Isolation and Characterization of Plastidic Glucose-6-Phosphate Dehydrogenase (G6PDH) from Castor (Ricinus communis L.)

by

Ka-Yu Law

A thesis submitted to the Department of Biology in conformity with the requirements for the degree of Master of Science

Queen’s University
Kingston, Ontario, Canada
September, 2007

Copyright© Ka-Yu Law, 2007
Abstract

Plant cells contain plastids, organelles dedicated to performing specific biochemical processes including photosynthesis, starch and oil biosynthesis. Fatty acid biosynthesis in oil seeds occurs in one type of plastid termed the leucoplast. Anabolic metabolism in leucoplasts includes the production of fatty acids and amino acids that depend on the availability of reductants such as NADPH. NADPH can be generated in plastid by glucose 6-phosphate dehydrogenase (G6PDH) which is the chief control enzyme and first step in the Oxidative Pentose Phosphate Pathway (OPPP). G6PDH catalyses the reaction of NADP⁺ and glucose 6-phosphate to NADPH and 6-phosphogluconate. At least two compartment-specific isoforms of G6PDH exist in plants, a cytosolic and a plastidic form. In this study, castor oil seed (COS) (Ricinus communis L.) was used as a model enzyme system for the ongoing study of oil biosynthesis in plants. This is the first ever report of the full-length clone of the plastidic isoform of G6PDH being isolated from a castor cDNA library using polyclonal potato plastidic G6PDH antiserum. The full-length cDNA was sequenced and compared to other G6PDH genes from higher plants, the castor sequence reveals conserved regions and conserved cysteine residues similar to other higher plant G6PDH. Over expression of the recombinant cleaved fusion protein in an E. coli expression system from the isolation of the cDNA clone shows it is enzymatically active, stable and unlike other plastid G6PDH’s dithiothreitol insensitive. In fact this G6PDH shows increased activation in the presence of dithiothreitol. Initial kinetic characteristics shows that it behaves in a similar fashion enzymatically when compared to other higher plant chloroplast G6PDH. The gene sequence and initial kinetic findings for castor G6PDH concur with other higher plant, non-photosynthetic, plastidic isoforms.
Acknowledgements

I sincerely thank my supervisors Dr. David Turpin and Dr. David Dennis for excellent guidance and supervision throughout the course of my project research, especially Dr. Dennis for his patience. I also thank Dr. Antje von Schaewen for the antibodies to the potato G6PDH enzymes used for the library screening. Dr. Mike Kuiper, an excellent teacher of protein chemistry, was instrumental in constructing and optimizing the G6PDH molecular model. Fellow lab mates Dr. Kelvin Cole and Fayek Negm are also greatly thanked for their patience and assistance during various aspects of this project. This work has been supported by the NSERC, Canada.
Table of Contents

Abstract ................................................................................................................................. i
Acknowledgements ............................................................................................................. ii
Table of Contents ............................................................................................................... iii
List of Figures ....................................................................................................................... iv
List of Tables ....................................................................................................................... v
List of Abbreviations ......................................................................................................... vi
Chapter 1: Introduction and Literature Review ................................................................... 1
  Fatty Acid Synthesis in Developing Seeds ..................................................................... 1
  The Oxidative Pentose Phosphate Pathway .................................................................... 4
  Glucose 6-Phosphate Dehydrogenase – General Properties ......................................... 7
  G6PDH in non-photosynthetic organisms .................................................................... 8
  G6PDH in photosynthetic organisms .......................................................................... 9
    G6PDH in Green Algae .......................................................................................... 10
    G6PDH in Cyanobacteria ..................................................................................... 10
    G6PDH in higher plants ....................................................................................... 11
  Ferredoxin-thioredoxin regulation of enzymes ............................................................. 13
  Research Objectives ..................................................................................................... 17
Chapter 2: Materials and Methods .................................................................................... 19
  Plant Material .............................................................................................................. 19
  Leucoplast Isolation .................................................................................................... 19
  Chemicals .................................................................................................................... 20
  Enzymatic Assays ....................................................................................................... 20
  Protein Determination ................................................................................................. 20
  Protein Electrophoresis and Western Blotting ............................................................. 21
  Castor cDNA Library Screening .............................................................................. 24
  Phage DNA Preparation and Subcloning ................................................................. 25
  Plasmid DNA Preparation .......................................................................................... 25
  DNA Electrophoresis and Southern Blotting .............................................................. 26
  Sequencing and Sequence Analysis .......................................................................... 27
  Computer analysis of sequence data ......................................................................... 27
  Molecular modeling .................................................................................................... 28
  Phylogenetic and Divergence Analysis ...................................................................... 29
  Recombinant DNA techniques, subcloning and sequencing ...................................... 29
  PGEX-4T-3 Expression Vector Construction ............................................................ 30
  Expression of GST fusion proteins from E. coli ....................................................... 30
  Purification of GST fusion proteins from E.coli ..................................................... 31
  Enzyme assays and determination of preliminary kinetics ........................................... 31
Chapter 3: Results ................................................................................................................. 32
Chapter 4: Discussion .......................................................................................................... 68
Chapter 5: Summary .......................................................................................................... 74
References .............................................................................................................................. 75
**List of Figures**

Fig 1 OPPP pathway ................................................................. 5
Fig 2 Comparison between Green Plastids, Leucoplasts and Green Plastids in Seeds......................................................... 16
Fig 3 Initial western blots using yeast and potato antibodies ................... 34
Fig 4 COS G6PDH nucleotide sequence with translation.......................... 38
Fig 5 Dendogram of 25 sequences to castor........................................ 41
Fig 6 Comparison of higher plant sequences to castor ........................... 44
Fig 7 Molecular model of castor G6PDH .......................................... 48
Fig 8 GST fusion western blot and purification western ......................... 52
Fig 9 Native gel activity with DTT .................................................. 54
Fig 10 Effects of DTT on G6PDH activity ....................................... 56
Fig 11 Effects of DTT on 6PG and overexpression ................................ 58
Fig 12 Kinetics of Km for G6PDH with NADPH ................................ 61
Fig 13 Kinetics of Km G6PDH with G-6-P ........................................ 64
List of Tables

Table 1 Comparison of Km values .................................................................67
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>AMP</td>
<td>adenosine monophosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>DHAP</td>
<td>dihydroxyacetone phosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E4P</td>
<td>erythrose 4-phosphate</td>
</tr>
<tr>
<td>F6P</td>
<td>fructose 6-phosphate</td>
</tr>
<tr>
<td>Fd red/ox</td>
<td>ferredoxin, reduced/oxidized</td>
</tr>
<tr>
<td>FNR</td>
<td>ferredoxin NADP⁺ oxidoreductoase (EC 1.18.1.2)</td>
</tr>
<tr>
<td>G-1-P</td>
<td>glucose 1-phosphate</td>
</tr>
<tr>
<td>G-3-P</td>
<td>glyceraldehyde 3-phosphate</td>
</tr>
<tr>
<td>G-6-P</td>
<td>glucose 6-phosphate</td>
</tr>
<tr>
<td>G6PDH</td>
<td>glucose 6-phosphate dehydrogenase (EC 1.1.1.49)</td>
</tr>
<tr>
<td>Gln</td>
<td>glutamine</td>
</tr>
<tr>
<td>Glu</td>
<td>glutamate</td>
</tr>
<tr>
<td>GS</td>
<td>glutamine synthetase (EC 6.3.1.2)</td>
</tr>
<tr>
<td>Km</td>
<td>Michaelis-Menten constant</td>
</tr>
<tr>
<td>NADP⁺</td>
<td>nicotinamide adenine dinucleotide phosphate, oxidized</td>
</tr>
<tr>
<td>NADPH</td>
<td>nicotinamide adenine dinucleotide phosphate, reduced</td>
</tr>
<tr>
<td>OPPP</td>
<td>oxidative pentose phosphate pathway</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>3-PGA</td>
<td>3-phosphoglycerate</td>
</tr>
<tr>
<td>PEP</td>
<td>phosphoenolpyruvate</td>
</tr>
<tr>
<td>PEPC</td>
<td>phosphoenolpyruvate carboxylase (EC 4.1.1.31)</td>
</tr>
<tr>
<td>6PG</td>
<td>6-phosphogluconate</td>
</tr>
<tr>
<td>6PGDH</td>
<td>6-phosphogluconate dehydrogenase (EC 1.1.1.43)</td>
</tr>
<tr>
<td>P$_i$</td>
<td>inorganic phosphate, orthophosphate (PO$_4^{3-}$)</td>
</tr>
<tr>
<td>PRK</td>
<td>phosphoribulokinase (EC 2.7.1.19)</td>
</tr>
<tr>
<td>Pyr</td>
<td>pyruvate</td>
</tr>
<tr>
<td>R5P</td>
<td>ribose 5-phosphate</td>
</tr>
<tr>
<td>RPPP</td>
<td>reductive pentose phosphate pathway</td>
</tr>
<tr>
<td>Ru5P</td>
<td>ribulose 5-phosphate</td>
</tr>
<tr>
<td>RuBP</td>
<td>ribulose 1,5-bisphosphate</td>
</tr>
<tr>
<td>S7P</td>
<td>sedoheptulose 7-phosphate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>TCA</td>
<td>tricarboxylic acid</td>
</tr>
<tr>
<td>Trx</td>
<td>thioredoxin</td>
</tr>
<tr>
<td>X5P</td>
<td>xyulose 5-phosphate</td>
</tr>
<tr>
<td>$V_{max}$</td>
<td>maximum velocity</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction and Literature Review

Fatty Acid Synthesis in Developing Seeds

Castor is a very valuable oilseed agricultural crop that provides a significant proportion of the world’s total supply of hydroxy fatty acids used to produce lubricants, paints, soaps, and some pharmaceuticals (Atsmon, 1989; Salunkhe and Desai, 1986). In the maturing seeds of many plants the main storage products produced are protein and oils (Murphy et al., 1989). Some seeds are non-photosynthetic and the synthesis of intermediary starch and fatty acids occurs in the plastids. Synthesis requires a supply of carbon precursors from the cytosol (Smith and Martin, 1993). Plant lipid metabolism depends critically upon plastids that are the only cellular site for de novo fatty acid biosynthesis. The plastids supply the entire plant cell with fatty acids for the synthesis of membrane components and storage lipids (Stumpf, 1984; Ohlrogge et al., 1993). Chloroplasts are the most important sites of glycerolipid biosynthesis in leaves synthesizing ~75% of the total leaf lipid (Harwood, 1980; Roughan and Slack, 1982). In contrast, there is less information elucidating lipid metabolism in nonphotosynthetic plastids.

There is also limited knowledge of the precursors required for oil synthesis in the seed plastidic tissues, or which cytosolic precursors are utilized by the plastids from developing embryos. It is known that starch synthesis uses ADP-glucose (ADPG), which is synthesized from glucose-1-phosphate (G-1-P) and ATP in the plastid (Smith and Martin, 1993). There is also evidence that hexose phosphate is imported as the carbon
source for ADPG synthesis in most plastids. Nonphotosynthetic tissues lack the capacity
to convert triose phosphate to hexose phosphate (Entwistle and ap Rees, 1990) but do
have the ability to incorporate metabolites into starch in isolated plastids. The metabolite
glucose-6-phosphate (G-6-P) is utilized by plastids from pea embryos and cauliflower
buds (Hill and Smith, 1991; Neuhaus et al., 1993a). In contrast, G-1-P is utilized by
plastids from wheat endosperm (Tyson and ap Rees, 1988), while those from maize
endosperm are able to use both G-1-P and G-6-P (Neuhaus et al., 1993b).

