

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

**Characterization of the *de novo* pyrimidine biosynthetic pathway in
*Arabidopsis thaliana***

by

Christopher W. Kafer

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Plant Physiology

Program of Study Committee:
Robert Thornburg, Major Professor
Martha James
Peter Keeling
Basil Nikolau
Martin Spalding

Iowa State University

Ames, Iowa

2002

UMI Number: 3073458

UMI[®]

UMI Microform 3073458

Copyright 2003 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

**Graduate College
Iowa State University**

**This is to certify that the doctoral dissertation of
Christopher W. Kafer
has met the dissertation requirements of Iowa State University**

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Dedication

For my parents, Frank and Charlene, who never lost faith in me.

TABLE OF CONTENTS

LIST OF NOMENCLATURE	v
ABSTRACT	vi
CHAPTER 1. GENERAL INTRODUCTION	1
Introduction	1
Dissertation organization	18
References	18
 CHAPTER 2. <i>ARABIDOPSIS THALIANA</i> CYTIDINE DEAMINASE 1 SHOWS MORE SIMILARITY TO PROKARYOTIC ENZYMES THAN TO EUKARYOTIC ENZYMES.	 38
Abstract	38
Introduction	39
Results	40
Discussion	47
Materials and Methods	50
Literature Cited	52
 CHAPTER 3. THYMINE-STARVATION UP-REGULATES A REGION OF THE <i>ARABIDOPSIS THALIANA</i> CHROMOSOME III CONTAINING BOTH THE UMP SYNTHASE GENE AND A PROXIMAL GENE ENCODING A NOVEL F-BOX PROTEIN	 66
Abstract	66
Introduction	66
Materials and Methods	68
Results	71
Discussion	77
Literature Cited	79
 CHAPTER 4. ANALYSIS OF THE <i>DE NOVO</i> PYRIMIDINE BIOSYNTHETIC PATHWAY IN <i>ARABIDOPSIS THALIANA</i>	 94
Abstract	94
Introduction	95
Results	96
Discussion	106
Materials and Methods	116
Literature Cited	125
 CHAPTER 5. GENERAL CONCLUSIONS	 152

LIST OF NOMENCLATURE

ADP: Adenosine 5' diphosphate

AMT: Aminopterin

ATCase: Aspartate Transcarbamoylase

ATP: Adenosine 5' triphosphate

DHOase: Dihydroorotase

DHODH: Dihydroorotate dehydrogenase

FOA: 5-fluoroorotate

IPTG: Isopropyl- β -D-thiogalactoside

kDa: Kilodalton

Km: Michaelis constant

MCO: Modified competitive oligonucleotide

MS media: Murashige and Skoog media

ORMV: Oil rapeseed mosaic virus

RT-PCR: Reverse transcription-polymerase chain reaction

T1 plants: The first generation of transgenic plants made by selfing

T2 plants: The second generation of transgenic plants made by selfing

TAE: Tris-acetate-EDTA

UMP: Uridine 5' monophosphate

UMPs: Uridine monophosphate synthase

ABSTRACT

Because pyrimidines are at the core of cellular metabolism, the study of pyrimidine metabolism is of intrinsic interest. We have characterized several aspects of pyrimidine metabolism in *Arabidopsis thaliana*. We have characterized the enzyme cytidine deaminase and determined that it is more similar to its prokaryotic homologues than its eukaryotic homologues. The full length cDNA was cloned into the expression vector pProEXHTb and contains an N-terminal 6 histidine tag. The expressed protein was purified to apparent homogeneity using a Ni-NTA agarose column. The K_m and V_{max} for cytidine were determined to be 226 μ M and 39.7 μ Moles/min respectively. The enzyme was also able to use 2'-deoxycytidine as substrate with an apparent K_m of 49 μ M and V_{max} of 24 μ Moles/min. The enzyme was unable to deaminate CMP, dCMP or cytosine. The high level of conservation between the *Arabidopsis* enzyme and the *E. coli* enzyme permitted us to model the active site of the *Arabidopsis* enzyme and to deduce that the mechanism is identical to the *E. coli* enzyme mechanism.

We have also developed a single tube semi-quantitative relative reverse transcription polymerase chain reaction (RT-PCR) assay that allows us to monitor expression of the *de novo* pyrimidine biosynthetic pathway in *Arabidopsis*. We have used this assay to quantitate the relative transcript levels of the UMP synthase gene and the SKIP5 F-box protein located 405 nucleotides upstream. We have determined that the 2 genes are coordinately transcriptionally regulated in organs of *Arabidopsis* and under conditions of pyrimidine starvation. The mechanism of this coordinate regulation is unknown as is the function of the SKIP5 protein.

We have used our quantitative PCR method to analyze transcript levels of the entire *de novo* pathway of pyrimidine biosynthesis relative to the 18S rRNA. We have synthesized an

oligonucleotide that is modified on its 3' end so it cannot be used by the polymerase in template extension. It does however compete for binding sites with the unmodified oligonucleotides and hence attenuates the level of amplification of the 18S transcript. We have used this method to quantitate the level of expression, relative to 18S rRNA, of each member of the pyrimidine biosynthetic pathway under various growth conditions. Under conditions of pyrimidine starvation, the entire pathway is transcriptionally upregulated. ATCase shows the highest level of upregulation at 20 fold while the rate limiting UMP synthase is upregulated by 5 fold. Under the abiotic stress conditions of phosphate starvation and saline the pathway is relatively unaffected. Likewise, the pathway is unaffected by conditions of flooding. However, under completely anaerobic conditions DHOase, DHODH and UMP synthase transcript levels are reduced to almost undetectable levels, while ATCase is unaffected. Under the biotic stress of ORMV infection DHODH and UMP synthase are both up regulated by two and four fold respectively.

We have generated lines of transgenic *Arabidopsis* plants containing the promoters of each member of the *de novo* biosynthetic pathway driving β -glucuronidase expression. Using these plants we have analyzed the expression patterns and levels of GUS expression. Two promoter GUS fusions were constructed for both ATCase and DHOase due to the presence of an intron within the 5' UTR of both these genes. The introns of both these genes were required for expression within roots of our transgenic plants. Further, the DHOase intron is required for the highest levels of expression within aerial portions of the plants. The DHODH and UMP synthase promoters were both relatively weak in comparison to the ATCase and DHOase genes. Regions within the UMP synthase coding region conferred greatly increased GUS expression.

CHAPTER 1. GENERAL INTRODUCTION

Introduction

Nucleotides

Within the cell there are two classes of nucleotides, purines and pyrimidines. Each class of molecule is synthesized by independent pathways. Nucleosides consist of a nitrogenous base linked from the N1 of the pyrimidine ring or the N9 of the purine ring to the C1 carbon of ribose in a β configuration. The two pathways of *de novo* nucleotide biosynthesis utilize the common precursor 5-phosphoribosyl-1-pyrophosphate (PRPP) as the activated ribose donor. In purine biosynthesis the purine ring is sequentially built on the ribose moiety. In contrast, the five carbon sugar is added to the pyrimidine ring after it has been synthesized. The pools of these molecules are tightly regulated and maintained in proportion to one another in plants (Dancer et al. 1990). The UTP/UDP ratio is kept similar to the ATP/ADP ratio through the action of nucleoside-5-diphosphate kinase.

Nucleotides can be synthesized *de novo* as well as from salvage of preformed bases or by recycling from nucleic acid turnover. (Figure 1) It is obviously beneficial from an energetic standpoint for an organism to recycle nucleotides from RNA degradation. Pyrimidine nucleotides can be converted back to UMP or simply reused without any conversion other than phosphorylation. Plants have all the enzymes necessary for this recycling with two notable exceptions, cytosine deaminase and dCMP deaminase.

The completion of the *Arabidopsis* genome sequencing project has allowed us to characterize the genes (Table 1) and pathways of pyrimidine metabolism as well as the genomic organization of the pathways. (Figure 2) Based on this genome analysis it appears that

pyrimidine biosynthesis is quite similar to other eukaryotic organisms with few exceptions.

Nucleotides are key components in a myriad of cellular reactions. They are the constituents of nucleic acids and components of the coenzymes FAD, NAD and CoA. UDP-Glucose, ADP-glucose, UDP-Galactose and CDP- diacylglycerol are the activated precursors in polysaccharide and phospholipid biosynthesis, respectively. The purine ATP is the energy currency of the cell while cyclic AMP is a regulator of many hormonal processes in animals. Because of their key role in almost every major biosynthetic pathway, their study is of intrinsic, and potentially practical, value.

De novo pyrimidine biosynthesis

Pyrimidine metabolism is evolutionarily conserved in all species examined. The *de novo* biosynthetic pathway consists of six enzymatic steps. (Figure 2) Carbamoyl phosphate synthase (CPSase) forms carbamoyl phosphate which is used in both pyrimidine and arginine biosynthesis. Aspartate is condensed with carbamoyl phosphate by the enzyme aspartate transcarbamoylase (ATCase) and is the first committed step in pyrimidine biosynthesis. The carbamoyl aspartate is cyclized by the enzyme dihydroorotase (DHOase) to yield the pyrimidine ring. Dihydroorotate is subsequently oxidized via the enzyme dihydroorotate dehydrogenase (DHODH) to yield orotate. The resulting orotate is used by the bi- functional enzyme UMP synthase (UMPs) along with phosphoribosyl pyrophosphate to form orotidine monophosphate (OMP). The OMP molecule is also decarboxylated by UMPs. The resulting uridine monophosphate (UMP) is the precursor to all other pyrimidine nucleotides within the plant cell.

There are some notable similarities and differences in the enzymes of the *de novo* pathway between plants and other organisms. In animals, the first three steps are carried out by

a large multifunctional polypeptide termed CAD. In plants these three steps are carried out by individually encoded proteins, as is found in prokaryotes. Like animals and other eukaryotes, UMP synthase is found as a single bifunctional polypeptide with discrete domains for each function, in plants. In prokaryotes these last two steps are found in individually encoded proteins.

Genes of the *de novo* pathway

CPSase

Two cDNAs have recently been cloned which encode the small and large subunit (carA and carB) of this enzyme (Brandenburg et al. 1998, Williamson et al. 1996). Like other enzymes of the *de novo* pathway, CPSase is more similar to a prokaryotic enzyme than the eukaryotic homologues. The small subunit shows 46% identity to the *E. coli* enzyme and 51% identity to the *Synechocystis* enzyme. The large subunit shares 66% identity to the *Synechocystis* enzyme. The C-terminus of the *Arabidopsis* large subunit contains a unique domain not found in other CPSases that may be important in subunit interaction (Williamson et al. 1996).

The *Arabidopsis* large subunit gene is found on chromosome I. It is 3800 bp long and is interrupted by 2 introns of 103 and 144 bp in length. The small subunit gene is found on chromosome III. This gene is 2600 bp long and contains nine introns of 89 to 175 bp.

Carbamoylphosphate Synthase (CPSase) is encoded within the nucleus, but the mature protein is targeted to and functions within the chloroplast. This enzyme is responsible for the production of carbamoylphosphate for both pyrimidine and arginine biosynthesis. Ornithine stimulates incorporation of NaHCO_3 into both UMP and arginine. However, in the presence of exogenously added uridine, incorporation of NaHCO_3 into UMP was reduced while

incorporation into arginine was unaffected (Lin et al. 1986).

ATCase

Several cDNAs encoding ATCase (pyrB1, pyrB2, pyrB3) have been cloned from *Pisum sativum* (Williamson et al. 1994), (Williamson et al. 1996) along with a single cDNA from *Arabidopsis* (Nasr et al. 1994). The *Arabidopsis* gene is 2.5 kb, located on chromosome III and is single copy in the *Arabidopsis* genome. The gene contains 5 introns ranging from 84 to 445 bp. An intron is found in the 5' UTR of this gene.

A motif within the ATCase protein has been identified as a putative pyrimidine binding site, based upon its homology to the known allosteric pyrimidine binding site of the *E. coli* regulatory ATCase subunit. This supports the argument that ATCase is the major regulated enzyme of pyrimidine biosynthesis. This binding site is also conserved in the *Arabidopsis* protein. The pea cDNAs contain N-terminal chloroplast targeting sequences which is in agreement with previous work reporting localization of ATCase activity to the chloroplast.

As in other organisms, this step is likely the major regulatory site. Indeed, studies in squash (*Cucurbita pepo*) tissue cultured cells have shown a 90% decrease in ATCase activity upon the addition of UMP to the growth media (Lin et al. 1986). Further, addition of exogenous uridine to the growth media was able to inhibit incorporation of radiolabeled NaHCO_3 into UMP but had no effect on the incorporation into arginine. This also lends support to the claim that CPSase is not the major regulatory step in pyrimidine biosynthesis (Lin et al. 1986; Lovatt et al. 1979).

DHOase

Like most other members of the pathway, DHOase is a nuclear encoded chloroplastic

enzyme. The cDNA encoding this enzyme has been cloned from *Arabidopsis* and shares 54.5% identity to the *E. coli* enzyme (Zhou et al. 1997) . This gene is also found as a single copy within the *Arabidopsis* genome on chromosome IV. The gene contains eight introns ranging from 68 to 310 bp. Like the ATCase gene, DHOase contains an intron in the 5' UTR.

Expression of DHOase has been examined during wheat germination. It was found that DHOase drops 50% during the first 30 min of imbibition and remains very low until approx. 12 hours after imbibition at which point, activity again rises. Two days after germination the activity is four fold higher than in the dry seeds (Mazus et al. 1972) and therefore, this enzyme must be developmentally regulated.

DHODH

To date the only dihydroorotate dehydrogenase homologue cloned from a higher plant was from *Arabidopsis* by complementation of a yeast auxotroph with plant cDNAs (Minet et al. 1992) . In this case, the enzyme appears to be similar to its eukaryotic homologues. This protein also has N-terminal targeting sequences which direct the protein to the mitochondrial intermembrane space.

The *Arabidopsis* DHODH gene is also single copy and is found on chromosome V. The 2298 base-pair gene is interrupted by 9 introns ranging in size from 87 to 193 base pairs. There are several interesting features of the promoter region of this gene. It is relatively short (290 bp) and contains no TATA box. Furthermore the promoter is, most likely bi-functional, as an expressed superoxide dismutase gene is found immediately upstream and in opposite orientation to the DHODH gene.

UMPs

In plants, UMP synthase is a bifunctional protein catalyzing the final 2 steps of *de novo* pyrimidine biosynthesis. As in most higher eukaryotes, the plant UMP synthase has both orotate phosphoribosyl transferase (OPRTase) and orotidylate decarboxylase (ODCase) activities.

The *Arabidopsis* UMPS gene has an interesting physical structure. The 5'-half of the gene encodes the OPRTase function while the 3'-half encodes the ODCase function. There are 5 introns found in the *Arabidopsis* gene ranging from 81 to 177 nucleotides. Interestingly, all five of these introns are found in the ODCase half and no introns are found in the OPRTase half. The *Nicotiana plumbaginifolia* gene contain introns in identical positions with the *Arabidopsis* gene; however there is an additional intron within the 5' half of the coding region of the tobacco gene that is not found in the *Arabidopsis* gene. Like the other genes in the *de novo* pathway UMP synthase is a single copy gene and located on chromosome III.

Intracellular organization of the *de novo* pathway

Almost the entire pathway of pyrimidine biosynthesis occurs within the chloroplast. (Figure 3) There is one notable exception, DHODH, which is localized in the mitochondria. The biosynthetic product of the first three *de novo* pathway enzymes, dihydroorotate, accumulates in the chloroplast. To produce UMP, dihydroorotate must be transported from the chloroplast and be taken up by the mitochondria where it is converted into orotic acid. Subsequently, orotic acid must be transported from the mitochondria and re-enter the chloroplast where it is converted into UMP by the enzyme UMP synthase.

Dihydroorotate is oxidized by DHODH within the mitochondria with concomitant

reduction of ubiquinone. Therefore, nucleotide biosynthesis is inextricably linked to respiration. Plants frequently undergo oxygen stress induced principally by flooding. In many parts of the globe, flooding is an annual or a semiannual occurrence. Pyrimidine biosynthesis would undoubtedly be affected by these events resulting in unknown consequences for the plant.

In maize, the shuttling of nucleotides between the chloroplast and the mitochondria appears to be mediated by a protein with similarity to microbial pyrimidine/ purine transport proteins. This protein is encoded by the *lpe1* gene (Schultes et al. 1996). Indeed, normal plant development appears to be dependent upon such pyrimidine transport. Mutations in the *lpe1* gene in maize result in abnormal chloroplasts and altered leaf development. Several homologues have been identified in the *Arabidopsis* EST databases (Accession numbers: AY093137, NP_199810, AAD14479, AAM47573) . Because the control of pyrimidine biosynthesis in plants involves this unusual subcellular compartment shuttling, these proteins may be an important and previously overlooked site of control for nucleotide metabolism in plants.

Pyrimidine Salvage and recycling

The yeast *FUR4* gene encodes the uracil permease. This protein has been shown to be down regulated by the addition of uracil. The mechanism of this down regulation is a decrease in *fur4* mRNA level and increased degradation of the permease protein. (Seron et al. 1999)

Pyrimidine nucleosides are principally taken up into plant cells via one of two routes. Wu et al., have identified 5-fluoro-2'-deoxyuridine FUDR resistant mutants of *Arabidopsis* that are deficient in uptake of FUDR from tissue culture media (Wu et al. 1994). This protein also transports thymidine and may be similar to thymidine transporters from yeast and other

eukaryotes. Uracil transport was unaffected in these mutants. In addition, work in our lab has demonstrated that the pyrimidine free bases orotic acid, uracil and cytosine and each of their fluorinated analogs are all taken up into plant cells (Santoso 1991). The proteins responsible for this transport have not been identified, but are apparently different from the transporter of Wu and King (1994).

Cytosine deaminase

It has been previously demonstrated that plants, like animals and unlike microorganisms, lack a cytosine deaminase activity (Stougaard 1993). This lack of cytosine deaminase points to the inability of plants to salvage the cytosine free base. Indeed, because of this inability to salvage the free base, fluorinated cytosine cannot be metabolized in plants and animals and is therefore an effective antimicrobial compound. This has been very effectively utilized in animal systems, but has been widely overlooked in plant systems, especially plant tissue cultures. Recently, a negative selection scheme utilizing a bacterial cytosine deaminase has been proposed (Serino et al. 1997). In this scheme the *codA* gene was introduced into the plastid genome of tobacco. A functional *codA* gene causes cell death in the presence of the normally non-toxic 5-fluorocytosine by conversion to the toxic 5-fluorouracil.

Cytidine deaminase

We have identified and characterized the *Arabidopsis* cytidine deaminase 1 (CDA1) (Kafer et al. 2000). *Arabidopsis* contains a gene family of seven genes encoding cytidine deaminase. Six of these genes CDA2 through CDA7 are present as a tandem repeat on a 15.7 kb stretch of chromosome IV (Faivre-Nitschke et al. 1998). Interestingly, all of the plant cytidine deaminases are more similar to the *E. coli* enzyme than they are to other eukaryotic

enzymes. The bacterial enzymes are α_2 dimers of 31 kDa subunits. The eukaryotic enzymes from humans, yeast, *Caenorhabditis*, and insects are tetramers of 15 kDa subunits. The plant enzyme is translated as a 32.5 kDa monomer that shows conserved identity with the *E. coli* enzyme throughout its length.

UPRTase

Plants appear to have several different uracil phosphoribosyl transferase (UPRTase) genes which are responsible for salvage of free uracil. We have recently identified one gene from *Arabidopsis thaliana* which is most similar to the *Toxoplasma gondii* and yeast enzymes rather than the bacterial enzymes (Weers et al. 1999). However, the *Nicotiana tabacum* enzyme has greater identity with the bacterial enzymes than to eukaryotic enzymes. At least five UPRTase genes or putative genes are found in the *Arabidopsis* genome. None have been well characterized. In addition, *Arabidopsis* appears to have an unusual bifunctional UPRTase/Uridine kinase. This gene has a single base deletion that causes a frame shift mutation. For this transcript to encode a functional protein with both enzymatic activities the ribosome must shift to maintain the correct reading frame. It is unknown if this particular gene product is relevant in pyrimidine biosynthesis since both activities are also carried out by other individually encoded genes. However, the transcript of this potentially novel bifunctional gene is found in the *Arabidopsis* EST database, so it is at least transcriptionally expressed.

Pyrimidine Modification

Based upon current knowledge, it appears that formation of TMP and CMP is essentially identical to animal cell systems. Because fluorouracil induces thymine starvation in plant cells (Santoso et al. 1998) as it does in animal cells, the biochemistry must be similar.

UMP/CMP kinase

An *Arabidopsis* cDNA has been cloned and shows 50% sequence homology to the mammalian enzyme (Zhou et al. 1998). A conserved nucleotide binding region (GGPGS/AGK) is located near the N-terminus. This motif is found in all eukaryotic monophosphokinases and anchors the gamma phosphate of the nucleotide (Pai et al. 1977; Zhou et al. 1998). Like its eukaryotic counterparts, UMP and CMP are equally acceptable as substrates for the plant protein.

NDP kinase

Nucleoside diphosphokinase catalyzes the conversion of the diphosphonucleosides to the triphosphonucleosides. Most of the work done on these kinases has centered around their involvement in signal transduction pathways. These important kinases have been shown to be autophosphorylating (Moisyadi et al. 1994; Tanaka et al. 1998), involved in phytochrome mediated light perception (Tanaka et al. 1998; Hamada et al. 1996), heat shock response (Moisyadi et al. 1994) and wound response (Harris et al. 1994). Most of the work has been centered around purine nucleoside diphosphokinase esp. GDP to GTP synthesis. A family of three NDP kinases have been identified in *Arabidopsis thaliana*.

CTP synthase, NDP reductase, dUTPase

CTP synthases have not been studied in plants. Our searches have identified at least three genes encoding CTP synthases from the *Arabidopsis* genome sequencing project. Likewise, the NDP reductase subunits have not been studied. In a search for meristem specific mRNAs, Pri-Hadash and coworkers isolated a small cDNA from tomato that encoded a dUTPase

(Pri-Hadash et al. 1992). This mRNA was preferentially expressed in the growing meristems, but its expression dramatically declined in tissues further from the meristems. This enzyme is similarly expressed in root meristems (Pardo et al. 1990) but is almost undetectable in mature root tissues.

Thymidylate synthase/ Dihydrofolate reductase

Thymidylate synthase (TS) and Dihydrofolate reductase (DHFR) are responsible for the production of dTMP and regeneration of tetrahydrofolate, respectively. Like UMP synthase, the enzyme in *Arabidopsis* and also carrots is a bifunctional polypeptide (Lazar et al. 1993; Luo et al. 1993) which apparently has arisen from a gene fusion. The N-terminus has identity with DHFR while the C-terminus has identity with TS.

The bifunctional nature of this enzyme is unusual. In bacteria and in most higher eukaryotes, including yeast, TS and DHFR proteins are separate enzymes encoded by distinct genes. Other bifunctional TS/DHFR enzymes occur in protozoans such as *Leishmania amazonensis* (Nelson et al. 1990), *Plasmodium falciparum* (Bzik et al. 1987), and *Paramecium tetraurelia* (Schlichtherle et al. 1996). Sequence analysis by Lazar and coworkers provides conflicting arguments on the two competing hypotheses for the origin of the plant and protozoan TS/DHFR enzyme (Lazar et al. 1993). These enzymes have identical fusion sites between the functional domains. Based upon this, it is possible that the protozoan and plant enzymes have a common ancestor. Alternatively, these authors indicate that pairwise amino acid sequence analysis shows that these bifunctional enzymes are each more similar to the monofunctional enzymes than they are to each other. This raises the possibility that these enzymes may have arisen by convergent evolution. There are two copies of the gene in *Arabidopsis* that appear to

have arisen by gene duplication (Lazar et al. 1993).

Pyrimidine catabolism

This pathway in plants has not been well characterized to our knowledge. However, genes encoding dihydrouracil dehydrogenase (uracil reductase), dihydropyrimidine dehydrogenase, and β -ureidopropionase have all been identified in *Arabidopsis*. We conclude that pyrimidine degradation is in all likelihood identical to that in other organisms. In rat liver the rate limiting step in this pathway is β -ureidopropionase (Wasternack 1978). It is unknown if this holds true for plants.

Degradation of the pyrimidine ring proceeds via the “reductive” pathway after nucleoside phosphorolysis by uridine phosphorylase, which yields the free base and ribose-1-phosphate. The reductive enzymes utilize thymine or uracil and cytidine must therefore be converted to uridine by cytidine deaminase before degradation. The end product of pyrimidine ring degradation is β -alanine.

The intracellular location of these enzymes in plants is not known for certain. In mammalian liver and *Euglena gracilis*, β -ureidopropionase was localized to the cytosol (Wasternack 1978).

Pyrimidine metabolism regulation

Control of *de novo* biosynthesis in *Saccharomyces cerevisiae*

Under conditions of pyrimidine starvation in *S. cerevisiae* the genes of the *de novo* biosynthetic pathway are transcriptionally upregulated three to eight fold. (Roy et al. 1990) The regulatory DNA binding protein Ppr1p is in part responsible for regulation of this pathway by up-regulating transcription of DHODH, ODCase, DHOase and OPRTase. (Flynn et al. 1999)

Under pyrimidine starvation conditions the dihydroorotate levels rise as ATCase expression is up-regulated under low UTP levels (Potier et al. 1990). When DHO levels rise to a critical concentration the DNA bound Ppr1p is converted to a transcriptionally active state by direct binding of DHO. (Flynn et al. 1999). Thus, a specific and sensitive mechanism which senses pyrimidine levels controls *de novo* pyrimidine biosynthesis in yeast is mediated by Ppr1p, the first eukaryotic regulatory gene to be characterized. There is no homologue of Ppr1p found in Arabidopsis so plants must regulate pyrimidine biosynthesis in a different manner.

Drosophila rudimentary gene expression

Because pyrimidines are co-factors in many other pathways, they may play an as yet undiscovered role in regulation of these pathways in plants. Examples of this regulation in other organisms are scattered throughout the literature. In *Drosophila*, the CAD complex (CPSase-ATCase-DHOase) is called *rudimentary* (*r*). Mutations in this gene cause developmental abnormalities and the eggs must be supplied with exogenous pyrimidine for survival. The wings of the surviving individuals are abnormal and show varying degrees of truncation and may also be wrinkled or blistered. (Fausto-Sterling et al. 1976) Homozygous females are sterile which can be partially alleviated by administering cytidine during development. These studies show that pyrimidine biosynthesis is required for normal ovarian and wing development. It is currently unknown what, if any, developmental requirements for pyrimidines are found in plants.

Mutants in *rudimentary* show truncated wings and undeveloped ovaries. Therefore, temporal and spatial expression of this gene is necessary for normal fly development. (Tsubota et al.) Two mutants have been identified that show this phenotype. One maps to a gene now

named *enhancer of rudimentary e(r)^{pl}*. This gene has been cloned and codes for a conserved protein also found in plants and animals (Gelsthorpe et al. 1997). In *Drosophila* this mutation enhances wing truncation in the *r* mutant phenotype. It is currently unknown how this interaction occurs but does not appear to be regulating *r* gene expression at the transcriptional level.

The *Arabidopsis* enhancer of rudimentary gene is located on chromosome five. It encodes a 109 amino acid protein. A protein database search reveals that homologues are also found in soybean and rice (Altschul et al. 1990). Threonine 24 of the *Arabidopsis* protein is phosphorylated by casein kinase II *in vitro* but lacks the second phosphorylation site found on the *er* proteins of other organisms (Gelsthorpe et al. 1997). The *Arabidopsis* amino acid sequence is 41% identical to the *Drosophila* sequence (Gelsthorpe et al. 1997). No experimental data on the function of this gene in any plant has been reported.

Plant pyrimidine metabolism control

Plants undergo developmental processes that theoretically require high levels of pyrimidine nucleotides. This includes processes such as germination, pollen tube growth, flowering, and seed set. In spite of these numerous processes that require large amounts of nucleotides, nucleotide metabolism has not been particularly well studied in plants.

Germination

Expression of the enzymes of *de novo* pyrimidine and pyrimidine salvage have been examined during germinating seedlings. Following imbibition, dry seeds typically have high concentrations of free bases that appear within a few hours (Grzelczak et al. 1975). These appear to arise by salvage from nucleosides stored within the seeds. Based upon incorporation

of labeled precursors into RNA, salvage appears to be the preferred source during the first few hours following germination. After the first day, *de novo* synthesis becomes the predominant source of nucleotides (Ashihara 1977).

Similarly, the *de novo* and salvage enzymes have been examined in dry seeds and most enzymes are present; however, following a brief decline in activity, the activity of the *de novo* enzymes dramatically increases within only a few hours following imbibition (Mazus et al. 1972). In contrast, the salvage enzymes UPRTase and Uridine kinase were high in dry seeds and showed little change during the first 24 hours of germination. The fact that relative high levels of salvage enzymes are maintained following germination along with the induction of the *de novo* pathway results in maximum funneling of all nucleotides towards pyrimidine accumulation.

Tissue Cultured Cells

In some tissue cultures the relative activities of the salvage and *de novo* pathways change markedly depending upon the stage of the culture. The key salvage pathway enzymes UPRTase and Uridine kinase, increase just after transfer of the cultures. Enzymatic reactions of the *de novo* pathway, namely CPSase and UMP synthase, increased after the initial post transfer lag phase with a maximum velocity during the active cell division phase (Kanamori Fukuda et al. 1981). The results are quite similar to the picture seen in germinating seeds where the initial phase is largely one of pyrimidine salvage which then quickly turns biosynthetic with salvage playing less of a role. The question of how the two pathways (*de novo* biosynthesis and salvage) are coordinately regulated still remains unanswered.

That the activities of CPSase and UMP synthase increased in *Vinca rosea* (Madagascar

periwinkle) cultures during the cell division phase led Kanamori- Fukuda and coworkers to postulate that the *de novo* pathway may not be feedback regulation but may rather be controlled by enzyme synthesis (Kanamori Fukuda et al. 1981). Likewise, these workers showed evidence that the salvage pathways were also controlled by the availability of the key salvage enzymes.

Studies with metabolic inhibitors

The pathways of arginine biosynthesis and pyrimidine biosynthesis share the carbamoylphosphate precursor pool and are tightly coordinately controlled. Metabolic inhibitors of each of these pathways should be useful in elucidating the regulatory mechanisms of each. The metabolite, 5-fluorouracil, is toxic to plant cells just as it is in animal cells (Sung et al. 1980), (Santoso et al. 1992).

Phaseolotoxin is a pathogenesis factor from *Pseudomonas syringae* pv phaseolicola, a pathogen of beans. It inhibits the condensation of carbamoylphosphate with ornithine and is capable of reversing the toxic effects of fluorouracil (Jacques et al. 1981). The result of treating cells with phaseolotoxin is the increase in intracellular ornithine which subsequently stimulates CPSase activity. The resulting carbamoylphosphate enters the *de novo* pyrimidine biosynthetic pathway thereby increasing the levels of UMP which overcomes the toxicity of fluorouracil.

In order to understand the regulation and expression of pyrimidine biosynthesis in plants, we have previously examined the effect of the metabolic inhibitor, 5-fluoroorotic acid, on UMP synthase expression in cell cultures of *Nicotiana plumbaginifolia* (Santoso et al. 1998). UMP synthase is the rate limiting step of pyrimidine biosynthesis in plants (Santoso et al. 1992). Addition of fluoroorotic acid causes an up-regulation of UMP synthase enzyme activity, resulting from transcriptional induction of the UMP synthase gene. Exogeneously added

thymine reversed this up-regulation. Methotrexate and aminopterin, which affect thymine levels by inhibiting DHFR also up-regulate UMP synthase in *N. plumbaginifolia* cells (Santoso et al. 1998).

Our lab has previously generated mutant *N. plumbaginifolia* cell lines which are capable of growing in the presence of 5-FOA. The most interesting of these mutant cell lines had constitutive high level expression of ATCase and UMP synthase. These cell lines have apparently lost the ability to regulate gene expression of members of the *de novo* pathway because addition of pyrimidines to the growth media does not repress expression of these two genes.

Pyrimidine biosynthesis and other pathways

Pyrimidines are co-factors phospholipid biosynthesis. In *Sacharomyces cerevisiae* it has been shown that phosphatidylcholine synthesis proceeds via the CDP-diacylglycerol and CDP-choline pathways. Overexpression of the CTP synthetase in *S. cerevisiae* led to a two fold increase in the utilization of the CDP-choline pathway and a reduction in phosphatidylserine synthetase activity (McDonough et al. 1995). Neither of these results were due to a transcriptional upregulation of the biosynthetic genes, however. CTP was shown to be the limiting reagent in the phosphocholine cytidyltransferase reaction and CTP had an allosteric inhibitory effect on the phosphatidylserine synthetase reaction. Similar experiments in plants have not been reported.

Recently, Loef et al. have shown that potato tuber disks incubated with orotate show an increase in starch biosynthesis (Loef et al. 1999). Orotate and uridine stimulated an increase in UTP, UDP and UDP-glucose levels. Feeding with orotate led to an increase in glucose-1-

phosphate levels by stimulating sucrose degradation two fold. These authors argue that their results show that uridine nucleotide levels are specifically stimulating sucrose degradation via sucrose synthase (SuSy) and concomitant production of hexose phosphate via the action of uridinedisphosphoglucose pyrophosphorylase (UGPase). The net rate of starch biosynthesis increased 30% due to the increase in hexose phosphates which are precursors of starch biosynthesis (Loef et al. 1999).

Dissertation organization

This dissertation is presented as three journal papers preceded by a general introduction and followed by a general conclusion.

References:

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J. Mol. Biol. **215**: 403-410.
- Ashihara, H. (1977). "Changes in activities of the *de novo* and salvage pathways of pyrimidine nucleotide biosynthesis during germination of black gram [mungo beans] (*Phaseolus mungo*) seeds." Z Pflanzenphysiol **81**: 199-211.
- Brandenburg, S. A., C. L. Williamson and R. D. Slocum (1998). "Characterization of a cDNA encoding the small subunit of Arabidopsis carbamoyl phosphate synthetase (Accession No. U73175)." Plant Physiol. **117**: 717.
- Bzik, D., W. Li, T. Horii and J. Inselburg (1987). "Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene." Proc Natl Acad Sci USA **84**: 8360-8364.
- Dancer, J., H. E. Neuhaus and M. Stitt (1990). "Subcellular compartmentation of uridine

- nucleotides and nucleoside-5'-diphosphate kinase in leaves." Plant Physiol. **92**: 637-641.
- Faivre-Nitschke, E. S., J. M. Grienberger and J. M. Gualberto (1998). "Cloning and characterisation of a cytidine deaminase gene family from *Arabidopsis thaliana*." GenBank AJ005811.
- Fausto-Sterling, A. and L. Hsieh (1976). "Studies on the female-sterile mutant rudimentary of *Drosophila melanogaster*. 1. An analysis of the rudimentary wing phenotype." Dev Biol **51**(2): 269-81.
- Flynn, P. and R. Reece (1999). "Activation of transcription by metabolic intermediates of the pyrimidine biosynthetic pathway." Mol Cell Biol **19**: 882-888.
- Gelsthorpe, M., M. Pulumati, C. McCallum, K. Dang Vu and S. Tsubota (1997). "The putative cell cycle gene, enhancer of rudimentary, encodes a highly conserved protein found in plants and animals." Gene **186**: 189-195.
- Gelsthorpe, M., M. Pulumati, C. McCallum, K. Dang-Vu and S. I. Tsubota (1997). "The putative cell cycle gene, enhancer of rudimentary, encodes a highly conserved protein found in plants and animals." Gene **186**(2): 189-95.
- Grzelczak, Z. and J. Buchowicz (1975). "Purine and pyrimidine bases and nucleosides of germinating *Triticum aestivum* [wheat] seeds." Phytochemistry **14**: 329-331.
- Hamada, T., N. Tanaka, T. Noguchi, N. Kimura and K. Hasunuma (1996). "Phytochrome regulates phosphorylation of a protein with characteristics of a nucleoside diphosphate kinase in the crude membrane fraction from stem sections of etiolated pea seedlings." J Photochem Photobiol B **33**: 143-151.

