arch. Biochem. Biophys* **283,** 388-396.

Lendzian, K.J. (1980) Modulation of glucose-6-phosphate-dehydrogenase by NADPH, NADP+ and dithiothreitol at variable NADPH-$NADP^+$ ratios in an illuminated reconstituted spinach (*Spinacia oleracea* L) chloroplast system. *Planta* **148,** 1-6.

Leonard, J.M., Knapp, S.J., Slabaugh, M.B. (1998) A *Cuphea* β-ketoacyl-ACP synthase shifts the synthesis of fatty acids towards shorter chains in Arabidopsis seeds expressing *Cuphea* FatB thioesterases. *Plant J.* **13,** 621-628.

Lin, Y., Cluette-Brown, J.E., Goodman, H.M. (2004) The peroxisome deficient Arabidopsis mutant *sse1* exhibits impaired fatty acid synthesis. *Plant Physiol* **135,** 814-827.

Lin, Y., Sun, L., Nguyen, L.V., Rachubinski, R.A., Goodman, H.M. (1999) The Pex16p homolog SSE1 and storage organelle formation in Arabidopsis seeds. *Science* **284,** 328-330.

Madoka, Y., Tomizawa, K., Mizoi, J., Nishida, I., Nagano, Y., Sasaki, Y. (2002) Chloroplast transformation with modified *accD* operon increases acetyl-CoA carboxylase and causes extension of leaf longevity and increase in seed yield in tobacco. *Plant Cell Physiol* **43,** 1518-1525.

Mansfield, S.G., Briarty, L.G. (1992) Cotyledon cell development in *Arabidopsis thaliana* during reserve deposition. *Can. J. Bot.* **70,** 151-164.

Marillia, E.F., Micallef, B.J., Micallef, M., Weninger, A., Pedersen, K.K., Zou, J., Taylor, D.C. (2003) Biochemical and physiological studies of Arabidopsis thaliana transgenic lines with repressed expression of the mitochondrial pyruvate dehydrogenase kinase. *J. Exp. Bot.* **54,** 259-270.

Martin, B.A., Wilson, R.F. (1984) Subcellular localization of triacylglycerol synthesis in spinach leaves. *Lipids* **19,** 117-121.

Martini, G., Ursini, M.V. (1996) A new lease of life for an old enzyme. *Bioessays* **18,** 631-637.

Meinke DW (1994) Seed Development in *Arabidopsis thaliana.* In *Arabidopsis* (Meyerowitz, E.M., Somerville, C.R., eds). New York: Cold Spring Harbor Laboratory Press, pp 253-295.

Mietkiewska, E., Giblin, E.M., Wang, S., Barton, D.L., Dirpaul, J., Brost, J.M., Katavic, V., Taylor, D.C. (2004) Seed-specific heterologous expression of a nasturtium *FAE* gene in Arabidopsis results in a dramatic increase in the proportion of erucic acid. *Plant Physiol* **136,** 2665-2675.

Millar, A.A., Smith, M.A., Kunst, L. (2000) All fatty acids are not equal: discrimination in plant membrane lipids. *Trends Plant Sci.* **5,** 95-101.

Millar, A.A., Wischer, M., Kunst, L. (1998) Accumulation of very-long-chain fatty acids in membrane glycerolipids is associated with dramatic alterations in plant morphology. *Plant Cell* **10,** 1889-1902.

Moire, L., Rezzonico, E., Goepfert, S., Poirier, Y. (2004) Impact of unusual fatty acid synthesis on futile cycling through β-oxidation and on gene expression in transgenic plants. *Plant Physiol* **134,** 432-442.

Moller, I.M. (2001) Plant mitochondria and oxidative stress: Electron transport, NADPH turnover, and metabolism of reactive oxygen species. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **52,** 561-591.

Moller, I.M., Rasmusson, A.G. (1998) The role of NADP in the mitochondrial matrix. *Trends Plant Sci.* **3,** 21-27.

Nemoto, Y., Sasakuma, T. (2000) Specific expression of glucose-6-phosphate dehydrogenase (G6PDH) gene by salt stress in wheat (*Triticum aestivum* L.). *Plant Sci.* **158,** 53-60.

Neuburger, M., Day, D.A., Douce, R. (1984) Transport of coenzyme A in plant mitochondria. *Arch. Biochem. Biophys* **229,** 253-258.

Neuhaus, H.E., Emes, M.J. (2000) Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **51,** 111-140.

Neuhaus, H.E., Wagner, R. (2000) Solute pores, ion channels, and metabolite transporters in the outer and inner envelope membranes of higher plant plastids. *Biochim Biophys Acta* **1465,** 307-323.

Niewiadomski, P., Knappe, S., Geimer, S., Fischer, K., Schulz, B., Unte, U.S., Rosso, M.G., Ache, P., Flugge, U.I., Schneider, A. (2005) The Arabidopsis plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. *Plant Cell* **17,** 760-775.

O'Neill, C.M., Gill, S., Hobbs, D., Morgan, C., Bancroft, I. (2003) Natural variation for seed oil composition in *Arabidopsis thaliana*. *Phytochemistry* **64,** 1077-1090.

Ohlrogge, J., Browse, J. (1995) Lipid biosynthesis. *Plant Cell* **7,** 957-970.

Ohlrogge, J.B., Browse, J., Somerville, C.R. (1991) The genetics of plant lipids. *Biochim Biophys Acta* **1082,** 1-26.

Ohlrogge, J.B., Jaworski, J.G. (1997) Regulation of fatty acid synthesis. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **48,** 109-136.

Ohlrogge, J.B., Kuhn, D.N., Stumpf, P.K. (1979) Subcellular localization of acyl carrier protein in leaf protoplasts of *Spinacia oleracea*. *Proc. Natl. Acad. Sci. U. S. A* **76,** 1194-1198.

Penfield, S., Rylott, E.L., Gilday, A.D., Graham, S., Larson, T.R., Graham, I.A. (2004) Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires *PHOSPHOENOLPYRUVATE CARBOXYKINASE1*. *Plant Cell* **16,** 2705-2718.

Plaxton, W.C. (1996) The organization and regulation of plant glycolysis. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **47,** 185-214.

Pleite, R., Pike, M.J., Garces, R., Martinez-Force, E., Rawsthorne, S. (2005) The sources of carbon and reducing power for fatty acid synthesis in the heterotrophic plastids of developing sunflower (*Helianthus annuus* L.) embryos. *J. Exp. Bot.* **56,** 1297-1303.

Poirier, Y., Ventre, G., Caldelari, D. (1999) Increased flow of fatty acids toward β-oxidation in developing seeds of Arabidopsis deficient in diacylglycerol acyltransferase activity or synthesizing medium-chain-length fatty acids. *Plant Physiol* **121,** 1359-1366.

Porterfield, D.M., Kuang, A., Smith, P.J., Crispi, M.L., Musgrave, M.E. (1999) Oxygen-depleted zones inside reproductive structures of *Brassicaceae*: implications for oxygen control of seed development. *Can. J. Bot.* **77,** 1439-1446.

Redinbaugh, M.G., Campbell, W.H. (1998) Nitrate regulation of the oxidative pentose phosphate pathway in maize (*Zea mays* L.) root plastids: induction of 6-phosphogluconate dehydrogenase activity, protein and transcript levels. *Plant Sci.* **134,** 129-140.

Regierer, B., Fernie, A.R., Springer, F., Perez-Melis, A., Leisse, A., Koehl, K., Willmitzer, L., Geigenberger, P., Kossmann, J. (2002) Starch content and yield increase as a result of altering adenylate pools in transgenic plants. *Nat. Biotechnol.* **20,** 1256-1260.

Reiser, J., Linka, N., Lemke, L., Jeblick, W., Neuhaus, H.E. (2004) Molecular physiological analysis of the two plastidic ATP/ADP transporters from Arabidopsis. *Plant Physiol* **136,** 3524-3536.

Rolletschek, H., Weber, H., Borisjuk, L. (2003) Energy status and its control on embryogenesis of legumes. Embryo photosynthesis contributes to oxygen supply and is coupled to biosynthetic fluxes. *Plant Physiol* **132,** 1196-1206.

Roughan, P.G., Holland, R., Slack, C.R. (1979) On the control of long-chain-fatty acid synthesis in isolated intact spinach (*Spinacia oleracea*) chloroplasts. *Biochem. J.* **184,** 193-202.

Routaboul, J.M., Benning, C., Bechtold, N., Caboche, M., Lepiniec, L. (1999) The TAG1 locus of Arabidopsis encodes for a diacylglycerol acyltransferase. *Plant Physiol. Biochem.* **37,** 831-840.

Ruuska, S.A., Girke, T., Benning, C., Ohlrogge, J.B. (2002) Contrapuntal networks of gene expression during Arabidopsis seed filling. *Plant Cell* **14,** 1191-1206.

Ruuska, S.A., Schwender, J., Ohlrogge, J.B. (2004) The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiol* **136,** 2700-2709.

Salati, L.M., Szeszel-Fedorowicz, W., Tao, H., Gibson, M.A., Amir-Ahmady, B., Stabile, L.P., Hodge, D.L. (2004) Nutritional regulation of mRNA processing. *J. Nutr.* **134,** 2437S-2443S.

Sasaki, Y., Kozaki, A., Hatano, M. (1997) Link between light and fatty acid synthesis: thioredoxin-linked reductive activation of plastidic acetyl-CoA carboxylase. *Proc. Natl. Acad. Sci. U. S. A* **94,** 11096-11101.

Scheibe, R. (1991) Redox-modulation of chloroplast enzymes. A common principle for individual control. *Plant Physiol.* **96,** 1-3.

Scheibe, R. (2004) Malate valves to balance cellular energy supply. *Physiol. Plant.* **120,** 21-26.

Scheibe, R., Anderson, L.E. (1981) Dark modulation of NADP-dependent malate-dehydrogenase and glucose-6-phosphate-dehydrogenase in the chloroplast. *Biochim. Biophys. Acta* **636,** 58-64.

Schnarrenberger, C., Flechner, A., Martin, W. (1995) Enzymatic evidence for a complete oxidative pentose-phosphate pathway in chloroplasts and an incomplete pathway in the cytosol of spinach leaves. *Plant Physiol.* **108,** 609-614.

Schnarrenberger, C., Oeser, A., Tolbert, N.E. (1973) 2 Isoenzymes each of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in spinach leaves. *Arch. Biochem. Biophys.* **154,** 438-448.

Schneider, A., Hausler, R.E., Kolukisaoglu, U., Kunze, R., van der, G.E., Schwacke, R., Catoni, E., Desimone, M., Flugge, U.I. (2002) An *Arabidopsis thaliana* knock-out mutant of the chloroplast triose phosphate/phosphate translocator is severely compromised only when starch synthesis, but not starch mobilisation is abolished. *Plant J.* **32,** 685-699.

Schwender, J., Goffman, F., Ohlrogge, J.B., Shachar-Hill, Y. (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* **432,** 779-782.

Schwender, J., Ohlrogge, J.B., Shachar-Hill, Y. (2003) A flux model of glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos. *J. Biol. Chem.* **278,** 29442-29453.

Shearer, H.L., Turpin, D.H., Dennis, D.T. (2004) Characterization of NADP-dependent malic enzyme from developing castor oil seed endosperm. *Arch. Biochem. Biophys* **429,** 134-144.

Shimakata, T., Stumpf, P.K. (1982) Purification and characterizations of β-ketoacyl-[acyl-carrier-protein] reductase, β-hydroxyacyl-[acylcarrier-protein] dehydrase, and enoyl-[acyl-carrier-protein] reductase from *Spinacia oleracea* leaves. *Arch. Biochem. Biophys* **218,** 77-91.

Shintani, D., Roesler, K., Shorrosh, B., Savage, L., Ohlrogge, J. (1997) Antisense expression and overexpression of biotin carboxylase in tobacco leaves. *Plant Physiol* **114,** 881-886.

Sindelar, L., Sindelarova, M. (2002) Correlation of viral RNA biosynthesis with glucose-6-phosphate dehydrogenase activity and host resistance. *Planta* **215,** 862-869.

Slabas, A.R., Fawcett, T. (1992) The biochemistry and molecular biology of plant lipid biosynthesis. *Plant Mol. Biol.* **19,** 169-191.

Slabas, A.R., Sidebottom, C., Kessell, R., Hellyer, A., Tombs, M.P. (1986) Oilseed rape NADH enoyl acyl-carrier protein reductase. *Biochem. Soc. Trans.* **14,** 581-582.

Smith, R.G., Gauthier, D.A., Dennis, D.T., Turpin, D.H. (1992) Malate- and pyruvate-dependent fatty acid synthesis in leucoplasts from developing castor endosperm. *Plant Physiol.* **98,** 1233-1238.

Soeda, Y., Konings, M.C., Vorst, O., van Houwelingen, A.M., Stoopen, G.M., Maliepaard, C.A., Kodde, J., Bino, R.J., Groot, S.P., van der Geest, A.H. (2005) Gene expression programs during Brassica oleracea seed maturation, osmopriming, and germination are indicators of progression of the germination process and the stress tolerance level. *Plant Physiol* **137,** 354-368.

Song, W., Koh, S., Czako, M., Marton, L., Drenkard, E., Becker, J.M., Stacey, G. (1997) Antisense expression of the peptide transport gene *AtPTR2-B* delays flowering and arrests seed development in transgenic Arabidopsis plants. *Plant Physiol* **114,** 927-935.

Sonnewald, U., Hajirezaei, M.R., Kossmann, J., Heyer, A., Trethewey, R.N., Willmitzer, L. (1997) Increased potato tuber size resulting from apoplastic expression of a yeast invertase. *Nat. Biotechnol.* **15,** 794-797.

Sreenivasulu, N., Altschmied, L., Panitz, R., Hahnel, U., Michalek, W., Weschke, W., Wobus, U. (2002) Identification of genes specifically expressed in maternal and filial tissues of barley caryopses: a cDNA array analysis. *Mol. Genet Genomics* **266,** 758-767.

Stacey, M.G., Koh, S., Becker, J., Stacey, G. (2002) AtOPT3, a member of the oligopeptide transporter family, is essential for embryo development in Arabidopsis. *Plant Cell* **14,** 2799-2811.

Stahl, U., Carlsson, A.S., Lenman, M., Dahlqvist, A., Huang, B., Banas, W., Banas, A., Stymne, S. (2004) Cloning and functional characterization of a

phospholipid:diacylglycerol acyltransferase from Arabidopsis. *Plant Physiol* **135,** 1324-1335.

Streatfield, S.J., Weber, A., Kinsman, E.A., Hausler, R.E., Li, J., Post-Beittenmiller, D., Kaiser, W.M., Pyke, K.A., Flugge, U.I., Chory, J. (1999) The phosphoenolpyruvate/phosphate translocator is required for phenolic metabolism, palisade cell development, and plastid-dependent nuclear gene expression. *Plant Cell* **11,** 1609-1622.

Suh, M.C., Kim, M.J., Hur, C.G., Bae, J.M., Park, Y.I., Chung, C.H., Kang, C.W., Ohlrogge, J.B. (2003) Comparative analysis of expressed sequence tags from *Sesamum indicum* and *Arabidopsis thaliana* developing seeds. *Plant Mol. Biol.* **52,** 1107-1123.

Taniguchi, M., Taniguchi, Y., Kawasaki, M., Takeda, S., Kato, T., Sato, S., Tabata, S., Miyake, H., Sugiyama, T. (2002) Identifying and characterizing plastidic 2-oxoglutarate/malate and dicarboxylate transporters in *Arabidopsis thaliana. Plant Cell Physiol* **43,** 706-717.

Tao, H., Szeszel-Fedorowicz, W., Amir-Ahmady, B., Gibson, M.A., Stabile, L.P., Salati, L.M. (2002) Inhibition of the splicing of glucose-6-phosphate dehydrogenase precursor mRNA by polyunsaturated fatty acids. *J. Biol. Chem.* **277,** 31270-31278.

Thom, E., Mohlmann, T., Quick, W.P., Camara, B., Neuhaus, H.E. (1998) Sweet pepper plastids: enzymic equipment, characterisation of the plastidic oxidative pentose-phosphate pathway, and transport of phosphorylated intermediates across the envelope membrane. *Planta* **204,** 226-233.

Thomaeus, S., Carlsson, A.S., Stymne, S. (2001) Distribution of fatty acids in polar and neutral lipids during seed development in *Arabidopsis thaliana* genetically engineered to produce acetylenic, epoxy and hydroxy fatty acid. *Plant Sci.* **161,** 997-1003.

Tjaden, J., Mohlmann, T., Kampfenkel, K., Henrichs, G., Neuhaus, H.E. (1998) Altered plastidic ATP/ADP-transporter activity influences potato (*Solanum tuberosum* L.) tuber morphology, yield and composition of tuber starch. *Plant J.* **16,** 531-540.

Trethewey, R.N., Geigenberger, P., Riedel, K., Hajirezaei, M.R., Sonnewald, U., Stitt, M., Riesmeier, J.W., Willmitzer, L. (1998) Combined expression of glucokinase and invertase in potato tubers leads to a dramatic reduction in starch accumulation and stimulation of glycolysis. *Plant J.* **15,** 109-118.

Trethewey, R.N., Riesmeier, J.W., Willmitzer, L., Stitt, M., Geigenberger, P. (1999) Tuber-specific expression of a yeast invertase and a bacterial glucokinase in potato leads to an activation of sucrose phosphate synthase and the creation of a sucrose futile cycle. *Planta* **208,** 227-238.

Van De Loo, F.J., Turner, S., Somerville, C. (1995) Expressed sequence tags from developing castor seeds. *Plant Physiol* **108,** 1141-1150.

Vigeolas, H., van Dongen, J.T., Waldeck, P., Huhn, D., Geigenberger, P. (2003) Lipid storage metabolism is limited by the prevailing low oxygen concentrations within developing seeds of oilseed rape. *Plant Physiol* **133,** 2048-2060.

Voelker, T., Kinney, A.J. (2001) Variations in the biosynthesis of seed-storage lipids. *Annu. Rev. Plant Physiol Plant Mol. Biol.* **52,** 335-361.

Voelker, T.A., Hayes, T.R., Cranmer, A.M., Turner, J.C., Davies, H.M. (1996) Genetic engineering of a quantitive trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. *Plant J.* **9,** 229-241.

Voelker, T.A., Worrell, A.C., Anderson, L., Bleibaum, J., Fan, C., Hawkins, D.J., Radke, S.E., Davies, H.M. (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. *Science* **257,** 72-74.

Vogel, G., Browse, J. (1996) Cholinephosphotransferase and diacylglycerol acyltransferase (Substrate specificities at a key branch point in seed lipid metabolism). *Plant Physiol* **110,** 923-931.

von Schaewen, A., Langenkamper, G.S.R., Graeve, K., Wenderoth, I., Scheibe, R. (1995) Molecular characterization of the plastidic glucose-6-phosphate dehydrogenase from potato in comparison to its cytosolic counterpart. *Plant Physiol.* **109,** 1327-1335.

Vulliamy, T., Mason, P., Luzzatto, L. (1992) The molecular basis of glucose-6-phosphate dehydrogenase deficiency. *Trends Genet* **8,** 138-143.

Wang, R., Okamoto, M., Xing, X., Crawford, N.M. (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiol* **132,** 556-567.

Wenderoth, I., Scheibe, R., vonSchaewen, A. (1997) Identification of the cysteine residues involved in redox modification of plant plastidic glucose-6-phosphate dehydrogenase. *J. Biol. Chem.* **272,** 26985-26990.

Wendt, U.K., Hauschild, R., Lange, C., Pietersma, M., Wenderoth, I., von Schaewen, A. (1999) Evidence for functional convergence of redox regulation in G6PDH isoforms of cyanobacteria and higher plants. *Plant Mol. Biol.* **40,** 487-494.

Wendt, U.K., Wenderoth, I., Tegeler, A., von Schaewen, A. (2000) Molecular characterization of a novel glucose-6-phosphate dehydrogenase from potato (*Solanum tuberosum* L.). *Plant J.* **23,** 723-733.

White, J.A., Benning, C. (2001) Genomic approaches toward the engineering of oil seeds. *Plant Physiol. Biochem.* **39,** 263-270.

White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., de Ilarduya, O.M., Jaworski, J.G., Ohlrogge, J.B., Benning, C. (2000) A new set of Arabidopsis expressed

sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. *Plant Physiol* **124,** 1582-1594.

Wiberg, E., Tillberg, E., Stymne, S. (1994) Substrates of diacylglyerol acyltransferase in microsomes from developing oil seeds. *Phytochemistry* **36,** 573-577.

Williams, J.F. (1980) A critical-examination of the evidence for the reactions of the pentose pathway in animal-tissues. *Trends Biochem. Sci.* **5,** 315-320.

Williams, M., Randall, D.D. (1979) Pyruvate dehydrogenase complex from chloroplasts of *Pisum sativum* L. *Plant Physiol* **64,** 1099-1103.

Wright, D.P., Huppe, H.C., Turpin, D.H. (1997) In vivo and in vitro studies of glucose-6-phosphate dehydrogenase from barley root plastids in relation to reductant supply for NO_2^- assimilation. *Plant Physiol.* **114,** 1413-1419.

Zheng, Z., Xia, Q., Dauk, M., Shen, W., Selvaraj, G., Zou, J. (2003) Arabidopsis *AtGPAT1*, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. *Plant Cell* **15,** 1872-1887.

Zou, J., Katavic, V., Giblin, E.M., Barton, D.L., Mackenzie, S.L., Keller, W.A., Hu, X., Taylor, D.C. (1997) Modification of seed oil content and acyl composition in the brassicaceae by expression of a yeast *sn-2* acyltransferase gene. *Plant Cell* **9,** 909-923.

Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G., Taylor, D.C. (1999) The *Arabidopsis thaliana TAG1* mutant has a mutation in a diacylglycerol acyltransferase gene. *Plant J.* **19,** 645-653.

Chapter 2

Genome-wide characterization of glucose-6-phosphate dehydrogenases in Arabidopsis[1]

[1] This chapter describes the work published in Wakao and Benning (2005) Genome-wide analysis of glucose-6-phosphate dehydrogenases in Arabidopsis. Plant Journal 41: 243-56 with some modifications.

2.1 Introduction

G6PDH has been studied in various organisms and with various interests. As discussed in Chapter 1, current studies on plant G6PDHs focus on the identification of its roles *in vivo* such as nitrogen assimilation. An interesting aspect of plant G6PDHs is the presence of multiple isoforms with distinct biochemical characteristics, tissue-specificity, and subcellular localization, which may reflect their specific *in vivo* roles. The discovery of the two groups of plastidic G6PDH isoforms, P1 and P2 with distinct characteristics, has brought up speculations and predictions for their specific roles *in vivo* (Chapter 1). It was expected that a comprehensive and comparative study facilitated by the genome sequence would contribute in correlating the isoform characteristics to their potential *in vivo* roles. The main purpose of this study was to obtain comparative information on the various G6PDH isoforms from a single plant, which would later serve as a foundation to the functional analysis of the individual isoforms using resources in Arabidopsis such as traditional mutants and T-DNA insertion lines. In this study, the G6PDH isoforms were examined for their *in vitro* biochemical characteristics, and their patterns of gene expression and activity in various tissues. These results as well as genomic and bioinformatic data were integrated to propose *in vivo* roles for individual G6PDH isoforms in Arabidopsis.

2.2 Material and methods

2.2.1 Bioinformatics

All Arabidopsis and rice genes annotated as G6PDH as well as their coding sequences and all reported expressed sequence tags (ESTs) encoding Arabidopsis G6PDHs were

searched through The Institute for Genomic Research website (http://www.tigr.org/tdb/e2k1/ath1/GeneNameSearch.shtml). Subcellular localization of proteins was predicted using ChloroP (Emanuelsson *et al.*, 1999) and TargetP (Emanuelsson *et al.*, 2000). Transit peptides were removed from the proteins predicted to be plastidic and the similarities of the mature protein sequences were determined by ClustalW (Li, 2003). The phylogenetic tree was assembled with the aid of Biology Workbench (San Diego Supercomputer Center, University of California, San Diego; http://workbench.sdsc.edu/) and Phylip DrawTree (Felsenstein, 1989). Active consensus sites were identified based on the PROSITE database (Hofmann *et al.*, 1999) using MOTIF (http://www.motif.genomenet.ad.jp).

2.2.2 Expression vector constructs and cloning

For tissues other than siliques, total RNA was extracted with RNeasy (QIAGEN, Valencia, CA). Silique RNA was extracted as previously described (Yu *et al.*, 2004). Following DNAse I treatment (Roche, Indianapolis, IN), cDNA was synthesized from 300 ng of RNA using a reverse transcriptase (RT) kit (Omniscript, QIAGEN, Valencia, CA). For cloning into expression vectors, regions coding for the mature G6PDH isoforms were amplified by RT-PCR from RNA of siliques 5-8 days after flowering (DAF). From the RT reaction, 5% was used as template for subsequent PCR reactions (Hotstartaq, QIAGEN, Valencia, CA) with 30 cycles. Primers as listed in Table 2.1 were designed to produce the mature proteins. The GenBank/EMBL accession numbers for the nucleotide sequences used for primer design of Arabidopsis G6PDHs are as follows: *AtG6PD1,* AY099561; *G6PD2,* AY065042; *AtG6PD3,* AY139768; *AtG6PD4,* AY056232;

AtG6PD5, AY065054; *AtG6PD6,* AJ010971. The PCR products for all genes except for *AtG6PD6* were sub-cloned into pCR2.1-TOPO (Invitrogen, Carlsbad, CA), then cloned into the expression vector pASK-IBA3 (IBA, St. Louis, MO) using either *Kpn*I or *Eco*RI sites. The orientation of the genes was confirmed by sequencing. The expression of the gene was controlled by a tetracyclin promoter, and the recombinant protein produced included an eight amino acid C-terminal tag that allowed affinity chromatography on a *Strep*-Tactin column (IBA, St. Louis, MO) and additional amino acids depending on the construct.

2.2.3 Recombinant protein expression

Isoforms G6PD1, G6PD2, and G6PD6 were produced in *E. coli* strain XL-1 Blue, and isoforms G6PD3, G6PD4, and G6PD5 in BL21 (DE3) to obtain higher protein yields. The expression strains were routinely cultured overnight at 37 °C in Luria-Bertani media (LB) ampicillin 100 μg/ml, inoculated at 1/250 dilution in LB ampicillin 100 μg/ml and cultured at 28 °C. Gene expression was induced by addition of 0.2 mg/l-LB anhydrotetracyclin at OD_{600}~0.3. At OD_{600} 0.7~0.8, cultures were cooled on ice and cells were harvested by centrifugation at 6000*g*. Cells from 250 ml of culture were resuspended in 1 ml of Wash buffer (0.1 M Tris pH8, 1 mM EDTA, 150 mM NaCl) and lysed by sonication. The protein was purified from cell lysate using *Strep*-Tactin columns (IBA, St. Louis, MO) following the manufacturer's instructions. Fractions with highest protein content were combined and the protein concentration was determined by Bio-Rad DC Protein Assay (Biorad, Hercules, CA).

2.2.4 Enzyme assays and kinetic analyses

All assays were performed in double beam mode using a DW-2RC UV-VIS spectrophotometer (SLM Instruments, Rochester, NY) detecting OD at 340 nm. A sample containing the enzyme was measured against a blank containing 0.1 M Tris (pH 8) with substrates or various dithiothreitol (DTT) and sodium tetrathionate (NaTT) concentrations as indicated in the text. The increase in absorbance at 340 nm was monitored, which is diagnostic for the generation of NADPH. Saturation curves were obtained by measuring activity under varying concentrations of $NADP^+$ and excess amount of glucose-6-phosphate (G6P), at least 10-fold that of saturating $NADP^+$. To determine K_{iNADPH}, saturation curves were obtained under three non-saturating concentrations of the inhibitor NADPH. The values of K_m, apparent K_m (K_m'), and V_{max} were deduced from best fitting hyperbolic curves using ORIGIN 5.0 (OriginLab Corporation, Northampton, MA). The K_m' values were plotted against NADPH concentration [I] for each curve. The K_i values were deduced from these plots by extrapolating to K_m'=0, in which K_i =-[I] based on the equation $K_m' = K_m (1+[I]/ K_i)$ (Segel, 1975). Each experiment was repeated two to three times. Buffers of 0.1 M 4-morpholineethanesulfonic acid (MES) for pH 6-7.5, 0.1 M Tris for pH 7-9.5, 0.1 M 3-(cycolohexylamino) propanesulfonic acid (CAPS) for pH 9-10 were used for determining the optimal pH under constant substrate concentrations. To study the effect of reduction by DTT, the enzyme was incubated with G6P and 0, 3, 5, 10 mM DTT at time zero. The reaction was initiated by adding constant non-saturating concentrations of $NADP^+$ to the

mixture and OD_{340} was measured at each time point. To examine its reversibility, 5 mM NaTT was added to reactions containing 3 mM DTT and the reaction was measured in a time-dependent manner. For DTT and NaTT, a 100 mM solution was freshly prepared in water for each experiment.

2.2.5 Semi-quantitative RT-PCR for gene expression analysis

Primers were designed to amplify the 3'-UTR of individual isoforms to distinguish the transcripts (Table 2.1). Semi-quantitative RT-PCR was performed using 3% of the RT reaction as template, with 1 U Taq polymerase (Roche, Indianapolis, IN) in a total volume of 15 µl with 28 cycles. Primer specificity and efficiency were confirmed by amplifying genomic fragments with serial dilution of genomic DNA. All RT-PCRs were repeated two to three times with similar relative levels detected. The *ACTIN1* gene (Genbank accession AY074873) was used as loading control. The relative densities of the bands detected for the transcript of an isoform in various tissues were calculated from the gel image with Molecular Analyst (BioRad, Hercules, CA), and normalized to the relative intensity calculated for the actin band from the respective tissue. All tissues were harvested from Col-2 plants grown in the same growth chamber for extracting RNA, except for roots, which were harvested from plantlets grown on agar plates.

2.2.6 Protein preparation and zymograms

Protein extraction from various tissues was carried out as previously described (Focks and Benning, 1998). Activity of G6PDH in protein extracts decreased during storage at – 80 °C. Therefore, all data were obtained with freshly prepared protein extracts. For

analyses of activity in different tissues in Figure 2.6a, protein samples were concentrated using an ultrafree-0.5 centrifugal filter device with a nominal MW limit of 50 kD (Millipore, Bedford, MA) and adjusted to approximately 6.3 mg/ml. All other experiments were carried out with protein extracts adjusted to 1-2 mg/ml if not stated otherwise. A recombinant G6PDH from *Leuconostoc mesenteroides* (Sigma, St. Louis, MO) was used as a positive control. Approximately 0.5 µl to 1 µl was applied onto Titan III cellulose acetate (CAE) plates (Helena Laboratories, Beaumont, TX) and electrophoresed for 20 min at 175 mV at 4 °C in electrophoresis buffer (180 mM Tris base, 100 mM boric acid, 20 mM dibasic sodium phosphate, 2 mM EDTA, 10 mM glucose). The CAE plates were overlaid with 5 m of dissolved (~40 °C) reaction medium (0.1 M Tris pH 9.2, 1 mM $NADP^+$, 10 mM G6P, 1% agar, 0.02 mg/ml phenazine methasulfate, 0.26 mg/ml nitrotetrazolium blue chloride) and incubated at 37°C until the color developed.

2.2.7 T-DNA insertion line isolation

T-DNA insertion lines were obtained from the SALK T-DNA insertion population (Alonso *et al.*, 2003). Genotypes of the plants were determined using PCR primers designed for gene-specific sequences flanking the insertion site and T-DNA left border using the i-sect tool (http://signal.salk.edu/tdnaprimers.html) with some modification to increase primer specificity. The homozygosity of the T-DNA insertion was indicated by the presence of a product of one gene-specific primer and the T-DNA left border primer, and the absence of a product of the two gene-specific primers. The insertion site was confirmed by sequencing the PCR product generated with the T-DNA left border primer

Gene		Primer
AtG6PD1	p(+)	5'-GGGGTACCATTCTTCGCCGAGAAACATTC-3'
	p(-)	5'-CGGAATTCAGCTTCTCCAAGATCTCCCC-3'
	e(+)	5'-AAAGTTCCTGGTCTTGGAAT-3'
	e(-)	5'-AATGCTCGATTTAGAACCAA-3'
AtG6PD2	p(+)	5'-CGGAATTCTTCCAGAGGAAGTCTGGTC-3'
	p(-)	5'-CCGGTACCCTGGTCAATACTGACGTCAC-3'
	e(+)	5'-CCCTGGTTTAGGAATGAGAT-3'
	e(-)	5'-TAAGAGAAACCCCTTTGGTT-3'
AtG6PD3	p(+)	5'-GGGGTACCTCGCAGAGGAGGTCCGTT-3'
	p(-)	5'-GGGGTACCCTGATCAAGACTTAGGTCTC-3'
	e(+)	5'-TCAAAGGAGATTCCAGATGC-3'
	e(-)	5'-CAAAAAGGGTGGCAAGAATA-3'
AtG6PD4	p(+)	5'-CGGAATTCCTCAACGGAGGAGGCTCGA-3'
	p(-)	5'-CGGAATTCGTCATCTGCCCATGGGACA-3'
	e(+)	5'-CCGGAGTTGTATGAATTTGG-3'
	e(-)	5'-TTCGTCATCTTCCTTCTCCA-3'
AtG6PD5	p(+)	5'-GGGGTACCATGGGTTCTGGTCAATGGCA-3'
	p(-)	5'-GGGGTACCCAATGTAGGAGGGATCCAAAT-3'
	e(+)	5'-GAAGTGAAATCGGTCCCATA-3'
	e(-)	5'-CCACAAAGCTGTTTTGCATT-3'
AtG6PD6	p(+)	5'-GGGGTACCATGGGATCTGGTCAATGGCACGTT-3'
	p(-)	5'-CCGGTACCTAGTGTAGGAGGGATCCAGATATAG-3'
	e(+)	5'-CAACCAGGTCTAGATATGAA-3'
	e(-)	5'-TGTCAAAGTGAACAGCTGAA-3'
ACTIN-1	e(+)	5'-AACAATCGATGGACCTGACTCG-3'
	e(-)	5'-TGCGACAATGGAACTGGAATGG-3'
g6pd5	t(+)	5'-GAAGTTTTTGGCTTTGCTGCG-3'
	t(-)	5'-GCAGTTATTCTGTGTGTATACGTCG-3'
g6pd6	t(+)	5'-CTAAACTCTCTTGCAACAGGTC-3'
	t(-)	5'-AAACACCAAACTCGAGGAGGGA-3'
T-DNA Lb[a]	t(-)	5'-GTTCACGTAGTGGGCCATCG-3'

[a] Modified from LBa1 primer suggested on SALK website (http://signal.salk.edu)

Table 2.1 Primer sequences
Primers used for cloning of recombinant protein production (p), gene expression analysis (e), and genotyping (t).

and the gene-specific primer. To confirm the loss of gene expression, RT-PCR was carried out similarly as described in previous sections except with 600 ng of initial total RNA and 35 cycles of PCR.

2.3 Results

2.3.1 Phylogenetic analysis

A previous phylogenetic analysis of G6PDH has been carried out by Wendt *et al.* (1999), which included three of the six Arabidopsis isoforms. The present analysis included plant protein sequences that were deposited to the database since then, and proteins translated from genes annotated as G6PDH found in genome sequences of Arabidopsis (Table 2.2) and rice (Figure 2.1). The Arabidopsis genes annotated as encoding G6PDHs are available in Genbank and were previously designated *AtG6PD1*-to-*AtG6PD6* (for accession numbers refer to Figure 2.1 legend). A search for additional G6PDH sequences using portions of the amino acid sequences of known G6PDHs did not yield new potentially G6PDH encoding genes for the Arabidopsis or rice genomes. The classifications of translated proteins in the phylogenetic tree (Figure 2.1) follow those designated by Wendt *et al.* (1999), plastidic (P1, P2), and cytosolic (Cy). The relations among the three already-described isoforms from Arabidopsis (G6PD1, G6PD3, and G6PD5) was in agreement with a previous report (Wendt *et al.*, 1999). Isoforms G6PD2 and G6PD6 were found to group with P2 and cytosolic isoforms respectively, each with highest similarity to another Arabidopsis isoform. Interestingly, G6PD4 did not group with any of the other Arabidopsis isoforms, but with a rice isoform.