A complete glycolytic pathway is present in non-photosynthetic plastids from pea
embryos, cauliflower buds, and wheat and castor bean endosperm (Entwistle and ap
Rees, 1988; Foster and Smith, 1993; Journet and Douce, 1985; Simcox et al., 1977). The
first two enzymes of the Oxidative Pentose Phosphate Pathway (OPPP) are present in
plastids isolated from pea roots, sycamore cells and cauliflower buds (Alban et al.,
1989b; Journet and Douce, 1985).

In most plants species, G6PDH has not been well characterized and research has
mainly concentrated on the cytosolic isoform. In the leucoplasts of castor oil seeds, no
significant activity of plastidic G6PDH had been detected from developing castor
endosperm (Simcox et al., 1977) but its presence is inferred from work on plastids that
has shown significant levels of 6-phosphogluconate dehydrogenase activity, the second
enzyme in the OPPP. The inability to detect plastidic G6PDH in COS maybe due to the
highly unstable nature of the enzyme and subsequent loss of activity during the
purification process and perhaps oxidizing agents were used in an attempt to activate the
enzyme during purification. The activity of the OPPP in non-photosynthetic plastids is
linked to nitrogen metabolism through the prerequisite of reductant (Emes and Bowsher, 1991) but its true involvement in fatty acid synthesis still remains poorly characterized.

The first committed step of fatty acid synthesis occurs in the plastid with the formation of malonyl-CoA from acetyl-CoA and bicarbonate catalysed by acetyl-CoA carboxylase (ACCase) (Harwood, 1988). Acetyl-CoA must be generated within the plastid as it cannot cross the plastid membrane and several substrates for acetyl-CoA synthesis have been indicated including pyruvate, malate, acetyl-camitine and acetate (Masterson et al., 1990; Miernyk and Dennis, 1983; Roughan et al., 1979; Smith et al., 1992). There are few studies of the synthesis of fatty acids derived from hexoses in non-photosynthetic plastids from a range of tissues (Entwistle and ap Rees, 1988; Journet and Douce, 1985; Simcox et al., 1977). Studies to determine potential carbon sources metabolized by the plastids for fatty acid synthesis were determined using $^{14}$C labeled G6P, fructose, glucose, F6P, and DHAP. Each of these metabolites were incorporated into starch in isolated plastids in an ATP-dependent manner. (Hill and Smith, 1991; Neuhaus et al., 1993a). Thus, plastids from seed embryos can utilize several different cytosolic carbon sources for both starch and fatty acid synthesis.

Most research on the enzymes and the metabolism of metabolites into starch and fatty acids uses isolated plastids and the degree of conversion depends greatly upon the intactness of the plastids suggesting that there are specific translocators that transmit metabolites across the intact plastid membrane. Translocators have been identified in a some plant species, for example small molecules like triose phosphate and G-6-P can be translocated in pea root plastids and cauliflower bud plastids (Journet and Douce, 1985 Batz et al., 1993; Alban et al., 1989a); glucose transporters in spinach chloroplasts have
also been reported (Weber et al., 2000). A translocator for ATP has been reported to be present in plastids from sycamore cells, pea roots, spinach chloroplasts and inferred in plastids from oilseed rape embryos by the dependence of starch synthesis on exogenous ATP (Flugge, 1998).

Pyruvate has been shown to be one of the most effective substrates for fatty acid synthesis by the plastids giving a rate five times better than acetate in castor bean endosperm and mustard cotyledons (Smith et al., 1992). Acetate is reported to be the better substrate in other plant plastids (Roughan et al., 1979). In contrast, for castor bean fatty acid synthesis by plastids, malate has proven to be a more effective substrate than either pyruvate or acetate. These substrates are important sources of carbon and reducing power for fatty acid synthesis in the castor bean endosperm (Smith et al., 1992).

NADPH, NADH and ATP can vary with the tissue source from which the plastids are isolated, and with the metabolites which are supplied as carbon sources (Smith et al., 1992; Sparace et al., 1988). The utilization of exogenous or internally generated reducing power and ATP to support fatty acid synthesis in plastids from castor oil seed embryos requires more study of the enzymes involved in the OPPP.

**The Oxidative Pentose Phosphate Pathway**

The OPPP provides an alternate route of carbohydrate catabolism to glycolysis. The major difference between glycolysis and the OPPP is that the latter produces only NADPH, whereas glycolysis generates NADH. Both pathways are not independent of each other as they share some common intermediates like G-3-P and G-6-P (Dennis et al., 1997). The OPPP is found in all prokaryotes and eukaryotes and its function is to generate
reductant and precursors for a variety of biosynthetic reactions. The steps of the OPPP are outlined in Figure 1. The OPPP is divided into two main phases namely oxidative or regenerative. The oxidative phase begins with G-6-P being oxidized by glucose 6-phosphate dehydrogenase (G6PDH) and 6-phosphogluconate dehydrogenase (6PGDH) producing ribulose 5-phosphate (Ru5P), two molecules of NADPH and one molecule of CO₂. The two reactions of the oxidative phase are irreversible and tightly regulated within the cell. The oxidative phase takes place in both the plastids and cytosol of plants and photosynthetic algae (Stitt and ap Rees 1979; Schnarrenberger et al., 1995; Klein 1986).

The regenerative phase leads to a series of rearrangements of the pentose phosphates, by various enzymatic reactions including epimerase, isomerase, transketolase and transaldolase. The enzymes of the regenerative phase are near equilibrium and regulatory mechanisms are not known. The regenerative phase can generate molecules containing three, four, six or seven carbon molecules, intended for a number of biosynthetic uses. Carbon inputs may be recycled in the regenerative phase of the OPPP and recycled back into the oxidative portion of the pathway, by the activity of hexose phosphate isomerase. The enzymes of the regenerative phase have been found only in the plastids of higher plants (Stitt and ap Rees 1979, Schnarrenberger et al., 1995). The OPPP generates NADPH and provides carbon precursors. The intermediates of the regenerative phase may be taken and shunted to other synthetic pathways. For example ribose 5-phosphate is required for nucleotide synthesis in both plants and animals and Erythrose 4-phosphate (E4P) is incorporated into aromatic amino acids and phenolic compounds in plants (Dennis et al. 1997, Pandolfi et al. 1995).
Figure 1. Enzymatic reactions of the Oxidative Pentose Phosphate Pathway (OPPP) (Dennis et al., 1997).
OPPP generated NADPH has been demonstrated to be important in providing large reserves of reductants for fatty acid biosynthesis in mammalian liver, plants and developing endosperm tissue (Eggleston and Krebs 1974; Agrawal and Canvin, 1971). The regeneration of hemoglobin from methemoglobin in animals is also critically dependent upon OPPP produced reductant (Luzzatto 1995). Reductant created from the OPPP may also be important to combat oxidative stress (Juhnke et al., 1996). Regulation of the OPPP occurs by induction and synthesis of G6PDH. Finer control of G6PDH and the OPPP activity in photosynthetic plants is accomplished through competitive inhibition by NADPH and redox modulation by thioredoxin. Control of the OPPP is important in plants as it essentially is the opposite of the Calvin cycle. Inactivation in the light is imperative to prevent futile cycling of C02, which would be fixed and then released by the two opposing pathways.

**Glucose 6-Phosphate Dehydrogenase – General Properties**

Glucose-6-phosphate 1-dehydrogenase (G6PDH, EC 1.1.1.49) is an enzyme that catalyses the rate-limiting step of the OPPP. G6PDH is the main focus of regulation of this pathway in both plants and animals. Almost all forms of G6PDH studied thus far are highly specific for their substrates, G-6-P and NADP\(^+\) for this reason, many of the properties of G6PDH have been preserved across evolutionary lines (Levy 1979). All G6PDH's have a single subunit type with the native enzyme existing as a dimer, tetramer or hexamer. The subunit size is also highly conserved between prokaryotes and eukaryotes between 50 to 67 kDa. The majority of characterized G6PDH enzymes
show hyperbolic kinetics with some cyanobacterial enzymes showing sigmoidal kinetics (Schaeffer and Stanier, 1978).

A number of bacterial forms of the enzyme are capable of using NAD\(^+\) in place of NADP\(^+\) (Anderson et al., 1997). All G6PDH's are inhibited by NADPH, which acts as a competitive inhibitor. Metabolic effectors of G6PDH include Mg\(^{2+}\) (increases activity), ATP, ADP, AMP, long-chain acyl-CoA's (in mammalian liver, adipose tissue and erythrocytes), steroids, CoA and acetylCoA (in some plant species and bacteria this acts as a negative effector) (Gordon et al., 1995; Levy 1979).

**G6PDH in non-photosynthetic organisms**

Microbial and mammalian forms of this enzyme have been isolated from a number of species and G6PDH ranges in size from 50,000-67 000 Da. The enzyme is dimeric, tetrameric or may form higher order oligomeric forms (Levy, 1979). Enzymatic activity for these forms is generally maximal around pH 7.4 with Michaelis constants of 0.86mM and 0.042 mM for glucose-6-phosphate and NADP\(^+\) respectively. In animal systems the OPPP is important in tissues like the liver that undergo fluctuations in metabolism (Eggleston and Krebs 1974).

Animal tissues responsible for high levels of fatty acid biosynthesis control the OPPP at the level of G6PDH regulation. In liver tissue, G6PDH is induced in response to high dietary carbohydrate, and is repressed in response to starvation (Levy 1979). The onset of lactation in mammals causes G6PDH activity to increase up to 60 times. Elevated G6PDH activity has been noted in mammary and hepatic tumors.
G6PDH in photosynthetic organisms

The regulation of G6PDH has been studied in plants, eukaryotic algae and cyanobacteria, but the gene encoding G6PDH has been isolated from only a few of these sources (Graeve et al., 1994, Fickenscher and Scheibe 1986, Srivastava and Anderson 1983, Gleason 1996). Plant G6PDH is an unstable enzyme and kinetic characterization has usually been performed on either crude extracts or partially purified protein preparations. All plant cells contain plastids, specialized organelles that perform various biochemical processes including photosynthesis, biosynthesis of starch and oil biogenesis. The latter processes in oil seeds, occur in a plastid termed the leucoplast. In leucoplasts, anabolic metabolism including the production of fatty acids and oil, mevalonate and amino acids, is dependent on the availability of reductants such as NADPH. G6PDH is the first and a key control enzyme in the Oxidative Pentose Phosphate Pathway (OPPP) where NADPH is concomitantly produced from NADP⁺ when G6PDH converts glucose-6-phosphate to 6-phosphogluconate. The OPPP serves as an alternative to glycolysis for the metabolism of glucose.

At least two compartment-specific isoforms of G6PDH, a cytosolic and a plastidic form exist in green plant tissues. G6PDH has been purified to apparent homogeneity from bacteria, yeast and animal tissues. In most cases it exists as a homodimer or a homotetramer. In higher plants, at least two G6PDH enzymes exist in two different cellular compartments, the cytosol and plastids. Each of these proteins in plants is encoded by a separate corresponding nuclear gene and share higher protein homology to each other (65%) than to cyanobacterial homologues (about 55%) (von Schaewen et al., 1995).
**G6PDH in Green Algae**

Two isoforms of G6PDH were purified by a five step chromatography purification from green algae, Chlorella vulgaris C-27. G6PDH1 is a cytosolic isoform and G6PDH2 is a chloroplastic type as determined by differing DTT sensitivity and sequence similarity of N-terminal and internal amino acid sequences to other known homologues. The partially pure proteins showed specific activity of 14.4 and 26.0 U/mg protein, respectively. G6PDH1 showed an apparent molecular weight of 200,000 with a 58,000 kDa subunit whereas G6PDH2 was larger at 450,000 kDa with a subunit of 52,000 kDa (Honjoh *et al.*, 2003). Elevated G6PDH activity as a means of regulation has been linked with inorganic nitrogen assimilation in unicellular green algae. The transfer of algal cultures to nitrogen free media results in an increase of G6PDH activity (Hipkin and Cannons 1985, Huppe and Turpin 1996).