- Harris, N., J. Taylor and J. Roberts (1994). "Isolation of a mRNA encoding a nucleoside diphosphate kinase from tomato that is up-regulated by wounding." Plant Mol Biol **25**: 739-742.
- Jacques, S. L. and Z. R. Sung (1981). "Regulation of pyrimidine and arginine biosynthesis investigated by the use of phaseolotoxin and 5- fluorouracil." Plant Physiology **67**: 287-291.
- Kafer, C. and R. Thornburg (2000). "*Arabidopsis thaliana* cytidine deaminase I shows more similarity to prokaryotic enzymes than to eukaryotic enzymes." Journal of Plant Biology **43**: 162-170.
- Kanamori Fukuda, I., H. Ashihara and A. Komamine (1981). "Pyrimidine nucleotide biosynthesis in *Vinca rosea* cells: changes in the activity of the *de novo* and salvage pathways during growth in a suspension culture." J Exp Bot **32**: 69-78.
- Lazar, G., H. Zhang and H. M. Goodman (1993). "The origin of the bifunctional dihydrofolate reductase-thymidylate synthase isogenes of *Arabidopsis thaliana*." Plant J, **3**: 657-668.
- Lin, F.-F. and C. J. Lovatt (1986). "Biosynthesis of pyrimidine nucleotides in cultured root callus of *Cucurbita pepo*." In Vitro Cell. Dev. Biol. **22**: 1-5.
- Loef, I. I., M. Stitt and P. Geigenberger (1999). "Orotate leads to a specific increase in uridine nucleotide levels and a stimulation of sucrose degradation and starch synthesis in discs from growing potato tubers." Planta **209**(3): 314-23.
- Lovatt, C. J., L. S. Albert and G. C. Tremblay (1979). "Regulation of pyrimidine biosynthesis in intact cells of *Cucurbita pepo*." Plant Physiol. **64**: 562-569.
- Luo, M., P. Piffanelli, L. Rastelli and R. Cella (1993). "Molecular cloning and analysis of a

- cDNA coding for the bifunctional dihydrofolate reductase-thymidylate synthase of *Daucus carota*." Plant Mol Biol **22**: 427-435.
- Mazus, B. and J. Buchowicz (1972). "Activity of enzymes involved in pyrimidine metabolism in the germinating wheat grains." Phytochem **11**: 77-82.
- McDonough, V. M., R. J. Buxeda, M. E. Bruno, O. Ozier-Kalogeropoulos, M. T. Adeline, C. R. McMaster, R. M. Bell and G. M. Carman (1995). "Regulation of phospholipid biosynthesis in *Saccharomyces cerevisiae* by CTP." J Biol Chem **270**(32): 18774-80.
- Minet, M., M.-E. Dufour and F. Lacroute (1992). "Complementation of *Saccharomyces cerevisiae* auxotrophic mutants by *Arabidopsis thaliana* cDNAs." Plant J. **2**(3): 417-422.
- Moisyadi, S., S. Dharmasiri, H. Harrington and T. Lukas (1994). "Characterization of a low molecular mass autophosphorylating protein in cultured sugarcane cells and its identification as a nucleoside diphosphate kinase." Plant Physiol **104**: 1401-1409.
- Nasr, F., N. Berthauche, M.-E. Dufour, M. Minet and F. Lacroute (1994). "Heterospecific cloning of *Arabidopsis thaliana* cDNAs by direct complementation of pyrimidine auxotrophic mutants of *Saccharomyces cerevisiae*. I. Cloning and sequence analysis of two cDNAs catalysing the second, fifth and sixth steps of the *de novo* pyrimidine biosynthesis pathway." Mol. Gen. Genet. **244**: 23-32.
- Nelson, K., G. Alonso, P. Langer and S. Beverley (1990). "Sequence of the dihydrofolate reductase-thymidylate synthase (DHFR-TS) gene of *Leishmania amazonensis*." Nucleic Acids Res **18**: 2819.
- Pai, E. F., W. Sachsenheimer, R. H. Schirmer and G. E. Schulz (1977). "Substrate positions and

- induced-fit in crystalline adenylate kinase." J. Mol. Biol. **114**: 37-45.
- Pardo, E. and C. Gutierrez (1990). "Cell cycle- and differentiation stage-dependent variation of dUTPase activity in higher plant cells." Exp Cell Res **186**: 90-98.
- Potier, S., F. Lacroute, J. C. Hubert and J. L. Souciet (1990). "Studies on transcription of the yeast URA2 gene." FEMS Microbiol Lett **60**(1-2): 215-9.
- Pri-Hadash, A., D. Hareven and E. Lifschitz (1992). "A meristem-related gene from tomato encodes a dUTPase: analysis of expression in vegetative and floral meristems." Plant Cell **4**: 149-159.
- Roy, A., F. Exinger and R. Losson (1990). "cis- and trans-acting regulatory elements of the yeast URA3 promoter." Mol Cell Biol **10**(10): 5257-70.
- Santoso, D. (1991). Production and characterization of UMP Synthase mutants from haploid cell suspensions of *Nicotiana tabacum*, Iowa State University.
- Santoso, D. and R. W. Thornburg (1992). "Isolation and Characterization of UMP synthase mutants from haploid cell suspensions of *Nicotiana tabacum*." Plant Physiol. **99**: 1216-1225.
- Santoso, D. and R. W. Thornburg (1998). "UMP synthase is transcriptionally regulated by pyrimidine levels in *Nicotiana plumbaginifolia*." Plant Physiol **116**: 815-821.
- Schlichtherle, I., D. Roos and J. Van Houten (1996). "Cloning and molecular analysis of the bifunctional dihydrofolate reductase-thymidylate synthase gene in the ciliated protozoan *Paramecium tetraurelia*." Mol Gen Genet **250**: 665-673.
- Schultes, N., T. Brutnell, A. Allen, S. Dellaporta, T. Nelson and J. Chen (1996). "Leaf permease1 gene of maize is required for chloroplast development." Plant Cell **8**: 463-75.

- Serino, G. and P. Maliga (1997). "A negative selection scheme based on the expression of cytosine deaminase in plastids." Plant J **12**(3): 697-701.
- Seron, K., M. O. Blondel, R. Haguenaue-Tsapis and C. Volland (1999). "Uracil-induced down-regulation of the yeast uracil permease." J Bacteriol **181**(6): 1793-800.
- Stougaard, J. (1993). "Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene." Plant J, **3**: 755-761.
- Sung, Z. R. and S. Jacques (1980). "5-fluorouracil resistance in carrot cell cultures: Its use in studying the interaction of the pyrimidine and arginine pathways." Planta **148**: 389-396.
- Tanaka, N., T. Ogura, T. Noguchi, H. Hirano, N. Yabe and K. Hasunuma (1998). "Phytochrome-mediated light signals are transduced to nucleoside diphosphate kinase in *Pisum sativum* L. cv. Alaska." J Photochem Photobiol B **45**: 113-121.
- Wasternack, C. (1978). "Degradation of Pyrimidines - Enzymes, Localization and Role in Metabolism." Biochem. Physiol. Pflanzen **173**: 467-499.
- Weers, B. and R. W. Thornburg (1999). "Characterization of the cDNA and gene encoding the *Arabidopsis thaliana* Uracil Phosphoribosyltransferase (AF116860)." Plant Physiol. **119**: 1567.
- Williamson, C. L., M. R. Lake and R. D. Slocum (1996). "A cDNA encoding carbamoyl phosphate synthetase large subunit (*carB*) from *Arabidopsis* (Accession No. U40341)." Plant Physiol. **111**: 1354.
- Williamson, C. L. and R. D. Slocum (1994). "Molecular cloning and characterization of the *pyrB1* and *pyrB2* genes encoding aspartate transcarbamoylase in pea (*Pisum sativum* L.)." Plant Physiol. **105**: 377-384.

- Williamson, C. L., L. To and R. D. Slocum (1996). "Characterization of a cDNA (Accession No. U05293) encoding a third aspartate transcarbamoylase (*pyrB3*) from pea." Plant Physiol. **112**: 446.
- Wu, K. and J. King (1994). "Biochemical and genetic characterization of 5-fluoro-2'-deoxyuridine-resistant mutants of *Arabidopsis thaliana*." Planta **194**: 117-122.
- Zhou, L., F. Lacroute and R. Thornburg (1998). "Cloning, expression in *Escherichia coli* and characterization of *Arabidopsis thaliana* uridine 5'-monophosphate/cytidine 5'-monophosphate kinase." Plant Physiol **117**: 245-253.
- Zhou, L., F. Lacroute and R. W. Thornburg (1997). "DNA sequence of dihydroorotase from *Arabidopsis thaliana* (AF000146)." Plant Physiol. **114**: 1569.
- Zhou, L. and R. W. Thornburg (1998). "Site specific mutagenesis of conserved residues in the phosphate binding loop of the *Arabidopsis* UMP/CMP kinase alter ATP and UMP binding." Arch. Biochem. Biophys. **358**: 297-302.

Figure 1. Interconnection of the pathways of pyrimidine metabolism.

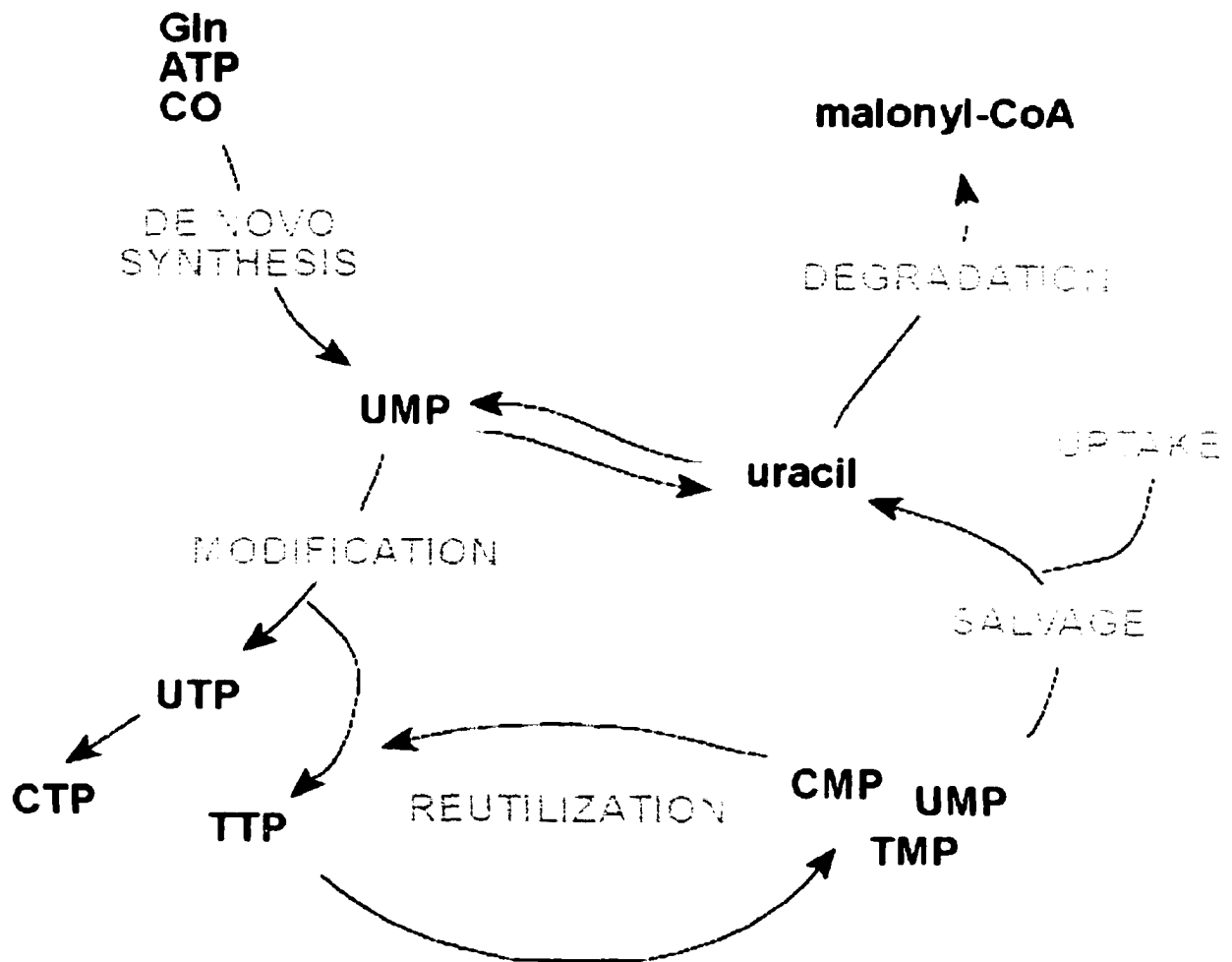


Table 1. Genes of pyrimidine metabolism.

Table 1. Genes of Pyrimidine Metabolism in *Arabidopsis thaliana*

Step	Enzyme	Gene Symbol	cDNA		Gene	Chrom	Genomic Location
			Genbank Accession Number	GenBank Accession Number			
Biosynthesis							
1a	Carbamoyl phosphate synthase (small subunit)	carA	U73175	AB018114	MGF10	III	48 cM
1b	Carbamoyl phosphate synthase (large subunit)	carB	U40341	AC008030	F1N18	I	49 cM
2	Aspartate transcarbamoylase	pyr2	X71843	AB024036	MQC12	III	28 cM
3	Dihydroorotase	pyr3	AF000146	AL161558 b	F7H19	IV	69 cM
4	Dihydroorotate dehydrogenase	pyr4	X62909	AB007648	MKD15	V	40 cM
5 & 6	UMP synthase	pyr56	X71842	AF276887 b	φAtUMPS	III	91.5 cM
Modification							
7	UMP/CMP kinase UMP/CMP kinase-like	pyr7	AF000147	NI ^a AL138658	-- T209	--	--
8	Nucleotide diphosphokinase	ndpla ndp2	AF058391 AF017640	c NI ^a	c --	V --	120 cM --

Table 1 (continued)

Step	Enzyme	Gene Symbol	Gene				Genomic Location
			CDNA Genbank Accession Number	GenBank Accessi on Number	genomic clone	Chrom	Position
9	CTP synthase	ctps1		AF001308	T10M13	IV	12 cM
		ctps2		AL022224	F1C12	IV	64 cM
		ctps3		AC004238	F19I3	II	67 cM
		ctps4		AC004135	T17H7	I	49 cM
				AL161552		II	
				AL161493			
10a	Ribonucleotide reductase (lsu)	rnrl	AF092841	AC007019 AB023036	F7D8	II	39 cM
10b	Ribonucleotide reductase (ssu)	rrc					
11	dUTPase	dut1	NI ^a	AL096859 b	T6H20	III	61 cM
12	Thymidylate synthase	ts	L08593	AC007119	F2G1	II	39 cM
13	Thymidylate kinase	tk		AL161585 AL035540 ?			
	dCMP kinase	dck		AC009176			

Table 1. (continued)

		Gene Symbol	Gene			Genomic Location	
Step	Enzyme		cDNA Accession Number	GenBank Accession Number	genomic clone	Chrom	Position
Salvage							
14	Uracil phosphoribosyltransferase A	uprt	AF116860	AL161565			
		urk1		AB024238	K1G2	III	43 cM
15	uridine kinase	urk2		AC016662	F2P9		
16	5'-nucleotidase	5nt					
Table 1. (continued)							
		cyd1	AF134487	AJ005687			
17	Cytidine deaminase			+	F27F23	II	35 cM
		cyd2-9		AC003058			
		cyd10-11		AF080676			
				AL079344	T16L4	IV	76 cM
		cyd???		+			
				AL161575			
				AC008075			
18	Cytosine deaminase	cod	NPd	NPd	--	--	--
Degradation							
19	Dihydropyrimidine dehydrogenase						
20	dihydropyrimidinase						

Table 1. (continued)

Step	Enzyme	Gene Symbol	Gene			
			cDNA Accession Number	GenBank Accession Number	genomic clone	Chrom Position
21	β -ureidopropionase	β UP		AB008268	MSJ1	
22	β -alanine synthase	bas				

^aNI = Not identified

^bThese genes have been sequenced and deposited in GenBank two or more times. The duplicate accessions are: DHODH, AL161558 (8) and AL031018 (24); UMPS, AF276887 ([Kafer, in preparation #2853]) and AL138656 (25); dUTPase, AL096859 (18) and AL133292 (19).

^c These genes span two BAC clones. The GenBank Accessions are: NDPK1a, AB023035 (K9H21) & AB008265 (MDC12);

^dNP = This enzyme is not present in plants

Figure 2. Chromosomal organization of pyrimidine metabolism

Arabidopsis chromosomes

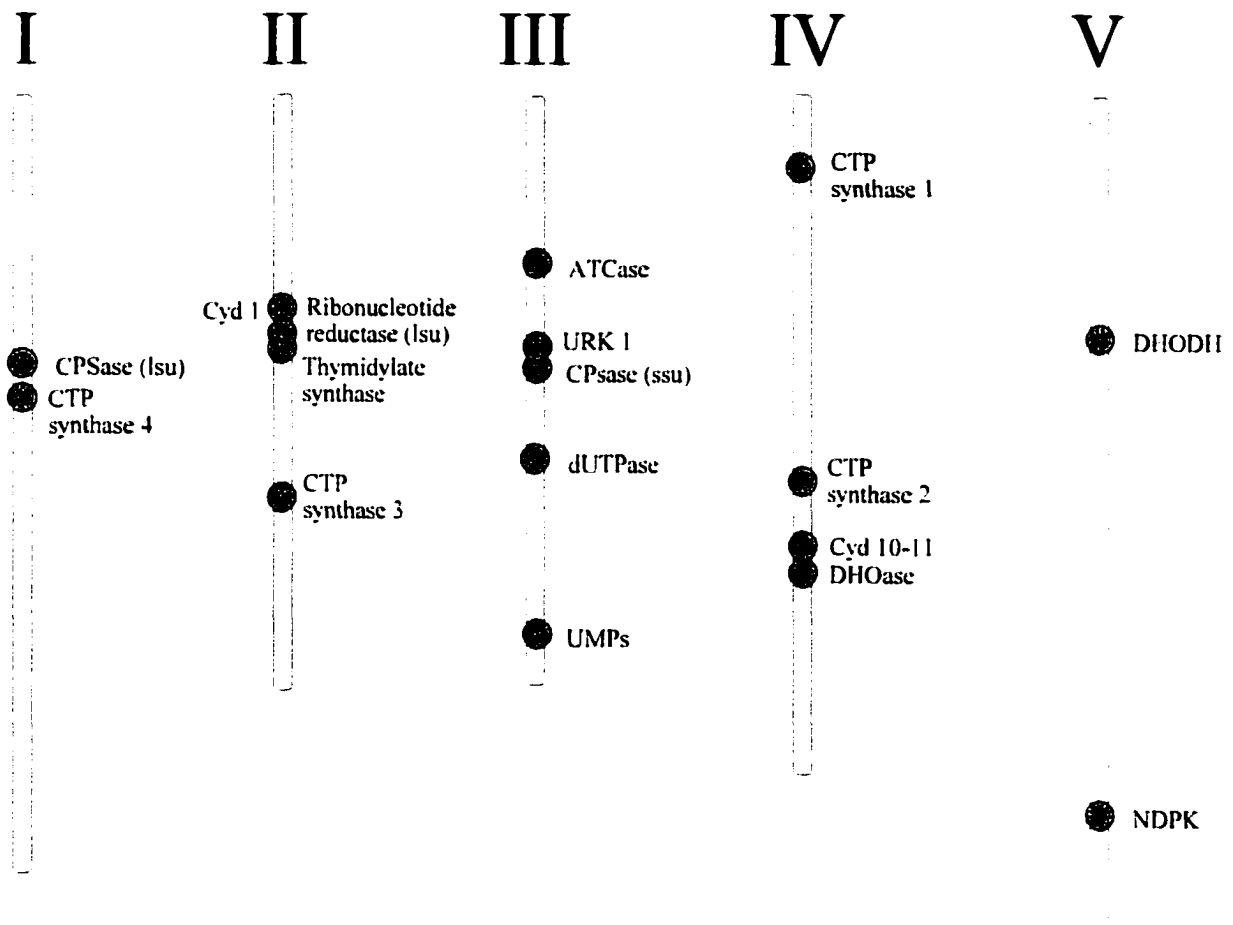


Figure 3. The *de novo* pathway of pyrimidine biosynthesis

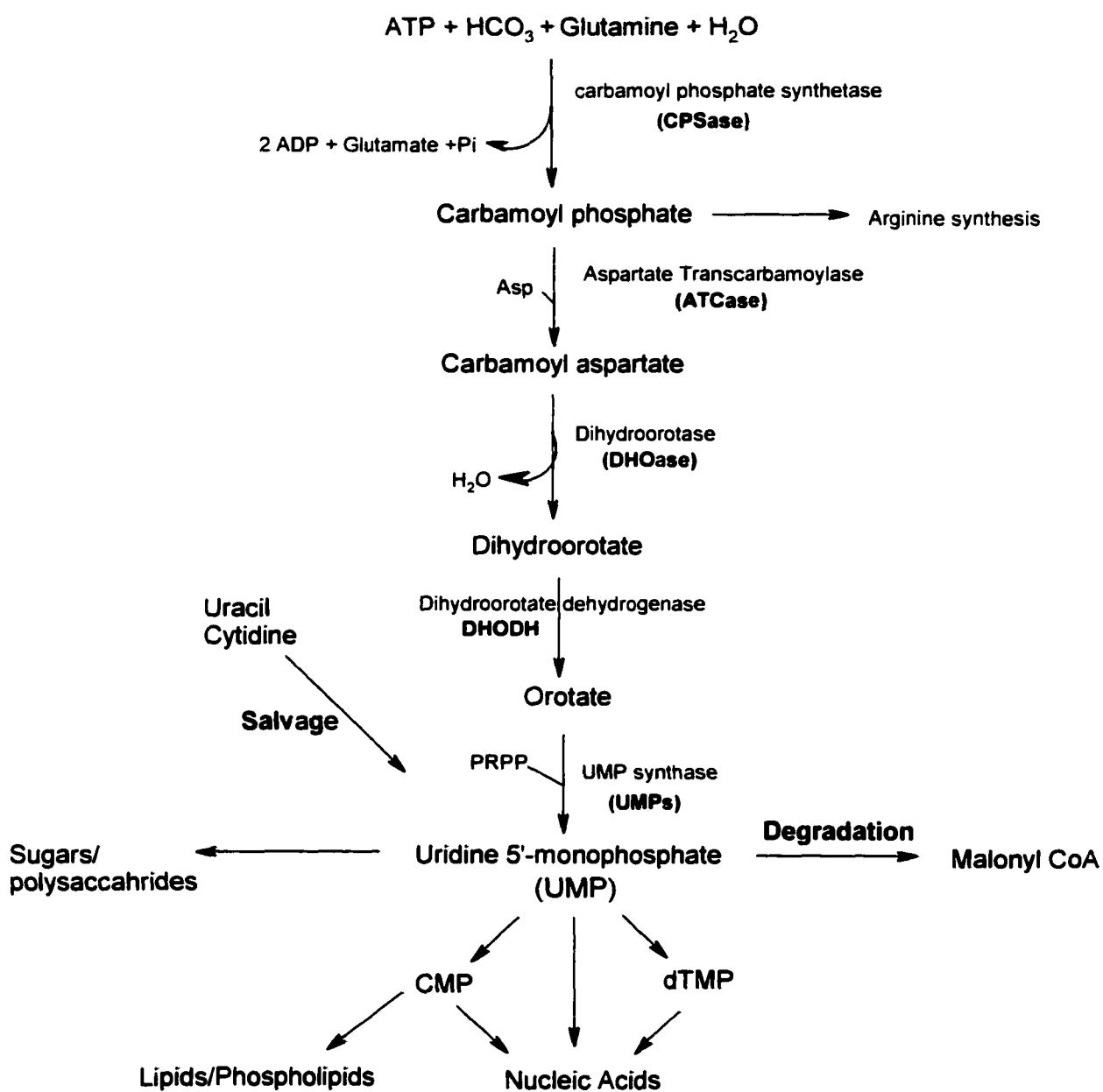
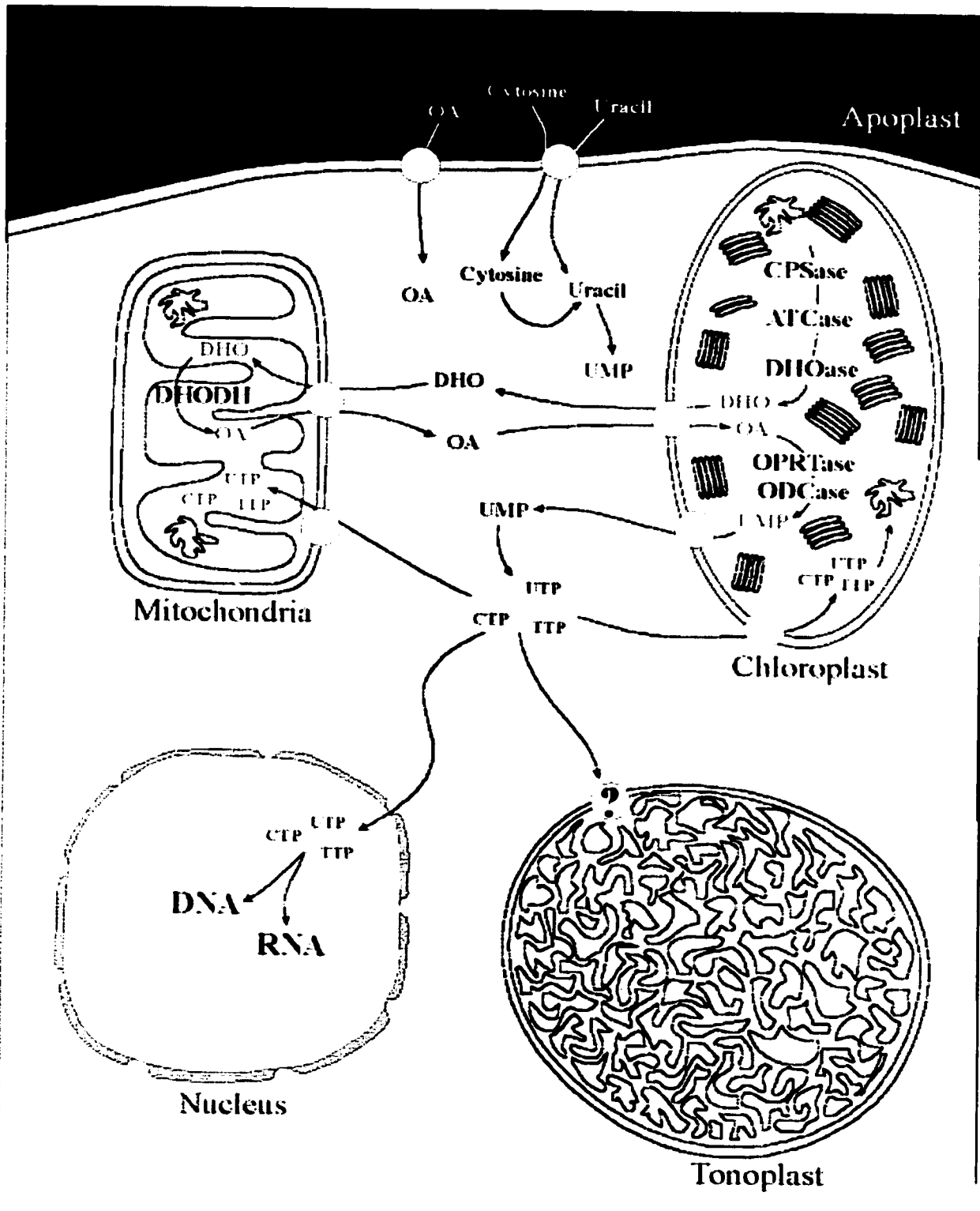


Figure 4. Intracellular localization of the pyrimidine biosynthetic pathway.



CHAPTER 2. *ARABIDOPSIS THALIANA* CYTIDINE DEAMINASE 1 SHOWS MORE SIMILARITY TO PROKARYOTIC ENZYMES THAN TO EUKARYOTIC ENZYMES

A paper published in the Journal of Plant Biology 43:162-170 (2000)

Chris Kafer & Robert Thornburg

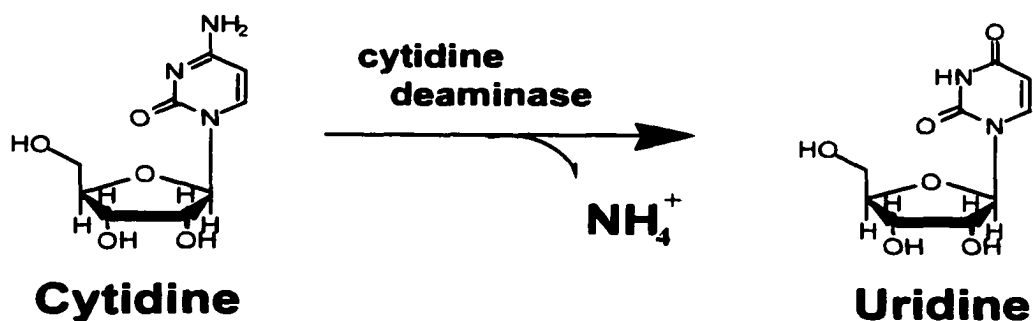
Abstract

Two Expressed Sequence Tagged (EST) clones were identified from the Arabidopsis database as encoding putative cytidine deaminases. Sequence analysis determined that the two clones overlapped and encoded a single cDNA. This cytidine deaminase corresponds to the *Arabidopsis thaliana* gene, *cdal*. The deduced amino acid sequence was more closely related to prokaryotic cytidine deaminases than to eukaryotic enzymes. The cDNA shares 44% amino acid identity with the *Escherichia coli* cytidine deaminase but only 26 and 27% identity with human and yeast enzymes. A unique zinc-binding domain of the *E. coli* enzyme forms the active site. A similar putative zinc-binding domain was identified in the Arabidopsis enzyme based upon primary sequence similarities. These similarities permitted us to model the active site of the Arabidopsis enzyme upon that of the *E. coli* enzyme. In this model, the active site zinc is coordinated by His⁷³, Cys¹⁰³, Cys¹⁰⁷, and an active site hydroxyl. Additional residues that participate in catalysis, Asn⁶⁴, Glu⁶⁶, Ala⁷⁸, Glu⁷⁹, and Pro¹⁰², are conserved between the Arabidopsis and *E. coli* enzymes suggesting that the Arabidopsis enzyme has a catalytic mechanism similar to the *E. coli* enzyme. The two overlapping ESTs were used to prepare a single, full-length clone corresponding to the *A. thaliana cdal* cDNA. This cDNA was subcloned into pProExHtb and expressed as a fusion protein with an N-terminal His₆ tag. Following purification on a Ni-NTA-Agarose column, the protein was analyzed for its kinetic

properties. The enzyme utilizes both cytidine ($K_m = 226 \text{ M}$) and 2'-deoxycytidine ($K_m = 49 \text{ M}$) as substrates. The enzyme was unable to deaminate cytosine, CMP or dCMP.

Introduction

Pyrimidines are at the core of cellular metabolism. They are components of DNA and RNA, they activate monosaccharides, they modify regulatory proteins, and they are involved in basic aspects of cellular physiology. Deficiencies in pyrimidine metabolism completely block cellular growth and development in both bacteria (Yamanaka et al., 1992) and yeast (Jong and Campbell, 1984, Prakash et al., 1979).



Cytidine deaminase (EC 3.5.4.5) catalyzes the first step in the salvage of pyrimidine nucleotides by hydrolytic deamination of cytidine into uridine. Cytidine deaminase has been identified and characterized from a wide variety of sources from bacteriophage (Maley, 1978) to mammals (Rothman et al., 1978). Our interest in cytidine deaminase stems from its role in pyrimidine salvage; however, cytidine deaminase also generates significant interest due to a variety of other observations. This makes a comparative analysis of the plant enzyme important in understanding the role of this enzyme in cell biology. First, in humans, cytidine deaminase

is responsible for the inactivation of useful chemotherapeutic agents such as cytidine arabinoside that are widely used in the treatment of leukemia (Meyers et al., 1973, Steuart and Burke, 1971). Second, cytidine deaminase catalyzes RNA editing. The best characterized example of RNA editing is the C to U editing of nucleotide 6666 of apolipoprotein B (apo-B) RNA in mammalian intestine catalyzed by a novel cytidine deaminase (Navaratnam et al., 1993, Teng et al., 1993). This modification generates a new stop translation codon (UAA) from a glutamate codon (CAA) producing a truncated protein (Powell et al., 1987).

RNA editing also occurs in both the mitochondria and chloroplasts of plants from bryophytes (Malek et al., 1996) to gymnosperms (Wakasugi et al., 1996) and angiosperms including both monocots (Freyer et al., 1997) and dicots (Lippok et al., 1996). RNA editing has been shown to completely modify the coding region of genes by creation of new stop codons (Zanlungo et al., 1995), new start codons (Kadowaki et al., 1995), and even creating entirely new open reading frames (Wakasugi, et al., 1996).