Locus	Localization	a.a.	TP	MW (kD)	Designation
At5g35790	plastidic	626	50	60	G6PD1
At5g13110	plastidic	646	50	62	G6PD2
At1g24280	plastidic	656	57	61	G6PD3
At1g09420	plastidic	675	49	65	G6PD4
At3g27300	cytosolic	516		59	G6PD5
At5g40760	cytosolic	515		59	G6PD6

Table 2.2 The Arabidopsis G6PDH gene complement
Amino acids (a.a.) correspond to the preprocessed protein. Subcellular localization, transit peptide (TP) length in amino acids, and molecular weight (MW) of the mature protein were predicted *in silico.*

The subcellular localization of all plant G6PDH isoforms was predicted by evaluating the presence of a target peptide in the translated protein using ChloroP (Emanuelsson *et al.*, 1999) and TargetP (Emanuelsson *et al.*, 2000). The predictions correlated with the phylogenetic clades not only for Arabidopsis but for all species examined (data not shown). Translated proteins for *AtG6PD1*, *AtG6PD2*, *AtG6PD3*, and *AtG6PD4* were predicted as plastidic while those for *AtG6PD5*, and *AtG6PD6* were cytosolic (Table 2.2), consistent with previous results (Kruger and von Schaewen, 2003). The amino acid sequences of Arabidopsis isoforms showed high conservation with similarities ranging from 41 to 92% (Figure 2.2 and data not shown). The cysteine residues reported to be conserved among plant plastidic isoforms that confer redox sensitivity (Wenderoth *et al.*, 1997; Wendt *et al.*, 1999) were also present in all isoforms predicted to be plastidic (Figure 2.2). The G6PDH active site consensus sequence DHYLGKE as given by the PROSITE database (Hofmann *et al.*, 1999) was conserved in all except *AtG6PD4*, which differed in three out of seven amino acids (Figure 2.2). Interestingly, the most similar rice

isoform also had two amino acids changed in the active site in the same positions as in G6PD4 of Arabidopsis (data not shown).

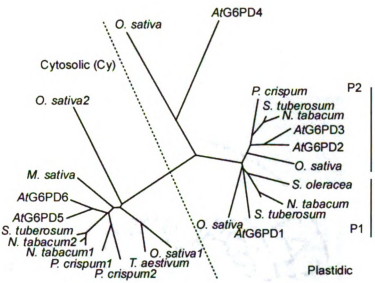
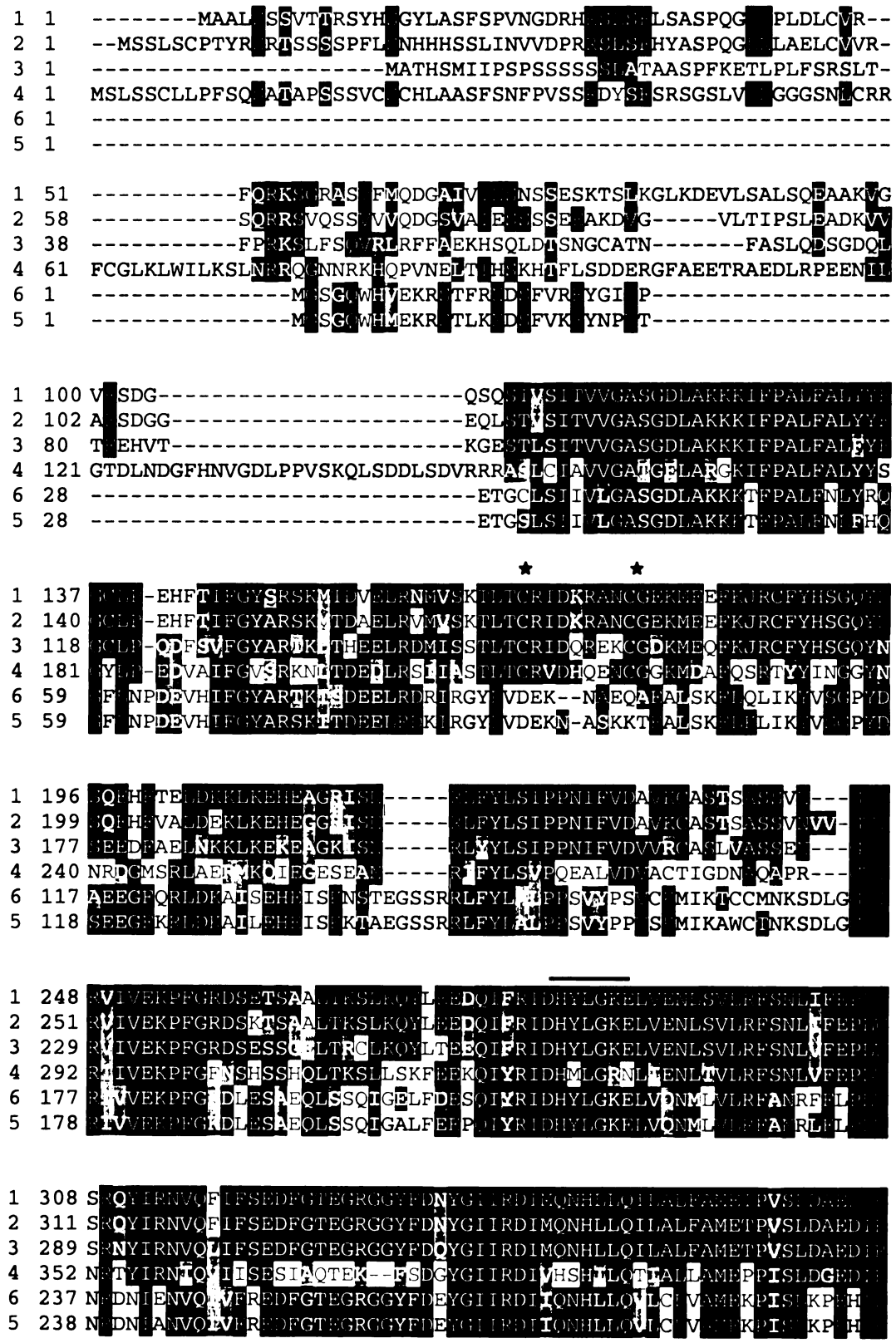


Figure 2.1 Phylogenetic analysis of plant G6PDHs

The mature protein sequences of G6PDHs were used. Abbreviations follow a previous report (Wendt et al., 1999, Figure 3), Cy, cytosolic; P1/P2, plastidic isoforms. SwissProt accession numbers are provided below. The ID numbers of DNA loci on the TIGR website (The Institute for Genomic Research http://www.tigr.org/tdb/e2k1/osa1/) are given for the two rice isoforms and their amino acid sequences were translated from available genomic sequence; *Arabidopsis thaliana AtG6PD1,* Q43727; *AtG6PD2,* Q9FY99; *AtG6PD3,* Q8L743; *AtG6PD4,* Q93ZW0; *AtG6PD5,* Q9LK23; *AtG6PD6,* Q9FJI5; *Medicago sativa,* AAB41552; *Nicotiana tabacum* P1, CAA04994; P2, AAF87216; Cy1, CAA04992; Cy2, CAA04994; *Oryza sativa*, P1, AAS07054; P2, BAC84352; predicted plastidic, translated protein of locus 7230.t00022; Cy1, translated protein of locus 4978.t0007; Cy2, CAC09489, CAE02006, CAE03156; *Petroselinum crispum,* Cy1, AAB69319; Cy2, AAB69318; P2, AAB69317; *Solanum tuberosum* Cy, CAA52442; P1, CAA58775; P2, CAB52708; *Spinacia oleracea,* P1, CAA03939; *Triticum aestivum,* Cy, BAA97663.



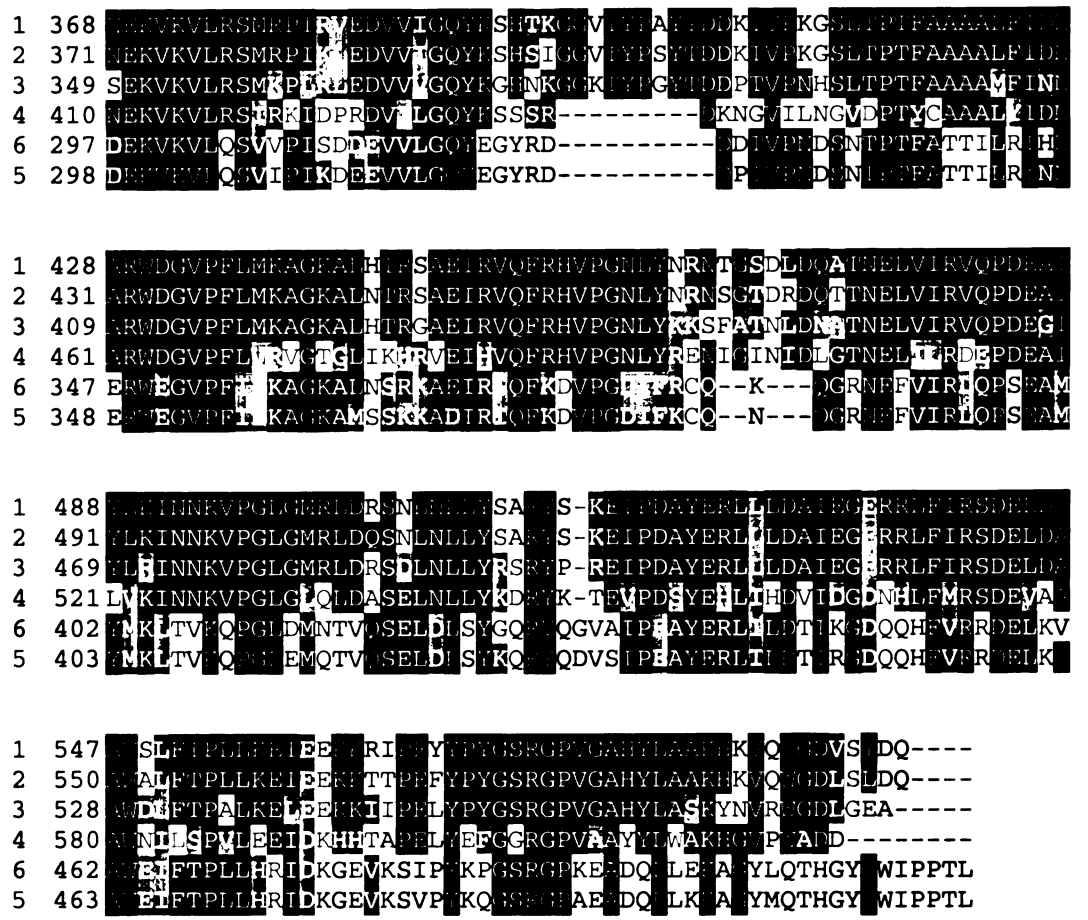


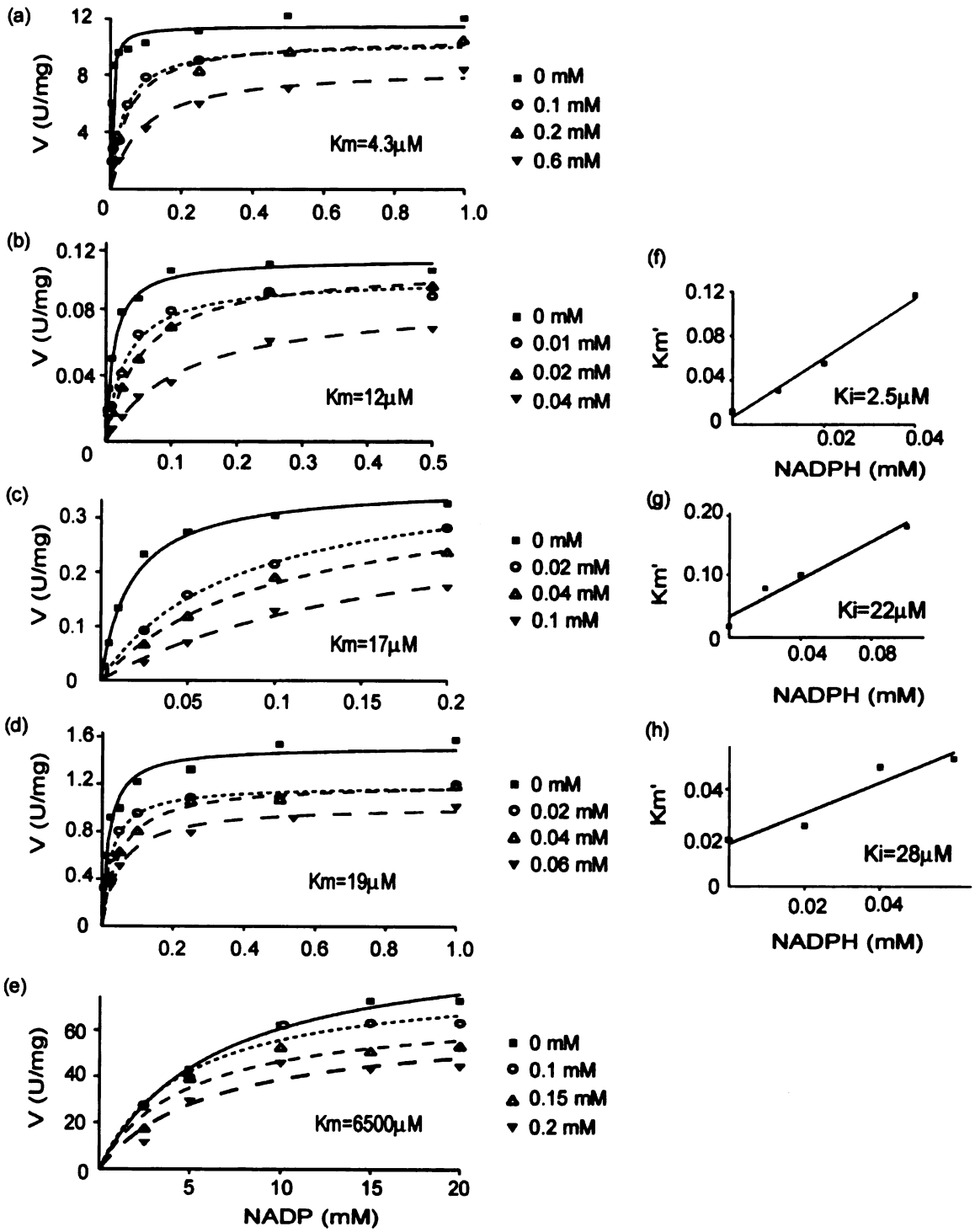
Figure 2.2 Protein alignment of Arabidopsis G6PDHs
Numbers preceding each row indicate 1, G6PD1; 2, G6PD2; 3, G6PD3; 4, G6PD4; 5, G6PD5; 6, G6PD6. A second set of numbers represents the amino acid positions on the respective protein. Identical amino acids are indicated by black boxes, similar ones are shaded gray. Asterisks indicate cysteine residues conserved among plastidic predicted isoforms. The bar indicates the G6PDH active motif DHYLGKE as defined by PROSITE database.

2.3.2 Biochemical characterization of recombinant proteins

The cDNAs for the mature G6PDH isoforms were expressed adding a C-terminal *Strep*-tag that allows subsequent purification from *E. coli*. Previously, the same system was successfully used for the expression of potato G6PDHs (Wendt *et al.*, 2000). Recombinant protein was purified by affinity chromatography after which one major band was observed by SDS-PAGE (data not shown). The yield of individual recombinant proteins varied among isoforms and individual preparations from 0.5 to 3 mg-protein/l-culture. All preparations lost activity rapidly upon storage at 4 °C, -20 °C with or without 20% glycerol, or –80 °C following snap-freezing (data not shown). Therefore all experiments described here were carried out immediately following purification. All experiments were repeated two to three times. In separate experiments of an isoform, K_{mNADP+} (K_m) values were reproducible whereas V_{max} values were lower when the time taken for purification process was long, or the column had been used repeatedly. In Figure 2.3 (a-e) an optimal repeat experiment for each isoform is shown in which the highest V_{max} was obtained. All isoforms had a similar pH optimum of 8-8.5 (data not shown), and utilized $NADP^+$ but not NAD^+ in a concentration-dependent manner. In all cases the plot of specific activity against substrate concentration followed a hyperbolic function. Isoform G6PDH4 exhibited very low, but measurable $NADP^+$ concentration-dependent activity indicating that $NADP^+$ is indeed a substrate (data not shown). The low activity was likely due to the G6PDH-atypical amino acid difference in the conserved active site (Figure 2.2) and thus, no further biochemical characterization of this isoform was pursued. Values for V_{max} were specific for each isoform but varied greatly among

the different recombinant proteins ranging from 0.1~100 U/mg-protein. All isoforms tested showed a concentration-dependent competitive inhibition by NADPH (Figure 2.3 a-e), as determined by Lineweaver-Burk plots (data not shown). A mixed inhibition mode might be suggested from the changing values of apparent V_{max} for G6PD5 (Figure 2.3d). However, this is possibly due to experimental limitation in completely saturating the enzyme with the substrate. Moreover, Lineweaver-Burk plots of repeated experiments indicated competitive inhibition (data not shown). The plot used to determine K_{iNADPH} (K_i) for each isoform is shown next to the substrate saturation curves (Figure 2.3f, g, h). The biochemical characteristics of the recombinant enzymes are summarized in Table 2.3. The different isoforms had K_m values for $NADP^+$ within the range of 1-20 μM, except for isoform G6PD6, which had a considerably higher K_m of 6.5 mM (Figure 2.3e, Table 2.3). The application of high substrate concentrations to this isoform and the addition of NADPH resulted in a high assay background, and therefore K_i was not determined for G6PD6. For any given isoform, a small K_i as compared to the K_m is indicative for a tight feed-back inhibition by NADPH. Only the K_i of G6PD2 was significantly lower than its K_m, whereas the two kinetic constants were comparable for G6PD3 and G6PD5 (Figure 2.3, Table 2.3). A precise K_i value could not be obtained for G6PD1, but results from several experiments were reproducibly in the range of 30-70 μM (Table 2.3).

Figure 2.3 Kinetic parameters of five active G6PDH isoforms
(a-e) Substrate ($NADP^+$) saturation curves under various inhibitor (NADPH) concentrations. The NADPH concentrations used for each enzyme are shown with their symbols accompanying their saturation curves. (f-h) Plots of apparent K_m (K_m') under different concentrations of the inhibitor used to determine K_i. Values of K_i are shown only for those that could be obtained reproducibly. (a) G6PD1; (b, f) G6PD2; (c, g) G6PD3; (d, h) G6PD5; (e) G6PD6.


(a)
V (U/mg)
Km=4.3μM
0 mM
0.1 mM
0.2 mM
0.6 mM
(b)
V (U/mg)
Km=12μM
0 mM
0.01 mM
0.02 mM
0.04 mM
(c)
V (U/mg)
Km=17μM
0 mM
0.02 mM
0.04 mM
0.1 mM
(d)
V (U/mg)
Km=19μM
0 mM
0.02 mM
0.04 mM
0.06 mM
(e)
V (U/mg)
Km=6500μM
NADP (mM)
0 mM
0.1 mM
0.15 mM
0.2 mM
(f)
Km'
Ki=2.5μM
NADPH (mM)
(g)
Km'
Ki=22μM
NADPH (mM)
(h)
Km'
Ki=28μM
NADPH (mM)

Isoforom	pH optimum	V_{max} (U/mg)	K_{mNADP} (μM)	K_{iNADPH} (μM)
G6PD1	8	11 (0.3)	4.3 (0.1)	30-70
G6PD2	8.5	0.11 (0.01)	12 (1)	2.5 (2)
G6PD3	8-8.5	0.36 (0.01)	17 (2)	22 (11)
G6PD4	8.5	<0.04	N.D.[a]	N.D.[a]
G6PD5	8.5	1.5 (0.05)	19 (3)	28 (14)
G6PD6	8	100 (5)	6500 (900)	N.D.[a]

[a] Not defined

Table 2.3 Biochemical characteristics of recombinant G6PDH isoforms
Values shown in parenthesis indicate standard deviation

This was an order of magnitude higher than its K_m (4.3 μM) suggesting that G6PD1 is less prone to feed-back inhibition. Taken together, these results indicate the different sensitivities to NADPH for each isoform, which may reflect the differential regulation required in different photosynthetic and redox environments, where the respective isoforms are present.

2.3.3 Sensitivity of the isoforms to reduction and oxidation

For the five active isoforms, the sensitivity to DTT which reduces disulfide bonds to dithiol (Johnson, 1972; Lendzian, 1980) was examined (Figure 2.4). The plastidic isoforms exhibited concentration-dependent inactivation by DTT over time (Figure 2.4a, b, c). This inactivation was reversed by addition of NaTT only for G6PD1 (P1 isoform), although the activity did not completely recover (Figure 2.4a). For the P2 isoforms G6PD2 and G6PD3 a significant reactivation by NaTT was not observed.

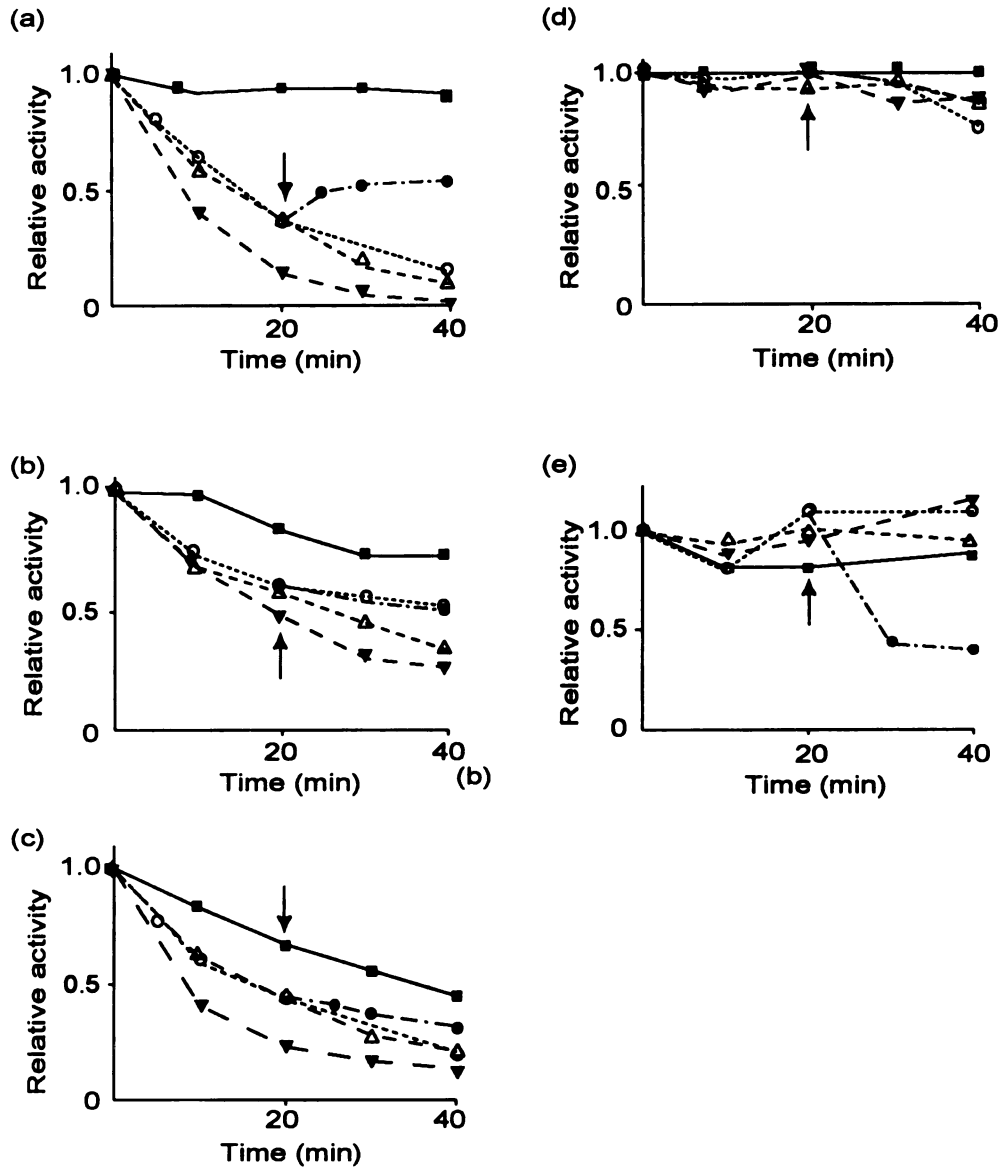


Figure 2.4 Sensitivity of G6PDH isoforms to reduction by DTT
Time dependent change in G6PDH activity following incubation with various concentrations of DTT. Activity is described as relative value starting with 1 at time zero. Arrows indicate the time when NaTT was added. Symbols indicate different concentrations of DTT and NaTT added, solid square, 0 mM DTT; open circle, 3 mM DTT; open triangle, 5 mM DTT; reverse solid triangle, 10 mM DTT; solid circle, 5mM NaTT added to 3 mM DTT. (a) G6PD1; (b) G6PD2; (c) G6PD3; (d) G6PD5; (e) G6PD6.

This could be due to an irreversible inactivation by DTT as was reported for the potato P2 isoform, which was reactivated only when incubated with DTT and thioredoxin but not with DTT alone (Wendt *et al.*, 1999), or it could be related to decreased stability during longer incubation times required for this experiment (Figure 2.4b, c). The inactivation by reduction is in agreement with previous reports for G6PDH isoforms from potatoes, which showed that only the plastidic isoforms possess the conserved cysteine residues and were sensitive (von Schaewen *et al.*, 1995; Wenderoth *et al.*, 1997; Wendt *et al.*, 2000). The cytosolic isoforms G6PD5 and G6PD6 were not affected by DTT (Figure 2.4d, e). Application of the oxidant NaTT to G6PD6 resulted in inactivation (Figure 2.4e), which was irreversible (data not shown), whereas G6PD5 was not affected by either DTT or NaTT (Figure 2.4d).

2.3.4 Expression of G6PDH genes varies greatly in different tissues

The frequency of reported ESTs in different cDNA libraries was examined for each isoform (as of July 2004) (Figure 2.5a). Isoform *AtG6PD6* was the most represented with 24 ESTs and was the only isoform reported to be present in many different library sources including whole plant, seedling, silique, seed, flower, and root. All other isoforms had ESTs in libraries from whole plants, seed or siliques. This *in silico* analysis allowed us to estimate the abundance of transcripts derived from each gene but the tissue-specificity for the isoforms was still unknown except for *AtG6PD6*. Therefore we examined the mRNA abundance for each isoform in different tissues using RT-PCR.

The primers were designed to amplify specific 3'-UTRs for each isoform. The primer pairs for each isoform were adjusted and tested for similar efficiencies with

genomic template (Figure 2.5b), allowing comparisons of transcript abundance between different isoforms. In Figure 2.5c, the gel photos are intended to allow comparisons of mRNA abundance across different isoforms whereas the graphs show relative levels of mRNA of a single isoform normalized to actin in different tissues. Judging from the intensity of amplified PCR products detected on the gel, overall the mRNA was most abundant for *AtG6PD6* (Figure 2.5c). In specific tissues mRNA was abundant for *AtG6PD1*, *AtG6PD2*, *AtG6PD3* and *AtG6PD5* (Figure 2.5c). The mRNA for *AtG6PD4* was detectable but its abundance was considerably lower compared to other isoforms (Figure 2.5c). This finding together with the amino acid difference in the active site and lack of activity suggests that *AtG6PD4* does not encode for a functional G6PDH. The mRNA of the highly expressed *AtG6PD6* gene encoding a predicted cytosolic isoform was found with similar levels in all tissues examined, suggesting it to encode the major G6PDH isoform throughout the plant. In contrast, the mRNA from *AtG6PD5,* predicted to encode the other cytosolic isoform, was detected at high level in leaves. The mRNA of the plastidic P1 isoform encoded by *AtG6PD1* was present in all tissues except for roots. Higher levels of mRNA were found in developing tissues such as buds, flowers, and early siliques compared to photosynthetic tissues (leaves and stems). This expression pattern was similar to that for the potato P1 isoform, for which the mRNA was predominant in green tissues. On the contrary, mRNAs for the P2 isoforms encoded by *AtG6PD2* and *AtG6PD3* were detectable in all tissues but their abundance was considerably higher in roots. It has been reported that *AtG6PD2* and *AtG6PD3* are induced in roots and shoots by nitrate (Wang *et al.*, 2003). In potato, the P2 gene was also found to be induced under

N-sufficiency in the roots when grown in hydroponic cultures, although expression was ubiquitous in soil-grown potato plants (Wendt *et al.*, 2000).

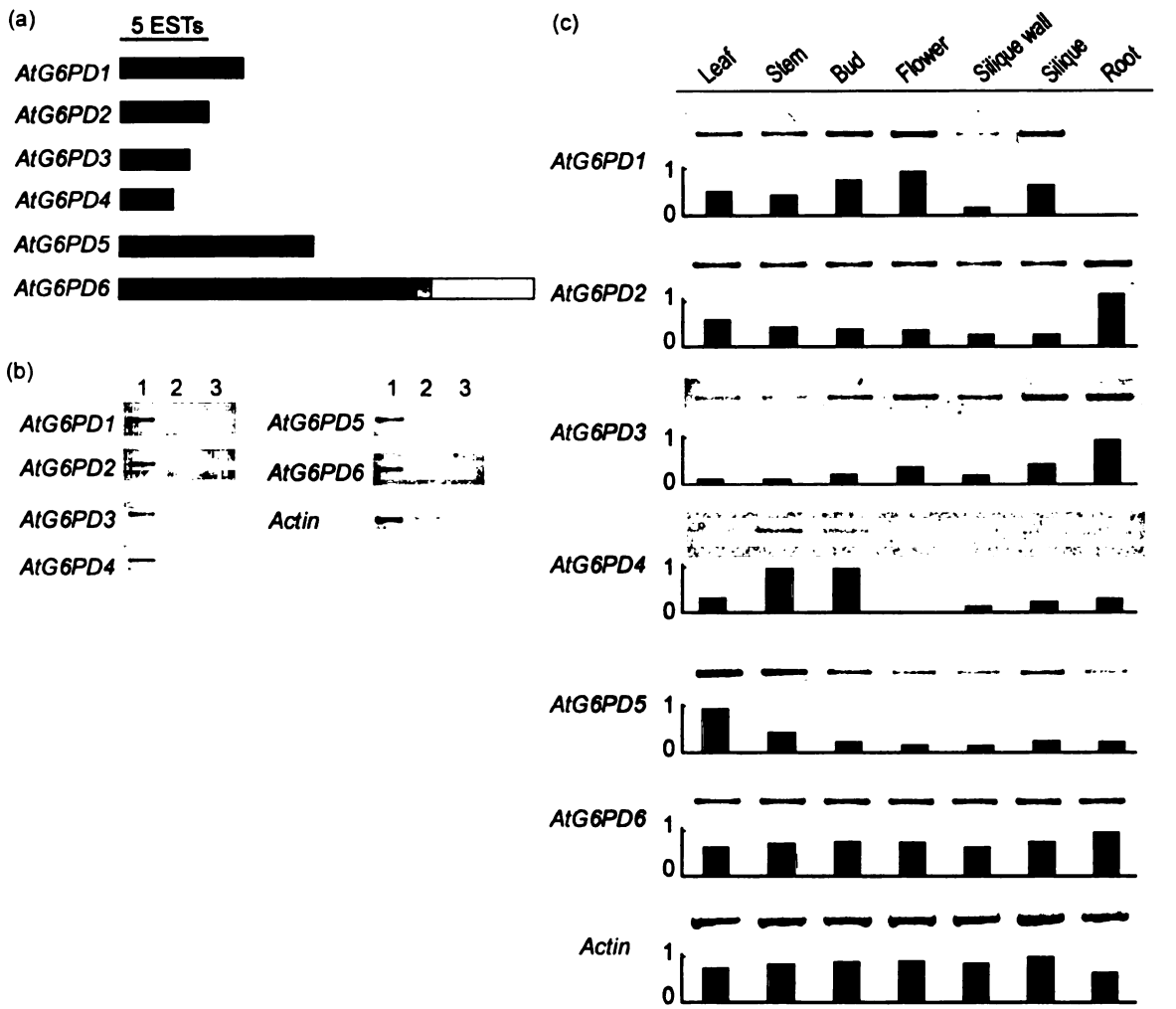


Figure 2.5 Expression of G6PDH isoforms in various tissues
(a) Number of ESTs reported. Bars indicate different sources of EST libraries. Solid, whole plant; striped, green siliques; dark gray, developing seeds; light gray, flowers; open, roots. (b) Efficiency of PCR primers tested in serial dilutions of genomic DNA template. PCR performed with genomic DNA template (lane 1); 1/10 dilution (lane 2); 1/100 dilution (lane 3). (c) RT-PCR from various tissues. Bar graphs indicate relative intensities in different tissues normalized to the actin fragment intensity from the respective tissue. The highest normalized intensity obtained defined as value 1.

2.3.5 G6PDH activity in different tissues

Three major isoforms of G6PDH activity are detected in vivo

As shown in Figures 2.3 and 2.4, G6PDH enzymatic activity was regulated post-translationally by different factors *in vitro* (substrate and inhibitor concentrations, and reductive inactivation). To begin exploring the roles for each isoform *in vivo,* a zymogram method separating native proteins based on their isoelectric point and detecting their activities was chosen rather than the detection of protein amounts, because enzyme activity can be considered the final level of gene expression. Zymograms are classic tools in the studies of G6PDH polymorphisms and isozyme identifications (Steele *et al.*, 1968; Eanes, 1983). On a given gel, higher mobility of a band indicates a lower isoelectric point of the protein. A bacterial G6PDH was loaded as a positive control. Three bands representing G6PDH activity (isoforms) were detected in stems and roots, of which two were also present in leaves (Figure 2.6a solid arrowheads, A, B, C). Three common isoforms (two more and one less active) with lower mobility were detected in buds, flowers, siliques and seed (Figure 2.6a open arrowheads, A', B', C'). No activity was detected when G6P was excluded (data not shown). With similar amounts of total protein loaded, of all tissues roots had three highly active isoforms (Figure 2.6a, A, B, C), whereas in photosynthetic leaves and stems, activity staining was less intense (Figure 2.6a, A, B) suggesting a significant difference in total G6PDH activity per total protein in these tissues. Developing tissues such as buds, siliques and seeds had the two isoforms with higher mobility showing relatively higher activity (Figure 2.6a, B', C'), whereas flowers had most activity in the second isoform (Figure 2.6a, B'). In all developing tissues, the lowest mobility isoform (Figure 2.6a, A') was low in activity. In a separate

experiment, less concentrated tissue extracts (approx. 1.7 mg-protein/ml) were assayed for G6PDH activity by zymograms and by measuring the total G6PDH activity in cell extracts spectrophotometrically (Figure 2.6b). The specific activity showed good correlation with band intensities observed on zymograms, indicating that this method is quantitative (Figure 2.6b, compare zymogram and graph). In addition, the proteins detected in this control experiment had similar mobility across different tissues, suggesting that the shift in bands observed in Figure 2.6a (A, B, C and A', B', C') is due to high sample concentrations (approx. 6.3 mg-protein/ml) applied to obtain the clearest possible banding pattern on the zymogram. Furthermore, to independently address whether the two sets of three bands with shifted mobility visible in Figure 2.6a were representative of the same three isoforms, extracts from wild-type buds and roots were mixed and analyzed (Figure 2.6c). The resulting zymogram showed an intermediate banding pattern. This together with the zymogram pattern of less concentrated extracts (Figure 2.6b), strongly suggests that isoforms detected as A, B, C are presumably equal to A', B', C' in Figure 2.6a. Because some isoforms were observed to be more labile as recombinant enzymes (G6PD2 and G6PD3, Figure 2.4b, c), it was considered that they could be preferentially lost under zymogram conditions. To address this possibility, the stability of the individual recombinant enzymes was analyzed by zymograms (Figure 2.6d). Similar amounts, approximately 0.2-1 mU, for each isoform were applied. All isoforms were detected except for G6PD4 (data not shown) indicating that none was significantly less stable under these conditions. Different extraction methods were tested, e.g. excluding DTT or protease inhibitors from the extraction buffer. These did not result in changes visible on the zymograms (data not shown) indicating that the lack of

detection of other isoforms is not due to inactivation by DTT or protease inhibitors during extraction.

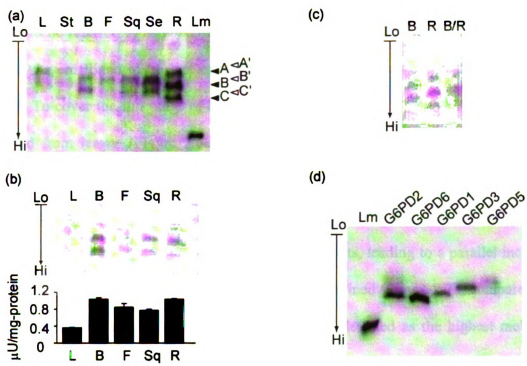


Figure 2.6 G6PDH activity in various tissues
G6PDH activity was detected by zymogram from (a, b, c), wild type tissue; (d), purified recombinant G6PDHs. (a), concentrated extracts; (b), dilute extracts also assayed spectrophotometrically as shown in the graph; (c), dilute and mixed extracts. Lanes, L, leaves; St, stems; B, buds; F; flowers; Sq, siliques; Se, seeds; R, roots; Lm, G6PDH from *Leuconostoc mesenteroides* (positive control); B/R, mixed buds and roots. Solid and open arrow heads indicate positions of bands detected in L, St, R (A, B, C) and B, F, Sq, Se (A', B', C'), respectively. Direction of electrophoresis is indicated by the arrow, the start of which is the origin. Lo, Hi, low and high protein mobility, respectively.