**G6PDH in Cyanobacteria**

G6PDH has been purified and characterized from *Anabaena sp.* and shows a specific activity of 2.80µmol NADP⁺ reduced/mg protein/min which is the amount of NADP⁺ reduced in micromoles for every milligram of protein every minute. (Gleason 1996). Like other G6PDH's, this enzyme is specific for its substrates and inhibited by NADPH. Divalent cations, Ca²⁺ and Mg²⁺ cause an enhancement of activity. ATP inhibits the enzyme while Gln, as well as other free amino acids enhance activity, particularly at low G-6-P concentrations (Schaeffer and Stanier 1978; Gleason 1996; Rowell and Simpson 1990). G6PDH from *Anabaena is* redox modulated. Activity is
inhibited by thioredoxin with maximal inhibition at pH 7.0. (Rowell and Simpson 1990).

**G6PDH in higher plants**

There are two compartment-specific isozymes of G6PDH in higher plants: one in the cytosol, the second in the plastids. The cytosolic isozyme has been purified from pea leaves (Fickenscher and Scheibe 1986) and potato tubers (Graeve *et al.*, 1994). The chloroplast protein has been purified from pea leaves (Srivastava and Anderson 1983). Another plastidic protein was cloned, but not directly purified, from potato shoots (von Schaewen *et al.*, 1995). Attempts to purify G6PDH from barley root plastids were unsuccessful, as the protein was highly unstable (Wright *et al.*, 1997). The two isozymes are immunologically distinct, but compartment-specific forms share similarity between plant tissues (Fickenscher and Scheibe 1986).

In pea chloroplasts, G6PDH exists as a homotetramer of 56,000 kDa monomers and has been purified to homogeneity (Srivastava and Anderson, 1983). It exhibits a pH optimum around 8.2 and has Km values of 2.4 uM and 370 uM for NADP⁺ and glucose-6-phosphate respectively. In potato and tobacco, the cDNA sequence of both cytosolic and plastidic isozymes have been reported (Graeve *et al.*, 1994).

The cytoplasmic G6PDH forms in potato are tetrameric, with the leaf form having a subunit mass of 60 kDa, and the tuber form slightly smaller at 55 kDa. The enzymatic activities of the two cytoplasmic G6PDH proteins are similar, as are kinetic constants, one notable difference is that the tuber enzyme has a greater affinity for NADP⁺. The cytoplasmic enzyme form does not show any redox sensitivity to thioredoxin or
dithiothreitol (DTT). Other effectors of activity like Mg$^{2+}$ do not appear to improve the activity. While the potato tuber enzyme is inhibited by ATP, GTP, UTP, acetyl-CoA and CoA, the leaf enzyme was shown not to be altered by any of these metabolites, regardless of the pH.

The plastidic enzyme has been more difficult to purify from higher plants and most of the characterization has been performed on isolated plastid extracts (Scheibe et al., 1989; Srivastava and Anderson 1983). The plastidic isoform from pea chloroplasts is similar to other higher plant G6PDH (Srivastava and Anderson 1983). The native protein is a tetramer composed of four 56 kDa subunits, with a lower affinity for G-6-P and a greater affinity for NADP$^{+}$ when compared to its cytosolic counterpart.

The regulation of G6PDH in higher plants occurs at the level of translation and transcription. In alfalfa, increases in both G6PDH transcription and enzyme activity occur in response to fungal elicitors in roots and nodules (Fahrendorf et al., 1995). Increase in plastidic G6PDH expression has been observed to occur when light deprived potato shoots are illuminated (von Schaewen et al., 1995). Regulation by light/dark transitions is one way to study the metabolic changes and activity of chloroplast G6PDH (Lendzian 1978; Lendzian 1980). The role of the ferredoxin/thioredoxin system has been examined in the light dependent inactivation of G6PDH. Isolated spinach chloroplasts follow a cycle of dark activation and light inactivation of the enzyme (Scheibe et al., 1989) as a way to regulate activity. G6PDH in spinach chloroplasts is inactivated by DTT (a sulfhydryl reagent that imitates the effects of thioredoxin). Light-dependent inactivation and DTT inactivation of G6PDH may not, be achieved through the same mechanism (Scheibe et al., 1989). Root plastid G6PDH is also redox regulated in a similar fashion to spinach
chloroplasts (Wright et al., 1997). There has been limited study of the effects of metabolites on the chloroplast enzyme but the root plastid enzyme shows inhibition by ATP, UDP, citrate and acetyl-CoA (Wright et al., 1997).

**Ferredoxin-thioredoxin regulation of enzymes**

Photosynthesis fixes CO$_2$ into carbohydrates, a process that occurs in the light, while the majority of carbohydrate degradation occurs in the dark. In the light, enzymes of the Calvin cycle, or the Reductive Pentose Phosphate Pathway (RPPP) need to be activated and at the same time those of the OPPP need to be inactivated (Huppe and Turpin 1994). There are three possible mechanisms of light-dependent regulation of chloroplast enzymes 1) effector-mediated, 2) ion-mediated and 3) protein-mediated. The effector-mediated mechanism leads to an allosteric activation or inactivation of enzymes, based on light-induced changes of relevant effector metabolites. For example Rubisco (RuBP-ribulose-1,5-bisphosphate) is activated in the light, by an increased level of NADPH and ATP. RuBP is further activated by intermediates of the Calvin cycle. Enzymes regulated by an ion-mediated mechanism change activity by altering pH or regulatory cations like Mg$^{2+}$. Some enzymes of the Calvin cycle have strict pH optima. In chloroplasts, favourable alkaline conditions occur only in the light. The third mechanism of protein-mediated regulation involve enzyme modification by another protein thioredoxin (Trx) that alters activity. Light regulation of enzymes using the ferredoxin-thioredoxin system is an example where the change in the redox balance occurs, increasing reduced reductant molecule levels upon illumination. (Buchanan, 1980). Thioredoxin regulates enzyme activity
by changing the redox status, specifically of disulfide bridges between cysteine residues, of target proteins.

The ferredoxin-thioredoxin system consists of ferredoxin, thioredoxin and ferredoxin-thioredoxin reductase (Buchanan et al., 1994). Ferrodoxin is the terminal protein electron acceptor of the photosynthetic electron transport chain. Excitation of chlorophyll reaction centers namely photosystem one (PSI) by light, transfers electrons to ferredoxin and then to target acceptors like Fd-dependent enzymes and \( \text{NADP}^+ \) which are reduced. Electrons are passed to ferredoxin-thioredoxin reductase reducing a disulfide bond to its dithiol form. In a similar event, reduced ferredoxin-thioredoxin reductase then reduces thioredoxin which acts as the protein mediator of light regulation of a variety of enzymes, by changing their redox status.

Thioredoxin (12 kDa) is a highly soluble protein that has been detected in plant, animal and bacterial systems; it regulates many cellular processes in photosynthetic and heterotrophic systems (Buchanan et al., 1994). In photosynthetic systems, reduction by thioredoxin generally leads to activation of an enzyme. One notable, yet not surprising, exception is G6PDH, in the light, reduced thioredoxin inactivates this enzyme of the OPPP.

This type of redox regulation by thioredoxin can be reversed. The target enzymes must be re-oxidized in the dark in order inactivate the light-dependent pathways and to activate dark-dependent pathways. The mechanism of oxidation of these proteins in the dark is not well characterized and may occur through different mechanisms, including the oxidation of thioredoxin and subsequent oxidation of reduced enzymes (Buchanan et al., 1994).
Wenderoth et al., (1997) identified the cysteine residues that were responsible for redox regulation in a higher plant chloroplast G6PDH by using site-directed mutagenesis.

Chloroplast G6PDH homologues all show the inactivation by redox modification of the ferredoxin-thioredoxin system (Scheibe, 1990; Buchanan, 1991). The redox chain pathway guarantees that specific target enzymes are activated and kept in the reduced state in the light. G6PDH activity is turned off and is only active in its oxidized state in the dark when photosynthetic electron flow ceases. The molecular basis for this control mechanism is disulfide-dithiol interchange of certain regulatory cysteine residues in the target enzymes. The reason for such regulation prevents carbohydrate breakdown by the OPPP that would work against concurrent CO$_2$ fixation by the Calvin cycle (Wendt et al., 1999).

The ferredoxin-thioredoxin system in green plastids, leucoplasts and green plastids in seeds is compared in Fig. 2. As described previously, PSI initiates a cascade of redox events that leads to an reduced G6PDH and a cessation of NADP$^+$ production. Conversely in the dark, G6PDH would be oxidized and resume production of NADP$^+$. Green plasids in seeds, also have this control mechanism present. Fig. 2 shows the additional biochemical steps to produce fatty acids with the initial molecules of glucose and pyruvate being transported across the plastid membrane and G6PDH being light dependent through PSI and Ferredoxin. Lastly in leucoplasts, sucrose is converted in the cytosol to malate, transported across the membrane and then converted to pyruvate where it follows similar steps to NADH production.
Figure 2. Comparison between Green Plastids, Leucoplasts and Green Plastids in Seeds.
Research Objectives

My objective was to clone and overexpress one leucoplast glucose-6-phosphate dehydrogenase (G6PDH) from castor oil seed (COS). Studying this enzyme is important for a better understanding of oil biosynthesis in an agriculturally important, commercial plant. By cloning and characterizing this enzyme, the pathways responsible for oil production may be better understood and eventually altered through genetic manipulation. In addition, this work is important to confirm that G6PDH is present and active in COS leucoplast as previous research was not successful in finding significant amounts of this enzyme. This may be due to immediate protease degradation of the enzyme during purification procedures which rendered it inactive before assays could be conducted. However, the second controlling enzyme in the OPPP, namely 6PG was found to be very active in leucoplast so it should stand to reason that the major rate determining enzyme in the OPPP would be present. This work was aimed at confirming this.

The plastidic enzyme for G6PDH has been cloned from potato plants (Solanum tuberosum L.) (Graeve K.et al, 1994) This work also led to cloning, purification and characterization of the second G6PDH cytosolic form. Antibodies were raised to the potato plastidic G6PDH and these antibodies are available for use. The G6PDH antibodies from potato should cross react and provide a means of isolating castor oil seed (COS) plastidic G6PDH from a cDNA expression library. The goal is to isolate a full length clone of this enzyme from COS tissue.

Once the full-length clone is isolated, the complete sequence could be determined and compared to the sequences of other known G6PDH from higher plants using protein
alignment and phylogeny. Finally, a theoretical protein model was generated to visualize the COS enzyme. The molecular modeling will be presented based on the coordinates from *Leuconostoc mesenteroides* (Adams *et al.*, 1993).