Results

Characterization of the Cytidine Deaminase mRNA

The full sequence of the *Arabidopsis thaliana* cytidine deaminase mRNA was constructed from two independent EST clones. Initially, we determined the entire nucleotide sequence of each clone. The comparison of these two clones, 149H23T7 and 150P17T7, revealed an identical overlapping region of 672 nucleotides (see Figure 1, Panel A). Clone 150P17T7 is incomplete at the 3' end. It contains the 5' untranslated region, start codon and the majority of the coding region extending to nucleotide 1003 (Ser²⁸⁸). Clone 149H23T7 is incomplete at the 5' end. It begins in the coding region at nucleotide 331 (Val⁶⁵) and extends

through the stop codon and a short 3' untranslated region to the poly A tail of the cytidine deaminase. A full-length clone was prepared from the two incomplete cytidine deaminase clones as outlined in the Materials and Methods section (Figure 1, Panel B).

mRNA Features

The mRNA encoding the putative cytidine deaminase deduced from the overlapping clones is 1161 nucleotides long. Based upon DNA matrix analysis, there were no significant internal duplications within the mRNA. A 141-nt 5' untranslated region was identified. There was no nucleotide bias of either AT (52%) or GC (48%) within the 5' untranslated region. The first ATG was found at nucleotides 142 to 144 of the mRNA. The identification of a purine nucleotide at nucleotide 139 and a G at nucleotide 145 (-3 and +4 relative to the ATG start codon) indicate that this mRNA is likely to be efficiently translated (Kozak, 1986).

The mRNA contains a single, 301 codon-long open reading frame that encodes a protein with high identity to the *E. coli* cytidine deaminase (see below). An amber stop codon (TAG) is located at nucleotides 1045 to 1047 of the mRNA and a 113 bp 3' untranslated sequence extends beyond. The polyadenylation site was found at nucleotide 1136. A consensus polyadenylation signal (AATAAA) was observed 46 nucleotides upstream from the polyadenylation site at nucleotides 1085 to 1090 of the mRNA. Further analysis of this region of the mRNA revealed that an unusual hairpin structure overlapped this polyadenylation signal (Figure 2). The ΔG of formation of this hairpin structure was -38.64 kJ/mol at 20°C, indicating that this structure is likely to form spontaneously within the mRNA. Thus, this polyadenylation signal is likely to be located within a hairpin structure. Uniquely, this hairpin structure results in the formation of a new consensus polyadenylation site that spans the base of the hairpin. This

bipartite polyadenylation signal is located only 14 nucleotides from the polyadenylation site (see Figure 2).

Protein Features

A single open reading frame was identified within the putative cytidine deaminase mRNA that encodes a 301 amino acid protein. The full-length protein has a calculated molecular weight of 32,582 and a pI of 5.52. To identify this protein, BLAST searches (Altschul et al., 1990) were performed using the translated amino acid sequence. This analysis identified the most closely related sequences in the databases as cytidine deaminases. Interestingly, the most closely related sequence was the *E. coli* cytidine deaminase rather than other eukaryotic homologues (Figure 3). Recently a family of seven cytidine deaminase cDNAs from Arabidopsis was deposited in the GenBank. All of these proteins show greater similarity to the *E. coli* enzyme than to other eukaryotic enzymes. The cytidine deaminase cDNA identified in this report corresponds to the cytidine deaminase *cdal* gene.

To further verify the relationship between the Arabidopsis and other enzymes, we performed a protein matrix analysis comparing the Arabidopsis cDNA with sequences from *E. coli* (GenBank Accession Number M60916), *Homo sapiens* (L27943), and *Saccharomyces cerevisiae* (U20865). This analysis is shown in Figure 4. While identity is clearly present between the Arabidopsis enzyme and all other cytidine deaminases, the most extensive identity is with the *E. coli* enzyme, and this identity extends throughout the full length of the enzyme. Further, the length of these enzymes differs significantly. Other eukaryotic cytidine deaminases are approximately half the size of the *E. coli* and Arabidopsis enzymes.

There are seven cysteine residues in the Arabidopsis cytidine deaminase, so there must be at least one unpaired cysteine residue in the cytidine deaminase protein as is found in other cytidine deaminases (Ipata and Cercignani, 1978). Both acidic and basic residues are scattered throughout the protein; however, the distribution of the negatively charged residues is a bit unusual. Both the N-terminal third (amino acids 1 to 120) and the C-terminal third of the protein (amino acids 225 to 301) are rich in glutamate (7/9 and 8/10 residues respectively). In contrast the middle third (amino acids 121 to 224) contains 16 negatively charged amino acids of which 15 are aspartates.

This protein does not contain readily discernible N-terminal targeting sequences that would direct the cytidine deaminase towards the secretory pathway, nucleus, or mitochondria. However, this protein does contain some features of a chloroplast protein including a high percentage of serine and alanine (27%) in the N-terminal 30 amino acids (Nakai and Kanehisa, 1992).

It has previously been suggested that the mRNA editing cytidine deaminase in mammalian cells forms a heterodimer with the 60 kDa RNA binding component by virtue of a leucine zipper (Navaratnam, et al., 1993). Numerous leucine residues are present throughout the middle portion of the Arabidopsis protein (amino acids 143 to 179). The spacing; however, of these residues in the Arabidopsis cytidine deaminase does not form a perfect heptad repeat, suggesting that this portion of the protein may not form a leucine zipper. Further, structural predictions (Rost et al., 1994) on this portion of the cytidine deaminase indicate that this region is not likely to form an α -helix.

Model of the Active Site

Other cytidine deaminase enzymes contain a zinc-binding, active site that is required for catalysis (Driscoll and Zhang, 1994). This active site has been best studied in the *E. coli* enzyme (Betts et al., 1994, Carlow et al., 1996 Carlow et al., 1998 . Xiang et al., 1995, Xiang et al., 1996, Xiang et al., 1997). In the *E. coli* enzyme a zinc-binding pocket exists between two alpha helices. In the *E. coli* enzyme the zinc is coordinated by His¹⁰², Cys¹²⁹, Cys¹³², and an active site water molecule (Betts, et al., 1994 Carlow, et al., 1996, Carlow, et al., 1998, Xiang, et al., 1995, Xiang, et al., 1996, Xiang, et al., 1997). An alignment of the Arabidopsis sequence with that of the *E. coli* enzyme and other cytidine deaminases is presented in Figure 5, Panel A. The zinc-binding amino acids are indicated with arrows. These zinc-binding residues are conserved in the Arabidopsis enzyme as residues His⁷⁷, Cys¹⁰³, and Cys¹⁰⁷.

Structural predictions on the Arabidopsis cytidine deaminase were performed using the PredictProtein server at EMBL-Heidelberg. The region from amino acids 77 to 110 was predicted to contain two α -helices separated by a β -sheet. This same pattern is observed in the 3D structure of the *E. coli* cytidine deaminase (Betts, et al., 1994). Therefore, a 24 amino acid sequence corresponding to the zinc-binding domain was modeled using the *E. coli* cytidine deaminase crystal structure (PDB accession # 1AF2) using a "first approach" method with the Swiss Model program at <http://expasy.hcuge.ch/swissmod/SWISS-MODEL.html>. The results of this are shown in Figure 5 panel B. In this model, the sidechains of selected amino acids are shown. As indicated above, His⁷⁷, Cys¹⁰³, and Cys¹⁰⁷ are involved in the coordination of the active site zinc atom. In addition, a hydrogen bond between the carboxylate group of the conserved Glu⁷⁹ and zinc-bound hydroxyl group apparently is critical in the stabilization of the

transition state intermediate. It also serves as a proton shuttle from the zinc-bound hydroxyl group to the leaving amino group (Carlow, et al., 1996, Xiang, et al., 1997). As can be seen by comparing the model with the *E. coli* enzyme structure, each of these functional sidechains in the Arabidopsis enzyme is in the appropriate location to mediate the hydroxylation/deamination reaction. In addition, several other amino acids that are known to participate in substrate binding in the *E. coli* enzyme such as the Ala¹⁰³ backbone amide that forms a hydrogen bond with the 2-keto group of cytidine, Asp⁸⁹ and Glu⁹¹ that form hydrogen bonds with the ribosyl 3' hydroxyl, and Pro¹²⁸ that orients the amino leaving group are all conserved in the Arabidopsis enzyme as Ala⁷⁸, Asn⁶⁴, Glu⁶⁶, and Pro¹⁰². Thus, based upon the high degree of conservation of structure and amino acids that are involved in the active site chemistry, we postulate that the mechanism of cytidine deamination in the Arabidopsis enzyme is similar to that of the *E. coli* enzyme. Figure 5, Panel C shows the conservation of interacting residues between the *E. coli* and the Arabidopsis enzyme.

Expression, Purification, and Analysis of Recombinant Cytidine Deaminase

To conclusively demonstrate that this cDNA encodes a cytidine deaminase, we expressed a recombinant form of the protein in *E. coli*, purified it, and evaluated its enzymatic activity. A full-length cytidine deaminase clone was constructed as described in Materials and Methods. This construct contained a 32 amino acid extension at the N-terminus of the recombinant protein that contains a His₆ affinity tag to facilitate purification. A secondary structural analysis of the recombinant protein with this N-terminal extension was performed. This extension was not expected to significantly alter the protein structure or its localization in *E. coli*. This clone was placed under the control of the powerful *Trc* promoter. This plasmid.

pRT439, was then transformed into *E. coli* BL21 cells, and the recombinant cytidine deaminase was expressed. After an initial growth to log phase, the cells were induced with 1 mM IPTG for 4 h. The resulting cells were concentrated by centrifugation, resuspended, and broken by sonication. After another centrifugation to pellet cell debris, the crude lysate was applied to a Ni-NTA column and allowed to flow through under gravity. The column was washed and subsequently eluted with 100 mM EDTA. The purity of the protein preparation was evaluated by SDS PAGE as shown in Figure 6. This figure shows the specific induction of a 36 kDa protein following IPTG treatment (compare lanes 1 and 2). The recombinant protein expressed from pRT439 is predicted to have a molecular mass of 36.4 kDa. Lane 3 shows the protein fraction bound to the Ni-NTA affinity column and specifically eluted with imidazole. This purified enzyme fraction was utilized for all subsequent protein methods.

Initially we evaluated the deamination of the cytosine-free base, cytidine nucleosides, and cytidine nucleotides. This analysis, shown in Table 1, demonstrates that the recombinant enzyme shows no activity against either the free base, cytosine, or the cytidine nucleotides, CMP or dCMP. The lack of activity in deaminating cytosine has been previously reported for the *E. coli* enzyme where the removal of the ribose reduces catalytic activity by 10^8 (Carlow, et al., 1998). These authors indicated that the electron-withdrawing nature of the substituent ribose activates the cytosine ring for nucleophilic attack. The lack of activity for CMP and dCMP also can be explained by the enzyme structure. An important hydrogen bond is formed between the ribose 5'-hydroxyl and the backbone carbonyl oxygen atom Ala631 of the other monomer (Betts, et al., 1994). The presence of a phosphate group on the ribose 5'-hydroxyl would therefore not permit this substrate to fit the binding pocket.

In contrast, the *Arabidopsis* enzyme showed high activity in the deamination of cytidine and an almost twofold higher activity in the deamination of 2'-deoxycytidine (see Table 1). This also was observed with the *E. coli* enzyme (Ashley and Bartlett, 1984) and can be explained from the crystal structure. The 2'-hydroxyl does not make hydrogen bonds within the active site. In order for cytidine to bind, the 2'-hydroxyl must be desolvated, and this desolvation of the 2'-hydroxyl destabilizes the cytidine protein complex. Consequently, the deoxycytidine is more strongly bound (Betts, et al., 1994). To assess this with the *Arabidopsis* enzyme, we also determined kinetic constants for the recombinant enzyme with both of the active substrates. As shown in Table 1, the K_m values for cytidine and deoxycytidine indicate that deoxycytidine binds approximately fivefold more tightly to the *Arabidopsis* enzyme than does cytidine (49 M vs 226 M). Again, this is similar to the *E. coli* enzyme.

Based upon these analyses, we conclude that this cDNA does indeed encode a cytidine deaminase with similar structure and enzymatic activity to the *E. coli* cytidine deaminase. Further, this cDNA appears to be a member of a closely related gene family with at least seven members that are all more similar to the prokaryotic cytidine deaminase than to the eukaryotic enzymes.

Discussion

We have identified two overlapping EST clones as *Arabidopsis* cytidine deaminase clones. From these two clones we have reconstructed the full-length cDNA and analyzed the mRNA and protein. We have expressed the cDNA in *E. coli* and verified that the expressed protein has cytidine deaminase activity. The mRNA encoding the cytidine deaminase has an

unusual feature in its 3' end that may function as a novel polyadenylation site. Such structural features are not common, and to our knowledge this unusual hairpin appears to be unique.

Structurally, the Arabidopsis protein is more similar to the *E. coli* than it is to either the yeast or the human cytidine deaminase. The *E. coli* enzyme is known to be a dimer of 31 kDa subunits (Betts, et al., 1994), while the human enzyme is a tetramer of 13 kDa subunits (Cacciamani et al., 1991). In our case, the Arabidopsis protein is translated with a molecular weight of 32.5 kDa and has identity with the *E. coli* enzyme throughout its sequence.

The substrate specificity also is similar to that found in the *E. coli* enzyme. Both enzymes are unable to use cytosine as a substrate. It has been previously demonstrated that plants, like animals and unlike microorganisms, lack a cytosine deaminase activity (Stougaard, 1993). This lack of cytosine deaminase points to the inability of plants to salvage the cytosine-free base. Indeed, because of this inability to salvage the free base, fluorinated cytosine cannot be metabolized in higher eukaryotes and is an effective antimicrobial compound. This has been very effectively utilized in animal systems, but has been widely overlooked in plant systems, especially plant tissue cultures.

The inability of this enzyme to deaminate CMP or dCMP is also noteworthy and leads to several unanswered metabolic questions. The Arabidopsis UMP/CMP kinase is very effective in converting CMP to CDP (Zhou et al., 1998). Thus, it appears that cytidine nucleotides are recycled to CDP and subsequently CTP rather than to uridine nucleotides. This, however, does not appear to be the case for dCMP. The Arabidopsis UMP/CMP kinase is 30-fold less effective in the conversion of the dCMP into dCDP than in the conversion of CMP into CDP (Zhou, et al., 1998). Whether dCMP is metabolized into dUMP or dCDP remains

unresolved. The cytidine deaminase, however, is twice as active in the conversion of deoxycytidine into deoxyuridine than in the conversion of cytidine into uridine. It also is unclear whether deoxyuridine is metabolized to uracil and deoxyribose, to dUMP, or both, as is the case in *E. coli* (Neuhard and Nygaard, 1987). Conversion to dUMP via phosphorylation with thymidine kinase would provide an alternative route to TMP biosynthesis. Deribosylation to uracil would recycle the ring system to either UMP by way of uracil phosphoribosyltransferase (Weers and Thornburg, 1999) or to degradation by dihydrouracil dehydrogenase. In bacteria, deoxyribose-5-phosphate is metabolized to glyceraldehyde-3-phosphate and acetaldehyde (Ackermann et al., 1974). Whether deoxyribose metabolism occurs in plants has not, to our knowledge, been evaluated. Resolution of these metabolic questions will require further investigation and isolation of additional salvage enzymes.

That all of the *Arabidopsis* cytidine deaminases are much more similar to the *E. coli* enzyme than to the eukaryotic enzymes from either humans or *Saccharomyces* is surprising. This similarity is both structural and functional. The structural similarities have permitted us to model the active site of the *Arabidopsis* enzyme based on the active site of the *E. coli* enzyme. Based upon this analysis, all features of the *E. coli* enzyme that have been identified to participate in the bacterial enzymatic mechanism are conserved in the *Arabidopsis* enzyme. The functional similarities between the bacterial and plant enzyme also are quite remarkable and lead us to the conclusion that the bacterial form of the cytidine deaminase has been conserved in plants.

Materials and Methods

Bacterial Strains

Escherichia coli strains XL1 Blue or DH5 were used for routine subcloning and sequencing DNA preparations. *E. coli* BL21 was used for expression of the recombinant protein due to the lack of *omp* and *lon* proteases.

DNA sequencing

DNA sequence reactions were performed at the Iowa State University Nucleic Acid Facility using the Applied Biosystems Prism Dye-deoxy Cycle Sequencing Kit. The reactions were run on an Applied Biosystems Prism 377 DNA sequencer (Perkin-Elmer Corp., Norwalk, CT). Sequencing was initiated from known vector sequences. On the basis of these runs, primers specific to the cytidine deaminase sequence were constructed. Both strands were sequenced in duplicate or triplicate. RNA structural calculations were conducted with the program RNAdraw (Matzura and Wennborg, 1996).

Construction of Expression Cassette

Two Expressed Sequence Tagged (EST) clones (150P17T7 and 149H23T7) were identified as potential cytidine deaminase cDNAs by Blast searches (Altschul, et al., 1990) and obtained from the Arabidopsis Biological Resource Center stock center (ABRC, 1995). The complete sequence of each clone showed that they were partial cDNAs that together encoded a single open reading frame of a putative cytidine deaminase.

A single, full-length cytidine deaminase clone was constructed from these two cDNAs. The 150P17T7 EST was digested with *Xba*I and *Bgl*II and ligated into the *Xba*I/*Bgl*II sites of pRT146 (a pUC derivative in which a *Bgl*II linker had been inserted into the *Hinc*II site). This

resulting plasmid was named pRT428. The EST clone, 149H23T7, was digested with *Bgl*II and *Sph*I and ligated into the *Bgl*II/*Sph*I sites of pRT428. The resulting clone, pRT429, contained the full-length cytidine deaminase cDNA.

The expression plasmid, pRT439, was constructed by polymerase chain reaction (PCR) amplification of the coding region with the pair of oligonucleotides CYDP2 and CYDP3: CYDP3 (5'-GCGCGGATCCATCGAAGGTCGT **AT**GGATAAGCCAAGCTTCGTA-3') contained the start codon (bold), a Factor Xa protease site (italicized), and a *Bam*HI site (underlined); CYDP2 (5'-CGGGATCCCTCGAGCTAAGCTTCATAGCAATGAA ACAC-3') was complementary to the cDNA, included the stop codon (bold), and contained *Bam*HI and *Xho*I sites (underlined). The resulting PCR product was digested with *Bam*HI and *Xho*I and ligated into the *Bam*HI/*Xho*I sites of pProEXHtb (Gibco, Gaithersburg, MD).

Expression of Cytidine Deaminase and Purification of Recombinant Protein

Individual *E. coli* BL21 transformants carrying pRT439 were picked from ampicillin selection plates and used to inoculate liquid media. The culture was grown in 50 to 100 ml NZCYM media (10 g/L casein hydrolysate, 5.0 g/L yeast extract, 1.0 g/L casamino acids, 2.0 g/L MgSO₄·7H₂O) w/ 75 g/ml ampicillin at 37°C until approximately 0.6 A₆₀₀. A 1 ml aliquot was taken to serve as an uninduced control. IPTG was then added to a final concentration of 1 mM, and the cultures were incubated for another 4 h.

Cells were harvested by centrifugation at 5000g for 10 min at 4°C. The pellet was resuspended in 10 ml binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-Cl pH 8.0) containing 1 mg/ml lysozyme and sonicated on ice for 30 sec. Cellular debris was removed by

centrifugation at 10,000g for 20 min at 4°C. The supernatant was applied to a 1 ml Ni-NTA-affinity column (Qiagen). The column was washed extensively (20 to 30 column volumes) with wash buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-Cl pH 8.0) to remove nonspecifically bound proteins. The bound, recombinant His₆ tagged protein was eluted with 1.5 ml of column-stripping buffer (100 mM EDTA, 20 mM Tris-Cl pH 8.0, 0.5 M NaCl). The protein was stored at -20°C until used and was stable through several freeze/thaw cycles.

Cytidine deaminase Assay

The cytidine deaminase assay was performed essentially as described (Wentworth and Wolfenden, 1978). Each reaction included 0.1 ml 1.67 mM cytidine, 0.5 ml 0.1 M Tris-HCl pH 8.0, enzyme, and water to 1 ml final vol for specific activity measurements. The decrease in absorbance at 290 nm was followed with time. As was observed for the *E. coli* enzyme (Wentworth and Wolfenden, 1978) no loss of activity was found when the enzyme was treated with high levels of EDTA. Assays for 2'-deoxycytidine deaminase were performed essentially as described (Ipata and Cercignani, 1978). One unit of activity is defined as the amount of enzyme required to catalyze the deamination of 1 Mol of substrate per min at 27°C in the above assay.

Literature Cited

ABRC (1995) *Arabidopsis* Biological Resource Center: Seed and DNA Stock List. Library information for ESTs. pp. 239

Ackermann RS, Cozzarelli NR, Epstein W (1974) Accumulation of toxic concentration of methylglyoxal by wild-type *Escherichia coli* K-12. *J Bacteriol* **119**:357-362

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local alignment search tool. *J Mol Biol* **215**:403-410
- Ashley GW, Bartlett PA** (1984) Inhibition of *Escherichia coli* cytidine deaminase by a phospho-pyrimidine nucleoside. *J Biol Chem* **259**:13621-13627
- Betts L, Xiang S, Short SA, Wolfenden R, Carter C** (1994) Cytidine deaminase. The 2.3 Å crystal structure of an enzyme: transition-state analog complex. *J Mol Biol* **235**:635-656
- Bradford MM** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. *Anal Biochem* **72**:248-254
- Cacciamani T, Vita A, Cristalli G, Vincenzetti S, Natalini P, Ruggieri S, Amici A, Magni G** (1991) Purification of human cytidine deaminase: molecular and enzymatic characterization and inhibition by synthetic pyrimidine analogs. *Arch Biochem Biophys* **290**:285-292
- Carlow DC, Short SA, Wolfenden R** (1996) Role of glutamate-104 in generating a transition state analogue inhibitor at the active site of cytidine deaminase. *Biochemistry* **35**:948-954
- Carlow DC, Short SA, Wolfenden R** (1998) Complementary truncations of a hydrogen bond to ribose involved in transition-state stabilization by cytidine deaminase. *Biochemistry* **37**:1199-1203
- Driscoll DM, Zhang Q** (1994) Expression and characterization of p27, the catalytic subunit of apolipoprotein B mRNA editing enzyme. *J Biol Chem* **269**:19843-19847
- Freyer R, Kiefer-Meyer MC, Kossel H** (1997) Occurrence of plastid RNA editing in all major lineages of land plants. *Proc Natl Acad Sci USA* **94**:6285-6290
- Ipata P, Cercignani G** (1978) Cytosine and cytidine deaminase from yeast. *Methods-Enzymol* **51**:394-401

- Jong AY, Campbell JL** (1984) Characterization of *Saccharomyces cerevisiae* thymidylate kinase, the CDC8 gene product. General properties. kinetic analysis. and subcellular localization. *J Bio Chem* **259**:14394-14398
- Kadowaki K, Ozawa K, Kazama S, Kubo N, Akihama T** (1995) Creation of an initiation codon by RNA editing in the *coxI* transcript from tomato mitochondria. *Curr Genet* **28**:415-422
- Kozak M** (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283-292
- Lippok B, Brennicke A, Unseld M** (1996) The *rps4*-gene is encoded upstream of the *nad2*-gene in *Arabidopsis* mitochondria. *Biol Chem Hoppe Seyler* **377**:251-257
- Malek O, Lattig K, Hiesel R, Brennicke A, Knoop V** (1996) RNA editing in bryophytes and a molecular phylogeny of land plants. *EMBO J* **15**:1403-1411
- Maley GF** (1978) Deoxycitidylate deaminase from T2-infected *Escherichia coli*. *Methods Enzymol* **51**:412-418
- Matzura O, Wennborg A** (1996) RNAdraw: an integrated program for RNA secondary structure calculation and analysis under 32-bit Microsoft Windows. *Computer Applications in the Biosciences (CABIOS)* **12**:247-249
- Meyers R, Malathi VG, Cox RP, Silber R** (1973) Studies on nucleoside deaminase. Increase in activity in HeLa cell cultures caused by cytosine arabinoside. *J Biol Chem* **248**:5909-5913
- Nakai K, Kanehisa M** (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**:897-911

- Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng B, Davidson NO, Scott J** (1993) The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J Biol Chem* **268**:20709-20712
- Neuhard J, Nygaard P** (1987) Purines and Pyrimidines. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. Neidhardt FC, Ingraham JL, Low BK, Magasanik B, Schaechter M and Umberger HE *ed.* American Society for Microbiology, Washington, DC. pp. 445-473
- Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J** (1987) A novel form of tissue-specific RNA processing produces Apolipoprotein-B48 in intestine. *Cell* **50**:831-840
- Prakash L, Hinkle D, Prakash S** (1979) Decreased UV mutagenesis in *cdc8*, a DNA replication mutant of *Saccharomyces cerevisiae*. *Mol Gen Genet* **172**:249-258
- Rost B, Sander C, Schneider R** (1994) PHD - an automatic mail server for protein secondary structure prediction. *CABIOS* **10**:53-60
- Rothman IK, Malathi VG, Silber R** (1978) Cytidine deaminase from leukemic mouse spleen. *Methods Enzymol* **51**:408-412
- Steuart CD, Burke PJ** (1971) Cytidine deaminase and the development of resistance to arabinosyl cytosine. *Nature, New Biol* **233**:109
- Stougaard J** (1993) Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene. *Plant J* **3**:755-761
- Teng B, Burant CF, Davidson NO** (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260**:1816-1819

Wakasugi T, Hirose T, Horihata M, Tsudzuki T, Kossel H, Sugiura M (1996) Creation of a novel protein-coding region at the RNA level in black pine chloroplasts: the pattern of RNA editing in the gymnosperm chloroplast is different from that in angiosperms. *Proc Natl Acad Sci USA* **93**:8766-8770

Weers B, Thornburg RW (1999) Characterization of the cDNA and gene encoding the *Arabidopsis thaliana* Uracil Phosphoribosyltransferase (AF116860). *Plant Physiol* **119**:1567

Wentworth DF, Wolfenden R (1978) Cytidine deaminases (from *Escherichia coli* and human liver). *Meth Enzymol* **51**:401-407

Xiang S, Short SA, Wolfenden R, Carter CW (1995) Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase. *Biochemistry* **34**:4516-4523

Xiang S, Short SA, Wolfenden R, Carter CW (1996) Cytidine deaminase complexed to deazacytidine: A "valence buffer" in zinc enzyme catalysis. *Biochemistry* **35**:1335-1341

Xiang S, Short SA, Wolfenden R, Carter CW (1997) The structure of the Cytidine deaminase - Product complex provides evidence for efficient proton transfer and ground-state destabilization. *Biochemistry* **36**:4768-4774

Yamanaka K, Ogura T, Niki H, Hiraga S (1992) Identification and characterization of the *smbA* gene, a suppressor of the *mukB* null mutant of *Escherichia coli*. *J Bacteriol* **174**:7517-7526

Zanlungo S, Quinones V, Moenne A, Holuigue L, Jordana X (1995) Splicing and editing of *rps10* transcripts in potato mitochondria. *Curr Genet* **27**:565-571

Zhou L, Lacroute F, Thornburg R (1998) Cloning, expression in *Escherichia coli* and characterization of *Arabidopsis thaliana* uridine 5'-monophosphate/cytidine 5'-monophosphate kinase. *Plant Physiol* 117:245-253

Figure legends

Figure 1. *Arabidopsis thaliana* cytidine deaminase cDNAs. Panel A. Physical structure of the cytidine deaminase EST clones 150P17T7 and 149H23T7 are presented as black lines. Conserved restriction enzyme sites between the two clones are indicated. The stippled box below the clones shows the structural features (5' UTR, coding region, and 3' UTR) of the cDNA. A scale showing the size of the clones in nucleotides is presented below. Panel B. Construction of the full-length cytidine deaminase cDNA. Features of this panel are the same as in Panel A.

Figure 2. Stem-loop structure forms a novel polyadenylation signal. The sequence of the *Arabidopsis thaliana* cytidine deaminase mRNA from 1029 to 1090 is presented. This structure was calculated to form with a ΔG of -38.64 kJ/mol at 20°C.

Figure 3. Phylogenetic analysis of cytidine deaminases. GenBank Accessions for the individual sequences used in this analysis are: *E. coli* GenBank Accession M60916; *Saccharomyces cerevisiae* U20865; *H. sapiens* L27943; *Bacillus subtilis* P19079; *Cenorhabditis elegans* U61949; *Brugia pahangi* X91065; *B. malayi* U80980. *A. thaliana* CYD1 AF134487; *A. thaliana* CYD2 through CYD7, AF080676.

Figure 4. Protein identity matrix comparison of the *Arabidopsis thaliana* cytidine deaminase (GenBank Accession AF1334487) with cytidine deaminase from *E. coli* (M60916), *H.*

sapiens (L27943), and *S. cerevisiae* (U20865). The analysis was performed with 30% identity in a window of 23 amino acids.

Figure 5. *Arabidopsis thaliana* cytidine deaminase active site. Panel A. Alignment of zinc-finger domains from various cytidine deaminases. GenBank Accessions for the individual sequences used in this analysis are presented in the legend to Figure 3. Panel B. Model of the *A. thaliana* cytidine deaminase zinc-finger domain from amino acids 77 to 111. The alpha carbon chain for the peptide is presented. The side chains of the amino acids that participate in the enzyme mechanism are presented. The zinc-coordinating His⁷⁷ is shown in black, and Cys¹⁰³ and Cys¹⁰⁷ are stippled. The active site Glu⁷⁹ is shown in hatched. Panel C. Structural elements at the cytidine deaminase active site. The active site of the *E. coli* is taken from (Betts, et al., 1994). The conserved active site residues of the *Arabidopsis* enzyme are shown interacting with the tetrahedral active site reaction intermediate.

Figure 6. Purification of the recombinant cytidine deaminase on the Ni-NTA-affinity column. Lane 1, proteins from uninduced *E. coli* BL21 cells containing pRT439; Lane 2, proteins from IPTG-induced BL21 cells containing pRT439; Lane 3, purified recombinant cytidine deaminase. Molecular weight marker sizes in kDa are indicated to the left of the figure. Molecular weight markers used were: Myosin, 200 kDa; -galactosidase, 116 kDa; Phosphorylase B, 97 kDa; BSA, 66 kDa; Ovalbumin, 45 kDa; Carbonic Anhydrase, 31 kDa; Soybean Trypsin Inhibitor, 21.5 kDa; Lysozyme, 14.4 kDa; Aprotinin, 6.5 kDa.

Table 1

Substrate specificity

The standard enzyme reaction contained 167 M cytidine analog, 0.05 M Tris-HCl pH 8.0, enzyme, and water in a final volume of 1 ml. The decrease in absorbance at 290 nm was followed with time. Experiments were performed in triplicate or quadruplicate. For determination of the K_m values the standard enzyme reaction contained variable amounts of cytidine or 2'-deoxycytidine ranging from 50 to 500 M. One unit of cytidine deaminase is defined as the amount of enzyme required to deaminate 1 Mole of cytidine per minute in the above standard reaction mixture at 25°C. Specific activity is units per milligram of protein as determined by the method of Bradford Bradford, 1976

substrate	specific activity ^a	K_m ^b	V_{max} ^c
cytidine	12.0 \pm 1.6	226.1 \pm 45	39.7 \pm 3.8
2'-deoxycytidine	20.3 \pm 3.6	49.3 \pm 6.7	24.4 \pm 1.3
CMP	nd ^d		
dCMP	nd		
cytosine	nd		

^a units per mg protein

^b μ molar

^c μ moles/min

^d no deaminase activity was detected

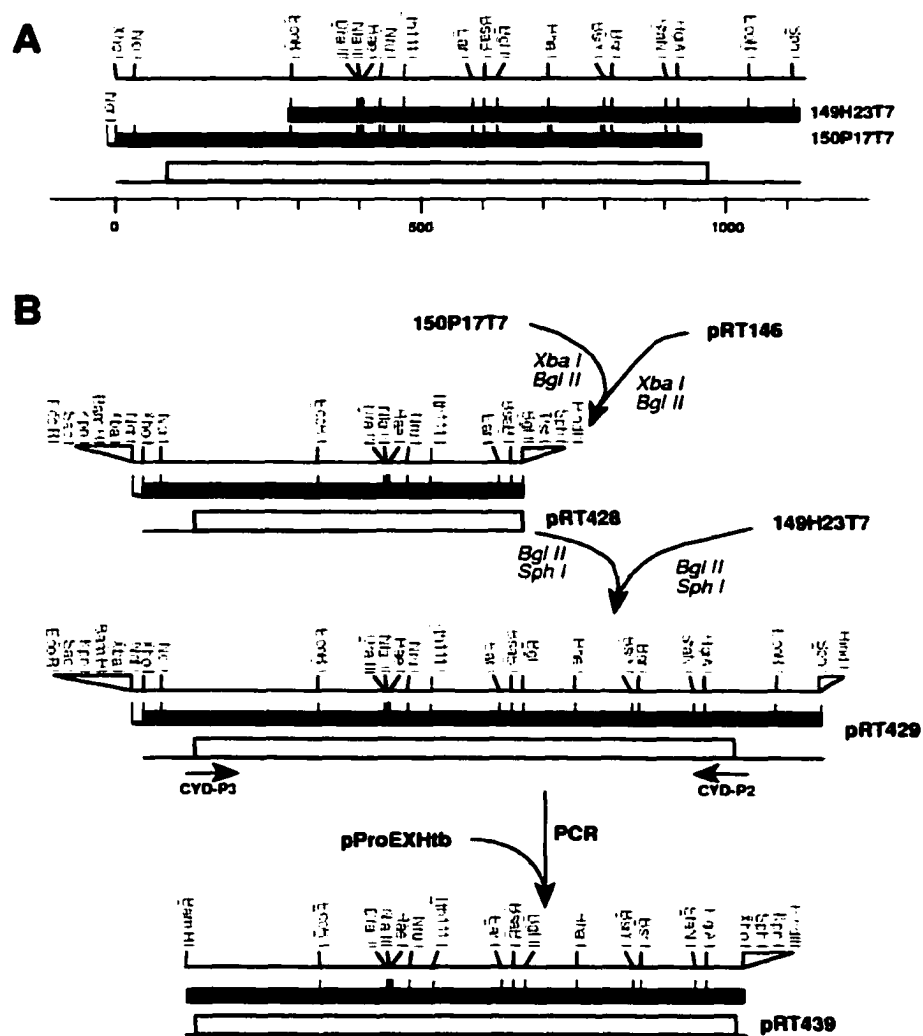


Figure 1.

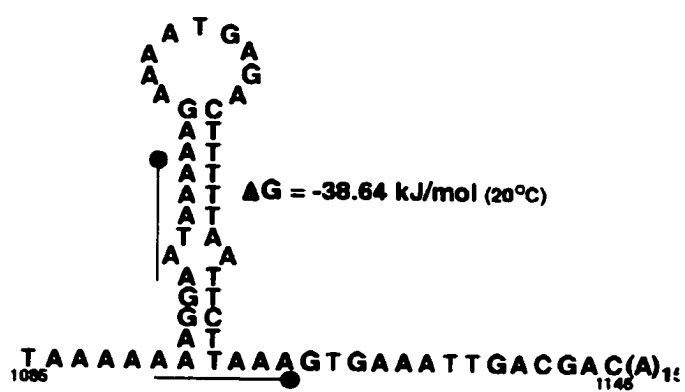


Figure 2.