2.3.6 Identification of the genes encoding active G6PDH isoforms *in vivo*

To identify the genes coding for the active isoforms detected on zymograms, we took advantage of T-DNA insertion lines obtained from the SALK institute (Alonso *et al.*, 2003). The informative SALK lines available and their confirmed T-DNA insertion sites are shown in Figure 2.7a. The homozygous T-DNA insertion lines for the respective genes will be referred to as *g6pd5* and *g6pd6*. The 3'-UTR transcripts of *AtG6PD5* and *AtG6PD6* were found to be lost in leaves of *g6pd5* and *g6pd6* by RT-PCR (data not shown). To cover the full range of banding patterns observed (Figure 2.6a), protein was extracted from leaves (Figure 2.6a, A, B), buds (Figure 2.6a, A', B', C'), and roots (Figure 2.6a, A, B, C) from wild-type, *g6pd5*, and *g6pd6*, and was analyzed by zymograms. Protein extracts from *g6pd6* showed two isoforms, lacking the highest mobility isoform band present in wild-type buds and roots, leading to a parallel increase in intensity of a low mobility isoform in all tissues examined (Figure 2.7b, compare WT and *g6pd6*). This result indicated that (1) the isoform detected as the highest mobility band in wild-type plants that was lost from its original relative position in *g6pd6* is a product of *AtG6PD6*, and (2) although detected with different mobilities in Figure 2.6a, band C' in buds and band C in roots represent the same protein encoded by *AtG6PD6*. The isoform with higher mobility than that represented by an intensified band seemed reduced in activity especially in root tissues (Figure 2.7b *g6pd6* lane R), which could be partly due to the difficulty in its separation from the highly active isoform. Similarly, *g6pd5* lacked the isoform with the lowest mobility in all three tissues compared to wild-type, although residual activity was detected in leaves suggesting this T-DNA insertion line is an incomplete knock-out (Figure 2.7b compare WT and *g6pd5*, lane L). This result

indicates that the lowest mobility bands that were lost from all tissues of *g6pd5* were due to proteins encoded by *AtG6PD5*. The loss of the low mobility isoform was accompanied by an increase in intensity of a high mobility isoform with a similar mobility as G6PD6 in all three tissues examined (Figure 2.7b under *g6pd5*). The question arose whether G6PD6 and G6PD5 activities are reciprocally induced in *g6pd5* and *g6pd6*, respectively. The mobilities of the induced isoforms were similar (if not overlapping), to G6PD5 (lowest mobility) and G6PD6 (highest mobility) which suggests that in *g6pd5* and *g6pd6*, the isoforms G6PD6 and G6PD5 were induced, respectively. Nonetheless, it cannot be ruled out that an isoform that was previously not detected with the same mobility was induced in *g6pd5* and *g6pd6* or that the T-DNA insertion generated an active protein with altered structure that migrated together with G6PD5 and G6PD6 in both *g6pd6* and *g6pd5*, respectively. However these latter possibilities seem rather unlikely, as one would statistically expect a different migration for an altered or different protein in at least one of the two T-DNA insertion lines.

Results from multiple experiments suggest that the activity of only three isoforms were detected on the zymogram, which leaves one isoform, presumably a plastidic isoform, to be identified. Plants homozygous for a T-DNA insertion in an exon of *AtG6PD3* (SALK_020160) did not lose any activity visible on the zymogram although the 3'-UTR transcript of the gene was found to be aberrant (data not shown). Preferential instability of this isoform on zymograms can be ruled out as well, as shown in Figure 2.6d. These results suggest that this isoform has little or no activity *in vivo* even in roots where it is predominantly expressed under normal conditions (Figure 2.5c). However, these results do not rule out its conditional role as the gene coding for the G6PD3 has

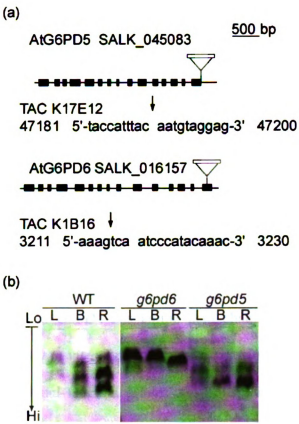


Figure 2.7 Identification of genes encoding G6PDH activity detected on zymograms using T-DNA insertion lines
(a) Schematic gene structure of *AtG6PD5* and *AtG6PD6*. The T-DNA insertion sites are shown as the nucleotide position in the respective genomic BAC clones (GenBank accession numbers, AP000381 (K17E12); AB015470 (K1B16)). Black boxes indicate the coding region. (b) G6PDH activity detected by zymogram in tissues from wild type (WT) and T-DNA insertion plants (*g6pd5* and *g6pd6*). Lanes L, leaves; B, buds; R, roots.

been found to be induced by nitrate in Arabidopsis (Wang *et al.*, 2003). Plants homozygous for T-DNA insertions several hundreds of base-pairs outside of the coding region for *AtG6PD1* (SALK_019159; upstream insertion) and *AtG6PD2* (SALK_003605; upstream insertion, SALK_005846; downstream insertion) were isolated and analyzed. All retained either the mRNA of the target gene or the wild-type G6PDH zymogram

pattern. Therefore the gene corresponding to the isoform that migrates in between G6PD5 and G6PD6 remains uncertain.

2.4 Discussion

2.4.1 The G6PDH complement in rice and Arabidopsis

The phylogenetic analysis gave new information on G6PD2 and G6PD6 belonging to P2 and cytosolic groups respectively. We also showed that the rice genome has loci presumably encoding for five G6PDHs, which represented members from each group as in Arabidopsis. Isoform G6PD4 has amino acid changes in the otherwise conserved active site, little activity *in vitro*, and low gene expression *in vivo*, properties suggesting that it is a non-functional G6PDH. There is a potential ortholog in rice that also has amino acid changes in the active site. Two of these positions are the same within the two proteins, although the replaced amino acids are different (data not shown). Both have plastidic G6PDH characteristics, N-terminal transit peptides and conserved cysteine residues, making it possible that these isoforms were once active plastidic G6PDHs. Whether G6PD4 requires other cofactors or has activity different from typical G6PDH enzymes is not known at this time.

2.4.2 Cytosolic G6PDH isoforms of Arabidopsis

The characterization of G6PDH isoforms is summarized in Table 2.4. The two cytosolic isoforms have distinct characteristics despite their sequence similarity (92% amino acid, data not shown). The abundance of mRNA and *in vivo* activity suggests that G6PD6 is the major cytosolic isoform of G6PDH in Arabidopsis. Isoform G6PD6 is unique in that

both values of V_{max} (100 U mg^{-1}) and K_m (6.5 mM) are at least an order of magnitude greater than those of the other isoforms. The cytosolic $NADP^+$ concentration in pea leaves in dark or light has been reported to be 140 μM, and a similar concentration of NADPH has been observed (Igamberdiev and Gardestrom, 2003). This suggests that under physiological conditions, G6PD6 is greatly sub-saturated and responds linearly to the change in $NADP^+$, i.e., redox levels, whereas the specific activity of the other cytosolic isoform, G6PD5, will not change (Figure 2.8). Isoform G6PD5 is the major cytosolic isoform in leaves of wild-type plants, and its activity is low in other tissues. The specific role of G6PD5 which appears to replace G6PD6 only in leaves is not known. The fact that G6PD5 is tolerant to both reduction and oxidation while G6PD6 is inactivated by oxidation (Figure 2.4d, e) may allow this isoform to carry out unique functions in leaves.

The activities of the two cytosolic isoforms appear to be coordinated. In wild type plants, all tissues except roots had only one of the two cytosolic activities (Figure 2.6a), with G6PD5 being detected in leaves and G6PD6 in the other parts of the plant. A reciprocal, but not equal, induction of their activities was observed in the two T-DNA insertion lines, assuming that the two induced isoforms observed in *g6pd5* and *g6pd6* were G6PD6 and G6PD5, respectively (Figure 2.7b). No induction at the mRNA level was detected in the T-DNA insertion lines (data not shown). It is not known whether this coordination and regulation occurs at the level of enzyme amount, substrate concentration, or post-translational modification. Cooperation between cytosolic and plastidic OPPP (in the provision of NADPH) has been observed in maize mutants lacking

cytosolic OPPP (Averill *et al.*, 1998). Increased rates of flux through OPPP were observed in these lines in response to nitrate although both genes encoding cytosolic 6-phosphogluconate dehydrogenase were disrupted. In this study, an increase in activity of what appeared to be the remaining cytosolic isoform was observed when one was disrupted, but none of the activities related to the plastidic G6PDH isoforms was stimulated. Whether an increase in plastidic G6PDH induction occurs when both cytosolic G6PDH isoforms are disrupted remains to be seen. Nonetheless our results suggest that cytosolic G6PDHs are coordinated in their activity, at least to an extent, presumably to maintain the redox balance in the cell.

2.4.3 Plastidic G6PDH isoforms in Arabidopsis

The P2 isoform encoding genes *AtG6PD2* and *AtG6PD3* showed similar abundance of ESTs, expression levels and patterns; lower in photosynthetic tissues and highest in roots (summarized in Table 2.4). Their proteins had similar kinetic parameters V_{max}, K_m, although G6PD3 may be less sensitive to feed-back inhibition. In contrast, *AtG6PD1*, belonging to the P1 group is highly expressed in developing tissues, to a lesser extent in green tissues, and absent in roots (Figure 2.5c, Table 2.4). The kinetic parameters for the G6PD1 protein were unique, especially its reduced sensitivity to feed-back inhibition (Table 2.3). All plastidic isoforms were sensitive to reductive inactivation by DTT although the degree of the sensitivity could not be directly compared due to their lability at room temperature (Figure 2.4a, b, c). The activity of the plastidic isoforms was measured with no prior treatment. We cannot entirely rule out the possibility that the recombinant plastidic enzymes as prepared were not completely oxidized and fully

active. However, because incubation with NaTT caused only a marginal activation (data not shown), we conclude that a large portion (if not all) of the protein was purified in its most active state.

Group	Gene	Expression	Feed-back inhibition	Reductive inactivation	*in vivo* activity[a]
Cy	*AtG6PD5*	Prevalent in leaf	$K_m \sim K_i$	Insensitive	Leaf and root
	AtG6PD6	Ubiquitous	N.D.[b]	Insensitive	Ubiquitous except leaf
P1	*AtG6PD1*	High in developing organs, absent in root	$K_m < K_i$	Sensitive	N.D.[b]
P2	*AtG6PD2*	Highest in root	$K_m > K_i$	Sensitive	N.D.[b]
	AtG6PD3	Highest in root	$K_m \sim K_i$	Sensitive	N.D.[c]
Other	*AtG6PD4*	Low	N.D.[b]	N.D.[b]	N.D.[b]

[a] Presence was determined by loss of a band in KO compared to wild-type

[b] not defined

[c] not detected

Table 2.4 Summary of the Arabidopsis G6PDH isoform properties
Cy, cytosolic; P1, plastidic P1; P2, plastidic P2. $K_m > K_i$, $K_m \sim K_i$, feed-back inhibited; $K_m < K_i$, no feed-back inhibition.

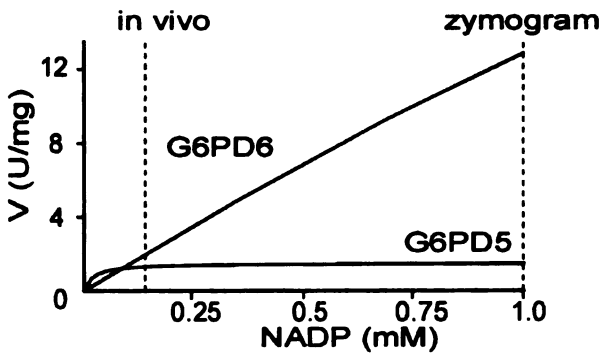


Figure 2.8 G6PDH activities in Arabidopsis.
Schematic representation of activities of the two cytosolic isoforms under a range of predicted concentrations of cytosolic $NADP^+$. The dotted line indicates the $NADP^+$ concentration reported for the cytosol (0.14 mM) and that used for the zymogram (1.0 mM).

Previous studies in potato and barley supported the view that P1 isoforms, which are predominantly present in green tissues, are feed-back inhibited by NADPH, whereas ubiquitous P2 isoforms are less sensitive to feed-back inhibition and to reductive inactivation (Wendt *et al.*, 2000; Esposito *et al.*, 2003). Our results show an alternative case in Arabidopsis where G6PD1 (P1) is the least NADPH sensitive of all isoforms and ubiquitously expressed except for roots, while G6PD2 and G6PD3 (P2) gene expression was not ubiquitous, but prevalent in roots (Figure 2.5c, Table 2.4). There are some possible explanations for these species-specific differences. The model proposed by Wendt *et al.* (2000) assumes the presence of two ancestral G6PDHs that arose to have different roles and characteristics in different tissues (P1 and P2), and predicts that similar features are maintained among the P1 and P2 isoforms in different plants. Since expression patterns are more often determined by regulatory sequences outside of the

coding region, which is not reflected in the phylogenetic tree, it is possible that in the case of Arabidopsis G6PDHs, changes in regulatory sequences occurred that altered expression patterns. Subsequently this was accompanied by amino acid changes leading to altered tertiary structures that rendered the enzymes more suitable for their roles in new tissues. In addition, we have shown that there is little correlation between mRNA levels and activity of G6PDH isoforms, which may indicate that expression patterns are not strong indicators of the *in vivo* roles of the native gene products at least in the case of G6PDH isoforms.

Plastidic isoforms G6PD1 and G6PD2 are the likely candidates for the ubiquitous isoform detected in zymograms (because T-DNA insertion in *AtG6PD3* did not lead to activity loss on a zymogram and G6PD4 is inactive). The overall transcript abundance throughout the plant was similar for both, but in neither case did the expression patterns match the activity present in all tissues as detected by zymograms. These results also suggest that the major regulatory mechanisms of Arabidopsis G6PDH isoform activity are post-transcriptional, at the levels of protein amount, substrate or inhibitor concentrations and post-translational redox modification. This was contrast to the case of potato, in which the induction of cytosolic G6PDH activity and increased mRNA abundance was observed in response to sugar, while the activity of plastidic (P1) isoform was induced by oxidative conditions, though protein levels remained the same (Hauschild and von Schaewen, 2003).

We initially detected in zymograms two sets of three apparent isoforms whose migration was slightly shifted (Figure 2.6a), two of which were identified to be pairs of G6PD5 (Figure 2.6a, A and A') and G6PD6 (Figure 2.6a, C and C') using T-DNA

insertion plant extracts. Our results suggest that bands B and B' in Figure 2.6a represent one G6PDH isoform, hence there remains only one isoform to be identified on zymograms. We favor the possibility of G6PD1 being the unidentified plastidic isoform visible on zymograms for the following reasons, (1) its V_{max} was relatively high as for the two cytosolic isoforms for which activity was detected on zymograms, and (2) its reduced sensitivity to feed-back inhibition (Table 2.3) allows this enzyme to be active in photosynthetic tissues. The activity of G6PD3, a P2 isoform, was not detected in the plant under normal conditions in agreement with its relatively low gene expression and V_{max} of the recombinant protein. Nonetheless, the possibility of G6PD3 and/or G6PD2 being active under specific conditions remains as *AtG6PD2* and *AtG6PD3* have been reported to be induced by nitrate in Arabidopsis (Wang *et al.*, 2003). Furthermore, the correlation of activity of a P2 isoform with oxidative stress has been shown in a recent study, in which transgenic tobacco plants expressing anti-sense cDNA for the plastidic P2 isoform unexpectedly showed increased tolerance to methylviologen (Debnam *et al.*, 2004). Whether Arabidopsis P2 isoforms are involved in oxidative- or other stress responses is not known at this time.

2.4.4 Interplay of different G6PDH isoforms *in vivo*

It has been widely accepted that native forms of G6PDH occur in homodimers or homotetramers, the latter shown to be the active form in maize leaves (Kahler and Kirkman, 1983; Srivastava and Anderson, 1983; Valenti *et al.*, 1984). Thus we assumed that the native active G6PDH enzymes in Arabidopsis are homomers, presumably in the form of homotetramers. This was supported by the activity of the recombinant proteins

expressed from a single gene in *E. coli*. We also observed the reduction of only one band of defined mobility (i.e., isoelectric point) by gene disruption indicating that G6PD5 and G6PD6 are each involved in a single complex for which activity could be detected on a zymogram.

The $NADP^+$ concentration (1 mM) applied to the zymograms is higher than *in vivo* concentrations, 0.14 mM in the cytosol (Igamberdiev and Gardestrom, 2003), and 0.4 mM total NADP(H) in the chloroplast stroma (Igamberdiev *et al.*, 2001). Therefore maximal activity is detected for all isoforms except for G6PD6, for which the K_m is higher (6.5 mM), a fact which leads to the over-representation of this isoform on a zymogram compared to other isoforms (Figure 2.8). Nonetheless, our zymogram experiments revealed many new findings such as the different levels of total G6PDH proteins present in active forms in a tissue as well as distinct combinations of isoforms. In summary, we observed three patterns of active isoforms as follows (Figure 2.6a and depicted in Figure 2.9); (1) isoforms G6PD5 and a plastidic isoform in green tissues, (2) G6PD6 and a plastidic isoform in developing tissues (mostly a plastidic isoform in flowers), and (3) all three isoforms in non-photosynthetic roots. Although G6PDH has conventionally been described as a house-keeping enzyme or dark enzyme, these results suggest that no single isoform is capable of all functions and that each has its individual roles. These patterns and combinations perhaps reflect the metabolism of the subcellular compartments and photosynthetic activity of the tissue, suggesting that G6PDH activity patterns may infer metabolic states brought about by differences in rates of reductive and oxidative reactions and the demand for NADPH.

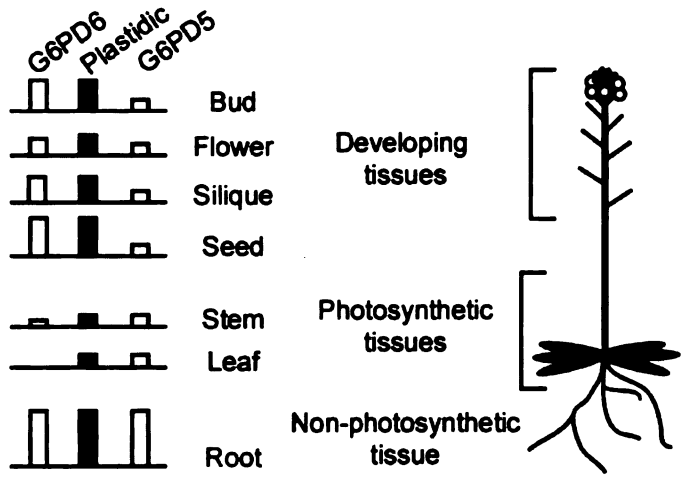


Figure 2.9 Schematic representation of relative activity of isoforms in plant tissues detected in zymograms.

2.5 Conclusions and perspectives

In this post-genomic era the deciphering of genomic sequences is routinely accomplished and gene expression at the mRNA level can be readily determined using multiparallel approaches. However, the parallel activity analysis of multiple enzymes still presents great challenges, even if the enzymes in question catalyze similar or identical reactions. This case study showed that none of the isoforms of G6PDHs in Arabidopsis were exactly the same in activity, sensitivity to inhibitors or with regard to the expression of the respective genes. As G6PDHs are enzymes of the central metabolism, what can be expected for the much larger enzyme families of secondary metabolism, which might have highly specialized functions? The difficulty in the comparison of gene family

members and their gene products arises from the need to detect levels of mRNA, proteins, or enzyme activity with high specificity and equal sensitivity for each member to allow meaningful conclusions. For example, RT-PCR conditions need to be carefully calibrated for each set of specific primers, different antisera against individual isoforms need to be highly specific and equally sensitive in Western blots, and enzyme isoforms need to be equally stable under the chosen assay conditions. With the exception of the generation of specific antisera, all the conditions were met in the current G6PDH study. Unlike the generation of specific antisera, a time-consuming process, zymograms are well suited for a genomic approach, because they provide parallel information and give an instant overview of the isoform activity for a given tissue or under a given environmental stimulus. However, they need to be carefully controlled and will not be applicable to all classes of enzymes. Other potential pitfalls of a global enzyme analysis approach as described here arise from the fact that recombinant proteins with affinity tags are analyzed, which might differ in activity and sensitivity to regulators as compared to their native forms. Furthermore, many enzymes have a heteromeric subunit composition, are integral to membranes, or are part of larger complexes and their activity cannot be reconstituted from a single protein. While there is much to be learned from the systematic activity analysis of protein families, the technical feasibility and the appropriateness of biochemical approaches need to be carefully evaluated in each case. The systematic activity analysis possibly in parallel with global structural analysis of proteins will be required to more fully understand the functioning of plant cells. After all, gene expression becomes manifested in the biochemical activity of the encoded gene product, in many cases an enzyme involved in central or secondary metabolism.

Reference

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H.M., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., Ecker, J.R. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301,** 653-657.

Averill, R.H., Bailey-Serres, J., Kruger, N.J. (1998) Co-operation between cytosolic and plastidic oxidative pentose phosphate pathways revealed by 6-phosphogluconate dehydrogenase-deficient genotypes of maize. *Plant J.* **14,** 449-457.

Debnam, P.M., Fernie, A.R., Leisse, A., Golding, A., Bowsher, C.G., Grimshaw, C., Knight, J.S., Emes, M.J. (2004) Altered activity of the P2 isoform of plastidic glucose 6-phosphate dehydrogenase in tobacco (*Nicotiana tabacum cv. Samsun*) causes changes in carbohydrate metabolism and response to oxidative stress in leaves. *Plant J.* **38,** 49-59.

Eanes, W.F. (1983) Genetic localization and sequential electrophoresis of glucose-6-phosphate dehydrogenase in *Drosophila melanogaster*. *Biochem. Genet.* **21,** 703-711.

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

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

Esposito, S., Massaro, G., Vona, V., Rigano, V.D., Carfagna, S. (2003) Glutamate synthesis in barley roots: the role of the plastidic glucose-6-phosphate dehydrogenase. *Planta* **216,** 639-647.

Felsenstein, J. (1989) PHYLIP -- Phylogeny Inference Package (Version 3.2). *Cladistics* **5,** 164-166.

Focks, N., Benning, C. (1998) *wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* **118,** 91-101.

Hauschild, R., von Schaewen, A. (2003) Differential regulation of glucose-6-phosphate dehydrogenase isoenzyme activities in potato. *Plant Physiol.* **133,** 47-62.

Hofmann, K., Bucher, P., Falquet, L., Bairoch, A. (1999) The PROSITE database, its status in 1999. *Nucleic Acids Res.* **27,** 215-219.

Igamberdiev, A.U., Bykova, N.V., Lea, P.J., Gardestrom, P. (2001) The role of photorespiration in redox and energy balance of photosynthetic plant cells: A study with a barley mutant deficient in glycine decarboxylase. *Physiol. Plant.* **111,** 427-438.

Igamberdiev, A.U., Gardestrom, P. (2003) Regulation of NAD- and NADP-dependent isocitrate dehydrogenases by reduction levels of pyridine nucleotides in mitochondria and cytosol of pea leaves. *Biochim. Biophys. Acta* **1606,** 117-125.

Johnson, H.S. (1972) Dithiothreitol - inhibitor of glucose-6-phosphate-dehydrogenase activity in leaf extracts and isolated chloroplasts. *Planta* **106,** 273-277.

Kahler, S.G., Kirkman, H.N. (1983) Intracellular glucose-6-phosphate-dehydrogenase does not monomerize in human-erythrocytes. *J. Biol. Chem.* **258,** 717-718.

Kruger, N.J., von Schaewen, A. (2003) The oxidative pentose phosphate pathway: structure and organisation. *Curr. Opin. Plant Biol.* **6,** 236-246.

Lendzian, K.J. (1980) Modulation of glucose-6-phosphate-dehydrogenase by NADPH, NADP+ and dithiothreitol at variable NADPH-$NADP^+$ ratios in an illuminated reconstituted spinach (*Spinacia oleracea* L) chloroplast system. *Planta* **148,** 1-6.

Li, K.B. (2003) ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics* **19,** 1585-1586.

Segel IH (1975) Enzymes. In *Biochemical calculations* John Wiley & Sons, pp 208-323.

Srivastava, D.K., Anderson, L.E. (1983) Isolation and characterization of light-modulatable and dithiothreitol-modulatable glucose-6-phosphate-dehydrogenase from pea-chloroplasts. *Biochim. Biophys. Acta* **724,** 359-369.

Steele, M.W., Young, W.J., Childs B. (1968) Glucose-6-phosphate dehydrogenase in *Drosophila melanogaster:* Starch gel electrophoretic variation due to molecular instability. *Biochem. Genet.* **2,** 159-175.

Valenti, V., Stanghellini, M.A., Pupillo, P. (1984) Glucose-6-phosphate-dehydrogenase isozymes of maize leaves - some comparative properties. *Plant Physiol.* **75,** 521-526.

von Schaewen, A., Langenkamper, G.S.R., Graeve, K., Wenderoth, I., Scheibe, R. (1995) Molecular characterization of the plastidic glucose-6-phosphate dehydrogenase from potato in comparison to its cytosolic counterpart. *Plant Physiol.* **109,** 1327-1335.

Wang, R., Okamoto, M., Xing, X., Crawford, N.M. (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiol* **132,** 556-567.

Wenderoth, I., Scheibe, R., vonSchaewen, A. (1997) Identification of the cysteine residues involved in redox modification of plant plastidic glucose-6-phosphate dehydrogenase. *J. Biol. Chem.* **272,** 26985-26990.

Wendt, U.K., Hauschild, R., Lange, C., Pietersma, M., Wenderoth, I., von Schaewen, A. (1999) Evidence for functional convergence of redox regulation in G6PDH isoforms of cyanobacteria and higher plants. *Plant Mol. Biol.* **40,** 487-494.

Wendt, U.K., Wenderoth, I., Tegeler, A., von Schaewen, A. (2000) Molecular characterization of a novel glucose-6-phosphate dehydrogenase from potato (*Solanum tuberosum* L.). *Plant J.* **23,** 723-733.

Yu, B., Wakao, S., Fan, J., Benning, C. (2004) Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in Arabidopsis. *Plant Cell Physiol.* **45,** 503-510.

Wakao, S., Benning, C. (2005) Genome-wide characterization of glucose-6-phosphate dehydrogenases in Arabidopsis. *Plant J.* **41,** 243-256.

Chapter 3

Roles of G6PDH isoforms in the plant and seeds of Arabidopsis

3.1 Introduction

G6PDH in organisms other than plants has been implicated in oxidative stress response. In humans, certain alleles of the G6PDH gene are associated with hypersensitivity in erythrocytes to oxidative stress such as exposure to drugs, infection and ingestion of fava beans (reviewed by Vulliamy *et al.*, 1992; Martini and Ursini, 1996). In yeast and *E. coli*, genes coding for G6PDH are up-regulated by oxidative stress (Kletzien *et al.*, 1994). The involvement of G6PDH in oxidative stress response is presumed to be through more than one pathway, consistent with the complexity of the process. For example, increased G6PDH gene expression results in resistance to oxidative stress through increasing glutathione levels in human cell cultures (Salvemini *et al.*, 1999), and the loss of G6PDH affects the binding of a DNA repair protein to DNA (Ayene *et al.*, 2002). Not only does G6PDH generate NADPH, it is also a part of the pentose phosphate pathway that produces precursors for nucleotide and aromatic amino acid biosynthesis (in plants). Therefore the analyses of the functional roles of G6PDH by enzyme activity alteration requires considering the possibility that more than one metabolic process are affected.

Current studies on the roles of G6PDH isoforms in plants have focused mostly on nitrogen assimilation (Chapter 1). Increased gene expression of G6PDHs in various conditions such as viral infection, drought and salt stresses has been reported in plants (Chapter 1, and references therein). But to date, only one report addresses the functional role of a plant G6PDH by alteration of its activity, in which the activity of a tobacco isoform was repressed by antisense expression (Debnam *et al.*, 2004). The transgenic tobacco plants had reduced G6PDH activity and unexpectedly had higher tolerance to

methylviologen, an inhibitor of the electron transport chain. Apparently, the role of G6PDH cannot be simply explained as in other organisms.

In plants, the presence of multiple isoforms in the cytosol and plastids complicates the studies addressing the roles of G6PDHs. In the present study, the availability of the genome sequence and Arabidopsis T-DNA insertion lines are exploited to characterize the *in vivo* roles of G6PDH in plants with altered activity of a single isoform. First, the effect on the plant of the loss of one or two isoforms was analyzed with focus on oxidative stress. Second, the contributions of G6PDH isoforms in seed oil accumulation were examined by characterizing the gene expression and enzyme activity during seed development and then the effect of the loss of a G6PDH isoform on seed oil content.

3.2 Material and methods

3.2.1 Plant growth conditions

All seeds were surface sterilized by incubating in 20% bleach, 0.05% Triton-X. The tubes containing the seeds were inverted for 15 min and washed three times with water. The seeds were suspended in 0.1% agar and plated onto 1X MS medium (pH 5.8) (Murashige and Skoog, 1962) with 1% sucrose, 0.9% agar unless stated otherwise. The duration of stratification was usually one to two days at 4 °C. Plants were grown in incubators with photon flux density of 60-80 μmol m^{-2} sec^{-1}, 16-8 h light-dark photoperiod, 22 and 18 °C, light and dark temperatures.

3.2.2 Plant transformation

All transformations were performed by Agrobacterium-mediated floral dipping (Clough and Bent, 1998).

3.2.3 Transient expression of *G6PDH::GFP* for subcellular localization analysis

The coding region of *G6PD5* and *G6PD6* were amplified with the following primers; for *G6PD5,* (+) 5'-GGACTAGTATGGGTTCTGGTCAATGGCA, (-) 5' GGACTAGTCAATGTAGGAGGGATCCAAA, and for *G6PD6,* (+) 5'-GGACTAGTATGGGATCTGGTCAATGGCA, (-) 5'-GGACTAGTTAGTGTAGGAGGGATCCAG. The cDNAs were cloned into the *Spe*I site of pCAMBIA1302 (Genbank accession no. AF234298). The orientation was confirmed by restriction analysis and sequencing. Onion epidermal peels were bombarded with the above constructs following the methods previously described (Varagona *et al.*, 1992) except that 0.1 M spermidine was used instead of thiamine to precipitate the DNA on to tungsten particles. The tungsten particles were resuspended in 30 µl 100% ethanol, and 10 µl was spotted onto a macrocarrier disc (Biorad, Heracules, CA). The DNA was bombarded using 1100 psi (pounds per square inch) rupture discs at approx. 4 cm distance using a biolistic gene delivery system (Dupont). For each construct, three peels were bombarded and incubated overnight at 22 °C in the dark. The peels were observed with a Leica DMR A2 microscope in the fluorescence mode with the L5 filter cube (Leica Microsystems, Wetzlar, Germany).

3.2.4 Oxidative stress experiments

After 12 days of incubation in the growth chamber, the seedlings were transferred to plates with various concentrations of 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) and sucrose (Suc) and photographed two weeks after the transfer. MS plates were prepared with 0, 100, 250, 500 nM DCMU. Serial dilutions of DCMU stock solutions in dimethyl sulfoxide (DMSO) were prepared such that all plates had an equal concentration of DMSO.

3.2.5 Sugar response experiments

WT, single and double mutant seeds were germinated and grown in the presence of 1% Suc for 10 days, after which they were transferred to plates with or without Suc. Seedlings were collected 8 days after transfer and the mRNA of *G6PD5* and *G6PD6* were examined by RT-PCR. G6PDH activity was examined by zymograms as described in Chapter 2 Material and methods.

3.2.6 RT-PCR for gene expression analysis

Total RNA was extracted using a kit (RNeasy, QIAGEN, Valencia, CA) from seedlings for the gene expression analysis in the single and double mutants. Primers were designed against the 5'-UTR in which the mRNAs of *G6PD5* and *G6PD6* showed sequence variation (primers agains the 3'-UTR were not used because the T-DNA had inserted in this region). The following primers were used for *G6PD5*, (+) 5' AATTTTGTCGCCACCGTCG, (-) 5'-CGTGACTGGATTGTATTCTTTT and for *G6PD6*, (+) 5'-AGTGGGAGAAAATGACGGAA-3', (-) 5'-

TGGAACTATACCATATTCTCTC. For gene expression analysis in developing seeds and siliques, total RNA extraction and subsequent RT-PCR from developing seeds and siliques were carried out as previously described in Chapter 2 and in Yu *et al.* (2004).

3.2.7 Construction of complementation vectors for *G6PD5* and *G6PD6*

To complement the T-DNA insertion lines, *G6PD5* cDNA for the coding sequence was inserted into pBI121 (Genbank accession no. AF485783), after removing the region encoding β-glucuronidase by digesting with *Xba*I and *Sac*I. The same sites were used to insert the cDNA. Primers used for amplification of the cDNA and introduction of the restriction sites were 5'-GCTCTAGAATGGGTTCTGGTCAATGGC-3', and 5'-CCGGAGCTCTTACAATGTAGGAGGGATCC-3'. The cDNA of the coding sequence for *G6PD6* was inserted into pBinAR-Hyg (Becker, 1990) using the *Kpn*I site. The primers used were 5'-GGGGTACCATGGGATCTGGTCAATGGCACGTT-3' and 5'-CCGGTACCTTATAGTGTAGGAGGGATCCAG-3'. The orientation was confirmed by sequencing. The reverse primers for both genes include the stop codons to avoid extra amino acids being included in the protein from the vector.

For complementation of the T-DNA insertion lines with a genomic fragment, the genomic region containing the full-length cDNA and approximately 1 kb upstream of the transcription initiation site of *G6PD5* and *G6PD6* was amplified by PCR using Expand High Fidelity PCR system (Roche, Indianapolis, IN). The 4.8 kb and 5.1 kb PCR will be inserted into the *Kpn*I and *Sma*I sites in pCAMBIA1300, respectively. The primers used for *G6PD5* and *G6PD6* are (+) 5'-GGGGTACCGAGAATGAAACAGTCGTAATGTG-3', (-) 5'-GGGGTACCCTTAGAAAGCTGGTTAAATTAGG-3' and (+) 5'-

GCTATATCAAGACTCATATGTATA-3', (-) 5'-GATGACTATGTAAAGACGGCGT-3', respectively.

3.2.8 RNA interference (RNAi) plants of *G6PD1* and *G6PD2*

Fragments of approximately 150 nucleotides in the 3'-UTR of *G6PD1* and *G6PD2* were selected as a target for RNA interference (RNAi). No more than eight consecutive bases matched between the two fragments when analyzed using the program "Calculate Optimal Local Sequence Alignments" (http://workbench.sdsc.edu/) (data not shown). The primers were designed to contain two restriction sites each to allow insertions in both sense (*Asc*I/*Swa*I sites) and antisense (*Spe*I/*Xba*I) orientation into pGSA1285 using a single primer set (http://chromdb.org/info/plasmids/vectors2.html) (Figure 3.6 for schematic representation of the constructs). For *G6PD1,* due to difficulty with the antisense insertion, a second set of primers with only *Spe*I sites was used. For the sense insertion of *G6PD1,* the following primers were used; (+) 5'-GGACTAGTGGCGCGCCTTAAGCTTTCGACAGAAACAA-3', (-) 5'-GCTCTAGAATTTAAATAAATGCTCGATTTAGAACCAA-3', and for the antisense insertion (+) 5'-GGACTAGTTTAAGCTTTCGACAGAAACA-3', and (-) 5'-GGACTAGTAAATGCTCGATTTAGAACCA-3'. For the insertion of *G6PD2* in both orientations, primers (+) 5'-GGACTAGTGGCGCGCCCGTCAGTATTGACCAGTAAAA-3', and (-) 5'-GCTCTAGAATTTAAATTATTTATTTATTAAATGGTACATAA-3' were used. For the construction of a seed-specific RNAi vector, the 35S promoter was removed from pGSA1285 by digesting with *Bgl*II and *Nco*I. A 1148 bp fragment upstream of a

Arabidopsis 2S albumin storage protein (At4g27160, found on BAC T24A18; Genbank accession no. AL035680) that has been found to drive gene expression in Arabidopsis seeds (Ohlrogge J.B. et al., United States Patent Application 20030005485) was amplified from genomic DNA template with primers (+) 5'-GAAGATCTCAAGAGTGTAAAACGTACC, (-) 5'-CATGCCATGGTTTTGCTATTTGTGTATGTTT, and inserted into the same sites.

3.2.9 Zymogram with seed extracts

Seed protein extraction and electrophoresis of cellulose acetate (CAE) plates were performed as described in Chapter 2 Material and methods, except approximately 6- to 8-fold (volume/weight-seeds) extraction buffer was used. Seed protein extracts were adjusted to approximately 1 to 2 mg-protein/ml. CAE was overlaid with 0.2 mM 5'-cyano-2,3-ditolyl tetrazolium chloride (CTC; excitation wavelength 450 nm, emission wavelength 630 nm) (Polysciences, Warrington, PA) instead of nitroblue tetrazolium chloride (NBT) as in zymograms for other tissue extracts. A stock solution of 50 mM CTC was prepared in water, 20 µl of which was added to the 5 ml of overlay solution (Chapter 2 Material and methods). The fluorescence was scanned using a fluorescence scanner (Molecular Imager FX, Biorad, Hercules, CA) with excitation wavelength of 532 nm and detection with a 640 nm filter, and the image was processed with PDQuest (Biorad, Hercules, CA).