Subcloning of the full length potato G6PDH gene into a pET-166 vector using the N-terminal histidine residues for over expression in E. coli strain BL21 (DE3) pLysS has been successful (von Schaewen *et al.*, 1995). Similarly the full length COS G6PDH was subcloned into the expression vector pGEX-4T-3 (Promega, Madison, WI). The overexpressed protein was purified using the Glutathione-S-Transferase terminal residues for affinity purification from E. coli strain BL21 (pLys). The GST tag was cleaved with thrombin prior to initial enzyme kinetic analyses. Characterization of this key enzyme of the OPPP in COS will further our understanding of fatty acid biosynthesis in this model system and in future manipulations of industrially important crops.
Chapter 2: Materials and Methods

Plant Material

Leucoplasts were isolated from developing castor seeds. (*Ricinus communis* L cv Baker 296) For propagation seeds were placed in lukewarm water for 24 hours, then planted in vermiculite (day 0). Seeds were maintained at 30°C in an incubator. On day 6, seedlings were transferred and grown under natural light supplemented with 16 hours of fluorescent light. The ambient daytime temperature was 25°C.

Leucoplast Isolation

Endosperm, stages 4-7 (20-35 days postanthesis), was ground in a mortar with sand in two volumes of homogenization buffer (50 mM HEPES-KOH (pH 7.5), 0.4 M sorbitol, 2 mM EDTA, 1 mM M902, 20 mM NaF, 1 mM DTT, 1 % (w/v) BSA, 1 % (w/v) Ficoll, 10 mM PMSF, 1 mM benzamidine, 5 mM E-caproic acid, 10 µg/mL leupeptin, and 10 µg/mL chymostatin). The pulverized endosperm was filtered through two layers of cheesecloth and centrifuged at 500 x g for five minutes. The supernatant was transferred to a fresh tube and centrifuged at 6,000 x g for 15 minutes. Lipid and supernatant were removed, and the pellet of organelles was frozen in liquid N2 and stored at -80°C (Greenwood and Bewley, 1982). In some cases, Leucoplasts were further purified from the organelle pellet before freezing following the procedure of Miernyk (1989) using the following additional steps. Purified intact leucoplasts were isolated directly instead of freezing the pellet, requiring the following additional steps. The pellet of organelles was gently resuspended in 5 mL of homogenizing buffer, and
layered on to a 10%, 22%, 35% discontinuous PFB-Percoll gradient (PFB-Percoll containing 3% (w/v) PEG 8000, 1% (w/v) BSA, and 1% (w/v) Ficoll). The organelles were centrifuged in a JS-13 rotor at 8,000 rpm for 15 minutes. The band of leucoplasts collects at the 22%-35% interface. The leucoplasts were harvested and resuspended in 20 mL of homogenization buffer (without BSA or Ficoll) before being re-pelleted at 10,000 x g for 10 minutes. The pellet was resuspended delicately in 0.5 mL of homogenization buffer (without BSA and Ficoll). Leucoplasts were lysed either by the addition of Triton X-100 to a concentration of 0.1% (w/v) or by the addition of two volumes of water (1 mL) before being gradually frozen (Boyle et al., 1986). All steps for leucoplast isolation were carried out at 4°C.

**Chemicals**

All chemicals used were from either ICN Biomedicals Inc. (St. Laurent, Quebec) or Sigma-Aldrich (Oakville, Ontario). Biochemicals used in enzymatic restriction digests, ligations, and sequence extension were from Pharmacia LKB Biotech (Dorval, Quebec). All solutions were prepared with Milli-Q purified water (Millipore, Bedford, MA, USA). Reagents used to create hybridization probes were supplied by Promega (Madison, WI, USA) Oligonucleotides used for sequencing were prepared by Cortec Inc. (Kingston, Ontario).

**Enzymatic Assays**

Enzyme activity was measured spectrophotometrically by monitoring NADPH production of G6PDH at 340 nm using a SpectraMax 250 plate reader and
SOFTmax®Pro software (Molecular Devices, Sunnyvale CA, USA). The standard assay for G6PDH contained 100 mM Tris-HCl (pH 8.0), with saturating concentrations of 2 mM G-6-P and 0.2 mM NADP⁺. Reduction of NADP⁺ to NADPH was measured by monitoring the change in absorbance at 340 nm. All assays were carried out at 25°C in a final volume of 200 µL. One unit of enzyme activity (U) is defined as the amount that catalyses the reduction of 1 µmol of NAD⁺ to NADPH per minute. The path length used for calculations was 0.60 cm (for 200 µL of liquid in a microtitre plate) (Huppe and Turpin, 1996).

**Protein Determination**

Protein concentration was determined by the method of Bradford (1976) using the Bio-Rad assay kit and bovine gamma-globulin as a standard, according to the instructions provided by the manufacturer. Bio-Rad Bradford reagent was diluted in four parts water. Diluted reagent, 230 µL, was added to 20 µL of protein solution (diluted with water as appropriate to a concentration of 0.025-0.50 µg/µL) and the absorbance was measured at 595 nm, using a Spectramax 250 plate reader.

**Protein Electrophoresis and Western Blotting**

All electrophoresis and electroblotting was performed with the Bio-Rad Mini Protean II unit and the Mini Trans-Blot cell. Molecular weight markers, polyvinylidene difluoride (PVDF) membranes, and bis-acrylamide/acrylamide were also from BioRad. Ammonium persulfate, bovine serum albumin (BSA), TEMED and bromophenol blue
were from Sigma-Aldrich (Oakville, Ontario). Methanol and glycerol were from BDH.

SDS-PAGE was performed using the Bio-Rad Mini-PROTEAN II dual slab cell according to the instructions provided by the manufacturer, based on the protocol developed originally by Laemmli (1970). A 0.75 mm thick 10% acrylamide resolving gel with a 4% stacking gel was used in all cases. Molecular weight protein standards were supplied by Bio-Rad (Richmond, CA, USA). Before loading, protein samples were boiled for 5 min, in the presence of 1% (w/v) sodium dodecyl sulfate (SDS) and 5% (v/v) 2-mercaptoethanol to ensure denaturation. Proteins were subject to electrophoresis under a constant voltage of 200 V until the bromophenol blue dye front was at bottom of the gel.

After electrophoresis, proteins were stained using Coomassie brilliant blue G-colloidal concentrate according to the manufacturer instructions (Sigma-Aldrich, Oakville, Canada). Gels were fixed for one hour in a solution of 10% (v/v) acetic acid in 40% (v/v) methanol and then stained in Coomassie solution and methanol (1:5) for 1-2 hours. To visualize proteins the gel was quickly destained with 10% (v/v) acetic acid in 25% (v/v) methanol for 60 seconds to remove excessive dye and then destained in 25% (v/v) methanol for another 24 hours.

Western blotting was performed using the Bio-Rad Mini Trans-Blot electrophoretic transfer cell. After electrophoresis, the gel, Immobilon-P nitrocellulose membranes (Millipore), filter paper and foam pads were equilibrated in blot transfer buffer for 15 minutes. Blot transfer buffer contains 10 mM CAPS in 10% methanol adjusted to pH 10.6 with KOH. The apparatus was assembled according to the
manufacturer's instructions and run at 100 V for 1 hour. After electrotransfer, the lane containing standards was cut from the membrane and Coomassie stained. The remainder of the membrane was soaked overnight in blocking buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.9, 0.05% Tween, 3% (w/v) skim milk powder, 1% (w/v) BSA). The blot was rinsed several times with TBST (150 mM NaCl, 50 mM Tris-HCl pH 7.9, 0.05% Tween). G6PDH-IgG anti-serum (diluted 1:5000 (v/v) in TBST) was added for 1 hour. The blot was rinsed several times in 1X TBST. The secondary antibody, alkaline phosphatase conjugated anti-rabbit IgG (Promega), was added for one hour. The membrane was then rinsed several times in TBST. The blot was developed with a solution containing 100 mM Tris-HCl (pH 9.8), 4 mM MgCl2, 100 mM NaCl, 0.1 mg/mL NBT and 0.05 mg/mL BCIP. Development was stopped with water.

Native PAGE was performed in a manner identical to SDS-PAGE, with the following exceptions: electrophoresis was carried out at 4°C, SDS was omitted from all solutions, and instead of using SDS loading buffer, protein samples containing 20% glycerol were loaded directly on to the gel, and samples were run for 1 hour. Native gel electrophoresis used 1 mm thick gels. Samples were kept on ice while in sample buffer, for no longer than 20 min, prior to loading onto the gel. Gels were run at 200 V constant voltage for 5 h. The protein was visualized by adding 5 mg of 3-(4, 5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) and 2.5 mg of phenazine methosulfate (PMS) (Research Organics). The NADPH produced by G6PDH reduces PMS that, in turn, reduces MTT. Reduction converts MTT from a yellow salt to a dark purple, insoluble product which precipitates onto the gel, over the native protein (Gabriel and Gersten 1992; Huppe and Turpin, 1996).
**Castor cDNA Library Screening**

The λgt 11 castor cDNA library was screened with potato polyclonal plastidic-G6PDH rabbit antiserum. Dr. D.T. Dennis (Queen’s University, Biology) received 50uL of polyclonal plastidic G6PDH antiserum from Dr. Antje von Schaewen at Universitat Osnabruck, Germany to be used as a hybridizing probe during library screening. To test the cross reactivity of this antiserum for COS G6PDH it was first used in western blot analysis of crude extracts of castor bean endosperm. This antibody cross-reacted specifically with a protein of the appropriate size, suggesting the presence of a COS plastidic G6PDH protein. The library was screened at a density of approximately 5,000-10,000 pfu/plate on *E. coli* Y1090 host strain. Plaque lifts were performed using Hybond-N+ nylon membrane (Amersham Biosciences) according to the instructions provided by the manufacturer.

The results of this screening identified multiple clones but all of these clones were not full length. The castor library was rescreened using a radiolabelled 5’ end fragment. The fragment was labeled using α(⁶³²P)-dCTP-label using the Prime-a-Gene random priming kit (Promega, Madison, WI) according to the manufacturer's instructions. Plaque lifts were performed using Hybond N+ nylon membrane (Amersham Biosciences) according to the instructions provided by the manufacturer. Hybridization was carried out for 16 hours and washes were performed using SDS and SSPE according to instructions provided by the membrane manufacturer (Amersham Biosciences). Membranes were exposed to film (Kodak) at -80°C, with intensifying screen and the film was developed after 3 days. Secondary and tertiary plaque screening was carried out to
isolate plaque pure clones. A clone of approximately 2100 base pairs was finally identified as full length.

**Phage DNA Preparation and Subcloning**

High titre stocks of the isolated λgt11 clones were used to prepare large scale liquid lysates of phage DNA (Sambrook *et al.*, 1989). Phage DNA (1-2 µg) was then digested with restriction enzymes *NotI* and *EcoRI* to release the insert. Digests were performed in a reaction mix containing 2X 1-Phor-All buffer (Amersham-Pharmacia), 0.5 mg/mL RNAase, and 0.5 µL of restriction enzyme per 20 µL digest. Digests were incubated for 3 hours at 37°C. The digested phage DNA was analyzed by agarose gel electrophoresis and fragments were ligated into *EcoRI* and *NotI* restriction digested vector, pGEM11 (Amersham Biosciences) as described by Sambrook *et al.* (1989). These ligated plasmids were then used to transform either *E. coli* JM 109 cells (for maintenance and long term storage). Transformed Colonies were grown on nutrient-rich agar plates with X-gal, IPTG and ampicillin antibiotic selection. Colonies carrying plasmids with an insert, indicated by a white colour, were selected for further analysis.