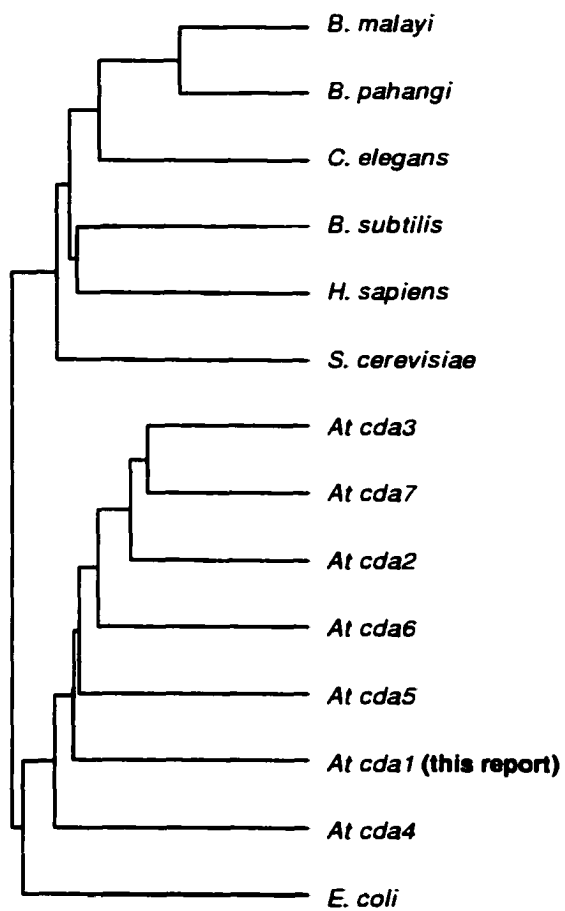


Figure 3.

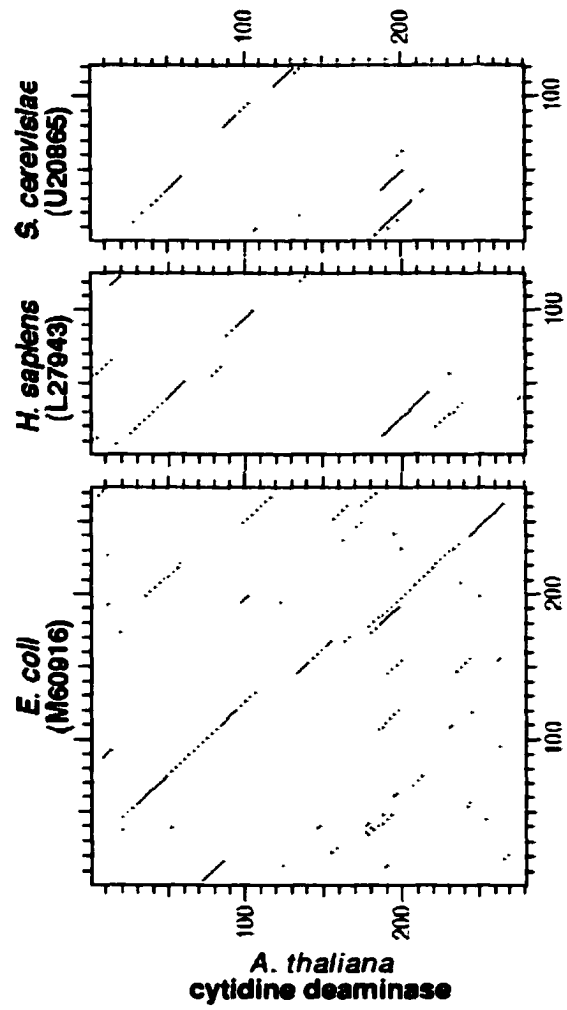


Figure 4.

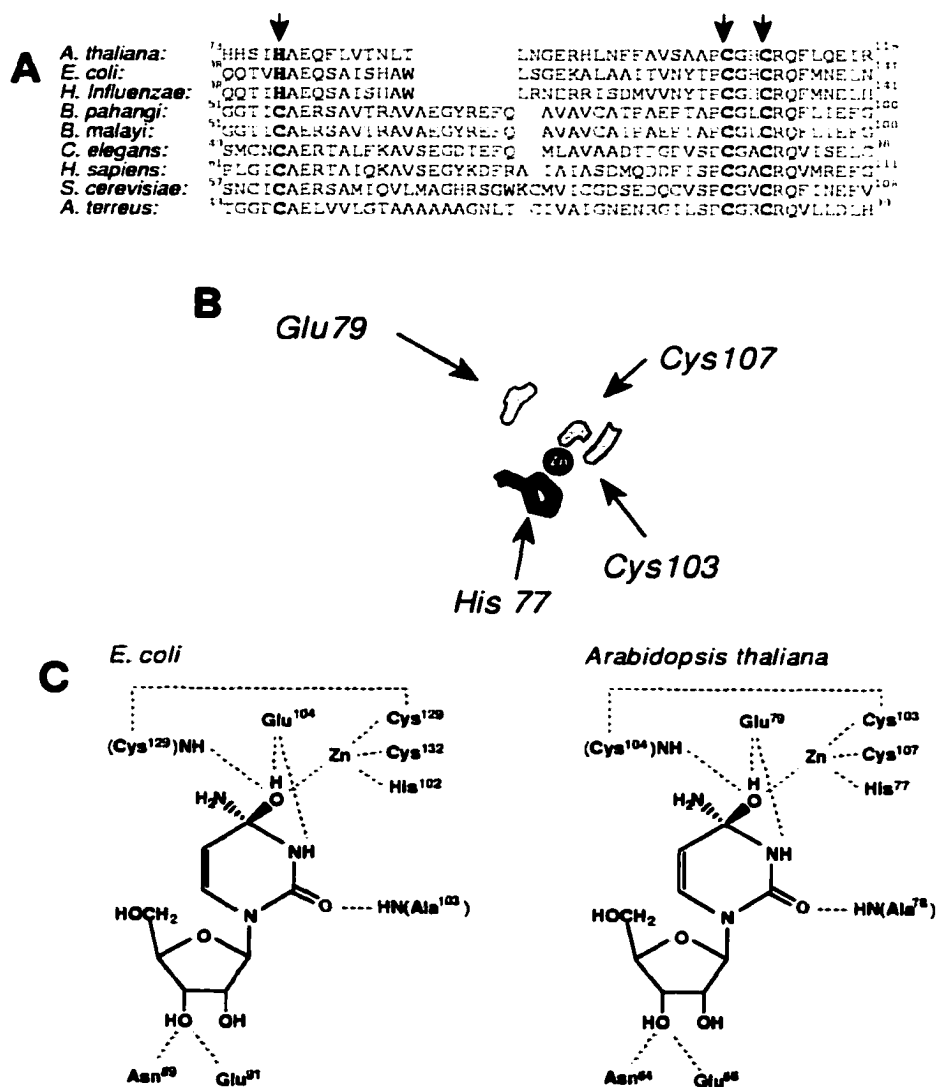


Figure 5.

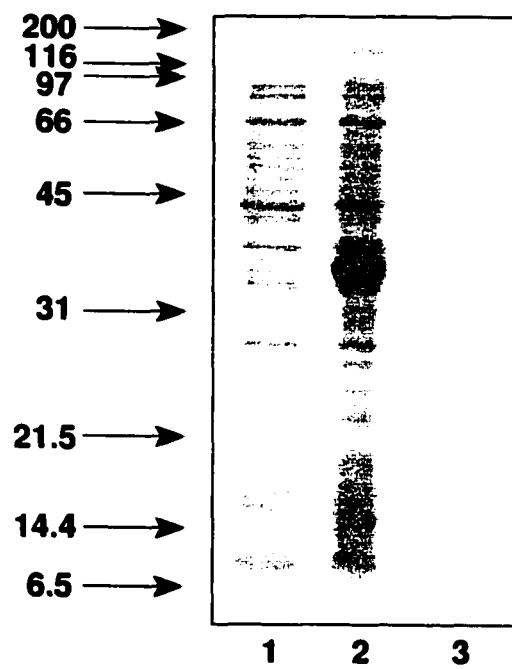


Figure 6.

CHAPTER 3. THYMINE-STARVATION UP-REGULATES A REGION OF THE *ARABIDOPSIS THALIANA* CHROMOSOME III CONTAINING BOTH THE UMP SYNTHASE GENE AND A PROXIMAL GENE ENCODING A NOVEL F-BOX PROTEIN

A paper published in the Journal of Plant Biology 45:161-169

Chris Kafer & Robert Thornburg

Abstract

We have cloned the *Arabidopsis thaliana* uridine 5'-monophosphate synthase (UMP synthase) locus and characterized transcript levels of this gene as well as the transcript encoding a novel F-box protein found upstream of UMP synthase, *Skp1*-interacting partner 5 (*SKIP5*). The 3' end of the *SKIP5* gene is 405 bp from the 5' end of the UMP synthase gene. To determine whether these proximate genes are coordinately or independently regulated, we have used an RT-PCR method to quantitate the transcript levels of each gene relative to 18S ribosomal RNA. Previous work has demonstrated an up-regulation of the UMP synthase gene in tobacco callus in the presence of compounds such as 5-fluoroorotate and aminopterin that induce thymine starvation. Here, we present results showing that both the UMP synthase and *SKIP5* genes are coordinately up-regulated in the presence of fluoroorotic acid.

Introduction

Pyrimidine metabolism is fundamental in all life forms. Pyrimidines are required for DNA and RNA biosynthesis, for the interconversion of carbohydrates, for polysaccharide and glycoprotein biosynthesis, for glycolipid and phospholipid biosynthesis, and many other cellular processes. When pyrimidine nucleotides become limiting, growth ceases (Jones and Hahn, 1979, Löffler, et al., 1997, Santoso and Thornburg, 1998, Sung and Jacques, 1980).

Therefore, understanding nucleotide biosynthesis is essential to gain a complete understanding of primary cellular metabolism.

De novo pyrimidine biosynthesis in plants is catalyzed by 6 enzymatic steps. cDNAs encoding every step have been cloned from *Arabidopsis* (Brandenburg, et al., 1998, Minet, et al., 1992, Nasr, et al., 1994, Williamson, et al., 1996, Zhou, et al., 1997). The final two steps of the pathway result in the conversion of orotic acid into UMP. In the first of these steps, the ribosyl-5'-phosphate group from phosphoribosyl pyrophosphate (PRPP) is transferred onto orotic acid to form orotidylate monophosphate (OMP) by the chloroplastic enzyme orotate phosphoribosyltransferase. OMP is subsequently decarboxylated to form UMP by OMP decarboxylase. In prokaryotes and lower eukaryotes (yeast), these two enzymes are separate and distinct. However, in higher eukaryotes, including plants, the genes for these two enzymes have been fused to form a single transcriptional unit that produces a single bifunctional protein, termed UMP synthase. The plant UMP synthase thus harbors both orotate phosphoribosyl transferase and OMP decarboxylase activity (Santoso and Thornburg, 1998, Santoso and Thornburg, 1992). UMP synthase is also the rate limiting step in pyrimidine biosynthesis in plants (Santoso and Thornburg, 1992).

While there is considerable information regarding the biochemistry of pyrimidine biosynthesis, regulation of expression of the pyrimidine biosynthetic genes has lagged far behind. UMP synthase is the best studied of the *de novo* pyrimidine biosynthetic genes in plants. It is transcriptionally-regulated by thymine levels within cells (Santoso and Thornburg, 1998). When thymine becomes limiting, cells respond by activation of a transcriptional program that results in the up-regulation of several genes in the *de novo* pyrimidine biosynthetic pathway, including UMP synthase. Recently, we have

demonstrated that some fluoroorotic acid-selected *Nicotiana plumbaginifolia* cell lines have lost this thymine-regulated transcriptional program and are unable to respond to alteration in thymine levels (Santoso and Thornburg, 2000). In an effort to better understand the regulation of the UMP synthase, we have isolated the locus encoding UMP synthase and have characterized gene expression from this locus.

Materials and Methods

Materials

The *Escherichia coli* strain, BL21:p33, expressing *Taq* polymerase (Engelke, et al., 1990) was kindly supplied to us by Dr. Dan Voytas, Iowa State University. The *Taq* polymerase was purified by the method of Desai and Pfaffle (Desai and Pfaffle, 1995). α -[^{32}P]-dCTP was obtained from New England Nuclear (Boston, MA, USA). MMLV reverse transcriptase was purchased from Promega, (Madison, WI, USA) and used according to the manufacturer's instructions. Random hexamers were obtained from the Iowa State University DNA synthesis facility. All other materials were of the highest quality obtainable and were obtained from Fischer Scientific Co. (Pittsburgh, PA, USA) or from Sigma Chemical Co. (St. Louis, MO, USA).

Gene isolation and Sequencing

To isolate the *Arabidopsis* UMP synthase gene, an *Arabidopsis thaliana* genomic library (Voytas, et al., 1990) was screened with the *Arabidopsis* UMP synthase cDNA (Nasr, et al., 1994). Positive plaques were taken through three rounds of screening. This yielded two positive clones called *fAt-UMPS1* and *fAt-UMPS2*. Both bacteriophage contained a 7 kb *XbaI* insert that hybridized with the UMP synthase cDNA. The hybridizing band was subcloned into the *XbaI* sites pBlueScript (SK+). The resulting clone,

pRT357, contained the full-length UMP synthase gene. All sequencing was performed at the Iowa State University DNA facility using the Applied Biosystems Prism Dye-deoxy Cycle Sequencing Kit (Perkin-Elmer Corp., Norwalk, CT, USA). The reactions were run on an Applied Biosystems Prism 377 DNA sequencer (Perkin-Elmer Corp., Norwalk, CT, USA). Sequencing was initiated from known vector sequences. On the basis of these runs, primers specific to the UMP synthase sequence were constructed. Both strands were sequenced in duplicate or triplicate.

Plants and Plant Growth

Arabidopsis thaliana cv. Columbia plants were used for all studies. Aerial tissues (leaves, stems, flowers and siliques) were harvested from young, wild type plants grown in soil. Tissue was either harvested and used immediately or frozen at -70°C until use. Root tissue was obtained from plants grown *in vitro* as described (Ostin, et al., 1998). For *in vitro* grown plants, approximately 25 surface sterilized (50% NaOCl, 10 min; 2 washes in sterile water) *Arabidopsis* seeds were used to inoculate 250 ml flasks containing 25 ml of MS medium + 2% glucose (Murashige and Skoog, 1962). The flasks were maintained with constant shaking at 100 rpm under continuous illumination for various times. After approximately 14 days of growth, the mass of plants was aseptically removed from the flask and briefly blotted to remove excess growth medium. The root tissue was excised and either used immediately for RNA isolation or frozen in aliquots at -70°C until use.

RNA Isolation

RNA was isolated using a modified guanidine thiocyanate method (Chomczynski and Sacchi, 1987). Frozen or freshly harvested tissue (100-300 mg) was ground in 1 ml of

extraction buffer (4M guanidine thiocyanate, 25 mM sodium citrate, 0.5% sarkosyl, 0.1 M 2-mercaptoethanol, 0.3 M sodium acetate pH 4.8) and extracted with 0.5 ml of 0.1 M sodium acetate saturated phenol:chloroform (5:2) pH 4.8 by vigorous vortexing in a 1.5 ml microfuge tube followed by a 3 min centrifugation at 17,000 x g. The upper aqueous phase was extracted the same way once more followed by a final chloroform extraction. RNA was precipitated by adding an equal volume of 2-propanol to the aqueous phase, mixing by inversion and then room temperature incubation for 5 min followed by a 10 min centrifugation at 17,000 x g. The RNA pellet was washed with 70% ethanol and centrifuged for 1 min. The pellet was dried and resuspended in 90 mL water. Contaminating DNA was removed by incubation with 2 units of RNase-free DNase in the manufacturers buffer (Promega, Madison WI, USA) at 37°C for 45 min followed by a single phenol:chloroform extraction and 2 propanol precipitation as before. The final dried pellet of RNA was resuspended in 35 mL of RNase free water. The RNA was quantitated by measuring absorbance at 260 nm on a Beckman DU-7400 spectrophotometer. RNA was stored at -70°C until use.

Reverse Transcription

Approximately 2.5 µg of RNA was incubated in a total volume of 15 ml with 2.5 mM random hexanucleotides and 2.5 mM each dNTP at 75°C for 5 min. then cooled to room temperature. Four ml of the manufacturers 5x MMLV buffer and 200 units of MMLV reverse transcriptase (Promega, Madison WI, USA) were added to each tube and mixed gently. The reverse transcription reaction was carried out at 37°C for 1 h and the reaction products were stored at -20°C until use.

Polymerase Chain Reaction

For all analyses a 50 µl “Hot Start” PCR was used (Klebe, et al., 1999). The reaction mixture contained 50 mM KCl, 10 mM Tris pH 9.0, 0.1% Triton X-100, 0.2 mM of each dNTP, 0.5 mM each primer, 1 µl Taq polymerase (Desai and Pfaffle, 1995), 2.5 mM MgCl₂ and 1 µl of the reverse transcription reaction. The PCR was performed in an MJ Research PTC-100 thermal cycler (MJ Research, Incline Village, NV, USA) with the following cycling parameters: denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 60 s. The number of cycles varies with each experiment and is presented in the individual figure legends.

[³²P]-labeled nucleotide triphosphates are incorporated in the PCR reaction to quantitate the PCR product yields. Following PCR, the reaction products were separated on an 8% polyacrylamide gel. The radiolabeled PCR products within the gel were visualized and quantitated using a Molecular Dynamics PhosphorImager and associated ImageQuant software (Amersham Biosciences, Piscataway, NJ, USA), .

Oligonucleotides

Table I shows the various oligonucleotides used in this study. Non-modified oligonucleotides were synthesized at the Iowa State University Nucleic Acid Facility. Oligonucleotides modified at the 3' end were synthesized by Sigma/Genosys (St. Louis, MO, USA).

Results

Cloning of the *Arabidopsis* UMP synthase Locus

The recent completion of the sequence of the *Arabidopsis thaliana* genome has confirmed our experimental characterization of the UMP synthase locus as well as filling some gaps in our knowledge. We originally cloned this region (Fig. 1) via traditional

genomic library screening (Voytas, et al., 1990) with the *Arabidopsis* UMP synthase cDNA (Nasr, et al., 1994). This locus was found to be single copy by Southern blot analysis (Fig. 2). The full length UMP synthase gene as well as 2090 bp upstream and 448 bp downstream were isolated. The most interesting feature of this locus is the proximity of an open reading frame immediately upstream of the coding region of the UMP synthase gene. The proximity of this gene was verified in genomic DNA by PCR analysis (Fig. 3). This ORF was originally identified as EST clone T88668. Recently a routine BLAST search of the T88668 protein sequence has revealed that this gene codes for the F-box protein Skp1 Interacting partner 5 (SKIP5) isolated by Koncz and coworkers using SKP1 (ASK1) as bait (Farras, et al., 2000). Due to the release of the *Arabidopsis* genome sequence, we have been able to verify that our clone was in fact correct and not a rearrangement during library construction, to place the gene on chromosome III at 92.5 cM and to verify our Southern analysis, that this gene is in fact single copy within the *Arabidopsis* genome.

Arabidopsis UMP synthase protein

A multiple sequence alignment of the available UMP synthase protein sequences using the "pileup" program in the Wisconsin GCG package was performed. The sequences used are: *Arabidopsis thaliana*, *Bos taurus*, *Caenorhabditis elegans*, *Dictyostelium discoideum*, *Drosophila melanogaster*, *Homo sapiens*, *Mus musculus*, *Naegleria gruberii*, *Nicotiana plumbaginifolia*, *Nicotiana tabacum*, *Oryza sativa* UMP synthase1 & 2 and *Zea mays* UMP synthase1. This program grouped the proteins into 3 clades with 2 outliers (Fig. 4). The mammalian sequences and the plant sequences formed 2 individual clades with the *Drosophila* and *Dictyostelium* sequences forming the third. The plant clade was further

branched into dicot and monocot sequences. The *C. elegans* and *N. gruberii* did not fall into any of the three clades formed by these sequences.

The bifunctional *Arabidopsis* UMP synthase is similar to other UMP synthases from higher eukaryotes. The N-terminal half of the gene (amino acids 1-220) encodes the orotate phosphoribosyl transferase function and shows 23% identity to the yeast *ura5* protein while the C-terminal half of the gene (amino acids 220-477) encodes the orotidine decarboxylase function and shows 49% identity to the yeast *ura3* protein.

When we examined the putative amino acid sequence of the *Arabidopsis* UMP synthase, we found at least 9 highly conserved residues (K27, R93, K94, K97, I104, E119, D120, R151, G425) that have been identified as catalytically important in crystallographic studies of prokaryotic orotate phosphoribosyl transferase and yeast orotidine decarboxylase (Smiley and Jones, 1992, Grubmeyer, et al., 1993, Scapin, et al., 1994, Scapin, et al., 1995, Ozturk, et al., 1995, Suchi, et al., 1997). These comparisons strongly suggest that the catalytic mechanism of the plant UMP synthases is identical to that of the yeast, Salmonella and Human UMP synthases.

Arabidopsis SKIP5

A routine blast search using the sequence of T88668 from our *Arabidopsis* UMP synthase clone has identified 2 overlapping EST clones (190E9T7 and 214F2T7) and a genomic BAC (T14E10). From these sequences we have determined a full length sequence of this transcript as well as the gene structure. The full length transcript is 1001 nucleotides in length. The gene contains 4 introns of variable length (85 to 911 nucleotides). Comparison of the EST clones to the genomic sequence revealed that clone 190E9T7 contained an unspliced intron 4. To test whether this was a cloning artifact or a form of

alternative splicing we performed a quantitative RT-PCR analysis using oligos that flank intron 4. We tested RNA isolated from root, leaf, stem and flowers as well as thymine starved whole plants. In every case we observed a band that corresponded to unspliced transcript. Quantitation of this band relative to spliced transcript shows $5.4 \pm 1.2\%$ of all transcripts were unspliced. We conclude that this intron may be poorly spliced and the EST clone 190E9T7 simply reflects this.

The full length cDNA encodes a 274 amino acid protein of 30624 Da and a predicted pI of 8.59. A PROSITE analysis of the protein reveals 2 motifs. Amino Acids 12-29 are a predicted bipartite nuclear localization signal (Dingwall and Laskey, 1986), amino acids 32-79 encode a predicted F-box motif (Bai et al. 1996) and amino acids 154-175 encode a single imperfect leucine rich repeat. This suggests that T88668 may be a nuclear protein involved in transcriptional regulation of an as yet unknown factor.

An ortholog of this protein was identified in the tomato EST databases. This cDNA was obtained and sequenced (GenBank Accession # AY056053). The tomato homolog is 18 amino acids shorter at its N-terminus but has a 15 amino acid extension at its C-terminus. The central portions of the two proteins showed 68% amino acid identity. No orthologs are found in any vertebrate (human, mouse, rat, pig, cattle or zebrafish), invertebrate (*Drosophila* or *C. elegans*), yeast or prokaryotic databases. Thus, this sequence may represent a novel plant specific F-box protein.

Koncz and coworkers recently isolated this clone in a yeast two-hybrid using *ASK1* (*SKIP1*) as a bait and have subsequently renamed it *SKIP5* (*SKiP1 Interacting Partner 5*) (Farras, et al., 2000). The fact that this protein interacts with *ASK1*, a member of the SCF

complex, is further evidence that it may be involved in transcriptional regulation of unknown factor(s) via an SCF complex.

Expression of *SKIP5* and UMP synthase mRNAs

To evaluate expression from this locus, we developed a semiquantitative RT-PCR assay to monitor the mRNA levels of the *SKIP5* and UMP synthase genes. This assay was developed using total RNA from 7-day old *Arabidopsis* plants grown *in vitro* as described in Experimental Procedures.

For RT-PCR methods to provide useful information, the measurements must be taken during the linear phase of the PCR process. Starting with the oligonucleotides specific for UMP synthase, we examined the accumulation of the UMP synthase PCR products from reverse transcribed leaf RNA. The UMP synthase PCR product began to accumulate after 18 cycles of amplification and the accumulation was linear through cycle 28 (data not shown).

A second concern is the use of an internal standard. Ideally, the internal standard should be invariant in all tissues and under all treatments. Because almost 80% of total RNA is rRNA, the level of the rRNA remains constant between samples and has a nearly constant steady state level. We have, therefore, chosen to compare all reactions to the 18S rRNA. This RNA species, however, is so abundant that it is difficult to find conditions that allow both the standard 18S and gene of interest to be simultaneously evaluated. To solve this problem, we are using modified competitive oligonucleotides (MCO).

These modified 18S rRNA oligonucleotides are identical in sequence with the non-modified 18S rRNA oligonucleotides (see Table I) with the exception that the 3' terminal

residue is modified so that cDNA cannot be extended from these oligonucleotides. These oligonucleotides can compete for binding to the 18S ribosomal RNA, however the 3' modifications block their ability to serve as primers in the polymerase chain reaction.

Using a mixture of the normal 18S oligonucleotides and the modified 18S oligonucleotides, we can reduce 18S PCR product to levels that are similar to that of the UMP synthase PCR products. Increasing the ratio of the modified to normal 18S oligonucleotides significantly reduces the accumulation of the 18S product. Over a range of ratios from 0.05 to 0.15 normal to modified 18S oligonucleotides, the levels of 18S rRNA to UMP synthase PCR products were roughly equivalent. Therefore, all further studies were performed at or near a ratio of 0.05 normal to modified oligonucleotides.

We next determined that all PCR products (UMPS, SKIP5 and 18S rRNA) could be measured in the linear PCR range as a function of cycle number. In this analysis, the ratio of normal to modified 18S rRNA oligonucleotides was evaluated at both 0.05 and 0.10. In both cases, the accumulation of all PCR products was linear from cycles 20 to 28. All additional analyses were performed at cycle 26.

Fluoroorotic Acid Induction

We initially evaluated whether the RT-PCR assay was capable of detecting changes in the UMP synthase mRNA levels. In our previous work with tobacco we have shown that UMP synthase levels are transcriptionally up-regulated by thymine starvation (Santoso and Thornburg, 1998). Therefore, we grew *Arabidopsis* plants in tissue culture flasks (Ostin, et al., 1998) and induced the plants with the addition of 20 mg/ml fluoroorotic acid (Santoso and Thornburg, 1998). Fig. 5 shows the RT-PCR profiles of a representative experiment for uninduced and FOA-induced plants. The bands from these gels were quantitated by

PhosphorImager and these data are presented in Table II. In these studies, the UMP synthase mRNA was up-regulated greater than 3-fold. This induction of the UMP synthase mRNA is in-line with the previously published levels for the induction of this gene (Santoso and Thornburg, 1998). Thymine starvation, induced by the addition of FOA, also caused the up-regulation of the *SKIP5* mRNA. The *SKIP5* mRNA was induced greater than 10-fold. Because pool sizes of pyrimidines are large within cells, the induction requires several days to achieve maximal effect (Santoso and Thornburg, 2000). For both genes, maximal induction was observed on day 5 and by day 7 the mRNA levels had begun to decline, implying a coordinated induction of both genes. A similar up-regulation of the UMP synthase and *SKIP5* mRNAs were observed when plants were thymine starved by the addition of 25 mg/ml aminopterin for 7 days (data not shown).

Tissue Specific Expression

Since we had verified that the RT-PCR assay successfully monitored the levels of UMP synthase in whole plants following the addition of fluoroorotic acid, we next examined the level of expression in various plant organs. Table III shows the results of a representative experiment. The highest levels of UMP synthase mRNA were observed in leaves and flowers. Stems and siliques were significantly lower and roots showed very low levels of UMP synthase mRNA. The tissue-specific pattern of expression for the *SKIP5* mRNA was similar. The highest levels were observed in leaves and flowers with stems and siliques significantly lower and roots showing the poorest levels of expression.

Discussion

UMP synthase is the rate limiting and final step of pyrimidine biosynthesis in plants (Santoso and Thornburg, 1992). It is transcriptionally up-regulated in response to thymine

starvation (Santoso and Thornburg, 1998). The UMP synthase locus in *Arabidopsis thaliana* harbors two genes in close proximity, UMP synthase and *SKIP5*. To evaluate the expression of these two genes, we have developed a semi-quantitative RT-PCR assay using attenuated 18S rRNA as a standard. Because 18S rRNA comprises 80% of total cellular RNA, we designed modified competitive oligonucleotides (MCOs) to attenuate the signal from the 18S rRNA. These MCOs are identical in sequence to the unmodified primer and will compete for the binding site but have a 3' C3 linker. Therefore these MCOs cannot initiate chain elongation. Using the native 18S rRNA oligonucleotides plus the 18S rRNA MCOs together with the gene specific oligonucleotides in ratios that permit the linear amplification of each of the PCR products from reverse transcribed RNA has permitted us to develop a single pot semiquantitative RT-PCR assay to evaluate gene expression.

The results of these studies confirm in *Arabidopsis*, our earlier observations on the induction of UMP synthase during nucleotide starvation in tobacco. These results also extend these observations by demonstrating that the F-box protein, *SKIP5*, is similarly regulated by thymine starvation. The FOA-induced patterns of expression for both *SKIP5* and UMP synthase implies coordinated regulation of these genes. However, a comparison of the 5' flanking regions of these genes failed to identify any sequences sharing significant identity between these two genes. Alternatively, FOA-induction of UMP synthase appears to result from derepression of that gene (Santoso and Thornburg, 2000). Therefore, it may be possible that a coordinated silencing mechanism is active at this locus to maintain low levels of expression of both mRNAs. Then as a result of thymine starvation a single event could result in the derepression of both genes.

F-box proteins are an increasingly large family of proteins (Xiao and Jang, 2000) that function in the transcriptional regulation of a wide variety of developmental and biosynthetic pathways. In plants, F-box proteins have been shown to function in jasmonate defense responses (Xie, et al., 1998), auxin response (Ruegger, et al., 1998), floral organ development (Samach, et al., 1999) and circadian oscillations (Somers, et al., 2000, Nelson, et al., 2000).

F-box proteins have two distinct functions. First, they interact with target proteins that are frequently transcriptional repressor molecules. A great diversity of sequences associated with F-box proteins, WD40 repeats, Ankyrin repeats, Kelch repeats, and leucine rich repeats, results in a large number of molecules that can be recognized by the F-box proteins. Second, the F-box motif links the F-box protein to an SCF (Skp1-cullin-F-box protein) complex. These SCF complexes comprise a new class of ubiquitin E3 ligases that ubiquitinate proteins recruited to the F-box protein via specific protein-protein interactions of the diverse repeat sequences. The result is the targeted degradation of specifically recruited proteins (transcriptional regulators) at the 26S proteasome (del Pozo and Estelle, 1999, del Pozo and Estelle, 2000).

Skip5 is an F-box protein that interacts with the *Arabidopsis Skp1* homolog (*Ask1*). Its expression is up-regulated in response to nucleotide starvation; however, the target of the *Skip5* protein is unknown. Studies to ascertain the function of *Skip5* are in progress.

Literature Cited

Brandenburg, S. A., C. L. Williamson and R. D. Slocum. 1998. Characterization of a cDNA encoding the small subunit of *Arabidopsis* carbamoyl phosphate synthetase (Accession No. U73175). *Plant Physiol.* **117**: 717.

- Chomczynski, P. and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal Biochem.* **162**: 156-9.
- del Pozo, J. and M. Estelle.** 1999. Function of the ubiquitin-proteosome pathway in auxin response. *Trends Plant Sci.* **4**: 107-112.
- del Pozo, J. C. and M. Estelle.** 2000. F-box proteins and protein degradation: an emerging theme in cellular regulation. *Plant Mol Biol.* **44**: 123-8.
- Desai, U. and P. Pfaffle.** 1995. Single step purification of a thermostable DNA polymerase expressed in *Escherichia coli*. *BioTechniques.* **19**: 780-784.
- Engelke, D. R., A. Krikos, M. E. Bruck and D. Ginsburg.** 1990. Purification of *Thermus aquaticus* DNA polymerase expressed in *Escherichia coli*. *Anal Biochem.* **191**: 396-400.
- Farras, R., A. Ferrando, J. Jasik, L. Okresz, A. Tiburcio, J. Schell, K. Salchert and C. Koncz.** 2000. *Arabidopsis thaliana SKP1* interacting partner 5 (*SKIP5*) mRNA.. GenBank Accession No. AF263381. pp.
- Grubmeyer, C., E. Segura and R. Dorfman.** 1993. Active site lysines in orotate phosphoribosyltransferase. *J Biol Chem.* **268**: 20299-20304.
- Jones, G. E. and J. Hahn.** 1979. *Haplopappus gracilis* cell strains resistant to pyrimidine analogues. *Theor. Appl. Genet.* **54**: 81-87.
- Klebe, R., S. Roriguez, M. VerBeek and T. Giambernardi.** 1999. Simple method for "Hot Starting" RT-PCR. *BioTechniques.* **27**: 1108-1110.
- Löffler, M., J. Jockel, G. Schuster and C. Becker.** 1997. Dihydroorotat-ubiquinone oxidoreductase links mitochondria in the biosynthesis of pyrimidine nucleotides. *Mol Cell Biochem.* **174**: 125-129.

- Minet, M., M.-E. Dufour and F. Lacroute.** 1992. Complementation of *Saccharomyces cerevisiae* auxotrophic mutants by *Arabidopsis thaliana* cDNAs. *Plant J.* **2**: 417-422.
- Murashige, T. and F. Skoog.** 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant.* **15**: 473-497.
- Nasr, F., N. Berthauche, M.-E. Dufour, M. Minet and F. Lacroute.** 1994. Heterospecific cloning of *Arabidopsis thaliana* cDNAs by direct complementation of pyrimidine auxotrophic mutants of *Saccharomyces cerevisiae*. I. Cloning and sequence analysis of two cDNAs catalysing the second, fifth and sixth steps of the *de novo* pyrimidine biosynthesis pathway. *Mol. Gen. Genet.* **244**: 23-32.
- Nelson, D., J. Lasswell, L. Rogg, M. Cohen and B. Bartel.** 2000. FKF1, a clock-controlled gene that regulates the transition to flowering in *Arabidopsis*. *Cell.* **101**: 331-340.
- Ostin, A., M. Kowalczyk, R. Bhalerao and G. Sandberg.** 1998. Metabolism of indole-3-acetic acid in *Arabidopsis*. *Plant Physiol.* **118**: 285-296.
- Ozturk, D., R. Dorfman, G. Scapin, J. Sacchettini and C. Grubmeyer.** 1995. Locations and functional roles of conserved lysine residues in *Salmonella typhimurium* orotate phosphoribosyltransferase. *Biochemistry.* **34**: 10755-10763.
- Ruegger, M., E. Dewey, W. M. Gray, L. Hobbie, J. Turner and M. Estelle.** 1998. The TIR1 protein of *Arabidopsis* functions in auxin response and is related to human SKP2 and yeast *grr1p*. *Genes Dev.* **12**: 198-207.
- Samach, A., J. Klenz, S. Kohalmi, E. Risseuw, G. Haughn and W. Crosby.** 1999. The UNUSUAL FLORAL ORGANS gene of *Arabidopsis thaliana* is an F-box protein required for normal patterning and growth in the floral meristem. *Plant J.* **20**: 433-445.