3.2.10 Seed oil content measurement with gas chromatography

Plants of different genotypes were grown in the same growth chamber and the seeds were harvested after having dried in the siliques attached to the plant. Single mature seeds from a single plant of which the genotype was confirmed by PCR, were ground using a 6 mm glass bead (Fischer Scientific, Pittsburgh, PA) in a glass tube using a paint shaker for three min. Lipid extraction and sample preparation for quantification by gas chromatography (GC) was carried out as described previously (Focks and Benning, 1998).

3.2.11 Elemental (carbon and nitrogen) analysis in seeds

Seeds were sent to the Duke Environmental Stable Isotope Laboratory (DEVIL) for elemental analysis. Between 2-5 mg of seeds were sent in three replicates for each plant, and for some of the plants that seed oil content was analyzed by GC.

3.3 Results

3.3.1 Subcellular localization of G6PD5 and G6PD6

The subcellular localization of the two isoforms G6PD5 and G6PD6 was examined by transient expression of the G6PDH genes with an N-terminal fusion to a green fluorescence protein (GFP). For both constructs containing *G6PD5* and *G6PD6*, the green fluorescence was observed dispersed in the cytosol and surrounding what is presumably the nucleus (Figure 3.1). The same patterns were observed in multiple experiments. A similar pattern was observed in cells expressing GFP alone, which is localized to the cytosol and to the nucleus. This result together with the lack of a potential

targeting sequence in the proteins suggests that G6PD5 and G6PD6 are both likely to be cytosolic proteins.

3.3.2 Generation of the double mutant *g6pd5 g6pd6*

The T-DNA insertion lines for the two genes, *G6PD5* and *G6PD6*, did not have any obvious morphological phenotypes (Figure 3.2a). To address whether this was because the two G6PDH isoforms have redundant functions, crosses between the two T-DNA insertion lines were performed to generate plants homozygous for both T-DNA insertions. The insertion sites of the T-DNA and the primers used to genotype the offsprings by PCR are indicated in Figure 3.2b.

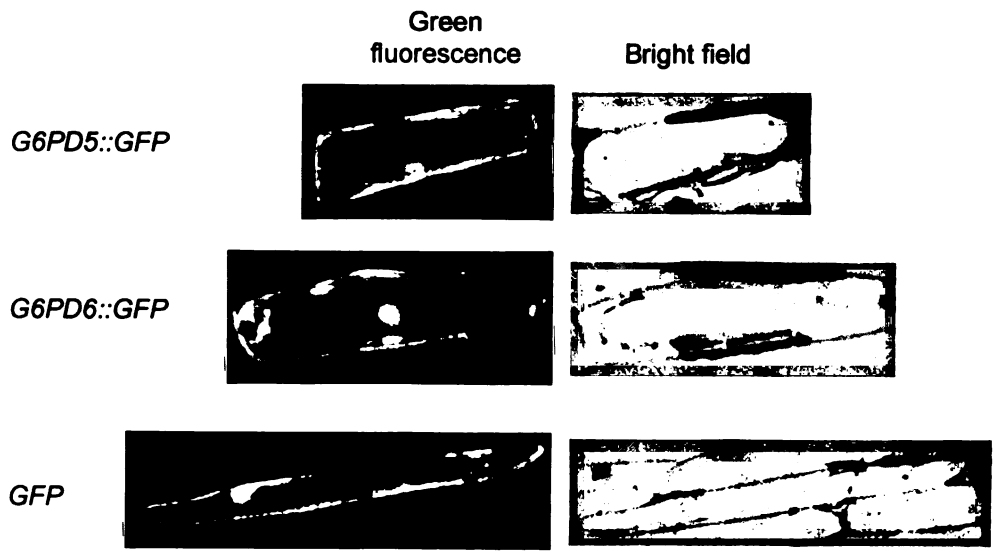
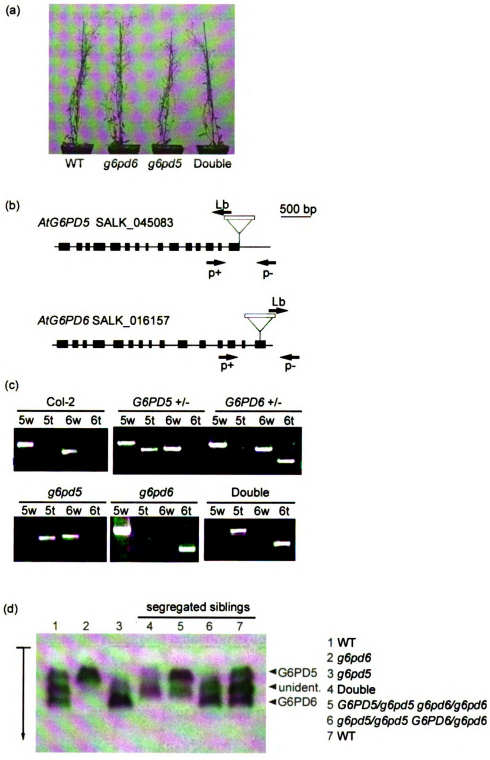


Figure 3.1 Subcellular localization of G6PD5 and G6PD6
Onion cells were bombarded with either pCAMBIA1302 or that inserted with *G6PD5* or *G6PD6* coding sequence. Bright field and fluorescence images are shown.

Figure 3.2 Single mutants *g6pd5, g6pd6* and the generation of a double mutant
(a) Morphological phenotypes of the single and double mutants. (b) Gene structure of *G6PD5, G6PD6* and T-DNA insertion sites. Primers used for genotyping PCR were designed as shown. (c) PCR genotyping. 5w, 6w; gene-specific primer pairs (p+, p-) for *G6PD5* and *G6PD6*; 5t, 6t, T-DNA insertion specific primer pairs, for *G6PD5* (p+, Lb) and *G6PD6* (Lb, p-). (d) G6PDH zymogram patterns of different G6PDH genotypes. The arrow indicates origin and direction of electrophoresis. Arrowheads indicate the position of G6PDH isoforms previously identified (Chapter 2), and the unidentified (unident.) isoform. Genotype nomenclature, *g6pd5, g6pd5/g6pd5 G6PD6/G6PD6*; *g6pd6, G6PD5/G6PD5 g6pd6/g6pd6*; Double, *g6pd5/g6pd5 g6pd6/g6pd6*; *G6PD5* +/-, *G6PD5/g6pd5 G6PD6/G6PD6*; *G6PD6* +/-, *G6PD5/G6PD5 G6PD6/g6pd6*.


(a)
WT
g6pd6
g6pd5
Double
(b)
AtG6PD5 SALK_045083
Lb
500 bp
p+
p-
AtG6PD6 SALK_016157
Lb
p+
p-
(c)
Col-2
G6PD5 +/-
G6PD6 +/-
5w 5t 6w 6t
g6pd5
g6pd6
Double
(d)
segregated siblings
1 2 3 4 5 6 7
G6PD5
unident.
G6PD6
1 WT
2 g6pd6
3 g6pd5
4 Double
5 G6PD5/g6pd5 g6pd6/g6pd6
6 g6pd5/g6pd5 G6PD6/g6pd6
7 WT

All possible genotypes were detected by PCR in the F2 population (Figure 3.2c). From hereon, the single mutants will be referred to as *g6pd5* (*g6pd5/g6pd5 G6PD6/G6PD6*), *g6pd6* (*G6PD5/G6PD5 g6pd6/g6pd6*), double mutant (*g6pd5/g6pd5 g6pd6/g6pd6*) for convenience.

G6PDH activity was examined using zymograms (Figure 3.2d). The zymogram pattern of the plants homozygous for one T-DNA insertion and heterozygous for the other looked identical to the respective single mutants, which was distinguishable by the loss of one band and the induction of another (Figure 3.2d lanes 5 and 6). This result suggests there is no gene dosage effect on cytosolic G6PDH activity. In the double mutant, the induction of both bands was lost, indicating that the intensified bands observed in *g6pd5* and *g6pd6* were indeed G6PD6 and G6PD5. Residual activity of G6PD5 was observed in the double mutant, consistent with previous observations indicating *g6pd5* is not a null mutant, in that the site of the T-DNA insertion is 3 bp downstream of the stop codon (Figures 3.2a, 2.7a). The zymogram pattern observed in the double mutant proves the reciprocal induction of the cytosolic isoforms (and not an induction of an alternative G6PDH).

3.3.3 Gene expression and activity of the cytosolic isoforms and the effect of sucrose

To address whether the reciprocal induction of activity of the cytosolic isoforms occurs at the mRNA level, their gene expression was examined by RT-PCR. To gain insight into their roles, the effect of sucrose (Suc) on their mRNA levels and activity was examined because the transcript of a cytosolic G6PDH in potato was found to increase in response to sugars (Hauschild and von Schaewen, 2003). The 5'-region of the mRNA was

amplified for both genes since the T-DNA insertion in *g6pd5* and *g6pd6* was located within the 3'-end of the genes, which was used to detect mRNA in other RT-PCR experiments. The overall amplification of the 5'-ends was less efficient than that of the 3'-ends (data not shown). This is likely to be due to the incomplete reverse transcription reaction, but the patterns obtained by RT-PCR of the 5'- and the 3'-ends were similar when compared in WT plants (data not shown). The mRNA of *G6PD5* and *G6PD6* was detected at low levels in *g6pd5, g6pd6*, and the double mutant, which may indicate that the T-DNA insertion does not abolish the presence of the mRNA completely, and the region distant from the insertion site was detected (Figure 3.3a). Additionally, the *G6PD5* transcript in mutant *g6pd5* was not as reduced as that for *G6PD6* in mutant *g6pd6.* This is consistent with the fact that *g6pd5* is an incomplete KO mutant, in which a residual activity was detected by zymogram previously (Figure 3.2d and 2.7b). No significant reciprocal induction of the mRNA of the two isoforms was observed (*G6PD6* in *g6pd5,* and *G6PD5* in *g6pd6*) in the presence or absence of Suc (Figure 3.3a). No induction of a single mRNA was observed by addition of Suc was observed (compare + and – Suc Figure 3.3a). Rather, a possible decrease in mRNA was observed in the presence of Suc (Figure 3.3a), and needs further investigation. When G6PDH activity was examined by zymograms in the same seedlings, the reciprocal induction in the single mutants was detected as previously (Figure 3.3b, compare with Figures 3.2d and 2.7b) and the presence of Suc did not affect overall or relative activity of the isoforms (Figure 3.3b). These results exhibit that the reciprocal induction in activity observed when one cytosolic isoform is lost in the single mutants is not accompanied by increased levels of mRNA,

and the presence of Suc does not have a large effect on the activity or mRNA levels of any of the isoforms examined.

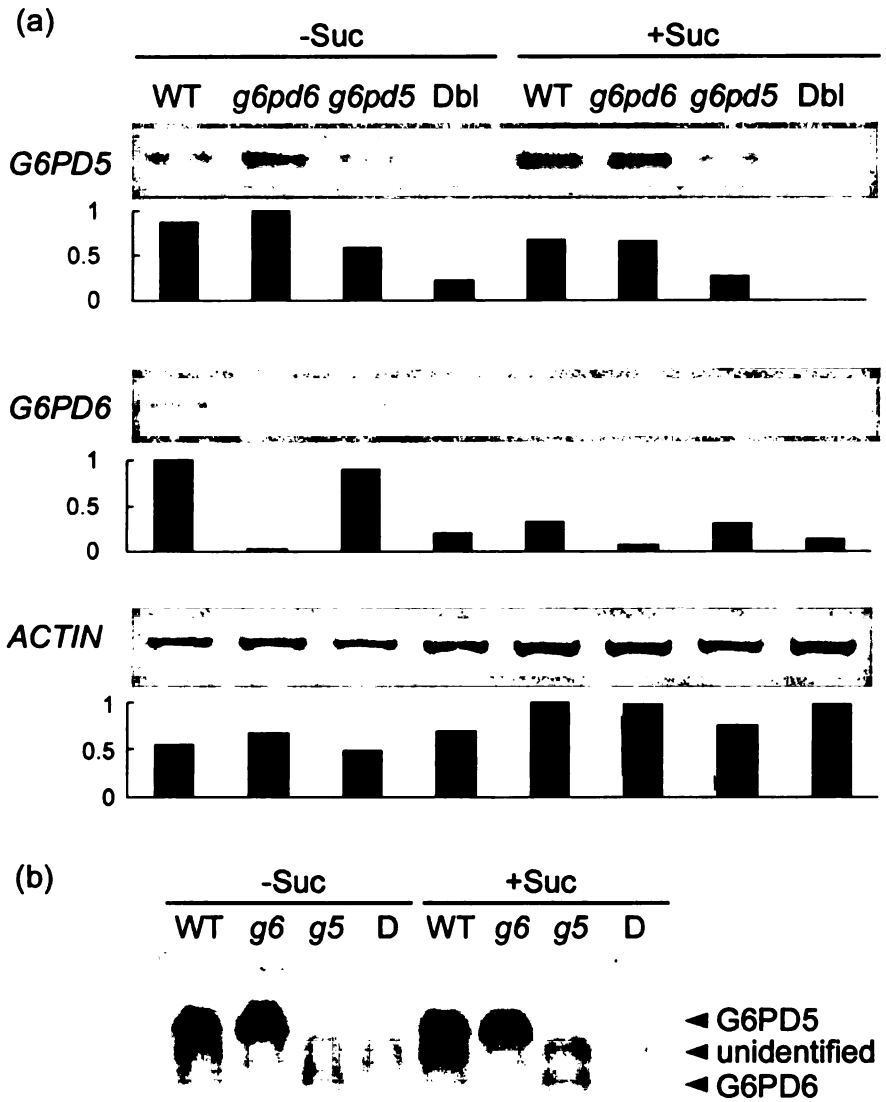


Figure 3.3 Gene expression and activity of cytosolic isoforms in single and double mutants and the effect of sucrose
(a) Levels of mRNA of *G6PD5* and *G6PD6* were analyzed by RT-PCR in the different genotypes. The graphs show the relative signal intensity normalized to that of *ACTIN*. Cycle numbers of 28, 27, and 24 were applied for *G6PD5*, *G6PD6*, and *ACTIN*, respectively. (b) Zymogram of seedlings grown on presence and absence of sucrose. *g5*, *g6pd5*; *g6*, *g6pd6*.

3.3.4 Oxidative stress sensitivity of the mutants

The double mutant lacked any obvious morphological phenotype as did the single mutants (Figure 3.2a). The possible roles of the two cytosolic isoforms under oxidative stress were examined by exposing the plants to DCMU, an uncoupler of the photosynthetic electron transport chain. This was examined in the presence or absence of Suc to test for any correlation between the responses to the two compounds.

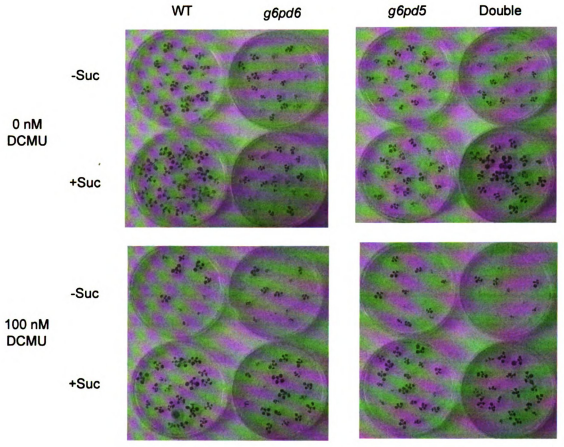


Figure 3.4 Effect DCMU and sucrose on single and double mutants of cytosolic G6PDHs
Seedlings were transferred to DCMU plates at 12 day-old, photographed 15 days after transfer. Genotype nomenclature is as defined in the legend of Figure 3.2.

At 100 mM of DCMU, all genotypes had similarly increased fraction of dead seedlings compared to those grown on plates without DCMU (Figure 3.4). No genotype was visibly more sensitive to DCMU than WT. The same was observed for plants heterozygous one of the T-DNA insertion (data not shown). The presence of Suc alleviated the inhibitory effect of DCMU; the plants of a genotype grown with or without DCMU have similar sizes when grown in the presence of Suc (Figure 3.4). Increased concentrations of DCMU led to increasingly inhibited growth of all genotypes (data not shown). None of the genotypes survived a concentration of 2.5 μM DCMU, but the plants bleached more slowly when Suc was present (data not shown). To test another type of oxidative stress, plants of different genotypes were grown on plates with high salt concentration (150 mM NaCl) but no difference was observed (data not shown). None of the T-DNA insertion lines were more sensitive to oxidative stress than WT plants under the conditions tested.

3.3.5 Expression of *G6PD5*, *G6PD6* cDNAs under the control of 35S promoter does not complement the mutant plants

The first attempt to complement the loss of a G6PDH band on a zymogram in the single mutants, *g6pd5* and *g6pd6,* was carried out with the respective cDNAs of the coding region inserted upstream of a GFP encoding gene in the plasmid pCAMBIA1302, the vectors used for subcellular localization experiments. Subsequently the full-length cDNA was expressed in the single mutants without the GFP gene and driven by the 35S promoter. Transformants were genotyped by PCR to confirm that they are homozygous for the T-DNA insertion, and that they contain the cDNA. Primers complementary to the cDNA will give rise to the unspliced, larger PCR product in WT plants. Protein extracts

from buds, in which the endogenous G6PD6 is active, were analyzed in *g6pd6* transformants transformed with a construct designed to constitutively express the *G6PD6* cDNA (*g6pd6* 35S::cDNA). There was no recovery of the lost band, and the induced G6PD5 band was present as in the non-transformed *g6pd6* plants (Figure 3.5a Buds). Because it is possible that the strength of expression by the 35S promoter differs in tissues, G6PDH activity in leaves was analyzed by zymograms and similar results were obtained (Figure 3.5a Leaf). The transcript from the introduced cDNA was detected by RT-PCR in leaves (the cDNA from the endogenous allele cannot be amplified by PCR due to the T-DNA insertion).

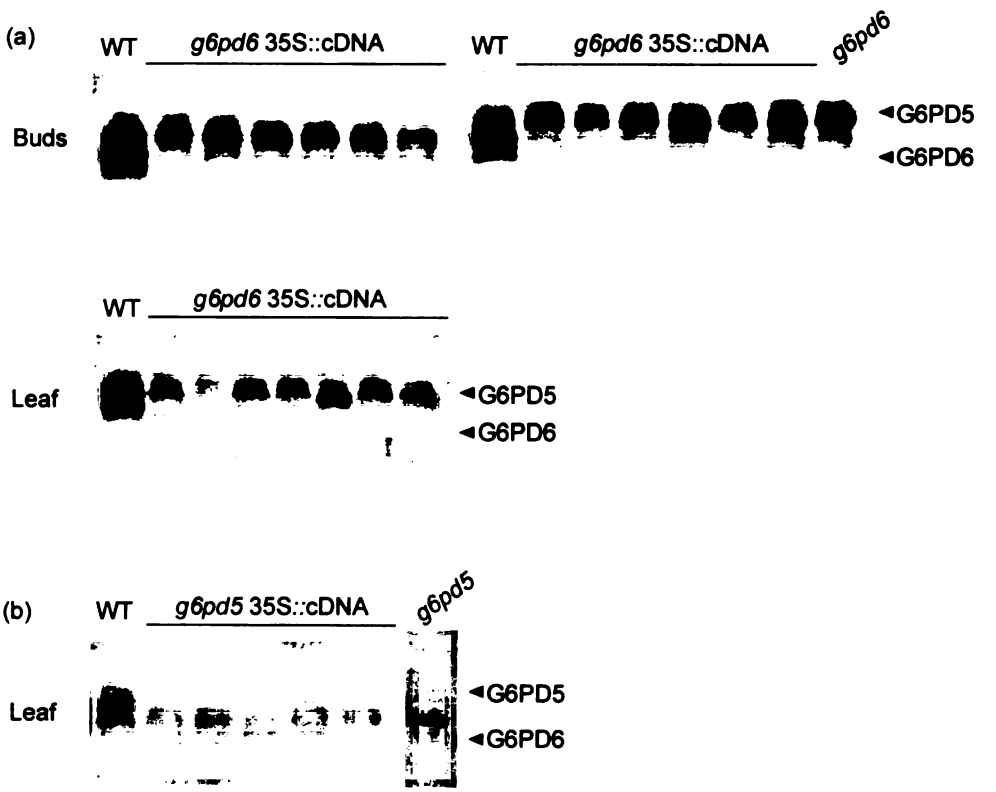


Figure 3.5 Zymogram analysis of single mutant transformed with the respective cDNAs
(a) *g6pd6* transformed with a *G6PD6* cDNA. (b) *g6pd5* transformed with a *G6PD5* cDNA. Each lane represents an individual transformant.

Sequencing results confirmed that no mutations were introduced into the transgene cDNA during the cloning procedure, indicating that it codes for a functional G6PD6 protein (data not shown). Similarly, when the cDNA for *G6PD5* was introduced into *g6pd5* plants, the zymogram pattern was identical to the original mutant (Figure 3.5b). It has been shown in Chapter 2 that the cDNA for *G6PD5* and *G6PD6* can be translated into an active enzyme in *E. coli.* The vectors used in this experiment, pBI121 and pBinAR-Hyg, have successfully resulted in protein production in many studies but it is possible that in the case of Arabidopsis G6PDH genes, the introns and/or untranslated regions (UTRs) contain important information for targeting and translation of the protein. To avoid these problems, complementation with a genomic region that includes the coding sequence (CDS), UTRs, and approximately 1 kb upstream of the transcription initiation site is under progress.

3.3.6 Analyses of T-DNA insertion lines in search for the genes coding for the ubiquitous G6PDH isoform

To analyze the function of the plastidic G6PDH isoforms, and to identify the third band detected on zymograms, T-DNA insertion lines for the predicted plastidic isoforms were examined. The presence of the T-DNA was detected by PCR for each line similarly to *g6pd5* and *g6pd6*. The insertion sites were confirmed by sequencing the PCR products amplified by a gene-specific primer and the T-DNA left border primer. The positions are listed in Table 3.1. In SALK_020160 line the T-DNA insertion was located in an exon of *G6PD3*. When the homozygous plants were analyzed by zymogram, no alteration of the G6PDH patterns was observed (data not shown). Similarly SALK_131208 has the T-

DNA insertion in an exon of *G6PD4*. The homozygous plants had no visible phenotypes either as seedlings on plates with or without Suc or mature plants on soil (data not shown). Zymogram analysis was not carried out for these plants because G6PD4 was found to have little activity (Chapter 2).

For *G6PD1* and *G6PD2* several T-DNA insertion lines were isolated and examined, none of which had T-DNA insertions in the CDS. SALK_19159 has a T-DNA insertion several hundred basepairs (bps) upstream of the CDS of *G6PD1*. In this line, the zymograms were identical to that of WT and the full-length cDNA for *G6PD1* was detected by RT-PCR (data not shown). SALK_19323 was not tested because the T-DNA insertion was reported to be similarly distant from the CDS of *G6PD1* as was the case of SALK_19159. SALK_025911 and SALK_003605 were found to have T-DNA insertion in the same position, several hundred bps upstream of the CDS, and SALK_005846 also have T-DNA insertions several hundred bps downstream of the CDS (Table 3.1). The full-length cDNA for *G6PD2* was detected in both of these lines by RT-PCR (data not shown) and therefore the plants were not analyzed by zymogram. The gene coding for the third band on the zymogram is still unknown. From hereon, this activity represented by the identified band on the zymogram will be referred to as the unidentified (/ubiquitous) isoform.

Because there were no informative T-DNA insertion lines analyzed by zymograms, RNAi lines for *G6PD1* and *G6PD2* are being generated. A 150 base stretch in the 3'-UTR of both genes was targeted for RNAi, using pFGC1285 (Figure 3.6, 3.7). The nucleotide sequences of the two genes are 66% identical (data not shown) and divergent

regions larger than few tens of bases were found only in the 5'-UTR and the 3'-UTR (Figure 3.7).

Gene	T-DNA lines	Clone	Location of CDS	Insertion site	Position relative to CDS
G6PD1	SALK_019159	MIK22 (P1)	7353-4546	7673	320bp upstream of start
	SALK_019323				
G6PD2	SALK_025911	T19L5 (BAC)	32278-34966	31644	634 bp upstream of start
	SALK_003605			31644	634 bp upstream of start
	SALK_005846			35180	214 bp downstream of stop
G6PD3	SALK_020160	F3I6 (BAC)	88807-91696	89586	exon 3
G6PD4	SALK_131208	F14J9 (BAC)	23318-20243	22591	exon 3

Table 3.1 T-DNA insertion lines for predicted plastidic G6PDH isoforms
The SALK lines that were analyzed are listed. The insertion sites are shown by the nucleotide position on a BAC or P1 clone, in which the gene is present.
Genbank accession numbers, MIK22, AB005236; T19L5, AL391711; F3I6, AL002396; F14J9, AC003970.

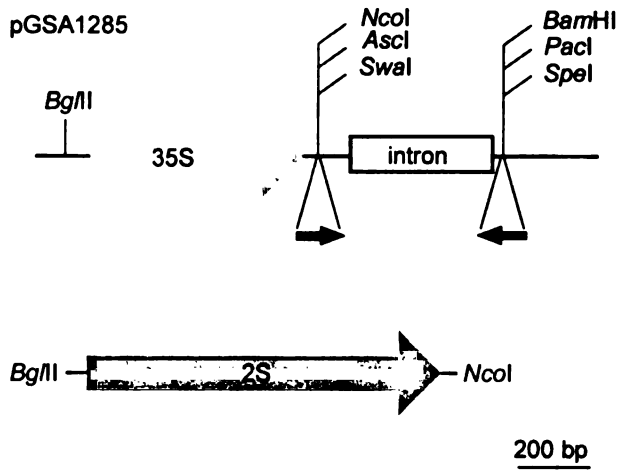


Figure 3.6 Constitutive and seed-specific RNAi vector for *G6PD1* and *G6PD2*
The black arrows indicate the 3'-UTR fragments targeted for RNAi and their orientation and gray arrows indicate promoters. Restriction sites used are shown.

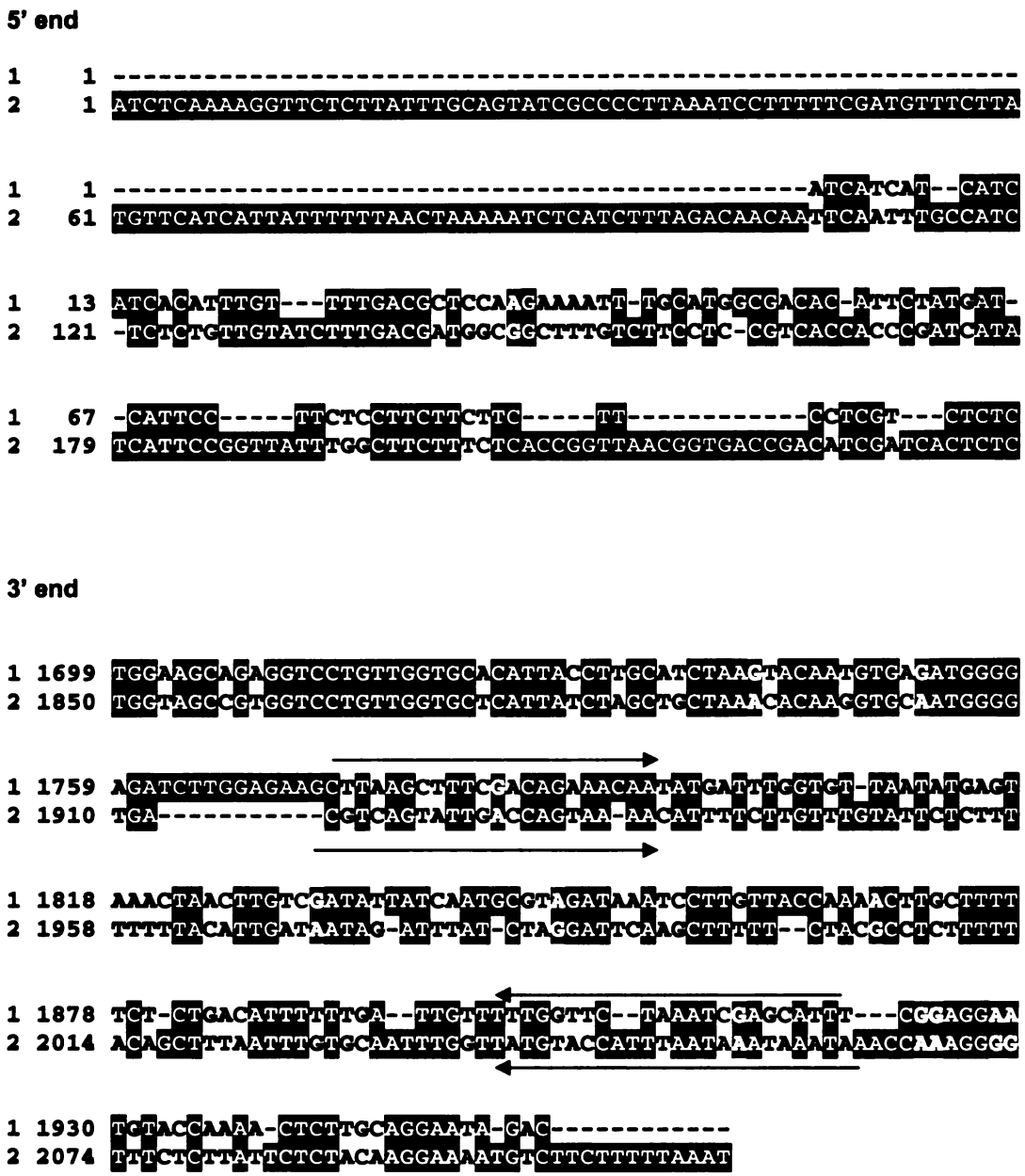


Figure 3.7 5'- and 3'-ends of *G6PD1* and *G6PD2* mRNA
The numbers preceding each row indicates *G6PD1* or *-2*, and the nucleotide positions in each mRNA (Genbank accession nos., *G6PD1*, AY099561; *G6PD2*, AY065042). Arrows indicate the position of the primers that were used to generate the fragment targeted for RNAi.

Seed-specific RNAi constructs were generated by replacing the 35S promoter with an upstream sequence of a 2S albumin storage protein gene from Arabidopsis. Transgenic Arabidopsis expressing a hairpin RNA that targets a stretch of 120 bases in the 3'-UTR of a FA desaturase gene has been reported to reduce enzyme activity as low as a null mutant (Stoutjesdijk *et al.*, 2002). The upcoming transgenic lines are expected to result in plants with reduced activity specifically in G6PD1 or G6PD2, hopefully leading to the identification of the third band on the zymogram, the unidentified/ubiquitous isoform. The contribution of the isoforms in seed oil accumulation is likely to be displayed in the seed-specific RNAi lines.

3.3.7 Roles of G6PDH during seed development

3.3.7.1 Gene expression in developing seeds and siliques

To begin to analyze the roles of G6PDH during seed development, the steady-state mRNA levels of G6PDH isoforms was examined by semi-quantitative RT-PCR. The expression patterns were compared between whole siliques and seeds (Figure 3.8). Whole siliques are easier to harvest especially at younger stages (2-5 DAF) but the RNA extracted from whole siliques represents that from seeds and silique walls, the latter of which takes up a larger portion at younger stages. It was tested whether expression patterns of G6PDH genes in siliques resembled that in seeds to simplify future experiments.

Under the same extraction and RT-PCR conditions, RNA prepared from whole siliques resulted in greater amplification of the G6PDH genes than that from seeds. This could be due to a combination of the transcripts being more abundant in silique walls as

well as the lower quality of seed RNA that presumably contains contamination of storage compounds. Whole siliques of later stages (14-16 DAF) had lower amplification efficiency than the younger siliques, which is also likely to be due to seed storage compounds contaminating the RNA (data not shown). Therefore RNA samples from siliques were analyzed and compared between 2- and 12 DAF (Figure 3.8). In whole siliques, two major patterns were observed. *G6PD1* and *G6PD5* had higher levels of mRNA earlier during development and declined later, while *G6PD2* and *G6PD6* initially had lower levels and increased. *G6PD3* had higher levels between 6 and 10 DAF (mid-stage development) and lower at other times.

In seeds, common expression patterns were not observed among isoforms. *G6PD1* had a transient peak at 9DAF, *G6PD2* was expressed through out development, *G6PD3, -5* had relatively low expression, and *G6PD6* seemed to have the most abundant transcript levels. Transcript for *G6PD4* was almost undetectable in both whole siliques and seeds as was with all other tissues examined in Chapter 2. These differences in the patterns observed in siliques and seeds indicate that gene expression analysis with interest in seeds should be carried out using isolated seeds, although there is the limitation of not being able to analyze young stages. Additionally, large changes in the tissue composition of whole siliques during development complicate the analysis of gene expression. The genes for most G6PDH isoforms were expressed in developing seeds, most abundantly detected being the transcripts of *G6PD2* and *G6PD6* (PCR primers have been shown to have similar amplification efficiencies in Chapter 2). The transient peak of *G6PD1* transcript at 9 DAF coincides with the time of which TAG accumulation begins to occur at a high rate in seeds.

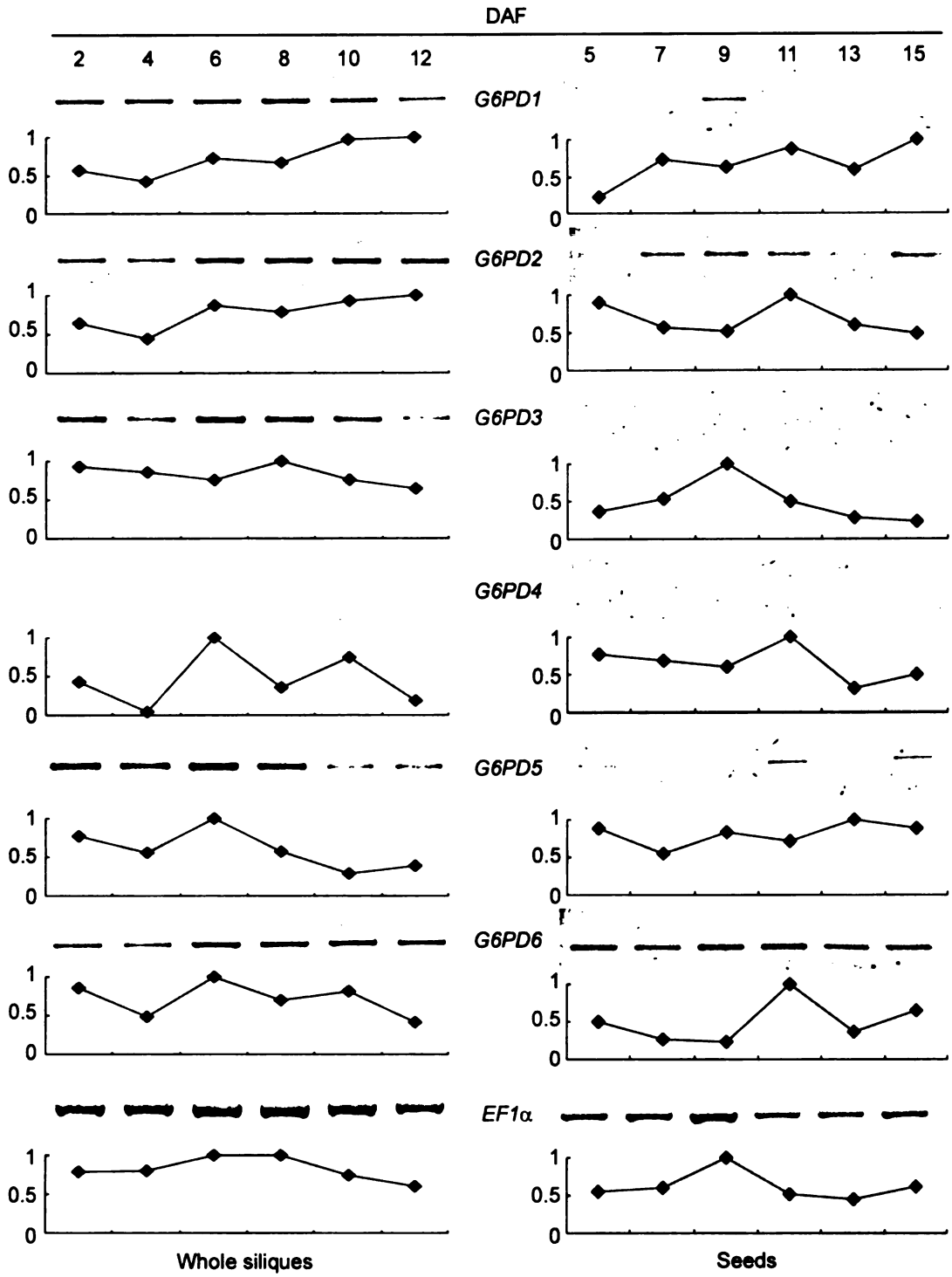


Figure 3.8 Gene expression of G6PDHs in developing siliques and seeds
RT-PCR was carried out as described in Material and methods. 28- and 26 cycles were applied for G6PDH genes and for *EF1α*, respectively. Graphs show the relative values of the intensity normalized to that of *EF1α*.