**Plasmid DNA Preparation**

The identity of inserts was confirmed by mini-plasmid preparation and plasmid isolation. Briefly, 3 mL *E.coli* cultures were grown from single plate colonies maintaining antibiotic selection, overnight. Cells were pelleted by centrifugation for 1 minute at 12,000 xg in 1.5 mL eppendorf tubes. Plasmid was isolated following the
alkaline lysis method of Birboim and Doly (1979). Inserts were liberated from plasmid backbone by performing EcoRI-Not I double restriction digests of the plasmid and performing agarose gel electrophoresis to determine which inserts were of the greatest molecular weight. For sequencing double stranded DNA of the recombinant plasmid with the largest insert size was prepared as described in the Pharmacia T7 sequencing kit handbook (Pharmacia, Montreal, PQ).

**DNA Electrophoresis and Southern Blotting**

Agarose gels, 1.0% (w/v), were used to separate DNA fragments electrophoretically as described by Sambrook *et al.* (1989). Agarose was prepared by heating dissolved agarose in 1X TrisBorate-EDTA (TBE) buffer (28 mM Tris, 28 mM boric acid, 2 mM EDTA). After cooling the solution to below 60°C, 5 µL ethidium bromide stock (10 mg/mL) was added to allow visualization under ultraviolet light. The solution was poured into a gel casting tray with comb and allowed to set. The gel was then placed in a buffer tank containing 1X TBE buffer. DNA was mixed with one fifth loading dye and loaded on to the gel. Loading dye consisted of 50% (v/v) glycerol, 100 mM Tris-HCl (pH 7.5), 10 mM EDTA (pH 8.0), and 0.05% (w/v). Lambda phage DNA digested with Hind III served as a molecular weight standard. Electrophoresis was performed at 80-100 V until the bromophenol blue dye had reached 80% of the gel length.

For Southern blot analysis the DNA was run on agarose gels prior to transfer to Hybond-N membrane (Amersham Biosciences) following the method of Southern (1975). After electrophoresis, the gel was treated in 0.25 M HCl for 30 minutes, then
in denaturation buffer for 30 minutes (1.5 M NaCl and 0.5 M NaOH) and finally in neutralization buffer (1.5 M NaCl, 0.5 M Tris-HCl pH 7.2, and 0.001 M EDTA) for 30 minutes. The gel was placed face up in contact with an overlaid Hybond-N+ membrane, two layers of blotting paper, and several cm of absorbent paper towels facilitating capillary action. The blotting sandwich was left for 12 hours under a 0.5 kilogram weight. Hybridization of the membrane after transfer was performed for 16 hours using standard stringency washes and exposure to film according to Sambrook et al. (1989).

**Sequencing and Sequence Analysis**

The dideoxy chain termination method was used to sequence double stranded DNA (Sanger et al., 1977) using α-[³⁵S] dATP label. Sequencing reactions were carried out using the T7 polymerase kit (Pharmacia) according to the manufacturer's instructions. Dideoxy terminated reactions were electrophoresed on 6.0% (w/v) polyacrylamide wedge gels and autoradiographed for up to 72 hours at room temperature. For full length castor G6PDH cDNA the sequence on each strand was repeated 2 times in both directions. Oligonucleotide primers used to extend the sequence were obtained from Cortec (Kingston, Ontario).

**Computer analysis of sequence data**

Routine sequence manipulation was performed using web-based software tools. Sequence alignments were generated by CLUSTAL W version 1.8 ([ftp://ftp.ebi.ac.uk/pub/software/unix/clustalw/](ftp://ftp.ebi.ac.uk/pub/software/unix/clustalw/)) (Thompson et al., 1994) and manually
corrected by eye. Phylogenetic relationships between the AFP sequences were established with MEGA software (http://www.megasoftware.net/) (Kumar et al., 2001). Gamma-corrected distances (a = 2) were used to compute the Neighbour joining tree. The branchpoint percentages of confidence were routinely obtained after bootstrapping 1000 times.

**Molecular modeling**

Molecular modeling of the castor G6PDH amino acid sequence was performed using a Silicon Graphics 02 workstation running SYBYL™ molecular modeling software (version 6.5, Tripos Associates, St Louis, MI, USA). A model of the protein was constructed by using the protein structure of the highest homology orthologue to castor. A comparison of the amino acid sequence of castor G6PDH to the PDB databank through The Biology Work Bench (http://biology.ncsa.uiuc.edu) revealed protein crystal co-ordinates of *Leuconostoc mesenteroides* G6PDH (Genbank accession: P11411; PDB accession: 1DPG) from which the model was built. Amino acid substitutions were introduced manually using the biopolymer menu to manually replace amino acid residues to construct the castor molecular model. The Tripos force field was used in conjunction with Gasteiger and Marsili derived charges and a distant dependent dielectric constant for the calculations (Gasteiger and Marsili, 1980). The sequence was minimized and geometrically optimized until the gradient between successive iterations was less than 0.05 kcal/mol. Molecular modeling of castor G6PDH was performed with the help of Dr. Michael Kuiper.
Phylogenetic and Divergence Analysis

Alignments of nucleotide or amino acid sequence data were carried out initially using Clustal W version 1.8 (ftp://ftp.ebi.ac.uk/pub/software/unix/clustalw/) (Thompson et al., 1994) with minor adjustments to optimize alignments. Primer sequence regions were excluded from phylogenetic analyses and distance estimations. Phylogenetic analysis of G6PDH sequences was performed with MEGA software (http://www.megasoftware.net/) (Kumar et al., 2001) using the Neighbour-Joining method (Saitou and Nei, 1987). A pair-wise deletion option for gaps was used and groupings in the Neighbor-Joining analyses were supported by bootstrap values of one thousand times. Homologue sequences used in the alignment were obtained from NCBI genbank and aligned using Clustal W. The Neighbour-Joining tree was produced using Mega v 3.0 and minimum evolutionary distance was calculated using pair-wise alignments.

Recombinant DNA techniques, subcloning and sequencing

All recombinant DNA techniques were performed using standard methods Sambrook et al., (1989). All G6PDH clones plaque purified from library screening were subcloned into plasmid pGEM11 and subsequently into pGEX-4T-3 (Amersham Biosciences) and sequencing performed on double stranded DNA using the T7 sequencing kit (Amersham Biosciences) according to the manufacturers instructions.
PGEX-4T-3 Expression Vector Construction

The castor G6PDH cDNA was sub-cloned into the pGEX-4T-3 bacterial fusion protein expression system to produce an in-frame fusion protein with the GST-tag and plant castor G6PDH. Constructs were designed such that the cleaved fusion protein would resemble the imported and processed mature protein. For subcloning castor G6PDH cDNA was digested with EcoRI and NotI and was ligated to the EcoRI and NotI digested pGEX-4T-3 vector.

Expression of GST fusion proteins from E. coli

Expression, purification and thrombin cleavage of GST fusion proteins was performed as described by the manufactures instructions (Amersham Biosciences, Pharmacia). Plasmid constructs in pGEX-4T-3 were used to transform BL21 (pLys) (Novagen) for overexpression studies by electroporation using Bio-Rad electroporation cuvettes (0.2 cm gap) under standard conditions for E. coli strains. Transformants were identified by growth on plate media containing antibiotic. Individual colonies were picked and grown in liquid culture until log phase was achieved (between 0.6- 1.0 Absorbance OD600). Protein was induced by addition of Isopropyl B-D galactoside (1 mM final IPTG) and induced for 2-6 hours. For enzyme assays, the bacteria were crushed by French Press, washed as per manufactures instructions and immediately brought up to 50% glycerol (BDH) and stored at –20°C until further purification was performed.
Purification of GST fusion proteins from *E. coli*.

Purification of GST fusion proteins from *E. coli* was carried out by crushing the bacteria using a French press apparatus and binding the lysate to an affinity matrix, Glutathione Sepharose 4B (Pharmacia) for 30 min. Unbound proteins were washed away with several volumes of 1X PBS buffer and eluted with Glutathione elution buffer (10 mM Glutathione, 50 mM Tris-HCl, pH 8.0) according to manufacturer’s directions.

Purified GST fusion protein was cleaved with bovine thrombin prepared at a concentration of 1U/uL in 1X PBS. Thrombin solution was added to purified fusion protein (1 U for every estimated 100 µg of fusion protein) and incubated for up to 16 hours at room temperature with gentle agitation. An alternate purification method bound the fusion protein to the matrix and instead of elution Thrombin was added directly to the Glutathione Sepharose pellet prior to incubation, the Sepharose beads were removed afterwards by centrifugation at 500 xg for 5 min.

Enzyme assays and determination of preliminary kinetics

Assays were done in 5mM MgCl₂, and 50 mM Tris-HCl (pH 8.0) and were carried out on GST purified (eluted from Glutathione Sepharose matrix) protein (non-cleaved) or thrombin cleaved protein. A small amount of enzyme (4uL in 50% glycerol) was used for assays as the concentration of the crude bacterial extracts was high. (17.7u/ml/min).
Chapter 3: Results

G6PDH has so far not been found in the leucoplasts of castor oil seeds, but its presence is inferred because plastids have significant levels of 6-phosphogluconate dehydrogenase activity, the second enzyme in the OPPP. The inability to detect plastidic G6PDH in COS could probably be due to the enzyme being unstable with the subsequent loss of activity during the purification process.

The instability of the leucoplast G6PDH has made it impossible to study its kinetics (Simcox et al., 1977). Initially, enzyme kinetics were attempted with partially purified protein samples from leucoplasts with little success. In order to get enough purified protein for enzyme kinetics the gene for castor seed G6PDH was isolated, sequenced and subcloned into an overexpression system as a means to overcome these problems.

Antibody library screening is sensitive and polyclonal antibodies erroneously cross reacting to abundantly expressed proteins could give a high background of false positive clones. Two different antibodies were available for library screening antibodies against yeast G6PDH and potato G6PDH. Prior to antibody screening of the λgt11 cDNA library the specificity of the two antibodies were evaluated by western blot analysis to test antibodies to G6PDH from potato and yeast to see which would cross react best to castor oil seed homogenates. Figure 3 shows the results of western blot analysis. Both antibodies cross reacted well with G6PDH from purified leucoplast preparations (Figure 3A, lane 1 and figure 3B lane 1). Antibodies also cross reacted with
Figure 3. Western blot to assess two different antibodies cross reactivity to castor homogenate prior to cDNA library screening. A) Potato anti-G6PDH hybridized to various protein samples, lanes are: prestained protein molecular weight markers (M), Leucoplast (1), isolated envelope (2), Chloroplast extract (3), Castor endosperm (4), Castor root (5), Young leucoplast membrane (6), Yong leucoplast stroma (7), Yong leucoplast whole (8), Brassica napas 2.5 week old seed (9). B) Yeast anti-G6PDH hybridized to various protein samples, lanes are: Leucoplast (1), isolated envelope (2), Chloroplast extract (3), Castor endosperm (4), Castor root (5), Yong leucoplast membrane (6), Yong leucoplast stroma (7), Yong leucoplast whole (8), Brassica napas 2.5 week old seed (9). The arrow to the right points to the correctly sized band for G6PDH, the angled arrow on the blot points to the castor seed cross reacting protein band.
castor endosperm but showed little reaction to castor root (Figure 3 A/B lanes 3 and 4).
Both potato and yeast antiserum reacted with young leucoplast membrane (Figure 3A lane 6 and figure 3B lane 6), young leucoplast stroma (Figure 3 A lane 7 and figure 3B lane 7) and young whole leucoplast extract (Figure 3A lane 8 and figure 3B lane 8) as expected. Brasica napas 2.5 week old seed also showed weak cross reactivity to both potato and yeast G6PDH antiserum (Figure 3A lane 9 and figure 3B lane 9).