- Santoso, D. and R. W. Thornburg.** 1992. Isolation and Characterization of UMP synthase mutants from haploid cell suspensions of *Nicotiana tabacum*. *Plant Physiol.* **99**: 1216-1225.
- Santoso, D. and R. W. Thornburg.** 1998. UMP synthase is transcriptionally regulated by pyrimidine levels in *Nicotiana plumbaginifolia*. *Plant Physiol.* **116**: 815-821.
- Santoso, D. and R. W. Thornburg.** 2000. Fluoroorotic acid-selected *Nicotiana plumbaginifolia* cell lines with a stable thymine starvation phenotype have lost the thymine-regulated transcriptional program. *Plant Physiol.* **123**: 1517-1524.
- Scapin, G., C. Grubmeyer and J. Sacchettini.** 1994. Crystal structure of orotate phosphoribosyltransferase. *Biochemistry.* **33**: 1287-1294.
- Scapin, G., D. Ozturk, C. Grubmeyer and J. Sacchettini.** 1995. The crystal structure of the orotate phosphoribosyltransferase complexed with orotate and alpha-D-5-phosphoribosyl-1-pyrophosphate. *Biochemistry.* **34**: 10744-10754.
- Smiley, J. and M. Jones.** 1992. A unique catalytic and inhibitor-binding role for Lys93 of yeast orotidylate decarboxylase. *Biochemistry.* **31**: 12162-12168.
- Somers, D. E., T. F. Schultz, M. Milnamow and S. A. Kay.** 2000. ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell.* **101**: 319-29.
- Suchi, M., H. Mizuno, Y. Kawai, T. Tsuboi, S. Sumi, K. Okajima, M. Hodgson, H. Ogawa and Y. Wada.** 1997. Molecular cloning of the human UMP synthase gene and characterization of point mutations in two hereditary orotic aciduria families. *Am J Hum Genet.* **60**: 525-539.
- Sung, Z. R. and S. Jacques.** 1980. 5-fluorouracil resistance in carrot cell cultures: Its use in studying the interaction of the pyrimidine and arginine pathways. *Planta.* **148**: 389-396.

Voytas, D. F., A. Konieczny, M. P. Cummings and F. M. Ausubel. 1990. The structure, distribution and evolution of the *Tal* retrotransposable element family of *Arabidopsis thaliana*. *Genetics*. **126**: 713-721.

Williamson, C. L., M. R. Lake and R. D. Slocum. 1996. A cDNA encoding carbamoyl phosphate synthetase large subunit (*carB*) from *Arabidopsis* (Accession No. U40341). *Plant Physiol.* **111**: 1354.

Xiao, W. and J. Jang. 2000. F-box proteins in Arabidopsis. *Trends Plant Sci.* **5**: 454-457.

Xie, D., B. Feys, S. James, M. Nieto-Rostro and J. Turner. 1998. COI1: an Arabidopsis gene required for jasmonate-regulated defense and fertility. *Science*. **280**: 1091-1094.

Zhou, L., F. Lacroute and R. W. Thornburg. 1997. DNA sequence of dihydroorotase from *Arabidopsis thaliana* (AF000146). *Plant Physiol.* **114**: 1569.

Footnotes

1) This work was sponsored by funds from the U.S. Department of Agriculture (91-37301-6208), from the Carver Trust, the Hatch Act and State of Iowa Funds.

2) Journal Paper No. J-19127 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 3202.

4) Abbreviations: Cf - final concentration; MMLV – Maloney Murine Leukemia Virus; MCO, modified competitive oligonucleotides;

5) Keywords: UMP synthase, *Nicotiana plumbaginifolia*, fluoroorotic acid, thymine, pyrimidine metabolism, F-Box protein.

6) GenBank Accessions for sequences reported in these studies are: *Arabidopsis thaliana* UMP synthase genomic locus, AF276887; *Nicotiana plumbaginifolia* UMP

synthase gene, AF277455; *Arabidopsis thaliana* SKIP5 cDNAs, AF276888, AF347971, and AF347972; *Lycopersicon esculentum* SKIP5 cDNA, AY056053.

Figure legends

Figure 1. Structure of the UMP synthase genomic locus. A 5.2 kb stretch of DNA was sequenced and analyzed. The location of the SKIP5 cDNA (hatched boxes) and the UMP synthase cDNA (open boxes) are indicated.

Figure 2. Genomic Southern. Ten μ g of genomic DNA isolated from *A. thaliana* cv. Columbia plants was digested to completion with the following restriction endonucleases: Lane 1, Eco RI; lane 2, Hind III; lane 3, Bam HI; and lane 4, Xba I. The digested DNA was run on a 1% agarose gel. Following electrophoresis, the DNA was blotted to a nylon membrane and hybridized with the radiolabeled *Arabidopsis thaliana* UMP synthase cDNA. A 1 kb ladder was used for molecular mass markers.

Figure 3. PCR amplification of the T88668/UMP synthase intergenic region.

Lane 1 is HindIII cut bacteriophage lambda standards. Lane 2 is amplified from genomic DNA isolated from *A. thaliana* cv. Columbia plants. Lane 3 is from the lambda clone fAt-UMPS1. The oligonucleotides used for this amplification were T88668-oligo F (5'-GAGACAAACCCTTTGGATCAT-3') and UMPS-oligo 6 (5'-GTCTTATTTTGGTGTTC-3').

Figure 4. Phylogenetic analysis of the UMP synthases present in the GenBank.

The amino acid sequences were compared using the Wisconsin GCG tool, "pile-up".

The sequences used for this study were Ng-UMPS, *Naegleria gruberii* (GenBank Accession # L08073); Ce-UMPS, *Caenorhabditis elegans* (Z29443); Os-UMPS1,

Oryza sativa UMPS1 (AF210323); *Os*-UMPS2, *Oryza sativa* UMPS2 (AF210325); *Zm*-UMPS1, *Zea mays* UMPS1 (AF277454); *At*-UMPS, *Arabidopsis thaliana* (X71842); *Nt*-UMPS, *Nicotiana tabacum* (U22260); *Np*-UMPS, *Nicotiana plumbaginifolia* (AF277455); *Bt*-UMPS, *Bos taurus* (X65125); *Mm*-UMPS, *Mus musculus* (P13439); *Hs*-UMPS, *Homo sapiens* (NM_000373); *Dm*-UMPS, *Drosophila melanogaster* (L00968); *Dd*-UMPS, *Dictyostelium discoideum* (X07560).

Figure 5. RT-PCR analysis of UMP synthase and SKIP5 genes from normal or thymine starved (20 mg/ml, 3 days) plants. Replicates show the reproducibility of the assay.

Table I. Oligonucleotides used in the RT-PCR assays

Identity	Sequence	n	T _m ^a
UMPS-1	5'-GATCCTGAGAGATGGCTGAG-3'	20	62°C
UMPS-2	5'-CAACGAGACATGAGTCTTAAAA-3'	22	60°C
T88668-1	5'-GAGACAAACCCTTTGGATCAT-3'	21	60°C
T88668-2	5'-GAAGATGGATCATTTCCCATC-3'	21	60°C
18S-1	5'-AACTTACCAGGTCCAGACATA-3'	21	60°C
18S-2	5'-TAGGAGCGACGGGCGGTG-3'	18	62°C
MCO-18S-1	5'-AACTTACCAGGTCCAGACATA-] ^b	21	60°C
MCO-18S-2	5'-TAGGAGCGACGGGCGGTG-] ^b	18	62°C

$$^aT_m = 2(A+T) + 4(G+C) \text{ } ^\circ\text{C}$$

^b3' end of the MCO oligonucleotides is not free. The oligonucleotides are modified at the 3' end and cannot elongate DNA.

Table II. Expression of *SKIP5* and UMP synthase in response to FOA

FOA		<i>SKIP5</i>			UMP synthase		
		<i>SKIP5</i>			<u>UMPS</u>		
concentration	day	18S	Average ^a	ratio ^b	18S	Average ^a	ratio ^b
None	7	0.12			0.42		
		0.11			0.44		
		0.12	0.12 ± 0.00	1.00	0.45	0.44 ± 0.02	1.00
20 ug/ml	3	0.41			0.84		
		0.77			1.07		
		0.68	0.62 ± 0.19	5.16	0.89	0.93 ± 0.12	2.11
20 ug/ml	5	1.68			1.87		
		0.93			1.25		
		1.23	1.28 ± 0.38	10.66	1.54	1.56 ± 0.31	3.55
20 ug/ml	7	0.56			1.09		
		0.53			0.94		
		0.53	0.54 ± 0.02	4.50	1.03	1.02 ± 0.08	2.32

^aAverage +/- standard deviation^bLevels of mRNAs were normalized to those level found in plants incubated in the absence of FOA.

Table III. Tissue specific expression of *SKIP5* and UMP synthase

Tissue	<i>SKIP5</i>			UMP synthase		
	<i>SKIP5</i>			<u>UMPS</u>		
	18S	Average ^a	ratio ^b	18S	Average ^a	ratio ^b
Leaf	1.04			0.32		
	1.08	1.06 ± 0.02	1.00	0.27	0.29 ± 0.03	1.00
	1.08			0.28		
Stem	0.54			0.24		
	0.51	0.54 ± 0.04	0.51	0.25	0.25 ± 0.01	0.87
	0.58			0.26		
Root	0.04			0.04		
	0.03	0.03 ± 0.00	0.03	0.04	0.04 ± 0.00	0.13
	0.04			0.04		
Flower	1.04			0.52		
	1.07	0.83 ± 0.38	0.78	0.52	0.43 ± 0.15	1.50
	0.40			0.26		
Silique	0.44			0.20		
	0.41	0.41 ± 0.02	0.39	0.18	0.19 ± 0.01	0.65
	0.39			0.18		

^aAverage +/- standard deviation^bLevels of mRNAs were normalized to the level of mRNA found in leaves.

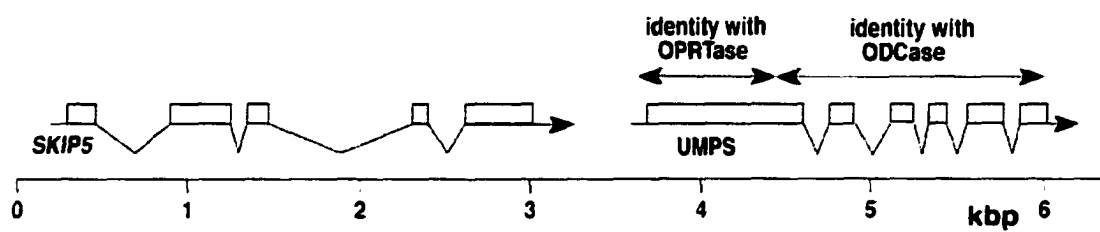


Figure 1.

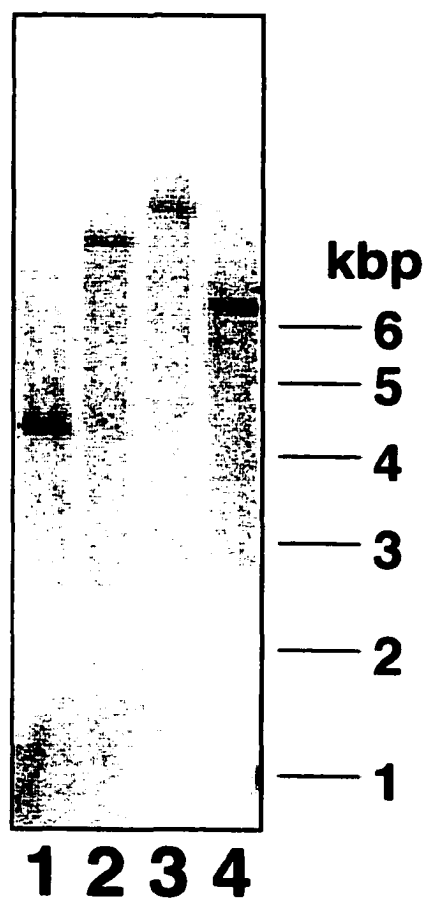


Figure 2.

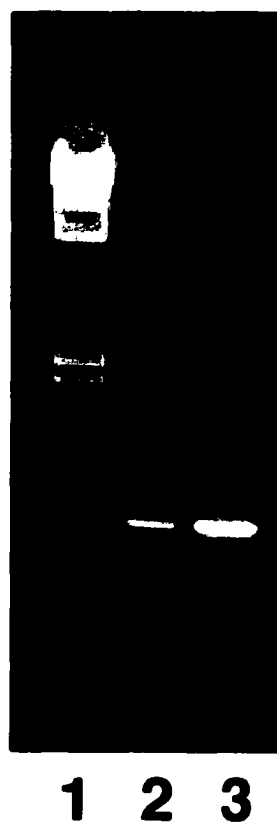


Figure 3.

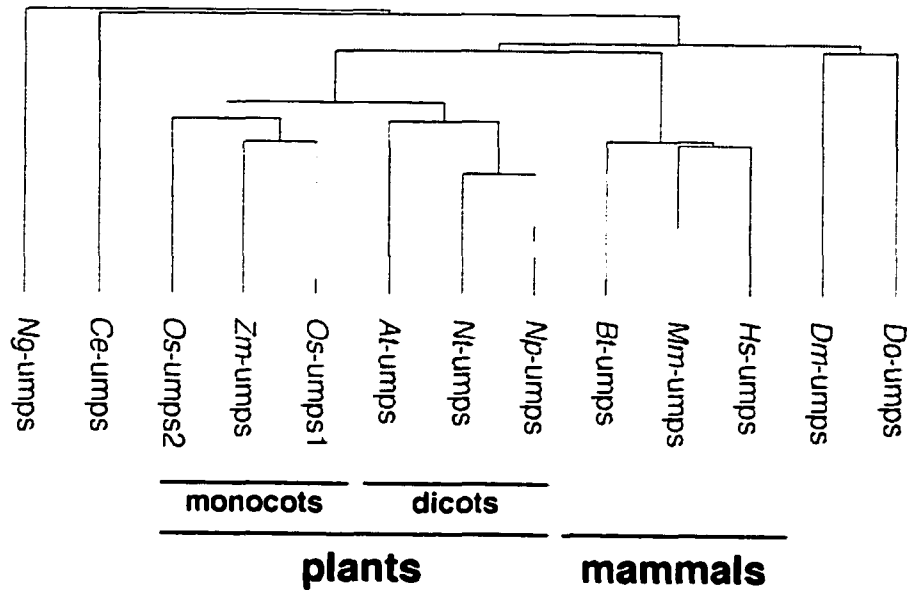


Figure 4.

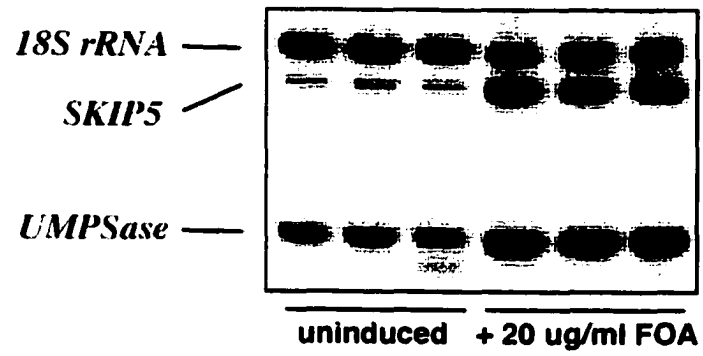


Figure 5.

CHAPTER 4. ANALYSIS OF THE *DE NOVO* PYRIMIDINE BIOSYNTHETIC PATHWAY IN *ARABIDOPSIS THALIANA*

A manuscript to be submitted to The Plant Cell

Chris Kafer & Robert Thornburg

Abstract

We have developed a single-tube relative-quantitative, multiplex RT-PCR method to simultaneously monitor the transcript levels, relative to 18S rRNA, of the entire pyrimidine *de novo* biosynthetic pathway. We have used this RT-PCR method to monitor transcript levels in *Arabidopsis* plants treated with the anti-folate Aminopterin, which induces a state of pyrimidine starvation. The levels of all transcripts of the *de novo* pathway are up regulated to various degrees in response to this compound, confirming that *Arabidopsis* responds similarly to tobacco. We have also analyzed the relative transcript level of each of the members of the *de novo* pathway in various tissues of *Arabidopsis* as well as in whole plants under stress. Curiously, plants grown under completely anaerobic conditions (anoxic) show a down regulation of three members of this pathway, yet under a flood condition the pathway appears to be unaffected. Under the abiotic stress of short-term phosphate starvation or salt stress the pathway is relatively unaffected. Under the biotic stress of infection with Oil Rapeseed Mosaic Virus the pathway is transcriptionally up-regulated within 7 days. To further characterize the expression patterns of these genes we have generated lines of transgenic *Arabidopsis* carrying the β -Glucuronidase gene (GUS) under control of each of the promoters of the *de novo* pathway genes. Expression patterns correlate well with the RT-PCR results and indicate that the majority of pyrimidine biosynthesis is predominantly localized to the vascular tissue and meristematic regions. Two of the *de novo* pathway genes contain rather short promoters and show the

weakest GUS expression in all lines generated. However, when coding regions of the UMP synthase gene are included in the construct, GUS expression is raised considerably.

Introduction

Pyrimidine biosynthesis is a core metabolic process which produces uridine, cytidine and thymidine nucleotides. These nucleotides are required not only for RNA and DNA synthesis but they are also required as co-factors for phospholipid, polysaccharide, and glycoprotein biosynthesis. Thus, pyrimidine nucleotides play a central and critical role in many diverse aspects of cellular metabolism. Because of this central role in metabolism we are interested in the molecular mechanisms that control the expression of the *de novo* pyrimidine biosynthetic pathway genes, as well as the reutilization and degradative pathway genes. We are especially interested in how these pathways are coordinately regulated in higher plants to precisely control intracellular pyrimidine pools.

We have previously shown that regulation of this pathway is at the level of transcription and that UMP synthase is the rate limiting step (Santoso and Thornburg, 1995; Santoso and Thornburg, 1998; Santoso and Thornburg, 2000; Kafer and Thornburg, 2002). In this study we report an analysis of the pyrimidine biosynthetic pathway in plants under normal and in plants undergoing abiotic and biotic stresses. We also report the analysis of transgenic plants containing promoter fragments of the genes of the entire *de novo* pathway and the respective expression profiles of these promoters driving GUS expression.

The *de novo* biosynthetic pathway is highly conserved across widely divergent species. The pathway produces pyrimidine nucleotides via five enzymatic steps. The first committed step of the pathway is Aspartate Transcarbamoylase (ATCase) which condenses aspartate with

carbamoylphosphate. Dihydroorotase (DHOase) carries out the cyclization of carbamoyl aspartate to yield the pyrimidine ring. ATCase and DHOase are both chloroplastic enzymes encoded in the cell nucleus. Dihydroorotate dehydrogenase (DHODH), found in the mitochondria, oxidizes the ring to produce the double bond between carbon five and six, while concomitantly reducing ubiquinone. In this regard, pyrimidine biosynthesis is tied to respiration. The final two steps are carried out by the enzyme UMP synthase (UMPs) which is also a nuclear encoded chloroplastic enzyme. This enzyme transfers the pyrimidine ring to phosphoribosyl pyrophosphate (PRPP) and finally decarboxylates the ring to yield uridine 5' monophosphate (UMP). The resulting UMP is the precursor to all other pyrimidine nucleotides within the cell.

In vertebrates and insects the first 3 enzymatic steps are carried out by a single protein named CAD and the final two steps are carried out by the bifunctional protein UMP synthase. In plants the first three enzymatic steps are carried out by separate proteins, however, plants have acquired the bifunctional UMP synthase. In *Arabidopsis* each enzyme of the *de novo* pathway is encoded in the genome by a single copy gene; however, in other species these genes may be part of small multigene families (Williamson and Slocum, 1993; Williamson and Slocum, 1994).

Results

Pyrimidine starvation

We have previously demonstrated in tobacco callus cultures that pyrimidine starvation causes a two to three fold up-regulation of transcription of the rate limiting step in pyrimidine biosynthesis, UMP synthase (Santoso, 1995; Santoso and Thornburg, 1998). Aminopterin

(AMT) inhibits thymidine biosynthesis by blocking regeneration of tetrahydrofolate from dihydrofolate, thereby causing pyrimidine starvation. To investigate whether other members of the pathway are also regulated by pyrimidine starvation, we grew plants in the presence of various concentrations of AMT and evaluated the expression of the *de novo* pyrimidine biosynthetic pathway using a novel RT-PCR assay. Seven days after AMT addition, RNA was isolated from whole plants and the multiplex RT-PCR was performed. Panel A of Figure 1 shows the phosphorimager scan of the radiolabeled RT-PCR products separated on an 8% TAE polyacrylamide gel. The triplicate PCR shows very good reproducibility. As can be observed, the amount of initial transcript relative to the 18S transcript increases as a function of increasing AMT concentration (ie an increasing state of pyrimidine starvation). We were particularly interested in UMP synthase because this enzyme is the rate limiting step in pyrimidine biosynthesis (Santoso, 1995). The UMP synthase transcript shows a three to five-fold increase in transcript level under conditions of pyrimidine starvation. These results are in good agreement with previous reports of transcript induction following nucleotide starvation (Santoso and Thornburg, 1998; Kafer and Thornburg, 2002).

As can be seen in Figure 1, not only was UMP synthase up-regulated, but each of the biosynthetic pathway genes was similarly up-regulation. The highest level of induction appears to be with ATCase which shows more than a 20-fold increase in transcript level relative to 18S under these conditions. DHOase, DHODH, and UMP synthase are up-regulated fifteen-fold, three-fold and five-fold, respectively. Thus, a dramatic and coordinated up-regulation of the entire biosynthetic pathway occurs under pyrimidine nucleotide starvation in *in vitro* grown plants. Further, a dose dependent correlation is observed implying that more stringent starvation

leads to stronger gene up-regulation. Thus, we conclude that plants have a molecular mechanism to recognize cellular pyrimidine levels and to modify expression of the *de novo* biosynthetic pathway accordingly. Mutant tobacco cell lines which have lost this pyrimidine sensing ability have been previously isolated (Santoso, 1991; Santoso and Thornburg, 1992; Santoso, 1995).

Organ specific expression

The finding that this multiplex RT-PCR assay showed similar up-regulation of UMP synthase and the other pyrimidine biosynthetic genes in whole *Arabidopsis* plants as was previously observed for UMP synthase in tobacco, confirmed the methodology of our assay method. We next used our RT-PCR method to quantitate transcript levels in various organs and treatments of whole *Arabidopsis* plants. The multiplex RT-PCR was performed in triplicate on RNA from roots, leaves, stems, flowers and siliques. Comparison of the level of each amplicon to the level of 18S amplicon again allows us to evaluate the levels of initial transcript found in different organs. Figure 3 shows the relative transcript level in each organ.

From this analysis we found some striking differences in transcript level between the various organs tested. Most notably the fact that roots contain over 2-fold higher levels of ATCase transcript than are found in leaves but the levels of DHOase, DHODH and UMP synthase in roots are quite low. With the exception of roots, all other tissues show basically the same pattern of expression throughout the plant organs. Siliques contained less than half the level of each transcript relative to that found in leaves. Stems also show lower transcript levels compared to that of leaves while flowers show similar levels of each transcript relative to leaves.

Germination

It has previously been shown that in black gram (*Phaseolus mungo*) and wheat the *de novo* biosynthetic pathway is activated very early in the germination process (Mazus and Buchowicz, 1972; Ashihara, 1977). Salvage of stored nucleotides is thought to be the predominant source of pyrimidines early in germination but within 24 hours of sowing the *de novo* enzymes can be detected in these seeds (Ashihara, 1977). To determine when during germination the *de novo* pathway of pyrimidine biosynthesis is activated we have used our RT-PCR method to monitor transcript levels during germination of *Arabidopsis* seeds over a five day period in 24 hour increments.

Figure 4 shows that transcripts for each of the *de novo* enzymes can be detected within 24 hours of germination. Again, by comparing the intensity of each *de novo* pyrimidine gene amplicon to the 18S amplicon we can make conclusions about the initial relative levels of each transcript. As can be seen in Figure 4, ATCase is up-regulated more than eight fold by day three. However, the other *de novo* biosynthetic transcripts slowly decline over the next several days and they maintain a relatively low level of expression thereafter.

Oxygen tension

The *de novo* pyrimidine biosynthetic pathway is linked to respiration by the inner mitochondrial-bound dihydroorotate dehydrogenase (DHODH). This redox enzyme uses ubiquinone as the electron acceptor. Therefore, it is expected that any process that disrupts electron transport would affect pyrimidine biosynthesis. To test the affect of oxygen tension on the pyrimidine biosynthesis we have subjected *Arabidopsis* plants to two different anoxic scenarios. In the first condition we transferred seven day old plants to a completely anaerobic condition. For the second condition, plants were submerged for up to 24 hours in water. RNA

was extracted from these plants at the indicated time points and used in our RT-PCR assay to monitor transcript levels of the pyrimidine biosynthetic genes. The results of these experiments are shown in Figures 5 and 6.

Figure 5 shows that under extreme anaerobic conditions ATCase expression is relatively unchanged. However, there is a severe reduction in the levels of DHOase, DHODH and UMP synthase transcripts. Figure 6 shows the results of a less severe flood stress. By comparing relative levels of each amplicon to the 18S amplicon it appears that the pathway is relatively unaffected until 24 hours of being submerged. After 24 hours the transcript of the rate-limiting UMP synthase gene shows an approximately two fold up-regulation. These two figures show that a difference in severity of oxygen tension has significantly different affects on pyrimidine biosynthesis. A totally anaerobic environment disrupts the plants ability to produce pyrimidines while a flood situation for 24 hours does not.

Abiotic stress

Phosphate is required for normal plant growth and is required for nucleoside phosphate production. Therefore, a limitation of this nutrient would be expected to alter biosynthesis of pyrimidines. Phosphate starvation can be induced in plants by growth in the presence of mannose (Edwards and Walker, 1983; Ciereszko et al., 2001). We have used two different strategies to induce a relatively short term phosphate deficiency in whole *Arabidopsis* plants. We have grown the plants in the presence of 10 or 20 mM mannose or in media completely lacking in phosphate. Figures 7 and 8 show the results of these experiments.

RNA was extracted from plants germinated and grown in the presence of the indicated concentrations of mannose. We then analyzed the transcript level of the pyrimidine

biosynthetic pathway found in these plants relative to 18S. As can be seen in Figure 7, mannose induced phosphate starvation resulted in little consistent change in the levels of these transcripts except for UMP synthase which under conditions of 20 mM mannose for 5 days fell to 25% of untreated levels.

Phosphate starvation was also induced by transferring seven day old plants grown in MS media to fresh MS media which lacked any phosphate. These plants were allowed to grow for another five to seven days. As before, RNA was extracted from these plants and used in our RT-PCR method. Figure 8 shows the results of this experiment. There appears to be no alteration in the relative transcript levels of the *de novo* pathway genes between treated and untreated plants. Therefore, we conclude that although phosphate starvation clearly has strong effects on plant growth and development, these effects do not appear to be mediated through effects on the pyrimidine biosynthetic pathway.

Other stresses that limit plant growth would be expected to alter core metabolic processes such as pyrimidine biosynthesis. A cessation of growth would presumably lower the requirements for pyrimidine biosynthesis. A high level of sodium chloride in the soil is a stress that land plants must cope with in many areas of the world (Hasegawa et al., 2000). To test the affect of NaCl on *in vitro* grown *Arabidopsis* plants, RNA was extracted from plants grown in the presence or absence of 150mM NaCl. As seen in Figure 9 there is no discernible significant alteration in pyrimidine biosynthesis.

Biotic stress

Virus replication within a plant cell theoretically requires additional pyrimidines for viral nucleic acid synthesis above that required for the normal functioning of the plant cell. If virus

infection leads to a depletion, or even a perturbation of nucleotide pools, it is likely that the plant cell would recognize this change in nucleotide levels and modify expression of the biosynthetic pathways accordingly. To evaluate whether virus infected plants show changes in *de novo* pyrimidine transcripts, we have analyzed soil grown *Arabidopsis* inoculated with Oil Rapeseed Mosaic Virus (ORMV). RNA was extracted from plants seven days after infection with ORMV and analyzed by RT-PCR. Figure 9 shows that there is a low level of induction (~70%) for ATCase and DHOase. However, DHODH and UMP synthase show a two fold and 4.5 fold increase respectively, when compared to mock inoculated plants.

Transgenic plants

To better understand the patterns of gene expression for the *de novo* pyrimidine biosynthetic pathway we prepared transgene constructs to express β -glucuronidase from each of the promoters of the *de novo* pathway. Figure 10 shows diagrams of each of the *de novo* biosynthetic genes and promoter fragment that was tested in these studies. Using these constructs we have generated transgenic *Arabidopsis* plants.

ATCase

Because the ATCase gene contains an intron in the 5' UTR, two constructs were made for this gene, with and without the intron. The TR498 transformed plants lacked the intron in the 5' UTR of the ATCase gene. When transgenic TR498 plants were stained for GUS activity we observed relatively low but significant level of GUS activity present in the foliage of 6 day old plants (Figure 11A). However, very little GUS staining was observed in the roots of the TR498 plants. As the plants developed through the 10 day and 17 day stages, similar patterns of staining were observed. Again, little GUS activity was observed in the roots. The expression

of GUS in the foliage was primarily restricted to the vascular tissue but low background levels were observed throughout the leaf blade (Figure 11E). GUS activity was apparent in flowers only as a very light background stain and at the base of siliques (Figure 11F). The staining observed in Figure 11 is representative of each of the 10 lines of Tr498 and Tr499 plants prepared in this study.

The TR499 plants contain the intron from the 5' UTR of the ATCase gene and had a mostly similar pattern of expression (Figure 11 G-L), but in addition these plants also show a significant level of expression in roots (Figure 11J). GUS activity was evident in floral organs as a light background stain (Figure 11L) and at the base of siliques as in the TR498 plants. Thus the presence of the intron in the 5'UTR is necessary for significant expression of ATCase within roots.

DHOase

The DHOase gene also contains an intron in the 5' UTR. Therefore we prepared two constructs with the promoter of the DHOase gene, with and without the intron. TR494 plants lack the intron in the 5' UTR of this gene. When transgenic plants were stained for GUS activity we observed a quite low level of activity in the foliage of six day old plants (Figure 12A). As the plants matured through day 10 and 17 the pattern of expression was unchanged (Figure 12 B&C). No GUS activity was observed within the roots of these TR494 plants at any time (Figure 12 A-C). The minor amount of GUS activity that we observed the foliage was localized to the vascular tissue and meristematic regions (Figure 12 A-F). Only a small amount of staining was observed in flowers and developing siliques of TR494 plants (Figure 12 D). The staining observed in Figure 12 is representative of each of the 10 lines of Tr494 and Tr495

plants prepared in this study.

TR495 plants contain the promoter region of the DHOase gene and 5' UTR intron. When the plants were stained for GUS activity we observed a significant level of GUS expression in the foliage of 6 day old plants (Figure 12 E). This pattern of expression was consistent through day 10 and 17 (Figure 12 F&G). The majority of GUS activity in the foliage was localized to the vascular tissue. We also observed a light overall background staining throughout the leaf parenchyma. In addition, a moderate level of GUS activity was also observed in flowers and siliques of TR495 plants (Figure 12 H).

DHODH

Tr510 plants contain the bi-functional promoter of dihydroorotate dehydrogenase (DHODH). A construct containing the entire region between the upstream SOD (found on the opposite strand of DHODH) driving GUS was made in the DHODH promoter orientation to assess expression of this gene. When transgenic TR510 plants were stained for GUS activity we observed a quite low level of GUS expression in the foliage of six day old plants (Figure 13a). As the plant developed through day 10 and day 17 the expression level of GUS decreased markedly (Figure 13 B&C). The expression of GUS in the foliage was primarily restricted to the vascular tissue. No GUS expression was detected within the roots of TR510 plants. Flowers and developing siliques show a relatively high level of GUS gene expression (Figure 13 D).

UMP synthase

The final set of transgenic plants that we have generated contain 3 different promoter-GUS fusions. The UMP synthase gene contains a relatively short promoter region at 405 bp

between the coding region of UMP synthase and the 3' end of SKIP5 (Kafer and Thornburg, 2002). The TR415 line contains the short proximal promoter fused to GUS (See Figure 10). When TR415 plants were stained for GUS activity we observed no GUS expression at day six, 10 or 17, no expression was detected within the flowers, roots or developing siliques of these plants (Figure 14 A-F).

TR446 contains the 405 bp proximal promoter and approximately one half of the upstream SKIP5 gene found in the same orientation as UMP synthase (See Figure 10). When TR446 plants were stained for GUS activity on day 6, we were able to detect a low level of activity in the apical meristem (Figure 14 G). This pattern of expression was unchanged through day 10 and 17 plants (Figure 14 H-I). A low level of GUS expression was also detectable in floral organs (Figure 14 J). No GUS activity was detectable in roots. When GUS activity was quantitated in 10 day old plants, it was below the detection level of our spectrophotometric assay.

TR457 lines of plants contain the same upstream region as TR446 plants with the addition of approximately one-third of the coding region of the UMP synthase gene (through codon 177). When TR457 plants were stained for GUS activity we observed a significant level of GUS expression in the leaves of six day old plants (Figure 14 K). Similar expression patterns were observed on day 10 and 17 plants (Figure 14 K-M). The GUS expression in foliage was present throughout the leaf blade and not restricted to the vascular tissue. A significant level of GUS activity was observed in the root tips of TR457 plants (Figure 14N) GUS activity was also detected within floral organs as a light overall staining (Figure 14O).

Quantitation of GUS activity

GUS activity was quantitated in two to three lines of transgenic plants. Each assay contained total protein from 15 to 20 plants. The assay was performed as in Materials and Methods. The results of this analysis are shown in Figure 16 as a histogram. The relative strength of the promoters can be assessed from this figure. TR495 plants had over 100 fold higher GUS activity than the TR494 plants. TR499 plants have approximately twice the activity of GUS as TR498 plants. The promoters of DHODH and UMP synthase showed similar levels of GUS gene expression.

Discussion

Here we report the development and use of a quantitative RT-PCR assay to evaluate the expression patterns of all four genes of the *de novo* biosynthetic pathway relative to 18S ribosomal RNA. Quantitative RT-PCR has become a widely accepted way to measure the level of specific transcripts, provided internal controls are used. The ratios of gene specific to control PCR product can be compared across samples and treatments and conclusions about initial levels of transcript can be drawn. We have previously developed a similar single tube multiplex, relative quantitative RT-PCR assay to evaluate the expression pattern of UMP synthase and its closely flanking neighbor SKIP5. (Kafer et al., 2002)

Pyrimidine starvation

We first used our RT-PCR method to test the anti-folate Aminopterin and its ability to cause pyrimidine starvation in whole *in vitro* grown plants as it does in callus cultures of tobacco. This compound inhibits regeneration of tetrahydrofolate from dihydrofolate and specifically causes a state of thymidine starvation. Under pyrimidine starvation conditions we

found that the entire pathway is transcriptionally up-regulation in a dose dependent manner after treatment of whole plants with Aminopterin. This confirms our previous findings in tobacco and supports the use of our relative quantitative RT-PCR method to monitor *de novo* pyrimidine biosynthesis. It also lends support to the argument that plants have a molecular mechanism to detect intracellular pyrimidine pool levels (Kafer and Thornburg, 1999) and to modify gene expression in a manner that correlates with severity of pyrimidine starvation.