After this experiment was performed, expression analysis using the Affymetrix ATH1 array of various tissues and developmental stages of Arabidopsis has been reported together with a publicly available database, AtGenExpress atlas (Schmid *et al.*, 2005). The results from that study for the G6PDH genes are shown in Appendix 2. Compared to the result obtained in this study the overall patterns are similar in most tissues as has been mentioned in Chapter 2. However, for seeds there are some inconsistencies which presumably are due to the difficulty in preparation of clean RNA from seed tissues as discussed in this chapter. The data by Schmid *et al.*, (2005) result from triplicates and stringent controls of data analyses, hence likely to be more reliable for gene expression data from seeds. Nonetheless, this section was written for the purpose of documentation.

3.3.7.2 G6PDH activity in seeds

Zymogram analysis in various tissues of Arabidopsis indicated the activity of two isoforms to be present in seeds of 5 DAF, one of which was found to be G6PD6 and the other remains unidentified but presumed to be a plastidic isoform (Figure 2.6a). It had been difficult to reproducibly detect G6PDH activity by zymogram in protein extracts from seeds even when loaded at similar protein concentrations as extracts from other tissues (Chapter 2, Figure 3.9). Interestingly, G6PDH activity could be detected in protein extracts prepared from seeds younger than 5 DAF but not from older seeds. Protein extracts from older seeds showed uneven electrophoresis and smearing of nitroblue tetrazolium (NBT) (Figure 3.9a). This smearing also occurred with the standard G6PDH added to the seed protein extracts (Figure 3.9a), suggesting that the cause is inherent to

the extract. It was suspected that the accumulated storage compounds such as lipids and polysaccharides were interfering with the electrophoresis on CAE. Therefore more diluted extracts were prepared and subjected to isoelectric focusing (IEF), which allows larger volumes to be loaded per sample (15 µl, versus 0.5-1 µl for CAE) to compensate for the diluted activity in the extracts. Similar results were observed with IEF with smearing of NBT in seeds older than 5 DAF (Figure 3.9b). G6PDH activity could not be detected on CAE zymograms when seed extracts were treated with a lipid-removing agent (Cleanascite, CPG, Lincoln Park, NJ), which could be due to dilution of the sample and inactivation of the enzymes by the compound (data not shown). Alternatively, the extracts were eluted through Sephadex using a spin column to remove sugars and gave similar results (data not shown). Only when the seeds were ground in a larger volume of buffer (between 6- to 8-fold (v/w)), but not when the extract was diluted after preparation, was activity detected using a more sensitive fluorescence probe, 5'-cyano-2,3-ditolyl tetrazolium chloride (CTC) (Figure 3.10). Using dilutions of a standard purified G6PDH, 1/10 of activity was detected in zymograms using CTC compared to NBT (data not shown). No fluorescence was detected when G6P was omitted (data not shown). Extraction in small volumes of buffer affected total G6PDH activity as measured by a liquid assay and not only electrophoresis on the CAE plates (Figure 3.10). From duplicate seed samples, the two extraction methods resulted in 10-fold difference in specific activity of total G6PDH detected by liquid assay, but the overall pattern was consistent between the two methods (Figure 3.10). These results suggest that the difficulty in detecting G6PDH activity in CAE zymograms was due to a compound(s) interfering with both electrophoresis and activity. The concentration of the protein extract

does not decrease proportionally to the increase in volume of the extraction buffer used. This suggests that the protein extraction efficiency is improved with the larger volume of extraction buffer. It may have other effects such as solubilization of lipids that solve the problems in electrophoresis and detection of activity.

During seed development, the zymograms and liquid assay showed that G6PDH activity decreases from 5 to 7 DAF and increases again reaching highest activity at 11-13 DAF (Figure 3.10b). No new isoform appeared during seed development and no change in the relative intensity of the two bands on the zymogram was observed (Figure 3.10a). It was predicted that G6PDH activity is lower in greener seeds due to NADPH feed-back inhibition and inactivation by thioredoxin, but in contrast G6PDH activity increased after 7 DAF as seeds become greener (seeds of 5 DAF are still transparent to the eye).

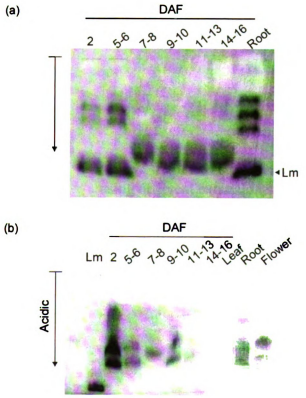


Figure 3.9 G6PDH activity in developing seeds using different zymogram methods
(a) CAE, protein samples were adjusted to 1 mg/ml and loaded approx. 1 µl, arrow indicates origin and direction of electrophoresis. Standard G6PDH from *Leuconostoc mesenteroides* was mixed with tissue protein samples as indicated by arrowhead.
(b) isoelectric focusing gel protein samples were adjusted to 1 mg/ml and loaded 10 µl. Lm, bacterial (*L. mesenteroides*) G6PDH, Whole siliques were extracted for 2 DAF, and the others from isolated seeds.

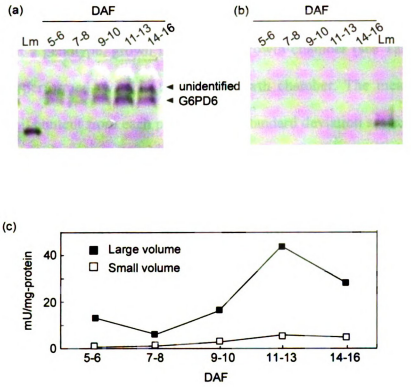


Figure 3.10 G6PDH activity measured in developing seeds using different extraction methods
(a, b) Zymograms. (c) Liquid assays.
Protein extracts were prepared with (a) >6-fold (v/w) buffer and adjusted to 1.8 mg/ml, solid squares; or with (b) approx. 2-fold (v/w) buffer and adjusted to 2.0 mg/ml, open squares. Lm, positive control with G6PDH from *L. mesenteroides*.

3.3.7.3 Oil content in the seeds of cytosolic G6PDH mutants

Oil content in seeds of WT, *g6pd5, g6pd6* and double mutants was measured to examine the contribution of the cytosolic G6PDHs. In a single experiment, a total of approximately 50 seeds were analyzed from a single plant of the four genotypes, two or three plants from a genotype, grown in the same growth chamber (see Table 3.2). This experiment was performed twice in the same growth chamber. The mean of the ~50 values for seed oil content from each plant and the standard deviation are shown in Table 3.2.

To treat these values statistically, the following assumptions were made according to Kimble (1978). It can be assumed that the mean value of the seed oil content of individual plants of the same genotype follows a normal distribution. The mean values from individual plants from Experiment 1 and 2 were pooled together since they originate from plants grown in the same growth chamber with the same conditions and therefore can be assumed as replicates. With these assumptions, the five or six mean values can be treated as random data points that belong to a normal distribution of a population (different plants of the same genotype with various seed oil content means). Therefore, each genotype (group) contains a data set with five (WT, *g6pd5,* and g6pd6) or six (Dbl) values (see numbered plants in Table 3.2).

The data sets were tested for variance with the application ANOVA in Microsoft Excel and http://www.physics.csbsju.edu/stats/. In both analyses, the probability of assuming the null hypothesis (that there is no difference between the data sets) was 0.32. This is not a significant score to accept that the data sets are different: there is a 0.32 probability that the difference in the data occured by random, but does not rule out that

they are different because the failure to reject a null hypothesis does not accept the null hypothesis.

			Oil (µg/seed)	C (%)	N (%)	Weight (mg)				Oil (µg/seed)	Weight (mg)
Exp.1	WT	1	7.24 (0.89)	59.3 (0.32)	3.2 (0.01)	9.5	Exp.2	WT	3	6.82 (0.64)	8.3
		2	6.36 (1.9)			8.0			4	6.73 (0.75)	8.7
									5	7.81 (0.88)	8.6
	g6pd5	1	6.93 (0.87)	60.0 (1.2)	3.4 (0.3)			*g6pd5*	3	7.14 (1.1)	
		2	6.79 (0.08)						4	7.31 (0.69)	
									5	7.80 (0.82)	
	g6pd6	1	7.44 (1.4)	58.9 (0.82)	3.3 (0.3)			*g6pd6*	3	8.83 (1.6)	
		2	6.34 (1.3)						4	7.14 (0.7)	
									5	7.51 (0.90)	
	Dbl	1	7.44 (0.63)	59.2 (0.8)	2.7 (0.1)	9.8		Dbl	4	7.74 (0.80)	8.7
		2	8.45 (1.0)			10.1			5	7.44 (0.69)	9.1
		3	7.12 (1.0)			9.0			6	7.73 (0.89)	9.4

Table 3.2 Seeds of single and double mutants
All the plants were grown in parallel in each experiment (Exp.). The same growth chamber and conditions were applied for the two experiments. Numbers following genotype represent individual plants. Standard deviation is shown in parenthesis. Carbon and nitrogen content is shown in % (w/w). Weight is indicated per 500 seeds.

	Mean	Stdv.
WT	6.99	0.56
g6pd5	7.19	0.39
g6pd6	7.45	0.90
Dbl	7.65	0.45

Table 3.3 Mean of the means of seed oil content
The means and standard deviations (Stdv.) of the mean values of seed oil content shown in Table 3.2.

Student's t-test was performed between WT and the three mutants. By using WT as a reference and analyzing the variance between two groups instead of four, the effect of the variance in the total data set (such as for *g6pd6,* see Table 3.3) is not as prominent. For *g6pd5, g6pd6,* and the double mutant the probability of assuming the null hypothesis (that they are not different from WT) were, 0.52, 0.35, and 0.056, respectively. This result suggests that there may be a difference between WT and the double mutant. Typically 95% or 99% confidence is the cut-off usually applied in rejecting the null-hypothesis (Kimble, 1971) and in this case, 94% confidence (from 1-0.056) is seen in t-test between WT and the double mutant. The confidence of these analyses is determined at least in part by the number of data in the group (genotype), and size of variance in a group (Kimble, 1971) and a larger number of individual plants to obtain more mean seed oil content will strengthen this analysis.

To address whether the possible increase in oil in the double mutant occurred at the expense of storage protein accumulation, the content of total carbon and nitrogen in the seeds was analyzed in experiment 1 (Table 3.2). The double mutant seeds had a lower content of nitrogen (approx. 15%) than the WT seeds while the carbon content was unchanged. Since the elemental analysis was carried out only in seeds from one plant, it cannot be concluded whether this particular to this plant or inherent to the genotype. The seeds of WT and the double mutant were counted 500 each and weighed (Table 3.2). A

correlation between the seed weight and the oil content was observed; seeds with higher oil content mean were heavier (r = +0.70, data not shown). These results suggest that if the seeds of the double mutant indeed contain more oil, it could be due to larger seeds rather than the effect being specific to oil accumulation.

Because FA elongation beyond C_{16} and C_{18} requires NADPH in the cytosol, it was possible that the accumulation of long-chain FAs would be affected in the mutant lines that lack cytosolic G6PDHs. The FA compositions of seeds of the four genotypes closely resembled each other (Table 3.4). Reciprocal induction of the cytosolic G6PDHs was observed in seeds of the single mutant lines (data not shown). Therefore it is possible that in the single mutants, the compensating G6PDH activity was sufficient to support FA elongation and leading to no change in oil content or FA composition. But in the case of the double mutant, this result suggests a reducing equivalent generating process other than G6PDHs exists in the cytosol because the plastidic isoform was observed to be unaltered in non-seed tissues (Figure 3.2d, 3.3b).

genotype	16:0	18:0	18:1	18:2	20:0	18:3	20:1	22:0	22:1	24:1
WT	10.1 (0.9)	4.0 (0.3)	13.9 (1.0)	26.8 (0.7)	2.4 (0.2)	20.9 (0.8)	19.5 (0.6)	0.30 (0.04)	2.2 (0.3)	0.14 (0.03)
g6pd5	9.8 (0.8)	4.0 (0.5)	13.3 (1.6)	26.9 (0.7)	2.5 (0.2)	20.9 (1.1)	19.9 (0.6)	0.32 (0.04)	2.2 (0.3)	0.16 (0.06)
g6pd6	9.9 (0.7)	4.0 (0.3)	13.8 (1.3)	26.7 (0.7)	2.3 (0.2)	21.2 (1.3)	19.4 (0.8)	0.29 (0.03)	2.3 (0.3)	0.15 (0.02)
Double	10.0 (0.8)	4.1 (0.3)	13.8 (1.0)	27.0 (0.9)	2.4 (0.2)	20.9 (1.2)	19.3 (1.5)	0.29 (0.04)	2.0 (0.3)	0.14 (0.04)

Table 3.4 FA composition in WT and KOs for cytosolic G6PDHs
Values indicate mol%-FA. Standard deviation is shown in parentheses.

3.3.8 The role of photosynthesis and G6PDH to provide reducing equivalents for seed oil biosynthesis: The *pds1* mutant

The *pds1* (*phytoene desaturase1*) mutant is defective in a *p*-hydroxyphenylpyruvate dioxygenase, which catalyzes an intermediary step in the synthesis of tocopherols and plastoquinone. Plants homozygous for the mutation are white due to photobleaching (Norris *et al.*, 1995; Norris *et al.*, 1998). In a silique of a heterozygous plant (*PDS1/pds1*), approximately 25% of the seeds are white (Figure 3.11a). We used this system to compare the effect of the presence and absence of photosynthesis on seed oil accumulation. The oil content of the white seeds and WT (Wassilewskija, WS) seeds from 6 DAF to 15 DAF was analyzed by GC when the seed colors were discernible by eye (Figure 3.11b). By 15 DAF, the white seeds accumulated about 75% of the content of the WT seeds indicating that 75% of the oil accumulated without the supply of NADPH from photosynthesis. To address whether the loss of photosynthesis affected G6PDH activity, protein was extracted and subjected to zymogram and liquid assays (Figure 3.11c,d). There was no change in the pattern of G6PDH observed by zymogram. When G6PDH activity was measured in liquid assays, specific activity was slightly lower in the white seeds compared to the green (Figure 3.11d). The *pds1* seeds cannot be distinguished by their colors when they are mature because WT seeds lose chlorophyll and the seed coat is not transparent. The oil content of single seeds from a WT and a *PDS1/pds1* plant was measured (Figure 3.12a). Seeds from WT plants showed a normal distribution with a peak between 8 and 8.5 μg/seed whereas those from the *PDS1/pds1* plant had two peaks, one that overlaps with that of WT and the other with a peak between 5 and 5.5 μg/seed (Figure 3.12a).

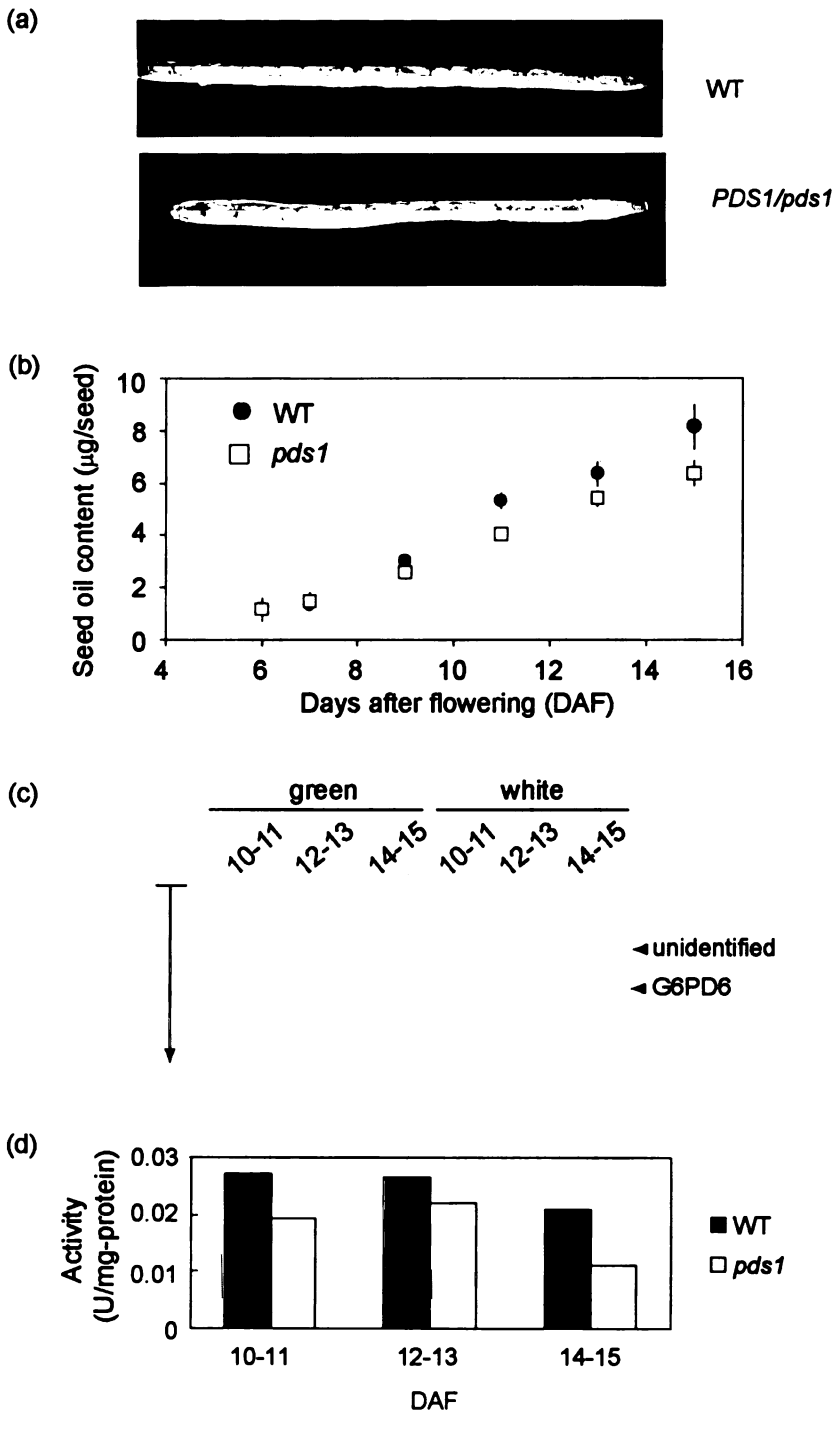


Figure 3.11 Developing *pds1* seeds
(a) Silique of a WT and *PDS1/pds1* plant. (b) White seeds accumulate oil at a lower rate than WT. (c,d) G6PDH activity in WT and white seeds measured by (c) zymogram and (d) liquid assay. Arrow indicates direction of electrophoresis and origin.

There were 49 seeds whose oil content was between 6.5-10.5 μg, and 15 seeds in the range of 4.5-6.5 μg. The ratio between the two populations is close to 3:1. Because *pds1* lacks the plastidic source of NADPH but not that in the cytosol, it was predicted that the relative rates of FA synthesis (plastid) and FA elongation (cytosol) would be affected. The FA composition of seeds was correlated to their total oil content (Figure 3.12b). Seeds with less total oil content, which is presumed to be enriched in *pds1* seeds contained lower C_{16} and C_{18} FAs and higher C_{20} and C_{22} FAs in relative content. Desaturation of FAs also involves a reduction step in regenerating reduced ferredoxin in the plastid (Chapter 1). A difference in the relative amounts of saturated to unsaturated FAs was not detected between seeds with WT levels of oil content and those with less (Figure 3.12b), suggesting the lack of photosynthesis did not specifically affect FA desaturation.

3.4 Discussion

3.4.1 T-DNA insertion lines for cytosolic G6PDHs

The two T-DNA insertion lines *g6pd5* and *g6pd6* did not have any morphological phenotypes or increased sensitivity to the oxidative stress under the conditions tested. DCMU inhibits the photosynthetic electron transport chain by binding to the D1 protein in photosystem II, and this plastid-specific mode of action of the inhibitor may be a reason for the lack of differences between the mutants and the WT in sensitivity. Exposure of the plants to compounds such as H_2O_2, a global oxidative stress compound, or methylviologen, an inhibitor of electron transport (in plastid and in mitochondria) may reveal differences in the genotypes.

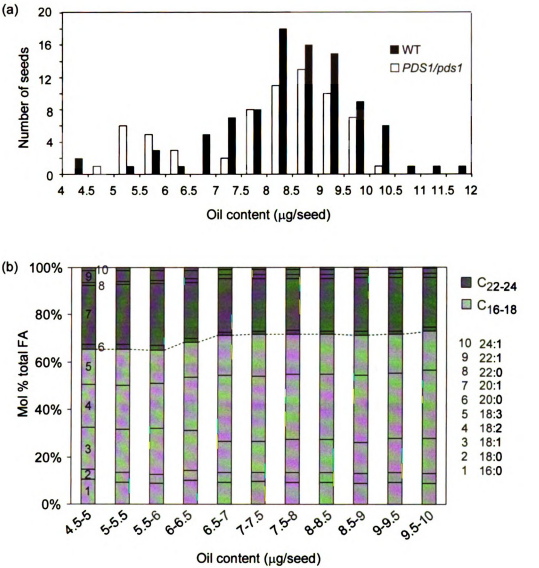


Figure 3.12 Oil in mature seeds from a *PDS1/pds1* plant
(a) Distribution of single seed oil content from WT and *PDS1/pds* plants.
(b) FA composition in seeds with different total oil content.

The lack of genotype-dependent difference in the present study also indicates that the oxidative stress generated in the plastid is not alleviated by the activity of G6PDH in the cytosol under the conditions tested. This suggests that redox levels in separate subcellular compartments are independently maintained to a certain level, assuming that G6PDH is a major player in the determining redox levels. Although shuttling mechanisms such as the triose-phosphate transporter and malate valve have been implicated in the transport of reducing equivalents across the plastidic membrane (Chapter 1), reducing power generated in the form of NADPH by cytosolic G6PDHs did not contribute to the resistance to DCMU in this study. Maize plants disrupted in all of the genes coding for cytosolic 6PGDHs (the second dehydrogenase in OPPP) have reduced nitrogen assimilation rates, suggesting that cytosolic OPPP influences the provision of NADPH in the plastid (Averill *et al.*, 1998). However such influence between the cytosol and the plastid on the provision of NADPH in response to oxidative stress was not observed in the present study.

A coordination of the activity of G6PD5 and G6PD6 was observed in WT plants; only one of the two is present in most tissues except roots (Chapter 2, Figure 2.6a) and in the single mutants, the activity of the remaining cytosolic G6PDH is induced (Figures 3.2d, 3.3b). G6PD6 activity was ubiquitous compared to that of G6PD5, which was leaf-specific (Figure 2.6a). It was surprising to find that the mutant lines had no obvious phenotypes (Figure 3.2a), especially so for *g6pd6*. Reciprocal induction of the two cytosolic G6PDHs suggested the redundancy of the two isoforms and the presence of a common mechanism through which the loss of an isoform is sensed, perhaps through changes in $NADPH/NADP^{+}$. The double mutant with clear reduction in both cytosolic

G6PDH activities still had no obvious phenotypes (Figure 3.2a). Interestingly, the plastidic isoform was not induced even after both cytosolic isoforms are lost (Figures 3.2d, 3.3b), suggesting that the coordination of G6PDH activity does not extend between the two subcellular compartments. The lack of both cytosolic G6PDHs is likely to be compensated by mechanisms other than plastidic G6PDHs, such as other enzymes that could utilize $NADP^+$, e.g. malate dehydrogenase and glycerladehyde-3-phosphate dehydrogenase.

The gene expression patterns and the *cis*-elements of pairs of non-essential paralogs have been examined in yeast (Kafri *et al.*, 2005). This study suggests that a paralog that normally is differentially expressed serves as a backup through transcriptional reprogramming when the other paralog is disrupted. It was proposed that molecular and metabolic changes as a result of one of a paralog being lost, such as the binding of transcription factors to a *cis*-element that is not mostly favored, allow transcriptional reprogramming of the paralog gene. In the present study the gene expression patterns of the two cytosolic G6PDHs did not all correlate with the activity detected in different tissues (Chapter 2). Moreover, in contrast to the transcriptional reprogramming of the paralog model, the reciprocal induction of the isoforms in single mutants was not observed at the level of mRNA (Figure 3.3a). The modes of coordination at the transcriptional and post-transcriptional level are not exclusive. But it is likely that the former does not have a major role in the regulation of the two cytosolic G6PDH isoforms in Arabidopsis, but rather, the activities of G6PD5, resistant to redox changes, and G6PD6, with high specific activity but inactivated by oxidation, are coordinated through changes in metabolite levels. G6PD5 and G6PD6 were found to differ in their kinetic

parameters K_m, V_{max} as well as their sensitivity to redox conditions (Chapter 2) despite the similarity in their sequences (92% amino acid similarity, data not shown). It is tempting to speculate that the substrate (and inhibitor) concentrations in certain tissues are such that only one of the two is active (Figure 2.8), except for roots in which both isoforms were active. However, this is unlikely to be a sole cause for one isoform to be predominantly active in a tissue. Regulation at a level beyond the abundance of mRNA, i.e., translation efficiency, protein stability, or post-translational modification, involving functions of other proteins or changes in metabolite levels, is likely to be playing an important role.

Cytosolic G6PDH has been shown to increase at the mRNA level in response to sugars in mammals and in potato (Hauschild and von Schaewen, 2003; Salati *et al.*, 2004). No increase in mRNA or activity in response to Suc was detected in Arabidopsis (Figure 3.3). Since Suc has been ruled out as an inducing metabolite, it is curious to which signals the two cytosolic isoforms respond, e.g. redox or metabolite levels, perhaps in central carbon metabolism. This could be addressed by measuring G6PDH activity in plants that are disrupted in a cytosolic reducing equivalent generating enzyme as mentioned above, or plants that are perturbed in carbon metabolism such as by disruption of a glycolytic enzyme.

3.4.2 Lack of complementation by a cDNA fused to a GFP gene and the cDNA alone

The complementation, as observed by the recovery of a band detected by zymograms, has not been achieved for both mutants. There were two possible limitations in the approach to complement with a GFP fused protein; first that the fusion protein has a different

isoelectric point resulting in an altered mobility on a zymogram, and second, that the attachment of GFP to a G6PDH may affect the activity because active G6PDHs form a homotetramer (Chapter 2 Discussion and references therein). Unexpectedly, the expression of the cDNA by itself in the mutants did not complement the mutants, despite the expression of the introduced gene. Although the vectors used, pBI121 and pBinAR-Hyg has been used commonly yielding the transgene product, active G6PDH was not recovered. For the subcellular localization experiment, the same cDNAs were expressed fused to a GFP gene in pCAMBIA1302. This vector contains a Kozak consensus to ensure efficient translation of the inserted gene. The lack of active G6PDH originating from the transformed DNA may be due to inefficient translation. The fact that the induction of cytosolic G6PDH activity did not correlate with the mRNA levels in the single mutants indicates the presence of a regulatory mechanism beyond transcript abundance such as translation efficiency of the mRNA and post-translational modification of the protein. It is possible that the 5'-UTR contains information that may increase the protein abundance, and hence the complementation with a genomic fragment of the cytosolic G6PDH genes is under way.

3.4.3 G6PDH activity in seeds

In WT seeds, the activity of two isoforms was detected: G6PD6 and the unidentified isoform (Figure 2.6a and Figure 3.10a). The activity in seeds between 5-16 DAF was lowest at 7-8 DAF and increased thereafter. Before 5 DAF, seeds are non-photosynthetic and exhibit rapid cell division, which is likely to require NADPH for FA synthesis to support membrane biosynthesis. A plastidic G6PDH is speculated to be important during

these stages, but could not be examined due to difficulty in isolating such young seeds in a sufficient amount to perform enzyme assays. Seeds of 5-6 DAF are still transparent to the eye and contained relatively high G6PDH activity (Figure 3.8), which may indicate the importance of G6PDH in the earlier non-photosynthetic stages of seed development. It is possible that the unidentified and presumably plastidic isoform is essential during embryogenesis, and because of this, no KO line was found for *G6PD1* and *G6PD2*. The upcoming RNAi lines contain the transgene controlled by the promoter of a storage protein gene. Conveniently, this construct may avoid such lethality due to the activity profile of the promoter, and allow the analysis of the loss of G6PDH during lipid accumulation, which occurs at overlapping times as that of storage proteins (Figure 1.1) (Ruuska *et al.*, 2002). Alternatively, the RNAi constructs designed to specifically silence *G6PD1* or *G6PD2* may silence both genes due to the similarity in their sequences (Figure 3.7). In such case, the loss of plastidic G6PDH can still be studied in the plant and in the seeds although the gene coding for the unidentified and ubiquitous isoform will remain uncertain.

There were preliminary indications that seeds of the double mutant accumulated more oil than those of WT, which may be accompanied by increase in seed mass and not specifically oil accumulation. This was unexpected since the lack of phenotype in whole plants of the single and double mutants indicated the presence of a compensatory mechanism to supply cytosolic NADPH, and also that the disruption of cytosolic G6PDHs does not affect the provision of plastidic NADPH. Moreover, it was surprising that the loss of cytosolic G6PDHs resulted in increased seed oil rather than the decrease. The composition of FAs was similar among the different genotypes. This suggests that

cytosolic OPPP does not contribute to FA synthesis or FA elongation and possibly in the double mutants, the loss of a cytosolic source of NADPH does not affect the seeds as much as does the alteration of carbon metabolism. The loss of cytosolic G6PDH, the first step in OPPP, may in turn increase the flux into cytosolic glycolysis, leading the generation of carbon substrates for various biosynthetic reactions not only for FAs (see Figure 1.4 for biochemical pathways and subcellular compartments). Assays on the activity for glycolytic enzymes may provide insights. But the limitations would be that the two subcellular compartments cannot be distinguished, and assays on selected enzymes with saturated substrate concentrations (as is the case with most liquid assays) will not represent the flux. It will require metabolic profiling or metabolic flux analyses in order to study the molecular details of the metabolic changes in the double mutant.

A lesson learned was the presence of a great variance between individual plants in contrast to a uniform distribution of seed oil content within a single plant. Future experiments should entail obtaining mean seed oil content from more individuals of the same genotype. Oil measurements of single seeds is not necessary since the strength of the statistical validation will rely on the number of samples in a certain genotype, and pooling a large enough number of seeds should yield reliable seed oil content values. Another important aspect that should be considered is that the seeds analyzed derive from plants that are homozygous for the T-DNA insertion. It is possible that the effect observed is not seed-specific, and is influenced by the alteration in the whole plant metabolism. Since G6PDH activity detected on zymograms suggested that there is no gene dosage effect (Figure 3.2d), a possible approach to address whether the increase in oil is seed-specific would be to analyze the oil content in single seeds from a plant that is

homozygous for T-DNA insertion in one cytosolic G6PDH gene and heterozygous for the other. If the loss of both cytosolic G6PDH activity and not one results in increased oil, as has been hinted in this study, the seed oil would be expected to show a binomial distribution with a ratio of 3:1; only the seeds that are homozygous for both T-DNA insertions will accumulate more oil.

This study has demonstrated the presence of a highly dynamic metabolic network that compensates for the loss of one or both of the cytosolic G6PDHs. In contrast to this flexibility, the network does not extend across subcellular compartments through G6PDH isoforms, at least under the conditions tested. This is suggested from the lack of difference between WT and mutants in response to plastidic oxidative stress and by the uninduced activity of the plastidic isoform in the double mutant. Overall, cytosolic G6PDHs are largely dispensable for plant growth under laboratory growth conditions. However, the increased seed oil content in the double mutant suggests there is indeed a change in metabolism as a result of the disruption in cytosolic G6PDHs, which may be elucidated by applying metabolic profiling and/or metabolic flux analyses. Additionally, whether some of the mutants have increased sensitivity to oxidative stress remains to be addressed using different stress agents that generate oxidative stress in the cytosol.

3.4.4 Contribution of photosynthesis and G6PDH in seed oil

The seeds homozygous for the *pds1* mutation have less oil during development and most likely in mature seeds. The genotypes of the mature seeds with less oil could not be determined because tissue could not be analyzed by PCR for genotype once the seeds were subjected to lipid extraction. In the seeds from a *PDS1/pds1* plant, two phenotypes

are observed in 1:3 ratio, approximately 25% of the seeds contain less oil and 25% of the seeds germinate into albino plants (Norris *et al.*, 1995). This together with the fact that white seeds had accumulated less oil by 15 DAF (Figure 3.12b), suggest that the mature seeds that accumulate less oil (33% reduction, compare two peaks in Figure 3.13a) are likely to be *pds1*. This indicates that the seeds accumulated up to 67% of seed oil in WT without photosynthesis and with WT levels of G6PDH activity (Figure 3.12c,d). The contribution of photosynthesis was examined by covering half of a silique with aluminum foil for 10 hours and seed oil content was compared in Brassica (Ruuska *et al.*, 2004). The FA accumulated in the light was found to be on average 2.5-fold of that in the dark (Ruuska *et al.*, 2004). As another comparison, it has been shown in a metabolic flux analysis of cultured Brassica embryos that OPPP provides 45% of the NADPH required for FA synthesis (Schwender *et al.*, 2003). Photosynthesis is likely to contribute more than just with NAPDH but also ATP, which together supports refixation of carbon as shown by Schwender *et al.*, (2005). Considering this, it is remarkable that the *pds1* mutant seeds still accumulate oil. Due to the lack of photosynthesis the *pds1* seeds are likely to have adapted their metabolic network to support biosynthetic reactions such as with increased OPPP, glycolysis and other enzymes as discussed in Chapter 1. The fact that the *pds1* mutant seeds had WT levels of G6PDH activity leads to the question of how much oil reduction would result from the disruption of the plastidic isoform. The oil may be reduced as much as 70%, or as has been repeatedly observed, the metabolic flexibility may make adjustments just so that the right amount of seed oil is made. This question still remains to be addressed.

The seeds with less oil, the fraction in which *pds1* seeds are presumably enriched, contain higher ratio of long-chain FAs (Figure 3.13b). This suggests that FA synthesis is affected more by the loss of photosynthesis than FA elongation (also shown by Bao *et al.* (1998)), which requires NADPH in the cytosol (Chapter 1). Similarly, the seeds of *wri1* seeds also accumulate higher ratio of long-chain FAs (Focks and Benning, 1998). It has not been examined whether the rate of FA elongation is enhanced in these seeds. The increased elongation of FA (if any) may be a compensating mechanism of seeds with less FAs to produce longer FAs, or to carry out a larger portion of FA synthesis to completion. It would be interesting to compare the relative rates of FA synthesis and elongation by feeding radiolabeled substrates for the two processes in these mutants.

3.4.5 More is not enough

Although the quantity of NADPH supply and oil production has been the focus of this study, the quality of oil is also an important factor that may affect seed viability and longevity. The seeds from a plant heterozygous for *pds1* give rise to albino *pds1* seedlings, but as the seeds are stored for longer periods, the fraction of germinated seeds decreases, along with the appearance of white seedlings (data not shown). The lack of tocopherol may well be the cause of the reduced germination rate as has been shown in *vte1*, a mutant lacking tocopherol that is affected in seed longevity (Sattler *et al.*, 2004). Seeds of *wri1* accumulate a higher ratio of long-chain FAs (Focks and Benning, 1998) and show similar reduction in germination rate over time (A. Cernac, unpublished data). The low longevity of *wri1* seeds could be due to the reduced content of lipids (20% of WT) or the altered FA composition. If the former is true, seeds with higher oil content

such as those of the double mutant could have increased seed longevity. The metabolic processes that the seeds go through during dormancy is largely unknown, but such variations in seed viability and longevity suggest that in the future, it is just as necessary to consider the quality of lipids when engineering the seeds of crop plants.

Reference

Averill, R.H., Bailey-Serres, J., Kruger, N.J. (1998) Co-operation between cytosolic and plastidic oxidative pentose phosphate pathways revealed by 6-phosphogluconate dehydrogenase-deficient genotypes of maize. *Plant J.* **14,** 449-457.

Ayene, I.S., Stamato, T.D., Mauldin, S.K., Biaglow, J.E., Tuttle, S.W., Jenkins, S.F., Koch, C.J. (2002) Mutation in the glucose-6-phosphate dehydrogenase gene leads to inactivation of Ku DNA end binding during oxidative stress. *J. Biol. Chem.* **277,** 9929-9935.

Bao, X., Pollard, M., Ohlrogge, J. (1998) The biosynthesis of erucic acid in developing embryos of brassica rapa. *Plant Physiol* **118,** 183-190.

Becker, D. (1990) Binary vectors which allow the exchange of plant selectable markers and reporter genes. *Nucleic Acids Res.* **18,** 203.

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

Debnam, P.M., Fernie, A.R., Leisse, A., Golding, A., Bowsher, C.G., Grimshaw, C., Knight, J.S., Emes, M.J. (2004) Altered activity of the P2 isoform of plastidic glucose 6-phosphate dehydrogenase in tobacco (*Nicotiana tabacum cv. Samsun*) causes changes in carbohydrate metabolism and response to oxidative stress in leaves. *Plant J.* **38,** 49-59.

Focks, N., Benning, C. (1998) *wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* **118,** 91-101.

Hauschild, R., von Schaewen, A. (2003) Differential regulation of glucose-6-phosphate dehydrogenase isoenzyme activities in potato. *Plant Physiol.* **133,** 47-62.

Kafri, R., Bar-Even, A., Pilpel, Y. (2005) Transcription control reprogramming in genetic backup circuits. *Nat. Genet* **37,** 295-299.