In total, 24 clones were isolated from the first antibody screening and all positive phage were plaque purified (small scale isolation) and phage DNA isolated for analysis. Inserts were removed from the λgt11 the phage by digestion with restriction enzymes EcoR1 and Not1. Inserts were ligated into pGEM 11 plasmid and transformed into JM109 bacteria. Analysis of the inserts from the plasmids revealed that most of the clones appeared to be shorter in base pair length and molecular weight as judged by agarose gel electrophoresis ranging in size between 700 and 1200 bases pairs, the longest insert was ~1400 base pairs. The longest phage insert was sequenced confirming that it was indeed G6PDH but it was not full length missing the 5’- end. The library was rescreened several times with antibody and plaque purified phage, however did not yield any inserts longer than the longest insert isolated from the first screening attempt. To find a cDNA clone with a complete 5’ coding sequence, a 200 base pair fragment was used from the 5’ end of the longest clone for library screening to avoid the shorter G6PDH cDNAs in the λgt11 library. The 5’ end fragment was labeled with a-32P-dCTP and used as a hybridization probe for library screening. This method proved successful with the isolation of a longer clone 2100 bp in length. In hindsight the inability to easily find a cDNA clone with the 5’- end was due to the cDNA sequence that harboured an
internal Eco R1 restriction site 162 bases pairs from the Met start site. This internal Eco R1 restriction site complicated the isolation of full length cDNA as many of the cDNAs were truncated during cDNA library construction as the cDNA insert pool were digested with both Eco R1 and Not 1 prior to ligation to the prepared phage arms.

The complete sequence of the castor G6PDH cDNA is shown in Figure 4. The coding sequence is 1803 base pairs in length encoding a protein that is 600 amino acids in length. The ChloroP server predicts the presence of chloroplast transit peptides (cTP) in protein sequences and the location of potential cTP cleavage sites (ChloroP 1.1 Server, http://www.cbs.dtu.dk/services/ChloroP/). The ChloroP algorithm predicts a chloroplast transit peptide of 61 amino acids, suggesting a cleavage between Val_{61} and Leu_{62}.

A second algorithm, TargetP, predicts the subcellular location of eukaryotic protein sequences (TargetP v.1.01 Server, http://www.cbs.dtu.dk/services/TargetP/). The subcellular location assignment is based on the predicted presence of any of the N-terminal presequences such as chloroplast transit peptide, mitochondrial targeting peptide or secretory pathway signal peptide. TargetP also predicts that the sequence contains a chloroplast transit peptide of 61 amino acids (Emanuelsson et al., 1999).

With its transit peptide the protein is predicted to have a theoretical pI of 7.18 and a molecular weight of 68,123.4 kDa. Without the signal sequence attached the length is 539 amino acids the pI is predicted to be 6.20 with a molecular weight of 61379.9 kDa (ProtParam Server, http://us.expasy.org/tools/protparam.html).

A number of G6PDH sequences have been reported in the literature. The full length castor sequence was compared using phylogenetic analysis to 25 other homolgous genes. The dendogram (Figure 5) shows that G6PDH evolution shows a
Figure 4. DNA sequence of a full length G6PDH cDNA isolated from the castor cDNA library. The restriction enzyme sites at the 5’ end (Eco RI) and the 3’ end (Not I) are shown underlined as well as the internal Eco RI site 162bp from the Met start site. Nucleotide sequence is in lower case with numbering to the right for the 1,803 base pair, full length cDNA. The amino acid translation in upper case letters is shown underneath. The Met amino acid start is bold and the nucleotides for the stop codon (tga) are bold.
Eco R1

gaattccgttgcgtgcgtgaagttgaaattgttttaaacccaaacccaccccaaggt
ccttgaaagaa

M A T L S S R S Y A N S N S Y S L
ttttcttttctttctacaatcaatggaagcaatcttcaggttgaatgctgttc 120
F S S S S S I N G K Q F Q R L N W I A
aataagttcttaacagctaaggttcttttcataactgctaagaaatattttta 180
N K F L T A K V Q A Q K N S Y P D V
gtttcagtcaggtgagttgaagctgcaccaccccaaggtgaardacagtctga 240
V L M Q D G A V A T P V N P V E N D S S
tttatgaaattgaaagatgttgttctcttttcatctacaggaattaaaagaa 300
F M K L D G L L S S I S S E E L K E
gaagttgttgtatatttaaaagatgctacatctctcagttggtgtg 360
E V G F D I N K D E S T V S I T V G A
tctggagaccttgccaaagaaaaagatttttcgacactttttggctttattattagaggtt 420
S G D L A K K I F P A L F A L Y Y E G
tgtctcccaacaggacacaccagctgtttggtttgtttatgctgcagtaagttgctg 480
C L P K H N T V F G Y A R S K M T D A E
cctagaaacatgattagcaagacctgcacgtaagataagagggagagttgttgtt 540
L R N M I S K T L T C R I D K R E N C G
gaaagatgagcaggtttcattataagatgtttctctacctctgctgtaagttccccg 600
E K M D E P L N R C F Y H S G Q Y D S Q
gacacttttcgaagactgaagagttgaggtgagttgccggtttctcag 660
E H F A E L D K K L K L E H E G G R V S N
cgcctttttttctcatcatacccccccacaataatttttgtagcagtataaaatgtgcgagc 720
R L F Y L S I P N I F V D A V K C A S
tctacagttcactctgtatgctctgtgacaccagcgttatttgagaaaccgtttgcgq 780
S S A S S G N G W T S V I V E K F P G R
gattcagaacctctcctgtctcttccaaagacatcattctagaaagagatcataaa 840
D S E S S A A L K T A L K Q Y L E D Q
atatttaggatatagcctcttggaagaaaaagactattcagtgaaatctttctgctcgc 900
I F R I D H Y L G K E L V E N L S V L R
tttccaaatctataacccactttttgttgcaggcaatgtatataaaagattcagctatga 960
F S N L I P E L W S R Q Y I R N V Q L
atattctcgaagatttttcgcaccagctgagttgaggttattttgataatattggata 1020
I F S E D F G T E G R G Y F D N Y G I
taaagagataataatcgacacatgtctcttttcaatcttcagctttgacccatgaaacc 1080
I R D I M Q N H L L Q I L A L F A M E T
ccctgctattttggcatcagatctatccagttttactaatagttttttttttttttttttgg 1140
P V S L D A E D I R N E K V K V L R S M
agccctatacgcctgaaagatgttctgagtgttctgcattacagactcaacactaaagggaggg 1200
R P I R E D V M I G Q Y K S H T K G G
attacactccacgctaccattgtagacaaaaagcttccatactgcccaaaccc 1260
I T Y P A I D K T V P K D S L T P T
tttgcaagcagcagccctctctctatatcagacaaacggaagatggtgaccttccctatg 1320
F A A A A L F I D N A R W D G V P F L M
aaagctgccaagcattacatataaagagcgactctagctagctgctacagctgtgtg 1380
K A G K A H N K R T E I R V Q F R H V
ccctggaacttatataaaacgcgtcttcctgatgcttatgatgataaaagctatcataat 1440
P G N L Y N R N F G S D I D K A T N E L
gttatcgcgagttcgcgtgagctatcacttgaagctacattaaacaggttccccgqgt 1500
V I R V Q P D E A I Y L K I N N K V P G
aaagttgtgacattgatatgctgagactttagaaggaacagagaaacttaatgtatttata
catgaatgttttattagatcacttaacaagagctctctactatagtaactatccatgggaagt
caggagcaagcaagtagataatccacatttcactatcttccacttttaagttac
agtatatatttttttgggctttctttctttctttccacgcagttctgtacattacgtgtaaagagctcagctactatcgttttttgggctttctttctttttgggctttctttttttgggctttctttttttgggctttctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Figure 5. Dendogram alignment of 25 different G6PD amino acid sequences compared to castor oil seed G6PD. A Neighbour joining tree (unrooted) was constructed using amino acid sequences p-distance estimation, a bootstrap value =1000 and with the pairwise deletion option. Bootstrap frequencies are shown at branchpoints. The initial alignment was produced using Clustal W v 1.8 and the dendogram tree was made with MEGA v 3.0. Cytosolic (CYT) and plastidic (P1 or P2) forms are indicated. Common group names are shown bracketed to the right of the alignment. Accession numbers used are: Anabaena, (Anabaena sp. PCC 7120) P48992; Arabidopsis_CYT (Arabidopsis thaliana) AJ010970; Arabidopsis_P1 (Arabidopsis thaliana) AJ001359; Arabidopsis_P2 (Arabidopsis thaliana) AC002396; Cyanidium (Cyanidium caldarium) AJ006246; Drosophila (Drosophila melanogaster) M26674; Dunaliella (Dunaliella bioculata) AJ132346; Escherichia (Escherichia coli) P22992; Haemophilus (Haemophilus influenzae) P44311; Human (Homo sapiens) P11413; Leuconostoc (Leuconostoc mesenteroides) P11411; Nicotiana_CYT (Nicotiana tabacum) AJ001770; Nicotiana_P1 (Nicotiana tabacum) AJ001772; Nicotiana_P2 (Nicotiana tabacum) X99405; Nostoc (Nostoc punctiforme) P48848; Rat (Rattus norvegicus) X07467; Saccharomyces (Saccharomyces cerevisiae) P11412; Schizosaccharomyces (Schizosaccharomyces pombe) AL023595; Solanum_CYT (Solanum tuberosum) X74421; Solanum_P1 (Solanum tuberosum) X83923; Solanum_P2 (Solanum tuberosum) AJ010712; Spinacia_P1 (Spinacia oleracea) O24357; Synechoccus (Synechococcus elongatus PCC 7942) P29686; Synechocystis (Synechocystis sp. PCC 6803) P73411; Parsley_P1 (Petroselinum crispum) AAB69317.
major split in homology between bacteria and cyanobacteria and eukaryotes. Among eukaryotes yeast, invertebrates, vertebrates and higher plant cytosolic isoforms appear to have diverged from the higher plant plastidic sequences. The higher plant plastidic branch itself shows a marked split between plastidic P1 and P2 forms.

Eight plastidic G6PDH clones have been previously reported from higher plants which fall into two isozyme classes, P1 and P2 based on their signal sequence identity (minus transit peptide). G6PDH isoforms between tobacco and potato share about 92 to 95% identity much higher than when comparing either of these two to the castor amino acid sequence. The full length clone isolated by library screening from castor seed falls into the P2 class of G6PDH isozymes as it has higher homologies to the tobacco (TOB P2), potato (POT P2), parsley (PAR P2) and arabidopsis (ARA P2) isozymes with 78%, 78%, 76%, 75% amino acid similarity respectively. There is less homology to the second isoform class of tobacco (TOB P1), potato (POT P1), arabidopsis (ARA P1) and spinach (SPI) with sharing 69%, 70%, 68% and 70% similarity respectively to the castor sequence.