Abiotic stress and pyrimidine biosynthesis

We expected that conditions that reduce plant growth would reduce the requirement for pyrimidine nucleotides. We found that treatment with 150mM NaCl for several days did indeed slow growth *in vitro* as expected, but did not result in any discernible down regulation of the biosynthetic pathway. While this was unexpected it can be explained by the fact that NaCl is not directly related to pyrimidine biosynthesis. Pyrimidine biosynthesis would be affected only secondarily, if at all. It is also possible that under longer periods of salinity stress the pathway would be down-regulated to a noticeable extent.

Phosphate starvation also had no noticeable affect on pyrimidine biosynthesis. In contrast to saline stress, the availability of phosphate is directly related nucleotide biosynthesis, therefore this result was somewhat unexpected. The high energy molecule phosphoribosyl pyrophosphate (PRPP) is the first intermediate in which phosphate is utilized in pyrimidine biosynthesis by the enzyme UMP synthase. A limitation of phosphate would ostensibly lead to a limitation in PRPP and hence a limitation in UMP biosynthesis.

We created a state of phosphate starvation by two different methods using either mannose at varying concentrations or by growing *Arabidopsis* in completely phosphate free

media for five to seven days. Neither of these two treatments led to a discernible alteration in pyrimidine biosynthetic gene transcript levels. It is possible that a more stringent or a longer term phosphate starvation would significantly alter pyrimidine biosynthesis. Other authors have used periods of up to three weeks on phosphate free media to monitor gene regulation (Ciereszko et al., 2001). However, in rapidly growing seven day old plants presumably requiring a continual supply of pyrimidines, it would be expected that short time periods lacking pyrimidine biosynthetic intermediates would affect nucleotide pools and hence expression of *de novo* genes. In these studies we used 2.5 to 5 fold higher concentrations of mannose than in previous reports to create a short term depletion of intracellular inorganic phosphate (Ciereszko et al., 2001). Further, the concentration of mannose in our experiments was 10 to 20 fold higher than the concentration of phosphate found in MS media. It is possible that under conditions of limiting pyrimidine intermediates that salvage and recycling of preformed bases and not *de novo* synthesis would become the predominant source of nucleotides, which would be sufficient to meet the short term needs of the plant. Further experiments using a “metabolomics” approach could be used to verify a change, if any, in nucleotide pools and their intermediates under the conditions of phosphate starvation we have used.

Another stress that plants frequently undergo in many areas of the world is flooding. Pyrimidine biosynthesis is inextricably linked to respiration at the enzyme dihydroorotate dehydrogenase. DHODH is located in the inner mitochondria and is responsible for production of orotate from dihydroorotate. The electrons from orotate are concomitantly used to reduce ubiquinone. Therefore, a functioning respiratory chain is required for pyrimidine biosynthesis. Other workers have argued that this link between pyrimidines and respiration is a molecular

switch that controls cellular proliferation of mammalian cells under hypoxic conditions. (Löffler, 1992; Löffler, 1994)

We have induced a simulated flood condition in our plants by two different methods. One is quite severe and the other is a much closer approximation to an actual flood situation. We have treated *in vitro* growing plants as described in Materials and Methods either by continuously purging the flask with nitrogen gas for 24 hours (severe anoxia) or by submerging plants in distilled water for up to 24 hours. The results of these different treatments were quite dissimilar. Under completely anaerobic conditions the final three enzymes were transcriptionally down-regulated. Under conditions of a less severe "flood" in which the plants were simply submerged for up to 24 hours there appears to be only limited change in the *de novo* biosynthetic pathway. The only gene that was up-regulation to any extent was UMP synthase by approximately two fold.

Perhaps the results can best be explained by the severity of the treatment. Completely anaerobic conditions are not an environment for which plants would have an adaptation and would not face. Completely anaerobic conditions would lead to a complete shut down of metabolism and death would be expected relatively quickly.

On the other hand, the plant would be expected to survive under a short term decrease of oxygen such as from being submerged. Further, under a relatively short duration of a flood situation that a plant is able to survive, it is likely that a minimal alteration of nucleotide pools would occur and therefore require minimal modulation of pyrimidine biosynthesis. Nevertheless, the rate limiting step, UMP synthase, does appear to be up-regulated by slightly more than two fold. If this stress is in fact causing a decrease in nucleotide pools by the inability

of DHODH to function, this relatively small increase in UMP synthase is in all likelihood the resulting attempt by the plant to produce more UMP.

Viral infection and pyrimidine biosynthesis

We have also analyzed the *de novo* biosynthetic pathway and whether it is significantly altered by biotic stress as caused by Oil Rapeseed Mosaic Virus. In cells undergoing viral infection it is likely that the intracellular pool of nucleotides would be depleted by the viral replication. To evaluate whether viral infection results in an up-regulation of the *de novo* pathway we compared mock infected plants to ORMV infected plants. We tested this using our RT-PCR method on plants that had been infected with ORMV or mock infected with distilled water. Seven days after treatment the rate limiting UMP synthase is up-regulated by more than four fold. DHODH is also marginally up-regulation by approximately two fold. These results are in good agreement with recent data from experiments using the Affimetrix "gene chip" by Dr. Steve Whitham. After five days of ORMV infection, gene chip analysis shows that UMP synthase and DHODH are both up-regulated approximately two fold (Dr. Steve Whitham, personal communication). ATCase and DHOase are not present on this chip.

Transgenic plant analysis

In order to further dissect the expression patterns of these genes we have generated lines of transgenic *Arabidopsis* which carry the GUS reporter gene under control of the individual promoter regions of the genes of the *de novo* pyrimidine biosynthetic pathway.

ATCase

The first set of promoter GUS fusions consists of the full length ATCase promoter region and 5'UTR. The 5' UTR of the ATCase gene contains a single intron, so two constructs

were made to assess whether the intron had any function. The results of these experiments were unexpected. The lines of plants that contain the full length promoter including the 5' intron (TR499), showed high levels of GUS staining in all organs of the plant, which is consistent with high levels of ATCase transcript observed with the RT-PCR assay. The highest level of GUS activity was observed in the meristematic regions and within the vascular tissues during development. The lines of plants that contain the promoter region but lack the intron from the 5' UTR (TR498) show significant levels of GUS expression in the aerial portions of the plant. However, these lines which lack the intron showed virtually no staining within the roots. (compare Figure 11 panels A,D,J and K). These results suggest that the ATCase intron 1 contains regulatory sequences that function in root specific regulation. Regulatory regions within introns that control tissue and developmental expression have been previously described (Callis et al., 1987). Also, Jeon et al. (2000) have described an intron in the rice α -tubulin A gene which shows similar expression patterns to that which is observed with the ATCase intron 1, in these studies. In these studies, GUS expression in root tissue is dependant on the first intron in the α -tubulin gene, which is found within the tubulin coding region. It has also been shown that introns can be required to achieve maximal levels of transgene expression and can also act as regulatory regions (Maas et al., 1991; Lou et al., 1993; Kato et al., 1998; Deyholos and Sieburth, 2000; Morello et al., 2002; Mun et al., 2002).

DHOase

The DHOase gene also contains an intron in the 5'UTR. To test whether this intron also affects transgene expression we generated lines of *Arabidopsis* carrying transgene constructs either with or without the 5' UTR intron. The results were strikingly similar to those found with

the ATCase transgenic lines. Transgenic lines containing constructs with the 5' intron (TR495) show relatively high levels of GUS expression throughout the plant. Transgenic lines containing constructs which lack the 5' the intron show no staining within any root tissue (compare Figure 12 panels A, D, G and J). Further, the staining in aerial portions of the plant are much reduced and almost completely lacking in mature plants. Like the TR498 and TR499 plants, the majority of the GUS activity was localized to the meristematic regions and vascular tissues within the plant. Thus, we conclude that the intron is a required element for native DHOase expression.

The observation that the introns of the DHOase and ATCase gene augment GUS expression to various degrees led us to analyze the sequences of these introns using the pileup feature of GCG (GCG, Version 9.0). The results of this analysis reveal that there are several highly conserved regions in these sequences. Overall the introns share 59% identity. This degree of identity is remarkable for genes that are functionally unrelated, except to share a common pathway. Presumably, the conserved regions of these introns play an integral role in the root specific expression pattern as well as, however the mechanism that connects these sequences to gene expression is unknown. Further experiments with promoter deletions should be conducted to further pinpoint the precise elements responsible for this root specific expression. It would also be informative to fuse these introns to other promoters (eg CaMV 35S) to assess the affect or alteration on known expression patterns.

DHODH

We have generated lines of transgenic *Arabidopsis* (TR510) that carry the GUS reporter gene under control of the DHODH promoter. This is a quite short bifunctional promoter with

only 380 nucleotides separating the DHODH gene from a superoxide dismutase found on the opposite strand (Figure 10). The identification of EST clones from the SOD gene reveals that the promoter is bi-functionally active. Of the four transgenic lines generated with the pRT510 construct, only two show detectable GUS gene expression. Like the ATCase and DHOase lines the majority of expression is found in the vasculature and rapidly dividing tissues such as meristematic regions. This promoter is quite weak and light staining is observed. This is in agreement with observations from the Affimetrix "gene chip" which shows basal expression of DHODH to be very low (Dr. Steve Whitham, personal communication). This promoter contains no identifiable TATA box and perhaps this expression level is a result. We cannot rule out the possibility that flanking sequences could influence expression of either or both DHODH or the SOD as was observed for the UMP synthase constructs (see below). The sequence of this promoter also has an internal repeat that is predicted to fold into an unusual cruciform structure (Figure 16). Whether this structure functions in the transcriptional activation of this gene, or even forms *in vivo*, is unknown. However, similar cruciform structures have been postulated to form in transposable elements such as the Human Alu element, within promoters and in yeast (Hambor et al., 1993; Sidorenko et al., 2000; Callejo et al., 2002).

UMP synthase

Finally, we have generated transgenic tobacco plants carrying the GUS reporter gene driven by the proximal promoter region of the *Arabidopsis* UMP synthase gene. The distance between UMP synthase and the upstream SKIP5 gene is also quite short (Figure 10) at only 405 nucleotides (Kafer and Thornburg, 2002). To evaluate whether this short promoter contained the complete UMP synthase regulatory sequences, the 405 bp promoter fragment was used to

prepare the transgene construct pRT415 (Figure 10). *Arabidopsis* lines containing this construct (Tr415) show little if any GUS activity. Further, this 405 nucleotide fragment was unable to drive GUS expression in any of the 10 transgenic tobacco lines prepared using this construct (data not shown). Because the core promoter seemed to lack GUS expression, we generated two additional promoter GUS fusions, pRT446 and pRT457 (Figure 10) to assess the possibility that regulatory sequences might be contained within the SKIP5 coding region or within the UMP synthase coding region itself.

While our lines of TR415 transgenic *Arabidopsis* plants showed no GUS expression the transgenic plants carrying a portion of the SKIP5 coding region (TR446) show low but significant staining for GUS expression localized to the meristematic and vascular regions. This pattern of expression is similar to the pattern seen in the plants carrying the ATCase and DHOase promoter-GUS fusions, Tr495 and Tr499. The highest level of staining is observed in the apical meristematic region and leaf primordia. This is not unexpected as rapidly dividing tissues would be expected to have a higher requirement for pyrimidines.

Our final line of transgenic plants, TR457, carries the SKIP5 coding region found in pRT446 plus approximately one third of the coding region of UMP synthase gene. The expression pattern in the TR457 plants is similar to the expression pattern seen in the Tr446 plants but the level of expression is significantly higher. GUS expression is localized to rapidly developing tissues, meristematic regions and the vascular tissue in these lines as well. We conclude that there are regulatory sequences within the SKIP5 gene and UMP synthase gene itself and that the 405 bp proximal promoter region requires these elements to express the native UMP synthase protein.

Summary

We have used two separate and complementary methods to analyze the expression pattern of the *de novo* pyrimidine biosynthetic pathway. In the first method we have quantitated relative transcript levels of each member of the pathway relative to the 18SrRNA transcript. It must be pointed out that this method is not useful for comparing levels of each transcript relative to each other. Differences in the efficiency of the reverse transcription of each gene of interest and differences in amplification efficiency during PCR are variables which can make drawing conclusions about relative levels of each transcript to one another difficult to make. Therefore, the RT-PCR method we have described provides information relating the relative levels of the *individual* genes between samples and/or treatments.

To assess the tissue specificity and the relative expression level of each gene in the pathway to one another we generated transgenic lines of *Arabidopsis* plants carrying the GUS gene under control of the individual promoters of the *de novo* pathway genes. This allows us to draw conclusions about the expression level of each gene in the pathway relative to one another. GUS activity was quantitated from several lines of transgenic plants derived from each construct. From this data we conclude that the ATCase and DHOase promoters are capable of driving GUS expression at up to 80 fold higher levels than the DHODH and UMP synthase promoters (Figure 17). However, caution is necessary when making conclusions about the relative amounts of *native protein* (ie ATCase, DHOase etc.) produced by the respective promoters because nothing is known about the degradative turnover rate of the *de novo* enzymes.

Using the transgenic plants we have also been able to localize GUS, and presumably the

native gene expression, to specific areas of plant organs. The majority of GUS expression in all transgenic lines, except those with truncated 5' UTR regions, was localized to the vascular tissue of roots, leaves and flowers. GUS activity was also observed in rapidly dividing tissue such as meristematic regions.

Materials and Methods

Plant growth conditions

Approximately 30 *Arabidopsis* seeds were surface sterilized in a 25% bleach solution for 20 minutes and rinsed three times with sterile distilled water. The seeds were transferred to 20 ml of sterile MS media and grown with constant shaking at 100 RPM under 16 h light 8 h darkness for seven to ten days before treatment. In the case of the pyrimidine starvation experiments utilizing aminopterin, the MS media was supplemented with hypoxanthine (0.1 mM Cf) and all 20 amino acids (0.01 mg/ml Cf), since dihydrofolate reductase is also involved in biosynthesis of some amino acids and in purine biosynthesis (Voet and Voet, 1990).

The preceding *in vitro* conditions were altered slightly to create short term phosphate starvation conditions. In the first case, seeds were either sown and allowed to grow for 10 days in MS media containing 10 or 20 mM mannose or allowed to germinate for three days first before addition of mannose (the highest concentration of mannose used severely inhibited germination). In the second case of short term phosphate starvation, plants were allowed to grow in liquid MS media as described above for seven days. The mass of plants was aseptically removed from the flask, rinsed in sterile phosphate-free MS media and briefly blotted on sterile filter paper. The plants were then returned to fresh MS media containing or lacking phosphate for an additional seven days. At this point, RNA was extracted from at least seven to ten whole

plants.

For the Oil rapeseed Mosaic Virus (ORMV) experiments, *Arabidopsis* plants were grown in potting mix under 16 hours of light at 22 degrees C. until maturity (beginning to flower) Plants were inoculated with ORMV in water by high pressure spray with an airbrush. The mock inoculated plants were sprayed in the same way with distilled water. The plants were grown another seven days and tissue was collected. At least 4 leaves from different plants were pooled for RNA isolation.

For analysis of transcript levels in tissues plants were grown in pots in potting mix under 16 hours of light. Tissues were harvested at the appropriate times and stored at -70°C until use. Root tissue was excised from plants grown aseptically in liquid MS media as described above for at least 2 weeks.

For the flood stress experiments *Arabidopsis* seeds were surface sterilized and plated in magenta tissue culture boxes containing MS media +2% glucose and 0.8% agar. Growth conditions were identical to that described above. After seven to ten days of growth pre-flood plants were harvested and the vessel was filled with distilled water and returned to the growth chamber. Whole plants were removed at the indicated time points for RNA isolation.

RNA isolation

RNA was isolated using a modified guanidine thiocyanate method (Chomczynski and Sacchi, 1987). Frozen or freshly harvested tissue (100-300 mg) was ground in 1 ml of extraction buffer (4M guanidine thiocyanate, 25 mM sodium citrate, 0.5% sarkosyl, 0.1 M 2-mercaptoethanol, 0.3 M sodium acetate pH 4.8) and extracted with 0.5 ml of 0.1 M sodium acetate saturated phenol:chloroform (5:2) pH 4.8 by vigorous vortexing in a 1.5 ml microfuge

tube followed by a 3 min centrifugation at 13,000 RPM. The upper aqueous phase was extracted the same way once more followed by a final chloroform extraction. RNA was precipitated by adding an equal volume of 2-propanol to the aqueous phase, mixing by inversion and then room temperature incubation for 5 min followed by a 10 min centrifugation at 16,000 x g. The RNA pellet was washed with 70% ethanol and centrifuged for 1 min. The pellet was dried and resuspended in 90 µl water. Contaminating DNA was removed by incubation with 2 units of RNase-free DNase in the manufacturers buffer (Promega, Madison WI, USA) at 37°C for 45 min followed by a single phenol:chloroform extraction and 2-propanol precipitation as before. The final dried pellet of RNA was resuspended in 35-50 µl of RNase free water. The RNA was quantitated by measuring absorbance at 260 nm on a Varian Cary Bio 50 spectrophotometer. RNA was stored at -70°C until use.

Reverse Transcription

Approximately 2.5 µg of RNA was incubated in a total volume of 15 µl with 2.5 mM random hexanucleotides and 2.5 mM each dNTP at 75°C for 5 min. then cooled to room temperature. Four µl of the manufacturers 5x MMLV buffer and 200 units of MMLV reverse transcriptase (Promega, Madison WI, USA) were added to each tube and mixed gently. The reverse transcription reaction was carried out at 37°C for 1 h and the reaction products were stored at -20°C until use.

Polymerase Chain Reaction

For all analyses a 25 µl or 50 µl "Hot Start" PCR was used (Klebe. et al., 1999). The reaction mixture contained 50 mM KCl, 10 mM Tris pH 9.0, 0.1% Triton X-100, 0.2 mM of each dNTP, 0.5 mM each primer, 1 µl Taq polymerase (Desai and Pfaffle, 1995), 2.5 mM

MgCl₂ and 1 µl of the reverse transcription reaction. The PCR was performed in an MJ Research PTC-100 thermal cycler (MJ Research, Incline Village, NV, USA) with the following cycling parameters: denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 60 s.

The linear range of amplification for each transcript was determined to be between cycles 22 and 30 (data not shown). Therefore, cycle 26 was chosen as the optimal cycle for quantitation in all subsequent experiments. It was also determined empirically that a 5:95 ratio of 18S:18S modified competitive oligonucleotide (MCO) yielded a PCR product approximately equal to the most abundant gene of interest. MCOs are identical in sequence to the native 18S oligonucleotide except the 3' hydroxyl is modified with a C3 spacer group. MCOs were synthesized at Sigma-Genosys (Sigma, ST. Louis, MO, USA). PCR was always done in triplicate on the reverse transcribed RNA to assess reproducibility. [³²P]-labeled dCTP was incorporated in the PCR reaction to label the PCR products for quantitation.

The primers were designed to generate amplicons that differ in length for each gene transcript. (Table 1) The resulting “ladder” facilitates ease of separation and quantitation in a single lane on an acrylamide gel. The amplicons for 18S, ATCase, DHOase, DHODH and UMPs are 450, 280, 223, 171 and 110 nucleotides long, respectively. Each primer was checked in a BLAST search (Altschul et al., 1990) against the *Arabidopsis* genome to confirm that it was a unique sequence and matched only to our gene of interest. To further confirm their suitability we tested the ability of each primer pair to generate a single amplicon at the expected size from genomic DNA (Data not shown). All primers were designed with a similar T_m (60 to 62°C) to maximize efficiency of the PCR.

Following PCR, the reaction products were separated on an 8% polyacrylamide TAE gel.

The radiolabeled PCR products within the gel were visualized and quantitated using a Molecular Dynamics PhosphorImager and associated ImageQuant software (Amersham Biosciences, Piscataway, NJ, USA). Volume integration was performed on bands corresponding to each transcript. The total intensity of the gene of interest was divided by the total intensity (minus background) of the 18S band to yield a relative value which is then compared across samples.

Plasmid construction

Genomic DNA was isolated from leaf tissue of *Arabidopsis* Columbia plants and used as a template for PCR. (Metzlaff, 1990) The ATCase promoter was amplified with primers common at the 5' end but differing at the 3' end to amplify fragments with intron and without the intron found in the 5' UTR (Table 1 and Figure 10). The PCR mixture was identical to that described above but 1 µg of genomic DNA was used as the template. The cycling parameters are as follows: after a single denaturation step of 4 min. at 94°C, 30 cycles of [denaturation at 94°C for 60 s, annealing at 55°C for 60 s and elongation at 72°C for 180 s] was performed. The PCR products were electrophoresed on a 0.8% TAE agarose gel and extracted. The resulting products of 3167 and 2665 nucleotides respectively (see Figure 10) were digested and cloned into the *Xba*I and *Bam*HI sites of the binary vector pCambia3300-GUS, to make constructs pRT498 and pRT499, respectively.

Likewise, PCR was used to amplify the DHOase promoter region. Again, the 5' primer was the same for each reaction with the 3' primer differing so as to amplify the proximal promoter with intron and without the intron found in the 5' UTR. (Table 1) The cycling

parameters were identical to that for amplification of the ATCase promoter. The resulting fragments were 1314 and 983 nucleotides long respectively (Figure 10). These fragments were digested with *Xba*I and *Bam*HI and cloned into the *Xba*I and *Bam*HI sites of the binary vector pCambia3300-GUS, to make constructs pRT494 and pRT495, respectively.

The DHODH promoter was amplified using a single set of primers (Table 1). Reaction conditions were identical to those above but the cycling conditions were modified due to the short length of this promoter. The cycling parameters are as follows: after a single denaturation step of 4 min. at 94°C, 30 cycles of [denaturation at 94°C for 60 s, annealing at 55°C for 60 s and elongation at 72°C for 60 s] was performed. The 380 nucleotide PCR product (Figure 10) was agarose gel purified as before, digested with *Xba*I and *Bam*HI and ligated into the *Xba*I and *Bam*HI sites of pCambia 3300-GUS to make construct pRT510.

Three UMP synthase promoter GUS fusions were constructed using genomic clones we have previously isolated (Kafer, et al. 2002) (Table 1). The first construct contains the proximal 405 nt upstream of the UMP synthase. The 405 nt *Eco*RI fragment was ligated into the plasmid pRT386 which contains the GUS gene. The *Bgl*II and *Sac*I fragment containing the promoter GUS fusion from this plasmid was then ligated into the *Bam*HI and *Sac*I sites of pBI101.2. The resulting plasmid was named pRT415 (Figure 10) and used to transform *Arabidopsis* as described.

The second plasmid, pRT446, contains 2056 nt upstream of the UMP synthase gene. This construct contains a portion of the flanking SKIP5 gene as well as the core promoter region of UMPs contained in pRT415 (Figure 10). The plasmid pRT443 was previously constructed during our genomic library screening by digesting our UMP synthase genomic clone with *Xba*I

and *Bam*HI and ligating into the plasmid pRT386. The resulting plasmid pRT443 contains the 2056 nt upstream of UMPs region as well as the GUS gene. pRT443 was digested with *Xba*I and *Sac*I, the fragment was gel purified and ligated into the *Xba*I and *Sac*I sites of pBI101.2. This construct, pRT446, (Figure 10) was used to transform *Arabidopsis* as described.

The final UMPs promoter GUS fusion constructed contains the promoter region as in pRT446 as well as the entire OPRTase portion (through codon 177) of the UMP synthase gene (Figure 10). This plasmid was constructed by ligating the *Xba*I and *Hinc*II fragment from pRT357, which contains the UMP synthase gene from our genomic library screening, into the *Xba*I and *Sma*I sites of pBI101.1. This plasmid, pRT457, was also used in *Arabidopsis* transformation as described.

Transgenic plant production

Transgenic *Arabidopsis* plants were generated by the floral dip method essentially as previously described (Desfeux et al., 2000). Briefly, *Arabidopsis* were grown in 2" pots in soil at 22°C under 16 hr day length until flowering. When primary bolts were four to six inches tall and contained flowers of various stage they were used for dipping.

Agrobacterium tumefaciens C58C1 pMP90 was transformed with the various binary vectors containing our promoter GUS fusions, via electroporation. Single colonies were inoculated into 5 ml of LB media containing 100 µg/ml Rifampicin and the appropriate antibiotic for plasmid selection and grown overnight. Five ml of the overnight culture was used to inoculate 150 ml of LB media containing the appropriate antibiotics and grown until an OD₆₀₀ >2.0. The culture was spun down in a centrifuge for 10 min at 5000 RPM. The pellet was resuspended in 300 ml of 5% sucrose containing 0.03% Silwet L-77 (OSI Specialties).

Typically 10-15 pots containing 5-10 flowering *Arabidopsis* plants each were inverted into the *Agrobacterium* solution and gently swirled for 15-20 seconds. Pots were laid on their sides in flats and covered with plastic wrap for 24 hours. At this point plants were up righted and returned to the growth room and grown for another 3-4 weeks and gradually allowed to dry out. Seeds were pooled from these T₀ plants and stored until use.

Depending on the vector used, transgenic T₁ plants were selected either by kanamycin or glufosinate ammonium (BASTA) application (AgroEvo). For kanamycin resistance containing plants, seeds were surfaced sterilized in 25 % bleach for 15 min and rinsed with sterile water twice. Several hundred seeds were plated on 150mm petri dishes containing MS media plus 50µg/ml kanamycin. Surviving T₁ plants were transferred to soil, grown to maturity using the Arasystem (Lehle Seeds, 1102 South Industrial Blvd, Suite D, Round Rock TX, 78681 USA) and seeds were collected. In the case of BASTA resistance, several thousand seeds were distributed over flats containing potting mix. When plants were showing the first set of true leaves (usually within ten days after sowing) the flats were sprayed with a solution of 0.3% basta. T₂ seeds were collected from the surviving plants using the Arasystem. All analysis was carried out on the resulting T₂ generation of plants.

GUS staining

T₂ plants were used for staining to localize GUS expression. Plants at the indicated stages were placed in GUS staining buffer (100 mM Tris-HCl pH 7.5, 200 mM NaCl, 2mM K₃Fe(CN)₆, 2mM K₄Fe(CN)₆, 10 mM EDTA, 0.5% Triton X-100, 2 mM X-Gluc) and vacuum infiltrated for 15-30 minutes. Plants were then incubated at 37°C for three hours to overnight. Staining was stopped by addition of 70% ethanol. Ethanol was changed several times until the

tissue had completely cleared.

GUS assay

GUS activity was assayed in extracts of whole plants grown *in vitro* as described above. The continuous spectrophotometric assays for *para*-nitrophenyl (pNP) formation were carried out essentially as described (Aich et al., 2001). Approximately 5-10 plants from each line were ground in 1.5 ml of GUS assay buffer (50mM sodium phosphate pH 7.0, 10 mM beta-mercaptoethanol) and microcentrifuged for 10 min at 12,000xg to remove cellular debris. Protein was quantitated via the method of Bradford (Bradford, 1976) using the Bio-Rad (Bio-Rad, Hercules, CA, USA) protein assay reagent using BSA as standards.

For each assay, 25 or 50µg of total protein was used in a one ml assay consisting of 1 mM *para*-nitrophenol β-D glucuronide (pNPG), 50 mM sodium phosphate pH 7.0, 10 mM beta-mercaptoethanol. The reaction was started by addition of protein and absorbance was monitored at 405 nm. The reaction was followed for 15 minutes and pNP formation remained linear over this time period. The reported molar extinction coefficient for pNP of 9000 was used to calculate amount of product formation (Aich et al., 2001).

Acknowledgments

We wish to thank Dr. Steve Whitham for generously providing *Arabidopsis* tissue for our analysis of ORMV infected plants. We also thank Dr. Basil Nikolau for his generous gift of the pCambia3300 vector containing the GUS gene used in our transgene experiments.

Literature Cited

- Aich, S., L. T. Delbaere and R. Chen (2001). "Continuous spectrophotometric assay for beta-glucuronidase." Biotechniques **30**(4): 846-50.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J. Mol. Biol. **215**: 403-410.
- Ashihara, H. (1977). "Changes in activities of the *de novo* and salvage pathways of pyrimidine nucleotide biosynthesis during germination of black gram [mungo beans] (*Phaseolus mungo*) seeds." Z Pflanzenphysiol **81**: 199-211.
- Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding." Anal. Biochem. **72**: 248-254.
- Callejo, M., D. Alvarez, G. B. Price and M. Zannis-Hadjopoulos (2002). "The 14-3-3 protein homologues of *S. cerevisiae*, Bmh1p and Bmh2p, have DNA cruciform binding activity and associate in vivo with ARS307." J Biol Chem **6**: 6.
- Callis, J., M. Fromm and V. Walbot (1987). "Introns increase gene expression in cultured maize cells." Genes Dev **1**(10): 1183-200.
- Ciereszko, I., H. Johansson, V. Hurry and L. A. Kleczkowski (2001). "Phosphate status affects the gene expression, protein content and enzymatic activity of UDP-glucose pyrophosphorylase in wild-type and pho mutants of *Arabidopsis*." Planta **212**(4): 598-605.
- Desfeux, C., S. J. Clough and A. F. Bent (2000). "Female reproductive tissues are the primary target of *Agrobacterium*- mediated transformation by the *Arabidopsis* floral-dip

method.” Plant Physiol **123**(3): 895-904.

Deyholos, M. K. and L. E. Sieburth (2000). “Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron.” Plant Cell **12**(10): 1799-810.

Edwards, G. E. and D. Walker (1983). C3, C4 : mechanisms, and cellular and environmental regulation, of photosynthesis. Berkeley, University of California Press.

GCG (Version 9.0). Wisconsin Package. Madison, Wisc., Genetics Computer Group.

Hambor, J. E., J. Mennone, M. E. Coon, J. H. Hanke and P. Kavathas (1993). “Identification and characterization of an Alu-containing, T-cell- specific enhancer located in the last intron of the human CD8 alpha gene.” Mol Cell Biol **13**(11): 7056-70.

Hasegawa, P. M., R. A. Bressan, J.-K. Zhu and H. J. Bohnert (2000). “Plant cellular and molecular responses to high salinity.” Annu. Rev. Plant Physiol. Plant Mol. Biol. **51**: 463-499.

Kafer, C. and R. Thornburg (2002). “The thymine starvation transcriptional program coordinately up-regulates both the *Arabidopsis thaliana* UMP synthase gene and a novel F-box gene located at the same locus.” Journal of Plant Biology.

Kafer, C. and R. W. Thornburg (1999). “Pyrimidine Metabolism in Plants.” Paths to Pyrimidines **5**: 7-19.

Kato, T., E. Itoh, R. F. Whittier and D. Shibata (1998). “Increase of foreign gene expression in monocot and dicot cells by an intron in the 5' untranslated region of a soybean phosphoenolpyruvate carboxylase gene.” Biosci Biotechnol Biochem **62**(1): 151-3.

Loffler, M. (1992). “The "anti-pyrimidine effect" of hypoxia and brequinar sodium (NSC

- 368390) is of consequence for tumor cell growth." Biochem Pharmacol **43**: 2281-2287.
- Löffler, M. (1994). "Pyrimidines as pacemakers of cell proliferation at low oxygen tension - a general biological principle?" Paths to Pyrimidines **2**: 1-6.
- Lou, H., A. J. McCullough and M. A. Schuler (1993). "Expression of maize Adh1 intron mutants in tobacco nuclei." Plant J **3**(3): 393-403.
- Maas, C., J. Laufs, S. Grant, C. Korfhage and W. Werr (1991). "The combination of a novel stimulatory element in the first exon of the maize Shrunken-1 gene with the following intron 1 enhances reporter gene expression up to 1000-fold." Plant Mol Biol **16**(2): 199-207.
- Mazus, B. and J. Buchowicz (1972). "Activity of enzymes involved in pyrimidine metabolism in the germinating wheat grains." Phytochem **11**: 77-82.
- Metzlaff, H. J. a. M. (1990). "A simple and rapid method for the preparation of total plant DNA." Biotechniques **8**(2): 176.
- Morello, L., M. Bardini, F. Sala and D. Breviario (2002). "A long leader intron of the Ostub16 rice beta-tubulin gene is required for high-level gene expression and can autonomously promote transcription both in vivo and in vitro." Plant J **29**(1): 33-44.
- Mun, J. H., S. Y. Lee, H. J. Yu, Y. M. Jeong, M. Y. Shin, H. Kim, I. Lee and S. G. Kim (2002). "Petunia actin-depolymerizing factor is mainly accumulated in vascular tissue and its gene expression is enhanced by the first intron." Gene **292**(1-2): 233-43.
- Santoso, D. (1991). Production and characterization of UMP Synthase mutants from haploid cell suspensions of *Nicotiana tabacum*. Iowa State University.
- Santoso, D. (1995). 5-Fluoroorotic acid-selected cell lines and regulation of UMP synthase gene

expression in tobacco cells, Iowa State University.

- Santoso, D. and R. W. Thornburg (1992). "Isolation and Characterization of UMP synthase mutants from haploid cell suspensions of *Nicotiana tabacum*." Plant Physiol. **99**: 1216-1225.
- Santoso, D. and R. W. Thornburg (1995). "The use of a heterologous antibody to characterize UMP synthase mutant of tobacco cells." Menara Perkebunan **63**: 71-79.
- Santoso, D. and R. W. Thornburg (1998). "UMP synthase is transcriptionally regulated by pyrimidine levels in *Nicotiana plumbaginifolia*." Plant Physiol **116**: 815-821.
- Santoso, D. and R. W. Thornburg (2000). "Fluoroorotic acid-selected *Nicotiana plumbaginifolia* cell lines with a stable thymine starvation phenotype have lost the thymine-regulated transcriptional program." Plant Physiol. **123**: 1517-1524.
- Sidorenko, L. V., X. Li, S. M. Cocciolone, S. Chopra, L. Tagliani, B. Bowen, M. Daniels and T. Peterson (2000). "Complex structure of a maize Myb gene promoter: functional analysis in transgenic plants." Plant J **22**(6): 471-82.
- Voet, D. and J. G. Voet (1990). Biochemistry. New York. John Wiley & Sons.
- Williamson, C. L. and R. D. Slocum (1993). "Characterization of an aspartate transcarbamoylase cDNA from pea (*Pisum sativum* L.)." Plant Physiol **102**(3): 1055-6.
- Williamson, C. L. and R. D. Slocum (1994). "Molecular cloning and characterization of the pyrB1 and pyrB2 genes encoding aspartate transcarbamoylase in pea (*Pisum sativum* L.)." Plant Physiol **105**(1): 377-84.