Kletzien, R.F., Harris, P.K., Foellmi, L.A. (1994) Glucose-6-phosphate dehydrogenase: a "housekeeping" enzyme subject to tissue-specific regulation by hormones, nutrients, and oxidant stress. *FASEB J.* **8,** 174-181.

Martini, G., Ursini, M.V. (1996) A new lease of life for an old enzyme. *Bioessays* **18,** 631-637.

Norris, S.R., Barrette, T.R., DellaPenna, D. (1995) Genetic dissection of carotenoid synthesis in arabidopsis defines plastoquinone as an essential component of phytoene desaturation. *Plant Cell* **7,** 2139-2149.

Norris, S.R., Shen, X., DellaPenna, D. (1998) Complementation of the Arabidopsis *pds1* mutation with the gene encoding p-hydroxyphenylpyruvate dioxygenase. *Plant Physiol* **117,** 1317-1323.

O'Neill, C.M., Gill, S., Hobbs, D., Morgan, C., Bancroft, I. (2003) Natural variation for seed oil composition in *Arabidopsis thaliana*. *Phytochemistry* **64,** 1077-1090.

Ruuska, S.A., Girke, T., Benning, C., Ohlrogge, J.B. (2002) Contrapuntal networks of gene expression during Arabidopsis seed filling. *Plant Cell* **14,** 1191-1206.

Ruuska, S.A., Schwender, J., Ohlrogge, J.B. (2004) The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiol* **136,** 2700-2709.

Salati, L.M., Szeszel-Fedorowicz, W., Tao, H., Gibson, M.A., Amir-Ahmady, B., Stabile, L.P., Hodge, D.L. (2004) Nutritional regulation of mRNA processing. *J. Nutr.* **134,** 2437S-2443S.

Salvemini, F., Franze, A., Iervolino, A., Filosa, S., Salzano, S., Ursini, M.V. (1999) Enhanced glutathione levels and oxidoresistance mediated by increased glucose-6-phosphate dehydrogenase expression. *J. Biol. Chem.* **274,** 2750-2757.

Sattler, S.E., Gilliland, L.U., Magallanes-Lundback, M., Pollard, M., DellaPenna, D. (2004) Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. *Plant Cell* **16,** 1419-1432.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., Lohmann, J.U. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nature genetics.* **37,** 501-506.

Schwender, J., Ohlrogge, J.B., Shachar-Hill, Y. (2003) A flux model of glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos. *J. Biol. Chem.* **278,** 29442-29453.

Stoutjesdijk, P.A., Singh, S.P., Liu, Q., Hurlstone, C.J., Waterhouse, P.A., Green, A.G. (2002) hpRNA-mediated targeting of the Arabidopsis *FAD2* gene gives highly efficient and stable silencing. *Plant Physiol* **129,** 1723-1731.

Varagona, M.J., Schmidt, R.J., Raikhel, N.V. (1992) Nuclear localization signal(s) required for nuclear targeting of the maize regulatory protein Opaque-2. *Plant Cell* **4,** 1213-1227.

Vulliamy, T., Mason, P., Luzzatto, L. (1992) The molecular basis of glucose-6-phosphate dehydrogenase deficiency. *Trends Genet* **8,** 138-143.

Yu, B., Wakao, S., Fan, J., Benning, C. (2004) Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in Arabidopsis. *Plant Cell Physiol.* **45,** 503-510.

Chapter 4

Introduction of a bacterial G6PDH into Arabidopsis

4.1 Introduction

The regulatory mechanisms of plant G6PDH include feed-back inhibition by NADPH, and the redox regulation of the plastid isoforms by the thioredoxin-ferredoxin system. Therefore, increasing G6PDH activity through overexpression of an endogenous gene is not a promising strategy in engineering plants to produce more oil, since the enzyme would be subject to the same regulatory mechanisms. Introduction of an enzyme that is not regulated by these mechanisms is a better approach as has been demonstrated in transgenic potatoes. The introduction of a bacterial ADP-Glc pyrophosphorylase that is insensitive to the allosteric effectors, resulted in enhanced enzyme activity and higher starch content in the tubers (Stark *et al.*, 1992).

Some aerobic bacteria lack phosphofructokinase and cannot produce fructose-1,6-bisphosphate in glycolysis. Some of these bacteria strains possess the Entner-Doudoroff pathway that produces pyruvate (Pyr) and glyceraldehyde-3-phosphate (GAP) from glucose-6-phosphate (G6P) (Figure 4.1). G6P is oxidized by G6PDH to 6-phosphoglucono-lactone, then converted to 6-phosphogluconate (6PG) through the same reactions as those in the oxidative pentose phosphate pathway (Chapter 1, Figure 1.5), and dehydrated to 3-keto-2-deoxy-6-phosphogluconate (KDPG) (Figure 4.1). KDPG is then cleaved by aldolase to Pyr and GAP, the latter of which is subsequently converted to Pyr by a subset of the glycolytic reactions. The anaerobe *Zymomonas mobilis* has been extensively studied since the 1970's because of its unique property to ferment ethanol from Pyr and rapidly grow in the presence of high ethanol concentrations. The high ethanol productivity of *Z. mobilis* has been of great industrial interest in terms of alternative fuel and beverage production (Jeffries, 2005) and has led to the recent

sequencing of its genome (Seo *et al.*, 2005). The *Z. mobilis* G6PDH encoded by the *zwf* gene is inhibited by PEP and utilizes both $NADP^+$ and NAD^+ (Scopes, 1997). In contrast to the plant plastidic G6PDHs, its insensitivity to feed-back inhibition by NADPH and redox regulation makes it a good candidate for introduction into plants to enhance G6PDH activity.

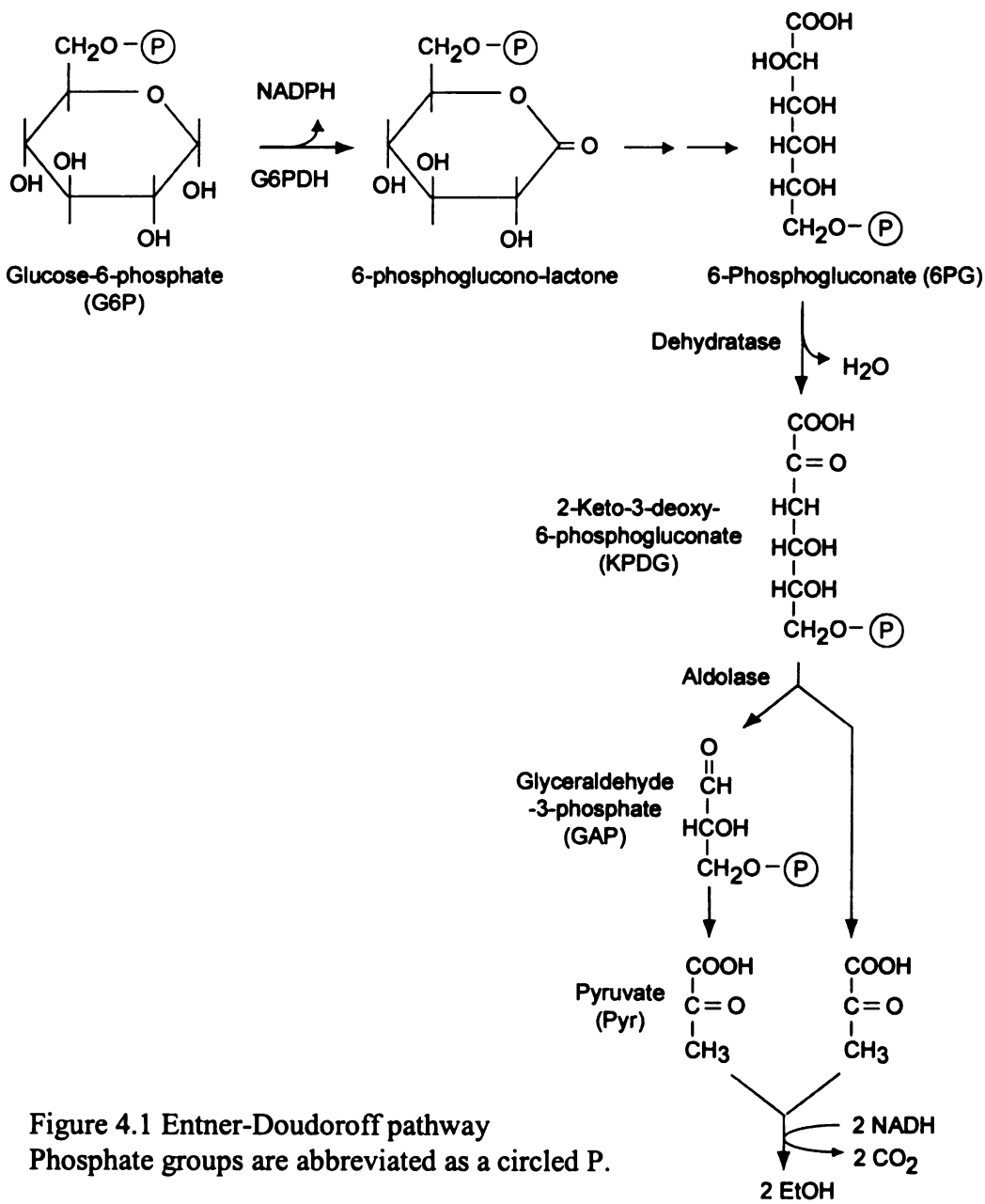


Figure 4.1 Entner-Doudoroff pathway
Phosphate groups are abbreviated as a circled P.

The working hypothesis in this study is that reducing equivalents are limiting for Arabidopsis seed oil accumulation (Chapter 1), and that increasing NADPH production through G6PDH activity will lead to increased seed oil content. Because endogenous G6PDHs are subject to regulatory mechanisms such as feed-back inhibition and inactivation by reduction, a bacterial gene coding for a G6PDH that is insensitive to the regulatory mechanisms was introduced to Arabidopsis to test whether the presence of a deregulated G6PDH activity will boost NAD(P)H production, and hence lead to increased seed oil accumulation. Transformants with low expression level of the bacterial G6PDH gene from *Z. mobilis* were generated, with no remarkable increase in G6PDH activity, or seed oil content. It is speculated that low expression was one of the reasons for the lack of high enzyme activity and seed oil content.

4.2 Material and Methods

4.2.1 Construction of transformation vector and plant transformation[2]

The *zwf* gene was amplified from the genomic DNA of *Z. mobilis* with primers (+) 5'-ATGCAGGTGTGGCCTCCAATGACAAATACCGTTTCG, (-) 5'-AGTCTAGATCAGTCATACCAAGT-3'. The nucleotide sequence of a chloroplast transit peptide of the Rubisco small subunit (RBCS) from Pea was amplified from a plasmid containing the gene, pPsprSS2-1 using primers, (+) 5'-GCTCTAGAAACCACAAGAACTAAGAA-3', and (-) 5'-AGTCTAGATCAGTCATACCAAGT-3'. The *zwf* gene was fused to the target sequence by PCR and subcloned into pPCR-ScriptAmp (Stratagene, La Jolla, CA).

[2] Construction of the transformation vector (USP-pBinAR-Hyg inserted with *RBCS::zwf*), the first transformation, and selection of the initial five transformants were carried out previously by Jamie Hubert.

RBCS::*zwf* was inserted into the *SpeI/SalI* sites of pBinAR-Hyg-USP (Bohmert *et al.*, 1998), in which the 35S promoter of pBinAR-Hyg (Becker, 1990) is replaced with a USP (Unknown Seed Protein) promoter of *Vicia faba* (Baumlein *et al.*, 1991The resulting plasmid was sequenced and confirmed to contain the *RBCS::zwf* cassette. The plasmid was introduced into Agrobacterium and Arabidopsis plants were transformed once by vacuum infiltration (Bechtold *et al.*, 1993) and once by floral dipping (Clough and Bent, 1998). Transformants were selected on MS agar plates supplemented with 25 μg/ml hygromycin.

4.2.2 Expression analysis by semi-quantitative RT-PCR

Semi-quantitative RT-PCR was performed from total RNA of developing seeds as previously described (Yu *et al.*, 2004). The primers used to amplify the transgene mRNA including the targeting sequence (RBCS) are, (+) 5'-GCTCTAGAAACCACAAGAACTAAGAA-3' (5' end of RBCS) and (-) 5'-AGTCTAGATCAGTCATACCAAGT-3' (3' end of *zwf*, Genbank accession no. AF313764). The mRNA for the control gene *EF1α* was amplified with primers described in Chapter 2 (Wakao and Benning, 2005). The PCR products were detected by Southern blotting using the *RBCS::zwf* as a probe as shown in Figure 4.6a. The radiolabeled DNA was detected by exposing the membrane to a phosphorimager screen and scanned with with a phosphorscanner (Storm 820, Molecular Dynamics/Amersham). The intensity of the detected PCR products was quantified by (Image Quant 5.2, Molecular Dynamics).

4.2.3 Protein extraction and G6PDH activity assay

Protein was extracted from developing seeds as previously described (Focks and Benning, 1998). Protein concentration was measured by Bradford assay (Biorad, Hercules, CA). G6PDH activity was measured by a liquid assay as described in Chapter 2.

4.2.4 Seed oil content analysis

Ten developing seeds were harvested from a staged silique (two siliques per plant, two plants per line) and were ground by a glass rod in a glass tube. Ten mature seeds from a single plant were collected and ground by a 6 mm glass bead (Fischer Scientific, Pittsburgh, PA) using a paint shaker for three min, four samples were prepared per line. Lipid extraction and sample preparation for quantification by gas chromatography was carried out as previously described (Focks and Benning, 1998).

4.2.5 Genotype analysis of T-DNA insertion lines

Approximately 10 μg of genomic DNA was digested with 200 U of restriction enzyme overnight, and analyzed by Southern blotting using the same probe as for RT-PCR shown in Figure 4.6a (*RBCS*::*zwf*). (For restriction sites in the T-DNA, also refer to Figure 4.6a).

4.3 Results

4.3.1 Generation of transgenic plants

After the initial transformation by vacuum infiltration, five plants were selected by hygromycin resistance, and will be referred to as lines 1 through 5. In the population of

T_3 generation of line 4, a plant appeared with a pale green phenotype with reticulate leaves and was named line s4 (data not shown). The pale green plants with reticulate leaves grew similarly to WT and gave rise to pale green- and non-green progeny. In a population of plants derived from a green plant of line 4, pale plants appeared at similar rates (data not shown). The segregation of the pale phenotype with the transgene was not analyzed and lines 4 and s4 were treated separately. Hence, six lines of plants that are assumed to originate from five independent transformants were analyzed further in detail. Additionally, five transformants were selected by hygromycin resistance after the second transformation and were named lines 6 to 10.

4.3.2 Gene expression analysis by semi-quantitative RT-PCR

To obtain quantitative information by RT-PCR, different PCR conditions were tested to select a range in which the amount of PCR product is proportional to the amount of the template. Three PCR cycle numbers (18, 21, 24) and four different amounts of starting RNA were tested (200, 400, 800, 1000 ng). When observed on an agarose gel, the PCR for *EF1α* yielded a single band of a size corresponding to a spliced mRNA indicating no residual genomic DNA was present in the prepared RNA (Figure 4.2a, agarose gel for *EF1α*). No bands were visible for *zwf* suggesting that the abundance of *zwf* transcript is lower than that of *EF1α* (Figure 4.2a, *zwf*). After Southern blotting both RT-PCR products were detected and the signal intensity was measured using Image Quant 5.2 (Molecular Dynamics). Figure 4.2b shows the plot of the signal intensity against starting RNA amount. When the plots were fit to a linear regression by minimum mean square method, the R square values were 0.993, 0.983, 0.964 for cycles 18, 21, 24 respectively.

From this result, 18 cycles and starting RNA amount of 300 ng were selected as conditions for further quantitative analysis of RT-PCR. Under identical conditions, the signal intensity detected for *zwf* transcripts was less than 1/50 of those of *EF1α* (data not shown). The blots shown in Figure 4.2a for *EF1α* and *zwf* were exposed for 15 min and 4 hrs, respectively. This together with the lack of a visible PCR product on an agarose gel suggests that the *zwf* transcript is significantly less abundant than that of *EF1α*.

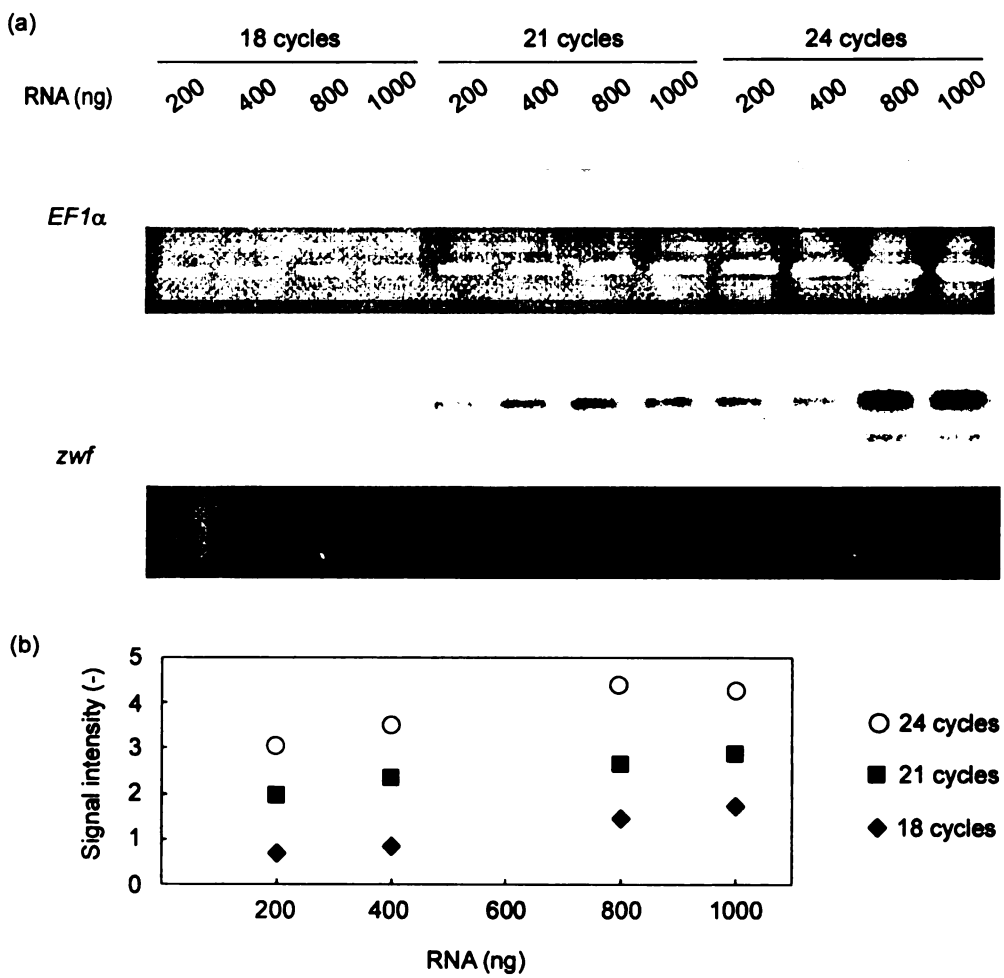


Figure 4.2 PCR product amplification with different cycle numbers
(a) For each gene, above, Southern blot of RT-PCR products and below, agarose gels before transfer to nylon membrane. Phosphoimager screen was exposed 15 min and 4 hours for *EF1α* and *zwf*, respectively.
(b) Signal intensity of *EF1α* calculated using a phosphorscanner.

The levels of *zwf* mRNA was examined during seed development in lines 1 to 5, and s4 (Figure 4.3). The transcripts of *zwf* was detected only in transgenic lines but with much less intensity than that of *EF1*α (Figure 4.3a). When the signal intensities were normalized to that of *EF1*α, lines 3 and s4 showed highest expression at earlier time points while the other lines had higher levels at later stages (Figure 4.3b).

4.3.3 G6PDH activity in developing seeds of transgenic plants

To examine whether the *zwf* transcript was translated to an active enzyme and led to higher total enzyme activity, G6PDH activity was measured in developing seeds (Figure 4.4a). There was no obvious increase in G6PDH activity in the transgenic lines compared to WT at any time of seed development. Because the product of *zwf* is specifically inhibited by PEP (Scopes, 1997), the effect of 1 mM PEP was examined to distinguish the bacterial and endogenous G6PDH activity in seeds of 9-10 DAF. PEP had a similar inhibitory effect on G6PDH activity in WT and transgenic lines (Figure 4.4b) suggesting that the activity detected in transgenic lines was not represented by the bacterial G6PDH.

4.3.4 Seed oil content in transgenic plants

The oil content in developing seeds was measured in transgenic plants. At later stages and in mature seeds, all transgenic seeds had less oil than WT when measured 10 seeds per sample. Lines 4 and s4 had higher oil content than WT at earlier stages (Figure 4.5, 7-8 and 9-10 DAF). The expression of the transgene in line s4 was high at 5-6 DAF but not in line 4.

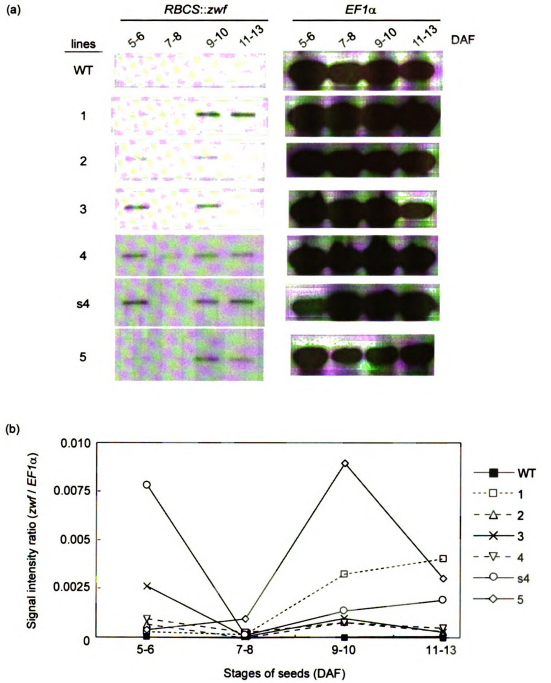


Figure 4.3 Expression of *zwf* in transgenic plants
(a) Southern blot of RT-PCR products.
(b) *zwf* mRNA levels normalized to that of *EF1α*.

The correlation between G6PDH activity and seed oil content could not be measured in younger seeds such as 5 DAF due to the small amount of tissue that could be harvested (Figure 4.4a shows activity from 9 DAF and on).

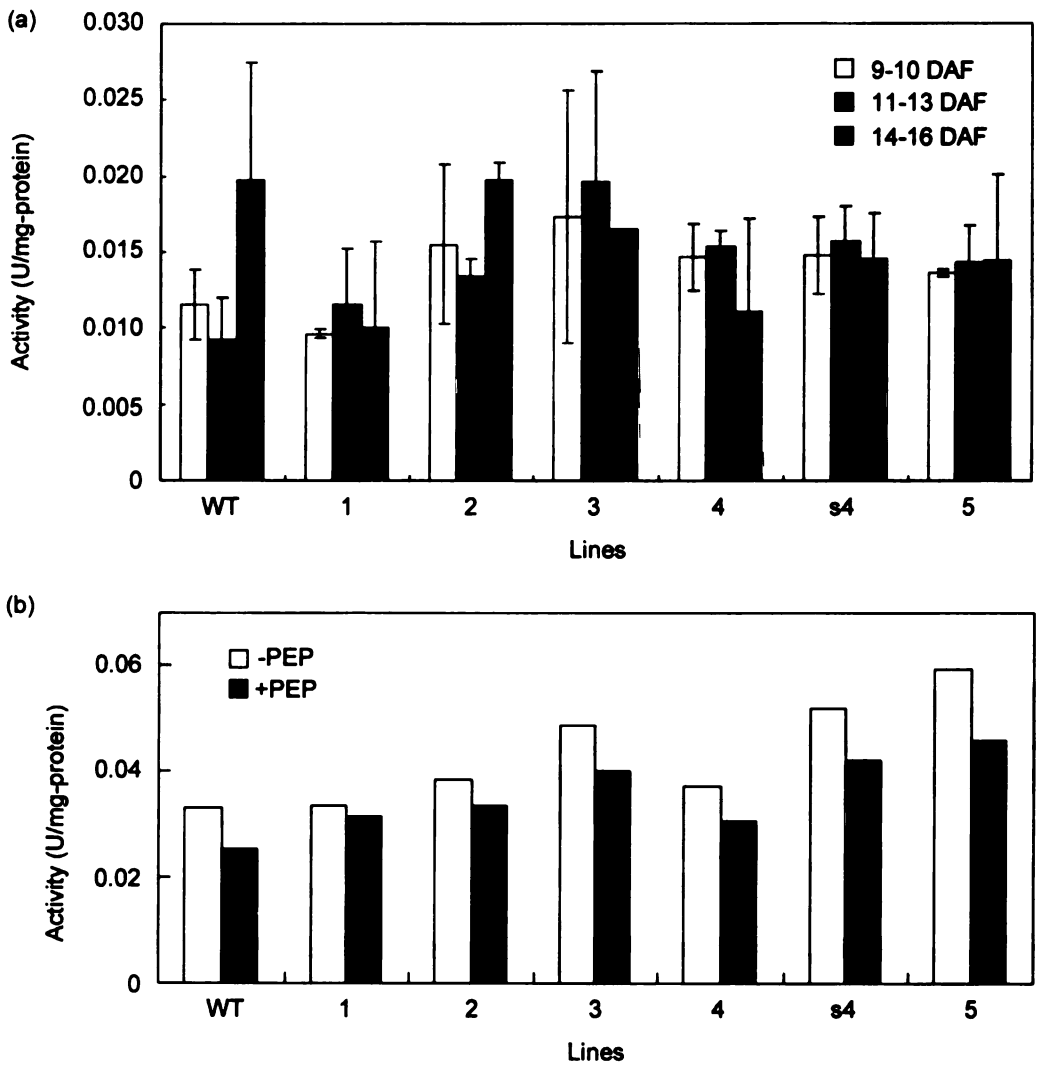


Figure 4.4 G6PDH activity in transgenic lines
(a) G6PDH activity in developing seeds. (b) Effect of PEP on G6PDH activity in seeds of 9-10 DAF. Activity was measured in protein extracts (a) freshly prepared or (b) stored at –80 °C.

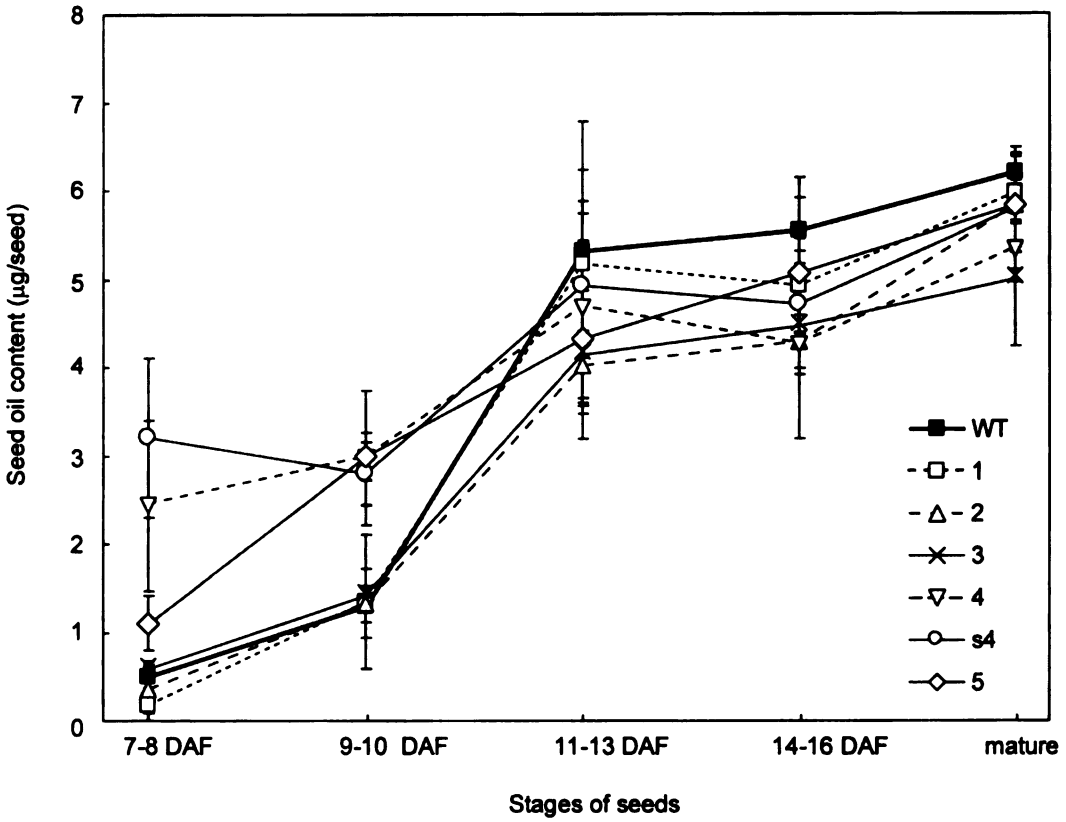


Figure 4.5 Oil accumulation in transgenic seeds

4.3.5 Genotype analysis of transgenic lines

Although the transgene was found to be expressed in the transgenic plants, because none of the transgenic lines had striking seed oil phenotype or large increase in G6PDH activity, their genotypes were examined by Southern blotting. Genomic DNA was extracted from WT and lines 1 to 5 and digested with *Eco*RI and *Sac*I. The expected restriction patterns of the T-DNA are shown in Figure 4.6a. No fragment was found to hybridize with the probe in WT in either blot (Figure 4.6b). All transgenic lines except for line 1 showed a single band slightly larger than 5700 bp when digested with *Sac*I (Figure 4.6b). In line 1, an additional band was seen between 5700 and 4500 bp (Figure 4.6b *Sac*I), possibly due to a second T-DNA inserted or scrambling of DNA accompanying the insertion events. When Southern analysis was performed with genomic DNA digested with *Xho*I, a large fragment (>30 kb) was detected in all transgenic lines (data not shown). These results could suggest that the transgenic lines have insertions in similar regions of the genome. When genomic DNA was digested with *Eco*RI, the two expected bands of approximately 1600 bp and 700 bp were both detected in all lines (Figure 4.6b, *Eco*RI). A 900 bp fragment was found in lines 1 and 3 (Figure 4.6b). Line 1 and 4 contained faint unique bands larger than 2500 and 1700 bp, respectively. These could be the genomic fragments that contain the small 100 bp fragment complementary the probe (downstream of the third *Eco*RI site in Figure 4.6a). From these results, line 1, 3, and 4 are likely to be independent lines, and 2 and 5 could not be distinguished. The origin of the 900 bp fragment in line 1 and 3 is unknown, but incomplete digestion of DNA is not likely to be the cause since 200 U were used for digestion. The results indicate that at least 4 independent lines are present.

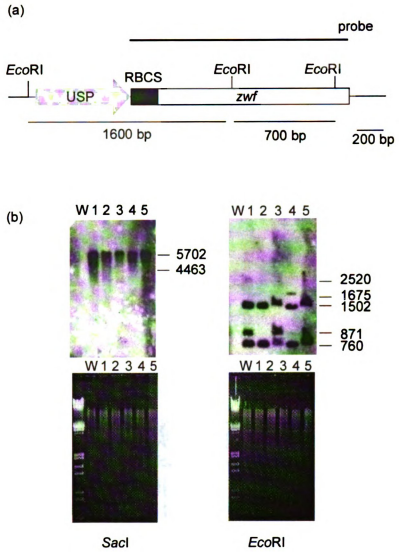


Figure 4.6 Southern analysis of transgenic lines
(a) Cassette containing USP promoter, RBCS::*zwf* gene and the restriction sites. The region used as the probe for Southern analyses is indicated by the dark bar. (b) Top row, autoradiograms; bottom row, agarose gel photos. Numbers next to the autoradiogram indicate expected DNA fragment sizes calculated from the agarose gel. Numbers above lanes indicate transgenic lines. W, WT.

Because it was possible that the actual number of independent T-DNA insertion lines was less than five, additional transgenic plants were generated by Agrobacterium-mediated floral dipping (Clough and Bent, 1998). Five seedlings were selected by hygromycin and grown to maturity (lines 6 to 10). To test for the presence and the abundance of the *zwf* mRNA, RT-PCR was carried out with seed RNA with excess number of cycles (40 cycles). No PCR product could be detected (Figure 4.7), suggesting low levels of transgene mRNA and thus the newly generated transgenic lines were not further characterized.

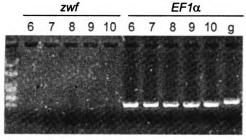


Figure 4.7 *zwf* expression in transgenic lines 6 to 10
Agarose gel photo of the RT-PCR. Numbers indicate individual transgenic lines. g, genomic DNA template.

4.4 Discussion

A gene coding for a bacterial G6PDH was expressed in Arabidopsis plants. Although the transgene was expressed in seeds, no large increase in total- or PEP-sensitive G6PDH activity was observed in the seeds of transgenic plants compared to WT. The oil content of mature seeds was similar in WT and transgenic plants. In this section, the possible limitations of the approach in the present study and potential strategies for future improvement with this approach are discussed.

Assuming similar PCR amplification and probe hybridization efficiencies for *zwf* and *EF1α*, the transgene mRNA had low abundance (less than 1/50 signal intensity of that of *EF1α*). As a comparison, under similar conditions RT-PCR products of endogenous G6PDH genes were visible on an agarose gel to similar intensities with two more cycles than *EF1α* (Chapter 2), which is equivalent to an approximately four-fold difference in signal intensity. Positional effects of the T-DNA insertion may be causing the low abundance of the transgene and perhaps can be solved by more rigorous screening for transformants. Indeed in a study where *Z. mobilis* glucokinase was introduced into potato plants, 50 transformants were screened yielding four lines with increase in glucokinase activity up to four-fold of that in WT (Trethewey *et al.*, 1998).

The addition of untranslated regions (UTRs) and introns of plant genes have been found to be effective in increasing a foreign gene product in plants (Bolle *et al.*, 1996; Koziel *et al.*, 1996; Bourdon *et al.*, 2001; Chaubet-Gigot *et al.*, 2001; Ali and Taylor, 2001). Such modifications include the addition of an intron in the 5'-UTR of a constitutively expressed histone H3 gene from plants (Chaubet-Gigot *et al.*, 2001) or 3'-

UTR of photosynthesis-related gene (Ali and Taylor, 2001). Presumably addition of such DNA increases mRNA stability and/or translation efficiency.

Another common limitation in transgene expression is the difference in the codon usage between the donor and acceptor organisms. A glucokinase from *Z. mobilis* has been introduced to a transgenic potato overexpressing a yeast invertase (Trethewey *et al.*, 1998). The transgenic tubers had increased glucokinase activity and hexose-phosphate pool, indicating the presence of a functional enzyme (Trethewey *et al.*, 1998). Similar to the present study, the unique biochemical properties of *Z. mobilis* glucokinase were exploited, such as low $K_{mGlucose}$ and K_{iATP} values that are not in the range of physiological concentrations in plants (Scopes *et al.*, 1985). Moreover, the codon usage of *Z. mobilis* is more different from potato (*Solanum tuberosum*) than from Arabidopsis (Table 4.1, http://www.kazusa.or.jp/codon/). Taken together, codon usage seems unlikely to be a cause of the lack of G6PDH activity increase in the transgenic Arabidopsis plants. In the potato study, there were no UTRs or introns were added to the coding sequence of the bacterial enzyme but required screening of 50 transgenic plants (Trethewey *et al.*, 1998).

In this study, the expression of *zwf,* G6PDH activity and seed oil accumulation could not be analyzed at the same stages during seed development, mostly due to the limited material Arabidopsis seeds offer. Because enzyme activity assays require larger amounts of tissue, they were performed in the range of 9-16 DAF (Figure 4.8). On the other hand, gene expression was examined in younger seeds, but due to difficulty in preparation of clean RNA from mature seeds, seeds older than 13 DAF were not analyzed (Figures 4.3, 4.8). Seed oil was measured from 7 DAF to mature seeds (Figure 4.8). Therefore the

correlation between the gene expression and G6PDH activity could not be characterized in detail, and it is not clear whether the lack of increase in G6PDH activity is solely due to the inefficiency in transcription and/or that in translation. With low mRNA levels, the temporal changes are also difficult to detect when normalized to a highly abundant transcript.

In the future, improvements to increase mRNA stability and translation efficiency, such as addition of UTRs or introns, could be made to the transformation construct to obtain robust expression and production of the bacterial G6PDH. Primary screening should be carried out with hygromycin resistance followed by a screen of a large number of transformants for plants with increased G6PDH activity. To reduce the number of samples to be assayed for G6PDH activity, seeds should be pooled from 9-11 DAF in which oil accumulates most rapidly and sufficient amount of tissue could be harvested relatively easily.