The cleavage site of the transit peptide in castor P2 G6PDH’s shown in Figure 6 reveals an almost conserved motif with TOB P2, POT P2 and PAR P2 to which it shares the highest degrees of homologies and thus it is likely that the cleavage site is positioned correctly where indicated for castor P2 in Figure 6. The castor sequence shows greatest homology to the other higher plant P2 plastidic forms grouping out with the tobacco, potato, parsley and arabidopsis P2 sequences (Figure 5 and 6).
All nine clones share the same sequence for both the coenzyme and active sites indicating the importance of these functional regions. Of interest are the cysteines which are involved in the redox regulation of this protein. Of the seven cysteines found in the castor sequence, those at position 141, 171, 179, and 189 are conserved throughout the other eight amino acid sequences compared. Cysteine 238 while conserved in all other sequences is not seen in parsley (PAR P2), it has a serine instead. Conserved cysteines at positions 171 and 179 have been shown to be involved in the redox regulation of potato G6PDH. The two cysteines at positions 524 and 596 are near the C-terminal end of the castor protein but are not conserved in the other clones. P1 forms appear to have several other cysteine residues that show little conservation to cysteines in P2 forms (Spinach P1 Cys 182; Potato P1 Cys 194; Tobacco P1 Cys 260 and Arabadopsis P1 Cys 249). The coenzyme binding site (GASGDL) and the active site (RIDHYLGKE) are absolutely conserved among the 9 higher plant G6PDH homologues (Figure 6).

A theoretical three dimensional model of castor G6PDH was determined for the full length sequence.. This was made possible by the discovery of the *Leuconostoc mesenteroides* G6PDH (Genbank accession: A39864, Swiss prot: 1DPG) whose 3D structure has been determined to two-angstrom resolution. (Adams et al., 1993) This allowed for homology modeling of the castor plastidic enzyme, which is the most reliable method of tertiary structure prediction (Figure 7). Comparisions of the active site, ligand properties and other binding sites can now be made.
Figure 6. Comparison of castor plastidic to other eight other higher plant G6PDH enzymes. The amino acid sequence was aligned with Clustal W v 1.8. Arrowsheads pointing to the bottom left hand corner show the area of signal sequence cleavage. Region A boxed is the conserved coenzyme binding site (GASGDL), region B boxed is the conserved active site (RIDHYLGKE). Cysteine residues are highlighted in red. Conserved amino acid are shown on the consensus sequence below the alignment. The sequences compared are castor.P2 (castor), Tob. (Tobacco), Pot. (Potato), Par (Parsley), Ara (Arabidopsis) and Spi (Spinach). P1 and P2 refer to the two plastidic forms identified in higher plants.
Figure 7. A) Theoretical three dimensional model of castor G6PDH modeled on *L. mesenteroides* enzyme backbone with amino acid substitutions/insertions. Alpha helices shown in red, Beta sheets in blue, Unassigned protein loops and turns in yellow. B) The three-dimensional structure of glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides* refined at 2.0 Å resolution.
The full length clone of castor plastidic G6PDH from the cDNA library was obtained in two fragments due to the internal Eco RI site at 162 bp. This gave rise to a 162 bp EcoRI fragment and a longer 1638 bp Eco RI/Not I piece. Since the internal EcoRI site was very close to the putative signal sequence cleavage site of the mature protein and no closer restriction site could be found or was possible, this fragment was inserted into the bacterial E.coli pGEX-4T-3 overexpression vector. The resulting production of the fusion protein with the addition of IPTG contains the glutathione s-transferase protein tag at the N-terminus (29kDa) and the mature plastidic G6PDH protein (65 kDa) adjacent to it. The castor G6PDH cDNA was sub-cloned into the pGEX-4T-3 bacterial fusion protein expression system to produce an in-frame fusion protein with the GST-tag and the castor G6PDH. Constructs were designed such that the cleaved fusion protein would resemble the imported and processed mature protein. When cleaved with thrombin the sequence is nearly exactly the same as the mature protein with the exception of the first 2 amino acids L,M that that are replaced by 3 overhanging amino acids P, N, and S in the new construct. The fusion protein (94 kDa) could be purified. In crude extracts, its activity could be maintained indefinitely in 50% glycerol at –20°C.

Purification of G6PDH from leucoplast castor endosperm has not been successful to date due to the highly unstable nature of this isoform. Initial attempts included leucoplasts being isolated from endosperm of castor and G6PDH activity being assayed. G6PDH in crude samples was complicated by the presence of endogenous 6-phosphogluconate dehydrogenase (6-PGDH), which oxidizes gluconate-6-phosphate with
the production of NADPH, leading to an overestimate of the G6PDH in the sample. It appears to have a pH optimal around 8.0.

The use of the pGEX overexpression system has allowed the purification of castor G6PDH overexpressed in E.coli. Crude bacterial extracts were prepared as described above in 50% glycerol and stored at \(-20^\circ\)C. Control values of just the pGEX-4T-3 vector in bacteria without the CB P2 fragment showed around 1/100\(^{th}\) the activity of the overexpressed extracts. (Data not shown). This crude extract was then passed over a PD 10 desalting column and immediately put onto the Sepharose 4B column that specifically binds the GST tag. The protein was then eluted off of the column with a volume of reduced glutathione and resulted in active protein that was once again stored in 50% glycerol at \(-20^\circ\)C. Preliminary findings on the activation control of this partially purified enzyme is contrary to all published data in that this G6PDH is activated in the presence of DTT and reducing agents including reduced glutathione and deactivated in the presence of oxidizing agents such as sodium tetrathionate. All assays were done in 5mM MgCl\(_2\) and 50 mM Tris-HCl (pH 8.0) and were carried out on the crude non-cleaved fusion protein.

Analysis of castor P2 G6PDH was performed on enzyme produced by overexpression in E. coli because the low amounts of detectable castor G6PDH in COS endosperm on western blots and failed isolation attempts from COS endosperm. (Figure 8). A small amount of overexpressed enzyme (4uL in 50% glycerol) could be used for assays as the concentration in the crude bacterial extracts was high. (17.7u/ml/min). Western blots were performed to show expression of the fusion protein (96 kDa) (Figure 8. A) or after thrombin cleavage (66 kDa) (Figure 8. B).
Native gel slices were used to determine the activity of overexpressed G6PD. As castor G6PDH appears to belong to the chloroplast-like P2 family the exposure to DTT which would simulate the effects of light exposure through the effects of thioredoxin should have decreased enzyme activity compared to the control native gel slices without DTT added. Surprisingly the addition of DTT appears to increase enzyme activity up to 48% (Figure 9). The effect of DTT was tested at various concentrations of DTT on purified G6PDH. Increasing the DTT concentration showed a major increase in G6PDH activity (Figure 10).

Finally, this DTT effect was tested on overexpressed protein and castor plastid extract to see whether the effect was an artifact of the overexpressed fusion protein. The stepwise addition of the following chemicals and enzyme (protein, NADP and buffer, then addition of 20mM DTT, then addition of 10mM 6PG then 6mM G6P) to the same reaction mixture in the same cuvette were performed. The addition of DTT in the absence of other substrates (Figure 11 lane 2) in plastid or expressed protein alone did not show any activity. Adding a saturating concentration of 6PG showed the plastids contained 6PGDH as determined by NADP+ reduction. The overexpressed enzyme did not have this activity. Finally the addition of saturating levels of 6 mM G6P to the reaction mixture of the purified overexpressed protein increased the activity almost ten fold compared to the slight increase in the plastid extract showing that the active amount of G6PDH in the native preparation was significantly less than the active amount in the overexpressed sample. (Figure 11)
Figure 8. Western blot analysis of bacterially expressed G6PDH.

A) Overexpression of G6PDH in BL21 bacteria transformed with pGEX-4T-3-castor G6PDH constructs. Lanes: (wt) wild type (untransformed) BL21 control lysate. Lanes 1-8 show bacterial transformants #1-8. Arrow to the right shows the location of the overexpressed GST-G6PDH fusion protein (96 kDa). Potato antiserum was used as the primary antibody at a 1:1000 dilution.

B) Western blot with potato antiserum for G6PDH control and overexpressed bacteria (transformant #8 from A). Lanes: (1) control BL21 lysate; (2) G6PDH expressed BL21 lysate; (3) control pellet (insoluble protein); (4) overexpression pellet (insoluble protein); (5) GST sepharose matrix unbound protein control; (6) GST sepharose matrix unbound protein overexpressed; (7) control matrix bound protein after treatment with thrombin (8) overexpressed matrix bound protein after treatment with thrombin (9) control matrix supernatant after thrombin digestion; (10) overexpression matrix supernatant after thrombin digestion. Potato antiserum was used as the primary antibody at a 1:1000 dilution.
A.

![Image of protein gel with molecular weight markers and labeled bands at 96 kDa and 66 kDa.]

B.

![Image of protein gel with molecular weight markers and labeled bands at 96 kDa and 66 kDa.]

96 kDa (Fusion Protein -with GST tag)

66 kDa (Cleaved protein -minus GST tag)
Figure 9. Native gel slices correlated to G6PDH activity. Inset shows the native gel lane with gel slices excised every 2 mm. Gel slices were analysed for G6PDH activity by monitoring the OD340 value.
Native Gel Activity

Slice Number vs umoles/min NAD(P)H

- Control
- 20 mM DTT

Native Gel Slices

1
2
3
4
5
6
7
8
9
10
11
12
13
14
Figure 10. Graph showing the effect of DTT on G6PDH activity. Increasing concentrations of DTT (0 to 20 mM) were assayed. Activity was measured by optical density at 340 nm, values for different DTT concentrations are shown in the table below. Several runs were completed but data is displayed from one specific run.
Effect of DTT on G6PDH Activity

<table>
<thead>
<tr>
<th>Conc. DTT (mM)</th>
<th>umoles/min NADPH formed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>71.25</td>
</tr>
<tr>
<td>1</td>
<td>143.47</td>
</tr>
<tr>
<td>5</td>
<td>159.64</td>
</tr>
<tr>
<td>10</td>
<td>171.34</td>
</tr>
<tr>
<td>20</td>
<td>176.67</td>
</tr>
</tbody>
</table>
Figure 11. Graph showing the effect of DTT or 6-P-G on G6PDH activity in overexpressed protein and plastid extracts. NOTE: The chemicals were added sequentially to the same reaction mixture in the same cuvette. Treatments included 1: NADP and buffer; 2) addition of 20 mM DTT ; 3) addition of 10 mM 6-P-G; 4) addition of 6 mM G-6-P. Activity was measured by optical density at 340 nm, exact values for extracts shown below converted to umoles/min. Several runs were completed but graph shows data from one specific run.
NADPH Production in Overexpressed Protein vs. Plastid Extract

1 = Protein + NADP + Buffer
2 = Above + 20mMDTT
3 = Above + 10mM 6PG
4 = Above + 6 mM G6P

<table>
<thead>
<tr>
<th>Addition</th>
<th>Overexpressed Protein umoles/min NADPH</th>
<th>Plastid Extract umoles/min NADPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1= Protein + NADP + Buffer</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2= Above + 20mMDTT</td>
<td>0.095</td>
<td>0.247</td>
</tr>
<tr>
<td>3= Above + 10mM 6PG</td>
<td>0.070</td>
<td>19.421</td>
</tr>
<tr>
<td>4= Above + 6 mM G6P</td>
<td>49.296</td>
<td>27.452</td>
</tr>
</tbody>
</table>
The purified overexpressed G6PDH protein enabled Km values to be obtained as shown in Figures 12 and 13 and are compared to other plant values in Table 1. The Km for NADP+ and G-6-P are very different compared to potato, pea leaves and black gram. (Table 1.) Castor G6PDH has a km of 38 uM for NADP+ as compared to only 2.4 uM for the chloroplast enzyme in pea leaves. The Km for G-6-P was 930 uM for castor G6PDH compared to only 370 uM for the chloroplast enzyme, again a much higher value. This significant difference suggest different interactions with substrates of G6PDH in castor leucoplasts compared to other plastid enzymes.
Figure 12. Substrate activity curve for G6PDH with NADP$^+$ (mM). The optical density at 340 nm was monitored for various concentrations to obtain the Km value. One run was completed for this determination.
Lineweaver-Burk NADPH

\[
y = 0.0004x + 0.0104
\]
Figure 13. Graph of substrate saturation of G6PDH for G-6-P (mM). Activity was monitored by optical density at 340 nm. Various concentrations of G-6-P were used to determine the Km value. One run was completed for this determination.
Km for Glucose-6-P

umoles/min NADPH

[G-6-P] mM
y = 0.0041x + 0.0078
Table 1. Comparison of Castor Bean Plastidic G6PDH kinetics with the enzyme from other plant species. Note, Kms were derived using Marquardt-Levenberg minimization