Figure Legends

Figure 1. Artificially induced pyrimidine starvation causes an upregulation of the entire *de novo*

biosynthetic pathway. *Arabidopsis* plants were grown as described in Materials and Methods. Aminopterin was added to the indicated final concentration. Plants were harvested five days later for RNA extraction and RT-PCR as described in Materials and Methods. (A) Phosphorimager scan of radiolabeled PCR products separated on an 8% TAE polyacrylamide gel. (B) Dose response of the indicated aminopterin concentration reported as the ratio of *de novo* transcript level to 18S rRNA level. Error bars are the standard deviation of 3 PCR replicates.

Figure 2. Organ specific expression of the *de novo* pyrimidine biosynthetic pathway. *Arabidopsis* was grown according to Materials and Methods and the indicated organs were harvested for RNA extraction. RT-PCR was performed as described in Materials and Methods. (A) Phosphorimager scan of radiolabeled PCR products separated on an 8% TAE gel. (B) Relative quantitative levels of each transcript expressed in each organ. Black bars show levels in roots, diagonal hatching shows level in leaves, solid grey bars show levels in stems, stippled bars show levels in flowers and vertical hatch bars show levels in siliques.

Figure 3. Analysis of the expression pattern of the *de novo* pyrimidine biosynthetic pathway. *Arabidopsis* seeds were grown according to Materials and Methods and harvested at the time points indicated for RT-PCR analysis. (A) Phosphorimager scan of radiolabeled PCR products separated on an 8% TAE polyacrylamide gel. (B) Ratio of pyrimidine gene intensity to 18S intensity. Solid black bars are 24 hours, gray bars are 48 hours, stippled bars are 72 hours, diagonal hatch are 96 hours and horizontal hatched bars are 120 hours after sowing. Error bars show the standard deviation of 3 PCRs.

Figure 4. Severe flood stress (anaerobic conditions) down regulates the final three steps of pyrimidine biosynthesis. *Arabidopsis* plants were grown according to Materials and Methods. Plants were harvested for RNA extraction and RT-PCR as described in Materials and Methods. (A) Phosphorimager scan of radiolabeled PCR products separated on an 8% polyacrylamide TAE gel. (B) Percent change in relative levels of each amplicon. The hypoxic levels were set to 100% and are the horizontal hatched bars. The solid black bars are anoxic levels.

Figure 5. Relative quantitative of pyrimidine biosynthetic genes from *Arabidopsis* plants grown in a simulated flood condition. Plants were grown as described in Materials and Methods and plants were harvested for RNA extraction and RT-PCR as described in Materials and Methods. (A) Phosphorimager scan of the radiolabeled PCR products and separation on an 8% polyacrylamide TAE gel. Amplicons are indicated to the left of the scan. (B) Relative levels of each amplicon under flood or preflood conditions. Solid bars show the preflood levels of each amplicon, stippled bars show 6 hours of flood, hatched bars show 12 hours of flood and solid white bars show levels after 24 hours of flood. Numbers are reported as gene of interest to 18S. Error bars are \pm s.d. of 3 PCR replicates.

Figure 6. Analysis of phosphate starvation stress induced by mannose feeding.

Plants were grown as described in Materials and Methods. The plants were harvested for RNA extraction and RT-PCR as described. (A) Phosphorimager scan of the radiolabeled PCR products separated on an 8% polyacrylamide TAE gel. (B) Relative ratios of each amplicon under phosphate starvation created by mannose feeding. Black

bars show levels in MS media, horizontal cross hatch bars show levels in MS media + 10mM mannose and solid grey bars show levels in MS media + 20mM mannose.

Figure 7. Relative quantitative of pyrimidine biosynthetic genes from *Arabidopsis* plants grown in media containing or lacking phosphate. Plants were grown as described in Materials and Methods and plants were harvested for RNA extraction and RT-PCR as described in Materials and Methods. (A) Phosphorimager scan of the radiolabeled PCR products and separation on an 8% polyacrylamide TAE gel. Amplicons are indicated to the left of the scan. (B) Relative levels of each amplicon in media lacking or containing phosphate. Solid bars show the media +PO₄ levels of each amplicon, hatched bars show media -PO₄. Numbers are reported as gene of interest to 18S. Error bars are +/- s.d. of 3 PCR replicates.

Figure 8. Analysis of saline stress on the de novo pathway of pyrimidine biosynthesis. Plants were grown *in vitro* as described in materials and methods. Plants were harvested and RNA was extracted for RT-PCR analysis as described in materials and methods. (A) Phosphorimager scan of radiolabeled PCR products after separation on an 8% polyacrylamide TAE gel. (B) Ratio of pyrimidine to 18S. Error bars are +/- sd of 3 PCRs.

Figure 9. Analysis of the de novo pathway under the biotic stress of infection with Oil Rapeseed Mosaic Virus. *Arabidopsis* plants were grown as described in Materials and Methods. RNA was extracted and RT-PCR performed as described. (A) Phosphorimager scan of the radiolabeled PCR products after separation on an 8% polyacrylamide TAE gel. (B) Ratios of pyrimidine gene to 18S. Error bars are std. dev.

of 3 PCR replicates.

Figure 10. Schematic representation of the promoters and genes of the pyrimidine biosynthetic pathway. Names of the plasmid constructs used in Arabidopsis transformation are shown below the respective gene.

Figure 11. GUS expression under control of the ATCase promoter with and without the intron from the 5' UTR of the ATCase gene. T2 seeds of pRT498 and pRT499 were grown as described in Materials and Methods. Plants were harvested at the indicated days and stained for GUS activity as described. TR498 A) Day 6 B) 10 day old plant C) 17 day old plant D) closeup of root and E) leaf F) flower TR499 G) 6 day old plant H) 10 day old plant I) 17 day old plant J) closeup of root and K) leaf L) flowers

Figure 12. Arabidopsis TR494 and TR495 carrying the DHOase promoter constructs. Plants were grown as in Materials and Methods and stained for GUS activity. TR494 A) 6 day old plant B) 10 day old plant C) 17 day old plant D) root E) leaf F) flowers TR495 G) 6 day old plant H) 10 day old plant I) 17 day old plant J) root K) leaf L) flower.

Figure 13. TR510 plants containing the DHODH promoter were grown as in Materials and Methods. Plants were harvested at the indicated time points and stained for GUS activity.

A) 6 day old plant B) 10 day old plant C) 17 day old plant D) flower

Figure 14. TR415, TR446, TR457 containing the UMP synthase promoter constructs. T2 seeds were grown as in Materials and Methods. Plants were harvested at 6, 10 and 17 days and stained for GUS activity. **A-D TR415 plants** A) 7 day old plant B) 10 day old plant C) 17 day old plant D) leaf E) root F) flower. **E-F TR446 plants** G) 7 day old plant H) 10

day old plant I) 17 day old plant. J) flowers. **I-M TR457 plants** K) 7 day old plants L) 10 day old plant M) 17 day old plant N) root tip O) flower.

Figure 15. The ATCase and DHOase introns contain conserved sequences. The sequences of the ATCase and DHOase introns found in the 5'UTR were analyzed using the "pileup" feature of GCG. Several regions of conserved identity are found. The red shading indicates the highest level of identity. Conserved nucleotides are indicated by the asterisks below the alignment.

Figure 16. Schematic of the Dihydroorotate dehydrogenase locus. DHODH and SOD are depicted with introns shown as lines and exons shown as boxes. The numbering is the position in base pairs on BAC MKD15. The cruciform structure which is predicted to fold within the promoter region is shown in the lower panel.

Figure 17. Histogram showing relative expression levels between transgenic plants. GUS assays were performed as in Materials and Methods. The units of activity are expressed as picomole pNP produced $\cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ total protein.

Table 1. Genes and Oligonucleotides used in this study
Oligonucleotides used to amplify promoter regions

Aspartate transcarbamoylase

ATCase P1	5'-GCTCTAGAAGCAATAGGTGAGAGAGATCG-3'	5' end of ATCase promoter
ATCase P2	5'-GCGGATCCGGATGCTAATTCGTTGAAGGG-3'	5'UTR of ATCase mRNA upstream of intron 1
ATCase P3	5'-GCGGATCCGTCGATTGAATCAGTAAAGCGA-3'	5'UTR of ATCase mRNA downstream of intron 1

Dihydroorotase

DHOase P1	5'-GCTCTAGACTCTTGAAACCCAAATCAAAGC-3'	5' end of DHOase promoter
DHOase P2	5'-GCGGATCCAGCTGCTCAAGGCTTTGCAA-3'	5' UTR of DHOase mRNA upstream of intron 1
DHOase P3	5'-GCGGATCCAAATGCACAACAGGGATGCTT-3'	5' UTR of DHOase mRNA downstream of intron 1

Dihydroorotate dehydrogenase

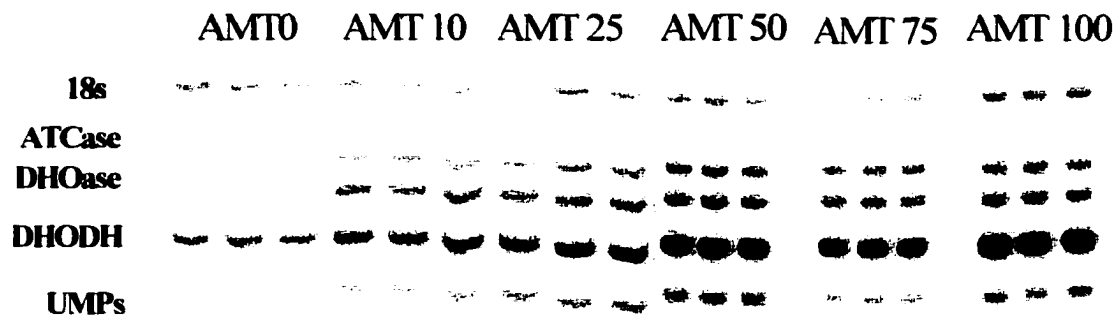
DHODH P1	5'-GGTCTAGAAGTTCCCACTGTTGTCG-3'	5' end of DHODH promoter
DHODH P2	5'-GGGGATCCTTAAATCGACAGTGAATCAAAC-3'	5' UTR of DHODH mRNA

Oligonucleotides used in quantitative PCR studies

Fragment size generated (nucleotides)

18S		450
18SF2	5'-AACTTACCAGGTCCAGACATA-3'	
18SR	5'-TAGGAGCGACGGGCGGTG-3'	
Aspartate transcarbamoylase		280
ATCaseF	5'-AGAACGGTTTGTTCATTAGAATG-3'	
ATCaseR	5'-AAATCAATTTATTCATTAAGTTCAAC-3'	
Dihydroorotase		223
DHOaseF	5'-GAAATCTCCTTGGAAGGTTCC-3'	
DHOaseR	5'-CAGTATCTTGCTTTTATTCCTC-3'	
Dihydroorotate dehydrogenase		171
DHODHF	5'-GGAAGTGGTGAAATGCTTAGA-3'	
DHODHR	5'-GTAGAAAGAGACAACCTGAATCA-3'	
UMP synthase		110
UMPsF	5'-GATCCTGAGAGATGGCTGAG-3'	
UMPsR	5'-CAACGAGACATGAGTCTTAAAA-3'	

A



B

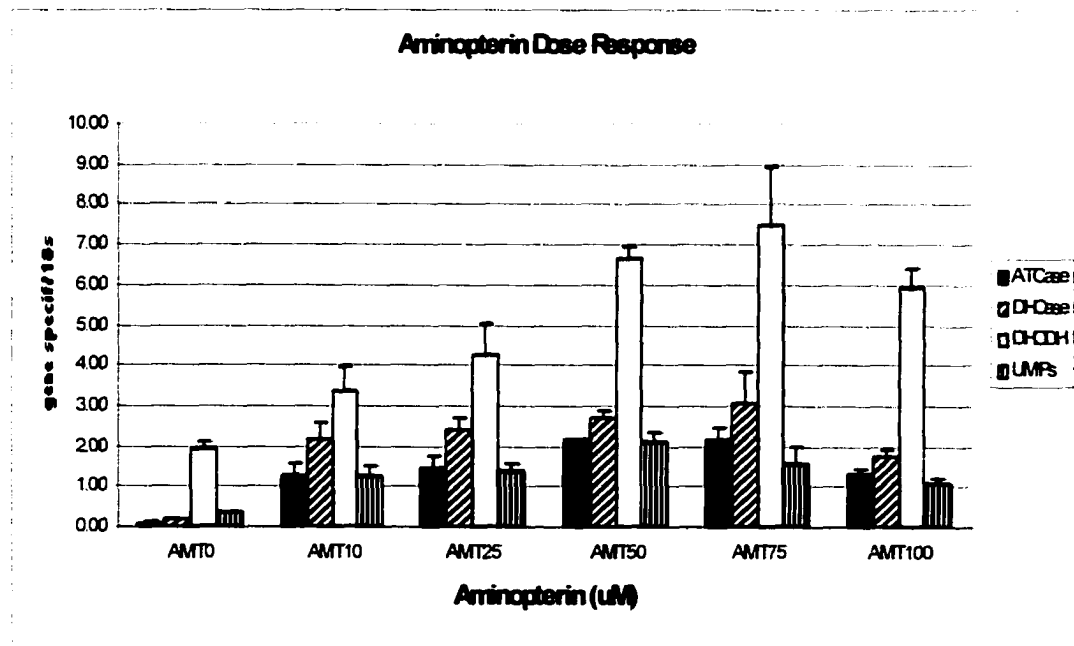
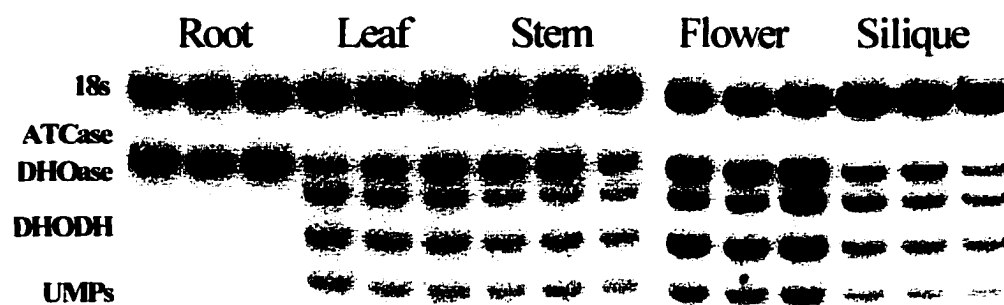


Figure 1.

A



B

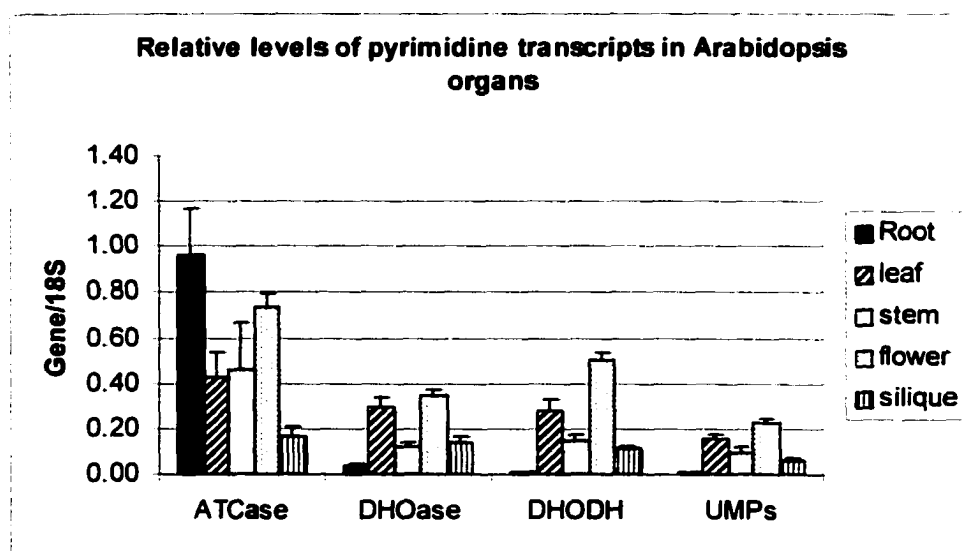


Figure 2.

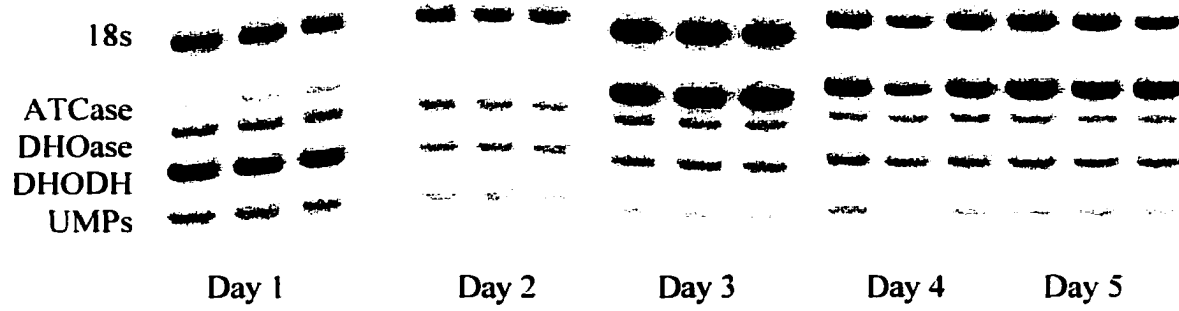
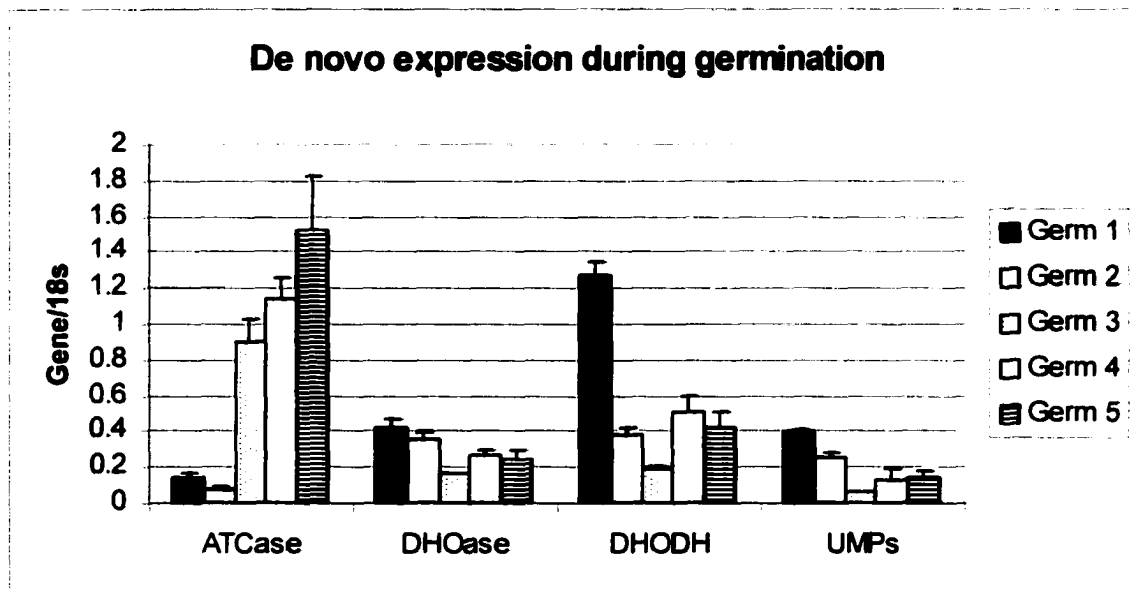
A**B**

Figure 3.

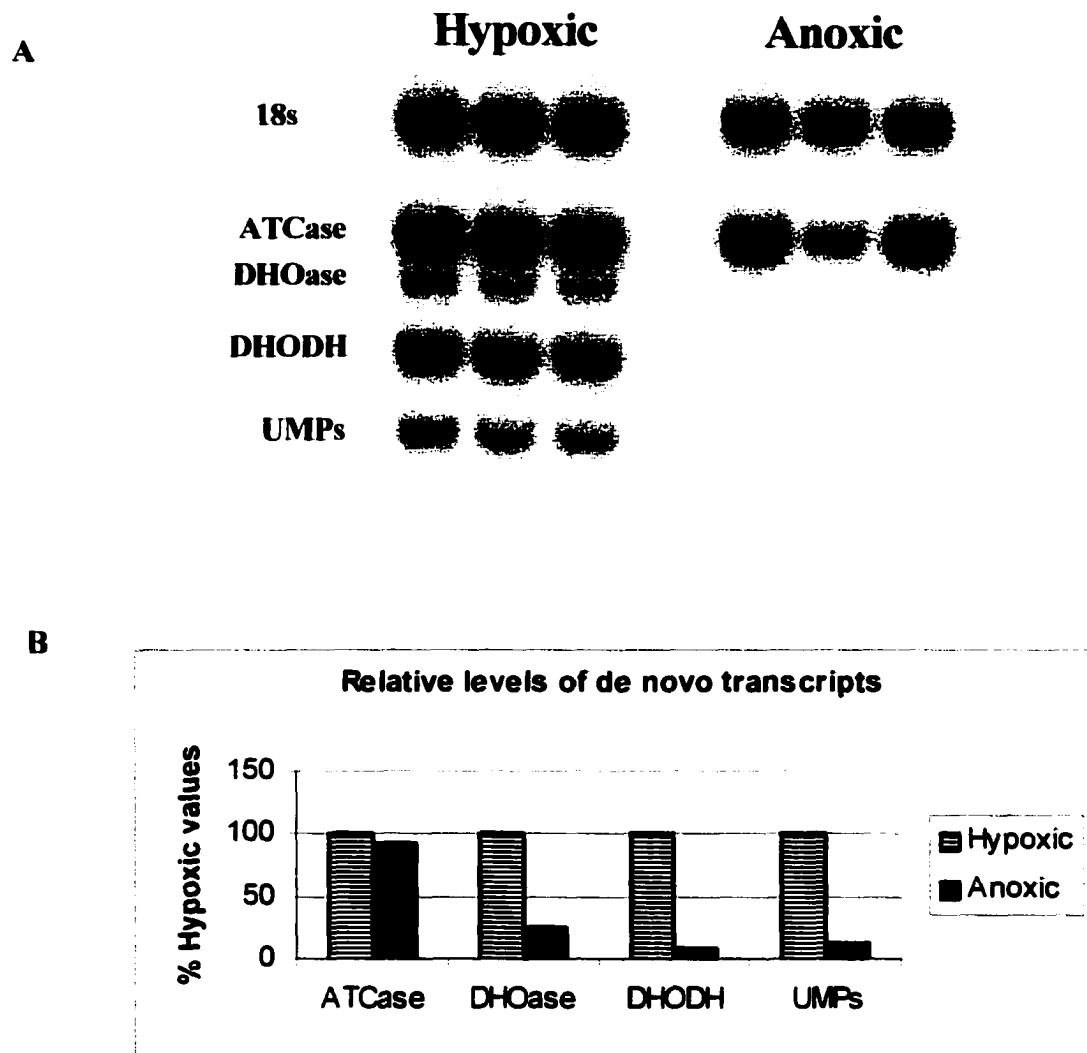
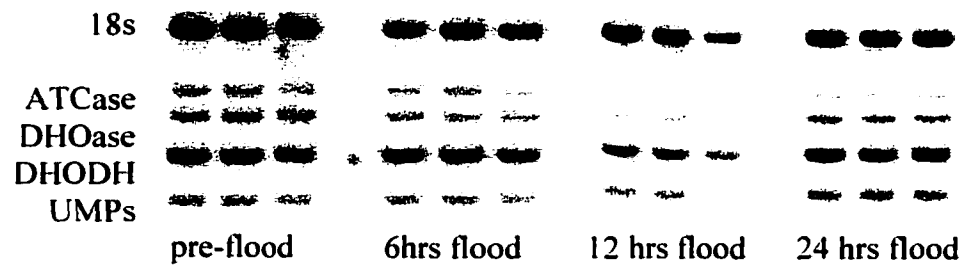


Figure 4.

A



B

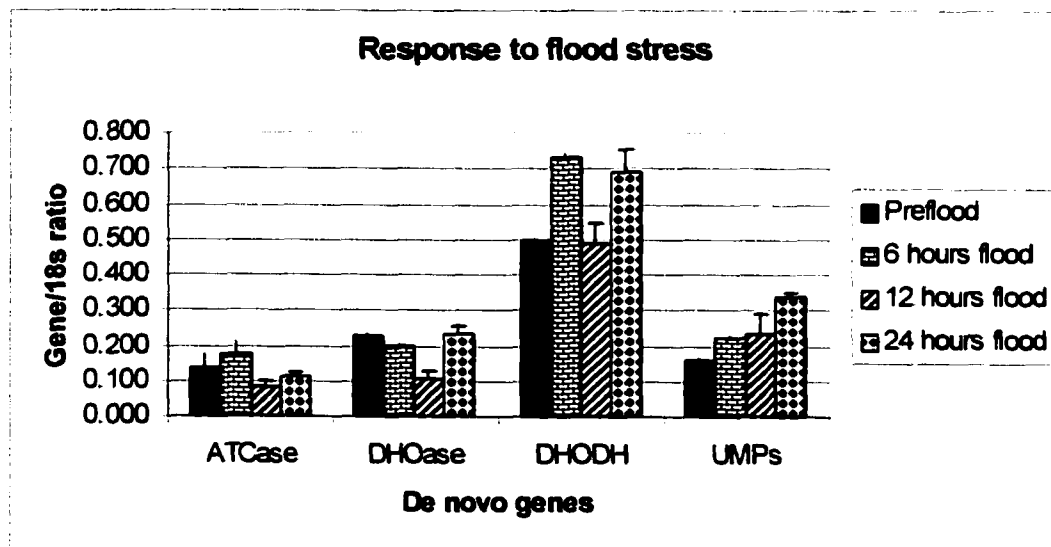


Figure 5.

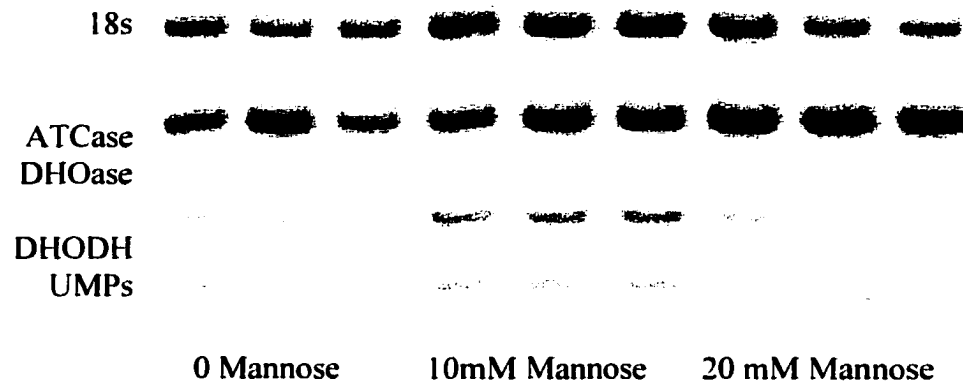
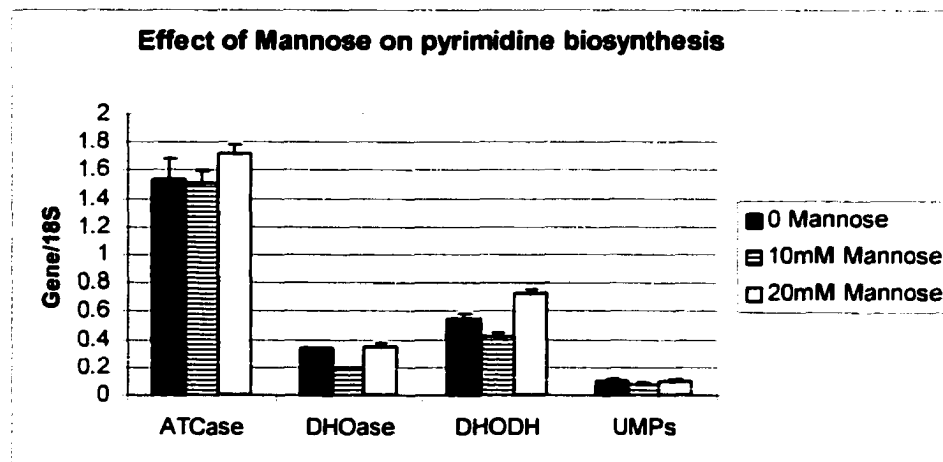
A**B**

Figure 6.

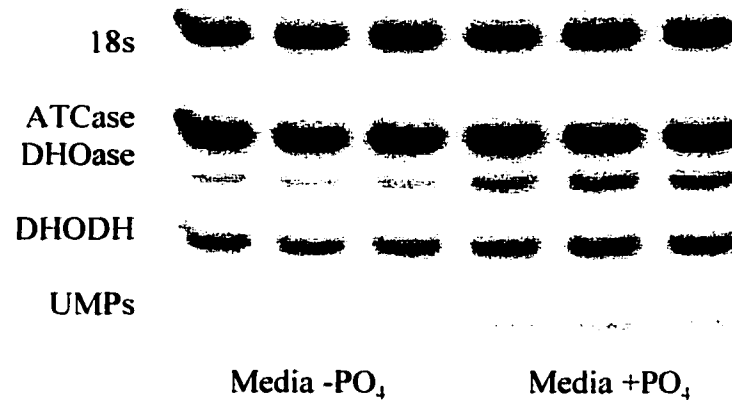
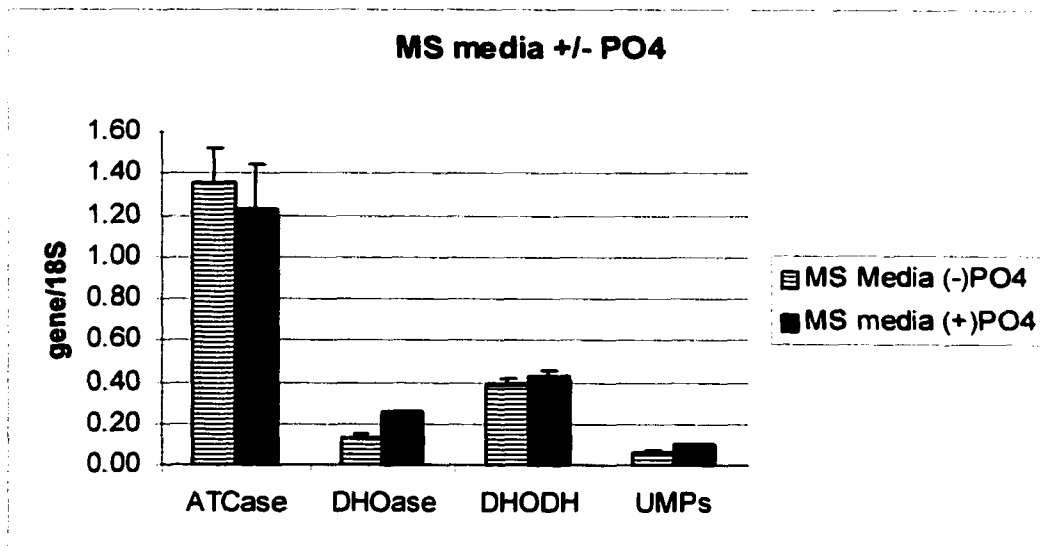
A**B**

Figure 7.

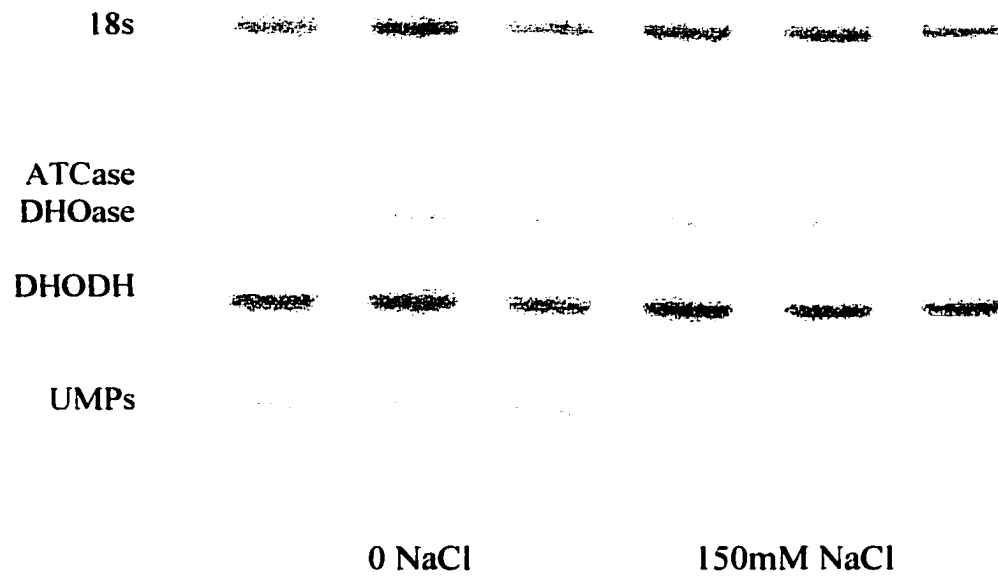
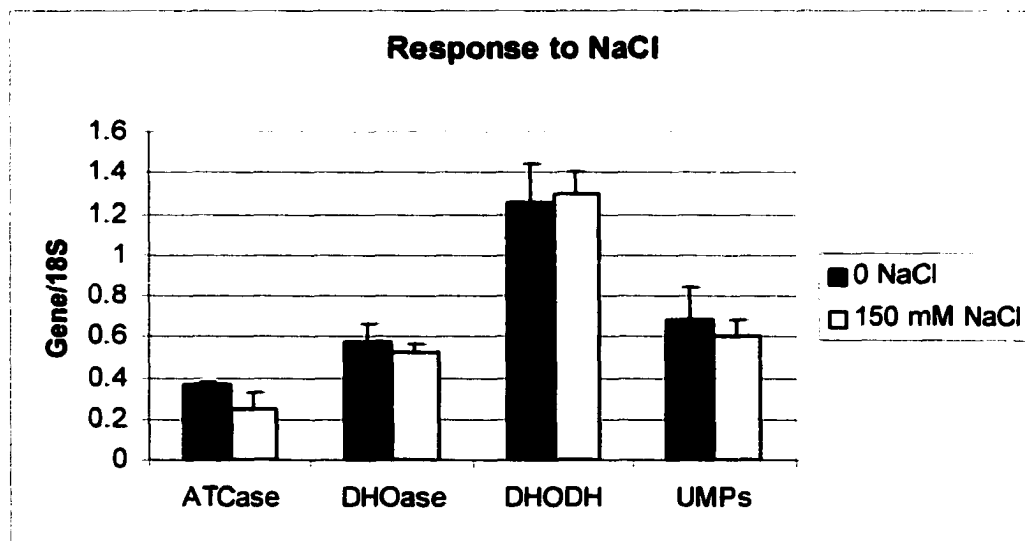
A**B**

Figure 8.