Organism	GC content	1st letter	2nd letter	3rd letter
Z. mobilis	47.57%	55.93%	41.76%	45.02%
A. thaliana	44.60%	50.88%	40.53%	42.37%
S. tuberosum	42.45%	50.62%	39.00%	37.65%

Table 4.1 GC content and codon usage in *Z. mobilis, A. thaliana, S. tuberosum*
The percentage of GC usage in the 3 letters in a codon is listed.
(http://www.kazusa.or.jp/codon/)

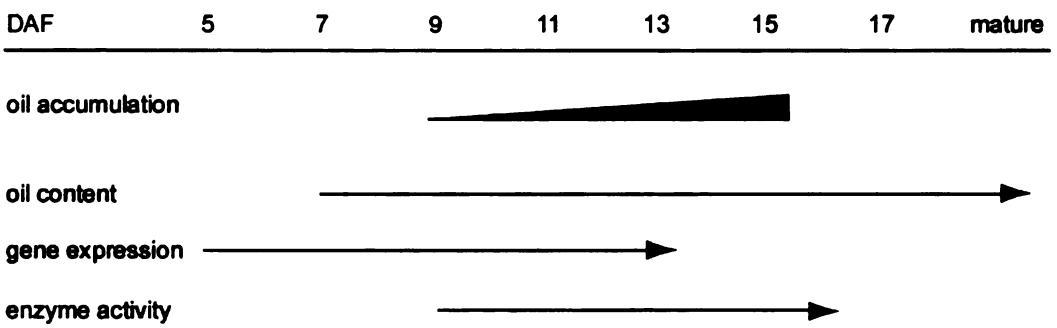


Figure 4.8 Time-frame of experimental procedures in this study

Reference

Ali, S., Taylor, W.C. (2001) The 3' non-coding region of a C4 photosynthesis gene increases transgene expression when combined with heterologous promoters. *Plant Mol. Biol.* **46,** 325-333.

Baumlein, H., Boerjan, W., Nagy, I., Bassuner, R., Van Montagu, M., Inze, D., Wobus, U. (1991) A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and Arabidopsis plants. *Mol. Gen. Genet* **225,** 459-467.

Bechtold N, Ellis.J, Pelletier G (1993) In planta Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. In *Life Sciences 316* Paris: C.R. Acad. Sci., pp 1194-1199.

Becker, D. (1990) Binary vectors which allow the exchange of plant selectable markers and reporter genes. *Nucleic Acids Res.* **18,** 203.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., Benning, C. (1998) *AGO1* defines a novel locus of Arabidopsis controlling leaf development. *EMBO J.* **17,** 170-180.

Bolle, C., Herrmann, R.G., Oelmuller, R. (1996) Different sequences for 5'-untranslated leaders of nuclear genes for plastid proteins affect the expression of the β-glucuronidase gene. *Plant Mol. Biol.* **32,** 861-868.

Bourdon, V., Harvey, A., Lonsdale, D.M. (2001) Introns and their positions affect the translational activity of mRNA in plant cells. *EMBO Rep.* **2,** 394-398.

Chaubet-Gigot, N., Kapros, T., Flenet, M., Kahn, K., Gigot, C., Waterborg, J.H. (2001) Tissue-dependent enhancement of transgene expression by introns of replacement histone H3 genes of Arabidopsis. *Plant Mol. Biol.* **45,** 17-30.

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

Focks, N., Benning, C. (1998) *wrinkled1*: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* **118,** 91-101.

Jeffries, T.W. (2005) Ethanol fermentation on the move. *Nat. Biotechnol.* **23,** 40-41.

Koziel, M.G., Carozzi, N.B., Desai, N. (1996) Optimizing expression of transgenes with an emphasis on post-transcriptional events. *Plant Mol. Biol.* **32,** 393-405.

Scopes, R.K. (1997) Allosteric control of *Zymomonas mobilis* glucose-6-phosphate dehydrogenase by phosphoenolpyruvate. *Biochem. J.* **326,** 731-735.

Scopes, R.K., Testolin, V., Stoter, A., Griffiths-Smith, K., Algar, E.M. (1985) Simultaneous purification and characterization of glucokinase, fructokinase and glucose-6-phosphate dehydrogenase from *Zymomonas mobilis*. *Biochem. J.* **228,** 627-634.

Seo, J.S., Chong, H., Park, H.S., Yoon, K.O., Jung, C., Kim, J.J., Hong, J.H., Kim, H., Kim, J.H., Kil, J.I., Park, C.J., Oh, H.M., Lee, J.S., Jin, S.J., Um, H.W., Lee, H.J., Oh, S.J., Kim, J.Y., Kang, H.L., Lee, S.Y., Lee, K.J., Kang, H.S. (2005) The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nat. Biotechnol.* **23,** 63-68.

Stark, D.M., Timmerman, K.P., Barry, G.F., Preiss, J., Kishore, G.M. (1992) Regulation of the amount of starch in plant-tissues by ADP-glucose pyrophosphorylase. *Science* **258,** 287-292.

Trethewey, R.N., Geigenberger, P., Riedel, K., Hajirezaei, M.R., Sonnewald, U., Stitt, M., Riesmeier, J.W., Willmitzer, L. (1998) Combined expression of glucokinase and invertase in potato tubers leads to a dramatic reduction in starch accumulation and stimulation of glycolysis. *Plant J.* **15,** 109-118.

Wakao, S., Benning, C. (2005) Genome-wide analysis of glucose-6-phosphate dehydrogenases in Arabidopsis. *Plant J.* **41,** 243-256.

Yu, B., Wakao, S., Fan, J., Benning, C. (2004) Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in Arabidopsis. *Plant Cell Physiol.* **45,** 503-510.

Chapter 5

Conclusions and perspectives

5 Conclusions and perspectives

The genome-wide characterization of Arabidopsis G6PDHs yielded information that allowed comparative analysis and predictions of their roles in the plant. The *in vivo* functions of some of the isoforms were addressed directly using the available T-DNA insertion plants. What has been learned from these analyses and perspectives on the functional analyses of the G6PDH isoforms are summarized in this section.

5.1 Six genes, three active isoforms in Arabidopsis

In different tissues of the plant, all except roots were found to contain the activity of two G6PDH isoforms, one cytosolic and one plastidic. In all tissues except roots, one of the two cytosolic isoforms was selectively active; G6PD5 in leaves and G6PD6 in others. Even in roots where photosynthesis is absent, an appearance of a new plastidic isoform was not observed. Similarly in the double mutant lacking both cytosolic isoforms, and in the T-DNA insertion line for *G6PD3*, no new isoform was observed by zymogram. This indicates that a majority of the *in vivo* functions are carried out by these three isoforms.

5.1.1 Coordination of cytosolic isoforms

Interestingly, when one cytosolic isoform was lost in T-DNA insertion plants the activity of the other cytosolic isoform was induced, indicating the existence of a common mechanism through which cytosolic G6PDH isoforms sense the presence or absence of the other. The differences between the two cytosolic isoforms, in terms of their biochemical characteristics, *in vivo* activity, and gene expression patterns have been discussed in detail in Chapters 2 and 3. What was clearly concluded in the present study

is that the activity of the two cytosolic isoforms is not primarily determined by the abundance of the transcript. The question remains; what are the regulatory steps that occur after the mRNA is made that ultimately leads to the presence of an active protein? Several indications hinted to the presence of a post-transcriptional regulation; the lack of correlation between transcript abundance and the activity of an isoform in tissue extracts, and the lack of a reciprocal induction of the mRNA in the single mutants. Possible mechanisms may involve post-translational regulation through protein modification or stability (e.g., G6PD6 is selectively inactivated by oxidation), or changes in metabolite levels that may affect enzyme activity. Because of the high conservation of the two isoforms, transgenic plants containing epitope-tagged version of the isoforms may be an approach to detect protein stability, and to identify other interacting proteins. When carrying out such experiments, the Kozak sequence should be included in the constructs to express G6PDH cDNA (until the *cis*-elements that facilitate translation efficiency are identified), and the tag fused to the protein should be small to prevent inhibition of the assembly of the active homotetramer.

5.1.2 Plastidic isoforms

The intriguing finding based on zymograms that only one plastidic isoform is active throughout the plant, raises the question about the roles of the three other plastidic isoforms. None of the gene expression patterns of the plastidic isoforms match that of the activity detected on zymograms. Analyses of T-DNA insertion plants narrowed the candidates down to G6PD1 and G6PD2.

The P1 plastidic isoform, G6PD1, is the only isoform that clearly has a higher Ki value compared to its Km, indicating that it is less prone to feed-back inhibition. The mRNA is ubiquitously present and to a relatively abundant level in all tissues except in roots. The high Ki suggests this isoform could potentially be active in all photosynthetic and non-photosynthetic tissues, but the low abundance of transcripts in roots argues against this possibility. It will be interesting to find the *in vivo* roles of this isoform, and how its tolerance to higher concentration of NADPH translates to its functions.

The P2 plastidic isoforms, G6PD2 and G6PD3, are highly similar to each other in sequence (63% nucleotide identity, 87% amino acid similarity, data not shown). The expression patterns of the two are similar with higher levels in roots, with higher abundance for *G6PD2* transcript than that for *G6PD3*. The transcripts of both genes are induced by nitrate suggesting some overlap in the roles of the two (Wang *et al.*, 2003). It also indicates that there is a regulation at the mRNA level of G6PDH isoforms. A T-DNA insertion line in *G6PD3* did not result in a loss of a band on the zymogram indicating that G6PD3 is not the isoform detected to be ubiquitous. G6PD1, -2 and -3 are all inactivated by reduction, although the sensitivity could not be compared. The RNAi lines for *G6PD1* and *G6PD2* will demonstrate which gene codes for the ubiquitous plastidic isoform and the characterization of the plants will provide insights into their *in vivo* roles.

G6PD4 was unique from any other G6PDHs. The recombinant protein had little activity when expressed in *E. coli*, which is likely due to the amino acid changes in the G6PDH active consensus site. Phylogenetic analysis showed this isoform to group separately from other G6PDH proteins deposited in the database. Interestingly a G6PDH isoform translated *in silico* from the rice genome was found to have changes in the active

consensus site and grouped relatively close to G6PD4. It is possible that there are genes similar to those coding for the two inactive G6PDHs in other plant species that have not been identified due to the lack of their activity. BLAST analysis using the *G6PD4* transcript against higher plant EST database results in matches to fragments of ESTs too short to analyze similarity to other G6PDH genes (data not shown). It is also possible that such genes have low expression as was observed for *G6PD4* in Arabidopsis, and therefore is under-represented in the EST database. It is puzzling how these genes, *G6PD4* and the rice gene, retained similarity to G6PHs while losing the G6PDH activity. They could possibly belong to a group of genes that code for proteins that have evolved from G6PDHs and acquired new functions (e.g., through binding G6P or other sugar molecules). Plants homozygous for a T-DNA insertion in *G6PD4* did not have any obvious phenotype in the presence or absence of sucrose (data not shown).

5.2 Redox communication between subcellular compartments

The strength of the genome-wide approach was not only that the G6PDH isoforms could be individually characterized but also that the redox communication between the plastid and cytosol could be inferred. The provision of NADPH in the plastid and cytosol seems to be independent of each other under the conditions tested, since the plastidic isoform was not induced in the double mutant lacking cytosolic G6PDHs. It is likely that other cytosolic enzymes that generate NADPH compensate for the loss of both cytosolic G6PDHs instead of the unidentified isoform that is presumed to be plastidic. Another surprising observation was that in *pds1* seeds the activity of both the plastidic isoform and G6PD6 was similar to that in WT. An induction of another isoform (perhaps G6PD5)

was expected since non-photosynthetic roots contained all three isoforms. Moreover, the plastidic isoforms were inactivated by reduction using DTT, hence the lack of photosynthesis was predicted to increase G6PDH activity. These results indicate that the relationship between photosynthesis and G6PDHs (both plastidic and cytosolic) is not as simple as predicted, and the regulation of G6PDH activity could involve tissue-specific metabolites and/or proteins. To start to understand the regulatory mechanisms of G6PDHs as well as their relationship with photosynthetic activity, it is essential to obtain information on the redox levels in subcellular compartments (e.g. by measuring the activation-state of plastidic malate dehydrogenases).

5.3 Role of G6PDH in seeds

Developing seeds have mainly two active G6PDH isoforms, the plastidic isoform and G6PD6. Disruption of G6PD6 or G6PD5 did not result in a large change in seed oil content. This could be explained by the reciprocal induction of G6PD5 and G6PD6 activity, which may have resulted in a similar level of cytosolic G6PDH activity in the seeds of the two single mutants. Surprisingly, seed oil analyses suggested that the seeds of the double mutant contained more oil than those from WT plants without affecting the composition of FAs, and with a possible increase in seed weight. These results implied that the effect of the loss of cytosolic G6PDHs was primarily due to the alteration in carbon metabolism; e.g. increased glycolytic flux leading to increase in substrate for biosynthetic reactions, rather than loss of reducing equivalents supply specifically for oil accumulation. These possibilities need to be verified by enzyme assays and/or radiolabeling experiments to measure metabolic flux. As discussed in Chapter 3, it should

be considered that the seeds come from plants homozygous for the T-DNA insertion, and that the effect may not be seed-specific, but a result of altered metabolism in the whole plant. This should be addressed by oil content analysis of seeds from plants that are homozygous in T-DNA insertion of only one cytosolic G6PDH and heterozygous for the other (see Chapter 3 Discussion).

An interesting question that remains is how metabolism is altered in other (if not all) parts of the plant, although not to an extent to become manifested in morphological phenotypes. The presence of a compensatory mechanism to generate cytosolic NADPH was implied by the lack of a whole-plant phenotype in the double mutant. The dynamic metabolic network was demonstrated in the plants analyzed in this study; the loss of one cytosolic G6PDH isoform is compensated by the induction of the other, and when both are lost, some other process presumably cytosolic, takes over.

The functions of the plastidic G6PDH could not be analyzed due to the lack of T-DNA insertion lines. It will be interesting to examine the loss of the plastidic isoform in the seed-specific RNAi lines of *G6PD1* and *G6PD2.* Loss of photosynthesis was shown to result in about 25% decrease in seed oil in *pds1,* while G6PDH activity analyzed by liquid assay and zymogram, did not greatly differ between WT and *pds1.* Whether the loss of the plastidic isoform reveals a backup system similar to that observed in the cytosolic mutants remains to be seen. However, the fact that the loss of photosynthesis directly results in oil reduction, suggests that a similar outcome can be expected if G6PDH is indeed serving as a major source of NADPH for seed oil accumulation.

Reference

Wang, R., Okamoto, M., Xing, X., Crawford, N.M. (2003) Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiol* **132,** 556-567.

Appendix 1

Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in Arabidopsis[3]

[3] This section describes a collaborative work published, to which I contributed by performing the gene expression analysis. Yu B., Wakao S., Fan J., Benning C. (2004) Loss of plastidic lysophosphatidic acid acyltransferase causes embryo-lethality in Arabidopsis. Plant Cell Physiol. 45, 503-510.

A.1.1 Summary

Phosphatidic acid is a key intermediate for chloroplast membrane lipid biosynthesis. *De novo* phosphatidic acid biosynthesis in plants occurs in two steps: First the acylation of the *sn*-1 position of glycerol-3-phosphate giving rise to lysophosphatidic acid; Second, the acylation of the *sn*-2 position of lysophosphatidic acid to form phosphatidic acid. The second step is catalyzed by a lysophosphatidic acid acyltransferase (LPAAT). Here we describe the identification of the *ATS2* gene of Arabidopsis encoding the plastidic isoform of this enzyme. Introduction of the *ATS2* cDNA into *E. coli* JC 201, which is temperature-sensitive and carries a mutation in its LPAAT gene *plsC*, restored this mutant to nearly wild-type at high temperature. A green-fluorescent protein fusion with ATS2 localized to the chloroplast. Disruption of the *ATS2* gene of Arabidopsis by T-DNA insertion caused embryo lethality. The development of the embryos was arrested at the globular stage concomitant with a transient increase in *ATS2* gene expression. Apparently, plastidic LPAAT is essential for embryo development in Arabidopsis during the transition from the globular to the heart stage when chloroplasts begin to form.

A.1.2 Introduction

The thylakoid membrane of photosynthetic organisms contains four major glycerolipids, mono- and digalactosyldiacylglycerol (MGDG and DGDG), phosphatidylglycerol (PG), and the sulfolipid sulfoquinovosyldiacylglycerol (SQDG). Two pathways contribute to the biosynthesis of membrane lipids in many plants including Arabidopsis (Roughan and Slack, 1982). According to this "two pathway hypothesis", fatty acids *de novo* synthesized in the chloroplast and bound to the acyl carrier protein (acyl-ACP) are either

directly incorporated into glycerolipids in the chloroplast envelopes or are exported to the endoplasmic reticulum (ER). At the ER, they serve in the form of acyl-CoAs as precursors for extraplastidic glycerolipids. A fraction of the diacylglycerol moieties derived from phosphatidylcholine is returned to the chloroplast and enters thylakoid lipid biosynthesis. As a consequence, diacylglycerol (DAG) moieties of plastidic and extra-plastidic origin are found in the glycerolipids of the thylakoid membranes of Arabidopsis.

A critical intermediate of both pathways in plants is phosphatidic acid (PA), which is also essential for the biosynthesis of glycerolipids and triacylglycerol in bacteria, yeast and animals. A deficiency for PA biosynthesis is lethal in *E. coli* consistent with an essential role for PA in this bacterium (Coleman, 1990). Unlike animals and yeast, in plants the glycerol-3-phosphate (G3P) pathway is thought to be the only pathway for the *de novo* PA biosynthesis (Athenstaedt and Daum, 1999). Two acylations are involved: First the transfer of an acyl group from either acyl-ACP or acyl-CoA to the *sn*-1 position of G3P catalyzed by a G3P acyltransferase leading to the formation of 1-acyl-*sn*-G3P (lysophosphatidic acid, LPA). This intermediate is further acylated to PA by an 1-acyl-*sn*-G3P acyltransferase (LPAAT). In plants, LPAAT activity is associated with multiple membrane systems, including chloroplasts, ER, and the outer membrane of mitochondria, suggesting the presence of several different isoforms.

Genes encoding the G3P acyltransferase have been isolated from different plant species. Generally, the *sn*-1 position of PA typically contains 16 or 18 carbons indicating that this enzyme does not discriminate between these two fatty acid substrate classes. However, in some plant species, G3P acyltransferase seems to be more specific. For instance, in Arabidopsis, the ATS1 (ACT1) protein, the plastidic isoform of G-3-P

acyltransferase, has considerable substrate preference for 18-carbon fatty acids (Yokoi *et al.*, 1998). Of the enzymes acting on LPA, the plastidic LPAAT prefers 16-carbon fatty acids and the ER form 18-carbon fatty acids. This substrate specificity provides the means to distinguish thylakoid lipid species derived from the plastid or ER-pathways based on the fatty acids at the *sn*-2 position. Furthermore, the distinct substrate specificities of the different LPAAT isoforms are a critical factor in determining the overall lipid acyl composition in plants. To date, several cDNAs encoding plant LPAATs have been isolated from coconut (Davies *et al.*, 1995; Knutzon *et al.*, 1995), the immature embryo of meadow foam (Brown *et al.*, 1995; Hanke *et al.*, 1995), the maize endosperm (Brown *et al.*, 1994) and from *Brassica napus* (Bourgis *et al.*, 1999). Among these enzymes, BAT2 of *B. napus* has been shown to be a plastid-localized isoform. However, the *in vivo* roles for this enzyme are still unknown due to the lack of lines with altered activity. In Arabidopsis, inactivation of the plastidic G3P acyltransferase in the *ats1* (*act1*) mutant led to the loss of the plastid pathway pathway for glycolipid biosynthesis (Kunst *et al.*, 1988). A surprising observation was that despite the drastic effects on plastidic glycolipid biosynthesis, PG biosynthesis was only mildly impaired in the *ats1* (*act1*) mutant. However, it is not clear whether the currently described alleles for *ats1* (*act1*) are leaky, a caveat that makes it difficult to draw definitive conclusions regarding the function of the enzyme and leaves the interpretation of the mutant phenotype ambiguous. This study was conducted to gain a better understanding of the origin of PA in plants and to determine the *in vivo* function of plastidic LPAAT. As will be described in detail, the respective null-mutant leads to embryo-lethality, an unequivocal demonstration of the essential function of plastidic LPAAT in Arabidopsis.

A.1.3 Material and methods

A.1.3.1 Plant materials and growth

Surface-sterilized seeds were germinated on 0.8% (w/v) agar-solidified Murashige and Skoog (MS) medium (Murashige and Skoog, 1962) containing 1% sucrose. Seedlings (10-days-old) of Arabidopsis wild-type and mutants were transferred to soil drenched with half-strength Arabidopsis nutrient solution (Estelle and Somerville, 1987), and grown under a photosynthetic photon flux density (PPFD) of 70-80 mol m^{-2} s^{-1} at 22/18°C (day/night) with a 14-h light/10-h dark period.

A.1.3.2 Bioinformatics

For routine sequence comparison, BLAST2 was used (Altschul *et al.*, 1997). Multiple sequence alignment was done by CLUSTALW software (Thompson *et al.*, 1994) at The Biology Work Bench (http://workbench.sdsc.edu/). Prediction of chloroplast transit peptides was accomplished using ChloroP (Emanuelsson *et al.*, 1999). Transmembrane spanning helices were predicted using TMHMM (Sonnhammer *et al.*, 1998).

A.1.3.3 T-DNA insertion analysis, *ATS2* cDNA cloning and expression in transgenic plants

The T-DNA insertional mutants, *ats2-1* (SALK_073445) and *ats2-2* (SALK_108812) were obtained from the Arabidopsis Stock Center (www.arabidopsis.org). The T-DNA left border primer P1 5'-GTTCACGTAGTGGGCCATCG-3', and gene-specific primers P2 5'-CAGGTACCTTAGAGATCCATTGATTCTGC-3' and P3 5'-GAGGATCCAGTGAAAAATTTATGGGCGA-3' were used to screen the plants for T-

DNA insertions (cf. Figure A.1.4). The T-DNA-bordering DNA fragments were ligated into the plasmid pPCR-script AMP SK(+) (Stratagene, La Jolla, CA), and sequenced to determine the localization of the T-DNA insertion. The *ATS2* open reading frame corresponding to Arabidopsis gene At4g30580 (GenBank accession no. NP_194787) was predicted from the DNA sequence of BAC F17I23_80 (GenBank accession no. AF160182) and amplified by reverse transcription-PCR using the primers 5'-GAGGATCCATGGATGTCGCTTCTGCTCG-3' and 5'-CAGGTACCTTAGAGATCCATTGATTCTGC-3'. For this purpose, total leaf RNA was isolated from 20-day-old plants (Col-2) according to the method by Logemann *et al.* (1987). Reverse transcription-PCR was performed by using the ProSTAR HF system from Stratagene (La Jolla, CA). The PCR product was inserted into the ligation-ready Stratagene vector pPCR-Script Amp SK(+) giving rise to pATS2. For plant transformation, the insert of pATS2 containing the full-length coding sequence of *ATS2* including the transit peptide was amplified by PCR using the primers 5'-GGACTAGTGATGTCGCTTCTGCTCGGA-3' and 5'-GGACTAGTGAGATCCATTGATTCTGCA-3'. This fragment was inserted into the binary vector pCAMBIA1304 (www.cambia.org) using the *Spe*I restriction site to give plasmid pCATS2. Based on the nature of pCAMBIA1304, this construct led to the expression of an ATS2-GFP (green fluorescent protein) fusion protein and the resulting plants were used in the complementation and localization analysis. Stable transformation of Arabidopsis was achieved using the vacuum infiltration method (Bechtold and Pelletier, 1998). Transgenic plants were selected in the presence of hygromycinB (25 µg/ml) on MS medium lacking sucrose.

A.1.3.4 Semi-quantitative RT-PCR

Total RNA was extracted from plant tissues other than developing siliques with an RNA extraction kit (RNeasy, QIAGEN, Valencia, CA). Silique RNA was extracted from frozen tissue (50-100 mg) which was ground in 1.5mL tubes. To the powder 350 μL of pre-heated (80°C) extraction buffer (0.02 M sodium borate, 30 mM EDTA, 30 mM EGTA, 1% sodium dodecylsulfate, 2% deoxycholate, 2% polyvinylpolypyrrolidone, 2% polyvinylpyrrolidone 40K, 100 mM dithiothreitol, 100 mM β-mercaptoethanol) were added. This mixture was incubated at 80 °C for 1min, then chilled on ice. After addition of 0.15 mg of Proteinase K, the extract was incubated at 37 °C for 1hr. Following the clearing by centrifugation, the supernatant was extracted twice with phenol/choloroform (1:1), and once with chloroform/isoamyl (24:1). Total RNA was precipitated with isopropanol, resolved in water, then precipitated with 2 M LiCl overnight, and washed with 70% ethanol. Following DNase treatment (DNase I, Roche, Indianapolis, IN), cDNA was synthesized from 300 ng of RNA using a reverse transcriptase kit (Omniscript, QIAGEN, Valencia, CA) and 1U of Taq DNA polymerase (Roche, Indianapolis, IN). The 3' region of the *ATS2* gene was amplified for gene expression analysis using the primers 5'-ACGCTAATGGGAACAGGCA-3' and 5'-AAGATCTCAACATTTAATTCTTC-3'. The coding region of a translation elongation factor EF1α (GenBank Accesion no. NM_125432) was used as a control (primers: 5'-ATGCCCCAGGACATCGTGATTTCAT-3' and 5'-TTGGCGGCACCCTTAGCTGGATCA-3') (Boisson *et al.*, 2001). Cycle numbers of 28 and 30 were applied for detection of *EF1α* and *ATS2* respectively. The amplified

fragments were quantified using Quantity One (BioRad, Hercules, CA). Each *ATS2* signal was normalized to the *EF1α* signal for calculation of relative amounts.

A.1.3.5 Heterologous complementation

The *ATS2* open reading frame lacking the predicted transit peptide was generated by RT-PCR using the primers 5'-GAGGATCCAGTGAAAAATTTATGGGCGA-3' and 5'-CAGGTACCTTAGAGATCCATTGATTCTGC-3'. The open reading frame was fused by blunt end ligation to the *lacZ* gene under the control of an IPTG-inducible promoter in the predigested plasmid pPCR-script AMP SK(+) (Stratagene, La Jolla, CA) giving rise to plasmid pATS2-s. The plasmid pATS2-s was used to complement the temperature-sensitive phenotype of *E. coli* JC201 ((Coleman, 1990); genotype: *plsC thr-1 ara-14 Δ(gal-attλ)-99 hisG4 rpsL136 xyl-5 mtl-1 lacY1 tsx-78 eda-50 rfbD1 thi-1*). This strain is unable to grow at 42°C due to the deficiency in lysophosphatidic acid acyltransferase encoded in the wild type by *plsC*, but can grow at 30°C. The control strain JC200 was isogenic to JC201 with exception of the *plsC* wild-type allele in JC200.

A.1.3.6 Microscopy

For the GFP fusion protein localization study, leaf samples of transgenic lines were directly examined using a Meridian Instruments Insight confocal laser scanning microscope (Okemos, MI). Excitation light was provided by an argon laser at 488 nm. GFP fluorescence was observed with a band-pass filter of 520-560 nm and chlorophyll fluorescence with a 670 nm cut-off filter.

Siliques of different developmental stages from heterozygous *ATS2/ats2* plants were dissected with hypodermic needles. Ovules from individual siliques were mounted on microscope slides in a clearing solution (chloral hydrate, water, glycerol, 8:2:1 v/v) and cleared for 1 hour at 4°C before microscopy. The ovule was observed with a Leica DMLB microscope (Leica Microsystems, Wetzlar, Germany).

A.1.4 Results

A.1.4.1 Identification of an Arabidopsis *ATS2* candidate gene

Taking advantage of the published *Arabidopsis* genome, we identified an Arabidopsis gene At4g30580 (GenBank accession no. NP_194787) on chromosome 4 as a putative ortholog of *BAT2* from *B. napus*. A wild-type cDNA corresponding to gene At4g30580 was isolated by reverse transcription-PCR and sequenced. This cDNA encoded 356 amino acids in agreement with the protein predicted from the genomic sequence. The protein was designated ATS2 following the gene designation at TAIR (www.Arabidopsis.org) for the plastidic G3P acyltransferase in Arabidopsis, ATS1(ACT1). Using the BLAST2 sequence alignment software (Altschul *et al.*, 1997), it showed 80% identity, and 85% similarity over 344 amino acid residues with BAT2 (Figure A.1.1a). Like for BAT2, three putative membrane spanning domains were predicted for ATS2 by the TMHMM software (Sonnhammer *et al.*, 1998).

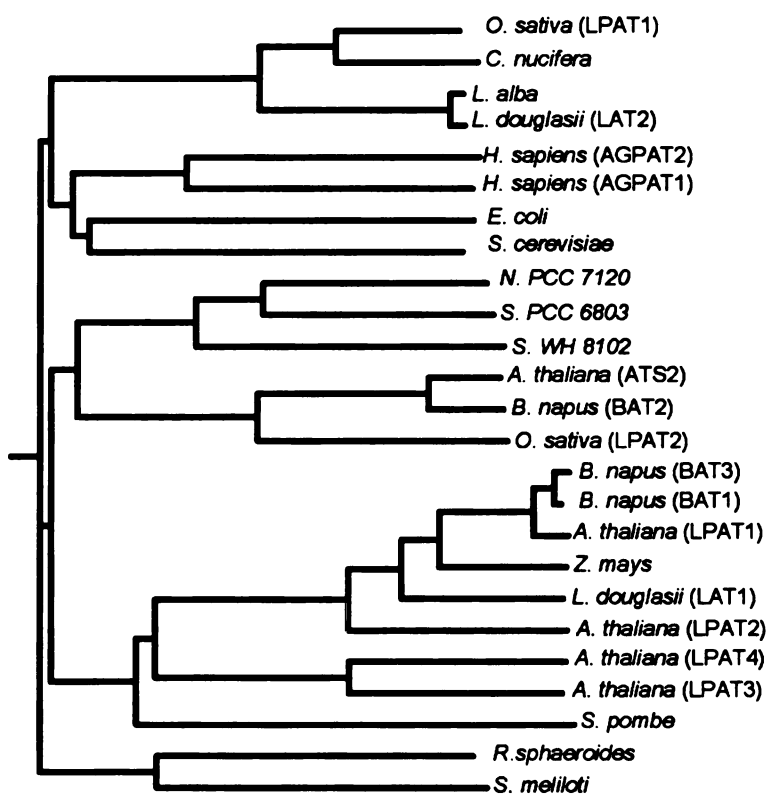
To explore the evolutionary origin of LPAATs, we performed a phylogenetic analysis with ATS2 and other putative and *bona fide* LPAATs.

Figure A.1.1 Evolutionary relationship between LPAATs
Sequence alignment of BAT2 and ATS2 (a). Identical amino acids are indicated by black boxes, similar ones are shaded grey. (b) Phylogenetic relationship of LPAATs. The following proteins were included in the analysis (GenBank accession no.): *Arabidopsis. thaliana ATS2* (NP_194787); *A. thaliana* LPAT1 (NP_567052); *A. thaliana* LPAT2 (NP_175537); *A. thaliana* LPAT3 (NP_565098); *A.thaliana* LPAT4 (NP_188515); *Brassica napus* BAT1 (CAA90019); *Brassica napus* BAT2(AF111161); Brassica napus BAT3 (CAB09138); *Oryza sativa* LPAT1 (CAE03516); Oryza sativa LPAT2 (AAL58271); Limnanthes alba (Q42868); *Limnanthes douglasii* LAT2 (CAA86877); *Limnanthes douglasii* LAT1 (CAA88620); *Cocos nucifera* (Q42670); *Synechococcus sp.* WH 8102 (NP_898339); *Nostoc sp.* PCC 7120 (NP_484285); Synechocystis sp. PCC 6803 (NP_441924); *Saccharomyces cerevisiae* (NP_010231); *Schizosaccharomyces pombe* (NP_595192); *Escherichia coli* (NP_417490); *Homo sapiens* AGPAT1 (AAH03007); *Homo sapiens* AGPAT2 (AAH07269); *Sinorhizobium meliloti* (NP_386764); *Rhodobacter sphaeroides* (ZP_00008109).

A


BAT2 1 MDVASARG SSHP YYSKPICSSQSSLIRI PISKGCCF RSSNL TSLHA
ATS2 1 MDVASARS SSHP SYYGKPICSSQSSLIRI SRDKVCCF RISNG TSFTT
BAT2 51 ASRGVTR------- RTSGVQWCYRSIRFDP VNDKNS------- RT TV
ATS2 51 SLHAVPSEKFMGETRRTGIQWSNRSIRHDPYR FLDKKSPRSSQLARD TV
BAT2 87 RSDLSGAATPESTYPEPEIKLSSRLRGICFCL VAGISAIVLIVLMIIGHP
ATS2 101 RADLSGAATPDSSFPEPEIKLSSRLRGIFFCV VAGISATFLIVLMIIGHP
BAT2 137 FVLLFDRYRRKFHHFIAKLWASISIYPFYKTDIQGLENLPSSDTPCVYVS
ATS2 151 FVLLFDPYRRKFHHFIAKLWASISIYPFYKINIEGLENLPSSDTPAVYVS
BAT2 187 NHQSFLDIYTLLSLGQSYKFISKTGIFVIPVIGWAMSMMGVVPLKRMDPR
ATS2 201 NHQSFLDIYTLLSLGKSFKFISKTGIFVIPIIGWAMSMMGVVPLKRMDPR
BAT2 237 SQVDCLKRCMELVKKGASVFFFPEGTRSKDGRLGPFKKGAFTIAAKTGVP
ATs2 251 SQVDCLKRCMELLKKGASVFFFPEGTRSKDGRLGSFKKGAFTVAAKTGVA
BAT2 287 VVPITLMGTGKIMPTGSEGILNHGDVRVIIHKPIYGSKADVLCEEARNKI
ATS2 301 VVPITLMGTGKIMPTGSEGILNHGNVRVIIHKPIHGSKADVLCNEARSKI
BAT2 337 AESMNLLS
ATS2 351 AESMDL--

B


O. sativa (LPAT1)
C. nucifera
L. alba
L. douglasii (LAT2)
H. sapiens (AGPAT2)
H. sapiens (AGPAT1)
E. coli
S. cerevisiae
N. PCC 7120
S. PCC 6803
S. WH 8102
A. thaliana (ATS2)
B. napus (BAT2)
O. sativa (LPAT2)
B. napus (BAT3)
B. napus (BAT1)
A. thaliana (LPAT1)
Z. mays
L. douglasii (LAT1)
A. thaliana (LPAT2)
A. thaliana (LPAT4)
A. thaliana (LPAT3)
S. pombe
R.sphaeroides
S. meliloti

The experimentally verified LPAATs included enzymes from *Escherichia coli* (Coleman, 1992), human (AGPAT1, AGPAT2) (Eberhardt *et al.*, 1997; West *et al.*, 1997), *Limnanthes douglasii* (LAT1, LAT2) (Brown *et al.*, 1995) and *L. alba* (meadowfoam) (Lassner *et al.*, 1995), coconut (Knutzon *et al.*, 1995), *B. napus* (BAT1 and BAT2) (Bourgis *et al.*, 1999), Yeast (Nagiec *et al.*, 1993), and maize (Brown *et al.*, 1994). Predicted LPAATs were those from *A. thaliana* (LPAT1-4), *Oryza sativa* (LPAT1-2), *B. napus* (BAT3), *Schizosaccharomyces pombe*, the bacteria *Rhodobacter sphaeroides* and *Sinorhizobium meliloti* and different cyanobacteria as indicated. Of these LPAATs, LPAT2 from *O. sativa* is a putative plastidic isoform based on the predicted presence of a chloroplast transit peptide. Figure A.1.1b shows a phylogenetic tree derived from protein sequence alignments. The ATS2 protein clustered with two plastidic LPAATs on a branch alongside the cyanobacterial cluster. ATS2 appears to represent a highly conserved plastidic LPAAT. It shared 68% and 91% identity, respectively, with *O. sativa* and *B. napus* over 207amino acids representing the domain aligning with the bacterial sequence encoded by *plsC*.

A.1.4.2 Expression of the *ATS2* cDNA complements an *E. coli* LPAAT mutant

To demonstrate directly that *ATS2* encodes an LPAAT, we tested if the expression of the *ATS2* cDNA rescued the temperature-sensitive phenotype of *E. coli* strain JC201. This strain is unable to grow at 42°C but grows well at 30°C due to the deficiency of LPAAT activity (Coleman, 1990). Thus, the test is based on the restoration of its growth at the non-permissive temperature. This strategy was successfully used to isolate LPAATs from many other organisms (Brown *et al.*, 1994; Davies *et al.*, 1995; Eberhardt *et al.*, 1997;

West *et al.*, 1997; Bourgis *et al.*, 1999). A truncated version of the *ATS2* cDNA lacking the sequence encoding the putative chloroplast transit peptide was fused to the *lacZ* gene which was under the control of an inducible promoter. The result shown in Figure A.1.2 indicates that the vector expressing *ATS2*, but not the empty vector rescued the temperature sensitivity of *E. coli* JC201. This result suggested that *ATS2* of Arabidopsis indeed encodes an LPAAT that can substitute for the inactive bacterial LPAAT in the mutant.

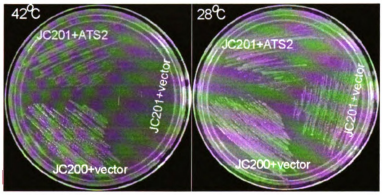


Figure A.1.2 Heterologous complementation of *E. coli* JC201 by expression of *ATS2* of Arabidopsis. Strain JC200 is the *E. coli* wild-type *plsC* control strain isogenic to JC201.