<table>
<thead>
<tr>
<th>Species</th>
<th>Location</th>
<th>Km for NADP(^+) (uM)</th>
<th>Km for G-6P (uM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castor Bean</td>
<td>Leucoplast</td>
<td>38</td>
<td>930</td>
</tr>
<tr>
<td><em>Ricinus communis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potato Tuber</td>
<td>Cytosol</td>
<td>6</td>
<td>260</td>
</tr>
<tr>
<td><em>Solanum tuberosum</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pea Leaves</td>
<td>Cytosol</td>
<td>13.5</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>Chloroplast</td>
<td>2.4</td>
<td>370</td>
</tr>
<tr>
<td>Black Gram</td>
<td>Cytosol</td>
<td>14</td>
<td>160</td>
</tr>
<tr>
<td><em>Phaseolus mungo</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4: Discussion

In developing endosperm of COS, Glucose-6-phosphate dehydrogenase (D-glucose 6-phosphate: NADP oxidoreductase, EC 1.1.1.49) catalyses the oxidation of glucose 6-phosphate to δ-glucono-1,5-lactone-6-phosphate (6-phosphogluconate) with the concomitant reduction of NADP⁺ to NADPH. This and the second step of the OPPP produce reducing equivalents in the form of NADPH and are probably the major means of providing reductant for biosynthesis under non-photosynthetic conditions, such as in darkened leaves or in non-photosynthetic tissues. In addition, the OPPP provides erythrose 4-phosphate, the precursor to the shikimate pathway (lignin) and ribose 5-phosphate for nucleic acid biosynthesis.

Numerous attempts over 20 years to detect glucose-6-phosphate dehydrogenase (G6PDH) in leucoplasts from developing castor seed endosperm have been unsuccessful even though the second enzyme of the pathway 6-phosphogluconate dehydrogenase (6PGDH) was present at high concentration. It was even suggested that 6-phosphogluconate was transported across the leucoplast membrane. The report was an attempt to re-evaluate the absence of the enzyme using more advanced techniques of molecular biology. It is clear from the results presented in this report that leucoplasts do contain G6PDH but the kinetics and regulation of the enzyme is different from other plastid G6PDH’s and this could account for the previous failures to detect it.

An important difference between the plastid and cytosolic isozymes is the thiol-mediated regulation of the chloroplast enzyme activity. The plastidic isozyme is rapidly inactivated in the light to prevent futile cycling between the photosynthetic reductive and
the OPPP. Inactivation occurs in the light via a thioredoxin-mediated mechanism in which key conserved cysteine residues are reduced and these cysteines are reoxidized in the dark, thereby reactivating this enzyme. This is an unusual characteristic of plastidic G6PDH in that most other chloroplast enzymes such as NADP-malate dehydrogenase and fructose-1,6-bisphosphatase are activated by thioredoxin mediated reduction. It is interesting to note that this method of regulation also occurs in some forms of cyanobacteria but none of the cysteine positions are conserved compared with the plant isoforms. This suggests functional convergence of redox regulation in G6PDH isoforms of cyanobacteria and higher plants (Wendt et al., 1999).

A small number of plant plastic G6PDH cDNA clones have been previously reported in the literature that fall into two distinct isozyme classes either P1 or P2 based on their signal sequence identity and by the degree of sequence homology. A sequence homology search using NCBI BLAST alignment algorithm (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi) with the castor oil seed G6PDH amino acid sequence showed that the best homology was to other higher plant plastidic forms although lower homologies were detected to many other homologues of the enzyme. Alignments of the deduced amino acid sequence to 25 other known G6PDH homologues show that the castor sequence groups with the higher plant P2 isoforms. The isolated, full length cDNA encoding castor bean plastidic G6PDH appears to belong to the P2 enzyme class based on its phylogeny to other higher plant isoforms. This was confirmed by using homology based searches to other transit sequence databases (ChloroP, TargetP) using the castor signal sequence region only.
Previously, isoforms of G6PDH have been characterized cloned, purified and sequenced from potato cytosol and potato leaf chloroplast (von Schaewen et al., 1995), tobacco chloroplast (Knight et al., 2001) and protein has been purified from cytoplasm and chloroplast from pea leaf (Anderson et al., 1983 and Semenikhina et al., 1999). Similarly, the cDNA from castor endosperm was isolated by screening of an expression library to obtain full length cDNA and subsequent construction of an over expression vector in *E. coli*. Preliminary kinetic data is presented on crude bacterial extracts. The amino acid sequence of castor plastidic G6PDH compared with plastidic clones from other higher plants reveal a high number of conserved residues.

As indicated by its signal sequence this G6PDH sequence belongs to the P2 class of isozymes and as predicted it has higher homologies (% identity and % similarity) to the tobacco (TOB P2), potato (POT P2), arabidopsis (ARA P2) and parsley (PET) P2 class of isozymes at 78%, 78%, 75% and 76% respectively. It has less homology to TOB P1, POT P1, ARA P1 and spinach (SPI) enzymes at 69%, 70%, 68% and 70% respectively. However, the G6PDH coding sequences from higher plants share the same conserved sequences for the coenzyme binding site (at aa 119-124) and the active site (at aa 283-291) and this is also the case with the cDNA isolated from castor. Of interest are the cysteine residues that are involved in the redox regulation of this protein. Seven cysteines, 141, 171, 179, and 189 are conserved throughout the nine clones while cysteine 238 is not seen in parsley (PET). Cysteines 171 and 179 have been shown to be involved in the redox regulation of potato G6PDH. In castor G6PDH the cysteines at positions 524 and 596 are found near the C-terminal end of the protein but these two sites are not conserved in any of the other higher plant amino acid
sequences. Molecular modeling would suggest that these two cysteines (524 and 596) likely are not free but form a disulfide bond due to their tertiary structure proximity.

Wendt et al., 1999 identified the cysteine residues that were responsible for redox regulation in a higher plant chloroplast G6PDH by site-directed mutagenesis (Wenderoth et al., 1997). Since G6PDH in cyanobacteria is also regulated by disulfide-dithiol interchange mediated by thioredoxin (Cossar et al., 1984) it was surprising that none of the cysteine positions in the potato sequence were conserved with those found in cyanobacterial G6PDH sequences (von Schaewen et al., 1995; Wenderoth et al., 1997). This raised the question of how redox regulation evolved in G6PDH enzymes of photoautotrophic eukaryotes. Molecular modeling shows that the mature castor coding sequence contains seven cysteine residues but some probably do not form disulfide bonds because of their distant location on the protein.

The DTT added to G6PDH assays did not decrease activity of the enzyme, in fact it increased the activity and this was contradictory to published reports of chloroplast G6PDH. It has been reported plastid G6PD is highly unstable and enzyme purification is nearly impossible from some tissues due to the labile nature of the enzyme. As result much of its characterization has been performed on crude or partially purified extracts (Lendzian 1980, Scheibe et al., 1989, Wright et al., 1997) Clearly overexpressed castor leucoplast G6PDH enzyme does not suffer from such instability reported for spinach chloroplast and barley root plastids. Overexpressed purified G6PDH enzyme from castor was routinely stored at –20°C and the enzyme easily survived multiple freeze-thaw cycles without loss of activity. In addition the purified protein was stored in 50% glycerol and did not require any additional stabilizing proteins (BSA) sometimes required for labile
proteins. From this we can conclude that the tertiary structure of castor G6PDH is quite stable.

When DTT is added to assays for leucoplast G6PDH, there is an increase in activity compared with that of other plastid forms of the enzyme in which the activity is decreased and this may have been the cause of the previous failure to detect this enzyme. Previous attempts were aimed at oxidizing the enzyme for activation and this report clearly shows that oxidation inactivates and reduction activates castor G6PDH. The three dimensional model suggests that the redox state of the disulfide bonds (whether reduced or oxidized should behave in a similar fashion to other plant G6PDH homologues. However, the presence or absence of disulfide bonds in castor enzyme seems to have little effect on its tertiary structure and thus the addition of DTT does not decrease it activity. That is to say the presence or absence of disulfide bonds in castor is not likely a method of controlling the enzyme activity.

In contrast, the addition of DTT appears to enhance activity of overexpressed G6PDH. This result may be an artifact of the bacterial overexpression system used to produce the castor enzyme but the data in Figures 10 and 11 suggest this is not the case. However, overexpressed proteins are, in many bacterially expressed systems, misfolded or may have improper disulfide bond formation when produced in prokayontic expression systems that not designed to fold eukaryotic proteins with disulfide linkages. Improperly folded enzymes may be adopting a folded conformation with active site distortions or pertubation and as a result show a reduced level of G6PDH activity. The addition of DTT may have released the improper disulfide bonding allowing the enzyme
to adopt a more favoured tertiary conformation, which in turn, resulted in higher activity levels when tested.

Castor G6PDH exhibits hyperbolic kinetics with respect to both of its substrates similar to the kinetics displayed by other higher plants.

Future work should continue this initial study of the OPPP and G6PDH in castor oil seeds. Studies should determine the pH optimum of the purified G6PDH. Also the enzyme should be tested for the effects of other effectors and inhibitors of activity; the NADPH/NADP$^+$ ratios could be tested on purified G6PDH as well. A Southern could be performed to determine the copy number to determine if a second isoform for G6PDH exists similar to the two plastidic isoforms in potato. A northern blot could also be performed to determine the developmental regulation and effects of light and thioredoxin on expression levels of G6PDH.

Additional work should include alternative attempts to purify and stabilize G6PDH from castor extracts and to also investigate why there is an instability with the enzyme in previous attempts. A careful look at protease degradation and gene copy number with respect to purification of this enzyme from castor extracts may provide additional information about this enzyme and system.
Chapter 5: Summary

Glucose 6-phosphate dehydrogenase is present in Castor (*Ricinus communis* L.) leucoplast. The full length clone of G6PDH with transit peptide was isolated from a Castor cDNA library using G6PDH polyclonal antibodies from potato. The coding sequence is 1803 base pairs in length encoding a protein that is 600 amino acids in length. This clone was inserted into an overexpression system and enzyme analysis of this cleaved recombinant protein showed a Km of 38uM for NADP$^+$ and 930uM for G-6-P. In addition, the purified recombinant protein exhibited a two fold increase in reducing NADP$^+$ to NADPH in the presence of the reducing agent DTT. This increase in activity in the presence of a reducing agent is opposite in regulation compared to other higher plant G6PDHs. In chloroplasts, reducing agents attenuate G6PDH activity. Compared to native leucoplast preparations, the recombinant preparations showed no detectable contamination from 6PGDH that contributed to the production of NADPH.
References