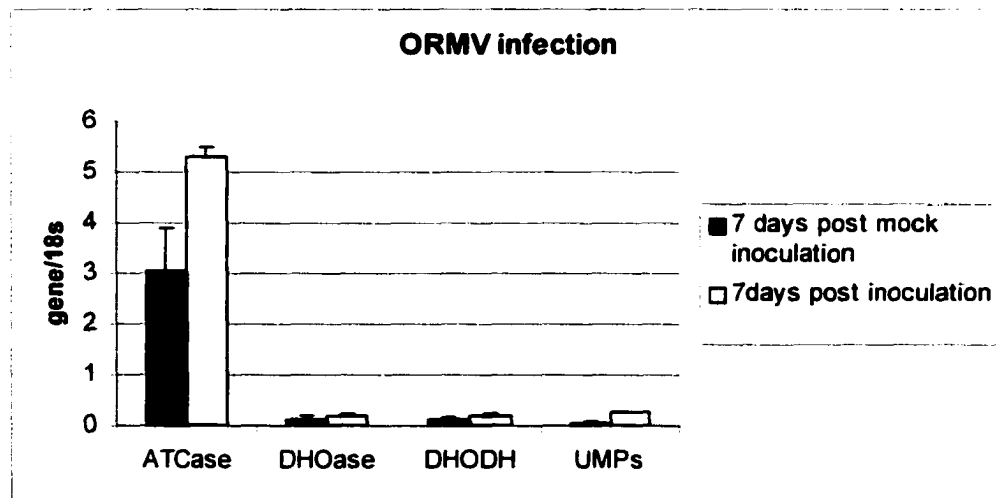
A**B**

Figure 9.

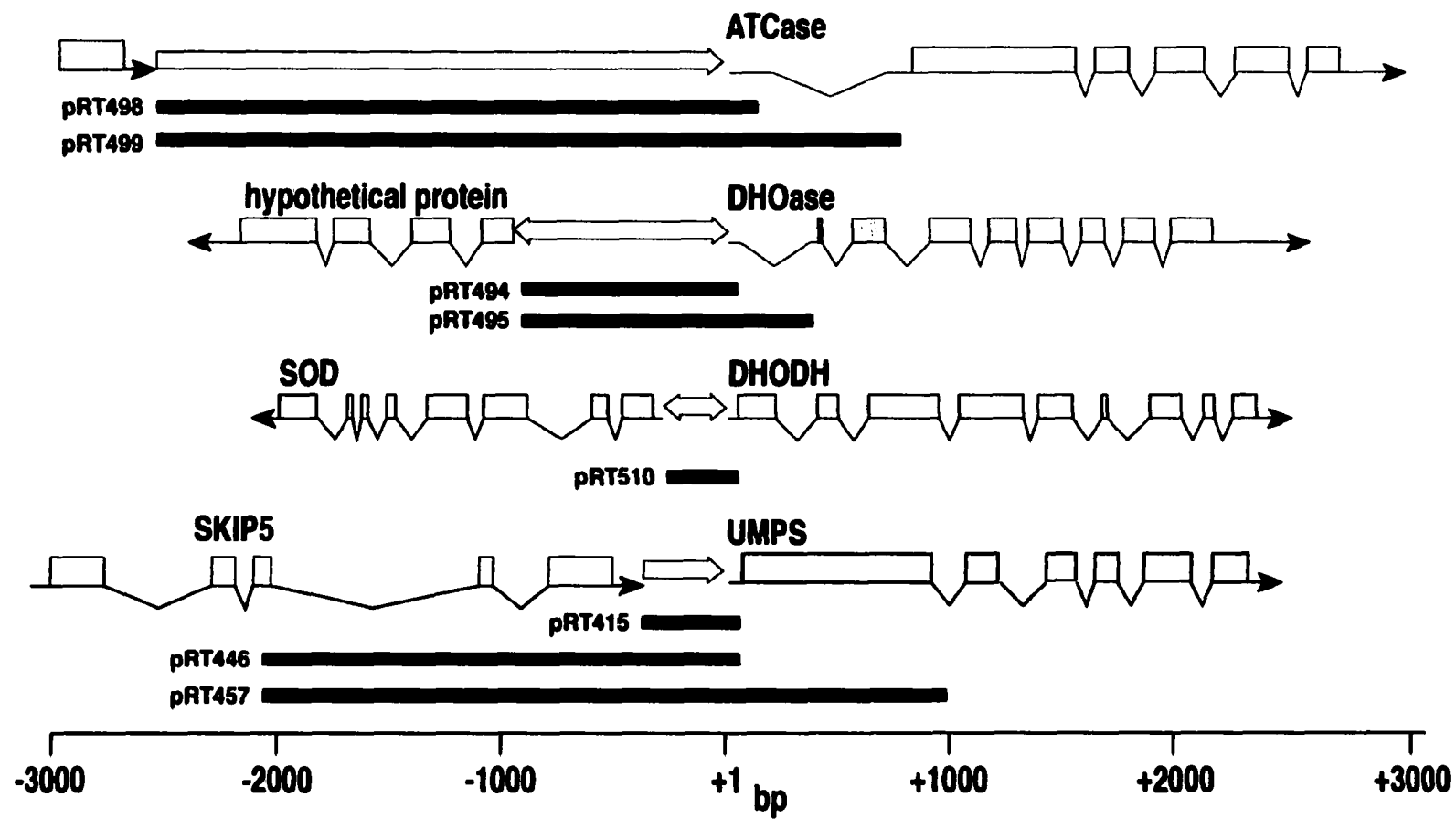


Figure 10.

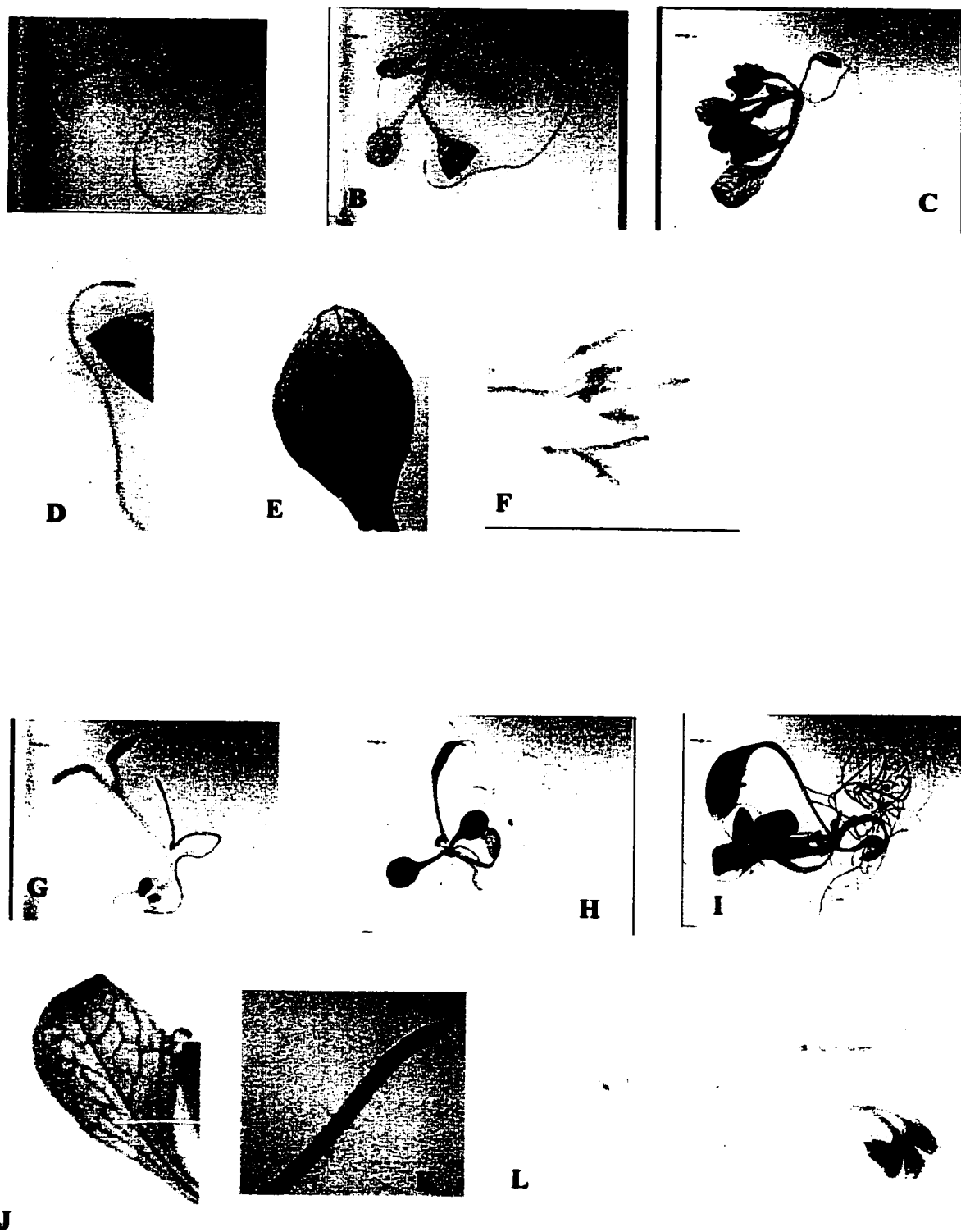


Figure 11.

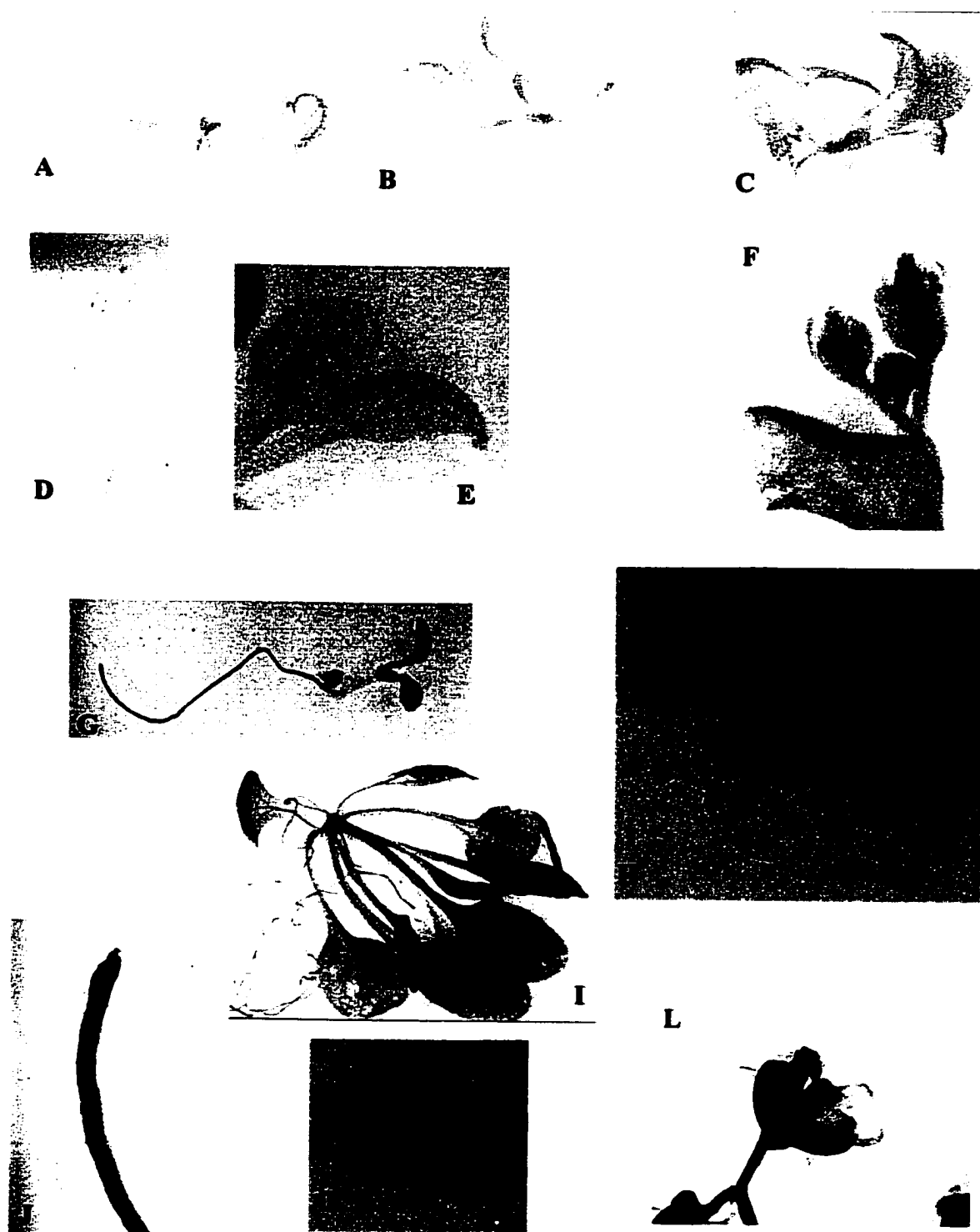


Figure 12.

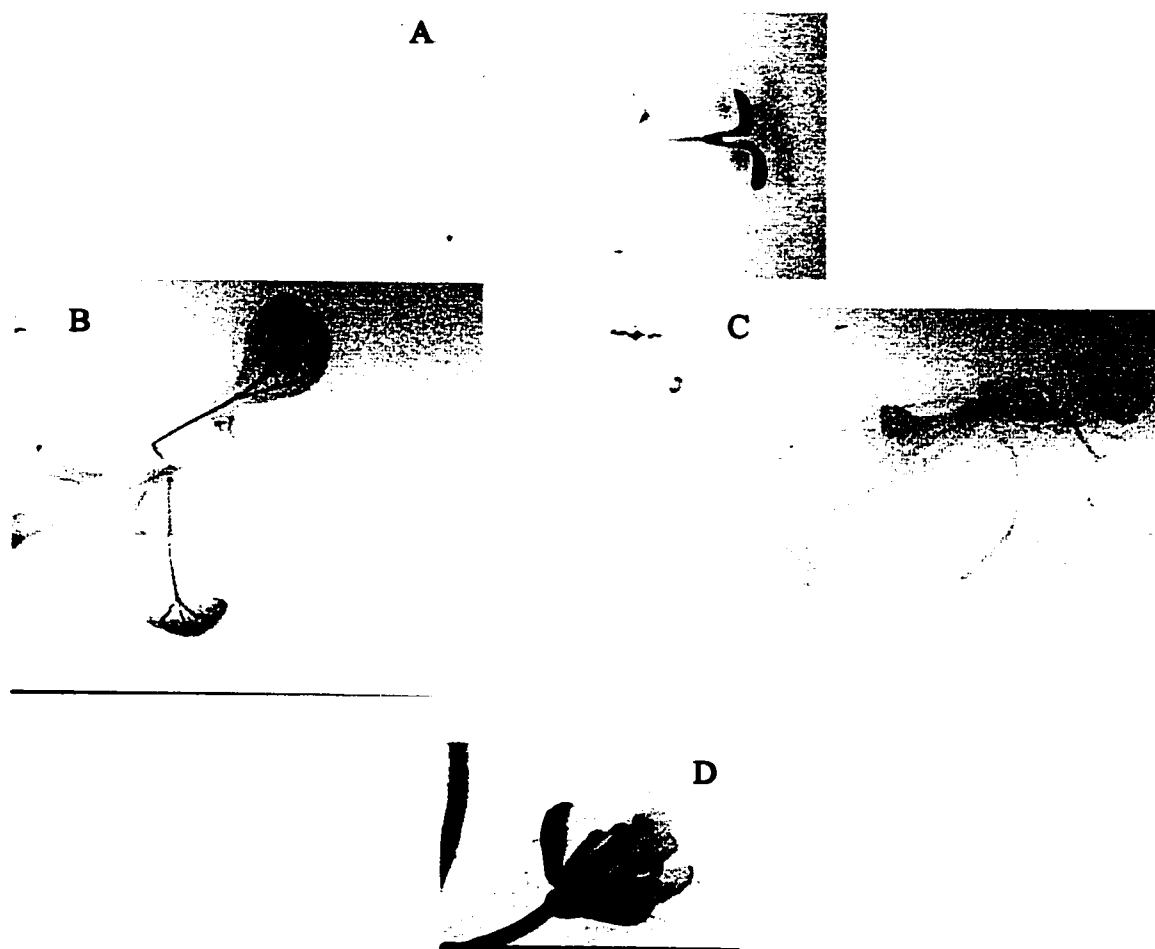


Figure 13.

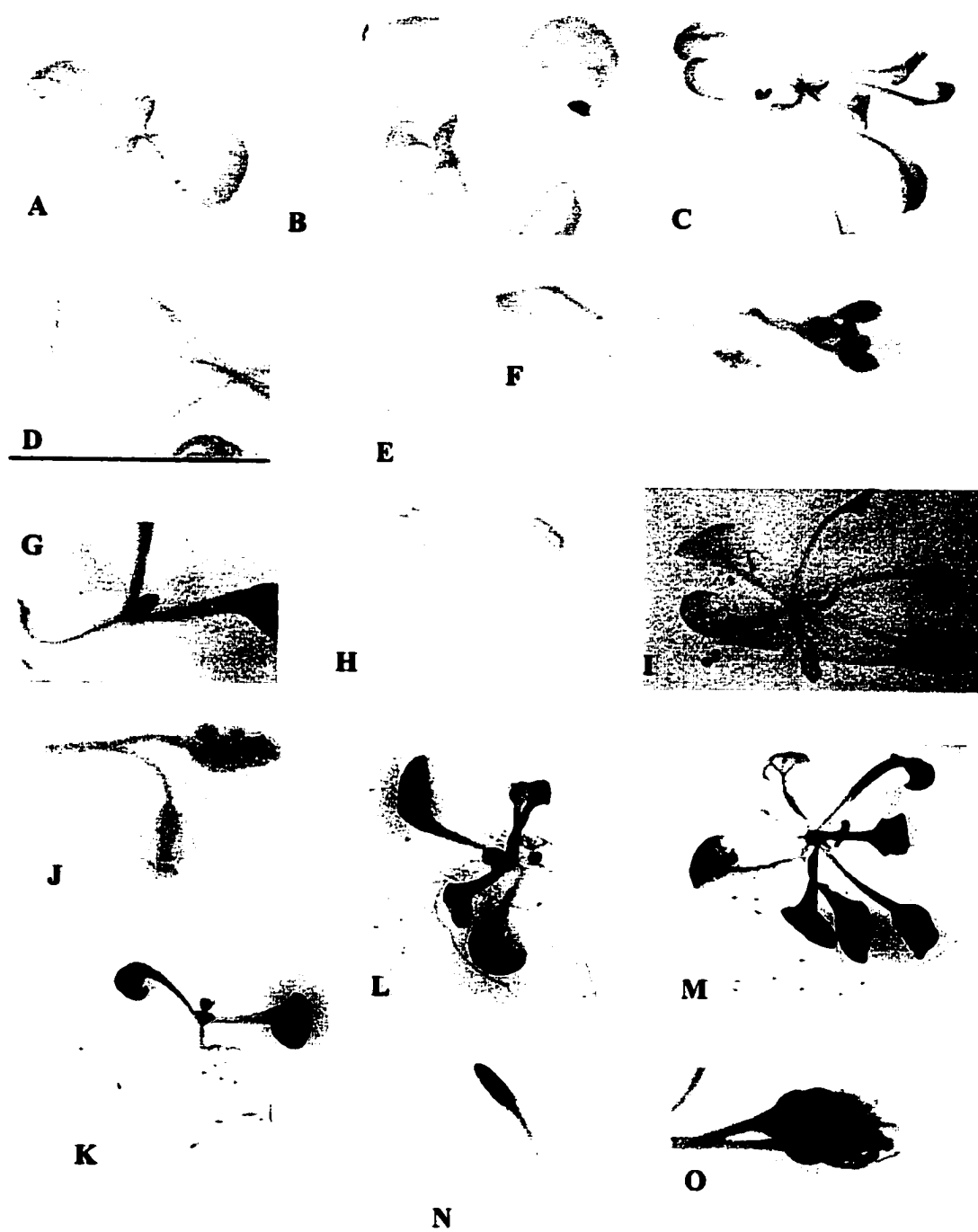


Figure 14.

```

                *          20      149      *          40          *
ATCase : AGGTCTGTGTAGGTTTTGAAATTCTGCGGATCAGATTCTGAATTTTTTTC : 50
DHOase : ----- : -

                60          *          80          *          100
ATCase : TTTTTCGATTTTAACTACGAAAGAAATTGAATTTCATAT : 100
DHOase : -----AGGTCGTAACTCTCCCTTAAACAAATTCAT : 34
                G TT      ATTG TT      A      AAA      GTT      AGTT

                *          120          *          140          *
ATCase : TCGAAATTAAGAAACTTCTCTTAACTATTAAATGAAAT : 149
DHOase : TCTAATTTCTTTAGGAGTTTACCGTTCGCCATTGTGTAT : 84
                T      GATT TGGT      CTA      T      T      A      A      TTT      AT      T      G      TTA

                160          *          180          *          200
ATCase : GAAATTCATTCCAAAGAAACCAAGGAAATTCCTCTCAACCA : 199
DHOase : TTTTGGATAATTTATTTTGTAAATCCGAAATCAATGTG : 134
                G      T      T      T      GA      T      C      G      TTTTG      T      T      CC      T      T      T      T

                *          220          *          240          *
ATCase : AAGACAAATGAATCAGGATCAAGTAATGTTTGAATCTATAT : 249
DHOase : ATAGAAATGATTTAACTGGATACTAAATTCGTTGAGA : 176
                A      AATT      ATT      AT      TA      TA      ATT      T      T      TG      A

                260          *          280          *          300
ATCase : CAGAGATATGTAAAGAGTTACCAAAATAACCGCAACAAGATCG : 299
DHOase : AACATTACCAAGGATTCGATATCTTTGAGATGATTGAAAGTTT : 226
                A      T      GA      A      GA      G      A      T      T      AT      A      A      TA      T

                *          320          *          340          *
ATCase : TGCATTATGTGAGGAGAGTGATCTTGTATCAATCAAGAG : 346
DHOase : CATGCTTAACTTTAATACGAGTCAATTTTGAAGCAT : 272
                T      G      C      GCT      AT      T      G      A      A      AAT      T      TTTA      GT

                360          *          380          *          400
ATCase : TATATAAAATTCATTTTCTCAAGTTCGAATTTAGTTGCTT : 395
DHOase : GTATAGATCTTCTACTGGAATTGTTGGAAAGG : 314
                TTATA      AAT      TG      AGT      T      TT      TCAT      CT      AAC

                *          420          *          440          *
ATCase : TCGATTATTCTTAAACTAACGGCATCATATATTGTTGTCTGTTATTTAG : 445
DHOase : ----- : -

ATCase : CAGA : 449
DHOase : ---- : -

```

Figure 15.

Dihydroorotate dehydrogenase Locus

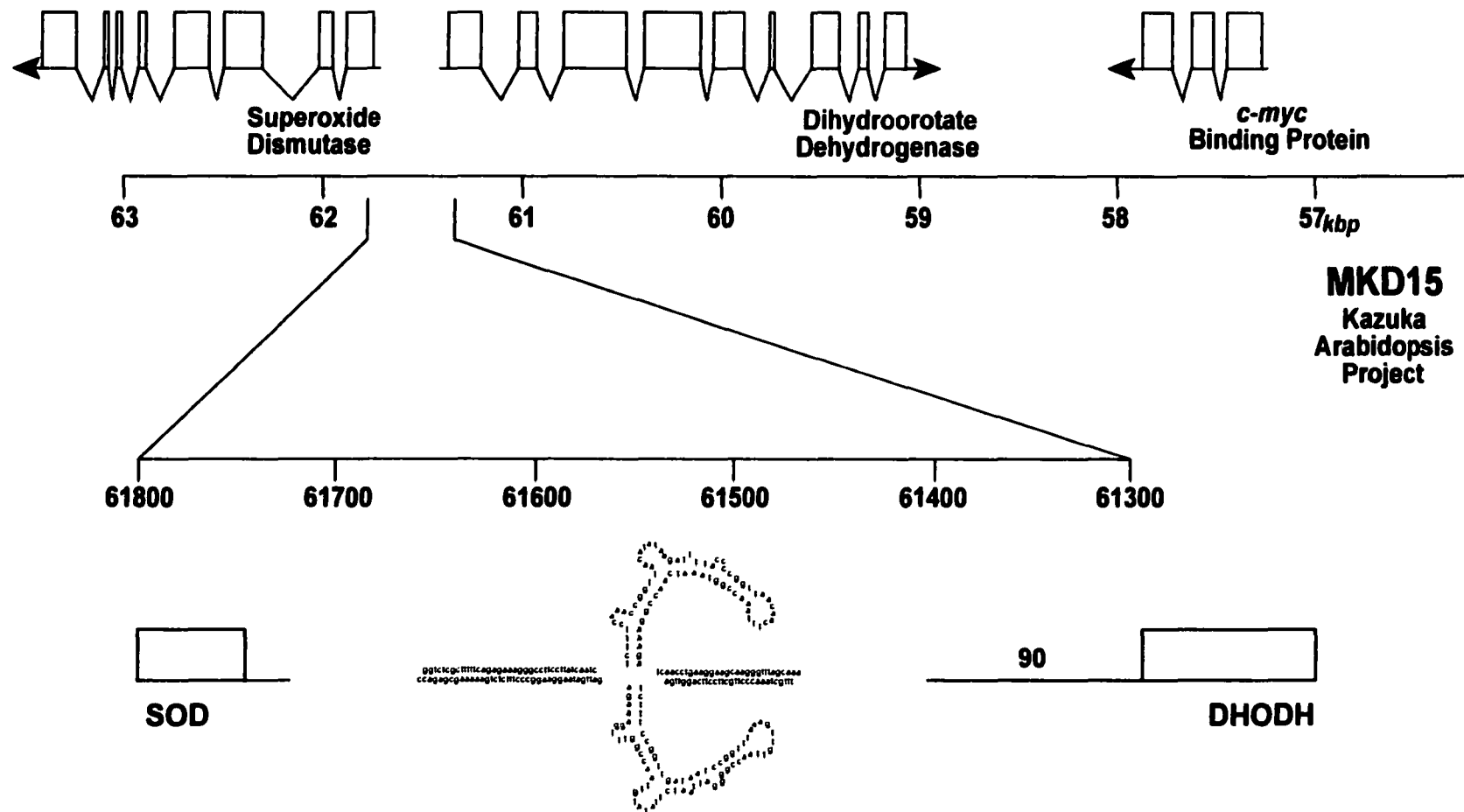


Figure 16.

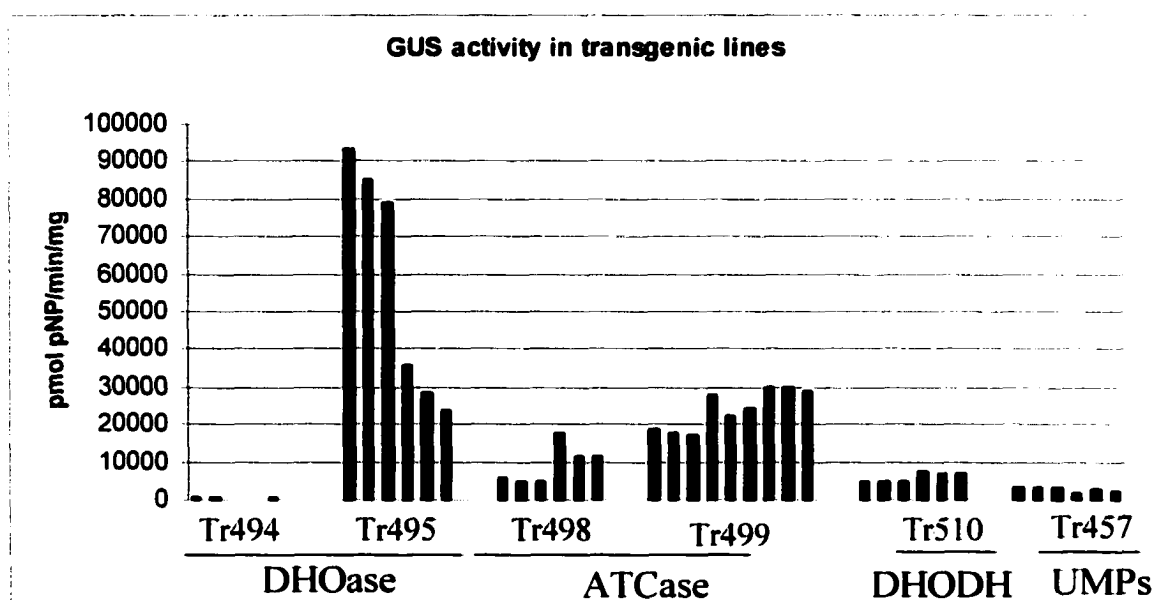


Figure 17.

CHAPTER 5. GENERAL CONCLUSIONS

Pyrimidines are at the core of cellular metabolism. They are a constituent of genetic material, co-factors in lipid metabolism, polysaccharide biosynthesis and may be involved in intracellular signalling. *De novo* pyrimidine biosynthesis is a highly conserved pathway that has changed little over millenia of evolution. Several of the enzymes of pyrimidine metabolism are much more closely related to prokaryotic homologues than to animal or insect homologues. While other enzymes show the opposite relationship. Plants have unquestionably maintained some of the pyrimidine metabolic genes from their cyanobacterial ancestors.

In Chapter 2, we describe the cloning, expression and characterization of a cytidine deaminase from *Arabidopsis thaliana*. Our clone, *cda1*, is a member of a large protein family of cytidine deaminases. There are at least 9 different members of this family and all show a high degree of conservation at the amino acid level. The *Arabidopsis* enzyme is more closely related to the prokaryotic forms than to eukaryotic forms. So much so that we were able to generate a model of the active site based on the *E. coli* enzyme. From this analysis we have determined that the mechanism of the *Arabidopsis* cytidine deaminase is in all likelihood identical to that of the *E. coli* enzyme.

To further characterize the *Arabidopsis cda1* protein the cDNA was subcloned into the bacterial expression vector pProEX Ht-b, which encodes a 6x histidine tag for purification. The enzyme was purified on Ni-NTA agarose. Further evidence for the argument that the *Arabidopsis* enzyme is evolutionarily more closely related to the prokaryotic enzymes was revealed when we determined the kinetic parameters of this enzyme. The purified *Arabidopsis* enzyme is able to utilize cytidine ($K_m=226$) and deoxycytidine ($V_m=49$) as substrates. The

enzyme was unable to utilize CMP, dCMP, or cytosine. The kinetic parameters are quite similar to that reported for the bacterial enzymes and quite different from those reported for the human or insect homologues.

To further our understanding of pyrimidine metabolism we have developed a single tube quantitative relative RT-PCR method to monitor transcriptional expression of the *de novo* pathway. We first used this method to monitor transcriptional regulation of the entire pathway under conditions of pyrimidine starvation. Plants grown in the presence of aminopterin or 5' Fluoroorotate show a co-ordinated up-regulation of all members of the *de novo* biosynthetic pathway. These results correlate quite well with our previous studies in which we have shown the up-regulation of UMP synthase under pyrimidine starvation conditions. In chapter 3 we report the use of this method showing that the closely 5' flanking gene to UMP synthase, SKIP5, appears to be transcriptionally co-ordinately up-regulated with UMP synthase. The levels of transcript show good correlation to one another under nearly all conditions tested.

SKIP5 is a member of the family of proteins called F-box proteins. F-Box proteins are responsible for recruiting intracellular proteins to the SCF complex where they are ubiquitinated. The ubiquitinated target of SKIP5 is then degraded by the 26S proteasome. It is tempting to speculate that SKIP5 is involved with pyrimidine biosynthesis because of the increased transcript level under conditions of pyrimidine starvation. However, the target of SKIP5 remains elusive.

In chapter 4 we report our analysis of *de novo* transcript level under various environmental conditions using our RT-PCR method. Under the abiotic stress conditions of high salinity and phosphate deficiency, expression of the pathway is relatively unchanged.

Under conditions of oxygen tension the severity of the treatment yields different results. Under a close approximation of flood conditions the pathway is relatively unchanged. Under completely anaerobic conditions the final three steps in the pathway are severely down-regulated to nearly undetectable levels. Under the biotic stress of viral infection DHODH and UMP synthase are both up-regulated. From this portion of the work we conclude that *Arabidopsis* plants have an intracellular mechanism to monitor nucleotide levels and regulate expression of the *de novo* biosynthetic pathway accordingly. We also conclude that processes that slow plant growth do not necessarily alter pyrimidine pools in such a manner as to require transcriptional regulation of the biosynthetic pathway.

In chapter 4 we also present our studies of the promoters of the entire *de novo* pyrimidine biosynthetic pathway. We have eight made constructs that contain promoter regions of the four *de novo* genes. Two of the genes, ATCase and DHOase, contain introns in the 5' UTR. To assess the function of the introns, two constructs were made for promoter, with and without the intron. The results of these studies reveal that the introns are indeed involved in the regulation of expression of the native genes. The intron in the ATCase 5' UTR is required for expression of the gene in root tissues. No root expression of GUS was observed in any of the more than 10 lines of transgenic plants harboring the promoter GUS fusion which lacked the ATCase intron. The analysis of the expression of the ATCase promoter also reveals that this promoter is relatively strong and drives GUS expression throughout the entire plant. This correlates quite well with our RT-PCR analysis showing high levels of transcript found in all organs of *Arabidopsis*.

The DHOase gene also contains an intron that is involved in the regulation of expression

of the native gene. Plants that contain the 5' UTR including the intron show significant GUS expression throughout the plant. Plants that have the DHOase promoter but lack the intron from the 5' UTR show a greatly reduced level of GUS gene expression.

Transgenic lines of plants that carry the short bi-functional DHODH promoter were also generated. This promoter construct was able to drive GUS expression rather weakly. We cannot rule out 5' or 3' flanking sequences are required for normal expression of this gene in *Arabidopsis*.

A UMP synthase promoter construct was made which contains the proximal promoter region between UMP synthase and the 3' end of SKIP5. This promoter construct showed no GUS activity in any of the lines of transgenic plants we tested. For this reason we made two more constructs with flanking regions. The construct which contains a portion of the SKIP5 gene is able to drive GUS expression weakly. The expression is limited to rapidly dividing tissues such as leaf primordia, meristems and developing anthers. The final construct contains the portion of SKIP5 and also approximately one third of the UMP synthase gene itself. This construct shows significant GUS staining throughout the plant. Thus we conclude that there are regulatory elements within UMP synthase necessary for expression of the native gene.

In all of the promoter-GUS transgenic lines we observed localization of high amounts of GUS activity to the vascular tissue with a light background staining of the leaf tissue. These result imply that pyrimidines may be transported via the vascular tissue to other organs and tissues to more precisely regulate global pyrimidine pools within the plant. GUS activity was also observed in significant levels within the meristems of leaves and in root tips. This result is expected as rapidly dividing tissue should theoretically have a higher requirement for nucleotides.