A.1.4.3 Localization of the ATS2 protein

The ATS2 protein sequence has a predicted 56 amino acid N-terminal chloroplast transit peptide (ChloroP) (Emanuelsson *et al.*, 1999). To experimentally verify the subcellular localization of ATS2, the *ATS2* cDNA was fused to the N- terminus of the green fluorescent protein (GFP) cDNA. The resulting construct was expressed in transgenic wild-type plants under the control of the 35S cauliflower mosaic virus (CMV) promoter. The GFP fluorescence in transgenic plants was observed using confocal microscopy. As

shown in Figure A.1.3, the GFP fluorescence was exclusively associated with chloroplasts. In control plants expressing only the GFP cDNA, green fluorescence was not associated with chloroplasts but appeared to be diffuse in the cell. Based on these results, it was concluded that ATS2 is localized in the plastid. These results were also consistent with findings of proteomics studies detecting the presence of this protein in chloroplast preparations (Ferro *et al.*, 2002; Ferro *et al.*, 2003).

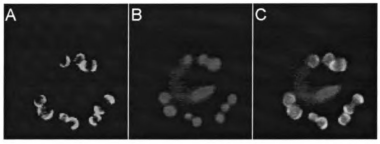


Figure A.1.3 Subcellular localization of the ATS2 protein
(a) ATS2-GFP fusion protein, (b) chlorophyll fluorescence, (c) overlay of (a) and (b).

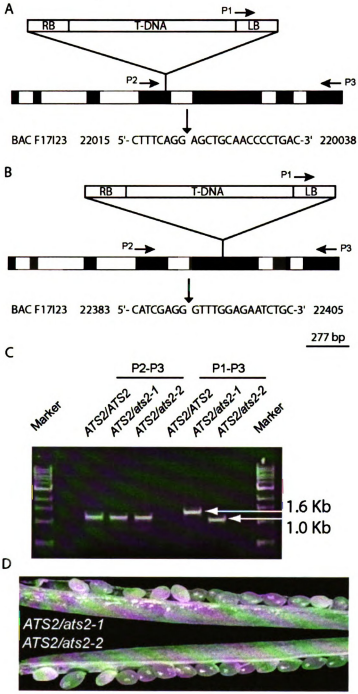
A.1.4.4 Isolation of mutants with T-DNA insertion into *ATS2*

In an attempt to investigate the function of ATS2 in the chloroplast, two T-DNA insertion lines, SALK_073445 and SALK_108812, (alleles *ats2-1* and *ats2-2* respectively) were identified in the Salk T-DNA insertion population (Alonso *et al.*, 2003). The T-DNA insertion sites were determined by sequencing PCR products obtained with a combination of gene- and T-DNA-specific primers as shown in Figure A.1.4. The *ats2-1* allele carried a T-DNA insertion in the fourth exon of the predicted At4g30580 gene (Figure A.1.4a) at base pair 22022 of the sequence of BAC F17I23 (GenBank accession no. AF160182) and the *ats2-2* allele in the fifth exon at base pair 22390 (Figure A.1.4b). Despite extensive

screening, no homozygous plants could be identified. Combinations of gene-specific and T-DNA-specific primers to generate PCR products diagnostic for mutant and wild-type chromosomes and a typical result for wild-type and *ATS2/ats2-1* and *ATS2/ats2-2* heterozygous lines are shown in 4c. Both mutant lines showed insert-specific as well as wild-type genomic DNA fragments. Furthermore, when immature siliques were opened which grew on heterozygous plants as confirmed by genotyping, approximately 25% of the seeds were clear white (Figure A.1.4d), indicative for the arrest of embryo development at an early stage. This observation raised the possibility that a homozygous *ats2* null-mutant was embryo-lethal. To quantify this phenomenon, one or two siliques from 10 heterozygous plants were opened and the ratio of viable to aborted seeds in all the siliques examined was obtained as shown in Table 1. These data were consistent with a 3:1 segregation ratio as expected for a recessive mutation in an essential nuclear gene. When viable plants derived from *ATS2/ats2-1* heterozygous lines or *ATS2/ats2-2* were analyzed, approximately two thirds (39 out of 61 for *ATS2/ats2-1* and 42 out of 62 for *ATS2/ats2-2*) produced aborted seeds consistent with a 2-to-1 ratio of heterozygous to wild-type genotypes among the surviving plants.

Complementation of the embryo-lethal phenotype was tested by introduction of the wild-type *ATS2* cDNA under the control of the 35S-CMV promoter into *ATS2/ats2* heterozygous plants. In the T1 generation, transgenic plants were selected on hygromycinB containing medium and tested for GFP expression. Three classes were expected among the transgenic plants: homozygous mutants (*ats2/ats2),* heterozygous plants (*ATS2/ats2*) and homozygous wild-type plants (*ATS2/ATS2*).

Figure A.1.4 Isolation and characterization of *ATS2* T-DNA insertion lines (a) Insertion into the *ATS2* gene in the heterozygous *ATS2-1/ats2-1* allele and (b) into the *ATS2-2/ats2-2* allele. Black boxes represent exons, open boxes introns. The exact insertion site (based on the sequence of BAC F17I23 GenBank Accession no. AF160182) is indicated by a vertical arrow. Horizontal arrows labeled P1, P2, P3 represent the binding site locations for PCR primers. (c) PCR results for genomic DNA isolated from the wild type (*ATS2/ATS2*) and the two allelic heterozygous mutant lines (*ATS2-1/ats2-1* and *ATS2-2/ats2-2*). The primer combination P2-P3 is diagnostic for the endogenous gene, the combination P1-P3 for the T-DNA flanking genomic DNA as indicated in (a) and (b). (d) Dissected green siliques of heterozygous mutant lines (alleles as indicated) showing the segregation of clear white seeds.


A
P1
RB
T-DNA
LB
P2
P3
BAC F17I23 22015 5'- CTTTCAGG AGCTGCAACCCCTGAC-3' 220038
B
P1
RB
T-DNA
LB
P2
P3
BAC F17I23 22383 5'- CATCGAGG GTTTGGAGAATCTGC-3' 22405
277 bp
C
P2-P3
P1-P3
Marker
ATS2/ATS2
ATS2/ats2-1
ATS2/ats2-2
ATS2/ATS2
ATS2/ats2-1
ATS2/ats2-2
Marker
1.6 Kb
1.0 Kb
D
ATS2/ats2-1
ATS2/ats2-2

Because a single copy of the transgene would behave like a second (dominant) genetic marker in this experiment, complementation would be in effect if homozygous mutant lines segregated 1/4 white seeds, heterozygous lines 1/16 white seeds, and homozygous wild-type lines no white seeds. All three classes were observed among 10 transgenic plants tested, indicating complementation. Because transgenic heterozygous plants were most informative, one plant for each allele was analyzed in detail. Genotyping by PCR (cf. Figure A.1.4c) confirmed these plants to be *ATS2/ats2* heterozygous and analysis of seeds in immature siliques indicated the presence of 1/16 white seeds (Tab. 1) as expected for complementation. Because of the embryo lethality, no further phenotypic analysis was feasible on homozygous lines. When heterozygous lines were analyzed, no changes in growth or leaf fatty acid composition were observed (data not shown).

	Green seeds	White seeds	X^2 (P< 0.1)
ATS2-1/ats2-1	413	135	0.04 (3:1)
ATS2-2/ats2-2	353	120	0.03 (3:1)
ATS2-1/ats2-1 (*ATS2 cDNA*)	498	35	0.09 (15:1)
ATS2-2/ats2-2 (*ATS2 cDNA*)	540	38	0.10 (15:1)

Table A.1.1 Segregation and complementation analysis of the two allelic heterozygous *ATS2/ats2* lines. Multiple siliques from single plants were analyzed. The heterozygous state of *ATS2/ats2* was confirmed by genotyping using PCR (cf. Figure A.1.4c).

A.1.4.5 A transient increase in ATS2 expression in developing siliques in the wild type corresponds to the arrest of embryo development in the mutant

Analyzing the expression of the *ATS2* gene, the mRNA was detected in all tissues tested (Figure A.1.5), as would be expected for a gene involved in an essential process. Interestingly, *ATS2* mRNA abundance was reduced in silique walls compared to leaves

and the abundance was higher in RNA samples from intact siliques. This result suggested an increased expression of *ATS2* in developing seeds. Because it is exceedingly difficult to obtain sufficient amounts of mRNA from very young embryos, a time course of developing intact siliques was analyzed. Comparison to the silique wall sample (5 days after flowering) should provide an indication of mRNA abundance in the developing seed. Interestingly, *ATS2* expression transiently increased in siliques four days after flowering (Figure A.1.5).

Embryo development was compared by Nomarski microscopy in different developing seeds of siliques growing on heterozygous plants (Figure A.1.6). Before the onset of visible chlorophyll accumulation at the transition from globular to the early heart stage (Mansfield *et al.*, 1991), the embryos were indistinguishable.

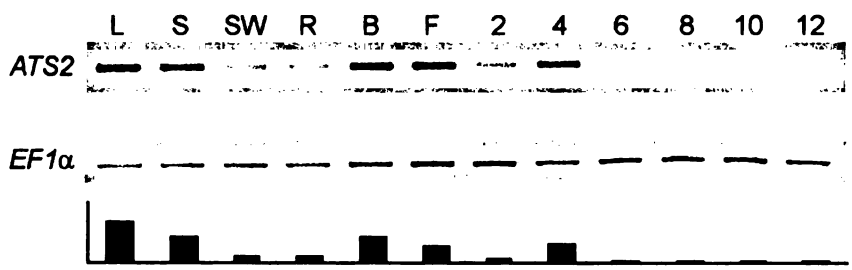


Figure A.1.5 Expression of *ATS2* in various tissues and during silique development
Semi-quantitative RT-PCR was used and products for *ATS2* are shown in comparison to the control gene *EF1α* (GenBank accession no. NM_125432). Relative levels of *ATS2* normalized to *EF1α* are shown in the bottom panel. Lanes from left to right: L, leaf; S, stem; SW, silique wall (age 5 days after flowering); R, root; B, bud; F, flower; 2-12, siliques of different development as indicated (days after flowering).

However in older siliques, seeds with arrested development became visible to the naked eye and closer examination revealed that the clear white seeds contained embryos arrested at the globular stage. Presumably, these seeds corresponded to dark brown aborted seeds apparent in mature siliques. It should be noted that the transient increase in *ATS2* expression observed for the wild type four days after flowering (Figure A.1.5) corresponded to the transition in embryo development from the globular to the heart stage, when the developmental arrest occurs in a quarter of the seeds developing in siliques on heterozygous plants.

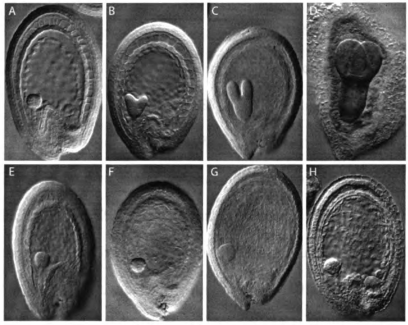


Figure A.1.6 Arrest of embryo development
Four siliques of different developmental age growing on heterozygous *ATS2/ats2-1-1* plants were dissected. (a, e), two representative embryos from the youngest silique, in which all embryos were at the globular stage and indistinguishable; (b, f), older silique with most embryos at the heart stage (b) and with about 25 % of the embryos arrested at the globular stage (f); (c) developing early torpedo stage embryo and (g) arrested embryo from the same silique at the globular stage; (d), oldest silique with developing embryo at the late torpedo stage and (h) globular arrested embryo from the same silique.

A.1.5 Discussion

With the availability of the Arabidopsis genome sequence (The Arabidopsis Genome Initiative, 2001) it has become feasible to predict genes encoding enzymes involved in lipid metabolisms in Arabidopsis based on sequence similarity to enzymes with known functions from bacteria and other plants. Ohlrogge and coworkers have used this approach to assemble a catalog of lipid metabolism genes in Arabidopsis (Beisson *et al.*, 2003). As we now know based on the present study, they correctly annotated Arabidopsis gene At4g30580 to encode a plastidic LPAAT. Functional proof was provided by heterologous complementation of the *E. coli* LPAAT-deficient *plsC* mutant (Figure A.1.2). Furthermore, subcellular localization of the ATS2-GFP fusion (Figure A.1.3) agreed with a plastidic localization of the protein as predicted. Independent experimental evidence for a plastid localization of ATS2 was also obtained by Ferro and colleagues (Ferro *et al.*, 2002; Ferro *et al.*, 2003). When confirmed and putative LPAAT protein sequences from different organism were aligned and grouped based on the amino acid sequence similarity, ATS2 of Arabidopsis clustered with BAT2 from *B. napus* and LPAT2 from *O. sativa*, both plastidic LPAATs (Figure A.1.3). Furthermore, these plastidic LPAATs formed a subcluster with cyanobacterial LPAATs providing supportive evidence for the cyanbacterial endosymbiont origin of chloroplasts. Four other putative Arabidopsis LPAATs were grouped in a distinct cluster representing microsomal LPAATs (Figure A.1.1) suggesting that their evolutionary origin is different from that of plastidic isoforms. Although in case of the well studied class of LPAATs the analysis of genomic information led to precise predictions, for which the experimental verification as shown seemed trivial, one should keep in mind that most of the information in current

genome databases is based only on conjecture. Corroborating genomic predictions is only the beginning to a broader understanding of the biological (physiological) function of gene products, in this case ATS2. One of the most versatile resources to shed light on the function of a particular gene in Arabidopsis is the broad availability of sequenced T-DNA insertion lines (Alonso *et al.*, 2003). In case of *ATS2*, two independent alleles were available, both of which gave rise to a striking phenotype, embryo-lethality. The presence of two independent alleles and restoration of embryo development following the transgenic expression of the *ATS2* cDNA (genetic complementation) linked the embryo defect to the T-DNA inactivation of the gene encoding plastidic LPAAT in the two allelic mutants. The result was surprising, because a mutant in the plastidic G3P acyltransferase gene, *ATS1*(*ACT1*), affecting the step prior to ATS2 in the pathway, had no effect on embryo development, despite the fact that the plastid pathway of thylakoid lipid biosynthesis was strongly impaired in this mutant (Kunst *et al.*, 1988). However, it should be noted that it is not clear at this time to which extent the *ats1*(*act1*) mutants is "leaky", if at all, because plastidic PG, which is assumed to be synthesized by the plastid pathway of membrane lipid biosynthesis, was only moderately reduced in the mutant. The complete loss of PG in the *ats2* mutants could not be the cause for the observed embryo arrest, because homozygous null-alleles of the *PGP1* gene encoding plastidic phosphatidylglycerolphosphate synthase gave rise to seedlings lacking plastidic PG (Hagio *et al.*, 2002; Babiychuk *et al.*, 2003). However, these mutant seedlings were completely white and non-photosynthetic. Therefore, either the absence of plastidic PA as a critical intermediate or signaling molecule, or the accumulation of lyso-PA harmful to membrane integrity might be the cause of the observed embryo lethality in the *ats2*

mutants. The fact that embryo development was arrested just at the transition from the globular to the heart stage, when thylakoid membranes start to develop and a corresponding peak in *ATS2* gene expression was observed (Figure A.1.5), suggests that the lack of PA as a central intermediate of lipid metabolism might be the main cause of embryo arrest in the *ats2* mutants. However, other possibilities cannot be ruled out at this time. While embryo-lethality is an easily observable phenotype, the respective null-mutants provide no detailed information on the role of plastidic ATS2, because homozygous lines could not be further analyzed and heterozygous lines lacked any phenotype. Mutations with leaky alleles in *ATS2* giving rise to viable plants with milder phenotypes will have to be isolated and studied to better understand the role of PA in plant metabolism.

Acknowledgments

We thank Hongbo Gao for his help with fluorescence microscopy. We would like to thank Dr. David S. Stephens for the *E. coli* strains JC201 and JC200. This work was supported in parts, by grants from the US Department of Energy Bioscience Program and the MSU Center for Novel Plant Products.

Reference

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., Ecker, J.R. (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301,** 653-657.

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

Athenstaedt, K., Daum, G. (1999) Phosphatidic acid, a key intermediate in lipid metabolism. *Eur. J. Biochem.* **266,** 1-16.

Babiychuk, E., Müller, F., Eubel, H., Braun, H.P., Frentzen, M., Kushnir, S. (2003) Arabidopsis phosphatidylglycerophosphate synthase 1 is essential for chloroplast differentiation, but is dispensable for mitochondrial function. *Plant J.* **33,** 899-909.

Bechtold N, Pelletier G (1998) In plant *Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. In *Arabidopsis Protocols* (Martinez-Zapater, J., Salinas, J., eds). Totowa, NJ: Humana Press, pp 259-266.

Beisson, F., Koo, A.J., Ruuska, S., Schwender, J., Pollard, M., Thelen, J.J., Paddock, T., Salas, J.J., Savage, L., Milcamps, A., Mhaske, V.B., Cho, Y., Ohlrogge, J.B. (2003) Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. *Plant Physiol.* **132,** 681-697.

Boisson, M., Gomord, V., Audran, C., Berger, N., Dubreucq, B., Granier, F., Lerouge, P., Faye, L., Caboche, M., Lepiniec, L. (2001) Arabidopsis glucosidase I mutants reveal a critical role of N-glycan trimming in seed development. *EMBO J.* **20,** 1010-1019.

Bourgis, F., Kader, J.C., Barret, P., Renard, M., Robinson, D., Robinson, C., Delseny, M., Roscoe, T.J. (1999) A plastidial lysophosphatidic acid acyltransferase from oilseed rape. *Plant Physiol.* **120,** 913-921.

Brown, A.P., Brough, C.L., Kroon, J.T.M., Slabas, A.R. (1995) Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-phosphate acyltransferase from *Limnanthes Douglasii. Plant Mol. Biol.* **29,** 267-278.

Brown, A.P., Coleman, J., Tommey, A.M., Watson, M.D., Slabas, A.R. (1994) Isolation and characterization of a maize cDNA that complements a 1-acyl sn-glycerol-3-phosphate acyltransferase mutant of *Escherichia Coli* and encodes a protein which has similarities to other acyltransferases. *Plant Mol. Biol.* **26,** 211-223.

Coleman, J. (1990) Characterization of *Escherichia coli* cells deficient in 1-acyl-sn-glycerol-3-phosphate acyltransferase activity. *J. Biol. Chem.* **265,** 17215-17221.

Coleman, J. (1992) Characterization of the *Escherichia coli* gene for 1-acyl-sn-glycerol-3-phosphate acyltransferase (Plsc). *Mol. Gen. Genet.* **232,** 295-303.

Davies, H.M., Hawkins, D.J., Nelsen, J.S. (1995) Lysophosphatidic acid acyltransferase from immature coconut endosperm having medium-chain length substrate-specificity. *Phytochemistry* **39,** 989-996.

Eberhardt, C., Gray, P.W., Tjoelker, L.W. (1997) Human lysophosphatidic acid acyltransferase - cDNA cloning, expression, and localization to chromosome 9q34.3. *J. Biol. Chem.* **272,** 20299-20305.

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

Estelle, M.A., Somerville, C. (1987) Auxin-resistant mutants of *Arabidopsis thaliana* with an altered morphology. *Mol. Gen. Genet.* **206,** 200-206.

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., Garin, J., Joyard, J., Rolland, N. (2003) Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana. *Mol. Cell. Prot.* **2,** 325-345.

Ferro, M., Salvi, D., Riviere-Rolland, H., Vermat, T., Seigneurin-Berny, D., Grunwald, D., Garin, J., Joyard, J., Rolland, N. (2002) Integral membrane proteins of the chloroplast envelope: Identification and subcellular localization of new transporters. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 11487-11492.

Hagio, M., Sakurai, I., Sato, S., Kato, T., Tabata, S., Wada, H. (2002) Phosphatidylglycerol is essential for the development of thylakoid membranes in Arabidopsis thaliana. *Plant Cell Physiol.* **43,** 1456-1464.

Hanke, C., Wolter, F.P., Coleman, J., Peterek, G., Frentzen, M. (1995) A plant acyltransferase involved in triacylglycerol biosynthesis complements an *Escherichia coli* sn-1-acylglycerol-3-phosphate acyltransferase mutant. *Eur. J. Biochem.* **232,** 806-810.

Knutzon, D.S., Nelsen, J.S., Lardizabal, K.D., Bleibaum, J.L., Hawkins, D.J., Davies, H.M., Metz, J.G. (1995) Purification and molecular cloning of a medium chain-preferring lysophosphatidic acid acyltransferase from immature coconut endosperm. *Plant Physiol.* **108,** 50.

Kunst, L., Browse, J., Somerville, C. (1988) Altered regulation of lipid biosynthesis in a mutant *Arabidopsis* deficient in chloroplast glycerol-3-phosphate acyltransferase activity. *Proc. Natl. Acad. Sci. U. S. A.* **85,** 4143-4147.

Lassner, M.W., Levering, C.K., Davies, H.M., Knutzon, D.S. (1995) Lysophosphatidic acid acyltransferase from meadowfoam mediates insertion of erucic-acid at the sn-2 position of triacylglycerol in transgenic rapeseed oil. *Plant Physiol.* **109,** 1389-1394.

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

Mansfield, S.G., Briarty, L.G., Erni, S. (1991) Early embryogenesis in *Arabidopsis thaliana*. I. The mature embryo sac. *Can. J. Bot.* 447-460.

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

Nagiec, M.M., Wells, G.B., Lester, R.L., Dickson, R.C. (1993) A suppressor gene that enables *Saccharomyces cerevisiae* to grow without making sphingolipids encodes a protein that resembles an *Escherichia coli* fatty acyltransferase. *J. Biol. Chem.* **268,** 22156-22163.

Roughan, P.G., Slack, C.R. (1982) Cellular organization of glycerolipid metabolism. *Ann. Rev. Plant Physiol.* **33,** 97-132.

Sonnhammer, E.L., von Heijne, G., Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6,** 175-182.

The Arabidopsis Genome Initiative (2001) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796-815.

Thompson, J.D., Higgins, D.G., 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.* **22,** 4673-4680.

West, J., Tompkins, C.K., Balantac, N., Nudelman, E., Meengs, B., White, T., Bursten, S., Coleman, J., Kumar, A., Singer, J.W., Leung, D.W. (1997) Cloning and expression of two human lysophosphatidic acid acyltransferase cDNAs that enhance cytokine-induced signaling responses in cells. *DNA Cell Biol.* **16,** 691-701.

Yokoi, S., Higashi, S., Kishitani, S., Murata, N., Toriyama, K. (1998) Introduction of the cDNA for Arabidopsis glycerol-3-phosphate acyltransferase (GPAT) confers unsaturation of fatty acids and chilling tolerance of photosynthesis on rice. *Mol. Breeding* **4,** 269-275.

Appendix 2

G6PDH gene expression patterns from AtGenExpress gene expression atlas

A.2 G6PDH gene expression patterns from AtGenExpress gene expression atlas

The gene expression patterns of G6PDH isoforms were analyzed in the AtGenExpress gene expression atlas (Schmid *et al.* 2005). The Excel worksheet containing formulas for calculation was constructed in the laboratory of Dr. John Ohlrogge. The y-axis represents the average of the triplicate values for the intensity detected on the array, normalized to that of the control genes. The tissues from which the RNA was extracted are shown on the x-axis. The bars to the right of the same group of tissues represent older developmental stages of the tissue.

The overall patterns and relative levels (by comparing y-axis values) are similar to what was observed in the present study (Chapter 2). Some differences are discussed here. *G6PD4* had low mRNA levels throughout the plant in both studies. Interestingly in seeds, high abundance of mRNA was detected (Figure A2.4). Plants homozygous for the T-DNA insertion in *G6PD4* were not examined by zymogram. If G6PD4 has a function other than as an active G6PDH, it would be interesting to find what seed-specific role it may carry out. *G6PD5* was found to be expressed predominantly in leaf, while activity was detected both in leaf and roots (Chapter 2). In contrast, AtGenExpress data for *G6PD5* shows the expression pattern to be similar to that of the activity detected in the present study (Figure A2.5). Finally, gene expression patterns from seeds are not similar between AtGenExpress and the present study. As mentioned in Chapter 3, this is probably due to the difficulty in preparing clean RNA from seeds. Considering that the AtGenExpress results are obtained from triplicates and with highly controlled data analyses, it may be more reliable for seed gene expression analysis.

Reference

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., Lohmann, J.U. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nature genetics.* **37,** 501-506.

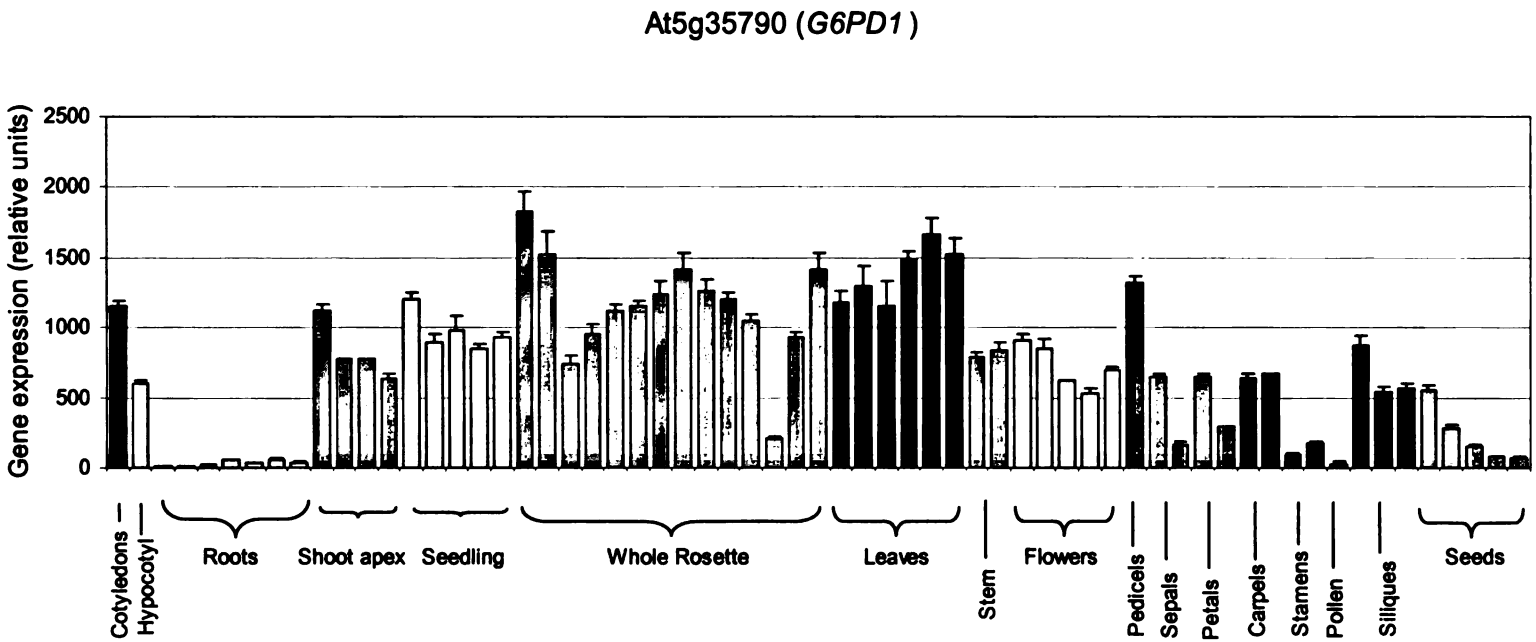


Figure A.2.1 Expression of *G6PD1* according to AtGenExpress expression atlas

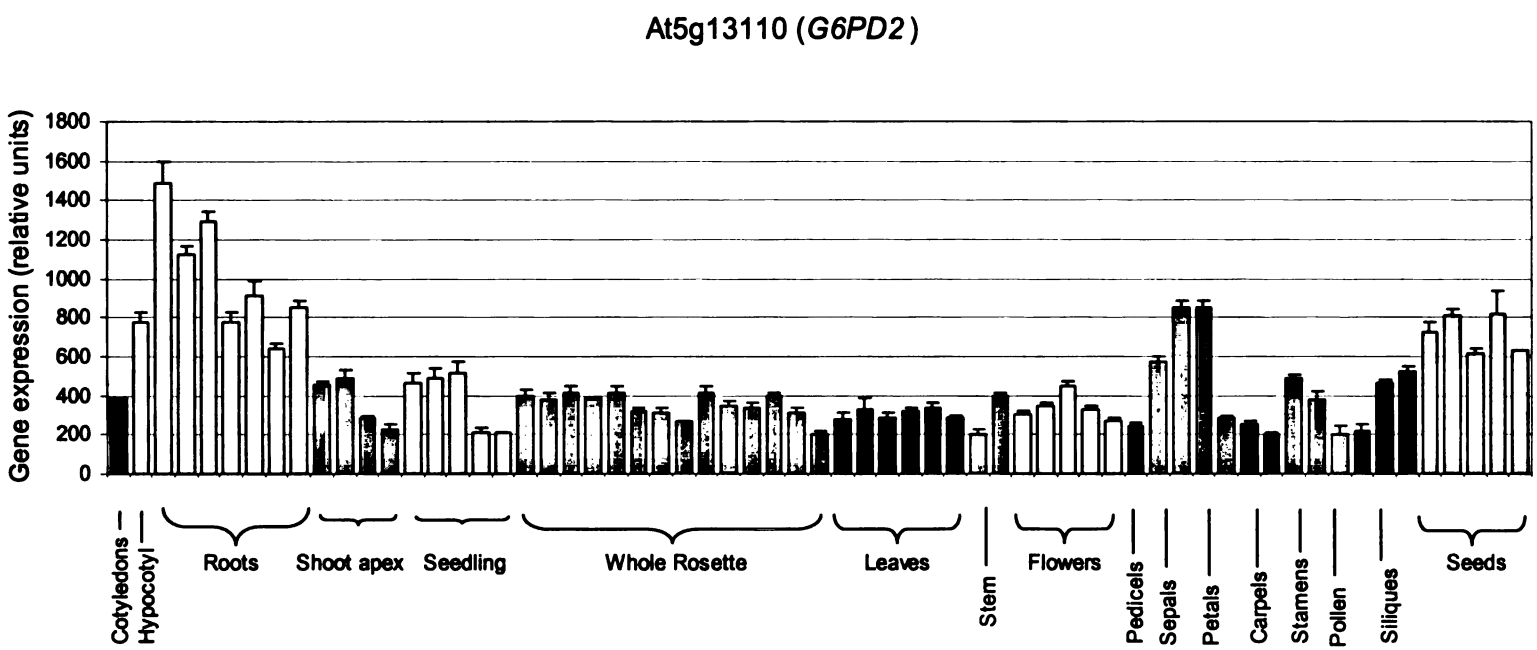


Figure A.2.2 Expression of *G6PD2* according to AtGenExpress expression atlas

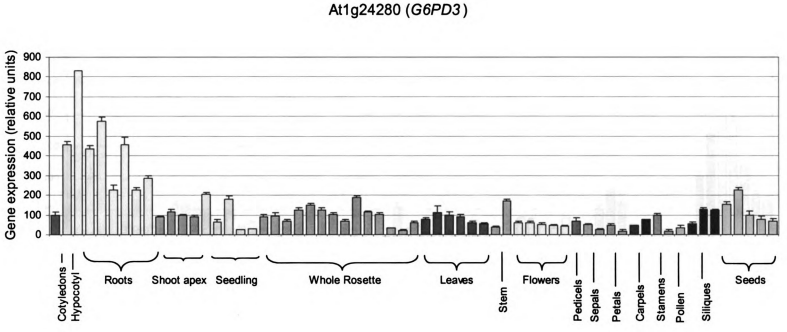


Figure A.2.3 Expression of *G6PD3* according to AtGenExpress expression atlas

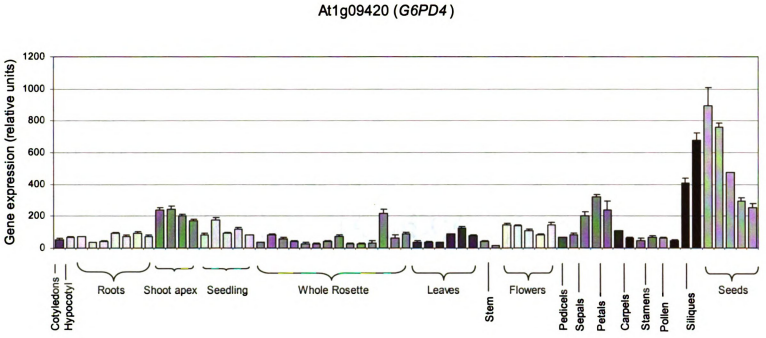


Figure A.2.4 Expression of *G6PD4* according to AtGenExpress expression atlas

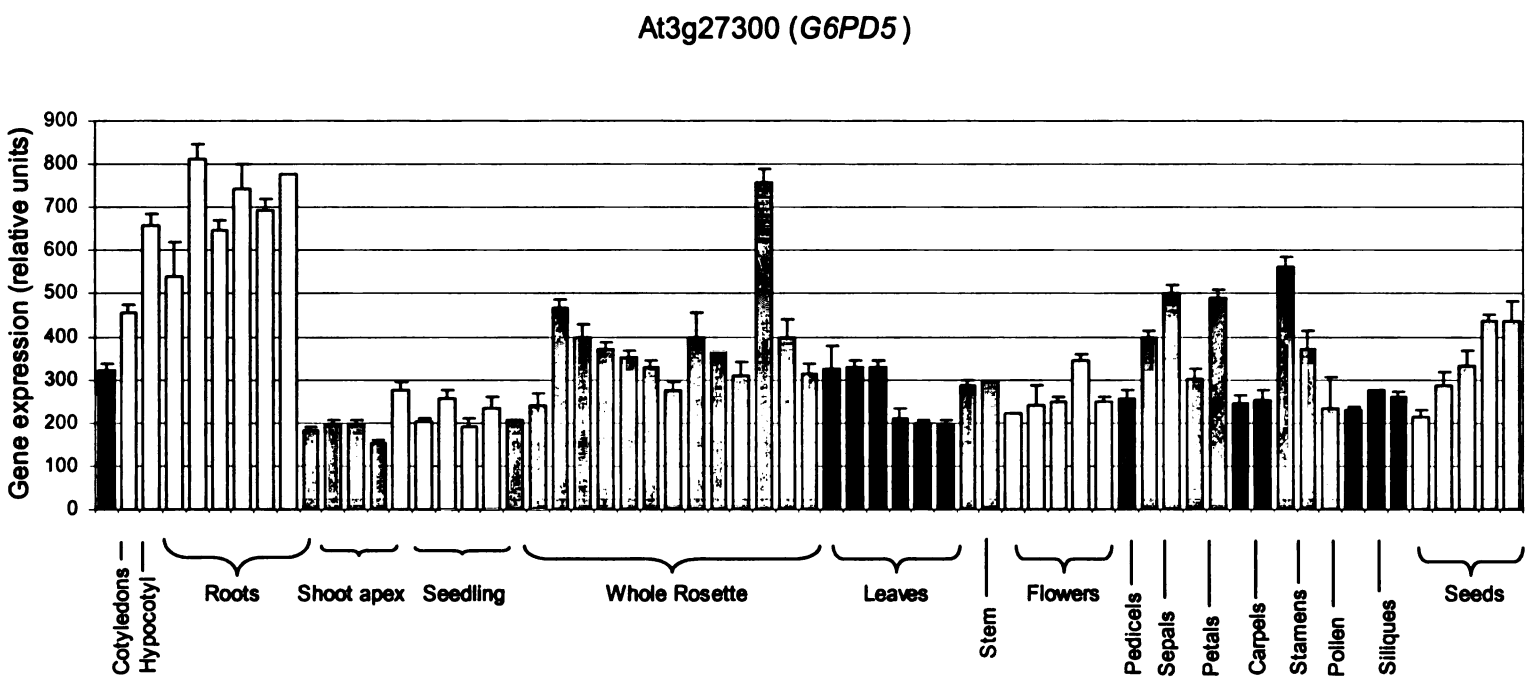


Figure A.2.5 Expression of *G6PD5* according to AtGenExpress expression atlas

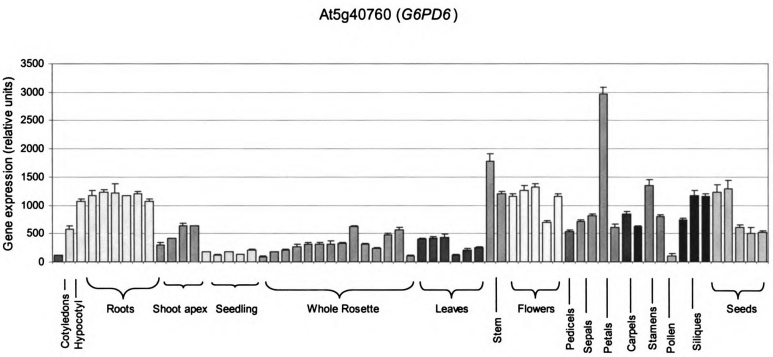


Figure A.2.6 Expression of *G6PD6* according to AtGenExpress expression atlas